ALGANT Master Thesis

The Witten-Kontsevich Theorem

David Klompenhouwer

Supervised by Prof. Paolo Rossi

Universiteit Leiden

Università degli Studi di Padova

22 September 2023

Abstract

The Witten-Kontsevich theorem relates intersection products of certain cohomology classes in the tautological ring of the moduli space of stable curves, to the KdV hierarchy of partial differential equations. In this thesis, a recent proof of this theorem is presented. Firstly, the ELSV formula relates such intersection products to simple Hurwitz numbers, which count branched covers of algebraic curves. Subsequently, the link between Hurwitz theory and integrable systems is made via the Sato Grassmannian construction for the KP hierarchy.

Acknowledgements

There are many people that I would like to thank for their friendship and support while undertaking my thesis.

Firstly, my supervisor Prof. Paolo Rossi, who introduced me to a fascinating research area that I am excited to continue exploring in the coming years. Secondly, Prof. Renzo Cavalieri, whose course in Hurwitz theory and our lively discussions helped me to find my way at the start of the research process. I would also like to express my gratitude to Prof. Harry Braden; what he taught me during my studies in Edinburgh ended up heavily influencing my academic interests.

Thank you to the many great friends that I made during my year in Leiden, and to the lads at the UVS football team. Grazie ai ragazzi del R.E.A.L. Padova, e ad Andrea, Lorenzo e Mario che mi hanno incoraggiato in questi ultimi mesi. Moreover, thank you to Ishan, who shares many interests with me and was generous enough to let me talk to him about the KP hierarchy for hours on end.

Despite the distance, Zoe has always been there for me, and for this I continue to be grateful to you. The same goes for Gabbo, Gio, Lollo and Sandy, my homeboys since day one.

Tenslotte wil ik mijn familie bedanken. Zonder jullie had ik het moeilijk gevonden om door te werken tot het einde. Met name Riccardo, Roman, en mijn ouders, ik zal jullie altijd dankbaar zijn.

Contents

1 Introduction

This thesis is about Witten's conjecture, or the Witten-Kontsevich theorem. Its main object of interest is the moduli space of stable genus g curves with n marked points, denoted by $\overline{\mathcal{M}}_{g,n}$. The theorem gives us topological information about $\overline{\mathcal{M}}_{g,n}$ by relating intersections in the top degree of certain natural cohomology classes to an integrable system of partial differential equations, from which one obtains recursive relations between these intersection numbers. It is the simplest example of an ongoing effort to understand the connection between Gromov-Witten invariants and classical integrable hierarchies of nonlinear PDEs.

The conjecture was first stated by theoretical physicist Edward Witten in [Wit90], where he noted that the statement of the conjecture followed by supposing that two approaches to two-dimensional quantum gravity coincide. It was first proved by Maxim Kontsevich in [Kon92], by combinatorializing the top intersections using ribbon graphs and applying Feynman diagram techniques. Various other proofs have appeared in the literature, many of which use different techniques to Kontsevich. For example, the approach taken in [KL06] uses virtual functorial localization on the moduli space of stable morphisms to $\mathbb{P}^1$, while the one in [Mir07] establishes a relationship with the Weil-Petersson volume of the moduli space of hyperbolic Riemann surfaces. On the other hand, the proofs in [OP01] and [KL07] go via the route of Hurwitz numbers. This is the route that we will focus on in this thesis, with particular attention paid to the latter paper by Kazarian and Lando.

We briefly present the statement of Witten's conjecture. Starting from the moduli space $\mathcal{M}_{g,n}$ of curves of genus g with n marked points, there is a natural way to compactify it, thus obtaining the moduli space $\overline{\mathcal{M}}_{g,n}$ of stable curves. Then, one considers a collection of geometrically significant classes on this space, called ψ-classes. Their Poincaré duals are denoted by ψ_i for $i = 1, \ldots, n$, and lie in $A^1(\overline{\mathcal{M}}_{g,n}; \mathbb{Q})$ or $H^2(\overline{\mathcal{M}}_{g,n}; \mathbb{Q})$, depending on whether one prefers to work in the Chow ring or cohomology ring. For $d_1 + \cdots + d_n = \dim_{\mathbb{C}} \overline{\mathcal{M}}_{g,n} = 3g - 3 + n$, the intersection product $\psi_1^{d_1} \cdots \psi_n^{d_n}$ in the top cohomological degree is denoted by

$$\langle \tau_{d_1} \cdots \tau_{d_n} \rangle := \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n} \in \mathbb{Q}.$$

The generating function in the formal variables $t_* = (t_0, t_1, t_2, \ldots)$ given by

$$F(t_*) = \sum_{n \geq 0} \sum_{(d_1, \ldots, d_n)} \langle \tau_{d_1} \cdots \tau_{d_n} \rangle \, \frac{t_{d_1} \ldots t_{d_n}}{|\mathrm{Aut}(d_1, \ldots, d_n)|}$$

encodes all such top intersection products. By the automorphisms of $(d_1, \ldots, d_n)$ we mean all the ways of permuting the d_j with the same value, e.g. $|\mathrm{Aut}(4,4,2,2,2,1)| = 2!\cdot 3!\cdot 1!$. Witten's conjecture states that $U = \frac{\partial^2 F}{\partial t_0^2}$ is a solution of the KdV hierarchy, or in other words F satisfies the following system of compatible PDEs:

$$(2n+1)\frac{\partial^3 F}{\partial t_n \partial t_0^2} = \frac{\partial^2 F}{\partial t_{n-1}\partial t_0}\frac{\partial^3 F}{\partial t_0^3} + 2\frac{\partial^3 F}{\partial t_{n-1}\partial t_0^2}\frac{\partial^2 F}{\partial t_0^2} + \frac{1}{4}\frac{\partial^5 F}{\partial t_{n-1}\partial t_0^4}, \quad n \geq 1.$$

In particular, U satisfies the famous KdV equation:

$$\frac{\partial U}{\partial t_1} = U\frac{\partial U}{\partial t_0} + \frac{1}{12}\frac{\partial^3 U}{\partial t_0^3}.$$

These PDEs, together with the string equation

$$\frac{\partial F}{\partial t_0} = \frac{t_0^2}{2} + \sum_{i=0}^{\infty} t_{i+1}\frac{\partial F}{\partial t_i}$$

give recursive relations that allow one to compute all the top intersections $\langle \tau_{d_1} \cdots \tau_{d_n} \rangle$.

The idea behind the proof in [KL07] is to relate the generating function F to the simple Hurwitz potential H, which is another generating function that encodes simple Hurwitz numbers. These numbers count ramified coverings of $\mathbb{P}^1$ with certain prescribed ramification profiles. The ELSV formula, first proved in [Eke+99; Eke+01], provides the bridge from F to H. Then, one can show that H is a solution of the KP hierarchy, another integrable hierarchy of PDEs of which the KdV hierarchy is a special case. To do this, one uses the cut-and-join equation from Hurwitz theory to express e^H in an explicit basis of polynomials, and then observes that the image of e^H under the linear isomorphism from the Boson-Fermion correspondence satisfies the Plücker relations from the Sato Grassmannian construction of the KP hierarchy. The goal of this thesis is to develop all the aforementioned theory in order to understand this proof, and to gain an appreciation of how Witten's conjecture inspired ongoing research into the connection between Gromov-Witten invariants and integrable hierarchies.

Section 2 introduces the moduli space of curves, its compactification, and its tautological ring $R^*(\overline{\mathcal{M}}_{g,n}; \mathbb{Q}) \subset H^*(\overline{\mathcal{M}}_{g,n}; \mathbb{Q})$. In some sense, this ring contains most of the geometrically relevant classes, of which the ψ-classes will be our main focus. We then show some preliminary results on top intersections of ψ-classes, which point towards Witten's conjecture. Most of the content of this section comes from Vakil's and Zvonkine's expository papers [Vak03; Vak08; Zvo14] and Witten's original paper [Wit90].

Section 3 introduces Hurwitz numbers and their links to algebra and combinatorics, following Cavalieri and Miles' textbook [CM16]. We prove that the cut-and-join operator annihilates the Hurwitz potential $H^\bullet$ for simple disconnected Hurwitz numbers. Next, following [Lan10], we express $H^\bullet$ as an expansion in a basis of symmetric functions called Schur functions. At the end, the ELSV formula and some of its consequences are presented.

Section 4 gives a brief but detailed overview of the KdV and KP hierarchies, following Buryak's lecture notes [Bur22]. Next, we explain how solutions of these hierarchies can be understood in the context of the Boson-Fermion correspondence, following the textbook [MJD00]. Next, we explain how the space of solutions in the Fermionic picture can be identified with an infinite-dimensional Grassmannian, the Sato Grassmannian, embedded into an ambient wedge space.

Section 5 puts all of the theory developed thus far together, to present the remaining parts of the proof of Witten's conjecture in [KL07]. We end the thesis by presenting some generalizations of Witten's conjecture and the essential objects of Gromov-Witten theory.

2 The moduli space of curves

2.1 Examples and dimension

The moduli space of curves is a central object in geometry, and efforts to understand it involve ideas from many areas in mathematics and physics. In everything that follows we work over $\mathbb{C}$, and by a curve we will mean a smooth, compact, complex curve; in other words a Riemann surface. We will also assume that curves are connected, although disconnected curves will enter the discussion in Section 3. By "dimension" we will mean the algebraic/complex dimension, which is half of the real dimension.

A natural question to ask is: can we classify curves up to isomorphism? This brings us to the first definition.

Definition 2.1.1. The *moduli space of curves* $\mathcal{M}_g$ is the set of isomorphism classes of genus g curves. For $2g-2+n > 0$, the moduli space $\mathcal{M}_{g,n}$ is the set of isomorphism classes of genus g curves with n distinct marked points. The elements of $\mathcal{M}_{g,n}$ are denoted by $(C, x_1, \dots, x_n)$.

Despite calling $\mathcal{M}_g$ and $\mathcal{M}_{g,n}$ by the name of "spaces", for the moment they are simply defined as sets. For them to be of any use, we would like to endow them with some sort of structure, ideally that of a projective variety or complex manifold. This is not always possible because of the presence of nontrivial automorphisms of curves, which give these spaces the structure of Deligne-Mumford stacks (in the algebro-geometric setting) or orbifolds (in the analytic setting). We will give more details on these structures shortly.

The condition $2g-2+n > 0$ in the second part of the definition is necessary for the orbifold structure to exist. This comes from the fact that a genus g curve has a finite group of automorphisms which preserve the n marked points if and only if its Euler characteristic is negative, $2-2g-n < 0$. We explain this briefly. For $g \geq 2$, the automorphism group of a genus g curve is finite due to a theorem by Hurwitz [Mir95, III.3]. A genus $g = 1$ curve C with one marked point is an elliptic curve, which has finite automorphism group [Sil09, III.10]. When one forgets the marked point, the automorphism group is infinite because for any $x_1, x_2 \in C$, there is an automorphism that sends x_1 to x_2 [Har77, IV.4]. For $g = 0$, we first remark that any genus zero curve C is isomorphic to the Riemann sphere $\mathbb{P}^1$ [Mir95, VII.1]; this is a consequence of the Riemann-Roch theorem. The automorphism group of $\mathbb{P}^1$ is the group of Möbius transformations $\mathrm{PGL}(2, \mathbb{C})$, acting by

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot x = \frac{ax+b}{cx+d}$$

Since $\mathrm{PGL}(2, \mathbb{C})$ has dimension three, we see that fixing less than three points on $\mathbb{P}^1$ allows infinitely many automorphisms, while fixing at least 3 points forces any automorphism to be the identity. In fact, given distinct points $x_1, x_2, x_3 \in \mathbb{P}^1$ there is a Möbius transformation sending the triple to $0, 1, \infty$ respectively, given by

$$x \longmapsto \frac{(x - x_1)(x_2 - x_3)}{(x - x_3)(x_2 - x_1)}. \tag{2.1}$$

Hence we see that the pairs (g, n) for which $(C, x_1, \ldots, x_n)$ does not have a finite group of automorphisms are $(0, 0), (0, 1), (0, 2)$ and $(1, 0)$.

The discussion above furnishes us with the first few examples of moduli spaces of curves.

Example 2.1.2. Let $g = 0$. Any rational curve with three marked points (C, x_1, x_2, x_3) is isomorphic to $(\mathbb{P}^1, 0, 1, \infty)$ using (2.1). Therefore $\mathcal{M}_{0,3} = \{*\}$ consists of a single point. Similarly any (C, x_1, x_2, x_3, x_4) is isomorphic to $(\mathbb{P}^1, 0, 1, \infty, t)$, where $t = \frac{(x_4 - x_1)(x_2 - x_3)}{(x_4 - x_3)(x_2 - x_1)}$. Hence $\mathcal{M}_{0,4} = \mathbb{P}^1 \setminus \{0, 1, \infty\}$. Generalizing this argument gives

$$\mathcal{M}_{0,n} = \{(t_1, \ldots, t_{n-3}) \in \left(\mathbb{P}^1 \setminus \{0, 1, \infty\}\right)^{n-3} : t_i \neq t_j\}.$$

We see that $\mathcal{M}_{0,n}$ has dimension $n - 3$.

Example 2.1.3. Consider $\mathcal{M}_{1,1}$, the moduli space of elliptic curves. Every elliptic curve is isomorphic to $\mathbb{C}/\Lambda$ for $\Lambda = z_1\mathbb{Z} \oplus z_2\mathbb{Z}$ a rank 2 lattice. Multiplying the basis of Λ by $\pm 1/z_1$ gives $\Lambda = \mathbb{Z} \oplus \tau\mathbb{Z}$ for some $\tau \in \mathbb{H}$. Two elliptic curves determined by τ and τ' are isomorphic if and only if τ' lies in the $\mathrm{SL}(2, \mathbb{Z})$-orbit of τ, namely

$$\tau' = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \tau = \frac{a\tau + b}{c\tau + d}.$$

Therefore $\mathcal{M}_{1,1} = \mathbb{H}/\mathrm{SL}(2, \mathbb{Z})$. In this case, $\mathcal{M}_{1,1}$ has dimension 1.

In the examples above, the dimension of the moduli space refers to how many parameters, also referred to as moduli, are required to uniquely determine a curve and its marked points. For example, for $\mathcal{M}_{0,n}$ we need $n-3$ parameters $t_1, \ldots, t_{n-3}$

while for $\mathcal{M}_{1,1}$ we need one parameter τ. This corresponds to the dimension of $\mathcal{M}_{g,n}$ as an orbifold. Before discussing the orbifold structure, we compute the dimension of $\mathcal{M}_{g,n}$ following [Hor+03, Exercise 23.2.1].

Proposition 2.1.4. *Let $2g-2+n>0$. The dimension of the space $\mathcal{M}_{g,n}$ is $3g-3+n$.*

Proof. We show that $\dim \mathcal{M}_g = 3g-3$. The result for $\mathcal{M}_{g,n}$ follows since fixing n distinct points imposes n more independent conditions. The following argument only works for $g \geq 2$, but it will nevertheless give the correct statement for $g=0$ and $g=1$. The idea is the following: fix an integer $d > 2g-2$, and compute the dimension of

$$\mathcal{C}_g^d = \{f : C \to \mathbb{P}^1 \text{ ramified cover} \mid g(C) = g, \deg f = d\},$$

the space of degree d covers of $\mathbb{P}^1$ by genus g curves, in two different ways. First, we compute $\dim \mathcal{C}_g^d$ as

$$\dim \mathcal{C}_g^d = \dim \mathcal{M}_g + \dim \mathcal{F}_C^d,$$

where $\mathcal{F}_C^d$ is the space of degree d covers of $\mathbb{P}^1$ by a *fixed* curve C of genus g. To compute $\dim \mathcal{F}_C^d$, consider a degree d line bundle L over C. The degree of the canonical bundle K_C of C is $2g-2$ [Mir95, V.1], so

$$\deg(K_C \otimes L^*) = \deg K_C - \deg L = 2g-2-d < 0.$$

So any nonzero section of $K_C \otimes L^*$ has a pole, which means $h^0(C, K_C \otimes L^*) = 0$. But $h^0(C, K_C \otimes L^*) = h^1(C, L)$ by Serre duality, so applying the Riemann-Roch theorem gives $h^0(C,L) = h^0(C,L) - h^1(C,L) = d+1-g$; hence L has a $(d+1-g)$-dimensional space of sections. Two independent sections $s, t \in H^0(C,L)$ determine a degree d cover $f : C \to \mathbb{P}^1$ by setting $f(x) = [s(x) : t(x)]$. Note that rescaling both sections gives the same cover. Conversely, any such cover f determines two sections s and t up to rescaling, by taking $\operatorname{div}(s) = f^{-1}([0:1])$ and $\operatorname{div}(t) = f^{-1}([1:0])$. Hence a cover in $\mathcal{F}_C^d$ is uniquely determined by a choice of line bundle $L \in \operatorname{Pic}^d C$ and two independent sections $s, t \in H^0(C,L)$ up to rescaling:

$$\begin{aligned} \dim \mathcal{F}_G^d &= \dim \operatorname{Pic}^d C + (2h^0(C,L) - 1) \\ &= g + 2(d-g+1) - 1 = 2d - g + 1. \end{aligned} \tag{2.2}$$

In the second equality we used the isomorphism $\operatorname{Pic}^0 C \cong \operatorname{Pic}^d C$ that twists the sheaf of sections d times at a point, and the fact that $\dim \operatorname{Pic}^0 C = g$ [Har77, B.5].

On the other hand, consider a degree d cover of $\mathbb{P}^1$ by a genus g curve. By the Riemann-Hurwitz formula [CM16, 4.4], the number of branch points on $\mathbb{P}^1$ (counted with multiplicity) is given by $b = 2d + 2g - 2$. By Riemann's existence theorem [CM16, 6.2], the branch points determine the cover, so

$$\dim \mathcal{C}_g^d = 2d + 2g - 2. \tag{2.3}$$

Putting (2.3) and (2.2) together we obtain $2d + 2g - 2 = \dim \mathcal{M}_g + 2d - g + 1$, which implies $\dim \mathcal{M}_g = 3g - 3$. $\square$

2.2 Complex orbifolds

As we have mentioned, it would be ideal if the moduli space $\mathcal{M}_{g,n}$ had the structure of a projective variety or complex manifold, but this is not always the case. To obtain a category in which these moduli spaces live, the notions of varieties or manifolds have to be relaxed to include Deligne-Mumford stacks or complex orbifolds. Intuitively, an orbifold generalises the notion of a manifold by being locally isomorphic to an open subset of $\mathbb{C}^n$ factored by the action of a finite group. Below we give the essential definitions and results following [Zvo14], without presenting proofs and without delving into the intricacies of the construction of $\mathcal{M}_{g,n}$ as an orbifold. Details of this can be found in [HM98].

Definition 2.2.1. Let X be a topological space. A *complex orbifold chart* on X consists of a homeomorphism $\varphi : U \to V/G$, where $U \subset X$ is an open set and $V \subset \mathbb{C}^n$ is a contractible open set endowed with a biholomorphic action of a finite group G.

The notions of subcharts and compatible charts are similar to the ones for manifolds, with the added condition that they behave well with respect to the group actions. Moreover, the formal definition of a complex orbifold using an atlas follows the same idea as for manifolds.

Definition 2.2.2. A chart $\varphi' : U' \to V'/G'$ is a *subchart* of $\varphi : U \to V/G$ if $U' \subset U$ and there is a group homomorphism $\sigma : G' \to G$ and a holomorphic embedding $i : V' \to V$ such that the stabilizers $G'_{y'}$ and $G_{i(y')}$ are isomorphic for all $y' \in V'$, and the following diagrams commute:

$$
\begin{tikzcd}
V' \times G' \arrow[r, "{(i,\sigma)}"] \arrow[d, "\rho'"'] & V \times G \arrow[d, "\rho"] \\
V' \arrow[r, "i"'] & V
\end{tikzcd}
\qquad
\begin{tikzcd}
U' \arrow[r, hook] \arrow[d, "\varphi'"'] & U \arrow[d, "\varphi"] \\
V'/G' \arrow[r, "j"'] & V/G
\end{tikzcd}
$$

where ρ' and ρ denote the group actions and j is the map induced by i, which is well-defined due to the commutativity of the first diagram.

Definition 2.2.3. Two orbifold charts (U, φ) and (U', φ') are *compatible* if every point in $U \cap U'$ is contained in a chart (U'', φ'') that is a subchart of both (U, φ) and (U', φ').

Definition 2.2.4. An *orbifold atlas* on a topological space X is a family of compatible orbifold charts covering X. A *complex orbifold* is a topological space together with a maximal orbifold atlas.

Definition 2.2.5. Let X be a complex orbifold. The *stabiliser* of $x \in X$ is the stabiliser subgroup $G_{\varphi(x)}$ of a representative of $\varphi(x)$ in V, where $\varphi : U \to V/G$ is some chart around x.

Notice that if the stabilizer of every point of an orbifold is trivial, then a complex orbifold structure is exactly the same as a complex manifold structure. We complete the definition of the orbifold category by defining a morphism of orbifolds as in [KL14].

Definition 2.2.6. A morphism of complex orbifolds $X_1 \to X_2$ is given by a continuous map $f : X_1 \to X_2$ of the underlying topological spaces with the additional data for every $x \in X_1$: charts $\varphi_1 : U_1 \to V_1/G_1$ and $\varphi_2 : U_2 \to V_2/G_2$ around x and $f(x)$ respectively, a holomorphic map $\hat{f} : V_1 \to V_2$, and a group homomorphism $\sigma : G_1 \to G_2$ making the following diagrams commute:

$$
\begin{tikzcd}
V_1 \times G_1 \arrow[r, "{(\hat{f},\sigma)}"] \arrow[d, "\rho_1"'] & V_2 \times G_2 \arrow[d, "\rho_2"] \\
V_1 \arrow[r, "\hat{f}"'] & V_2
\end{tikzcd}
\qquad
\begin{tikzcd}
V_1 \arrow[r, "\hat{f}"] \arrow[d] & V_2 \arrow[d] \\
V_1/G_1 \arrow[r, "\bar{f}"] & V_2/G_2 \\
U_1 \arrow[u, "\varphi_1"] \arrow[r, "f"'] & U_2 \arrow[u, "\varphi_2"']
\end{tikzcd}
$$

where $\bar{f}$ is the well-defined map induced by $\hat{f}$.

Before stating the main result about $\mathcal{M}_{g,n}$, we define a family of curves.

Definition 2.2.7. Let B be an orbifold. A *family of genus g curves with n marked points* over B is a morphism $p : \mathcal{C}_B \to B$ endowed with n disjoint sections $s_i : B \to \mathcal{C}_B$, such that every fibre of p is isomorphic to a smooth curve of genus g.

Intuitively, the section s_i picks out the i^{th} marked point in each fibre. We now state a simplified version of the result, which will suffice for us.

Theorem 2.2.8. *Suppose $2g-2+n > 0$. Then $\mathcal{M}_{g,n}$ has a $(3g-3+n)$-dimensional complex orbifold structure. Moreover, there exists a $(3g-2+n)$-dimensional complex orbifold $\mathcal{C}_{g,n}$ and an orbifold morphism $\pi : \mathcal{C}_{g,n} \to \mathcal{M}_{g,n}$ satisfying the following conditions for every $C = (C, x_1, \ldots, x_n) \in \mathcal{M}_{g,n}$:*

- *There is a chart $U \to V/G$ around C and a chart $\pi^{-1}(U) \to \mathcal{C}/G$ of $\mathcal{C}_{g,n}$;*
- *The map $\hat{\pi} : \mathcal{C} \to V$ is a family of genus g curves with n marked points;*
- *The fibre $C_0 = \hat{\pi}^{-1}(0)$ is isomorphic to C;*
- *The G-action on $\mathcal{C}$ preserves C_0 and acts as $\mathrm{Aut}(C)$;*
- *For any family of genus g curves with n marked points $p : \mathcal{C}_B \to B$ such that $p^{-1}(b) \cong C$ for some $b \in B$, there is a subset $B' \subset B$ containing b and a map $f : B' \to V$ such that the restriction of $p : \mathcal{C}_B \to B$ to B' is the pull-back of $\hat{\pi} : \mathcal{C} \to V$ by f. In other words,*

$$\begin{tikzcd} p^{-1}(B') \arrow[r] \arrow[d, "p"'] & \mathcal{C} \arrow[d, "\hat{\pi}"] \\ B' \arrow[r, "f"'] & V \end{tikzcd}$$

  is a pull-back square.

Because of the universal property in the last bullet point, we call the map $\pi : \mathcal{C}_{g,n} \to \mathcal{M}_{g,n}$ the *universal curve* over $\mathcal{M}_{g,n}$. Moreover, a consequence of the third and fourth bullet points is that the stabiliser of $(C, x_1, \ldots, x_n) \in \mathcal{M}_{g,n}$ is isomorphic to $\mathrm{Aut}(C, x_1, \ldots, x_n)$, the group of automorphisms of C which fix each x_i. From the discussion in Section 2.1, we know that n-marked curves of genus g have finitely many automorphisms if and only if $2g - 2 + n > 0$, or equivalently

$(g, n) \neq (0,0), (0,1), (0,2), (1,0)$. Therefore $\mathcal{M}_{g,n}$ has an orbifold structure if and only if $2g - 2 + n > 0$, which is why this condition is included at the start of the theorem. We have also discussed how $\mathbb{P}^1$ with $n \geq 3$ marked points has a trivial automorphism group. This is equivalent to the stabiliser of every point in $\mathcal{M}_{0,n}$ being trivial, which means that Theorem 2.2.8 endows $\mathcal{M}_{0,n}$ with the structure of a complex manifold. This is not true anymore when $g = 1$ because elliptic curves always have a nontrivial automorphism (the involution), so $\mathcal{M}_{1,1}$ is unavoidably a complex orbifold. We see that in general, the obstruction to $\mathcal{M}_{g,n}$ having the structure a complex manifold is the presence of nontrivial automorphisms of curves inside $\mathcal{M}_{g,n}$.

2.3 Deligne-Mumford-Knudsen compactification

In Example 2.1.2 we argued that $\mathcal{M}_{0,4} = \mathbb{P}^1 \setminus \{0, 1, \infty\}$, which shows that the space $\mathcal{M}_{g,n}$ need not be compact. A sensible compactification $\overline{\mathcal{M}}_{g,n}$ of $\mathcal{M}_{g,n}$ would be one where we add objects that look like curves, in such a way that $\overline{\mathcal{M}}_{0,4} = \mathbb{P}^1$. Such a compactification was given by Deligne and Mumford for $n = 0$ in [DM69], and by Knudsen for n general in [Knu83]. To explain it, we use $\mathcal{M}_{0,4}$ as guiding example.

The objects that we add to $\mathcal{M}_{g,n}$ to compactify it should arise as limits of sequences of curves in as natural a way as possible. For example, consider the four-pointed rational curve $(C, x_1, x_2, x_3, x_4) \cong (\mathbb{P}^1, 0, 1, \infty, t) \in \mathcal{M}_{0,4}$ as $t \to 0$. In the limit we obtain a curve where the points x_1 and x_4 coincide. But this is not independent of the local coordinate on C; if we change it via the map $x \mapsto x/t$, we would get $C(x_1, x_2, x_3, x_4) \cong (\mathbb{P}^1, 0, 1/t, \infty, 1)$ and in the $t \to 0$ limit the points x_2 and x_3 coincide. In order to make our choice independent of coordinates, we include both possibilities in the limit curve. With this description, a four-pointed rational

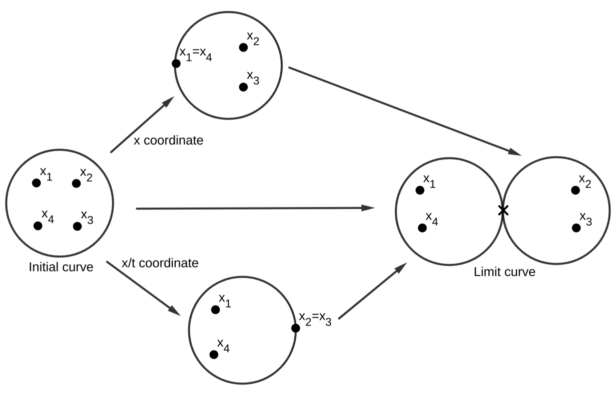

Figure 1: Limit curve of $(C, x_1, x_2, x_3, x_4) \in \mathcal{M}_{0,4}$.

curve tends to two two-pointed rational curves meeting along a simple node, see Figure 1. This node is analytically isomorphic to $xy = 0$ in $\mathbb{C}^2$, so the figure is somewhat misleading as the curves should meet transversally. This is a simple example of a stable curve with two components. We give the general definition.

Definition 2.3.1. A *stable curve* C is a curve whose points are either smooth or simple nodes, and such that $\mathrm{Aut}(C)$ is finite. A stable curve with n marked points $x_1, \ldots, x_n$ is a stable curve whose marked points are smooth. A *special point* on C is a point that is either a marked point or a node.

For a normal genus g curve with n marked points, its automorphism group is finite if and only if $2g - 2 + n > 0$. For a stable curve C, a similar statement is true when we consider its connected components. Let $C_1, \ldots, C_k$ be the connected components of C, which are smooth curves meeting at the simple nodes. Let g_i be the genus of C_i and n_i be the number of special points on C_i. Then $\mathrm{Aut}(C)$ is finite if and only if $2g_i - 2 + n_i > 0$ for every i. The genus g of C is the genus of the smooth curve that arises by "smoothening" all the nodes. A combinatorial argument involving Euler characteristics shows that

$$2 - 2g + 2\delta = \sum_{i=1}^{k} (2 - 2g_i)$$
$$\implies g = \sum_{i=1}^{k} g_i - k + 1 + \delta,$$

where δ is the number of nodes in C.

Definition 2.3.2. For $2g - 2 + n > 0$, the *the moduli space of stable curves* $\overline{\mathcal{M}}_{g,n}$ is the set of isomorphism classes of stable curves of genus g with n marked points.

A theorem analogous to Theorem 2.2.8 states that $\overline{\mathcal{M}}_{g,n}$ is a complex orbifold and has the properties that one would desire for a compactification of $\mathcal{M}_{g,n}$. A detailed exposition of this can be found in [HM98, Chapter 4].

Theorem 2.3.3. *Suppose $2g - 2 + n > 0$. Then $\overline{\mathcal{M}}_{g,n}$ is compact and has a $(3g-3+n)$-dimensional complex orbifold structure. Moreover, there exists a $(3g-2+n)$-dimensional complex orbifold $\overline{\mathcal{C}}_{g,n}$ and an orbifold morphism $\pi : \overline{\mathcal{C}}_{g,n} \to \overline{\mathcal{M}}_{g,n}$ satisfying the following conditions for every $C \in \overline{\mathcal{M}}_{g,n}$:*

- *$\mathcal{M}_{g,n} \subset \overline{\mathcal{M}}_{g,n}$ is an open dense sub-orbifold and $\pi^{-1}(\mathcal{M}_{g,n}) = \mathcal{C}_{g,n} \subset \overline{\mathcal{C}}_{g,n}$;*

- *There is a chart $U \to V/G$ around C and a chart $\pi^{-1}(U) \to \overline{\mathcal{C}}/G$ of $\overline{\mathcal{C}}_{g,n}$;*
- *The map $\hat{\pi} : \overline{\mathcal{C}} \to V$ is a family of stable genus g curves with n marked points;*
- *The fibre $C_0 = \hat{\pi}^{-1}(0)$ is isomorphic to C;*
- *The G-action on $\mathcal{C}$ preserves C_0 and acts as $\mathrm{Aut}(C)$;*
- *For any family of stable genus g curves with n marked points $p : \overline{\mathcal{C}}_B \to B$ such that $p^{-1}(b) \cong C$ for some $b \in B$, there is a subset $B' \subset B$ containing b and a map $f : B' \to V$ such that the restriction of $p : \overline{\mathcal{C}}_B \to B$ to B' is the pull-back of $\hat{\pi} : \overline{\mathcal{C}} \to V$ by f.*

We also refer to the family $\pi : \overline{\mathcal{C}}_{g,n} \to \overline{\mathcal{M}}_{g,n}$ as the universal curve over $\overline{\mathcal{M}}_{g,n}$.

There is a natural stratification of $\overline{\mathcal{M}}_{g,n}$, which we now outline. To each stable curve one can associate a graph consisting of numbered vertices, edges and half-edges, called its *dual graph.* The vertices correspond to the irreducible components and the number at each vertex corresponds to the genus of the component. An node between two components is represented by an edge between the corresponding vertices, and each marked point is represented by a half-edge incident to the appropriate vertex. We illustrate some examples in Figure 2. The vertices without

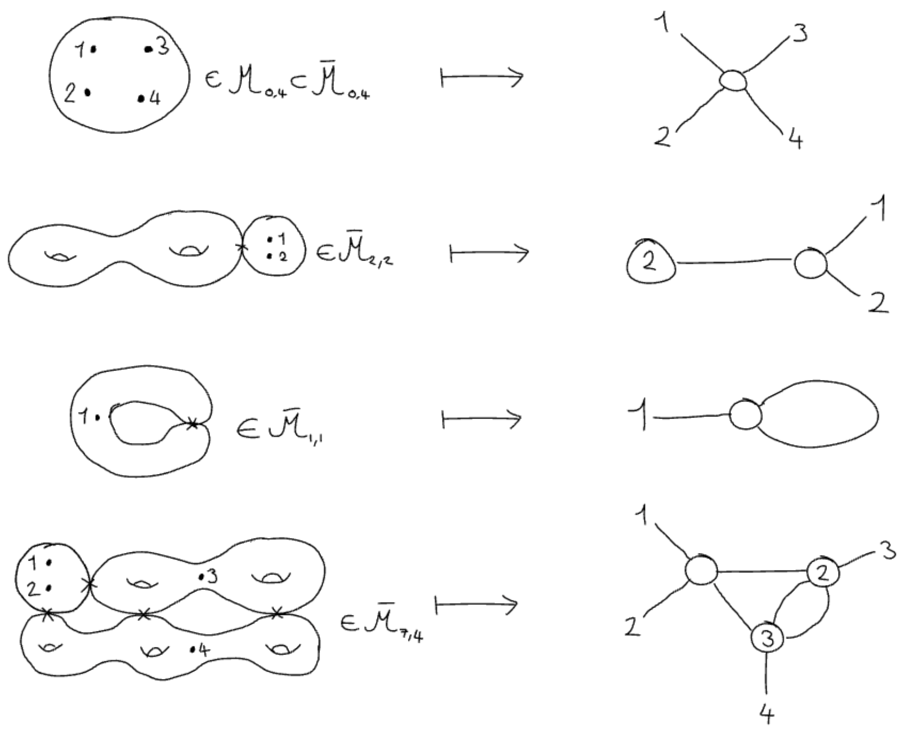

Figure 2: Dual graphs of stable curves.

a number inside them represent a genus 0 component. The moduli space $\overline{\mathcal{M}}_{g,n}$ is stratified by dual graphs. This means that each stratum $\mathcal{M}_\Gamma \subset \overline{\mathcal{M}}_{g,n}$ is labelled by a stable graph Γ, which represents the topological type of the curves in the stratum.

In Example 2.1.2 and Example 2.1.3 we argue that $\mathcal{M}_{0,4} = \mathbb{P}^1 \setminus \{0, 1, \infty\}$ and $\mathcal{M}_{1,1} = \mathbb{H}/\operatorname{SL}(2,\mathbb{Z})$. The only boundary stratum of $\overline{\mathcal{M}}_{1,1}$ is the third one in Figure 2, while the three boundary strata of $\overline{\mathcal{M}}_{0,4} = \mathbb{P}^1$ are given in Figure 3, and correspond to the missing points $0, 1$ and ∞.

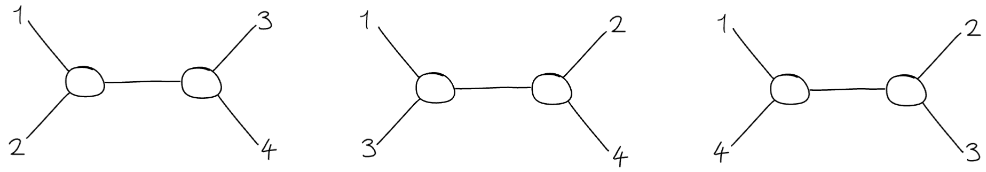

Figure 3: Boundary strata of $\overline{\mathcal{M}}_{0,4}$.

2.4 Interlude: Chow ring versus cohomology ring

Perhaps the most important reason for why one wants to compactify $\mathcal{M}_{g,n}$ is to be able to apply powerful tools from algebraic topology and geometry, such as intersection theory and cohomology. For example, a result in [Beh02] shows that an analogue of Poincaré duality [Bre93, VI.8] is true for compact complex orbifolds such as $\overline{\mathcal{M}}_{g,n}$:

$$H^k(\overline{\mathcal{M}}_{g,n};\mathbb{Q}) \cong H_{2d-k}(\overline{\mathcal{M}}_{g,n};\mathbb{Q}), \tag{2.4}$$

where $d = \dim \overline{\mathcal{M}}_{g,n} = 3g-3+n$. We remind that d denotes the complex/algebraic dimension, which is why we have written $2d$ in the equation above. Here, by the (co)homology groups of an orbifold X over $\mathbb{Q}$ we mean the singular (co)homology groups of the underlying topological space with coefficients in $\mathbb{Q}$. By the same reference, there is an associative multiplicative structure on $H^*(X;\mathbb{Q})$ making it into a graded ring. From now on, when working with the orbifold $\overline{\mathcal{M}}_{g,n}$ we will always consider its cohomology ring over $\mathbb{Q}$, and write it as $H^*(\overline{\mathcal{M}}_{g,n})$ instead of $H^*(\overline{\mathcal{M}}_{g,n};\mathbb{Q})$. The reason why we take coefficients in $\mathbb{Q}$ rather than in $\mathbb{Z}$ is because the orbifold structure involves factoring by finite groups of automorphisms, which gives rise to rational numbers when performing computations in enumerative geometry. For example, see equation (2.16) later.

We momentarily interrupt our discussion of the moduli space of curves to comment on the conventions that we will adopt when discussing cohomology classes on $\overline{\mathcal{M}}_{g,n}$, which arise from the difference between the cohomology ring and its algebro-

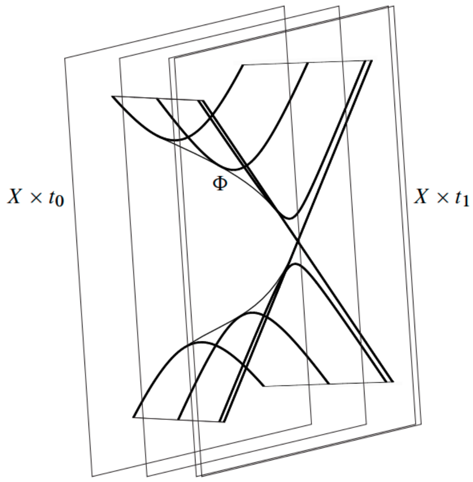

Figure 4: Rational equivalence between hyperbola and union of two lines in $\mathbb{P}^2$ [EH16, p. 17].

geometric counterpart, the Chow ring. We briefly present the Chow ring $A^*(X)$ of a smooth quasi-projective algebraic variety X over $\mathbb{C}$ following [EH16, Chapter 1] in order to be aware of these differences.

Firstly, denote by $Z(X)$ the free abelian group generated by the subvarieties of X. Elements of this group are called *cycles*, and are given by formal sums $\sum_i n_i Y_i$ with $n_i \in \mathbb{Z}$ and where each $Y_i \subset X$ is a subvariety. Two cycles $A, B \in Z(X)$ are *rationally equivalent* if there is a rationally parametrized family of cycles interpolating between them; this means that there is a cycle on $\mathbb{P}^1 \times X$ whose restrictions to two fibres $\{t_0\} \times X$ and $\{t_1\} \times X$ are A and B. See Figure 4 for an example. The *Chow group* of X is defined by identifying rationally equivalent cycles in X:

$$A(X) = Z(X)/\operatorname{Rat}(X).$$

The subgroup of $A(X)$ generated by dimension k subvarieties is denoted by $A_k(X)$, and $A^k(X)$ is the one generated by codimension k subvarieties. Two subvarieties $Y, Z \subset X$ are *generically transverse* if for every $p \in Y, Z$, the tangent spaces to Y and Z span the tangent space to X:

$$T_pY + T_pZ = T_pX.$$

This is equivalent to the condition

$$\operatorname{codim}(T_pY \cap T_pZ) = \operatorname{codim}(T_pY) + \operatorname{codim}(T_pZ). \tag{2.5}$$

This definition is extended to cycles by saying that $A = \sum_i n_i Y_i$ and $B = \sum_j m_j Z_j$

are generically transverse if each Y_i is generically transverse to eachg Z_j. The following are two fundamental results in intersection theory in algebraic geometry, with the second one following from the first.

Lemma 2.4.1 (Moving lemma)**.** *Let X be a smooth quasi-projective variety. Then for every $\alpha, \beta \in A(X)$ there are generically transverse cycles $A, B \in Z(X)$ such that $[A] = \alpha$ and $[B] = \beta$. Moreover, the class $[A \cap B] \in A(X)$ is independent of the choice of such cycles A and B.*

Theorem 2.4.2. *Let X be a smooth quasi-projective variety. There is a unique product structure on $A(X)$ defined by*

$$\alpha\beta = [A \cap B], \quad \forall\, \alpha, \beta \in A(X)$$

where we use the notation from the moving lemma. This makes

$$A^*(X) := \bigoplus_{k=0}^{\dim X} A^k(X)$$

into an associative commutative ring, graded by codimension.

The resulting ring $A^*(X)$ is the *Chow ring* of X. It is graded by codimension due to condition (2.5). The moving lemma was proved by Chow in [Cho56], whose first sentence is "It is well-known that there is a close analogy between the intersection theories in the abstract algebraic geometry and in topology". By comparing what we have presented above to a topology text such as [Bre93, VI. 6], one sees that the topological intersection product is similar to the algebraic intersection product (Theorem 11.9 in the reference is analogous to Theorem 2.4.2).

As before, one can construct this for a complex orbifold such as $\overline{\mathcal{M}}_{g,n}$ as well. In general the Chow ring is neither a stronger nor a weaker invariant than the cohomology ring. An important difference, however, is that the Chow ring only sees even-graded cohomology, since the index in $A^k(X)$ refers to the complex/algebraic codimension. Then there is a natural map $A^k(X) \to H^{2k}(X; \mathbb{Q})$. For our purposes, from this point onward, we will not worry about whether we need to work in $H^*(\overline{\mathcal{M}}_{g,n})$ or in $A^*(\overline{\mathcal{M}}_{g,n})$ and we will always use $H^*(\overline{\mathcal{M}}_{g,n})$ in our notation. Then, for example, the class of a codimension 1 sub-orbifold (a divisor) of $\overline{\mathcal{M}}_{g,n}$ will lie in $H^2(\overline{\mathcal{M}}_{g,n})$, since complex codimension 1 corresponds to real codimension 2. By Poincaré duality (2.4) this is the same as lying in $H_{2d-2}(\overline{\mathcal{M}}_{g,n})$, where $d = \dim \overline{\mathcal{M}}_{g,n} = 3g - 3 + n$.

2.5 The tautological ring

In this section we introduce some natural cohomology classes on the moduli space of stable curves $\overline{\mathcal{M}}_{g,n}$, which constitute its so-called tautological ring.

There is a *forgetful* morphism between moduli spaces of stable curves, which forgets the last marked point:

$$\begin{array}{ccc} \overline{\mathcal{M}}_{g,n+1} & \longrightarrow & \overline{\mathcal{M}}_{g,n}, \\ (C, x_1, \dots, x_n, x_{n+1}) & \longmapsto & (\hat{C}, \hat{x}_1, \dots, \hat{x}_n). \end{array} \tag{2.6}$$

We remind that every irreducible component C_i of C must satisfy $2g_i - 2 + n_i > 0$ for C to be a stable curve, where g_i is the genus of C_i and n_i is the number of special points on C_i. This may no longer be true after we have forgotten x_{n+1}, since the number of special points on the component containing x_{n+1} is decreased by 1. Hence by $(\hat{C}, \hat{x}_1, \dots, \hat{x}_n)$ we mean the stabilisation of $(C, x_1, \dots, x_n)$, which consists in shrinking the component containing x_{n+1} to a point if it has become unstable. An important fact, outlined in the proposition below, is that the forgetful morphism (2.6) is the universal curve over $\overline{\mathcal{M}}_{g,n}$.

Proposition 2.5.1. *Let $2g - 2 + n > 0$. The universal curve $\pi : \overline{\mathcal{C}}_{g,n} \to \overline{\mathcal{M}}_{g,n}$ and the forgetful map $\overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{M}}_{g,n}$ are isomorphic as families over $\overline{\mathcal{M}}_{g,n}$.*

Proof. For every stable curve $(C, x_1, \dots, x_n, x_{n+1}) \in \overline{\mathcal{M}}_{g,n+1}$, let $y \in (\hat{C}, \hat{x}_1, \dots, \hat{x}_n)$ be the image of x_{n+1} under the stabilization of $(C, x_1, \dots, x_n)$. This defines a map of families over $\overline{\mathcal{M}}_{g,n}$:

$$\begin{array}{ccc} \overline{\mathcal{M}}_{g,n+1} & \longrightarrow & \overline{\mathcal{C}}_{g,n}, \\ (C, x_1, \dots, x_n, x_{n+1}) & \longmapsto & ((\hat{C}, \hat{x}_1, \dots, \hat{x}_n), y), \end{array} \tag{2.7}$$

where by $((\hat{C}, \hat{x}_1, \dots, \hat{x}_n), y)$ we mean the point $y \in \overline{\mathcal{C}}_{g,n}$ lying over $(\hat{C}, \hat{x}_1, \dots, \hat{x}_n) \in \overline{\mathcal{M}}_{g,n}$. There are the following three possibilities for what y could be:

(i) If the curve $(C, x_1, \dots, x_n)$ is stable, then it equals its stabilization $(\hat{C}, \hat{x}_1, \dots, \hat{x}_n) \in \overline{\mathcal{M}}_{g,n}$, so $y = x_{n+1}$.

(ii) If $(C, x_1, \dots, x_n)$ is unstable, then x_{n+1} could lie on a genus 0 component of C with another marked point x_i and one node, and no other special points. Then, when stabilizing $(C, x_1, \dots, x_n)$ to obtain $(\hat{C}, \hat{x}_1, \dots, \hat{x}_n) \in \overline{\mathcal{M}}_{g,n}$, this component is shrunk to the point $\hat{x}_i$. Therefore $y = \hat{x}_i$.

(iii) Alternatively, if $(C, x_1, \dots, x_n)$ is unstable, then x_{n+1} could lie on a genus 0 component with no other marked points and two nodes. This component gets shrunk to a node in $(\hat{C}, \hat{x}_1, \ldots, \hat{x}_n)$, so y equals that node in $(\hat{C}, \hat{x}_1, \ldots, \hat{x}_n)$.

These are the only possibilities. To check this, notice that $(C, x_1, \ldots, x_n)$ is unstable if and only if the component C_i containing x_{n+1} satisfies $2g_i - 2 + (n_i - 1) \leq 0$, which can only happen if x_{n+1} falls in case (ii) or (iii). We use cases (i)-(iii) to construct the inverse map $\overline{\mathcal{C}}_{g,n} \to \overline{\mathcal{M}}_{g,n+1}$. Given $((C, x_1, \ldots, x_n), y) \in \overline{\mathcal{C}}_{g,n}$:

(i) If y is not a special point of $(C, x_1, \ldots, x_n)$, then send it to the stable curve $(C, x_1, \ldots, x_n, y) \in \overline{\mathcal{M}}_{g,n+1}$.

(ii) If y is one of the marked points x_i, then send it to the stable curve $(C', x_1, \ldots, x_{i-1}, x_i', x_{i+1}, \ldots, x_n, y') \in \overline{\mathcal{M}}_{g,n+1}$ obtained by replacing $x_i \in C$ by a genus 0 curve with two marked points x_i' and y', joined to the rest of C by a simple node.

(iii) If y is a node in C, then send it to the stable curve $(C'', x_1, \ldots, x_n, y'') \in \overline{\mathcal{M}}_{g,n+1}$ obtained by replacing the point y by a genus zero curve with one marked point y'', joined to its two adjacent components by two simple nodes.

Figure 5 illustrates this isomorphism with the dotted arrow for $(g, n) = (2, 2)$. $\square$

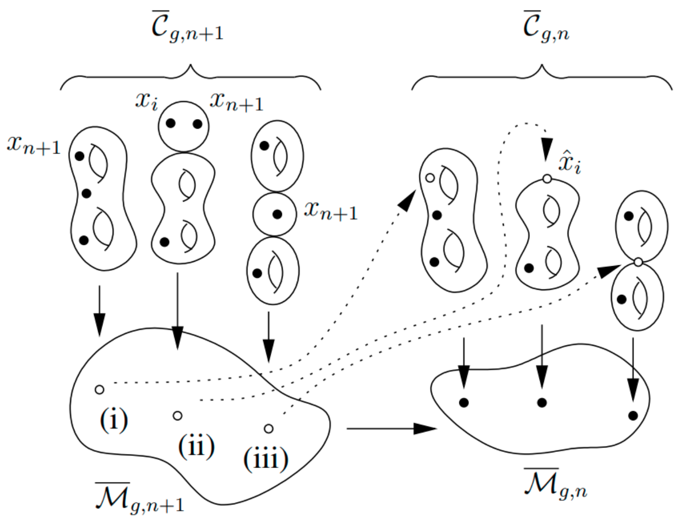

Figure 5: The universal curve over $\overline{\mathcal{M}}_{g,n}$ as the forgetful morphism [Zvo14, p. 18].

Because of this proposition, we will henceforth use the notations $\pi : \overline{\mathcal{C}}_{g,n} \to \overline{\mathcal{M}}_{g,n}$ and $\pi : \overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{M}}_{g,n}$ interchangeably to denote the universal curve.

There are two other morphisms of interests, called *gluing* morphisms. The first one

$$\mathrm{gl}_1 : \overline{\mathcal{M}}_{g_1,n_1+1} \times \overline{\mathcal{M}}_{g_2,n_2+1} \longrightarrow \overline{\mathcal{M}}_{g_1+g_2,n_1+n_2}$$

takes two marked curves $(C, x_1, \ldots, x_{n_1+1})$ and $(C', x'_1, \ldots, x'_{n_2+1})$ and identifies them along the points x_{n_1+1} and x'_{n_2+1} to create one curve with $n_1 + n_2$ marked points. The second one

$$\mathrm{gl}_2 : \overline{\mathcal{M}}_{g,n+2} \longrightarrow \overline{\mathcal{M}}_{g+1,n}$$

takes a marked curve $(C, x_1, \ldots, x_{n+1}, x_{n+2})$ and identifies the last two points x_{n+1} and x_{n+2}. This creates an extra 'hole' in C, thereby increasing its genus by one. The three morphisms described so far are called *natural* morphisms. They induce pushforwards on cohomology:

$$\begin{aligned}\pi_* : H^*(\overline{\mathcal{M}}_{g,n+1}) &\to H^*(\overline{\mathcal{M}}_{g,n}),\\ (\mathrm{gl}_1)_* : H^*(\overline{\mathcal{M}}_{g_1,n_1+1}) \otimes_{\mathbb{Q}} H^*(\overline{\mathcal{M}}_{g_2,n_2+1}) &\to H^*(\overline{\mathcal{M}}_{g_1+g_2,n_1+n_2}),\\ (\mathrm{gl}_2)_* : H^*(\overline{\mathcal{M}}_{g,n+2}) &\to H^*(\overline{\mathcal{M}}_{g+1,n}).\end{aligned}$$

Knowing the structure of the cohomology ring of a space allows one to understand its properties more deeply. In the case of $\overline{\mathcal{M}}_{g,n}$, it turns out that the entire cohomology ring $H^*(\overline{\mathcal{M}}_{g,n})$ is far too large and complicated to work with. For this reason, it is common to restrict one's attention to a certain subring of $H^*(\overline{\mathcal{M}}_{g,n})$ defined below.

Definition 2.5.2. The system of *tautological rings* $\Big(R^*(\overline{\mathcal{M}}_{g,n}) \subset H^*(\overline{\mathcal{M}}_{g,n})\Big)_{g,n}$ is the smallest system of $\mathbb{Q}$-algebras that is closed under pushforwards by the natural morphisms.

Notice that the tautological rings are defined for all (g, n) simultaneously. This definition is due to Faber and Pandharipande [FP03], and the idea behind it is that most of the classes in $\overline{\mathcal{M}}_{g,n}$ arising from geometry lie inside the tautological ring, although this is not immediately apparent. As it turns out, constructing nontautological classes in $\overline{\mathcal{M}}_{g,n}$ is highly nontrivial [GP01].

We now start giving examples of natural cohomology classes on $\overline{\mathcal{M}}_{g,n}$ arising from geometry. Firstly, denote the set of nodes in the singular fibres of $\overline{\mathcal{C}}_{g,n}$ by $\Delta \subset \overline{\mathcal{C}}_{g,n}$, and consider the line bundle over $\overline{\mathcal{C}}_{g,n} \backslash \Delta$ whose fibre at every point $y \in C$ is the cotangent line $T^*_y C$. This can be extended to a line bundle $\mathbb{L} \to \overline{\mathcal{C}}_{g,n}$, called the *relative cotangent line bundle* [Zvo14]. For $i = 1, \ldots, n$ let $s_i : \overline{\mathcal{M}}_{g,n} \to \overline{\mathcal{C}}_{g,n}$ be the section of the universal curve that selects the i^{th} marked point on each curve. Pulling back along these sections gives line bundles $\mathbb{L}_i := s_i^* \mathbb{L}$ over $\overline{\mathcal{M}}_{g,n}$, whose fibre at at $(C, x_1, \ldots, x_n)$ is $T^*_{x_i}C$.

Definition 2.5.3. The ψ-classes in $\overline{\mathcal{M}}_{g,n}$ are the first Chern classes of the line bundles $\mathbb{L}_i \to \overline{\mathcal{M}}_{g,n}$:

$$\psi_i := c_1(\mathbb{L}_i) \in H^2(\overline{\mathcal{M}}_{g,n}), \quad i = 1, \ldots, n.$$

Each ψ-class lies in the tautological ring: $\psi_i \in R^2(\overline{\mathcal{M}}_{g,n})$ [Sch20, Proposition 6.25]. A description of $R^*(\overline{\mathcal{M}}_{g,n})$ that is equivalent to Definition 2.5.2 was given in [GV05]:

Definition 2.5.4. The system of *tautological rings* $\left(R^*(\overline{\mathcal{M}}_{g,n}) \subset H^*(\overline{\mathcal{M}}_{g,n})\right)_{g,n}$ is the smallest system of $\mathbb{Q}$-vector spaces closed under pushforwards by the natural morphisms, such that all monomials in $\psi_1, \ldots, \psi_n$ are contained in $R^*(\overline{\mathcal{M}}_{g,n})$.

The tautological ring was first studied by Mumford in [Mum83] for $\overline{\mathcal{M}}_g$. In his paper, he defined two other types of tautological classes called the κ and λ-classes. The former are defined as pushforwards of powers of $\psi := c_1(\mathbb{L}) \in H^2(\overline{\mathcal{C}}_{g,n})$ along the universal curve $\pi : \overline{\mathcal{C}}_{g,n} \to \overline{\mathcal{M}}_{g,n}$:

$$\kappa_i := \pi_*\left(\psi^{i+1}\right) \in H^{2i}(\overline{\mathcal{M}}_{g,n}), \quad i = 0, 1, \ldots. \tag{2.8}$$

There is rank g vector bundle $\mathbb{E} \to \overline{\mathcal{M}}_{g,n}$, given by $\mathbb{E} := \pi_*(\mathbb{L})$. This is called the Hodge bundle, and its fibre over $(C, x_1, \ldots, x_n)$ is the space of abelian differentials on C. The λ-classes on $\overline{\mathcal{M}}_{g,n}$ are

$$\lambda_j := c_j(\mathbb{E}) \in H^{2j}(\overline{\mathcal{M}}_{g,n}), \quad j = 1, \ldots, g. \tag{2.9}$$

Using previous work by Looijenga [Loo95], Faber made various conjectures about the tautological ring of $\mathcal{M}_g$ [Fab99], including that it is a Gorenstein algebra with socle in codimension $g-2$. In other words, that it behaves like the cohomology ring of a compact $(g-2)$-dimensional manifold. In [Pan02] a similar conjecture is made for $\overline{\mathcal{M}}_{g,n}$, namely that $R^*(\overline{\mathcal{M}}_{g,n})$ is a Gorenstein algebra with socle in codimension $2(3g-3+n)$. This was proved in [FP03].

Proposition 2.5.5. *The degree $2d = 2(3g-3+n)$ of the tautological ring is one-dimensional: $R^{2d}(\overline{\mathcal{M}}_{g,n}) \cong \mathbb{Q}$. Moreover $R^i(\overline{\mathcal{M}}_{g,n}) = 0$ for $i > 2d$, and the pairing induced by the intersection product*

$$R^k(\overline{\mathcal{M}}_{g,n}) \times R^{2d-k}(\overline{\mathcal{M}}_{g,n}) \to R^d(\overline{\mathcal{M}}_{g,n}) \cong \mathbb{Q}$$

is nondegenerate.

This is nontrivial if the tautological ring is defined in terms of the Chow ring, whereas it is true if we work in $H^*(\overline{\mathcal{M}}_{g,n})$ (the degree is doubled in the statement of the conjecture because we work in cohomology). The notation for the isomorphism in the top degree is given by the integral,

$$\int_{\overline{\mathcal{M}}_{g,n}} : R^{2d}(\overline{\mathcal{M}}_{g,n}) \xrightarrow{\cong} \mathbb{Q},$$

in analogy with how integrating top-degree forms on a compact, connected, orientable manifold establishes an isomorphism $H_{dR}^{\dim M}(M) \cong \mathbb{R}$ [Spi99, Chapter 8].

Proposition 2.5.5 implies that one can recover the structure of the tautological ring of $\overline{\mathcal{M}}_{g,n}$ from knowledge of intersection products of classes in $\overline{\mathcal{M}}_{g,n}$ in the top degree. Therefore, since ψ-classes are central to the definition of the tautological ring, it is natural to study the following intersection products:

$$\langle \tau_{d_1} \cdots \tau_{d_n} \rangle := \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n} \in \mathbb{Q}. \tag{2.10}$$

These numbers are defined for any n-tuple of natural numbers $(d_1, \ldots, d_n)$, and the genus g is determined by the condition $d_1 + \cdots + d_n = \dim \overline{\mathcal{M}}_{g,n} = 3g - 3 + n$. If, for every $i \geq 0$, we say that n_i is the number of integers in $(d_1, \ldots, d_n)$ equal to i, we can write (2.10) as

$$\langle \tau_0^{n_0} \tau_1^{n_1} \cdots \rangle = \langle \underbrace{\tau_0 \cdots \tau_0}_{n_0 \text{ times}} \underbrace{\tau_1 \cdots \tau_1}_{n_1 \text{ times}} \cdots \rangle, \tag{2.11}$$

where $\sum_{i \geq 0} n_i = n$. In this case, the genus g is determined by $\sum_{i \geq 0} n_i(i-1) = 3g-3$. Witten's conjecture gives a recursive way to compute all such numbers (2.10) and (2.11).

2.6 Towards Witten's conjecture

The simplest example of an intersection product (2.10) for $2g - 2 + n > 0$ is

$$\langle \tau_0^3 \rangle = \int_{\overline{\mathcal{M}}_{0,3}} 1 = 1,$$

since $\overline{\mathcal{M}}_{0,3}$ consists of the single point $(\mathbb{P}^1, 0, 1, \infty)$. In $\overline{\mathcal{M}}_{0,n}$ it is possible to express each ψ_i as a linear combination of certain divisors [Zvo14, 2.2], which allows one to compute successively complicated intersection products on $\overline{\mathcal{M}}_{0,n}$. For example

$\langle \tau_0^3 \tau_1 \rangle = 1$.

There is an elegant and efficient way of computing intersection products of ψ-classes in $\overline{\mathcal{M}}_{0,n}$. The first step towards this is understanding the discrepancy between ψ-classes on $\overline{\mathcal{M}}_{g,n+1}$ and pull-backs of ψ-classes on $\overline{\mathcal{M}}_{g,n}$ by the forgetful morphism $\pi : \overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{M}}_{g,n}$. Before explaining this, we fix some notation. This notation is the same as in Definition 2.5.3, but with an added "prime" everywhere to indicate that we are working on the $(g, n+1)$ level. So we have sections $s_i' : \overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{C}}_{g,n+1}$ corresponding to the i^{th} marked points, the relative cotangent bundle $\mathbb{L}' \to \overline{\mathcal{C}}_{g,n+1}$, its pull-back by the sections $L_i' := s_i'^* \mathbb{L}' \to \overline{\mathcal{M}}_{g,n+1}$, and the ψ-classes $\psi_i' := c_1(\mathbb{L}_i') \in H^2(\overline{\mathcal{M}}_{g,n+1})$. Moreover, denote the forgetful map on the level of universal curves by $\pi_{\mathcal{C}} : \overline{\mathcal{C}}_{g,n+1} \to \overline{\mathcal{C}}_{g,n}$. This maps a point on a curve $(C, x_1, \ldots, x_n, x_{n+1})$ to the image of the same point on the stabilized curve $(\hat{C}, \hat{x}_1, \ldots, \hat{x}_n)$. Then there is a commutative diagram of orbifold morphisms

$$\begin{tikzcd} \overline{\mathcal{M}}_{g,n+1} \arrow[r, "s_i'"] \arrow[d, "\pi"'] & \overline{\mathcal{C}}_{g,n+1} \arrow[d, "\pi_{\mathcal{C}}"] \\ \overline{\mathcal{M}}_{g,n} \arrow[r, "s_i"'] & \overline{\mathcal{C}}_{g,n} \end{tikzcd}$$

for every $i = 1, \ldots, n$. We now prove the following result following [Wit90, 2b].

Lemma 2.6.1 (Comparison)**.** *Let $D_{0,\{i,n+1\}} \subset \overline{\mathcal{M}}_{g,n+1}$ be the boundary divisor corresponding to stable curves with one node, where one component is genus 0 and contains only the marked points x_i and x_{n+1} (see Figure 6). Then for $i = 1, \ldots, n$:*

$$\psi_i' = \pi^* \psi_i + D_{0,\{i,n+1\}}.$$

Proof. Let α be a local nonzero section of $\mathbb{L} \to \overline{\mathcal{C}}_{g,n}$. For each $1 \le i \le n$, consider the image of $D_{0,\{i,n+1\}}$ under s_i', denoted by $D_i \subset \overline{\mathcal{C}}_{g,n+1}$. This is the locus of marked points x_i belonging to an irreducible component that becomes unstable after forgetting x_{n+1} (this corresponds to possibility (ii) in the proof of Proposition 2.5.1). Therefore $\pi_{\mathcal{C}}$ maps every component of D_i to a point. Because of this, the local section $\pi_{\mathcal{C}}^* \alpha$ of the line bundle $\pi_{\mathcal{C}}^* \mathbb{L} \to \overline{\mathcal{C}}_{g,n+1}$ vanishes on D_i, with a simple zero. One can see this by expressing α in local coordinates, i.e. as a sum of differentials on $\overline{\mathcal{C}}_{g,n}$. On the other hand, the line bundles $\mathbb{L}' \to \overline{\mathcal{C}}_{g,n+1}$ and $\pi_{\mathcal{C}}^* \mathbb{L} \to \overline{\mathcal{C}}_{g,n+1}$ are isomorphic outside of $\bigcup_{i=1}^n D_i \subset \overline{\mathcal{C}}_{g,n+1}$, since $\pi_{\mathcal{C}}$ will not map a component outside of D_i to a point. Therefore

$$\mathbb{L}' \cong \pi_{\mathcal{C}}^* \mathbb{L} \otimes \bigotimes_{i=1}^{n} \mathcal{O}(D_i). \tag{2.12}$$

The pull-back of $\bigotimes_{i=1}^n \mathcal{O}(D_i)$ along the section $s_i' : \overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{C}}_{g,n+1}$ is just $\mathcal{O}(D_{0,\{i,n+1\}})$, since $\operatorname{im} s_i'$ and $\operatorname{im} s_j'$ do not intersect for $i \neq j$. Hence the pull-back of (2.12) along s_i' is

$$\begin{aligned} \mathbb{L}_i' = s_i'^* \mathbb{L}' &\cong s_i'^* \pi_{\mathcal{C}}^* \mathbb{L} \otimes \mathcal{O}(D_{0,\{i,n+1\}}) \\ &\cong \pi^* \mathbb{L}_i \otimes \mathcal{O}(D_{0,\{i,n+1\}}), \end{aligned}$$

where we used $\pi_{\mathcal{C}} \circ s_i' = s_i \circ \pi$. Taking the Chern class of the above expression yields $\psi_i' = \pi^* \psi_i + D_{0,\{i,n+1\}}$. $\square$

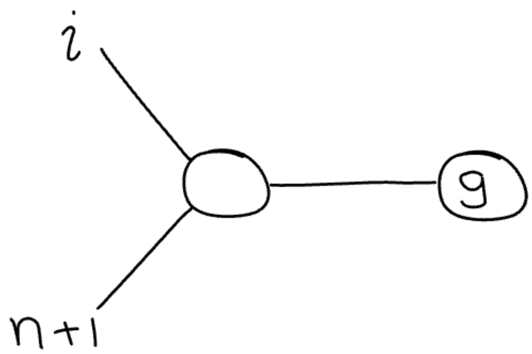

Figure 6: The dual graph of $D_{0,\{i,n+1\}} \subset \overline{\mathcal{M}}_{g,n+1}$.

Using the comparison lemma, we can express a power of a ψ-class in $\overline{\mathcal{M}}_{g,n+1}$ as follows.

$$\begin{aligned} \psi_i'^{\,d} &= \psi_i' \, (\pi^* \psi_i + D_{0,\{i,n+1\}})^{d-1} \\ &= \psi_i' \, (\pi^* \psi_i)^{d-1} + \psi_i' \, D_{0,\{i,n+1\}} (\cdots), \end{aligned}$$

where $(\cdots)$ represents other terms. The line bundle $\mathbb{L}_i'$ is trivial over $D_{0,\{i,n+1\}}$, since every point in $D_{0,\{i,n+1\}}$ is a rigid object. So the product $\psi_i' \, D_{0,\{i,n+1\}} = c_1(\mathbb{L}_i') \, D_{0,\{i,n+1\}} = c_1 \left(\mathbb{L}_i'|_{D_{0,\{i,n+1\}}} \right)$ is zero. This leaves us with

$$\begin{aligned} \psi_i'^{\,d} &= (\pi^* \psi_i + D_{0,\{i,n+1\}})(\pi^* \psi_i)^{d-1} \\ &= \pi^*(\psi_i^d) + \pi^*(\psi_i^{d-1}) D_{0,\{i,n+1\}}. \end{aligned}$$

Therefore the product $\psi_1'^{d_1}\cdots\psi_n'^{d_n}$ inside $\overline{\mathcal{M}}_{g,n+1}$ (excluding ψ'_{n+1}) is

$$\begin{aligned}&\left(\pi^*(\psi_1^{d_1})+\pi^*(\psi_1^{d_1-1})D_{0,\{1,n+1\}}\right)\cdots\left(\pi^*(\psi_n^{d_n})+\pi^*(\psi_n^{d_n-1})D_{0,\{n,n+1\}}\right)\\ &=\pi^*(\psi_1^{d_1}\cdots\psi_n^{d_n})+\sum_{i=1}^n\pi^*(\psi_1^{d_1}\cdots\psi_i^{d_i-1}\cdots\psi_n^{d_n})D_{0,\{i,n+1\}},\end{aligned}\tag{2.13}$$

where we have used $D_{0,\{i,n+1\}}\,D_{0,\{j,n+1\}}=0$ for $i\neq j$ because the divisors do not intersect. Using (2.13) we express the intersection product $\langle\tau_{d_1}\cdots\tau_{d_n}\tau_0\rangle$ on $\overline{\mathcal{M}}_{g,n+1}$ as a sum of similar products on $\overline{\mathcal{M}}_{g,n}$, by integrating along the fibres of $\pi:\overline{\mathcal{M}}_{g,n+1}\to\overline{\mathcal{M}}_{g,n}$ and applying the push-pull formula [EH16, Theorem 1.23]:

$$\begin{aligned}\langle\tau_{d_1}\cdots\tau_{d_n}\tau_0\rangle&=\int_{\overline{\mathcal{M}}_{g,n+1}}\psi_1'^{d_1}\cdots\psi_n'^{d_n}=\int_{\overline{\mathcal{M}}_{g,n}}\pi_*\left(\psi_1'^{d_1}\cdots\psi_n'^{d_n}\right)\\ &=\int_{\overline{\mathcal{M}}_{g,n}}\pi_*\left(\pi^*(\psi_1^{d_1}\cdots\psi_n^{d_n})\right)+\sum_{i=1}^n\int_{\overline{\mathcal{M}}_{g,n}}\pi_*\left(\pi^*(\psi_1^{d_1}\cdots\psi_i^{d_i-1}\cdots\psi_n^{d_n})D_{0,\{i,n+1\}}\right)\\ &=\int_{\overline{\mathcal{M}}_{g,n}}\psi_1^{d_1}\cdots\psi_n^{d_n}\,\pi_*(1')+\sum_{i=1}^n\int_{\overline{\mathcal{M}}_{g,n}}\psi_1^{d_1}\cdots\psi_i^{d_i-1}\cdots\psi_n^{d_n}\,\pi_*(D_{0,\{i,n+1\}}).\end{aligned}$$

In the sum it is implied that all terms with negative exponents are ignored. The fundamental class $1'=[\overline{\mathcal{M}}_{g,n+1}]\in H^0(\overline{\mathcal{M}}_{g,n+1})$ is the multiplicative identity. The first term above vanishes by definition because $\psi_1^{d_1}\cdots\psi_n^{d_n}\,\pi_*(1')$ is not a top intersection class. For the second term, we have already argued that π contracts the genus 0 component of every point in $D_{0,\{i,n+1\}}$, making the latter isomorphic to $\overline{\mathcal{M}}_{g,n}$ via π. Hence $\pi_*(D_{0,\{i,n+1\}})=[\overline{\mathcal{M}}_{g,n}]=1$, and we are left with

$$\int_{\overline{\mathcal{M}}_{g,n+1}}\psi_1'^{d_1}\cdots\psi_n'^{d_n}=\sum_{i=1}^n\int_{\overline{\mathcal{M}}_{g,n}}\psi_1^{d_1}\cdots\psi_i^{d_i-1}\cdots\psi_n^{d_n},\tag{2.14}$$

or equivalently

$$\langle\tau_{d_1}\cdots\tau_{d_n}\tau_0\rangle=\sum_{i=1}^n\langle\tau_{d_1}\cdots\tau_{d_i-1}\cdots\tau_{d_n}\rangle.\tag{2.15}$$

Equations (2.14) and (2.15) are called the *string equation*. They allow one to compute all intersection numbers of ψ-classes on $\overline{\mathcal{M}}_{0,n}$ inductively.

Proposition 2.6.2. *For any $n\geq 3$ and any n-tuple of nonnegative integers $(d_1,\dots,d_n)$ such that $d_1+\cdots+d_n=\dim\overline{\mathcal{M}}_{0,n}=n-3$:*

$$\int_{\overline{\mathcal{M}}_{0,n}}\psi_1^{d_1}\cdots\psi_n^{d_n}=\binom{n-3}{d_1,\dots,d_n}=\frac{(n-3)!}{d_1!\cdots d_n!}.$$

Proof. By induction on n. The $n=3$ case is $\langle\tau_0^3\rangle=1$. Assuming the result is true for n, consider

$$\int_{\overline{\mathcal{M}}_{0,n+1}} \psi_1^{d_1} \cdots \psi_n^{d_n} \psi_{n+1}^{d_{n+1}}.$$

Notice that $d_1 + \cdots + d_{n+1} = \dim \overline{\mathcal{M}}_{0,n+1} = n - 2$ implies that $d_{n+1} = 0$ without loss of generality. Then, using the string equation and the inductive hypothesis,

$$\begin{aligned}\int_{\overline{\mathcal{M}}_{0,n+1}} \psi_1^{d_1} \cdots \psi_n^{d_n} &= \sum_{i=1}^{n} \int_{\overline{\mathcal{M}}_{0,n}} \psi_1^{d_1} \cdots \psi_i^{d_i - 1} \cdots \psi_n^{d_n} = \sum_{i=1}^{n} \binom{n-3}{d_1, \ldots, d_i - 1, \ldots, d_n} \\ &= \binom{n-3}{d_1, \ldots, d_n} \sum_{i=1}^{n} d_i = \binom{n-2}{d_1, \ldots, d_n} = \binom{n-2}{d_1, \ldots, d_n, d_{n+1}},\end{aligned}$$

where we multiplied the top and bottom of the i^{th} fraction by d_i in the third equality. □

For genus $g = 1$, the first intersection product is

$$\langle \tau_1 \rangle = \int_{\overline{\mathcal{M}}_{1,1}} \psi_1 = \frac{1}{24}. \tag{2.16}$$

For two different proofs of this fact, see [Vak08, 3.13], [Ric22, Theorem 12.4.3], or [Zvo14, Proposition 2.26]. The factor of $\frac{1}{12}$ in $\frac{1}{24}$ arises because a rational elliptic fibration has 12 singular fibres, or because there exists an elliptic modular form of weight 12 with a simple zero at the cusp. The additional factor of $\frac{1}{2}$ is due to the fact that elliptic curves have two automorphisms, the identity and the involution. Obtaining the analogue of Proposition 2.6.2 for intersections of ψ-classes in $\overline{\mathcal{M}}_{1,n}$ follows a similar procedure to the one for $\overline{\mathcal{M}}_{0,n}$, which we outline below.

Denote by $f : \overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{C}}_{g,n}$ the isomorphism (2.7) between the forgetful morphism and the universal curve over $\overline{\mathcal{M}}_{g,n}$. Similarly to Lemma 2.6.1, we compute the 'discrepancy' between the line bundles $\mathbb{L}'_{n+1}$ and $f^*\mathbb{L}$ over $\overline{\mathcal{M}}_{g,n+1}$. Noticing that $f = s_i \circ \pi = \pi_{\mathcal{C}} \circ s'_i$ when restrictred to $D_{0,\{i,n+1\}} \subset \overline{\mathcal{M}}_{g,n+1}$, a similar argument to the lemma gives

$$\mathbb{L}'_{n+1} \cong f^*\mathbb{L} \otimes \bigotimes_{i=1}^{n} \mathcal{O}(D_{0,\{i,n+1\}}).$$

Therefore the degree of $\mathbb{L}'_{n+1}$ along the fibres of $\pi : \overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{M}}_{g,n}$ is the sum of the corresponding degrees of $\mathbb{L}$ and each $\mathcal{O}(D_{0,\{i,n+1\}})$, which is $2g - 2 + n$. Hence

$$\pi_*(\psi_{n+1}) = (2g - 2 + n)[\overline{\mathcal{M}}_{g,n}]. \tag{2.17}$$

As a side remark, notice that the left-hand side of (2.17) is the κ-class κ_0, which was defined in (2.8). Multiplying equation (2.13) for the intersection product

${\psi'_1}^{d_1} \cdots {\psi'_n}^{d_n}$ in $\overline{\mathcal{M}}_{g,n+1}$ by ψ'_{n+1} we obtain

$$
{\psi'_1}^{d_1} \cdots {\psi'_n}^{d_n}\, \psi'_{n+1} = \pi^*(\psi_1^{d_1} \cdots \psi_n^{d_n})\, \psi'_{n+1},
$$

since $D_{0,\{i,n+1\}}\psi_{n+1} = 0$ for every $1 \leq i \leq n$. Therefore, integrating along the fibres of $\pi : \overline{\mathcal{M}}_{g,n+1} \to \overline{\mathcal{M}}_{g,n}$ and using the push-pull formula,

$$
\begin{aligned}
\langle \tau_{d_1} \cdots \tau_{d_n} \tau_1 \rangle &= \int_{\overline{\mathcal{M}}_{g,n+1}} {\psi'_1}^{d_1} \cdots {\psi'_n}^{d_n}\, \psi'_{n+1} = \int_{\overline{\mathcal{M}}_{g,n}} \pi_* \left({\psi'_1}^{d_1} \cdots {\psi'_n}^{d_n}\, \psi'_{n+1} \right) \\
&= \int_{\overline{\mathcal{M}}_{g,n}} \pi_* \left(\pi^*(\psi_1^{d_1} \cdots \psi_n^{d_n})\, \psi'_{n+1} \right) = \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n}\, \pi_*(\psi'_{n+1}).
\end{aligned}
$$

Applying (2.17) one obtains

$$
\int_{\overline{\mathcal{M}}_{g,n+1}} {\psi'_1}^{d_1} \cdots {\psi'_n}^{d_n}\, \psi'_{n+1} = (2g - 2 + n) \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n}, \tag{2.18}
$$

or equivalently

$$
\langle \tau_{d_1} \cdots \tau_{d_n} \tau_1 \rangle = (2g - 2 + n) \langle \tau_{d_1} \cdots \tau_{d_n} \rangle. \tag{2.19}
$$

Equation (2.18) and (2.19) are known as the *dilaton equation*. Using both the string and dilaton equations, we can recursively compute top intersections of ψ-classes on $\overline{\mathcal{M}}_{1,n}$ starting from the initial condition $\langle \tau_1 \rangle = \frac{1}{24}$. This is because for every such intersection product

$$
\int_{\overline{\mathcal{M}}_{1,n}} \psi_1^{d_1} \cdots \psi_n^{d_n}, \tag{2.20}
$$

the dimension condition $d_1 + \cdots + d_n = \dim \overline{\mathcal{M}}_{1,n} = n$ ensures that $d_n \leq 1$ without loss of generality. One also obtains an expression for (2.20) similar to the one in Proposition 2.6.2. It is [LZ04, Proposition 4.6.11]

$$
\int_{\overline{\mathcal{M}}_{1,n}} \psi_1^{d_1} \cdots \psi_n^{d_n} = \frac{1}{24} \binom{n}{d_1, \ldots, d_n} \left(1 - \sum_{k=1}^{n} \frac{(k-2)!(n-k)!}{n!} e_k(d_1, \ldots, d_n) \right),
$$

where e_k is the k^{th} symmetric function

$$
e_k(d_1, \ldots, d_n) = \sum_{i_1 < \cdots < i_k} d_{i_1} \cdots d_{i_k}.
$$

From this discussion, one can deduce that the more relations one has such as the string and dilaton equations, the higher up in genus g one can go to compute intersections of ψ-classes recursively. A useful way of storing all this information is through a generating function. Let $t_* = (t_0, t_1, t_2 \ldots)$ be formal variables, and define

$$F(t_*) = \sum_{(d_1,\ldots,d_n)} \langle \tau_{d_1} \cdots \tau_{d_n} \rangle \frac{t_{d_1} \cdots t_{d_n}}{|\mathrm{Aut}(d_1,\ldots,d_n)|}.$$

This generating function was first introduced in [Wit90]. The sum runs over all positive integers n and all n-tuples of nonnegative integers $(d_1,\ldots,d_n)$, and we do not sum over the genus g since it is determined by $3g-3+n = \sum_{j=1}^{n} d_j$. By $\mathrm{Aut}(d_1,\ldots,d_n)$ we mean the group of automorphisms of $(d_1,\ldots,d_n)$, so that $|\mathrm{Aut}(d_1,\ldots,d_n)| = \prod_{i=0}^{\infty} n_i!$, where $n_i = |\{j : d_j = i\}|$. Therefore one can express F as a sum over $(n_0, n_1, \ldots)$ where only finitely many n_i are nonzero:

$$F(t_*, \ldots) = \sum_{(n_0,n_1,\ldots)} \langle \tau_0^{n_0} \tau_1^{n_1} \cdots \rangle \prod_{i=0}^{\infty} \frac{t_i^{n_i}}{n_i!}.$$

This time the genus is determined by $3g-3 = \sum_{i=0}^{\infty} n_i(i-1)$.

As a first example of the usefulness of this generating function, we show that the string equation (2.15) is equivalent to the following partial differential equation for F:

$$\frac{\partial F}{\partial t_0} = \frac{t_0^2}{2} + \sum_{i=0}^{\infty} t_{i+1} \frac{\partial F}{\partial t_i}. \tag{2.21}$$

Firstly, the term containing t_0^2 on the left is $\frac{\partial}{\partial t_0} \langle \tau_0^3 \rangle \frac{t_0^3}{3!} = \langle \tau_0^3 \rangle \frac{t_0^2}{2}$. Therefore (2.21) says that $\langle \tau_0^3 \rangle = 1$, as expected. In general, the term containing the monomial $\frac{t_{d_1} \cdots t_{d_n}}{|\mathrm{Aut}(d_1,\ldots,d_n)|} = \prod_{i=0}^{\infty} \frac{t_i^{n_i}}{n_i!}$ on the left is

$$\begin{aligned} \frac{\partial}{\partial t_0} \left(\langle \tau_0^{n_0+1} \tau_1^{n_1} \tau_2^{n_2} \cdots \rangle \frac{t_0^{n_0+1}}{(n_0+1)!} \prod_{i=1}^{\infty} \frac{t_i^{n_i}}{n_i!} \right) &= \langle \tau_0\, \tau_0^{n_0} \tau_1^{n_1} \tau_2^{n_2} \cdots \rangle \prod_{i=0}^{\infty} \frac{t_i^{n_i}}{n_i!} \\ &= \langle \tau_0\, \tau_{d_1} \cdots \tau_{d_n} \rangle \frac{t_{d_1} \cdots t_{d_n}}{|\mathrm{Aut}(d_1,\ldots,d_n)|}. \end{aligned} \tag{2.22}$$

On the right, the coefficient of the same monomial comes from terms with one more power of t_i and one less power of t_{i+1}, for every i:

$$\begin{aligned} &\sum_{i=0}^{\infty} t_{i+1} \frac{\partial}{\partial t_i} \left(\langle \cdots \tau_i^{n_i+1}\, \tau_{i+1}^{n_{i+1}-1} \cdots \rangle \cdots \frac{t_i^{n_i+1}}{(n_i+1)!} \frac{t_{i+1}^{n_{i+1}-1}}{(n_{i+1}-1)!} \cdots \right) \\ &= \sum_{i=0}^{\infty} n_{i+1} \langle \cdots \tau_i^{n_i+1}\, \tau_{i+1}^{n_{i+1}-1} \cdots \rangle \prod_{k=0}^{\infty} \frac{t_k^{n_k}}{n_k!} \\ &= \sum_{j=1}^{n} \langle \cdots \tau_{d_j-1} \cdots \rangle \frac{t_{d_1} \cdots t_{d_n}}{|\mathrm{Aut}(d_1,\ldots,d_n)|}. \end{aligned} \tag{2.23}$$

Comparing (2.22) and (2.23) gives the string equation (2.15). By a similar procedure, one can show that the dilaton equation (2.19) is equivalent to

$$\frac{\partial F}{\partial t_1} = \frac{1}{3}\sum_{i=0}^{\infty}(2i+1)t_i\frac{\partial F}{\partial t_i} + \frac{1}{24}.$$

A surprising fact is that the dilaton PDE can be deduced if one assumes that F obeys the string equation and $U = \frac{\partial^2 F}{\partial t_0^2}$ obeys the KdV equations, a particular system of partial differential equations (see [Koc01, Lemma 3.4.2]). This was one of the pieces of evidence that led Witten to make his conjecture [Wit90], which was later proved in [Kon92].

Theorem 2.6.3 (Witten-Kontsevich)**.**

(1) $U(t_*) := \frac{\partial^2 F}{\partial t_0^2}$ obeys the KdV equations:

$$\frac{\partial U}{\partial t_n} = \frac{\partial R_{n+1}}{\partial t_0}, \quad n \geq 0, \tag{2.24}$$

where the R_n are polynomials in $U, \frac{\partial U}{\partial t_0}, \frac{\partial^2 U}{\partial t_0^2}, \cdots$ determined inductively by

$$R_1 = U, \qquad \frac{\partial R_{n+1}}{\partial t_0} = \frac{1}{2n+1}\left(\frac{\partial U}{\partial t_0}R_n + 2U\frac{\partial R_n}{\partial t_0} + \frac{1}{4}\frac{\partial^3 R_n}{\partial t_0^3}\right).$$

(2) The generating function $F(t_*)$ obeys the string equation (2.21).

We have already proved (2). It is not immediately apparent how (1) and (2) uniquely determine F, so we explain this briefly. Firstly, the string equation provides us with the initial datum $U(t_0, 0, 0, \dots) = t_0$, which is equivalent to $\langle \tau_0^3 \rangle = 1$. This, together with the KdV equations, completely determines U (see Theorem 4.1.3). To show that F is also completely determined, it suffices to show that every $W := \langle \tau_{d_1} \cdots \tau_{d_n} \rangle$ can be recovered from knowledge of U and the string equation. When $n = 1$ we have

$$W = \langle \tau_{3g-2} \rangle = \langle \tau_{3g}\tau_0^2 \rangle = \left.\frac{\partial U}{\partial t_{3g}}\right|_{t_i=0},$$

where the second equality is obtained by applying the string equation (2.15) twice. For general $(d_1, \dots, d_n)$, one can assume $1 \leq d_j \leq 3g-2$ for each j. Then one inducts on $r := \max_j\{d_j\}$. The base case is $r = 3g-2$, which we have done above.

To deduce the $r-1$ case, assume $d_1 = r-1$ and introduce the known quantity

$$W' = \langle \tau_{d_1+2}\, \tau_{d_2} \cdots \tau_{d_n} \tau_0^2 \rangle = \left. \frac{\partial^n U}{\partial t_{d_1+2} \partial t_{d_2} \cdots \partial t_{d_n}} \right|_{t_i=0}.$$

If W' contains no τ_0 and τ_1 factors, applying the string equation twice expresses W' as a sum of W and other objects that are known by the inductive hypothesis. If W' does contain such factors, this procedure needs to be repeated a finite number of times with other suitable known objects $W'', W''', \ldots$, because other factors of τ_0 will appear.

In any case, knowledge of U and the string equation suffices to compute each intersection product $\langle \tau_{d_1} \cdots \tau_{d_n} \rangle$. Hence the statement of the Witten-Kontsevich theorem determines F uniquely. This establishes the surprising fact that the topology of the moduli space of curves encoded in its tautological ring is governed by an integrable hierarchy of partial differential equations. We now dedicate the next two sections to developing the necessary tools to understand the proof of this result given in [KL07].

3 Hurwitz theory

3.1 Hurwitz numbers

The proof of the Witten-Kontsevich theorem proposed in [KL07] relates the generating function

$$F(t_*) = \sum_{(d_1,\dots,d_n)} \langle \tau_{d_1} \cdots \tau_{d_n} \rangle \frac{t_{d_1} \cdots t_{d_n}}{|\mathrm{Aut}(d_1,\dots,d_n)|}$$

for top intersection products of ψ-classes in $\overline{\mathcal{M}}_{g,n}$ to another generating function, which encodes Hurwitz numbers. We introduce these now. For proofs of the subsequent statements about Riemann surfaces, see [CM16]. Much of this entire section is adapted from [CM16].

Consider two connected, compact Riemann surfaces X, Y and $f : X \to Y$ a non-constant holomorphic map between them. For every $x \in X$ there are charts centred at x and $f(x)$ such that the local form of f in these charts is $z \mapsto z^{k_x}$ for some integer $k_x \geq 1$. This integer is called the *ramification index* of f at x. If $k_x = 1$ then f is said to be *unramified* at x. If $k_x \geq 2$ then x is called a *ramification point* of f. The *ramification locus* $R \subset X$ is the subset of all ramification points of f, and it is finite. The image of R is the *branch locus* $B \subset Y$, and given any two points $y_1, y_2 \in Y \setminus B$ outside of the branch locus, the cardinalities of the fibers at those points agree: $|f^{-1}(y_1)| = |f^{-1}(y_2)|$. The degree d of f is defined as the cardinality of these fibers. For a branch point $y \in B$ we have $|f^{-1}(y)| < d$ instead.

Now for any point $y \in Y$, consider its preimage $f^{-1}(y) = \{x_1, \dots, x_\ell\} \subset X$ and the corresponding ramification indexes $(k_{x_1}, \dots, k_{x_\ell})$, ordered so that $k_{x_1} \geq k_{x_2} \geq \cdots \geq k_{x_\ell}$. The n-tuple $k = (k_{x_1}, \dots, k_{x_\ell})$ is called the *ramification profile* of f at y. It is always true that

$$\sum_{i=1}^{\ell} k_{x_i} = d. \tag{3.1}$$

Definition 3.1.1. Let d be a positive integer. A *partition* of d is a non-increasing tuple of positive integers $\mu = (\mu_1, \dots, \mu_\ell)$ such that $\sum_{i=1}^{\ell} \mu_i = d$, and this is indicated by the symbol $\mu \vdash d$. The sum d of the elements is called the *size* of μ and is denoted by $|\mu|$, and the number of elements ℓ is called the *length* of μ and is denoted by $\ell(\mu)$. The set of all partitions of d is denoted by $\mathcal{P}_d$, and $\mathcal{P} := \coprod_{d \geq 0} \mathcal{P}_d$.

With this definition in mind, equation (3.1) says that the ramification profile of f at any point in Y is a partition of the degree d of f. When $y \in Y \setminus B$ is not a branch point, the previous paragraph implies that $\ell = d$ and $(k_{x_1}, \dots, k_{x_d}) = (1, \dots, 1)$. In this case, f is said to be unramified over y. If, instead, $y \in B$ is a branch point, then $\ell < d$ and f is said to have ramification over y. We further distinguish between two types of ramification. If $(k_{x_1}, \ldots, k_{x_\ell}) = (2, 1, \ldots, 1)$ then f is said to have *simple ramification* over y. If not, f has *non-simple ramification* over y.

Given two such maps $f : X \to Y$ and $f' : X' \to Y$, we say that f and f' are isomorphic if there exists a biholomorphism $\varphi : X \to X'$ such that the diagram

$$\begin{tikzcd} X \arrow[rr, "\varphi"] \arrow[dr, "f"'] & & X' \arrow[dl, "f'"] \\ & Y & \end{tikzcd}$$

commutes. The group of automorphisms of f is denoted by $\mathrm{Aut}(f)$. We now present the definition of Hurwitz numbers, which count how many holomorphic maps there are between Riemann surfaces of a certain genus with prescribed ramification profiles, appropriately weighted by the number of automorphisms of such maps.

Definition 3.1.2. Let Y be a connected, compact Riemann surface of genus h with n marked points $b_1, \ldots, b_n \in Y$. Let $\mu_1, \ldots, \mu_n \in \mathcal{P}_d$ be partitions of a positive integer d, and let $g \geq 0$. The *connected Hurwitz number* for the data $g, h, d, \mu_1, \ldots, \mu_n$ is

$$H_{g \xrightarrow{d} h}(\mu_1, \ldots, \mu_n) = \sum_{[f]} \frac{1}{|\mathrm{Aut}(f)|}.$$

The sum runs over all isomorphism classes of holomorphic maps $f : X \to Y$, where X is a connected, compact Riemann surface of genus g, the branch locus of f is $\{b_1, \ldots, b_n\} \subset Y$ and the ramification profile of f at b_i is μ_i.

The first to discuss these numbers was Hurwitz in his 1891 paper [Hur91], where he wished to find a systematic way of computing all such numbers. When we allow the source curve X to be possibly disconnected in the above definition, we denote the corresponding disconnected Hurwitz number by

$$H^{\bullet}_{g \xrightarrow{d} h}(\mu_1, \ldots, \mu_n).$$

Note that the genus of a disconnected Riemann surface is determined by the additivity of the Euler characteristic. In other words, if X has connected components

$C_1, \ldots, C_n$ of genus $g_1, \ldots, g_n$, then the genus g of X is determined by

$$\begin{aligned} 2 - 2g &= \sum_{i=1}^{n} (2 - 2g_i) \\ \implies g &= 1 - n + \sum_{i=1}^{n} g_i. \end{aligned} \tag{3.2}$$

In order for a Hurwitz number to be well-defined, its data needs to satisfy the Riemann-Hurwitz formula

$$2 - 2g = d(2 - 2h) - \sum_{x \in R} (k_x - 1),$$

which in this case means

$$2 - 2g = d(2 - 2h) - nd + \sum_{i=1}^{n} \ell(\mu_i). \tag{3.3}$$

Hence the genus g does not need to be specified in the data because it is completely determined by the other pieces of data. By definition, a Hurwitz number whose data does not satisfy (3.3) is equal to 0.

In general, the Hurwitz number for a given data will be a rational number rather than just an integer, because we are weighting each isomorphism class of maps by its automorphism group. This might seem counter-intuitive, since the objective is to ‘count’ or ‘enumerate’ how many maps there are between Riemann surfaces. Nevertheless, there are several reasons for which it is important to retain the information about automorphisms at the cost of obtaining rational numbers. One of these reasons is the connection between Hurwitz numbers and the moduli space of curves. In Section 2.2, we outlined how the orbifold structure of moduli spaces is due to the presence of non-trivial automorphisms of curves. This is also the reason why we consider coefficients in $\mathbb{Q}$ rather than $\mathbb{Z}$ when studying the cohomology ring of $\overline{\mathcal{M}}_{g,n}$. Another reason for why we wish to take into account automorphisms will become apparent in the next section, where we show that computing Hurwitz numbers essentially boils down to counting equivalence classes in symmetric groups subject to certain conditions.

3.2 Monodromy representations

We now give a simple example of a computation of a Hurwitz number.

Example 3.2.1. Take $g = 0$ and $Y = \mathbb{P}^1$ with branch points $b_1 = 0, b_2 = \infty$ and ramification profiles $\mu_1 = \mu_2 = (d)$. Hence we wish to compute $H_{0 \xrightarrow{d} 0}((d),(d))$. Firstly, the source curve X is isomorphic to $\mathbb{P}^1$ because its genus is 0, so we can simply take it to be $\mathbb{P}^1$. A Hurwitz cover for this data is $p : \mathbb{P}^1 \to \mathbb{P}^1, x \mapsto x^d$. We now show that any other Hurwitz cover $f : \mathbb{P}^1 \to \mathbb{P}^1$ for this data is isomorphic to p. Since f is a rational function it is of the form

$$f(x) = a\,\frac{(x - x_1)^d}{(x - x_2)^d},$$

for some $a, x_1, x_2 \in \mathbb{C}$ (assuming $x_1 \neq \infty, x_2 \neq \infty$). Let η be a d^{th} root of a, and $\varphi : \mathbb{P}^1 \to \mathbb{P}^1$ be given by

$$\varphi(x) = \eta\,\frac{x - x_1}{x - x_2}.$$

Then the diagram

$$\begin{tikzcd} \mathbb{P}^1 \arrow[rr, "\varphi"] \arrow[dr, "f"'] & & \mathbb{P}^1 \arrow[dl, "p"] \\ & \mathbb{P}^1 & \end{tikzcd}$$

commutes, so f is isomorphic to p. By replacing f with p in the above discussion, we see that $\mathrm{Aut}(p)$ is the group of d^{th} roots of unity. Hence

$$H_{0 \xrightarrow{d} 0}((d),(d)) = \frac{1}{|\mathrm{Aut}(p)|} = \frac{1}{d}.$$

Another example is the Hurwitz number counting genus g hyperelliptic covers of $\mathbb{P}^1$:

$$H_{g \xrightarrow{2} 0}((2)^{2g+2}) = \frac{1}{2}.$$

In this case, every ramification profile is simple and the Riemann-Hurwitz formula forces there to be $2g + 2$ ramification points. One can show that only one Hurwitz cover contributes to the sum, and one uses the Riemann existence theorem to construct such a hyperelliptic cover.

Computing further examples with more exotic data becomes more challenging without a systematic framework to understand ramified covers algebraically. To begin developing such a framework, start with a degree d ramified covering $f : X \to Y$ as before. Take a basepoint $y_0 \in Y \setminus B$ together with a loop $\gamma : [0,1] \to Y \setminus B$ based at y_0. For any $x \in f^{-1}(y_0)$, the loop γ lifts to a unique path $\tilde{\gamma}_x : [0,1] \to X$ satisfying $\tilde{\gamma}_x(0) = x$ and $f \circ \tilde{\gamma}_x = \gamma$. Notice that $\tilde{\gamma}_x(1) \in f^{-1}(y_0)$ because $\gamma(1) = y_0$, so we get a map

$$\begin{aligned} \tilde{\sigma}_\gamma : \quad f^{-1}(y_0) &\longrightarrow f^{-1}(y_0), \\ x &\longmapsto \tilde{\gamma}_x(1). \end{aligned}$$

This is a bijection due to the existence and uniqueness of such path lifts. Since the basepoint y_0 does not belong to the branch locus $B \subset Y$, there is a bijection $L : f^{-1}(y_0) \to \{1, \dots, d\}$, called a y_0-labeling.

Definition 3.2.2. The pair consisting of a degree d ramified covering $f : X \to Y$ and a y_0-labeling $L : f^{-1}(y_0) \to \{1, \dots, d\}$ is called a *y_0-labeled map*. Two y_0-labeled maps $(f : X \to Y, L)$ and $(f' : X' \to Y, L')$ are isomorphic if there is a biholomorphism $\varphi : X \to X'$ such that the diagrams

$$\begin{tikzcd} X \arrow[rr, "\varphi"] \arrow[dr, "f"'] & & X' \arrow[dl, "f'"] \\ & Y & \end{tikzcd} \qquad \begin{tikzcd} f^{-1}(y_0) \arrow[rr, "\varphi"] \arrow[dr, "L"'] & & (f')^{-1}(y_0) \arrow[dl, "L'"] \\ & \{1, \dots, d\} & \end{tikzcd}$$

commute.

The composite map

$$\sigma_\gamma := L \circ \tilde{\sigma}_\gamma \circ L^{-1} : \{1, \dots, d\} \to \{1, \dots, d\}$$

is a permutation in the symmetric group S_d. Of course, it is essential that σ_γ be independent of the choice of labeling, or equivalently, invariant under isomorphisms of y_0-labeled maps. This is indeed the case: let (f, L) and (f', L') be isomorphic y_0-labeled maps as above, let $\gamma : [0, 1] \to Y \setminus B$ be a loop centred at y_0, and let σ_γ and σ'_γ be the resulting permutations. The lift of γ to $\varphi(x) \in (f')^{-1}(y_0)$ is $\tilde{\gamma}'_{\varphi(x)} = \varphi \circ \tilde{\gamma}_x$ for every $x \in f^{-1}(y_0)$. Therefore

$$\tilde{\sigma}'_\gamma(\varphi(x)) = \tilde{\gamma}'_{\varphi(x)}(1) = \varphi(\tilde{\gamma}_x(1)) = \varphi(\tilde{\sigma}_\gamma(x)),$$

so $\tilde{\sigma}'_\gamma = \varphi \circ \tilde{\sigma}_\gamma \circ \varphi^{-1}$. Then, using $\varphi|_{f^{-1}(y_0)} = (L')^{-1} \circ L$,

$$\begin{aligned} \sigma'_\gamma &= L' \circ \tilde{\sigma}'_\gamma \circ (L')^{-1} = L' \circ \varphi \circ \tilde{\sigma}_\gamma \circ \varphi^{-1} \circ (L')^{-1} \\ &= L' \circ (L')^{-1} \circ L \circ \tilde{\sigma}_\gamma \circ L^{-1} \circ L' \circ (L')^{-1} = L \circ \tilde{\sigma}_\gamma \circ L^{-1} \\ &= \sigma_\gamma, \end{aligned}$$

as claimed.

Moreover, the homotopy invariance of the path lift $\tilde{\gamma}_x$ of γ implies that the permutation σ_γ only depends on the homotopy class of γ. Thus an isomorphism class of y_0-labeled maps (f, L) induces a well-defined map

$$\begin{array}{rcll} \Phi: & \pi_1(Y \setminus B, y_0) & \longrightarrow & S_d, \\ & [\gamma] & \longmapsto & \sigma_\gamma. \end{array}$$

The map Φ is a group homomorphism. This is because the lift of a concatenation $\gamma * \eta$ of loops at x is $\tilde{\gamma}_x * \tilde{\eta}_{\tilde{\gamma}_x(1)}$, which implies that $\tilde{\sigma}_{\gamma*\eta}(x) = \tilde{\eta}_{\tilde{\gamma}_x(1)}(1) = (\tilde{\sigma}_\eta \circ \tilde{\sigma}_\gamma)(x)$. Such a homomorphism is called a *monodromy representation*, and the key result which shows how information about the ramification profiles of $f : X \to Y$ is encoded in Φ is given now.

Lemma 3.2.3. *Let $f : X \to Y$ be a degree d ramified covering and $b \in B \subset Y$ a branch point with ramification profile $(k_1, \ldots, k_\ell)$. Let ρ be a simple loop based at $y_0 \in Y \setminus B$ winding once around b. Then the cycle type of $\Phi(\rho) \in S_d$ is $(k_1, \ldots, k_\ell)$.*

Proof. Denote $f^{-1}(b) = \{x_1, \ldots, x_\ell\} \subset X$, so that k_j is the ramification index at x_j. In what follows, refer to Figure 7. For each $j = 1, \ldots, \ell$ let $U_j \subset X$ and $V_j \subset Y$ be local charts with coordinates z and w respectively, centred at x_j and b such that the local form of f is $w = z^{k_j}$. Let $y_j \in Y$ be the point corresponding to $w = 1$, and α_j a path connecting y_0 to y_j. Moreover let β_j be the simple loop based at y_j winding once around b along the circle $\{w : |w| = 1\}$ in V_j. We investigate the permutation $\Phi(\beta_j) = \sigma_{\beta_j}$, which is determined by the map $\tilde{\sigma}_{\beta_j} : f^{-1}(y_j) \to f^{-1}(y_j)$. In the chart U_j the preimage $f^{-1}(y_j)$ is $\{z : z^{k_j} = 1\} = \left\{\exp\left(\frac{2\pi i}{k_j} m_j\right) : m_j = 0, 1, \ldots, k_j - 1\right\}$, the set of k_j^{th} roots of unity. For every $m_j = 0, 1 \ldots, k_j - 1$ the path $\tilde{\beta}_{m_j} : [0,1] \to U_j \subset X$ defined by $\tilde{\beta}_{m_j}(t) = \exp\left(\frac{2\pi i}{k_j}(m_j + t)\right)$ is the lift of β_j at $\exp\left(\frac{2\pi i}{k_j} m_j\right)$. Hence the map $\tilde{\sigma}_{\beta_j}$ is

$$\exp\left(\frac{2\pi i}{k_j} m_j\right) \longmapsto \tilde{\beta}_{m_j}(1) = \exp\left(\frac{2\pi i}{k_j}(m_j + 1)\right).$$

for every $j = 1, \ldots, \ell$. Upon the choice of labeling $L_j : f^{-1}(y_j) \to \{1, \ldots, k_j\}$ given by $\exp\left(\frac{2\pi i}{k_j} m_j\right) \mapsto m_j + 1$, the resulting permutation $\sigma_{\beta_j} = (12 \cdots k_j)$ has cycle type (k_j). Now notice that ρ is homotopic to $\alpha_j * \beta_j * \alpha_j^{-1}$ for every $j = 1, \ldots, \ell$, so $\Phi(\rho)$ has the same cycle type as σ_{β_j} when restricted to the k_j numbers in $\{1, \ldots, d\}$ picked out by the labeling L_j. Hence $\Phi(\rho)$ is the product of the disjoint cycles $\sigma_{\beta_1}, \ldots, \sigma_{\beta_\ell}$, which has cycle type $(k_1, \ldots, k_\ell)$. □

This result implies that the image of a monodromy representation Φ induced by a ramified cover $f : X \to Y$ is determined by the ramification profiles of f.

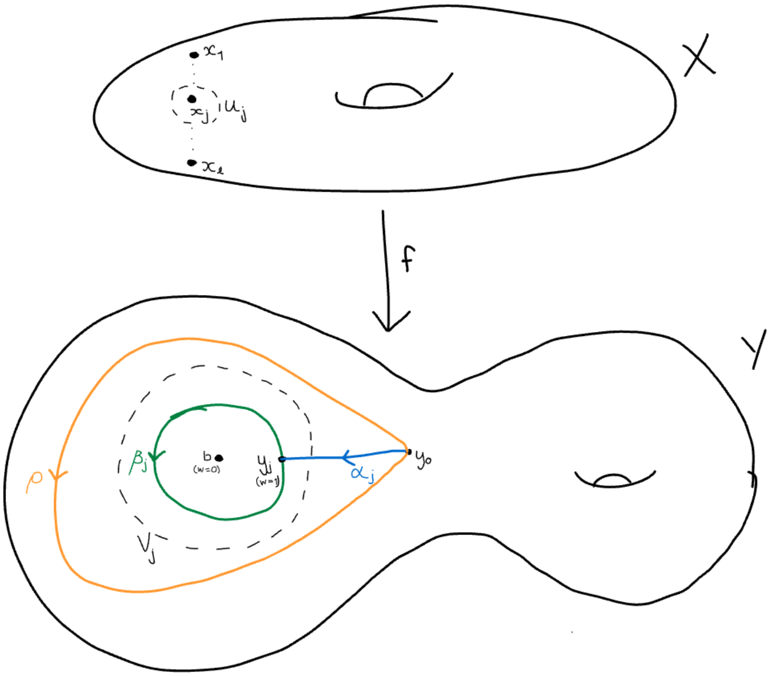

Figure 7: The proof of Lemma 3.2.3.

Definition 3.2.4. Let Y be a connected, compact Riemann surface of genus h with basepoint $y_0 \in Y$ and $B = \{b_1, \dots, b_n\}$, and let $\mu_1, \dots, \mu_n$ be partitions of a positive integer d. A *monodromy representation of type* $(h, d, \mu_1, \dots, \mu_n)$ is a group homomorphism $\Phi : \pi_1(Y \setminus B, y_0) \to S_d$ such that if ρ_i is the homotopy class of a simple loop around b_i, then $\Phi(\rho_i)$ has cycle type μ_i. If $\operatorname{im} \Phi$ acts transitively on $\{1, \dots, d\}$ then Φ is a *connected* monodromy representation.

So Lemma 3.2.3 essentially says that a degree d cover $f : X \to Y$ with Y a (connected) genus h Riemann surface and ramification profiles $\mu_1, \dots \mu_n$ induces a (connected) monodromy representation of type $(h, d, \mu_1, \dots, \mu_n)$. Remark that the genus g of the source curve X is not included in the data of the monodromy representation, since it is completely determined by the Riemann-Hurwitz formula (3.3). The set of all connected monodromy representations of type $(h, d, \mu_1, \dots, \mu_n)$ is denoted by $M_{h,d}(\mu_1, \dots, \mu_n)$, and the corresponding set for possibly disconnected representations is $M^{\bullet}_{h,d}(\mu_1, \dots, \mu_n)$.

We now have a natural map from the set of ramified coverings to the set of monodromy representations. We now show that one can go the other way. Given a monodromy representation $\Phi : \pi_1(Y \setminus B, y_0) \to S_d$ of type $(h, d, \mu_1, \dots, \mu_n)$, we wish to construct a degree d ramified cover $f : X \to Y$ with X of the correct genus and ramification profiles $\mu_1, \dots, \mu_n$ at the branch points $b_1, \dots, b_n \in B \subset Y$.

Firstly, we give a description of the fundamental group of an n-punctured surface $Y \setminus B$ of genus h. If $h = 0$ then Y is $\mathbb{P}^1 \cong S^2$, and the punctured sphere $S^2 \setminus B$ is homotopy equivalent to a flower graph with $n-1$ petals, which one may denote by the wedge sum $\bigvee_{i=1}^{n-1} S^1$. By the Seifert-Van Kampen theorem the fundamental group is the free group generated by $n-1$ elements, see for example [Hat01, 1.2]. We can present this more symmetrically by adding an extra generator and a relation:

$$\pi_1(S^2 \setminus B, y_0) = \langle \rho_1, \ldots, \rho_n \,|\, \rho_1 \cdots \rho_n \rangle, \tag{3.4}$$

where ρ_i represents a simple loop around b_i. If, instead, we take $B = \emptyset$ and Y a genus h surface, we may represent Y as an identification polygon with $4h$ sides given by $\alpha_1\beta_1\overline{\alpha}_1\overline{\beta}_1 \cdots \alpha_h\beta_h\overline{\alpha}_h\overline{\beta}_h$. Then the fundamental group is

$$\pi_1(Y, y_0) = \langle \alpha_1, \beta_1, \ldots, \alpha_h, \beta_h \,|\, [\alpha_1, \beta_1] \cdots [\alpha_h, \beta_h] \rangle, \tag{3.5}$$

where $[\alpha_k, \beta_k] = \alpha_k\beta_k\alpha_k^{-1}\beta_k^{-1}$, where α_k and β_k represent the loops around the k^{th} handle of the surface. For a discussion about this, see for example [Die08, 2.8]. Putting (3.4) and (3.5) together, we may represent an n-punctured genus h surface $Y \setminus B$ as the identification polygon $s_1\overline{s}_1 \cdots s_n\overline{s}_n\alpha_1\beta_1\overline{\alpha}_1\overline{\beta}_1 \cdots \alpha_h\beta_h\overline{\alpha}_h\overline{\beta}_h$ with n punctured vertices corresponding to the punctures. Therefore

$$\pi_1(Y \setminus B, y_0) = \langle \alpha_1, \beta_1, \ldots, \alpha_h, \beta_h, \rho_1, \ldots, \rho_n \,|\, [\alpha_1, \beta_1] \cdots [\alpha_h, \beta_h]\rho_1 \cdots \rho_n \rangle.$$

So a monodromy representation Φ of type $(h, d, \mu_1, \ldots, \mu_n)$ is equivalent to a choice of permutations $\Phi(\alpha_1), \Phi(\beta_1), \ldots, \Phi(\alpha_h), \Phi(\beta_h), \Phi(\rho_1), \ldots, \Phi(\rho_n) \in S_d$ such that each $\Phi(\rho_i)$ has cycle type μ_i and $[\Phi(\alpha_1), \Phi(\beta_1)] \cdots [\Phi(\alpha_h), \Phi(\beta_h)]\Phi(\rho_1) \cdots \Phi(\rho_n)$ equals the identity. In fact, we will often regard the sets of monodromy representations $M_{h,d}(\mu_1, \ldots, \mu_n)$ and $M^\bullet_{h,d}(\mu_1, \ldots, \mu_n)$ as *being* the sets of such collections of permutations, see for example Theorem 3.4.3.

To illustrate how one obtains a ramified covering from a monodromy representation of a certain type, we use an example.

Example 3.2.5. Let $Y = \mathbb{P}^1$ with basepoint y_0 and punctures $b_1, b_2, b_3 \in \mathbb{P}^1$. From the previous discussion, the fundamental group $\pi_1(Y \setminus \{b_1, b_2, b_3\}, y_0)$ is $\langle \rho_1, \rho_2, \rho_3 \,|\, \rho_1\rho_2\rho_3 \rangle$. Consider the monodromy representation given by $\Phi(\rho_1) = (123), \Phi(\rho_2) = (13), \Phi(\rho_3) = (12)$. We construct the source curve X as follows, see Figure 8. Choose another point $p \in \mathbb{P}^1$ and draw line segments from p to each b_i, so that $\mathbb{P}^1 \setminus B$ is homeomorphic to the identification polygon in the figure. Take three copies of the polygon labeled P_1, P_2, P_3 and consider the natural projection

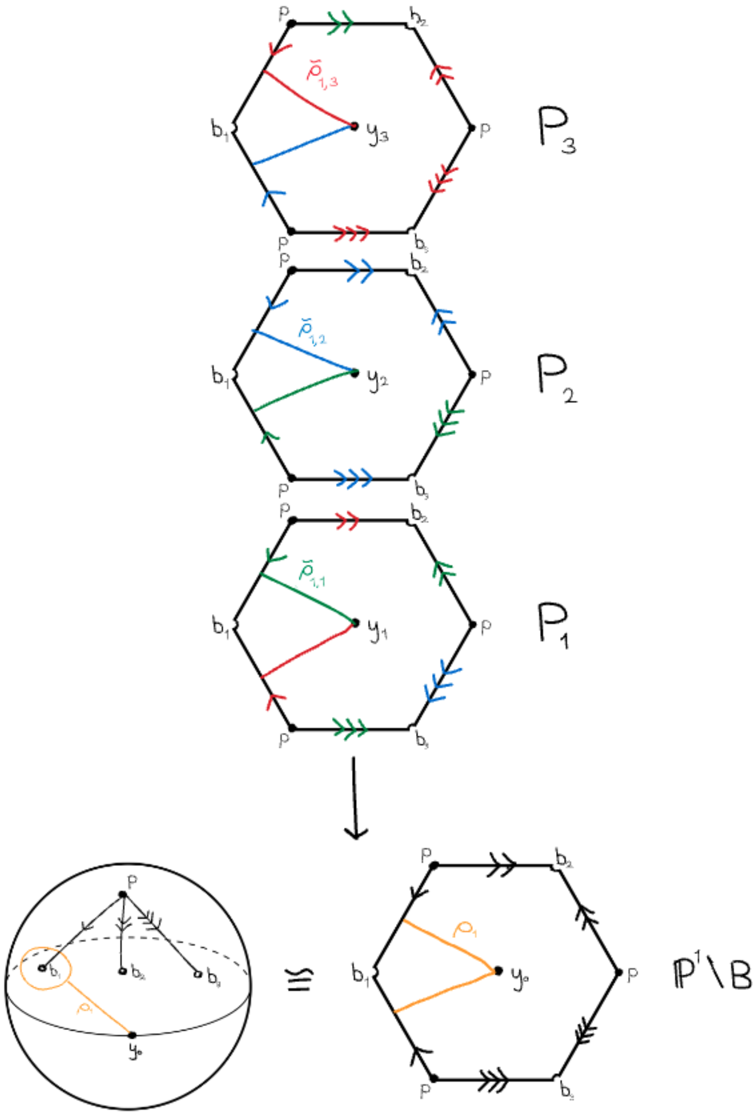

Figure 8: Constructing a Hurwitz cover from a monodromy representation.

$P_1 \sqcup P_2 \sqcup P_3 \to \mathbb{P}^1 \setminus B$. Let $y_i \in P_i$ be the three preimages of y_0 under this projection. Now use the permutations $\Phi(\rho_1), \Phi(\rho_2)$ and $\Phi(\rho_3)$ to dictate how the three polygons should be glued together. For example, consider $\Phi(\rho_1) = (123)$:

- Since $\Phi(\rho_1)$ sends $1 \mapsto 2$, the lift $\tilde{\rho}_{1,1}$ of ρ_1 starting at y_1 should end at y_2. Hence the top left side of P_1 and the bottom left side of P_2 should be identified (single green arrow).

- Since $\Phi(\rho_1)$ sends $2 \mapsto 3$, the lift $\tilde{\rho}_{1,2}$ of ρ_1 starting at y_2 should end at y_3. Hence the top left side of P_2 and the bottom left side of P_3 should be identified (single blue arrow).

- Since $\Phi(\rho_1)$ sends $3 \mapsto 1$, the lift $\tilde{\rho}_{1,3}$ of ρ_1 starting at y_3 should end at y_1. Hence the top left side of P_3 and the bottom left side of P_1 should be identified (single red arrow).

Repeating the procedure with $\Phi(\rho_2)$ and $\Phi(\rho_3)$ ensures that all the sides of the three polygons are identified according to Φ, as shown in the figure. Hence we obtain a topological surface $X^\circ = (P_1 \sqcup P_2 \sqcup P_3) / \sim$ and a topological cover $f^\circ : X^\circ \to \mathbb{P}^1 \setminus B$ induced by the projection. By Riemann's existence theorem, there is a unique compact Riemann surface X containing X° as a dense open subset such that f° extends to a holomorphic map $f : X \to Y$. By construction f is a ramified covering with branch points b_1, b_2, b_3 whose respective ramification indexes are given by the cycle types of $\Phi(\rho_1), \Phi(\rho_2), \Phi(\rho_3)$.

For a surface Y of genus h and an arbitrary monodromy representation Φ of type $(h, d, \mu_1, \ldots, \mu_n)$, one follows a similar procedure to construct the associated ramified cover. One represents the punctured surface $Y \setminus B$ as a $(4h+2n)$-sided identification polygon by drawing segments from an arbitrary point to the branch points. Then one takes d copies of this polygon and identifies their sides according to the permutations $\Phi(\rho_1), \ldots, \Phi(\rho_n)$. Note that for genus $h \neq 0$ one must also identify the sides labeled by α_i, β_i in (3.5) according to the monodromy representation.

Proposition 3.2.6. *Let Y be a Riemann surface of genus h with $B = \{b_1, \ldots, b_n\} \subset Y$, and let Φ be a monodromy representation of type $(h, d, \mu_1, \ldots, \mu_n)$. Then there is a y_0-labeled map $f : X \to Y$ whose branch locus is B and whose associated monodromy representation is Φ. This map is unique up to isomorphism of y_0-labeled maps.*

Given a Riemann surface Y of genus h, thanks to Lemma 3.2.3 and subsequently Proposition 3.2.6 there is a bijection between the set of isomorphism classes of (connected) y_0-labeled maps $f : X \to Y$ of degree d with branch locus B and ramification profiles $\mu_1, \ldots, \mu_n$, and the set of (connected) monodromy representations $\Phi : \pi_1(Y \setminus B, y_0) \to S_d$ of type $(h, d, \mu_1, \ldots, \mu_n)$. As a result, we arrive at the main theorem linking Hurwitz numbers $H_{g \xrightarrow{d} h}(\mu_1, \ldots, \mu_n)$ and the corresponding set of monodromy representations $M_{h,d}(\mu_1, \ldots, \mu_n)$.

Theorem 3.2.7. *Let Y be a Riemann surface of genus h with n marked points $b_1, \ldots, b_n \in Y$. The connected Hurwitz number for the data $h, d, \mu_1, \ldots, \mu_n$ is given by*

$$H_{g \xrightarrow{d} h}(\mu_1, \ldots, \mu_n) = \frac{1}{d!} |M_{h,d}(\mu_1, \ldots, \mu_n)|.$$

The corresponding disconnected Hurwitz number is

$$H^\bullet_{g \xrightarrow{d} h}(\mu_1, \ldots, \mu_n) = \frac{1}{d!}|M^\bullet_{h,d}(\mu_1, \ldots, \mu_n)|.$$

Proof. Pick a basepoint $y_0 \in Y$. The proof will be done for connected Hurwitz numbers, and the disconnected case follows analogously. By definition, we wish to compute the right-hand side of

$$H_{g \xrightarrow{d} h}(\mu_1, \ldots, \mu_n) = \sum_{[f]} \frac{1}{|\mathrm{Aut}(f)|},$$

where the sum runs over isomorphism classes of ramified covers $f : X \to Y$ satisfying the data. For any such f, let $S_f = \{L : f^{-1}(y_0) \to \{1, \ldots, d\}\}$ be the set of y_0-labelings. There is a free left group action of $\mathrm{Aut}(f)$ on S_f given by $\varphi \cdot L = L \circ \varphi^{-1}$. Notice that the y_0-labeled maps (f, L) and (f, L') are isomorphic if and only if L' lies in the orbit $\mathrm{Aut}(f) \cdot L$ of L, and by the Orbit-Stabilizer theorem the cardinality of $\mathrm{Aut}(f) \cdot L$ is simply $|\mathrm{Aut}(f)|$. Hence the number of isomorphism classes of y_0-labeled maps for the given map f is

$$m_f = \frac{|S_f|}{|\mathrm{Aut}(f)|} = \frac{d!}{|\mathrm{Aut}(f)|}.$$

As a result,

$$\sum_{[f]} \frac{1}{|\mathrm{Aut}(f)|} = \frac{1}{d!} \sum_{[f]} m_f.$$

Since isomorphic y_0-labeled maps give rise to the same monodromy representation, the number m_f is also equal to the number of distinct monodromy representations arising from f by different labelings of $f^{-1}(y_0)$. Hence, by Lemma 3.2.3 and Proposition 3.2.6, the sum $\sum_{[f]} m_f$ is equal to $|M_{h,d}(\mu_1, \ldots, \mu_n)|$. ☐

Theorem 3.2.7 allows one to easily compute many Hurwitz numbers. For example, all the connected and disconnected numbers for degree 3 covers of $\mathbb{P}^1$ are readily computed. We provide one such computation.

Example 3.2.8. Let $m, k \geq 1$ and consider the Hurwitz number

$$H^\bullet_{g \xrightarrow{3} 0}((3)^m, (2,1)^k).$$

The exponents of (3) and $(2,1)$ denote the number of ramification profiles of that type. Firstly, the Riemann-Hurwitz formula (3.3) forces k to be the even number $k = 2(g - m + 2) =: 2n$ and therefore the genus of the source is $g = m + n - 2$.

Secondly, the presence of a branch point with ramification profile (3) implies that the image of the associated monodromy representation will always contain a 3-cycle, which acts transitively on $\{1, 2, 3\}$. Hence the disconnected Hurwitz number is equal to the connected one $H_{g \xrightarrow{3} 0}((3)^m, (2,1)^{2n})$. By the theorem, this is equal to $\frac{1}{3!}|M_{0,3}((3)^m, (2,1)^{2n})|$. The number of distinct monodromy representations is simply the number of ways to pick 3-cycles $\rho_1, \ldots, \rho_m$ and 2-cycles $\tau_1, \ldots, \tau_{2n}$ in S_3 such that $\rho_1 \cdots \rho_m \tau_1 \cdots \tau_{2n}$ equals the identity. There are $2^m 3^{2n-1}$ ways of doing this, so

$$H^{\bullet}_{g \xrightarrow{3} 0}((3)^m, (2,1)^{2n}) = H_{g \xrightarrow{3} 0}((3)^m, (2,1)^{2n}) = \frac{1}{3!} 2^m 3^{2n-1} = 2^{m-1} 3^{2n-2}.$$

3.3 Simple Hurwitz potential

From now on we focus our attention on *simple* Hurwitz numbers, where the base curve is $\mathbb{P}^1$ and there is only one branch point with non-simple ramification. Without loss of generality, one can take this branch point to be $0 \in \mathbb{P}^1$. These numbers are

$$H_{g \xrightarrow{d} 0}(\mu, (2,1,\ldots,1)^b) \quad \text{and} \quad H^{\bullet}_{g \xrightarrow{d} 0}(\mu, (2,1,\ldots,1)^b), \tag{3.6}$$

where $\mu = (\mu_1, \ldots, \mu_\ell) \neq (2, 1 \ldots, 1) \in \mathcal{P}_d$ and b is the number of branch points with simple ramification. Notice that $d = |\mu|$ by definition, and $b = 2g - 2 + \ell + |\mu|$ by the Riemann-Hurwitz formula (3.3). We therefore simplify the notation for the simple Hurwitz numbers (3.6) to

$$H_g^b(\mu) \quad \text{and} \quad H_g^{\bullet,b}(\mu). \tag{3.7}$$

If we need to specify the individual components of the partition $\mu = (\mu_1, \ldots, \mu_\ell)$, we write $H_g^b(\mu_1, \ldots, \mu_\ell)$ and $H_g^{\bullet,b}(\mu_1, \ldots, \mu_\ell)$. We stress that in this case each μ_i is a component of a partition μ, rather than an individual partition as in Section 3.2. We write the corresponding sets of monodromy representations as

$$M^b(\mu) \quad \text{and} \quad M^{\bullet,b}(\mu). \tag{3.8}$$

Theorem 3.2.7 says that the numbers (3.7) and the cardinalities of (3.8) are related by multiplication by $1/|\mu|!$. Now, introduce the formal variables β and $p_* = (p_1, p_2, p_3, \ldots)$, and for a partition $\mu = (\mu_1, \ldots, \mu_\ell)$ denote the monomial $p_{\mu_1} \cdots p_{\mu_\ell}$ by the shorthand notation p_μ. Just as in Section 2.6, consider the following generating functions which encode all the connected and disconnected simple

Hurwitz numbers:

$$H(\beta; p_*) = \sum H_g^b(\mu)\, p_\mu \frac{\beta^b}{b!},$$
$$H^\bullet(\beta; p_*) = \sum H_g^{\bullet,b}(\mu)\, p_\mu \frac{\beta^b}{b!}.$$

These are called the connected and disconnected *simple Hurwitz potentials*, respectively. We take the sums over all nonnegatve integers $b \geq 0$ and $\ell \geq 0$ and all non-increasing ℓ-tuples $\mu = (\mu_1, \ldots, \mu_\ell)$ of nonnegative integers (for $\ell = 0$ the only choice is $\mu = (\emptyset)$). There is no need to sum over g because it is completely determined by b and $(\mu_1, \ldots, \mu_\ell)$ via the Riemann-Hurwitz formula $g = \frac{1}{2}(2 + b - |\mu| - \ell(\mu))$. Alternatively, one could sum over g and let b vary accordingly instead. The variables p_* parameterize the non-simple ramification data over $0 \in \mathbb{P}^1$, while the variable β keeps track of the number of simple branch points. Of course, one can define Hurwitz potentials for general Hurwitz numbers as in [CM16], with more variables to keep track of more pieces of data, but for our purposes the generating functions above will suffice.

For most of the rest of this section, we find an expansion for H and $H^\bullet$ in a particular basis of polynomials. Before we do this, we show how H and $H^\bullet$ are related. To gain some intuition first, consider an arbitrary cover associated to the *disconnected* simple Hurwitz number $H_g^{\bullet,b}(\mu)$. This consists of a disjoint union of connected covers $f_1 \sqcup \cdots \sqcup f_n : X_1 \sqcup \cdots \sqcup X_n \to \mathbb{P}^1$ such that:

- the genera $g_1, \ldots, g_n$ of the connected components $X_1, \ldots, X_n$ satisfy $1 - n + \sum_{i=1}^n g_i = g$ (see equation (3.2));
- each connected cover has only one non-simple branch point $0 \in \mathbb{P}^1$;
- the non-simple ramification profiles $\mu_1, \ldots, \mu_n \in \mathcal{P}$ of each cover are such that $(\mu_1; \ldots; \mu_n) = \mu$ (up to reordering);
- the number of simple branch points $b_1, \ldots, b_n$ for each cover satisfy $\sum_{i=1}^n b_i = b$.

Thus $H_g^{\bullet,b}(\mu)$ will be equal to the sum of all the possible products of *connected* Hurwitz numbers $\prod_i H_{g_i}^{b_i}(\mu_i)$ arising in this way, weighted by the automorphisms that arise from distributing the simple branch points and from permuting connected components with identical data. This is made precise in the following elegant result, which we adapt from [CM16, Theorem 10.2.1].

Proposition 3.3.1. *The connected and disconnected Hurwitz numbers are related* *by exponentiation:*

$$H^\bullet = e^H - 1.$$

Proof. Let $\mathfrak{s} = (g, b, \mu)$ with $\mu \in \mathcal{P}_{|\mu|}$ denote data for a simple Hurwitz number, and let $H_\mathfrak{s}$ and $H^\bullet_\mathfrak{s}$ denote the corresponding simple Hurwitz numbers $H^b_g(\mu)$ and $H^{\bullet,b}_g(\mu)$. The monomial associated to $\mathfrak{s}$ is

$$\mathrm{mon}(\mathfrak{s}) = p_{\mu_1} \cdots p_{\mu_\ell} \beta^b = p_\mu \beta^b,$$

so that $H_\mathfrak{s}$ and $H^\bullet_\mathfrak{s}$ are the coefficients of $\frac{1}{b!}\,\mathrm{mon}(\mathfrak{s})$ in H and $H^\bullet$ respectively. For another set of data $\mathfrak{s}' = (g', b', \mu')$ define

$$\mathfrak{s} + \mathfrak{s}' = (g + g' - 1, b + b', (\mu; \mu')),$$

so that $\mathrm{mon}(\mathfrak{s}+\mathfrak{s}') = \mathrm{mon}(\mathfrak{s})\,\mathrm{mon}(\mathfrak{s}')$. For a collection of such data $\mathfrak{s}_1, \ldots, \mathfrak{s}_n$ define the tuple

$$\vec{\mathfrak{s}} = (\mathfrak{s}_1^{k_1}, \ldots, \mathfrak{s}_n^{k_n}),$$

where $\mathfrak{s}_i^{k_i} = \overbrace{\mathfrak{s}_i, \ldots, \mathfrak{s}_i}^{k_i \text{ times}}$. Intuitively, this denotes a disjoint union of $N := \sum_{i=1}^n k_i$ connected curves whose Hurwitz data sums to $\sum_{i=1}^n k_i \mathfrak{s}_i =: |\vec{\mathfrak{s}}|$, which is the Hurwitz data for the entire disconnected curve. The disconnected Hurwitz number $H^\bullet_\mathfrak{s}$ for $\mathfrak{s} = (g, b, \mu)$ will therefore involve a sum over all tuples $\vec{\mathfrak{s}}$ such that $|\vec{\mathfrak{s}}| = \mathfrak{s}$, and products of connected Hurwitz numbers corresponding to such tuples:

$$H^\bullet_\mathfrak{s} = \sum_{|\vec{\mathfrak{s}}| = \mathfrak{s}} \binom{b}{b_1, \ldots, b_N} \frac{1}{k_1! \cdots k_n!} \prod_{i=1}^n (H_{\mathfrak{s}_i})^{k_i}.$$

The multinomial coefficient arises from the different ways in which one can distribute the simple branch points among the connected components, while the fraction after that arises from permuting the connected components with identical data. Hence the coefficient of $\mathrm{mon}(\mathfrak{s})$ in $H^\bullet$ is

$$\frac{1}{b!} H^\bullet_\mathfrak{s} = \sum_{|\vec{\mathfrak{s}}| = \mathfrak{s}} \frac{1}{b_1! \cdots b_N!} \frac{1}{k_1! \cdots k_n!} \prod_{i=1}^n (H_{\mathfrak{s}_i})^{k_i}. \tag{3.9}$$

On the other hand, since $|\vec{\mathfrak{s}}| = \mathfrak{s}$ implies that $\mathrm{mon}(|\vec{\mathfrak{s}}|) = \mathrm{mon}(\mathfrak{s})$, the coefficient of $\mathrm{mon}(\mathfrak{s})$ in

$$e^H - 1 = \sum_{N \geq 1} \frac{1}{N!} \left(\sum_{\mathfrak{s}'} \frac{1}{b!} H_{\mathfrak{s}'}\,\mathrm{mon}(\mathfrak{s}') \right)^N$$

is equal to

$$\sum_{|\vec{\mathfrak{s}}|=\mathfrak{s}} \frac{1}{N!}\binom{N}{k_1,\ldots,k_n}\prod_{i=1}^{n}\left(\frac{1}{b_i!}H_{\mathfrak{s}_i}\right)^{k_i}$$
$$=\sum_{|\vec{\mathfrak{s}}|=\mathfrak{s}} \frac{1}{k_1!\cdots k_n!}\frac{1}{b_1!\cdots b_N!}\prod_{i=1}^{k_i}(H_{\mathfrak{s}_i})^{k_i},$$

which is the same as (3.9). □

3.4 Cut-and-join

Now that we know that the disconnected simple Hurwitz potential is simply the exponential of the connected one, we work with $H^\bullet$. We develop recursive relations among disconnected simple Hurwitz numbers in the form of a partial differential equation for the generating function $H^\bullet$. This is done by analyzing what happens when one shrinks a loop surrounding two branch points, one of which is the branch point with non-simple ramification. From the algebraic point of view given us by Theorem 3.2.7, this amounts to analyzing what happens to the cycle type of a permutation in the symmetric group when composed with a transposition.

Lemma 3.4.1 (Cut)**.** *Let $\sigma \in S_d$ be a permutation with cycle type $\mu = (\mu_1, \ldots, \mu_\ell)$, written as a disjoint union of cycles $\sigma = \sigma_1 \cdots \sigma_\ell$. Let $\tau = (a\,b) \in S_d$ be a transposition. If a and b belong to the* same *cycle of σ, say σ_ℓ, then this cycle is* cut *in two upon composition with τ. In other words, $\tau\sigma$ has cycle type $\lambda = (\mu_1, \ldots, \mu_{\ell-1}, m', m'')$, where $m' + m'' = \mu_\ell$. If $m' \neq m''$ then there are μ_ℓ such transpositions τ giving rise to the cycle type λ, while if $m' = m''$ there are $\mu_\ell/2$.*

Proof. Write σ_ℓ as the cycle $(a\,a_2\,a_3 \cdots a_{m'}\,b\,a_{m'+2} \cdots a_{\mu_\ell})$ for some $2 \leq m' \leq \mu_\ell - 1$. Then

$$\tau\sigma_\ell = (a\,b)(a\,a_2\,a_3 \cdots a_{m'}\,b\,a_{m'+2} \cdots a_{\mu_\ell}) = (a\,a_2\,a_3 \cdots a_{m'})(b\,a_{m'+2} \cdots a_{\mu_\ell})$$

has cycle type (m', m'') where $m'' = \mu_\ell - m'$. If $m' \neq m''$, we may suppose $m' < m''$. Then the μ_ℓ transpositions that give rise to this cycle type are $(a\,b), (a_2\,a_{m'+2})$, $(a_3\,a_{m'+3}), \ldots, (a_{m'}\,a_{2m'}), (j\,a_{2m'+1}), \ldots, (a_{m''}\,a_{\mu_\ell}), (a_{m''+1}, a_1), \ldots, (a_{\mu_\ell}, a_{m'})$. If $m' = m'' = \mu_\ell/2$, then the list stops at $(a_{m'}, a_{2m'})$ since the subsequent cycles will already have been counted, so there are only $\mu_\ell/2$. □

Lemma 3.4.2 (Join)**.** *Let $\sigma \in S_d$ be a permutation with cycle type $\mu = (\mu_1, \ldots, \mu_\ell)$, written as a disjoint union of cycles $\sigma = \sigma_1 \cdots \sigma_\ell$. Let $\tau = (a\,b) \in S_d$ be a trans-*

position. If a and b belong to different *cycles of σ, say $\sigma_{\ell-1}$ and σ_ℓ, then these cycles are* joined *upon composition with τ. In other words, $\tau\sigma$ has cycle type $\lambda = (\mu_1, \ldots, \mu_{\ell-2}, \mu_{\ell-1} + \mu_\ell)$. There are $\mu_{\ell-1}\mu_\ell$ such transpositions τ giving rise to the cycle type λ.*

Proof. Write $\sigma_{\ell-1}$ and σ_ℓ as the cycles $(a\, a_2\, a_3 \cdots a_{\mu_{\ell-1}})$ and $(b\, b_2\, b_3 \cdots b_{\mu_\ell})$ respectively. Then

$$\tau\sigma_{\ell-1}\sigma_\ell = (a\, b)(a\, a_2\, a_3 \cdots a_{\mu_{\ell-1}})(b\, b_2\, b_3 \cdots b_{\mu_\ell}) = (a\, a_2\, a_3 \cdots a_{\mu_{\ell-1}}\, b\, b_2\, b_3 \cdots b_{\mu_\ell})$$

has cycle type $(\mu_{\ell-1} + \mu_\ell)$. Any choice of transposition $\tau = (a\, b)$ with a belonging to $\sigma_{\ell-1}$ and b belonging to σ_ℓ will give rise to this cycle type, so there are $\mu_{\ell-1}\mu_\ell$ such transpositions. □

With these two results in hand, we describe their consequence in terms of the simple disconnected Hurwitz potential $H^\bullet$, which was first described in [GJ97].

Theorem 3.4.3 (Cut-and-join)**.** *The simple disconnected Hurwitz potential $H^\bullet(\beta; p_1, p_2, \ldots)$ is annihilated by the cut-and-join differential operator:*

$$\frac{\partial}{\partial\beta} - \frac{1}{2}\sum_{i,j=1}^{\infty}\left(ijp_{i+j}\frac{\partial^2}{\partial p_i \partial p_j} + (i+j)p_i\, p_j \frac{\partial}{\partial p_{i+j}}\right).$$

Proof. Let $\mu = (\mu_1, \ldots, \mu_\ell)$ be a partition of $|\mu| = d$ and $b \geq 0$. The coefficient of the monomial $p_\mu \frac{\beta^b}{b!}$ in $\frac{\partial H^\bullet}{\partial\beta}$ is

$$H_g^{\bullet,b+1}(\mu) = \frac{1}{d!}|M^{\bullet,b+1}(\mu)| \tag{3.10}$$

by Theorem 3.2.7. It suffices to show that this is the same as the coefficient of $p_\mu \frac{\beta^b}{b!}$ in

$$\frac{1}{2}\sum_{i,j=1}^{\infty}\left(ijp_{i+j}\frac{\partial^2 H^\bullet}{\partial p_i \partial p_j} + (i+j)p_i\, p_j \frac{\partial H^\bullet}{\partial p_{i+j}}\right). \tag{3.11}$$

The elements of $M^{\bullet,b+1}(\mu)$ are collections of permutations $(\sigma, \tau_0, \tau_1, \ldots, \tau_b) \in S_d^{b+2}$ where σ has cycle type μ, the τ_j are transpositions and $\tau_b\tau_{b-1}\cdots\tau_0\,\sigma$ is the identity. Let Λ be the set of all partitions $\lambda \vdash d$ that are obtained by cutting a cycle in μ or joining two cycles in μ, as in the cut and join lemmas. The map

$$\begin{array}{cccc} \Psi : & M^{\bullet,b+1}(\mu) & \longrightarrow & \coprod_{\lambda\in\Lambda} M^{\bullet,b}(\lambda), \\ & (\sigma, \tau_0, \tau_1, \ldots, \tau_b) & \longmapsto & (\tau_0\sigma, \tau_1, \ldots, \tau_b), \end{array}$$

allows one to write equation (3.10) as

$$\frac{1}{d!}|M^{\bullet,b+1}(\mu)| = \frac{1}{d!}\sum_{\lambda\in\Lambda}|\Psi^{-1}(M^{\bullet,b}(\lambda))|. \tag{3.12}$$

So let $x := (\rho, \tau_1, \ldots, \tau_b) \in M^{\bullet,b}(\lambda)$ for some partition $\lambda = (\lambda_1, \ldots, \lambda_{\ell'}) \in \Lambda$. The cardinality of $\Psi^{-1}(x)$ is equal to the number of transpositions τ_0 in S_d such that $\tau_0^{-1}\rho$ has cycle type μ. We split the possibilities in three cases, depending on how μ is obtained from λ, and use the cut and join lemmas.

- Suppose μ is obtained from λ by joining two cycles of length i and j. Let $n_i = |\{k : \lambda_k = i\}|$ and define n_j similarly. Then, by Lemma 3.4.2 there are $in_i\, jn_j$ transpositions τ_0 that cause this to happen, so $|\Psi^{-1}(x)| = ijn_in_j$. Therefore

$$\begin{aligned}\frac{1}{d!}|\Psi^{-1}(M^{\bullet,b}(\lambda))| &= \frac{1}{d!}\sum_{x\in M^{\bullet,b}(\lambda)}|\Psi^{-1}(x)| = \frac{1}{d!}|M^{\bullet,b}(\lambda)|\, ijn_in_j \\ &= ijn_in_jH^{\bullet,b}_{g-1}(\lambda). \end{aligned} \tag{3.13}$$

  This is precisely the coefficient of the monomial $p_\mu\frac{\beta^b}{b!}$ in the term $ijp_{i+j}\frac{\partial^2 H^\bullet}{\partial p_i \partial p_j}$ of the cut-and-join operator (3.11). To see this, note that differentiating $H^\bullet$ with respect to p_i and p_j selects the monomials whose associated partition contains at least one i and one j, such as λ. Then, indeed,

$$ijp_{i+j}\frac{\partial^2}{\partial p_i\partial p_j}H^{\bullet,b}_{g-1}(\lambda)\,p_\lambda\frac{\beta^b}{b!} = ijn_in_j\,H^{\bullet,b}_{g-1}(\lambda)\,\underbrace{\frac{p_{i+j}\,p_\lambda}{p_i\,p_j}}_{=p_\mu}\,\frac{\beta^b}{b!},$$

  where the equality under the brace is due to how μ is obtained from λ.

- Suppose μ is obtained from λ by cutting a cycle of length $i+j$ into two cycles of lengths i and j, with $i \neq j$. Let $n_{i+j} = |\{k : \lambda_k = i+j\}|$, as before. By Lemma 3.4.1 we have $|\Psi^{-1}(x)| = (i+j)\,n_{i+j}$, so

$$\begin{aligned}\frac{1}{d!}|\Psi^{-1}(M^{\bullet,b}(\lambda))| &= \frac{1}{d!}\sum_{x\in M^{\bullet,b}(\lambda)}|\Psi^{-1}(x)| = \frac{1}{d!}|M^{\bullet,b}(\lambda)|\,(i+j)\,n_{i+j} \\ &= (i+j)\,n_{i+j}\,H^{\bullet,b}_{g}(\lambda). \end{aligned} \tag{3.14}$$

  This is the coefficient of $p_\mu\frac{\beta^b}{b!}$ in the term $(i+j)p_i\,p_j\frac{\partial H^\bullet}{\partial p_{i+j}}$ of the cut-and-join operator (3.11), with $i \neq j$. To see this,

$$(i+j)p_i\,p_j\frac{\partial}{\partial p_{i+j}}H_g^{\bullet,b}(\lambda)\,p_\lambda\frac{\beta^b}{b!} = (i+j)\,n_{i+j}\,H_g^{\bullet,b}(\lambda)\,\underbrace{\frac{p_i\,p_j\,p_\lambda}{p_{i+j}}}_{=p_\mu}\frac{\beta^b}{b!}.$$

- Suppose μ is obtained from λ by cutting a cycle of length $i+j$ into two cycles of lengths i and j, with $i = j$ this time. Then, by Lemma 3.4.1, the same as the previous case occurs but with a factor of $\frac{1}{2}$.

$$\frac{1}{d!}\left|\Psi^{-1}(M^{\bullet,b}(\lambda))\right| = \frac{1}{2}(i+j)\,n_{i+j}\,H_g^{\bullet,b}(\lambda). \tag{3.15}$$

  This is the coefficient of $p_\mu\frac{\beta^b}{b!}$ in the term $\frac{1}{2}(i+j)p_i\,p_j\frac{\partial H^\bullet}{\partial p_{i+j}}$ of the cut-and-join operator (3.11), with $i = j$, as above.

We have now considered all possible partitions $\lambda \in \Lambda$, so that the coefficient of $p_\mu\frac{\beta}{b!}$ in $\frac{\partial H^\bullet}{\partial \beta}$, which is given by equation (3.12), is the sum of the three terms (3.13), (3.14) and (3.15) over all possible values of i and j, with an additional factor of $\frac{1}{2}$ when $i \neq j$ to avoid double counting. By the arguments made in each of the three cases, this is precisely the coefficient of the same monomial in the cut-and-join operator (3.11). $\square$

We have established that the disconnected simple Hurwitz potential $H^\bullet(\beta; p_1, p_2, \dots)$ satisfies a partial differential equation that encodes recursive relations among its coefficients:

$$\frac{\partial H^\bullet}{\partial \beta} = \underbrace{\frac{1}{2}\sum_{i,j=1}^{\infty}\left(ijp_{i+j}\frac{\partial^2}{\partial p_i \partial p_j} + (i+j)p_i\,p_j\frac{\partial}{\partial p_{i+j}}\right)}_{=:A} H^\bullet. \tag{3.16}$$

Using the exponentiation relation $H^\bullet = e^H - 1$ from Proposition 3.3.1, it follows almost immediately that the corresponding cut-and-join equation for the connected Hurwitz potential, which was first presented in [GJV00], is

$$\frac{\partial H}{\partial \beta} = \frac{1}{2}\sum_{i,j=1}^{\infty}\left(ijp_{i+j}\left(\frac{\partial^2 H}{\partial p_i \partial p_j} + \frac{\partial H}{\partial p_i}\frac{\partial H}{\partial p_j}\right) + (i+j)p_i\,p_j\frac{\partial H}{\partial p_{i+j}}\right).$$

We now analyze the consequences of the cut-and-join equation (3.16), following the work in [Lan10]. With the aid of some representation theory of symmetric groups, we will arrive at an explicit expression for $H^\bullet(\beta; p_1, p_2, \dots)$ in terms of eigenfunctions of the cut-and-join operator A. First, expand the disconnected potential $H^\bullet$ in the variable β as follows:

$$H^\bullet(\beta; p_*) = \sum_{b=0}^{\infty} H^\bullet_{(b)}(p_*) \frac{\beta^b}{b!}.$$

The cut-and-join equation $\frac{\partial H^\bullet}{\partial \beta} = AH^\bullet$ can be written recursively as $H^\bullet_{(b+1)}(p_*) = AH^\bullet_{(b)}(p_*)$, so

$$H^\bullet_{(b)}(p_*) = A^b H^\bullet_{(0)}(p_*).$$

The term $H^\bullet_{(0)}(p_*)$ encodes all simple Hurwitz covers with no simple branch points. In the connected case, this can only happen if the genus of the source g and the single ramification profile μ satisfy $g = \frac{1}{2}(2 - |\mu| - \ell(\mu)) \geq 0$, from which one can deduce that either $\mu = (\emptyset)$ or $\mu = (1)$. The former is not possible, as it would imply the existence of an isomorphism from a genus $g = 1$ curve to $\mathbb{P}^1$. Hence the only possibility is the unique isomorphism class of isomorphisms $\mathbb{P}^1 \to \mathbb{P}^1$. In the disconnected case, one can take d copies of this isomorphism for every d, weighted by $\frac{1}{d!}$ automorphisms arising from permuting the copies of $\mathbb{P}^1$. Hence

$$H^\bullet_{(0)}(p_*) = \sum_{d \geq 0} H^{\bullet,0}_{1-d}((1^d)) p_1^d = \sum_{d \geq 0} \frac{1}{d!} p_1^d = e^{p_1}.$$

Next, we take a brief detour towards the representation theory of symmetric groups in order to further understand the cut-and-join equation. For a partition $\mu \vdash d$, denote by C_μ the equivalence class in S_d of permutations with cycle type μ. Consider the *group algebra* $\mathbb{C}[S_d]$ of S_d, whose elements are formal linear combinations $\sum_{\sigma \in S_d} a_\sigma \sigma$ with $a_\sigma \in \mathbb{C}$. Addition in $\mathbb{C}[S_d]$ is given by

$$\sum_{\sigma \in S_d} a_\sigma \sigma + \sum_{\sigma \in S_d} b_\sigma \sigma = \sum_{\sigma \in S_d} (a_\sigma + b_\sigma)\sigma,$$

while multiplication is the composition of permutations in S_d extended to $\mathbb{C}[S_d]$ by requiring it to be bilinear. Use the same notation C_μ to denote the element in $\mathbb{C}[S_d]$ given by the sum of all permutations of cycle type μ. This lies in the centre of $\mathbb{C}[S_d]$, since for every $\rho \in S_d$ we have

$$\rho\, C_\mu\, \rho^{-1} = \sum_{\sigma \in C_\mu} \rho\sigma\rho^{-1} = \sum_{\sigma' \in C_\mu} \sigma' = C_\mu.$$

The centre of $\mathbb{C}[S_d]$, denoted henceforth by $\mathcal{Z}\mathbb{C}[S_d]$, is called the *class algebra* of S_d and $\{C_\mu : \mu \vdash d\}$ forms a basis of it. When $\mu = (2, 1, \ldots, 1)$ is the cycle type of a transposition, we shorten the notation for $C_{(2,1,\ldots,1)}$ to C_2.

The irreducible representations of S_d are called Specht modules and there is one for each partition $\mu \vdash d$ [Sag91]. We denote the Specht module associated to μ by S^μ and its associated representation by $\rho_\mu : S_d \to GL\,(S^\mu)$. Notice that ρ_μ can be naturally extended to a representation of $\mathbb{C}[S_d]$, which we also denote by $\rho_\mu : \mathbb{C}[S_d] \to GL(S^\mu)$. There is another basis $\{\chi_\mu : \mu \vdash d\}$ of $\mathcal{Z}\mathbb{C}[S_d]$ where each χ_μ acts with as a scalar with trace 1 on the irreducible representation ρ_μ and trivially on all other irreducible representations ρ_λ ($\lambda \neq \mu$). Concretely, they are given by

$$\chi_\mu = \sum_{\lambda \vdash d} \frac{1}{d!}\, \chi^\mu_\lambda\, C_\lambda, \tag{3.17}$$

where $\chi^\mu_\lambda = \chi^\mu(C_\lambda) \in \mathbb{C}$ is the character of the representation ρ_μ evaluated at any element of the conjugacy class C_λ. Since C_2 is in the centre of $\mathbb{C}[S_d]$, it also acts as a scalar on every irreducible representation $\rho_\mu : \mathbb{C}[S_d] \to GL(S^\mu)$ by Schur's lemma. Denote this scalar by $f_2(\mu)$. Then the product $C_2\chi_\mu$ acts as a scalar with trace $f_2(\mu)$ on S^μ and trivially on S^λ ($\lambda \neq \mu$), so we must have

$$C_2\, \chi_\mu = f_2(\mu)\chi_\mu \in \mathcal{Z}\mathbb{C}[S_d].$$

In other words, the action of C_2 on $\mathcal{Z}\mathbb{C}[S_d]$ is diagonal in the basis $\{\chi_\mu : \mu \vdash d\}$. We find an expression for $f_2(\mu)$ as follows:

$$\begin{aligned} \chi^\mu(C_2) &= \frac{1}{|C_2|} \sum_{\sigma \in C_2} \chi^\mu(\sigma) = \frac{1}{|C_2|} \chi^\mu\Big(\sum_{\sigma \in C_2} \sigma\Big) \\ &= \frac{1}{|C_2|} \operatorname{Tr} \rho_\mu(C_2) = \frac{1}{|C_2|} f_2(\mu) \dim S^\mu, \end{aligned}$$

so that

$$f_2(\mu) = |C_2| \frac{\chi^\mu(C_2)}{\dim S^\mu}. \tag{3.18}$$

This agrees with the more general expression for $f_{C_\lambda}(\mu)$ in [Oko00]. The cardinality of C_2 is the number of transpositions in S_d, namely $d(d-1)/2$. The value of $\chi^\mu(C_2)$ can be computed using the Frobenius character formula [FH04, Lecture 4]. This states that if $\mu = (\mu_1, \ldots, \mu_\ell) \vdash d$, then $\chi^\mu(C_2)$ is the coefficient of the monomial $x_1^{\mu_1+\ell-1} x_2^{\mu_2+\ell-2} \cdots x_{\ell-1}^{\mu_{\ell-1}+1} x_\ell^{\mu_\ell}$ in

$$(x_1^2 + x_2^2 + \cdots x_\ell^2)(x_1 + x_2 + \cdots + x_\ell)^{d-2} \prod_{1 \le i < j \le \ell} (x_i - x_j).$$

The last term in (3.18) is $\dim S^\mu$. This is equal to the number of standard μ-

tableaux [Sag91, 2.6], which is in turn equal to

$$d!\det\left(\frac{1}{(\mu_i - i + j)!}\right) \tag{3.19}$$

by the determinantal formula [Sag91, 3.2]. The matrix in the determinant is taken to be $\ell(\mu) \times \ell(\mu)$, although we can take it to be arbitrarily large by setting $\mu_k = 0$ for $k > \ell(\mu)$. In the end, the expression (3.18) for $f_2(\mu)$ turns out to be

$$f_2(\mu) = \frac{1}{2}\sum_{i=1}^{\ell(\mu)} \mu_i(\mu_i - 2i + 1). \tag{3.20}$$

We now transfer this information to the world of Hurwitz theory. Firstly, there is an isomorphism of algebras

$$\begin{array}{ccc} \mathcal{Z}\mathbb{C}[S_d] & \longrightarrow & \mathbb{C}[p_*]_d, \\ C_\mu & \longmapsto & |C_\mu|\, p_\mu, \end{array} \tag{3.21}$$

between $\mathcal{Z}\mathbb{C}[S_d]$ and the algebra of weighted homogeneous polynomials of degree d in the variables $p_* = (p_1, p_2, \dots)$, where the variable p_i is given the weight i. In the literature this is known as the characteristic map [Sag91, 4.7]. Under this isomorphism, the basis elements $\chi_\mu \in \mathcal{Z}\mathbb{C}[S_d]$ are mapped to

$$\chi_\mu \longmapsto s_\mu(p_*) := \sum_{\lambda \vdash d} \frac{1}{d!}\, \chi^\mu_\lambda |C_\lambda|\, p_\lambda \tag{3.22}$$

due to equation (3.17). This is one of the many characterizations of the *Schur functions* $\{s_\mu(p_*) : \mu \vdash d\}$ given by a theorem of Frobenius [Sag91, Theorem 4.6.4]. The Schur functions $s_k(p_*) := s_{(k)}(p_*)$ for the partitions (k) containing only one number satisfy the expansion

$$\sum_{k\geq 0} s_k(p_*) z^k = \exp\left(\sum_{j\geq 1} p_j \frac{z^j}{j}\right), \tag{3.23}$$

and the Schur function for an arbitrary partition μ is determined by the Jacobi-Trudi determinant [Sag91, 4.5]

$$s_\mu(p_*) = \det\Big(s_{\mu_i - i + j}(p_*)\Big). \tag{3.24}$$

The matrix in the determinant can be taken to be $\ell(\mu) \times \ell(\mu)$ or larger, just as for equation (3.19). Setting $p_1 = 1$ and $p_j = 0$ for $j > 1$ in (3.23) we obtain $s_k(1, 0, 0, \dots) = 1/k!$ for every $k \geq 0$. Then, combining the Jacobi-Trudi determinant (3.24) with the determinantal formula (3.19) we get

$$s_\mu(1,0,0,\dots) = \frac{1}{d!}\dim S^\mu, \quad \mu \vdash d. \tag{3.25}$$

Moreover, the action of C_2 on $\mathcal{Z}\mathbb{C}[S_d]$ given by multiplying by C_2 is mapped to the action of the cut-and-join operator A on $\mathbb{C}[p_1, p_2, \dots]_d$, via the characteristic isomorphism (3.21). This is because the cut-and-join operator measures the way in which the cycle type of a permutation is changed upon composition with a transposition in C_2. Therefore the equation $C_2\,\chi_\mu = f_2(\mu)\chi_\mu$ in $\mathcal{Z}\mathbb{C}[S_d]$ translates to $\mathbb{C}[p_*]_d$ as

$$A\, s_\mu(p_*) = f_2(\mu)\, s_\mu(p_*), \tag{3.26}$$

in other words the Schur functions form a complete set of eigenfunctions for the cut-and-join operator. Lastly, inverting the formula (3.22) for the basis change from $\{p_\mu : \mu \vdash d\}$ to $\{s_\mu(p_*) : \mu \vdash d\}$ yields [GJ08]

$$p_\lambda = \sum_{\mu \vdash d} \chi^\mu_\lambda\, s_\mu(p_*). \tag{3.27}$$

We are now in a position to expand the disconnected Hurwitz potential $H^\bullet(\beta; p_*)$ in the basis of Schur functions for $\mathbb{C}[p_*] = \bigoplus_{d\geq 0} \mathbb{C}[p_*]_d$. We remind that we expanded $H^\bullet(\beta;p_*)$ in the variable β as $\sum_{b\geq 0} H^\bullet_{(b)}(p_*)\frac{\beta^b}{b!}$, with $H^\bullet_{(b)}(p_*) = A^b H^\bullet_{(0)}(p_*) = A^b e^{p_1}$. Using the fact that for every $d \geq 0$ the monomial p_1^d is the monomial in the variables p_* corresponding to the partition $(1,\dots,1) = (1^d)$, we compute

$$\begin{aligned} e^{p_1} &= \sum_{d\geq 0} \frac{1}{d!}\, p_{(1^d)} \overset{(3.27)}{=} \sum_{d\geq 0} \frac{1}{d!} \sum_{\mu\vdash d} \chi^\mu_{(1^d)}\, s_\mu(p_*) \\ &= \sum_{d\geq 0}\sum_{\mu\vdash d} \frac{1}{d!} \dim S^\mu\, s_\mu(p_*) \overset{(3.25)}{=} \sum_{d\geq 0}\sum_{\mu\vdash d} s_\mu(1,0,0,\dots)\, s_\mu(p_*) \\ &= \sum_\mu s_\mu(1,0,0,\dots)\, s_\mu(p_*). \end{aligned}$$

In the third equality we used the fact that the character χ^μ of the representation S^μ applied to the identity in S_d gives the dimension of S^μ. Therefore the disconnected simple Hurwitz potential is

$$
\begin{aligned}
H^\bullet(\beta; p_*) &= \sum_{b=0}^{\infty} H^\bullet_{(b)}(p_*)\frac{\beta^b}{b!} = \sum_{b=0}^{\infty} A^b e^{p_1}\frac{\beta^b}{b!} \\
&= \sum_{b=0}^{\infty}\sum_{\mu} s_\mu(1,0,0,\dots)A^b s_\mu(p_*)\frac{\beta^b}{b!} \\
&\overset{(3.26)}{=} \sum_{b=0}^{\infty}\sum_{\mu} s_\mu(1,0,0,\dots)f_2(\mu)^b\, s_\mu(p_*)\frac{\beta^b}{b!} \\
&= \sum_{\mu} s_\mu(1,0,0,\dots)s_\mu(p_*)e^{f_2(\mu)\beta}.
\end{aligned}
$$

3.5 ELSV formula

We now state a celebrated result that links Hurwitz theory to intersection theory on the moduli space of curves from Section 2. In their paper [GJV00], Goulden and Jackson conjectured that the connected simple Hurwitz number for $\mu = (\mu_1, \dots, \mu_n)$ should be of the form

$$H_g^b(\mu) = b!\prod_{i=1}^{n}\frac{\mu_i^{\mu_i}}{\mu_i!}P_{g,n}(\mu_1,\dots,\mu_n),$$

where $P_{g,n}$ is a symmetric polynomial of degree $3g-3+n$ with no term of degree less than $2g-3+n$, whose coefficient in degree d has sign $(-1)^{d-(3g-3+n)}$. This "polynomiality conjecture" was confirmed by Ekedahl, Lando, Shapiro and Vainshtein in [Eke+99; Eke+01], where they proved the following remarkable theorem.

Theorem 3.5.1 (ELSV formula)**.** *Let $\mu = (\mu_1, \dots, \mu_n)$ and $2g-2+n > 0$. Then*

$$H_g^b(\mu) = b!\prod_{i=1}^{n}\frac{\mu_i^{\mu_i}}{\mu_i!}\int_{\overline{\mathcal{M}}_{g,n}}\frac{1-\lambda_1+\lambda_2-\cdots+(-1)^g\lambda_g}{(1-\mu_1\psi_1)\cdots(1-\mu_n\psi_n)}. \tag{3.28}$$

In some references there is an additional multiplicative factor of $\frac{1}{|\mathrm{Aut}(\mu)|}$ on the right-hand side of (3.28); this arises when one chooses to mark the points in the preimage of the branch point with nonsimple ramification, which we do not do here. A few other comments on this formula are in order. Firstly, the number of simple ramification points b is determined by the Riemann-Hurwitz formula, namely $b = 2g-2+n+|\mu|$. The λ-classes in the numerator are the Chern classes of the Hodge bundle $\mathbb{E} \to \overline{\mathcal{M}}_{g,n}$, which were introduced in equation (2.9). The ψ-classes appearing in the denominator contribute to the integrand by considering each fraction $\frac{1}{1-\mu_i\psi_i}$ as its geometric series expansion. Hence the entire term in the integral is

$$\sum_{j=0}^{g}(-1)^j\lambda_j\prod_{i=1}^{n}\sum_{d\geq 0}(\mu_i\psi_i)^d = \sum_{j=0}^{g}\sum_{(d_1,\ldots,d_n)}(-1)^j\lambda_j\psi_1^{d_1}\cdots\psi_n^{d_n}\,\mu_1^{d_1}\cdots\mu_n^{d_n}.$$

Since the cohomology classes ψ_i and λ_j lie in $H^2(\overline{\mathcal{M}}_{g,n})$ and $H^{2j}(\overline{\mathcal{M}}_{g,n})$ respectively, and $\dim_{\mathbb{R}}\overline{\mathcal{M}}_{g,n} = 2(3g-3+n)$, the integral only picks out the monomials $\lambda_j\psi_1^{d_1}\cdots\psi_n^{d_n}$ which satisfy $j+d_1+\cdots+d_n = 3g-3+n$. Let us see what happens when we set $g=0$. The integral in (3.28) is

$$\int_{\overline{\mathcal{M}}_{g,n}}\frac{1}{(1-\mu_1\psi_1)\cdots(1-\mu_n\psi_n)} = \sum_{(d_1,\ldots,d_n)\vdash n-3}\mu_1^{d_1}\cdots\mu_n^{d_n}\int_{\overline{\mathcal{M}}_{g,n}}\psi_1^{d_1}\cdots\psi_n^{d_n}.$$

We have already computed the genus 0 intersection products of ψ-classes in Proposition 2.6.2. Hence we obtain

$$\sum_{(d_1,\ldots,d_n)\vdash n-3}\mu_1^{d_1}\cdots\mu_n^{d_n}\binom{n-3}{d_1,\ldots,d_n} = (\mu_1+\cdots+\mu_n)^{n-3} = |\mu|^{n-3}.$$

The ELSV formula then implies

$$H_0^b(\mu) = b!|\mu|^{n-3}\prod_{i=1}^{n}\frac{\mu_i^{\mu_i}}{\mu_i!}.$$

This was first observed by Hurwitz himself in his 1891 paper [Hur91], but it was only proved much later in [GJ97]. In general Hurwitz numbers are easier to compute than intersections of ψ-classes, so the ELSV formula is useful to solve problems in the moduli space of curves.

Subsequently, a special case of the ELSV formula was proved by Fantechi and Pandharipande using different techniques [FP02], followed by Graber and Vakil [GV03] who proved it in its general case using virtual localization. What is interesting to note is that Goulden and Jackson's initial polynomiality conjecture could not be proven without the ELSV formula until a decade lateer [Dun+15]. For further details on the ELSV formula and Graber and Vakil's proof, see [Cav16; OP01; Liu10]. We now describe one of its consequences, which is the first step in the proof of Witten's conjecture presented in [KL07].

Lemma 3.5.2. *For every nonnegative integer d and $\mu = 1,\ldots,d+1$, there are* *unique constants* c_μ^d *such that*

$$\sum_{\mu=1}^{d+1} \frac{c_\mu^d}{1-\mu\psi} = \psi^d + \mathcal{O}(\psi^{d+1}). \tag{3.29}$$

Proof. Expanding the fraction $\frac{1}{1-\mu\psi}$ as a geometric series, the condition (3.29) becomes

$$\sum_{k\geq 0}\left(\sum_{\mu=1}^{d+1} \mu^k\, c_\mu^d\right)\psi^k = \psi^d + \mathcal{O}(\psi^{d+1}).$$

This is equivalent to a linear system

$$B_d\,\mathbf{c}_d = (0,\dots,0,1)^T,$$

where $\mathbf{c}_d = (c_1^d,\dots,c_{d+1}^d)^T$ and B_d is the $(d+1)\times(d+1)$ Vandermonde matrix

$$\begin{pmatrix} 1 & 1 & 1 & \cdots & 1 & 1 \\ 1 & 2 & 3 & \cdots & d & d+1 \\ 1^2 & 2^2 & 3^2 & \cdots & d^2 & (d+1)^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 1^{d-1} & 2^{d-1} & 3^{d-1} & \cdots & d^{d-1} & (d+1)^{d-1} \\ 1^d & 2^d & 3^d & \cdots & d^d & (d+1)^d \end{pmatrix}.$$

Then the terms c_μ^d can be computed as

$$\begin{aligned} c_\mu^d = \left(B_d^{-1}\right)_{\mu,d+1} &= \frac{1}{\det B_d}\,(\operatorname{adj} B_d)_{\mu,d+1} \\ &= (-1)^{\mu+d+1}\frac{\det B_d\langle d+1,\mu\rangle}{\det B_d}, \end{aligned}$$

where $\langle d+1,\mu\rangle$ denotes removal of the $(d+1)^{\text{th}}$ row and μ^{th} column. Using the Vandermonde determinants

$$\det B_d = \prod_{1\leq i<j\leq d+1}(j-i) \quad\text{and}\quad \det B_d\langle d+1,\mu\rangle = \prod_{1\leq i<j\leq d}(a_j - a_i)$$

with $a_i = \begin{cases} i, & 1\leq i\leq \mu-1 \\ i+1, & \mu\leq i\leq d+1 \end{cases}$, one obtains $c_\mu^d = \frac{(-1)^{d-\mu+1}}{(d-\mu+1)!(\mu-1)!}$. □

Using this lemma, we can 'invert' the ELSV formula (3.28) to obtain an expression for the invariants $\langle \tau_{d_1}\cdots\tau_{d_n}\rangle = \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{d_1}\cdots\psi_n^{d_n}$ in terms of simple Hurwitz numbers. For an n-tuple $(d_1,\dots,d_n)$ consider all the numbers $c_{\mu_1}^{d_1},\dots,c_{\mu_n}^{d_n}$ with

$\mu_i = 1, \ldots, d_i + 1$ and the genus g such that $d_1 + \cdots + d_n = \dim \overline{\mathcal{M}}_{g,n} = 3g - 3 + n$. We compute the following sum

$$\sum_{\mu_1=1}^{d_1+1} \cdots \sum_{\mu_n=1}^{d_n+1} c_{\mu_1}^{d_1} \cdots c_{\mu_n}^{d_n} \int_{\overline{\mathcal{M}}_{g,n}} \frac{1 - \lambda_1 + \lambda_2 - \cdots + (-1)^g \lambda_g}{(1 - \mu_1 \psi_1) \cdots (1 - \mu_n \psi_n)}$$

by noticing that it involves n multiples of equation (3.29):

$$\begin{aligned}
&= \int_{\overline{\mathcal{M}}_{g,n}} \left(\sum_{\mu_1=1}^{d_1+1} \frac{c_{\mu_1}^{d_1}}{1 - \mu_1 \psi_1} \right) \cdots \left(\sum_{\mu_n=1}^{d_n+1} \frac{c_{\mu_n}^{d_n}}{1 - \mu_n \psi_n} \right) (1 - \lambda_1 + \lambda_2 - \cdots + (-1)^g \lambda_g) \\
&= \int_{\overline{\mathcal{M}}_{g,n}} \left(\psi_1^{d_1} + \mathcal{O}(\psi_1^{d_1+1}) \right) \cdots \left(\psi_n^{d_n} + \mathcal{O}(\psi_n^{d_n+1}) \right) (1 - \lambda_1 + \lambda_2 \cdots + (-1)^g \lambda_g) \\
&= \int_{\overline{\mathcal{M}}_{g,n}} \left(\psi_1^{d_1} \cdots \psi_n^{d_n} + \mathcal{O} \right) = \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n},
\end{aligned}$$

where $\mathcal{O}$ denotes tautological classes in degree $> d_1 + \cdots + d_n = \dim \overline{\mathcal{M}}_{g,n}$. We have shown that

$$\langle \tau_{d_1} \cdots \tau_{d_n} \rangle = \sum_{\mu_1=1}^{d_1+1} \cdots \sum_{\mu_n=1}^{d_n+1} \left(\prod_{i=1}^{n} c_{\mu_i}^{d_i} \right) \int_{\overline{\mathcal{M}}_{g,n}} \frac{1 - \lambda_1 + \lambda_2 - \cdots + (-1)^g \lambda_g}{(1 - \mu_1 \psi_1) \cdots (1 - \mu_n \psi_n)}.$$

Using the ELSV formula, we can express the integral in the above equation as a multiple of a Hurwitz number, thus obtaining the following proposition.

Proposition 3.5.3. *Let $d_1, \ldots, d_n$ be nonnegative integers. Then*

$$\langle \tau_{d_1} \cdots \tau_{d_n} \rangle = \sum_{\mu_1=1}^{d_1+1} \cdots \sum_{\mu_n=1}^{d_n+1} \left(\frac{1}{b!} \prod_{i=1}^{n} \frac{(-1)^{d_i - \mu_i + 1}}{(d_i - \mu_i + 1)! \mu_i^{\mu_i - 1}} \right) H_g^b(\mu_1, \ldots, \mu_n),$$

where $b = 2g - 2 + n + \mu_1 + \cdots + \mu_n$ in each term of the sum and $g = (d_1 + \cdots + d_n - n + 3)/3$.

4 Integrable hierarchies

We now have two generating functions in the variables $t_* = (t_0, t_1, \dots)$ and $p_* = (p_1, p_2, \dots)$

$$F(t_*) = \sum \langle \tau_{d_1} \cdots \tau_{d_n} \rangle \frac{t_{d_1} \cdots t_{d_n}}{|\mathrm{Aut}(d_1, \dots, d_n)|},$$
$$H^\bullet(\beta; p_*) = \sum H_g^{\bullet, b}(\mu)\, p_\mu \frac{\beta^b}{b!},$$

encoding top intersections of ψ-classes in the moduli space of stable curves and disconnected simple Hurwitz numbers, respectively. In Section 3.4 we obtained an expression for the second one in terms of Schur functions:

$$H^\bullet(\beta; p_*) = \sum_\mu s_\mu(1, 0, 0, \dots) s_\mu(p_*) e^{f_2(\mu)\beta}, \tag{4.1}$$

where $f_2(\mu) = \frac{1}{2} \sum_{i=1}^{\ell(\mu)} \mu_i(\mu_i - 2i + 1)$. On the other hand, in Section 2.6 we presented evidence for the Witten-Kontsevich theorem, which states that $U = \frac{\partial^2 F}{\partial t_0^2}$ satisfies the KdV equations. The goal of this thesis, as mentioned in the introduction, is to present a proof of this theorem given in [KL07]. To be able to do this, it remains to discuss integrable hierarchies of partial differential equations, of which the KdV hierarchy is a notable example.

In the first three sections, we introduce integrable hierarchies following [Bur22, Sections 2-5]. The first section involves integrable hierarchies with one dependent variable, such as the KdV hierarchy. In the next section, we generalize to the case of several dependent variables and introduce the KP hierarchy. We will see how the KdV hierarchy can be obtained as a reduction of the KP hierarchy. The third section is about encoding all the information about the KP hierarchy in a single function, called the tau-function. The next two sections focus on the rich algebraic structure that underpins integrable hierarchies. The main reference for these two sections is [MJD00], which is based on earlier work by Sato [SMJ78]. The last section brings all the results together to show that the simple disconnected Hurwitz potential (4.1) is a tau-function of the KP hierarchy.

4.1 KdV hierarchy

Before beginning, we generalize the definition of the set of partitions $\mathcal{P} = \coprod_{d \geq 0} \mathcal{P}_d$ to include zeroes, and we introduce an ordering on the set of generalized partitions.

Definition 4.1.1. A *generalized partition* $\mu = (\mu_1, \dots, \mu_\ell)$ is a partition as in

Definition 3.1.1 where the components are allowed to equal zero, i.e. $\mu_1 \geq \mu_2 \geq \cdots \mu_\ell \geq 0$. The set of all generalized partitions of d is denoted by $\widetilde{\mathcal{P}}_d$, and $\widetilde{\mathcal{P}} := \coprod_{d \geq 0} \widetilde{\mathcal{P}}_d$.

Definition 4.1.2. The *lexicographical order* $>$ on $\mathcal{P}$ and $\widetilde{\mathcal{P}}$ is given by $\mu > \lambda$ if and only if $|\mu| > |\lambda|$, or $|\mu| = |\lambda|$ and $\mu_1 = \lambda_1, \dots, \mu_k = \lambda_k, \mu_{k+1} > \lambda_{k+1}$ for some k.

This section is adapted from [Bur22, Section 2]. From now on, let u be a formal variable. Consider the formal variables u_n for $n \geq 0$ and identify $u_0 = u$. We will also write $u_x := u_1, u_{xx} = u_2, u_{xxx} := u_3$, etc. Let

$$\mathcal{R} := \mathbb{C}[u_0, u_1, u_2, \dots] = \mathbb{C}[u_n]_{n \geq 0}$$

be the algebra of polynomials in these formal variables. We think of $u = u_0$ as a variable which we evaluate at a power series $\omega(x) \in \mathbb{C}[[x]]$, and we think of $u_x = u_1, u_{xx} = u_2, \dots$ as variables which we evaluate at the derivatives $\partial_x \omega(x), \partial_x^2 \omega(x), \dots$. Therefore u denotes the *dependent* variable in our discussion, and x is the *independent* variable. We call an object $P = P(u, u_x, u_{xx}, \dots) \in \mathcal{R}$ a polynomial, since it becomes a differential polynomial in x when evaluated at $u_0 = \omega(x), u_1 = \partial_x \omega(x), u_2 = \partial_x^2 \omega(x), \dots$. We write $P|_{u_n = \partial_x^n \omega}$ for this differential polynomial.

There is an operator $\partial_x : \mathcal{R} \to \mathcal{R}$ which behaves how one would intuitively expect it to behave, by replacing each "n^{th} derivative" u_n by the "$(n+1)^{\text{th}}$ derivative" u_{n+1}:

$$\partial_x = \sum_{n \geq 0} u_{n+1} \frac{\partial}{\partial u_n}.$$

Clearly $\mathbb{C} \subset \ker \partial_x$, but in fact $\mathbb{C} = \ker \partial_x$. To prove this, suppose $P \in \mathcal{R}$ does not lie in $\mathbb{C}$. If $P = P(u)$, then $\partial_x P = \frac{\partial P}{\partial u} u_x \neq 0$ so $P \notin \ker \partial_x$. Otherwise we may suppose that it can be expressed as a finite sum

$$P = \sum_{\mu \in \mathcal{P}} c_\mu u_\mu,$$

where $c_\mu \in \mathbb{C}[u]$ and u_μ denotes $u_{\mu_1} \cdots u_{\mu_\ell}$. Let $\lambda = \max\{\mu \in \mathcal{P} : c_\mu \neq 0\}$, where the maximum is taken with respect to the lexicographical order. Then $|\lambda| \geq 1$, so

$$\partial_x P = c_\lambda u_{\lambda'} + \sum_{\mu < \lambda} c'_\mu u_\mu \neq 0,$$

where $\lambda' = (\lambda_1 + 1, \lambda_2, \lambda_3, \dots)$ and $c'_\mu \in \mathbb{C}[u]$. Hence $P \notin \ker \partial_x$, and we have proved that $\mathbb{C} = \ker \partial_x$. Some of the proofs in this section involve such arguments with partitions.

For $P \in \mathcal{R}$, define the *evolutionary operator* $D_P : \mathcal{R} \to \mathcal{R}$ by

$$D_P = \sum_{n\geq 0} (\partial_x^n P) \frac{\partial}{\partial u_n}$$

so that $D_P u = P$, and the associated *evolutionary PDE* by

$$\frac{\partial u}{\partial t} = P, \tag{4.2}$$

where we have introduced an additional independent variable t. Note that evolutionary operators satisfy $[D_P, \partial_x] = 0$ and the Leibniz rule $D_P(QR) = (D_P Q)R + Q(D_P R)$. In fact, any operator $H : \mathcal{R} \to \mathcal{R}$ satisfying those two conditions is an evolutionary operator [Bur22, Proposition 2.1]. In our discussion, solutions of evolutionary PDEs will always be formal power series $\omega(x,t) \in \mathbb{C}[[x,t]]$. For such a solution of (4.2) and any $Q \in \mathcal{R}$ we have

$$\frac{\partial}{\partial t}\left(Q|_{u_n = \partial_x^n \omega} \right) = (D_P Q)|_{u_n = \partial_x^n \omega}. \tag{4.3}$$

In particular $\frac{\partial \omega}{\partial t} = (D_P u)|_{u_n = \partial_x^n \omega}$. An example of an evolutionary PDE is

$$\frac{\partial u}{\partial t} = 3uu_x + \frac{1}{4}u_{xxx}. \tag{4.4}$$

This is known as the *Korteweg-de Vries (KdV) equation*, and we have already introduced it in Theorem 2.6.3, up to a change of independent coordinate $t \mapsto 3t$. It originates from physics in the description of shallow water waves.

For $P, Q \in \mathcal{R}$ a computation shows that $[D_P, D_Q] = D_R$, where $R = D_P Q - D_Q P$. The two evolutionary PDEs associated to P and Q

$$\frac{\partial u}{\partial t} = P, \quad \frac{\partial u}{\partial s} = Q, \tag{4.5}$$

are said to be *compatible* if $[D_P, D_Q] = 0$, or equivalently $D_P Q = D_Q P$. One says that the flows $\frac{\partial}{\partial t}$ and $\frac{\partial}{\partial s}$ *commute*, or that each flow is an *infinitesimal symmetry* of the other. In the following theorem, we introduce countably many independent variables $\{t_j : j \in J\}$, and we use the shorthand notation t_* to denote $(t_j)_{j\in J}$.

Theorem 4.1.3. *Let J be a countable set of indices and $P_j \in \mathcal{R}$ for each $j \in J$.*

Then the system of PDEs

$$\frac{\partial u}{\partial t_j} = P_j, \quad j \in J,$$

has a unique solution $\omega(x, t_*) \in \mathbb{C}[[x, t_*]]$ *for an arbitrary initial condition* $\omega|_{t_*=0} = f(x) \in \mathbb{C}[[x]]$ *if and only if* $[D_{P_j}, D_{P_{j'}}] = 0$ *for all* $j, j' \in J$.

Proof. We prove the theorem for $|J| = 2$, so denote $P_1 = P, P_2 = Q$ and $t_1 = t, t_2 = s$ as in (4.5). The general case follows analogously. Firstly, suppose $[D_P, D_Q] = 0$. If $\omega(x, t, s) \in \mathbb{C}[[x, t, s]]$ is a solution with $\omega(x, 0, 0) = f(x)$, then using equation (4.3) we have

$$\left.\frac{\partial^{m+n}\omega}{\partial t^m \partial s^n}\right|_{t=s=0} = \left.(D_P^m D_Q^n u)\right|_{u_n=\partial_x^n f(x)},$$

so Taylor expanding $\omega(x, t, s)$ around $(t, s) = (0, 0)$ gives

$$\omega(x, t, s) = \sum_{m,n \geq 0} \frac{t^m s^n}{m!n!} \left.(D_P^m D_Q^n u)\right|_{u_n=\partial_x^n f(x)} = \left.\left(e^{tD_P + sD_Q} u\right)\right|_{u_n=\partial_x^n f(x)}.$$

This proves the uniqueness of $\omega(x, t, s)$. For the existence, we show that $\omega(x, t, s)$ given by the above equation satisfies $\frac{\partial u}{\partial t} = P$ and $\frac{\partial u}{\partial s} = Q$.

$$\begin{aligned}
\frac{\partial \omega}{\partial t} &= \left.\left(\frac{\partial}{\partial t} e^{tD_P + sD_Q} u\right)\right|_{u_n=\partial_x^n f(x)} = \left.\left(D_P e^{tD_P + sD_Q} u\right)\right|_{u_n=\partial_x^n f(x)} \\
&= \left.\left(e^{tD_P + sD_Q} D_P u\right)\right|_{u_n=\partial_x^n f(x)} = \left.\left(e^{tD_P + sD_Q} P(u, u_x, u_{xx}, \dots)\right)\right|_{u_n=\partial_x^n f(x)} \\
&= P\left.\left(e^{tD_P + sD_Q} u, \partial_x e^{tD_P + sD_Q} u, \partial_x^2 e^{tD_P + sD_Q} u, \dots\right)\right|_{u_n=\partial_x^n f(x)} \\
&= P(\omega, \partial_x \omega, \partial_x^2 \omega, \dots),
\end{aligned}$$

and similarly for $\frac{\partial u}{\partial s} = Q$. Clearly, $\omega(x, t, s) = \left.\left(e^{tD_P + sD_Q} u\right)\right|_{u_n=\partial_x^n f(x)}$ also satisfies the initial condition $\omega(x, 0, 0) = f(x)$. For the converse direction, suppose $\omega(x, t, s) \in \mathbb{C}[[x, t, s]]$ satisfies the system (4.5). We wish to show that $[D_P, D_Q] = 0$, or equivalently $R := D_P Q - D_Q P = 0$ since $[D_P, D_Q] = D_R$. Notice that

$$\begin{aligned}
(D_P Q)|_{u_n=\partial_x^n \omega} = (D_P D_Q u)|_{u_n=\partial_x^n \omega} &= \frac{\partial}{\partial t}\frac{\partial \omega}{\partial s} = \frac{\partial}{\partial s}\frac{\partial \omega}{\partial t} \\
= (D_Q D_P u)|_{u_n=\partial_x^n \omega} &= (D_Q P)|_{u_n=\partial_x^n \omega},
\end{aligned}$$

so $R|_{u_n=\partial_x^n \omega} = 0$. We can write R as a finite sum

$$R = \sum_{\mu \in \widetilde{\mathcal{P}}} c_\mu u_\mu, \quad c_\mu \in \mathbb{C}.$$

If $\omega(x,0,0) = f(x) = \sum_{i\geq 0} b_i x_i$ ($b_i \in \mathbb{C}$) is the initial condition, then

$$0 = \left(R|_{u_n = \partial_x^n f(x)} \right)\Big|_{x=0} = \sum_{\mu \in \widetilde{\mathcal{P}}} c_\mu \prod_{i\geq 0} (i! b_i)^{n_i(\mu)}$$

for any values of $b_0, b_1, b_2, \ldots$, where $n_i(\mu) = |\{k : \mu_k = i\}|$. Hence all the numbers $c_\mu \prod_{i\geq 0} (i!)^{n_i(\mu)}$ vanish. So $c_\mu = 0$ for all $\mu \in \widetilde{\mathcal{P}}$ and $R = 0$, as required. $\square$

One way of understanding the statement of the theorem is as follows: given an initial condition $\omega(x,0,0) \in \mathbb{C}[[x]]$ and a system of two PDEs $\frac{\partial u}{\partial t} = P$, $\frac{\partial u}{\partial s} = Q$, the compatibility condition $[D_P, D_Q] = 0$ is equivalent to the following diagram being "commutative":

$$\begin{tikzcd} \omega(x,t,0) \arrow[r, "B"] & \omega(x,t,s) \\ \omega(x,0,0) \arrow[u, "A"] \arrow[r, "B"'] & \omega(x,0,s) \arrow[u, "A"'] \end{tikzcd}$$

where A and B denote the operations 'solve $\frac{\partial u}{\partial t} = P$' and 'solve $\frac{\partial u}{\partial s} = Q$' respectively.

Definition 4.1.4. An *integrable hierarchy* is a countable sequence of pairwise compatible evolutionary PDEs

$$\frac{\partial u}{\partial t_j} = P_j, \quad j \in J,$$

where the $P_j \in \mathcal{R}$ are linearly independent.

The space of *local functionals* is

$$\mathcal{I} := \mathcal{R}/(\operatorname{im} \partial_x \oplus \mathbb{C}),$$

and the natural surjection $\mathcal{R} \to \mathcal{I}$ is denoted by $P \mapsto \int P\,dx$. The evolutionary operator $D_P : \mathcal{R} \to \mathcal{R}$ descends to a well-defined operator $D_P : \mathcal{I} \to \mathcal{I}$ because $[D_P, \partial_x] = 0$.

Definition 4.1.5. An element $\overline{h} = \int Q\,dx \in \mathcal{I}$ is called a *conserved quantity* for the PDE $\frac{\partial u}{\partial t} = P$ if $D_P \overline{h} = 0$.

The reason for this definition is that if $\omega(x,t)$ is a solution of $\frac{\partial u}{\partial t} = P$ and $\int Q\,dx$ is a conserved quantity of this PDE, then the function $Q(x,t) = \int Q|_{u_n = \partial_x^n \omega}\,dx$ does not depend on t:

$$\frac{\partial}{\partial t}\int Q|_{u_n=\partial_x^n\omega}\,dx = \int \frac{\partial}{\partial t}\left(Q|_{u_n=\partial_x^n\omega}\right)dx = \int (D_P Q)|_{u_n=\partial_x^n\omega}\,dx$$
$$= \left(D_P\int Q\,dx\right)\Big|_{u_n=\partial_x^n\omega} = 0.$$

As an example, we show that $\int u^2\,dx$ is a conserved quantity of the KdV equation $\frac{\partial u}{\partial t} = P := 3uu_x + \frac{1}{4}u_{xxx}$.

$$\begin{aligned} D_P\int u^2\,dx &= \int D_P u^2\,dx = \int\sum_{n\geq 0}(\partial_x^n P)\frac{\partial}{\partial u_n}u^2\,dx = \int(3uu_x + \frac{1}{4}u_{xxx})2u\,dx \\ &= 6\int u^2u_x\,dx + \frac{1}{2}\int uu_{xxx}\,dx = 2\int\partial_x(u^3)\,dx + \frac{1}{2}\int(\partial_x(uu_{xx}) - u_xu_{xx})\,dx \\ &= -\frac{1}{4}\int\partial_x(u_x^2)\,dx = 0. \end{aligned}$$

Here we have used $\int \operatorname{im}\partial_x\,dx = 0 \in \mathcal{I}$ repeatedly. The KdV equation has infinitely many conserved quantities, of which the above one is the first [Bur22, Theorem 2.16]. They are given by $\int H_j\,dx, j\geq 1$, where the $H_j \in \mathcal{R}$ are determined recursively by

$$H_j = \begin{cases} \frac{1}{2}u, & j = 1, \\ \frac{i}{2\sqrt{2}}\partial_x H_{j-1} - \frac{1}{2}\sum_{a+b=j-1}H_aH_b, & j\geq 2. \end{cases}$$

It turns out that the even-indexed conserved quantities are irrelevant since $\int H_{2j}\,dx = 0$, and that the odd-indexed ones span the space of all conserved quantities of the KdV equation [Bur22, Corollary 2.36]. Hence we rename (and renormalize) the conserved quantities as follows:

$$\overline{h}_j^{\mathrm{KdV}} = (-1)^j 2^{j+1}(2j+1)\int H_{2j+1}\,dx, \quad j\geq 0. \tag{4.6}$$

We have chosen to start the indexing from $j = 0$, while [Bur22] starts at $j = 1$. This is done in order to be consistent with the indexing choice in [Wit90; KL07]. We list the first few:

$$\begin{aligned} \overline{h}_0^{\mathrm{KdV}} &= \int u\,dx, \\ \overline{h}_1^{\mathrm{KdV}} &= \int\frac{3}{2}u^2\,dx, \\ \overline{h}_2^{\mathrm{KdV}} &= \int\left(\frac{5}{2}u^3 + \frac{5}{8}uu_{xx}\right)dx, \\ \overline{h}_3^{\mathrm{KdV}} &= \int\left(\frac{35}{8}u^4 + \frac{35}{16}u^2u_{xx} + \frac{7}{32}uu_{xxxx}\right)dx. \end{aligned}$$

The *variational derivative* is the operator

$$\frac{\delta}{\delta u} := \sum_{n\geq 0} (-\partial_x)^n \circ \frac{\partial}{\partial u_n} : \mathcal{R} \to \mathcal{R},$$

where $\circ$ denotes composition of operators. One can check that $\operatorname{im} \partial_x \oplus \mathbb{C} \subset \ker \frac{\delta}{\delta u}$, and moreover it is true that $\operatorname{im} \partial_x \oplus \mathbb{C} = \ker \frac{\delta}{\delta u}$ [Bur22, Theorem 2.22], so it induces a well-defined inclusion $\frac{\delta}{\delta u} : \mathcal{I} \to \mathcal{R}$. Define a bracket $\{\cdot,\cdot\} : \mathcal{I} \times \mathcal{I} \to \mathcal{I}$ by

$$\{\overline{h}, \overline{g}\} = \int \frac{\delta \overline{h}}{\delta u} \partial_x \frac{\delta \overline{g}}{\delta u}\, dx.$$

By associating to $\overline{h} \in \mathcal{I}$ the evolutionary operator

$$D_{\overline{h}} := D_{\partial_x \frac{\delta \overline{h}}{\delta u}} = \sum_{n\geq 0} \partial_x^{n+1} \left(\frac{\delta \overline{h}}{\delta u} \right) \frac{\partial}{\partial u_n},$$

one can recover the bracket $\{\cdot,\cdot\}$ on $\mathcal{I}$ by noticing that for $\overline{g} = \int g\, dx \in \mathcal{I}$ we have

$$\begin{aligned} D_{\overline{h}} \overline{g} &= \int \sum_{n\geq 0} \partial_x^{n+1} \left(\frac{\delta \overline{h}}{\delta u} \right) \frac{\partial g}{\partial u_n}\, dx = -\int \sum_{n\geq 0} \partial_x^{n} \left(\frac{\delta \overline{h}}{\delta u} \right) \partial_x \left(\frac{\partial g}{\partial u_n} \right) dx \\ &= \cdots = \int \sum_{n\geq 0} \partial_x \left(\frac{\delta \overline{h}}{\delta u} \right) (-\partial_x)^n \left(\frac{\partial g}{\partial u_n} \right) dx = \int \partial_x \left(\frac{\delta \overline{h}}{\delta u} \right) \frac{\delta g}{\delta u}\, dx = \{\overline{g}, \overline{h}\}. \end{aligned}$$

The bracket $\{\cdot,\cdot\}$ on $\mathcal{I}$ is the infinite-dimensional analogue of the Poisson bracket on a manifold M, and the evolutionary operators $D_{\overline{h}}$ are analogous to the Hamiltonian vector fields $X_h \in \mathfrak{X}(M)$. For a reference, see [MR99, 5.5]. The equation above is similar to the Poisson bracket-Lie derivative identity $\{h, g\} = X_g(h)$. Moreover, the commutator identity $[X_h, X_g] = X_{\{g,h\}}$ is true in our case as well, namely $[D_{\overline{h}}, D_{\overline{g}}] = D_{\{\overline{g},\overline{h}\}}$ [Bur22, Theorem 2.29]. These facts imply that $\{\cdot,\cdot\}$ endows $\mathcal{I}$ with the structure of a Lie algebra and that the map $\mathcal{R} \times \mathcal{I} \to \mathcal{R}, (P, \overline{h}) \mapsto D_{\overline{h}} P$ is a right Lie algebra action of $\mathcal{I}$ on $\mathcal{R}$.

Definition 4.1.6. A *Hamiltonian* PDE is an evolutionary PDE of the form

$$\frac{\partial u}{\partial t} = \partial_x \frac{\delta \overline{h}}{\delta u}. \tag{4.7}$$

The local functional $\overline{h} \in \mathcal{I}$ is called the *Hamiltonian.*

Notice that if $\overline{g} \in \mathcal{I}$ is a conserved quantity of the Hamiltonian PDE (4.7), then

the PDE
$$\frac{\partial u}{\partial s} = \partial_x \frac{\delta \overline{g}}{\delta u}$$
is compatible with (4.7), since $0 = D_{\overline{h}}\overline{g} = \{\overline{g}, \overline{h}\}$ implies $[D_{\overline{h}}, D_{\overline{g}}] = D_{\{\overline{g},\overline{h}\}} = 0$. This means that two Hamiltonian PDEs with Hamiltonians $\overline{h}$ and $\overline{g}$ are compatible if and only if $\{\overline{g}, \overline{h}\} = 0$. We return to our example of the KdV equation $\frac{\partial u}{\partial t} = 3uu_x + \frac{1}{4}u_{xxx}$ and show that it is Hamiltonian:

$$\begin{aligned}
\frac{1}{5}\partial_x \frac{\delta \overline{h}_2^{\mathrm{KdV}}}{\delta u} &= \partial_x \frac{\delta}{\delta u}\left(\frac{1}{2}u^3 + \frac{1}{8}uu_{xx}\right) \\
&= \partial_x \sum_{n\geq 0}(-\partial_x)^n \frac{\partial}{\partial u_n}\left(\frac{1}{2}u^3 + \frac{1}{8}uu_{xx}\right) \\
&= \partial_x\left(\frac{3}{2}u^2 + \frac{1}{8}u_{xx} + \frac{1}{8}u_{xx}\right) = 3uu_x + \frac{1}{4}u_{xxx}.
\end{aligned}$$

Therefore, by the previous statement, the conserved quantities $\overline{h}_j^{\mathrm{KdV}}$ of the KdV equation give rise to infinitely many Hamiltonian PDEs

$$\frac{\partial u}{\partial t_j} = P_j^{\mathrm{KdV}} := \frac{1}{2j+3}\partial_x \frac{\delta \overline{h}_{j+1}^{\mathrm{KdV}}}{\delta u}, \quad j \geq 0, \tag{4.8}$$

that are compatible with the KdV equation. We remind that we have started the indexing from $j = 0$ instead of $j = 1$ to be consistent with [Wit90; KL07]. A little more work [Bur22, Corollary 2.36] shows that the evolutionary operators $D_{P_j^{\mathrm{KdV}}}$ pairwise commute and that they form a basis for the space of infinitesimal symmetries of the KdV equation. Therefore (4.8) forms an integrable hierarchy in the sense of Definition 4.1.4, which is called the *KdV hierarchy*. It is the same set of PDEs as (2.24) in Theorem 2.6.3, after the rescaling $t_j \mapsto (2j+1)t_j$. The first three equations are

$$\begin{aligned}
\frac{\partial u}{\partial t_0} &= u_x, \\
\frac{\partial u}{\partial t_1} &= 3uu_x + \frac{1}{4}u_{xxx}, \\
\frac{\partial u}{\partial t_2} &= \frac{15}{2}u^2u_x + \frac{5}{2}u_xu_{xx} + \frac{5}{4}uu_{xxx} + \frac{1}{16}u_{xxxxx}.
\end{aligned}$$

From the first two equations, we see that the independent variables t_0 and t_1 can be identified with x and t from the original KdV equation (4.4).

4.2 KP hierarchy

We now extend the formalism developed in Section 4.1 to include multiple dependent variables $(u_i)_{1\leq i\leq N}$, following [Bur22, Sections 3-4]. Similarly to before, we identify $u_{i,x} := u_{i,1}, u_{i,xx} := u_{i,2}, u_{i,xxx} := u_{i,3}, \dots$, and define the algebra

$$\mathcal{R}_u = \mathbb{C}[u_{i,n}]_{\substack{1\leq i\leq N\\ n\geq 0}}.$$

As before, we have the operator

$$\partial_x = \sum_{i=1}^{N}\sum_{n\geq 0} u_{i,n+1}\frac{\partial}{\partial u_{i,n}} : \mathcal{R}_u \longrightarrow \mathcal{R}_u$$

which satisfies $\ker \partial_x = \mathbb{C}$. The space of local functionals is

$$\mathcal{I}_u = \mathcal{R}_u/(\operatorname{im}\partial_x \oplus \mathbb{C})$$

and the variational derivatives

$$\frac{\delta}{\delta u_i} = \sum_{n\geq 0}(-\partial_x)^n \circ \frac{\partial}{\partial u_{i,n}} : \mathcal{I}_u \longrightarrow \mathcal{R}_u, \quad 1\leq i\leq N.$$

satisfy $\operatorname{im}\partial_x \oplus \mathbb{C} = \bigcap_{i=1}^{N} \ker \frac{\delta}{\delta u_i}$. To an N-tuple $\overline{P} = (P_1,\dots,P_N) \in \mathcal{R}_u^N$ we associate the evolutionary operator

$$D_{\overline{P}} = \sum_{i=1}^{N}\sum_{n\geq 0}(\partial_x^n P_i)\frac{\partial}{\partial u_{i,n}} : \mathcal{R}_u \longrightarrow \mathcal{R}_u$$

which satisfies $[D_{\overline{P}}, \partial_x] = 0$ and the Leibniz rule, and the evolutionary PDEs

$$\frac{\partial u_i}{\partial p} = P_i, \quad 1\leq i\leq N, \tag{4.9}$$

where we denote the independent variable by p now. Analogously to (4.3), if $\omega_1(x,p),\dots,\omega_N(x,p) \in \mathbb{C}[[x,p]]$ are solutions of (4.9) then for any $Q \in \mathcal{R}_u$ we have

$$\frac{\partial}{\partial p}\left(Q|_{u_{i,n}=\partial_x^n\omega_i}\right) = (D_{\overline{P}}Q)|_{u_{i,n}=\partial_x^n\omega_i}. \tag{4.10}$$

Two N-tuples of evolutionary PDEs associated to $\overline{P}$ and $\overline{Q}$ are called compatible if $D_{\overline{R}} = [D_{\overline{P}}, D_{\overline{Q}}] = 0$, where $R_i = D_{\overline{P}}Q_i - D_{\overline{Q}}P_i$. Analogously to Theorem 4.1.3, given a countable set of indices J and $\overline{P}_j = (P_{1,j},\dots,P_{N,j})$ for each $j \in J$, the system of PDEs

$$\frac{\partial u_i}{\partial p_j} = P_{i,j}, \quad 1 \leq i \leq N, \quad j \in J,$$

has a unique solution $\omega_1(x,p_*), \ldots, \omega_N(x,p_*) \in \mathbb{C}[[x,p_*]$ if and only if $[D_{\overline{P}_j}, D_{\overline{P}_{j'}}] = 0$ for every $j, j' \in J$. Such a system is also called an integrable hierarchy.

We now introduce the KP hierarchy, an important example of an integrable hierarchy with countably many ($N \to \infty$) dependent variables. First, consider composing $\partial_x : \mathcal{R}_u \to \mathcal{R}_u$ with the operator induced by $P \in \mathcal{R}_u$. By the product rule of differentiation, this is

$$\partial_x \circ P = \partial_x P + P \circ \partial_x.$$

Applying the product rule $n \geq 1$ times gives

$$\partial_x^n \circ P = \sum_{k=0}^{n} \binom{n}{k} (\partial_x^k P)\, \partial_x^{n-k}. \tag{4.11}$$

We may treat the operator ∂_x as a formal variable, and extend the ring of operators on $\mathcal{R}_u$ by introducing its formal inverse ∂_x^{-1}. With this in mind, we define a *pseudo-differential operator* as being a Laurent series of the form

$$A = \sum_{n \leq m} a_n \partial_x^n, \quad m \in \mathbb{Z}, \quad a_n \in \mathcal{R}_u.$$

Moreover the positive part, the negative part and the residue of A are defined as

$$A_+ = \sum_{0 \leq n \leq m} a_n \partial_x^n, \quad A_- = A - A_+, \quad \operatorname{res} A = a_{-1}.$$

We extend the multiplication (4.11) to pseudo-differential operators as follows:

$$\left(\sum_{n_1 \leq m_1} a_{n_1} \partial_x^{n_1}\right) \circ \left(\sum_{n_2 \leq m_2} b_{n_2} \partial_x^{n_2}\right) = \sum_{\substack{n_1 \leq m_1 \\ n_2 \leq m_2}} \sum_{k \geq 0} \binom{n_1}{k} a_{n_1} (\partial_x^k b_{n_2})\, \partial_x^{n_1+n_2-k}. \tag{4.12}$$

This endows the space $\mathcal{PO}_u$ of pseudo-differential operators with an algebra structure, which is associative [Bur22, Proposition 4.1]. The commutator of $A, B \in \mathcal{PO}_u$ is $[A,B] = A \circ B - B \circ A$. The action of an evolutionary operator $D_{\overline{P}}$ on $\mathcal{PO}_u$ is given coefficient-wise by

$$D_{\overline{P}} \left(\sum_{n \leq m} a_n \partial_x^n\right) = \sum_{n \leq m} (D_{\overline{P}}\, a_n) \partial_x^n, \tag{4.13}$$

from which the Leibniz rule $D_{\overline{P}}(A \circ B) = (D_{\overline{P}}A) \circ B + A \circ (D_{\overline{P}}B)$ for every $A, B \in \mathcal{PO}_u$ follows. Now consider the pseudo-differential operator

$$L = \partial_x + \sum_{i \geq 1} u_i \partial_x^{-i}.$$

The positive part is $L_+ = \partial_x$, while $L_- = \sum_{i \geq 1} u_i \partial_x^{-i}$ and $\operatorname{res} L = u_1$. For any $j \geq 1$ let $L^j = \underbrace{L \circ \cdots \circ L}_{j \text{ times}}$. As an example, let us compute the second and third powers of L using the multiplication rule (4.12).

$$\begin{aligned} L^2 &= \partial_x^2 + \partial_x \sum_{i \geq 1} u_i \partial_x^{-i} + \sum_{i \geq 1} u_i \partial_x^{-i+1} + \left(\sum_{i \geq 1} u_i \partial_x^{-i} \right) \left(\sum_{j \geq 1} u_j \partial_x^{-j} \right) \\ &= \underbrace{\partial_x^2 + 2u_1}_{(L^2)_+} + \underbrace{\sum_{i \geq 1} (\partial_x u_i + 2u_{i+1}) \partial_x^{-i} + \sum_{\substack{i,j \geq 1 \\ k \geq 0}} \binom{-i}{k} u_i (\partial_x^k u_j) \partial_x^{-i-j-k}}_{(L^2)_-}, \end{aligned} \tag{4.14}$$

in particular $\operatorname{res} L^2 = \partial_x u_1 + 2u_2$. The third power is given by the slightly longer expression

$$\begin{aligned} L^3 = &\overbrace{\partial_x^3 + 3u_1 \partial_x + 3u_2 + 3\partial_x u_1}^{(L^3)_+} + \sum_{i \geq 1} (\partial_x^2 u_i + 3\partial_x u_{i+1} + 3u_{i+2}) \partial_x^{-i} \\ &+ \sum_{\substack{i \geq 1 \\ \ell \geq 0}} \left(2 \binom{-i}{\ell} u_i (\partial_x^\ell u_1) + u_i u_{\ell+1} \right) \partial_x^{-i-\ell} \\ &+ \sum_{\substack{i,j \geq 1 \\ \ell \geq 0}} \binom{-i}{\ell} \left((\partial_x u_i)(\partial_x^\ell u_j) + 2u_i (\partial_x^\ell u_{j+1}) + \frac{\ell + 2 - i}{\ell + 1} u_i (\partial_x^{\ell+1} u_j) \right) \partial_x^{-i-j-\ell} \\ &+ \sum_{\substack{i,j,k \geq 1 \\ \ell,m,n \geq 0}} \binom{-i}{m} \binom{-j}{\ell} \binom{m}{n} (\partial_x^n u_j)(\partial_x^{\ell+m-n} u_k) \partial_x^{-i-j-k-\ell-m}, \end{aligned}$$

in particular $\operatorname{res} L^3 = \partial_x^2 u_1 + 3\partial_x u_2 + 3u_3 + 3u_1^2$. The positive parts of the commutators $[(L^j)_+, L]$ satisfy

$$\begin{aligned} [(L^j)_+, L]_+ &= [L^j - (L^j)_-, L]_+ = -[(L^j)_-, L]_+ \\ &= -[(L^j)_-, L_+ + L_-]_+ = -[(L^j)_-, L_+]_+ \\ &= -[(L^j)_-, \partial_x]_+ = 0, \end{aligned}$$

so $[(L^j)_+, L]$ is equal to its negative part $[(L^j)_+, L]_-$. The *KP (Kadomtsev-Petviashvili)* *hierarchy* is obtained from this commutator as the following system of evolutionary PDEs:

$$\frac{\partial u_i}{\partial p_j} = P_{i,j} := \mathrm{Coef}_{\partial_x^{-i}}[(L^j)_+, L], \quad i, j \geq 1, \tag{4.15}$$

where the expression on the right-hand side means that we select the coefficient of ∂_x^{-i} in $[(L^j)_+, L]$. This is a system of countably many PDEs for the countably many dependent variables $u_1, u_2, \ldots$, each of which depend on countably many independent variables $p_1, p_2, \ldots$. Notice that $L_+ = \partial_x$ implies that $[L_+, L] = \sum_{i\geq 1}(\partial_x u_i)\partial_x^{-i}$, so

$$\frac{\partial u_i}{\partial p_1} = \partial_x u_i, \quad i \geq 1. \tag{4.16}$$

Hence x is identified with the first independent variable p_1. If we write $\overline{P}_j = (P_{1,j}, P_{2,j}, \ldots)$ and denote the associated evolutionary operators by $D_j := D_{\overline{P_j}}$, then (4.15) is equivalent to

$$\frac{\partial L}{\partial p_j} = D_j L = [(L^j)_+, L],$$

where the action of $\frac{\partial}{\partial p_j}$ and D_j on L is coefficient-wise (4.13). Since L contains all the dependent variables u_i as its coefficients, it uniquely determines the KP hierarchy. The following result shows that the KP hierarchy is indeed an integrable hierarchy in the sense of Definition 4.1.4.

Theorem 4.2.1. *The operators D_j of the KP hierarchy pairwise commute: $[D_j, D_{j'}] = 0$. Moreover, the local functionals $\int \mathrm{res}\, L^i\, dx$ are conserved quantities for any flow $\frac{\partial}{\partial p_j}$ of the KP hierarchy.*

Proof. We first check that $[D_j, D_{j'}] = 0$ for any $j, j' \geq 1$, or equivalently $[D_j, D_{j'}]L = 0$. Write $L_+^j := (L^j)_+$ for brevity. By the Leibniz rule, we have

$$D_j L^{j'} = \sum_{a+b=j'-1} L^a \circ D_j L \circ L^b = \sum_{a+b=j'-1} L^a \circ [L_+^j, L] \circ L^b = [L_+^j, L^{j'}], \tag{4.17}$$

where the last equality can be proved by inducting on j'. Hence

$$\begin{aligned} D_j(D_{j'}L) &= D_j[L_+^{j'}, L] = [(D_j L^{j'})_+, L] + [L^{j'}, D_j L] \\ &= [[L_+^j, L^{j'}], L] + [L^{j'}, L_+^j]. \end{aligned}$$

Then $[D_j, D_{j'}]L = D_j(D_{j'}L) - D_{j'}(D_j L)$ equals the expression

$$[[L_+^j, L^{j'}], L] + [L^{j'}, L_+^j] - [[L_+^{j'}, L^j], L] - [L^j, L_+^{j'}],$$

which vanishes due to the Jacobi identity and using $L^j = L^j_+ + L^j_-$. The local functionals $\int \operatorname{res} L^i \, dx$ are conserved quantities of all the KP equations since

$$D_j \int \operatorname{res} L^i \, dx = \int D_j \operatorname{res} L^i \, dx = \int \operatorname{res} D_j L^i \, dx = \int \operatorname{res}[L^j_+, L^i] dx.$$

The last expression vanishes because $\int \operatorname{res}[A, B] dx = 0$ for any $A, B \in \mathcal{PO}_u$. For example, for $A = P\partial_x^n$ and $B = Q\partial_x^m$,

$$\operatorname{res}[P\partial_x^n, Q\partial_x^m] = \begin{cases} \binom{m}{n+m+1} P(\partial_x^{n+m+1} Q) - \binom{n}{n+m+1} (\partial_x^{n+m+1} P) Q, & n+m+1 \geq 0, \\ 0, & \text{else.} \end{cases}$$

For $n + m + 1 \geq 0$,

$$\begin{aligned} &\int \left(\binom{m}{n+m+1} P(\partial_x^{n+m+1} Q) - \binom{n}{n+m+1} (\partial_x^{n+m+1} P) Q \right) dx \\ &= \int \left(\binom{m}{n+m+1} + (-1)^{n+m} \binom{n}{n+m+1} \right) P(\partial_x^{n+m+1} Q) \, dx = 0, \end{aligned}$$

as claimed. □

We now wish to see how the KdV and KP hierarchies are related, and for this we give some general results on reductions of integrable hierarchies. In addition to the algebra $\mathcal{R}_u$ associated to the dependent variables $u_1, \ldots, u_N$, consider another set of dependent variables $v_1, \ldots, v_M$ and their associated algebra $\mathcal{R}_v$. Let $(D_j : \mathcal{R}_u \to \mathcal{R}_u)_{j \geq 1}$ and $(D'_j : \mathcal{R}_v \to \mathcal{R}_v)_{j \geq 1}$ be two collections of pairwise commuting evolutionary operators in $\mathcal{R}_u$ and $\mathcal{R}_v$ giving rise to the two integrable hierarchies

$$\frac{\partial u_i}{\partial p_j} = P_{i,j} = D_j u_i, \quad 1 \leq i \leq N, \ j \geq 1, \tag{4.18}$$

$$\frac{\partial v_k}{\partial p_j} = Q_{k,j} = D'_j v_k, \quad 1 \leq k \leq M, \ j \geq 1. \tag{4.19}$$

Let $\theta : \mathcal{R}_u \to \mathcal{R}_v$ be an algebra morphism that commutes with ∂_x and is compatible with the evolutionary operators, in other words $[\theta, \partial_x] = 0$ and $\theta \circ D_j = D'_j \circ \theta$.

Lemma 4.2.2. *Suppose a collection of formal power series $\nu_k(x, p_*) \in \mathbb{C}[[x, p_*]]$, $1 \leq k \leq M$, is a solution of the second integrable hierarchy (4.19). Then the collection*

$$\omega_i(x, p_*) := \theta(u_i)|_{v_{k,n} = \partial_x^n \nu_k} \in \mathbb{C}[[x, p_*]], \quad 1 \leq i \leq N,$$

is a solution of the first integrable hierarchy (4.18).

Proof. We use equation (4.10) to compute

$$\begin{aligned}\frac{\partial \omega_i}{\partial p_j} &= \frac{\partial}{\partial p_j}\left(\theta(u_i)|_{v_{k,n}=\partial_x^n \nu_k}\right) = D'_j(\theta(u_i))\Big|_{v_{k,n}=\partial_x^n \nu_k} = \theta(D_j(u_i))|_{v_{k,n}=\partial_x^n \nu_k} \\ &= \theta(P_{i,j})|_{v_{k,n}=\partial_x^n \nu_k} = \left(P_{i,j}|_{u_{i'},m=\partial_x^m \theta(u_{i'})}\right)\Big|_{v_{k,n}=\partial_x^n \nu_k} = P_{i,j}|_{u_{i,n}=\partial_x^n \omega_i},\end{aligned}$$

where in the second-to-last equality we used $[\theta, \partial_x] = 0$. □

As an example with $N = M = \infty$, we present a change of dependent variables for the KP hierarchy that resemble the action-angle variables from the Liouville-Arnold theorem [Arn89, Chapter 10].

Example 4.2.3. Since each $\operatorname{res} L^i$ has the form $\operatorname{res} L^i = iu_i + Q_i$ for some polynomials $Q_i \in \mathcal{R}_u$ depending only on $u_1, \dots, u_{i-1}$ and their derivatives, the algebra homomorphism $\theta : \mathcal{R}_u \to \mathcal{R}_v$ sending $\partial_x^n \operatorname{res} L^i$ to $v_{i,n}$ to is an isomorphism commuting with ∂_x. Let $D'_j : \mathcal{R}_v \to \mathcal{R}_v$ be the evolutionary operators induced by the KP operators D_j and the isomorphism θ, namely $D'_j = \theta \circ D_j \circ \theta^{-1}$. By Lemma 4.2.2, the solutions of the KP hierarchy (4.15) are in bijective correspondence with solutions of

$$\frac{\partial v_i}{\partial p_j} = D'_j v_i, \quad i, j \geq 1. \tag{4.20}$$

By Theorem 4.2.1 we know that $D_j \operatorname{res} L^i \in \operatorname{im} \partial_x$, so $D'_j v_i = \theta(D_j(\operatorname{res} L^i))$ is of the form $\partial_x Q_{i,j}$ with $Q_{i,j} \in \mathcal{R}_v$. The polynomials $Q_{i,j}$ are symmetric in i and j since

$$\begin{aligned} D_j \operatorname{res} L^i &= \operatorname{res}[L^j_+, L^i] = -\operatorname{res}[L^j_-, L^i] = -\operatorname{res}[L^j_-, L^i_+] \\ &= -\operatorname{res}[L^j, L^i_+] = \operatorname{res}[L^i_+, L^j] = D_i \operatorname{res} L^j \end{aligned}$$

implies $D'_i v_j = D'_j v_i$. Moreover, if we introduce a gradation on $\mathcal{R}_v$ by defining $\deg v_{i,n} := n$, then $Q_{i,j}$ does not contain monomials of degree d when d is odd [BRZ21, Lemma 3.9]. The dependent variables v_i are called the *normal coordinates* of the KP hierarchy, since they correspond to the conserved quantities $\operatorname{res} L^i$. Similarly to the first equations (4.16) of the original KP hierarchy we have

$$\frac{\partial v_i}{\partial p_1} = \frac{\partial v_1}{\partial p_i} = \partial_x v_i, \quad i \geq 1.$$

Using the previous computations for L, L^2 and L^3 we obtain

$$\begin{aligned}
v_1 &= \theta(\operatorname{res} L) = \theta(u_1), \\
v_2 &= \theta(\operatorname{res} L^2) = \theta(\partial_x u_1 + 2u_2), \\
v_3 &= \theta(\operatorname{res} L^3) = \theta(\partial_x^2 u_1 + 3\partial_x u_2 + 3u_3 + 3u_1^2),
\end{aligned}$$

which we use to compute the PDE in (4.20) corresponding to $i = j = 2$:

$$\begin{aligned}
\frac{\partial v_2}{\partial p_2} &= D_2' v_2 = \theta\left(D_2(\operatorname{res} L^2)\right) = \theta\left(\operatorname{res}[L_+^2, L^2]\right) \\
&= \theta\left(\operatorname{res}\left[\partial_x^2 + 2u_1, \sum_{i=1}^{3}(\partial_x u_i + 2u_{i+1})\partial_x^{-i} + \sum_{\substack{i,j\geq 1, k\geq 0 \\ i+j+k\leq 3}} \binom{-i}{k} u_i(\partial_x^k u_j)\partial_x^{-i-j-k}\right]\right) \\
&= \theta\left(\partial_x^3 u_1 + 4\partial_x^2 u_2 + 4\partial_x u_3 + 2\partial_x u_1^2\right) = \partial_x \theta\left(\frac{4}{3}\operatorname{res} L^3 - 2(\operatorname{res} L)^2 - \frac{1}{3}\partial_x^2 \operatorname{res} L\right) \\
&= \partial_x \underbrace{\left(\frac{4}{3}v_3 - 2v_1^2 - \frac{1}{3}\partial_x^2 v_1\right)}_{Q_{2,2}}. \qquad (4.21)
\end{aligned}$$

In the previous example we had an algebra isomorphism $\theta : \mathcal{R}_u \to \mathcal{R}_v$ commuting with ∂_x. In the following lemma, $\theta : \mathcal{R}_u \to \mathcal{R}_v$ is a surjective algebra morphism commuting with ∂_x such that $\ker \theta$ is invariant under each D_j.

Lemma 4.2.4. *Let $\theta : \mathcal{R}_u \to \mathcal{R}_v$ be a surjective algebra morphism commuting with ∂_x such that $D_j \ker \theta \subset \ker \theta$. Consider the integrable hierarchy*

$$\frac{\partial u_i}{\partial p_j} = P_{i,j} = D_j w_i, \quad i, j \geq 1. \qquad (4.22)$$

For any collection of formal power series $f_i(x) \in \mathbb{C}[[x]], i \geq 1$, such that $K|_{u_{i,n}=\partial_x^n f_i(x)} = 0$ for every $K \in \ker \theta$, we have

$$K|_{u_{i,n}=\partial_x^n \omega_i} = 0,$$

where $\omega_i(x, p_*) \in \mathbb{C}[[x, p_*]], i \geq 1$, is the solution to (4.22) with initial condition $\omega_i(x, 0) = f_i(x)$.

Proof. By the proof of Theorem 4.1.3, the solution $\omega_i(x, p_*)$ is

$$\omega_i(x, p_*) = \left(\exp\Big(\sum_{j\geq 1} p_j D_j\Big) u_i\right)\Bigg|_{u_{k,n}=\partial_x^n f_k(x)},$$

which implies that for any $Q \in \ker\theta$,

$$K|_{u_{i,d}=\partial_x^n \omega_i} = \left(\exp\Big(\sum_{j\geq 1} p_j D_j\Big) K\right)\Bigg|_{u_{i,n}=\partial_x^n f_i(x)} = 0,$$

where the last equality follows from $D_j \ker\theta \subset \ker\theta$. $\square$

Under the assumptions of the above lemma, θ induces an algebra isomorphism $\overline{\theta} : \mathcal{R}_u/\ker\theta \to \mathcal{R}_v$ that commutes with ∂_x. The evolutionary operators $(D_j : \mathcal{R}_u \to \mathcal{R}_u)_{j\geq 1}$ induce pairwise compatible evolutionary operators $(\overline{D}_j : \mathcal{R}_u/\ker\theta \to \mathcal{R}_u/\ker\theta)_{j\geq 1}$. For each $j \geq 1$, let $D'_j = \overline{\theta} \circ \overline{D}_j \circ \overline{\theta}^{-1}$ be the corresponding evolutionary operator in $\mathcal{R}_v$. Then the resulting integrable hierarchy

$$\frac{\partial v_k}{\partial p_j} = Q_{k,j} = D'_j v_k, \quad 1 \leq k \leq M,\ j \geq 1$$

is called a *reduction* of the initial hierarchy (4.22). By Lemma 4.2.2 and Lemma 4.2.4, solutions of the reduced hierarchy are in bijective correspondence with solutions of the initial hierarchy that vanish on $\ker\theta$. The polynomials $Q_{k,j}$ are given by $\theta(D_j P_k)$, where $P_k \in \mathcal{R}_u$ is an arbitrary polynomial such that $\theta(P_k) = v_k$.

There is a reduction from the KP hierarchy (4.15) to the KdV hierarchy (4.8), which we now illustrate [Bur22, Section 4.4].

Example 4.2.5. Firstly, consider the pseudo-differential operator $L = \partial_x + \sum_{i\geq 1} u_i \partial_x^{-i}$ from the definition of the KP hierarchy. From the computation in equation (4.14), its square L^2 has the form

$$L^2 = \partial_x^2 + \sum_{i\geq 0} R_i \partial_x^{-i},$$

where $R_i = 2u_{i+1} + \widetilde{R}_i$ and $\widetilde{R}_i$ only depends on $u_1, \ldots, u_i$ and their derivatives. Consider the single dependent variable u from Section 4.1 and its algebra $\mathcal{R} = \mathbb{C}[u_n]_{n\geq 0}$ (which is distinct from $\mathcal{R}_u = \mathbb{C}[u_{i,n}]_{i\geq 1, n\geq 0}$). Then there are unique polynomials $Z_i \in \mathcal{R}, i \geq 1$, such that

$$\left(\partial_x + \sum_{i\geq 1} Z_i \partial_x^{-i}\right)^2 = \partial_x^2 + 2u. \tag{4.23}$$

They are given recursively by

$$Z_i = \begin{cases} u, & i = 1, \\ -\frac{1}{2}\, \widetilde{R}_{i-1}\Big|_{u_{k,n} = \partial_x^n Z_k}, & i \geq 2. \end{cases}$$

Define an algebra morphism $\theta : \mathcal{R}_u \to \mathcal{R}$ by $u_{i,n} \mapsto \partial_x^n Z_i$, whose kernel $\ker\theta$ is the ideal generated by $\{\partial_x^n R_i : i \geq 1, n \geq 0\}$. Notice that for each KP operator D_j we have

$$\begin{aligned} \sum_{i \geq 1} D_j R_i\, \partial_x^{-i} &= D_j(L_-^2) = D_j(L^2)_- \overset{(4.17)}{=} [L_+^j, L^2]_- \\ &= [L_+^j, L_-^2]_- = \sum_{i \geq 1} [L_+^j, R_i \partial_x^{-i}]_-, \end{aligned}$$

so $\ker\theta = (\partial_x^n R_i)_{i \geq 1, n \geq 0}$ is invariant under the KP operators since $[L_+^j, R_i \partial_x^{-i}] \in (\partial_x^n R_i)_{i \geq 1, n \geq 0}$. So by Lemma 4.2.4 the morphism $\theta : \mathcal{R}_u \to \mathcal{R}$ defines a reduction of the KP hierarchy (4.15) to the integrable hierarchy in one dependent variable

$$\frac{\partial u}{\partial p_j} = Q_j, \quad j \geq 1, \tag{4.24}$$

where $Q_j = \theta(D_j u_1)$. Notice that

$$\theta(L)^2 = \Big(\partial_x + \sum_{i \geq 1} \theta(u_i) \partial_x^{-i}\Big)^2 = \Big(\partial_x + \sum_{i \geq 1} Z_i \partial_x^{-i}\Big)^2 \overset{(4.23)}{=} \partial_x^2 + 2u,$$

so $\theta(L)^{2j} = \theta(L)_+^{2j}$. Therefore the even-indexed equations of the reduced hierarchy (4.24) are trivial:

$$\begin{aligned} Q_{2j} &= \theta(D_{2j} u_1) = \theta(D_{2j} \operatorname{res} L) \overset{(4.17)}{=} \theta(\operatorname{res}[L_+^{2j}, L]) \\ &= \operatorname{res}[\theta(L)_+^{2j}, \theta(L)] = \operatorname{res}[\theta(L)^{2j}, \theta(L)] = 0. \end{aligned}$$

Correspondingly, the even-indexed conserved quantities $\int \operatorname{res} L^{2i} dx$ of the KP hierarchy are mapped to zero under θ:

$$\theta\left(\int \operatorname{res} L^{2j}\, dx\right) = \int \operatorname{res} \theta(L)^{2j} dx = \int \operatorname{res} \left(\partial_x^2 + 2u\right)^j dx = 0.$$

On the other hand, the odd-indexed equations correspond to the ones for the KdV hierarchy (4.8) since $Q_{2j+1} = P_j^{\mathrm{KdV}}$ for $j \geq 0$ [Bur22, Proposition 4.18] (we remind that we have started the indexing for the KdV hierarchy at $j = 0$ instead of $j = 1$), while the odd-indexed conserved quantities of the KP equations are mapped to the conserved quantities of the KdV equation:

$$\theta\left(\int \operatorname{res} L\,dx\right) = \int \operatorname{res}\theta(L)\,dx = \int \operatorname{res}\Big(\partial_x + \sum_{i\geq 1} Z_i\partial_x^{-i}\Big)dx$$
$$= \int Z_1\,dx = \int u\,dx = \overline{h}_0^{\mathrm{KdV}},$$

and for $j \geq 1$

$$\theta\left(\int \operatorname{res} L^{2j+1}\,dx\right) = \int \operatorname{res}\theta(L)^{2j}\theta(L)\,dx = \int \operatorname{res}\left(\partial_x^2 + 2u\right)^j\Big(\partial_x + \sum_{i\geq 1} Z_i\partial_x^{-i}\Big)dx$$
$$= (-1)^j 2^{j+1}(2j+1)\int H_{2j+1}\,dx = \overline{h}_j^{\mathrm{KdV}}.$$

From this example we conclude that, under the identification of independent variables $t_j = p_{2j+1}$, there is a bijective correspondence between solutions $\omega(x, t_*) \in \mathbb{C}[[t_*]]$ of the KdV hierarchy (4.8) and solutions $L|_{u_i=\omega_i} = \left(\partial_x + \sum_{i\geq 1} u_i\partial_x^{-1}\right)\Big|_{u_i=\omega_i}$ of the KP hierarchy (4.15) satisfying $L^2_- = 0$, i.e. $L^2 = \partial_x^2 + 2u$.

4.3 Tau-functions

In the previous section, we introduced the KP hierarchy

$$\frac{\partial u_i}{\partial p_j} = \mathrm{Coef}_{\partial_x^{-i}}[(L^j)_+, L], \quad i, j \geq 1, \text{ or equivalently}$$
$$\frac{\partial L}{\partial p_j} = [(L^j)_+, L], \quad j \geq 1, \quad L = \partial_x + \sum_{i\geq 1} u_i\partial_x^{-i}.$$

The first PDEs of this hierarchy are $\frac{\partial u_i}{\partial p_1} = \partial_x u_i$, so that we can identify x with the independent variable p_1. Therefore, by a solution of the KP hierarchy we mean a collection of power series $\omega_1(p_*), \omega_2(p_*), \cdots \in \mathbb{C}[[p_*]]$, or equivalently a pseudo-differential operator with coefficients in $\mathbb{C}[[p_*]]$ given by

$$L = \partial_x + \sum_{i\geq 1} \omega_i(p_*)\partial_x^{-i},$$

which satisfy the PDEs above. The algebra $\mathcal{PO}_p$ of pseudo-differential operators with coefficients in $\mathbb{C}[[p_*]]$ has the same multiplication (4.12) as in $\mathcal{PO}_u$, but with differentiation by x replaced by differentiation by p_1. The positive part, negative part, and residue of elements of $\mathcal{PO}_p$ are defined in the same way as $\mathcal{PO}_u$.

The KP hierarchy has many equations: one for every pair consisting of a dependent variable and an independent variable. It would be useful if we could store all this information in a single equation for a single function. In this section, we develop the necessary tools to do that, following [Bur22, Section 5]. First, we introduce yet another set of formal variables $w_1, w_2, \dots$ and their associated algebra $\mathcal{R}_w$. Let

$$W := 1 + \underbrace{\sum_{i\geq 1} w_i \partial_x^{-i}}_{\widetilde{W}} \in \mathcal{PO}_w.$$

The inverse of W in $\mathcal{PO}_w$ is the well-defined pseudo-differential operator $W^{-1} = 1 - \widetilde{W} + \widetilde{W}^2 - \widetilde{W}^3 + \cdots$. A careful computation [Bur22, Lemma 5.5] shows that

$$W \circ \partial_x \circ W^{-1} = \partial_x + \sum_{i\geq 1} Y_i \partial_x^{-i},$$

where the coefficients $Y_i \in \mathcal{R}_w$ are of the form $Y_i = -w_{i,x} + \widetilde{Y}_i$ and $\widetilde{Y}_i$ only depends on $w_1, \dots, w_{i-1}$ and their derivatives. Therefore the algebra homomorphism $\theta : \mathcal{R}_u \to \mathcal{R}_w$ given by $\theta(u_{i,n}) = \partial_x^n Y_i$ is injective. Moreover,

$$\theta(L) = \partial_x + \sum_{i\geq 1} \theta(u_i)\partial_x^{-i} = \partial_x + \sum_{i\geq 1} Y_i \partial_x^{-i} = W \circ \partial_x \circ W^{-1}.$$

For $j \geq 1$ let $\widetilde{D}_j : \mathcal{R}_w \to \mathcal{R}_w$ be the evolutionary operator defined by

$$\widetilde{D}_j W = -(W \circ \partial_x^j \circ W^{-1})_- \circ W = -\theta(L)^j_- \circ W.$$

Proposition 4.3.1. *Let $(D_j : \mathcal{R}_u \to \mathcal{R}_u)_{j\geq 1}$ be the evolutionary operators of the KP hierarchy, which are given by $D_j L = [L^j_+, L]$. Then*

$$\theta \circ D_j = \widetilde{D}_j \circ \theta.$$

Proof. Applying $\widetilde{D}_j$ to both sides of $W \circ W^{-1} = 1$ yields $\widetilde{D}_j W^{-1} = -W^{-1} \circ \widetilde{D}_j W \circ W^{-1} = W^{-1} \circ \theta(L)^j_-$. Then

$$\begin{aligned}(\widetilde{D}_j \circ \theta)L &= \widetilde{D}_j(W \circ \partial_x \circ W^{-1}) = (\widetilde{D}_j W) \circ \partial_x \circ W^{-1} + W \circ \partial_x \circ (\widetilde{D}_j W^{-1}) \\ &= -\theta(L)^j_- \circ W \circ \partial_x \circ W^{-1} + W \circ \partial_x \circ W^{-1} \circ \theta(L)^j_- = -\theta(L)^j_- \circ \theta(L) + \theta(L) \circ \theta(L)^j_- \\ &= -[\theta(L)^j_-, \theta(L)] = \theta(-[L^j_-, L]) = \theta([L^j_+, L]) = (\theta \circ D_j)L,\end{aligned}$$

as claimed. □

By the above proposition and repeated use of $L^j = L^j_+ + L^j_-$, one can also show that the operators $(\widetilde{D}_j : \mathcal{R}_w \to \mathcal{R}_w)_{j\geq 1}$ pairwise commute. Hence we obtain an integrable hierarchy

$$\frac{\partial W}{\partial p_j} = \widetilde{D}_j W = -(W \circ \partial_x^j \circ W^{-1}) \circ W, \quad j \geq 1, \tag{4.25}$$

known as the *SW (Sato-Wilson) hierarchy.* By Lemma 4.2.2, if $W \in \mathcal{PO}_p$ is a solution of the SW hierarchy then $L := W \circ \partial_x \circ W^{-1}$ is a solution of the KP hierarchy. Notice that the correspondence of solutions is not bijective, since the morphism $\theta : \mathcal{R}_u \to \mathcal{R}_w$ is injective but not surjective. One can prove [Bur22, Proposition 5.11] that given a solution L of the KP hiearchy, there exists a solution W of the SW hierarchy such that $L = W \circ \partial_x \circ W^{-1}$, which is defined up to the transformation

$$W \longmapsto W \circ \left(1 + \sum_{i\geq 1} c_i \partial_x^{-i}\right), \quad c_i \in \mathbb{C}. \tag{4.26}$$

The operator W is called a *dressing operator* of the KP hierarchy.

Consider the repeated action of ∂_x on e^{xz}:

$$\partial_x e^{xz} = ze^{xz}, \quad \partial_x^2 e^{xz} = z^2 e^{xz}, \quad \partial_x^3 e^{xz} = z^3 e^{xz} \dots.$$

One can extend this action to negative powers of ∂_x by defining $\partial_x^{-i} e^{xz} = z^{-i} e^{xz}$, so that a pseudo-differential operator $A = \sum_{n\leq m} a_n \partial_x^n \in \mathcal{PO}_p$ acts on e^{xz} as follows:

$$Ae^{xz} := \sum_{n\leq m} a_n z^n e^{xz} = \widehat{A} e^{xz}, \tag{4.27}$$

where on the right we are multiplying Laurent series. The Laurent series $\widehat{A} = \sum_{n\leq m} a_n z^n \in \mathbb{C}[[p_*]][[z, z^{-1}]]$ is the *symbol* of A, defined by identifying ∂_x^n with z^n. Notice that multiplication of Laurent series is not well-defined in general, for example $(\cdots + z^{-2} + z^{-1} + 1)(1 + z + z^2 + \cdots)$. In our discussion, however, the Laurent series that appear will always have a well-defined multiplication. We extend (4.27) in two ways: firstly, the action of another pseudo-differential operator $B \in \mathcal{PO}_p$ on Ae^{xz} is defined as $B(Ae^{xz}) = (B \circ A)e^{xz}$. Secondly, let

$$\xi(p_*, z) := \sum_{j\geq 1} p_j z^j \tag{4.28}$$

and define the corresponding action of $A \in \mathcal{PO}_p$ on $e^{\xi(p_*,z)}$ by identifying $p_1 = x$:

$$Ae^{\xi(p_*,z)} := e^{\sum_{j\geq 2} p_j z^j} \widehat{A} e^{p_1 z}.$$

Definition 4.3.2. Let $L \in \mathcal{PO}_p$ be a solution of the KP hierarchy, and W a dressing operator of L. The *wave function* associated to L and W is

$$\psi(p_*, z) := We^{\xi(p_*,z)} \in \mathbb{C}[[p_*]][[z, z^{-1}]].$$

The wave function ψ satisfies the basic properties

$$\begin{aligned} L\psi &= L(We^{\xi(p_*,z)}) = (L \circ W)e^{\xi(p_*,z)} = (W \circ \partial_x \circ W^{-1} \circ W)e^{\xi(p_*,z)} \\ &= W(\partial_x e^{\xi(p_*,z)}) = zWe^{\xi(p_*,z)} = z\psi \end{aligned} \tag{4.29}$$

and consequently

$$\begin{aligned} \frac{\partial \psi}{\partial p_j} &= \frac{\partial}{\partial p_j}(We^{\xi(p_*,z)}) = \frac{\partial W}{\partial p_j} e^{\xi(p_*,z)} + W\frac{\partial e^{\xi(p_*,z)}}{p_j} = -(L^j_- \circ W)e^{\xi(p_*,z)} + z^j We^{\xi(p_*,z)} \\ &= -L^j_-\psi + L^j\psi = L^j_+\psi. \end{aligned} \tag{4.30}$$

The *adjoint* of a pseudo-differential operator $A = \sum_{n \le m} a_n \partial_x^n \in \mathcal{PO}_p$ is defined as

$$A^\dagger := \sum_{n \le m} (-\partial_x)^n \circ a_n = \sum_{n \le m} (-1)^n \sum_{k \ge 0} \binom{n}{k} (\partial_x^k a_n)\, \partial_x^{n-k},$$

and it satisfies the properties that one would expect of it [Bur22, Lemma 5.16]:

$$(A \circ B)^\dagger = B^\dagger \circ A^\dagger, \quad (A^\dagger)^\dagger = A, \quad (A^\dagger)_\pm = (A_\pm)^\dagger.$$

The *adjoint wave function* associated to a solution L of the KP hierarchy and its associated dressing operator W is

$$\psi^\dagger(p_*, z) := (W^\dagger)^{-1} e^{-\xi(p_*,z)}.$$

Similarly to (4.29) and (4.30), the adjoint wave function satisfies

$$L^\dagger \psi^\dagger = z\psi^\dagger \quad \text{and} \quad \frac{\partial \psi^\dagger}{\partial p_j} = -(L^\dagger)^j_+ \psi^\dagger.$$

After the following lemma, we prove the defining property of the wave functions ψ and $\psi^\dagger$ of the KP hierarchy.

Lemma 4.3.3. *For any $A = \sum_i a_i \partial_x^j$ and $B = \sum_j b_j \partial_x^j$ the following holds:*

$$\operatorname{res}_z(Ae^{xz} \cdot Be^{-xz}) = \operatorname{res}_{\partial_x}(A \circ B^\dagger),$$

where res_z *denotes the coefficient of* z^{-1} *and* $\mathrm{res}_{\partial_x}$ *the coefficient of* ∂_x^{-1}.

Proof. The left-hand side is

$$\mathrm{res}_z \left(\left(\sum_i a_i z^i e^{xz} \right) \cdot \left(\sum_j b_j (-z)^j e^{-xz} \right) \right) = \mathrm{res}_z \left(\sum_{i,j} (-1)^j a_i b_j z^{i+j} \right) = \sum_{i+j=-1} (-1)^j a_i b_j,$$

while the right-hand side is

$$\begin{aligned} &\mathrm{res}_{\partial_x} \left(\left(\sum_i a_i \partial_x^i \right) \circ \left(\sum_j (-1)^j \partial_x^j \circ b_j \right) \right) = \mathrm{res}_{\partial_x} \left(\sum_{i,j} (-1)^j a_i \partial_x^{i+j} \circ b_j \right) \\ &= \mathrm{res} \left(\sum_{i,j} (-1)^j a_i \sum_{k \geq 0} \binom{i+j}{k} (\partial_x^k b_j)\, \partial_x^{i+j-k} \right) = \sum_{i+j+1 \geq 0} (-1)^j a_i \binom{i+j}{i+j+1} \partial_x^{i+j+1} b_j \\ &= \sum_{i+j+1=0} (-1)^j a_i b_j, \end{aligned}$$

where in the last equality we used the fact that $\binom{i+j}{i+j+1} = 0$ for $i+j+1 > 0$. $\square$

Theorem 4.3.4 (Bilinear identity)**.** *Let L be a solution of the KP hierarchy and W a dressing operator, and $\psi, \psi^\dagger$ the associated wave functions. Introduce the formal variables $\lambda_* = (\lambda_j)_{j \geq 1}$ and $p'_* := p_* + \lambda_*$. Then*

$$\mathrm{res}_z \left(\psi(p_*, z) \cdot \psi^\dagger(p'_*, z) \right) = 0. \tag{4.31}$$

The converse is also true: if $\gamma(p_*, z) = \left(1 + \sum_{i \geq 1} \gamma_i z^{-i}\right) e^{\xi(p_*, z)}$ and $\rho(p_*, z) = \left(1 + \sum_{i \geq 1} \rho_i z^{-i}\right) e^{-\xi(p_*, z)}$ satisfy

$$\mathrm{res}_z \left(\gamma(p_*, z) \cdot \rho(p'_*, z) \right) = 0,$$

then $\gamma = W e^{\xi(p_*, z)} = \psi$ and $\rho = (W^\dagger)^{-1} e^{-\xi(p_*, z)} = \psi^\dagger$ for a solution W of the SW hierarchy.

Proof. Since $p'_* = p_* + \lambda_*$, Taylor expanding the bilinear identity (4.31) around $\lambda_* = 0$ gives the equivalent condition

$$\mathrm{res}_z \left(\psi \cdot \frac{\partial^n \psi^\dagger}{\partial p_{j_1} \cdots \partial p_{j_n}} \right) = 0 \quad \text{for all } n \geq 0, j_1, \ldots, j_n \geq 1, \tag{4.32}$$

so we wish to prove (4.32). The property $\frac{\partial \psi^\dagger}{\partial p_j} = -(L^\dagger)^j_+ \psi^\dagger$ implies that

$$\frac{\partial^n \psi^\dagger}{\partial p_{j_1} \cdots \partial p_{j_n}} = \sum_{j=0}^{m} f_j \partial_x^j \psi^\dagger$$

for some $m \geq 0$ and $f_j \in \mathbb{C}[[p_*]]$, since only the positive part of each power of $L^\dagger$ is involved. So it suffices to prove $\operatorname{res}_z(\psi \cdot \partial_x^j \psi^\dagger) = 0$ for any $j \geq 0$. Using Lemma 4.3.3,

$$\begin{aligned}\operatorname{res}_z(\psi \cdot \partial_x^j \psi^\dagger) &= \operatorname{res}_z \left(W e^{\xi(p_*,z)} \cdot (\partial_x^j \circ (W^\dagger)^{-1}) e^{-\xi(p_*,z)} \right) \\ &\overset{(4.28)}{=} \operatorname{res}_z \left(W e^{xz} \cdot (\partial_x^j \circ (W^\dagger)^{-1}) e^{-xz} \right) = \operatorname{res}_{\partial_x} \left(W \circ (\partial_x^j \circ (W^\dagger)^{-1})^\dagger \right) \\ &= \operatorname{res}_{\partial_x} \left(W \circ W^{-1} \circ (\partial_x^j)^\dagger \right) = \operatorname{res}_{\partial_x}((-\partial_x)^j) = 0.\end{aligned}$$

Let us now prove the converse. Let $\widetilde{W} = 1 + \sum_{i \geq 1} \gamma_i \partial_x^{-i}$ and $\widetilde{V} = 1 + \sum_{i \geq 1} \rho_i \partial_x^{-i}$, so that $\gamma = \widetilde{W} e^{\xi(p_*,z)}$ and $\rho = \widetilde{V} e^{-\xi(p_*,z)}$. We have to prove that $\widetilde{V} = (\widetilde{W}^\dagger)^{-1}$ and that $\widetilde{W}$ is a solution of the SW hierarchy. For the first part, we have that γ and ρ satisfy (4.32), so the same computation that we have just done yields

$$0 = \operatorname{res}_z(\gamma \cdot \partial_x^j \rho) = (-1)^j \operatorname{res}_{\partial_x} \left(\widetilde{W} \circ \widetilde{V}^\dagger \circ \partial_x^j \right)$$

for every $j \geq 0$, which implies $(\widetilde{W} \circ \widetilde{V}^\dagger)_- = 0$. But $(\widetilde{W} \circ \widetilde{V}^\dagger)_+ = 1$, so $\widetilde{W} \circ \widetilde{V}^\dagger = 1$ as required. Next, we show that $\widetilde{W}$ satisfies the SW hierarchy (4.25)

$$\frac{\partial \widetilde{W}}{\partial p_j} = L^j_- \circ \widetilde{W}, \quad L := \widetilde{W} \circ \partial_x \circ \widetilde{W}^{-1}. \tag{4.33}$$

Firstly, notice that

$$\frac{\partial \widetilde{W}}{\partial p_j} e^{\xi(p_*,z)} = \frac{\partial}{\partial p_j}(\widetilde{W} e^{\xi(p_*,z)}) - (\widetilde{W} \circ \partial_x^j) e^{\xi(p_*,z)} = \frac{\partial}{\partial p_j}(\widetilde{W} e^{\xi(p_*,z)}) - L^j(\widetilde{W} e^{\xi(p_*,z)}),$$

so

$$\left(\frac{\partial \widetilde{W}}{\partial p_j} + L^j_- \circ \widetilde{W} \right) e^{\xi(p_*,z)} = \frac{\partial \gamma}{\partial p_j} - L^j_+ \gamma.$$

By assumption γ and ρ satisfy (4.32), so for all $k \geq 0$:

$$\begin{aligned}
0 &= \operatorname{res}_z \left(\partial_x^k \circ \left(\frac{\partial \gamma}{\partial p_j} - L_+^j \gamma \right) \cdot \rho \right) \\
&= \operatorname{res}_z \left(\partial_x^k \circ \left(\frac{\partial \widetilde{W}}{\partial p_j} + L_-^j \circ \widetilde{W} \right) e^{\xi(p_*,z)} \cdot (\widetilde{W}^{-1})^\dagger e^{-\xi(p_*,z)} \right) \\
&= \operatorname{res}_{\partial_x} \left(\partial_x^k \circ \left(\frac{\partial \widetilde{W}}{\partial p_j} + L_-^j \circ \widetilde{W} \right) \circ \widetilde{W}^{-1} \right).
\end{aligned}$$

Therefore $\left(\left(\frac{\partial \widetilde{W}}{\partial p_j} + L_-^j \circ \widetilde{W} \right) \circ \widetilde{W}^{-1} \right)_- = 0$. But the positive part of this expression also vanishes, so we obtain (4.33). $\square$

From now on, fix a solution $W = 1 + \sum_{i \geq 1} \omega_i(p_*) \partial_x^{-i} \in \mathcal{PO}_p$ of the SW hierarchy and $L = W \circ \partial_x \circ W^{-1}$ the corresponding solution of the KP hierarchy. Define $\widehat{W}^\dagger := \widehat{(W^\dagger)^{-1}}\Big|_{z \mapsto -z}$ so that the adjoint wave function $\psi^\dagger = \widehat{W}^\dagger e^{-\xi(p_*,z)}$. Introduce an operator $G_z : \mathbb{C}[[p_*]] \to \mathbb{C}[[p_*]][[z^{-1}]]$ defined by

$$f(p_*) \longmapsto f\left(p_1 - \frac{1}{z}, p_2 - \frac{1}{2z^2}, p_3 - \frac{1}{3z^3}, \dots \right). \tag{4.34}$$

The action of G_z on $\mathbb{C}[[p_*]][z, z^{-1}]]$ is given coefficient-wise, namely

$$G_z \cdot \sum_{n \leq m} f_n(p_*) z^n = \sum_{n \leq m} G_z(f_n) z^n.$$

Two consequences of the bilinear identity (4.31) are [Bur22, Propositions 5.24-5.25]

$$\widehat{W}^{-1} = G_z(\widehat{W}^\dagger) \quad \text{and} \quad \partial_x \log \widehat{W} = \omega_1 - G_z(\omega_1). \tag{4.35}$$

Moreover, if we differentiate (4.34) by z we get

$$\frac{\partial}{\partial z} G_z(f) = \sum_{j \geq 1} z^{-j-1} \frac{\partial}{\partial p_j} G_z(f),$$

so the operator $N_z := \frac{\partial}{\partial z} - \sum_{j \geq 1} z^{-j-1} \frac{\partial}{\partial p_j}$ satisfies $N_z(G_z(f)) = 0$ for every $f \in \mathbb{C}[[p_*]]$. With these facts and definitions in mind, we introduce the tau-function of the KP hierarchy in the following theorem [Bur22, Theorem 5.27].

Theorem 4.3.5. *Let $W \in \mathcal{PO}_p$ be a solution of the SW hierarchy. There exists a unique formal power series $\tau(p_*) \in \mathbb{C}[[p_*]]$ such that $\tau(0) = 1$ and*

$$\widehat{W} = \frac{G_z(\tau)}{\tau}. \tag{4.36}$$

Moreover we have

$$\frac{\partial \log \tau}{\partial p_j} = \operatorname{res}_z \left(z^j N_z(\log \widehat{W}) \right), \quad j \geq 1. \tag{4.37}$$

We make a few comments about this theorem. Firstly, given a solution $L = W \circ \partial_x \circ W^{-1} \in \mathcal{PO}_p$ of the KP hierarchy, equation (4.36) is taken as the definition of a *tau-function* of the KP hierarchy associated L. This definition is actually equivalent to the second condition (4.37), since

$$\begin{aligned}
\widehat{W} = \frac{G_z(\tau)}{\tau} &\iff \log(\widehat{W}) = \log G_z(\tau) - \log \tau = G_z(\log \tau) - \log \tau \\
&\iff N_z(\log \widehat{W}) = -N_z(\log \tau) = \sum_{k \geq 1} z^{-k-1} \frac{\partial \log \tau}{\partial p_k} \\
&\iff \operatorname{res}_z \left(z^j N_z(\log \widehat{W}) \right) = \operatorname{res} \left(\sum_{k \geq 1} z^{j-k-1} \frac{\partial \log \tau}{\partial p_k} \right) = \frac{\partial \log \tau}{\partial p_j}, \quad j \geq 1.
\end{aligned}$$

The second and third equivalences are valid because both expressions only contain negative powers of z. The condition (4.37) arises because of the *tau-symmetry*

$$\frac{\partial}{\partial p_i} \operatorname{res}_z \left(z^j N_z(\log \widehat{W}) \right) = \frac{\partial}{\partial p_j} \operatorname{res}_z \left(z^i N_z(\log \widehat{W}) \right), \quad i, j \geq 1.$$

Tau-symmetry plays an important role in the theory of Hamiltonian evolutionary PDEs, see [Dub+16] for example. We now see how to recover the conserved quantities of the KP hierarchy from the tau-function.

Proposition 4.3.6. *Let $\tau(p_*) \in \mathbb{C}[[p_*]]$ be a tau-function of the KP hierarchy and $L = W \circ \partial_x \circ W^{-1} \in \mathcal{PO}_p$ the corresponding solution. Then*

$$\operatorname{res} L^i = \frac{\partial^2 \log \tau}{\partial p_1 \partial p_i}, \quad i \geq 1. \tag{4.38}$$

Proof. Consider the computation

$$\begin{aligned}
\frac{\partial^2 \log \tau}{\partial p_1 \partial p_i} &= \frac{\partial}{\partial p_1} \operatorname{res}_z \left(z^i N_z(\log \widehat{W}) \right) = \operatorname{res}_z \left(z^i N_z \left(\frac{\partial}{\partial p_1} \log \widehat{W} \right) \right) \\
&\overset{(4.35)}{=} \operatorname{res}_z \left(z^i N_z(\omega_1 - G_z(\omega_1)) \right) = \operatorname{res}_z \left(z^i N_z(\omega_1) \right) \\
&= \operatorname{res}_z \left(- \sum_{j \geq 1} z^{i-j-1} \frac{\partial \omega_1}{\partial p_j} \right) = - \frac{\partial \omega_1}{\partial p_i}.
\end{aligned}$$

The last term is the coefficient of $-\partial_x^{-1}$ on the left-hand side of the equation for the SW hierarchy,

$$\frac{\partial W}{\partial p_i} = -L_-^i \circ W.$$

Since $L_-^i = \operatorname{res} L^i + \sum_{j\geq 2} \omega_j \partial_x^{-j}$ and $W = 1 + \sum_{j\geq 1} \omega_j \partial_x^{-j}$, the coefficient of $-\partial_x^{-1}$ on the right-hand side is $\operatorname{res} L^i$, proving the claim. $\square$

In Example 4.2.3 we switched to the normal coordinates of the KP hierarchy, and obtained

$$\frac{\partial}{\partial p_j} \operatorname{res} L^i = \operatorname{res}[L_+^j, L^i], \quad i, j \geq 1.$$

We computed the PDE corresponding to $i = j = 2$:

$$\frac{\partial}{\partial p_2} \operatorname{res} L^2 = \partial_x \left(\frac{4}{3} \operatorname{res} L^3 - 2(\operatorname{res} L)^2 - \frac{1}{3} \partial_x^2 \operatorname{res} L \right).$$

Using Proposition 4.3.6, we can write this in terms of the tau-function by identifying $x = p_1$:

$$\frac{\partial^2 \log \tau}{\partial p_2^2} = \frac{4}{3} \frac{\partial^2 \log \tau}{\partial p_1 \partial p_3} - 2 \left(\frac{\partial^2 \log \tau}{\partial p_1^2} \right)^2 - \frac{1}{3} \frac{\partial^4 \log \tau}{\partial p_1^4}. \tag{4.39}$$

This is a form of the original KP equation, which describes nonlinear wave motion [KP70].

In Example 4.2.5 we established a bijective correspondence between solutions of the KdV hierarchy and solutions $L \in \mathcal{PO}_p$ of the KP hierarchy satisfying $L_-^2 = 0$. By definition, a tau-function of the KdV hierarchy is a tau-function of the associated KP hierarchy. Therefore we see that the statement of the Witten-Kontsevich theorem (Theorem 2.6.3) is that the generating series F for intersections of ψ-classes is a tau-function of the KdV hierarchy.

Corollary 4.3.7. *Let $\tau \in \mathbb{C}[[p_*]]$ be a tau-function of the KP hierarchy. If τ is a tau-function of the KdV hierarchy then $\frac{\partial \log \tau}{\partial p_{2j}}$ is constant for all $j \geq 1$. On the other hand, if $\frac{\partial \log \tau}{\partial p_2}$ is constant then τ is a tau-function of the KdV hierarchy.*

Proof. Let L be the solution of the KP hierarchy associated to τ, and W a dressing operator of L. If τ is a tau-function of the KdV hierarchy, then $L_-^{2j} = 0$, so

$$\frac{\partial W}{\partial p_{2j}} = -L_-^{2j} \circ W = 0.$$

Applying $\frac{\partial}{\partial p_{2j}}$ to

$$\frac{\partial \log \tau}{\partial p_i} = \operatorname{res}_z \left(z^i N_z(\log \widehat{W}) \right), \quad i \geq 1, \tag{4.40}$$

gives $\frac{\partial^2 \log \tau}{\partial p_{2j} \partial p_i} = 0$ for all $i \geq 1$, so $\frac{\partial \log \tau}{\partial p_{2j}}$ is constant. On the other hand, suppose $\frac{\partial \log \tau}{\partial p_2}$ is constant. The dressing operator W is defined uniquely up to the transformation (4.26)

$$W \longmapsto W \circ \left(1 + \sum_{i \geq 1} c_i \partial_x^{-i}\right), \quad c_i \in \mathbb{C}.$$

Therefore, since $\widehat{W} = \frac{G_z(\tau)}{\tau}$, the tau-function τ is defined uniquely up to the transformation

$$\tau \longmapsto \tau \exp\left(\sum_{i \geq 1} d_i p_i\right), \quad d_i \in \mathbb{C}, \tag{4.41}$$

where $1 + \sum_{i \geq 1} c_i z^{-i} = \exp\left(-\sum_{i \geq 1} \frac{d_i}{i} z^{-i}\right)$. Therefore $\widetilde{\tau} = \tau \exp(-dp_2)$ is also a tau-function associated to L, where $d = \frac{\partial \log \tau}{\partial p_2}$. Then $\frac{\partial \log \widetilde{\tau}}{\partial p_2} = 0$, so the corresponding dressing operator $\widetilde{W}$ satisfies

$$0 = \frac{\partial \widetilde{W}}{\partial p_2} = -L_-^2 \circ \widetilde{W},$$

so $L_-^2 = 0$, i.e. τ is a tau-function of the KdV hierarchy. □

Thanks to the above corollary and the transformation (4.41) for the tau-function, we deduce that solutions of the KdV hierarchy are in bijective correspondence with tau-functions of the KP hierarchy that do not depend on the even variables. This makes sense, because in the last section we identified the KdV variables t_j with the odd KP variables p_{2j+1}. Moreover, the statement of the Witten-Kontsevich theorem (Theorem 2.6.3) says that the generating series $F(t_*)$ for intersection numbers of ψ-classes is a tau-function of the KdV hierarchy.

To conclude this section we give an answer to the problem that we introduced at the start, namely how to encode the countably many PDEs for the countably many variables of the KP hierarchy into a single equation for a single function. It would make sense for this single function to be the tau-function, since the equality

$$\operatorname{res} L^i = \frac{\partial^2 \log \tau}{\partial p_1 \partial p_i}, \quad i \geq 1,$$

from Proposition 4.3.6 implies that one can obtain all the values of the normal coordinates $\operatorname{res} L^i$ from τ. The equation that we want will be a corollary of the bilinear identity (4.31). Firstly, since the dressing operator W associated to τ

satisfies $\widehat{W} = \frac{G_z(\tau)}{\tau}$, the corresponding wave function is

$$\psi(p_*, z) = \frac{G_z(\tau)}{\tau} e^{\xi(p_*,z)} = \frac{\tau\left(p_1 - \frac{1}{z}, p_2 - \frac{1}{2z^2}, p_3 - \frac{1}{3z^3}, \dots\right)}{\tau(p_1, p_2, p_3, \dots)} e^{\xi(p_*,z)}. \tag{4.42}$$

Now introduce an operator $G_z^\dagger : \mathbb{C}[[p_*]] \to \mathbb{C}[[p_*]][[z^{-1}]]$ inverse to (4.34), namely

$$f(p_*) \longmapsto f\left(p_1 + \frac{1}{z}, p_2 + \frac{1}{2z^2}, p_3 + \frac{1}{3z^3}, \dots\right). \tag{4.43}$$

The property $G_z(W^\dagger) = \widehat{W}^{-1}$ from equation (4.35) implies that

$$\widehat{W}^\dagger = G_z^\dagger(G_z(\widehat{W}^\dagger)) \overset{(4.35)}{=} G_z^\dagger(\widehat{W}^{-1}) = G_z^\dagger\left(\frac{\tau}{G_z(\tau)}\right) = \frac{G_z^\dagger(\tau)}{\tau}.$$

Hence the adjoint wave function is

$$\psi^\dagger(p_*, z) = \frac{G_z^\dagger(\tau)}{\tau} e^{-\xi(p_*,z)} = \frac{\tau\left(p_1 + \frac{1}{z}, p_2 + \frac{1}{2z^2}, p_3 + \frac{1}{3z^3}, \dots\right)}{\tau(p_1, p_2, p_3, \dots)} e^{-\xi(p_*,z)}. \tag{4.44}$$

Corollary 4.3.8 (Hirota bilinear identity)**.** *Let $\tau(p_*) \in \mathbb{C}[[p_*]]$ satisfy $\tau(0) = 1$. Then τ is a tau-function of the KP hierarchy if and only if*

$$\operatorname{res}_z\left(\tau\left(p_1 - \frac{1}{z}, p_2 - \frac{1}{2z^2}, \dots\right) \cdot \tau\left(p_1' + \frac{1}{z}, p_2' + \frac{1}{2z^2}, \dots\right) e^{\xi(p_*,z) - \xi(p_*',z)}\right) = 0, \tag{4.45}$$

where $p_*' = p_* + \lambda_*$.

Proof. Suppose τ is a tau-function of the KP hierarchy. By Theorem 4.3.4, the wave functions $\psi(p_*, z)$ and $\psi^\dagger(p_*, z)$, given by equations (4.42) and (4.44), satisfy the bilinear identity

$$\operatorname{res}_z(\psi(p_*, z) \cdot \psi^\dagger(p_*', z)) = 0,$$

which is equivalent to (4.45). Conversely, suppose τ satisfies (4.45). Then $\gamma := \frac{G_z(\tau)}{\tau} e^{\xi(p_*,z)}$ and $\rho := \frac{G_z^\dagger(\tau)}{\tau} e^{-\xi(p_*,z)}$ satisfy the bilinear identity

$$\operatorname{res}_z\left(\gamma(p_*, z) \cdot \rho(p_*', z)\right) = 0.$$

Then by Theorem 4.3.4, γ and ρ are wave functions associated to a solution W of the SW hierarchy. Hence τ is a tau-function of the KP hierarchy. $\square$

We explain how the Hirota bilinear identity (4.45) encodes all the PDEs of the KP hierarchy (4.15), following the ideas in [MJD00, Chapter 3]. To do this, we use the *Hirota derivative.* For two functions $f(p)$ and $g(p)$ of a single variable p, the Hirota derivatives $D_p^i f \cdot g$, $i \geq 0$, are defined by the Taylor expansion of $f(p+q)g(p-q)$ around $q=0$:

$$\exp(qD_p)f \cdot g := \sum_{i\geq 0} \frac{q^i}{i!}(D_p^i f \cdot g) = f(p+q)g(p-q).$$

In other words

$$\frac{1}{i!}D_p^i f \cdot g = \sum_{\substack{n,m\geq 0\\ n+m=i}} \frac{(-1)^m}{n!m!}\frac{\partial^n f}{\partial p^n}\frac{\partial^m g}{\partial p^m}.$$

For many independent variables $p_1, p_2, \ldots$ the Hirota derivatives $D_{p_{j_1}}^{i_1}\cdots D_{p_{j_k}}^{i_k} f \cdot g := D_{j_1}^{i_1}\cdots D_{j_k}^{i_k} f \cdot g$ are defined in a similar way:

$$\exp\left(\sum_{j=1}^{\infty} q_j D_j\right) f \cdot g = f(p_1+q_1, p_2+q_2, \ldots)g(p_1-q_1, p_2-q_2, \ldots). \tag{4.46}$$

We warn that the D_j in (4.46) are not to be confused with the evolutionary operators of the KP hierarchy from last section. In this case the standalone symbol D_j does not mean anything, since it is not a differential operator. The notation established above requires the presence of two functions f and g in the definition, and so $D_j f \cdot g$ should be interpreted as a single symbol. Consequently, the expression $D_{j_1}D_{j_2} f \cdot g$ is not to be interpreted as $D_{j_1}(D_{j_2} f \cdot g)$, since this last expression has no meaning.

Now we shift the variables p_* in the Hirota bilinear identity (4.45) by $p_* \mapsto p_* + \frac{1}{2}\lambda_*$ and define $p'_* = p_* - \frac{1}{2}\lambda_*$, so that the product of tau-functions has the form of equation (4.46):

$$\tau\Big(p_j + \underbrace{\big(\frac{1}{2}\lambda_j - \frac{1}{jz^j}\big)}_{q_j}\Big)_{j\geq 1} \tau\Big(p_j - \underbrace{\big(\frac{1}{2}\lambda_j - \frac{1}{jz^j}\big)}_{q_j}\Big)_{j\geq 1}.$$

Then, using $\xi(p_* + \frac{1}{2}\lambda_*, z) - \xi(p_* - \frac{1}{2}\lambda_*, z) = \sum_{j\geq 1}\lambda_j z^j$, the Hirota bilinear identity can be written as

$$\operatorname{res}_z\left(\exp\Big(\sum_{j\geq 1}\lambda_j z^j\Big)\exp\Big(\sum_{j\geq 1}\Big(\frac{\lambda_j}{2} - \frac{1}{jz^j}\Big)D_j\Big)\tau\cdot\tau\right) = 0. \tag{4.47}$$

Expanding the two exponentials and taking the coefficient of z^{-1} yields all the equations of the KP hierarchy, which appear as coefficients of monomials in the λ_*

variables, which are of the form

$$P(D_1, D_2, D_3, \dots)\tau \cdot \tau = 0 \tag{4.48}$$

for some polynomial P. Define the weight of the variable λ_j to be j. If the monomial in λ_* has weighted degree d, then the polynomial P in its coefficient (4.48) has weighted degree $d+1$, where D_j is given the weight j.

As an example, we compute the coefficients of all the monomials in λ_* of weighted degree ≤ 3. Firstly,

$$\begin{aligned}\exp\Big(\sum_{j\geq 1}\lambda_j z^j\Big) =& 1 + (\lambda_1 z + \lambda_2 z^2 + \lambda_3 z^3 + \cdots) + \frac{1}{2!}(\lambda_1 z + \lambda_2 z^2 + \cdots)^2 \\ &+ \frac{1}{3!}(\lambda_1 z + \cdots)^3 + \cdots \\ =& 1 + \lambda_1 z + \left(\lambda_2 + \frac{1}{2}\lambda_1^2\right) z^2 + \left(\lambda_3 + \lambda_1\lambda_2 + \frac{1}{6}\lambda_1^3\right) z^3 + \cdots,\end{aligned} \tag{4.49}$$

where we have ignored all terms in z^4 and above. For the second exponential, first notice that the Hirota derivatives $D_{j_1}^{i_1}\cdots D_{j_k}^{i_k}\tau\cdot\tau$ are zero when $i_1 + \cdots + i_k$ is odd. For example,

$$\begin{aligned} D_j\tau\cdot\tau &= \frac{\partial\tau}{\partial p_j}\tau - \tau\frac{\partial\tau}{\partial p_j} = 0, \\ D_j^3\tau\cdot\tau &= \frac{\partial^3\tau}{\partial p_j^3}\tau - 3\frac{\partial^2\tau}{\partial p_j^2}\frac{\partial\tau}{\partial p_j} + 3\frac{\partial\tau}{\partial p_j}\frac{\partial^2\tau}{\partial p_j^2} - \tau\frac{\partial^3\tau}{\partial p_j^3} = 0. \end{aligned}$$

So when expanding the second exponential in (4.47), we need only take the terms of even power.

$$\begin{aligned} &\exp\left(\sum_{j\geq 1}\Big(\frac{\lambda_j}{2} - \frac{1}{jz^j}\Big)D_j\right) \\ =& 1 + \frac{1}{2!}\left(\left(\frac{1}{2}\lambda_1 - \frac{1}{z}\right)D_1 + \left(\frac{1}{2}\lambda_2 - \frac{1}{2z^2}\right)D_2 + \left(\frac{1}{2}\lambda_3 - \frac{1}{3z^3}\right)D_3 + \cdots\right)^2 \\ &+ \frac{1}{4!}\left(\left(\frac{1}{2}\lambda_1 - \frac{1}{z}\right)D_1 + \cdots\right)^4 + \cdots \\ =& -\frac{1}{2}\lambda_1 D_1^2 z^{-1} + \frac{1}{2}D_1^2 z^{-2} \\ &- \frac{1}{2}\lambda_2 D_1 D_2 z^{-1} - \frac{1}{4}\lambda_1 D_1 D_2 z^{-2} + \frac{1}{2}D_1 D_2 z^{-3} \\ &+ \left(-\frac{1}{2}\lambda_3 D_1 D_3 - \frac{1}{12}\lambda_1^3 D_1^4\right)z^{-1} + \left(-\frac{1}{4}\lambda_2 D_2^2 + \frac{1}{16}\lambda_1^2 D_1^4\right)z^{-2} \\ &+ \left(-\frac{1}{6}\lambda_1 D_1 D_3 - \frac{1}{12}\lambda_1 D_1^4\right)z^{-3} + \left(\frac{1}{8}D_2^2 + \frac{1}{3}D_1 D_3 + \frac{1}{24}D_1^4\right)z^{-4} + \cdots, \end{aligned}$$

where we have ignored all terms without z^{-1}, z^{-2}, z^{-3} or z^{-4}. When we multiply the above with (4.49) and only keep the z^{-1} term in accordance with the Hirota bilinear identity, all the coefficients of monomials in λ_* of weighted degree 1 and 2 (i.e. of λ_1, λ_1^2 and λ_2) vanish, and we are left with

$$\begin{aligned}0 =& \frac{1}{144}\lambda_1^3\left(D_1^4 + 3D_2^2 - 4D_1D_3\right)\tau\cdot\tau - \frac{1}{24}\lambda_1\lambda_2\left(D_1^4 + 3D_2^2 - 4D_1D_3\right)\tau\cdot\tau\\ &+ \frac{1}{24}\lambda_3\left(D_1^4 + 3D_2^2 - 4D_1D_3\right)\tau\cdot\tau.\end{aligned}$$

Hence we obtain the single Hirota bilinear PDE $(D_1^4 + 3D_2^2 - 4D_1D_3)\tau\cdot\tau = 0$. Using the readily checked facts

$$D_{j_1}D_{j_2}\tau\cdot\tau = 2\tau^2\frac{\partial^2\log\tau}{\partial p_{j_1}\partial p_{j_2}}, \quad D_j^4\tau\cdot\tau = 2\tau^2\frac{\partial^2\log\tau}{\partial p_j^4} + 6\tau^2\left(\frac{\partial^2\log\tau}{\partial p_j^2}\right)^2,$$

we see that this PDE is exactly

$$\frac{\partial^2\log\tau}{\partial p_2^2} = \frac{4}{3}\frac{\partial^2\log\tau}{\partial p_1\partial p_3} - 2\left(\frac{\partial^2\log\tau}{\partial p_1^2}\right)^2 - \frac{1}{3}\frac{\partial^4\log\tau}{\partial p_1^4}.$$

This is the original KP equation, which we already computed in (4.21) and (4.39).

In general, taking more terms in z when expanding the Hirota bilinear identity yields the subsequent equations of the KP hierarchy in the form $P(D_1, D_2, D_3, \dots)\tau\cdot\tau = 0$, where P is an even polynomial of weighted degree ≥ 5. In general [KM81], these Hirota bilinear equations have the form

$$\det\begin{pmatrix}\widetilde{p}_{\mu_1+1}\left(-\frac{\widetilde{D}}{2}\right) & \widetilde{p}_{\mu_1+1}\left(\frac{\widetilde{D}}{2}\right) & \cdots & \widetilde{p}_{\mu_1+\ell-1}\left(\frac{\widetilde{D}}{2}\right)\\ \widetilde{p}_{\mu_2}\left(-\frac{\widetilde{D}}{2}\right) & \widetilde{p}_{\mu_2}\left(\frac{\widetilde{D}}{2}\right) & \cdots & \widetilde{p}_{\mu_2+\ell-2}\left(\frac{\widetilde{D}}{2}\right)\\ \vdots & \vdots & \ddots & \vdots\\ \widetilde{p}_{\mu_\ell-\ell+2}\left(-\frac{\widetilde{D}}{2}\right) & \widetilde{p}_{\mu_\ell-\ell+2}\left(\frac{\widetilde{D}}{2}\right) & \cdots & \widetilde{p}_{\mu_\ell}\left(\frac{\widetilde{D}}{2}\right)\end{pmatrix}\tau\cdot\tau = 0, \qquad (4.50)$$

where $\mu_1 \geq \mu_2 \geq \cdots \geq \mu_\ell \geq 1$, $\ell \geq 2$, $\widetilde{D} = (D_1, D_2/2, D_3/3, \dots)$ and the polynomials $\widetilde{p}_i$ are defined by

$$\exp\Big(\sum_{j\geq 1}\lambda_j z^j\Big) = \sum_{i\geq 1}\widetilde{p}_i(\lambda_*)z^i.$$

The independent PDEs of the KP hierarchy obtained from (4.50) are indexed by two-part partitions $(\mu_1, \mu_2) \in \mathcal{P}_d$ of $d \geq 4$ (corresponding to the weighted degree of P) that do not contain a 1 [Mvo22]. For $d = 4$ there is a single equation corresponding to the partition $(2, 2)$, and for $d = 5$ there is also only one corresponding to $(3,2)$. For $d = 6$ there are two arising from $(4,2)$ and $(3,3)$, and so on.

4.4 Fock spaces

In the last section we introduced a way of encoding all the data of the KP hierarchy into a single equation for a single function. The function is the tau-function $\tau(p_*) \in \mathbb{C}[[p_*]]$, which is defined as the unique power series for which we can express the wave function $\psi(p_*, z) = \widehat{W} e^{\xi(p_*,z)}$ as

$$\psi(p_*, z) = \frac{G_z(\tau)}{\tau} e^{\xi(p_*,z)} = \frac{\tau\left(p_1 - \frac{1}{z}, p_2 - \frac{1}{2z^2}, \dots\right)}{\tau(p_1, p_2, \dots)} e^{\xi(p_*,z)}.$$

The equation is the Hirota bilinear identity from Corollary 4.3.8:

$$\operatorname{res}_z \left(\tau\left(p_1 - \frac{1}{z}, p_2 - \frac{1}{2z^2}, \dots\right) \tau\left(p_1' + \frac{1}{z}, p_2' + \frac{1}{2z^2}, \dots\right) e^{\xi(p_*,z) - \xi(p_*',z)} \right) = 0.$$

These statements are pertinent to the space $\mathbb{C}[p_*]$, or rather its completion $\mathbb{C}[[p_*]]$. In what follows, the space of polynomials in the variables p_* will be referred to as *charge zero Bosonic Fock space*. In general, Bosonic Fock space $\mathbb{C}[p_*][z, z^{-1}]$ includes an additional variable z. The differentiation and multiplication operators $\frac{\partial}{\partial p_n}$ and p_n act in the usual way on $\mathbb{C}[p_*]$. The commutation relations

$$\left[\frac{\partial}{\partial p_n}, \frac{\partial}{\partial p_m}\right] = 0, \quad [p_n, p_m] = 0, \quad \left[\frac{\partial}{\partial p_n}, p_m\right] = \delta_{nm} := \begin{cases} 1, & n = m, \\ 0, & n \neq m, \end{cases}$$

define an algebra structure on the space generated by these operators. This algebra is called the *Heisenberg algebra* $\mathcal{B}$. The element $1 \in \mathbb{C}[p_*]$ is called the *vacuum state* and it generates the charge zero Bosonic Fock space, in the sense that $\mathbb{C}[p_*] = \mathcal{B} \cdot 1$. The differentiation operators $\frac{\partial}{\partial p_n} \in \mathcal{B}$ are referred to as *annihilation operators* since they annihilate the vacuum state: $\frac{\partial}{\partial p_n} \cdot 1 = 0$. The multiplication operators $p_n \in \mathcal{B}$ are referred to as *creation operators* and $\mathbb{C}[p_*]$ has a linear basis obtained from the action of the creation operators on the vacuum state: $\{p_{n_1} p_{n_2} \cdots p_{n_r} \cdot 1\}$.

The reason for introducing this notation and nomenclature is because there is another Fock space $\mathcal{F}^0$, called *charge zero Fermionic Fock space*. The Bosonic representation of the Heisenberg algebra $\mathcal{B}$ can be identified with the Fermionic representation of the Clifford algebra $\mathcal{A}$, via the *Boson-Fermion correspondence*. In particular, this provides a linear isomorphism between $\mathbb{C}[p_*]$ and $\mathcal{F}^0$. In this section, we introduce the Clifford algebra $\mathcal{A}$ and its representation in Fermionic Fock space, and derive the Fermionic counterparts of the tau-function and Hirota bilinear identity obtained from the Boson-Fermion correspondence. This is useful because it gives one a relatively straightforward way of checking whether a function is actually a tau-function of the KP hierarchy (i.e. whether it satisfies the Hirota bilinear identity), which we will use to deduce that the disconnected Hurwitz potential

$$H^\bullet(\beta; p_*) = \sum_\mu s_\mu(1, 0, 0, \dots) s_\mu(p_*) e^{f_2(\mu)\beta} \in \mathbb{C}[[p_*]][[\beta]]$$

from Sections 3.3-3.4 is a tau-function of the KP hierarchy. Most of this section is adapted from [MJD00, Chapters 4-9], and a brief but detailed overview can be found in [JM83, §1-§2].

Consider the *Clifford algebra* $\mathcal{A}$ generated by $\{a_i, a_i^\dagger : i \in \mathbb{Z} + \frac{1}{2}\}$ with the anticommutation relations

$$\{a_i, a_j\} = 0, \quad \{a_i^\dagger, a_j^\dagger\} = 0, \quad \{a_i, a_j^\dagger\} = \delta_{i+j,0} = \begin{cases} 1, & i = -j, \\ 0, & i \neq -j. \end{cases} \tag{4.51}$$

The anticommutator $\{\cdot, \cdot\}$ is defined by $\{A, B\} := AB + BA$. We call the elements a_i and $a_i^\dagger$ *Fermions*, and they are indexed by half-integers. The relations above imply, in particular, that $a_i^2 = (a_i^\dagger)^2 = 0$. Note that the authors in [MJD00] denote the Fermions by ψ_i and ψ_i^*, but we have opted for the notation above to avoid confusion with the KP wave functions ψ and $\psi^\dagger$. By successively using the anticommutation relations to transpose the order of products of Fermions, one can write every element of the Clifford algebra $\mathcal{A}$ as a linear combination of monomials of the form

$$a_{i_1} \cdots a_{i_r} a_{j_1}^\dagger \cdots a_{j_s}^\dagger, \quad i_1 < \cdots < i_r \text{ and } j_1 < \cdots < j_s.$$

The set of elements of this form form a linear basis of $\mathcal{A}$ [Bou59, Section 9.3].

Just like the Heisenberg algebra $\mathcal{B}$ acts on Bosonic Fock space, the Clifford algebra $\mathcal{A}$ acts on Fermionic Fock space $\mathcal{F}$. The space $\mathcal{F}$ is generated by increasing sequences of half-integers $\lambda = (\lambda_1, \lambda_2, \lambda_3, \dots)$ such that $\lambda_{j+1} = \lambda_j + 1$ for all j sufficiently large. We write them as

$$|\lambda\rangle = |\lambda_1, \lambda_2, \lambda_3, \dots\rangle.$$

We can represent each generator $|\lambda\rangle$ by a *Maya diagram*: a row of black and white Go stones, where far away enough to the right all the stones are black, and far away enough to the left all the stones are white. The components $\lambda_1, \lambda_2, \lambda_3, \dots$

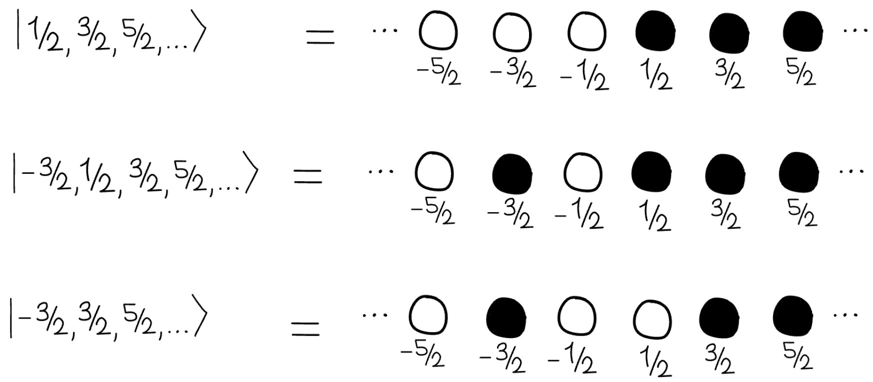

Figure 9: Maya diagrams.

of $|\lambda\rangle$ indicate the positions of the black stones, and all the other positions are occupied by white stones. See Figure 9 for some examples. The vacuum state $|\mathrm{vac}\rangle := |\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots\rangle$ corresponds to the Maya diagram where all the stones to the left of 0 are white, and all the stones to the right of 0 are black. The action of $a_{-i} \in \mathcal{A}, i \in \mathbb{Z} + \frac{1}{2}$, on a vector $|\lambda\rangle = |\lambda_1, \lambda_2, \lambda_3, \dots\rangle \in \mathcal{F}$ is given as follows:

$$a_{-i}|\lambda\rangle = \begin{cases} (-1)^{j-1}|\lambda_1, \dots, \lambda_{j-1}, \lambda_{j+1}, \dots\rangle, & \text{if } \lambda_j = i \text{ for some } j \geq 1, \\ 0, & \text{else.} \end{cases}$$

The first expression is to be understood as $|\lambda_2, \lambda_3, \dots\rangle$ if $\lambda_1 = i$. In terms of Maya diagrams, if $|\lambda\rangle$ contains a black stone at the i^{th} position then the action of the fermion a_{-i} replaces it with a white stone. Otherwise it sends the vector $|\lambda\rangle$ to zero. On the other hand the action of $a_i^\dagger \in \mathcal{A}$ is

$$a_i^\dagger|\lambda\rangle = \begin{cases} (-1)^{j}|\lambda_1, \dots, \lambda_j, i, \lambda_{j+1}, \dots\rangle, & \text{if } \lambda_j < i < \lambda_{j+1} \text{ for some } j \geq 1, \\ 0, & \text{else,} \end{cases}$$

The first expression is to be understood as $|i, \lambda_1, \lambda_2, \dots\rangle$ if $i < \lambda_1$. In this case if $|\lambda\rangle$ contains a white stone at the i^{th} position then the action of $a_i^\dagger$ replaces it with a black stone. If not, it sends $|\lambda\rangle$ to zero. Thanks to the powers of -1 in the above definitions these actions are compatible with the anticommuation relations (4.51), so that they define a representation of the Clifford algebra $\mathcal{A}$ on $\mathcal{F}$. As an example, the action of a_i and $a_i^\dagger$ on the vacuum state $|\mathrm{vac}\rangle = |\frac{1}{2}, \frac{3}{2}, \frac{5}{2}, \dots\rangle$ is

$$a_i|\mathrm{vac}\rangle = \begin{cases} (-1)^{i+1}|\frac{1}{2}, \dots, -i-1, -i+1, \dots\rangle, & \text{if } i < 0, \\ 0, & \text{if } i > 0, \end{cases}$$

$$a_i^\dagger|\mathrm{vac}\rangle = \begin{cases} (-1)^{i}|i, \frac{1}{2}, \frac{3}{2}, \dots\rangle, & \text{if } i < 0, \\ 0, & \text{if } i > 0. \end{cases}$$

From this we see that the fermions $\{a_i, a_i^\dagger\}_{i>0}$ indexed by positive half-integers annihilate the vacuum state, and for this reason they are called *annihilation operators*. On the other hand the fermions $\{a_i, a_i^\dagger\}_{i<0}$ indexed by negative half-integers are called *creation* operators. It is clear that every Maya diagram can be obtained by starting with the vacuum state diagram and switching white and black stones a finite amount of times, so that every generator $|\lambda\rangle \in \mathcal{F}$ is obtained (up to a sign) by applying creation operators to $|\mathrm{vac}\rangle$:

$$|\lambda\rangle = \pm\, a_{i_1} \cdots a_{i_r} a_{j_1}^\dagger \cdots a_{j_s}^\dagger |\mathrm{vac}\rangle, \quad i_1 < \cdots < i_r < 0 \text{ and } j_1 < \cdots < j_s < 0. \quad (4.52)$$

Such elements form a linear basis of $\mathcal{F}$. The *charge* of $|\lambda\rangle$ given above is $\ell := r - s \in \mathbb{Z}$, or equivalently $\ell := \lim_{j\to\infty}(\lambda_j - j + \frac{1}{2})$. Intuitively, the charge measures how many times a white stone has replaced a black stone when starting from the vacuum state. For each $\ell \in \mathbb{Z}$, we define $|\ell\rangle := |\ell + \frac{1}{2}, \ell + \frac{3}{2}, \ell + \frac{5}{2}, \dots\rangle$ as the charge ℓ Maya diagram obtained by shifting the vacuum state ℓ steps to the right, or $-\ell$ steps to the left if $\ell < 0$:

$$|\ell\rangle = \begin{cases} a_{\ell+\frac{1}{2}}^\dagger \cdots a_{-\frac{1}{2}}^\dagger |\mathrm{vac}\rangle, & \text{if } \ell < 0, \\ |\mathrm{vac}\rangle, & \text{if } \ell = 0, \\ a_{-\ell+\frac{1}{2}} \cdots a_{-\frac{1}{2}} |\mathrm{vac}\rangle, & \text{if } \ell > 0. \end{cases}$$

Let $\mathcal{F}^\ell \subset \mathcal{F}$ be the subspace spanned by the charge ℓ Maya diagrams. Then $\mathcal{F}$ decomposes as the direct sum

$$\mathcal{F} = \bigoplus_{\ell\in\mathbb{Z}} \mathcal{F}^\ell.$$

The dual Fermionic Fock space $\mathcal{F}^*$ is defined dually to $\mathcal{F}$. It is generated by decreasing sequences of half-integers $\eta = (\eta_1, \eta_2, \eta_3, \dots)$ such that $\eta_{j+1} = \eta_j - 1$ for j large enough. We write them as

$$\langle\eta| = \langle\dots, \eta_3, \eta_2, \eta_1|.$$

In terms of Maya diagrams, each η_j denotes the position of a white stone. The fermion $a_i \in \mathcal{A}, i \in \mathbb{Z}+\frac{1}{2}$, acts on $\langle\eta|$ by replacing a black stone at the i^{th} position (if there is one) with a white one:

$$\langle\eta|a_i = \begin{cases} (-1)^j \langle\ldots,\eta_{j+1},i,\eta_j,\ldots,\eta_1|, & \text{if } \eta_{j+1} < i < \eta_j \text{ for some } j \geq 1, \\ 0, & \text{else,} \end{cases}$$

where the first expression is to be understood as $\langle\ldots,\eta_2,\eta_1,i|$ if $\eta_1 < i$. The action of $a^\dagger_{-i}$ replaces a white stone at the i^{th} position with a black one:

$$\langle\eta|a^\dagger_{-i} = \begin{cases} (-1)^{j-1} \langle\ldots,\eta_{j+1},\eta_{j-1},\ldots,\eta_1|, & \text{if } \eta_j = i \text{ for some } j \geq 1, \\ 0, & \text{else,} \end{cases}$$

where the first expression is to be understood as $\langle\ldots,\eta_3,\eta_2|$ if $\eta_1 = i$. As before, the vacuum state $\langle\text{vac}| = \langle\ldots,-\frac{5}{2},-\frac{3}{2},-\frac{1}{2}|$ corresponds to the Maya diagram where all the stones to the left of 0 are white and all the stones to the right of 0 are black. The fermions $\{a_i, a_i^\dagger\}_{i<0}$ with negative index annihilate $\langle\text{vac}|$ in this case. Hence each Maya diagram $\langle\eta| \in \mathcal{F}^*$ is of the form

$$\langle\eta| = \pm\,\langle\text{vac}|a_{i_1}\cdots a_{i_r}a^\dagger_{j_1}\cdots a^\dagger_{j_s}, \quad 0 < i_1 < \cdots < i_r \text{ and } 0 < j_1 < \cdots < j_s.$$

The charge of $\langle\eta|$ given by the above equation is also defined as $\ell := r - s$. The standard charge ℓ Maya diagram $\langle\ell| := \langle\ldots,\ell-\frac{5}{2},\ell-\frac{3}{2},\ell-\frac{1}{2}|$ is therefore given by

$$\langle\ell| = \begin{cases} \langle\text{vac}|a_{\frac{1}{2}}\cdots a_{-\ell-\frac{1}{2}}, & \text{if } \ell < 0, \\ \langle\text{vac}|, & \text{if } \ell = 0, \\ \langle\text{vac}|a^\dagger_{\frac{1}{2}}\cdots a^\dagger_{\ell-\frac{1}{2}}, & \text{if } \ell > 0. \end{cases} \tag{4.53}$$

There is a pairing $\mathcal{F}^* \times \mathcal{F} \to \mathbb{C}$, which is given by

$$(\langle\eta|,|\lambda\rangle) \longmapsto \langle\eta|\lambda\rangle := \delta_{\lambda_1+\eta_1,0}\,\delta_{\lambda_2+\eta_2,0}\,\delta_{\lambda_3+\eta_3,0}\cdots$$

on the generators, and extended linearly to all elements $(\langle u|,|v\rangle) \in \mathcal{F}^* \times \mathcal{F}$. It satisfies $\langle\text{vac}|\text{vac}\rangle = 1$ and $(\langle u|a)|v\rangle = \langle u|(a|v\rangle)$ for all $a \in \mathcal{A}$, so we denote the latter by $\langle u|a|v\rangle$. Moreover, we denote the *vacuum expectation value* $\langle\text{vac}|a|\text{vac}\rangle$ by $\langle a\rangle$ for short. The first few examples of vacuum expectation values are

$$\langle 1\rangle = 1, \quad \langle a_i\rangle = 0, \quad \langle a_i^\dagger\rangle = 0, \quad \langle a_i a_j\rangle = 0, \quad \langle a_i^\dagger a_j^\dagger\rangle = 0, \tag{4.54}$$

since in each case (except the first) one of the fermions will act as an annihilation operator on the left or on the right; for example either $a_i|\mathrm{vac}\rangle = 0$ or $\langle\mathrm{vac}|a_i = 0$, so we must have $\langle a_i\rangle = 0$. The vacuum expectation value $\langle a_i a_j^\dagger\rangle$ can only be nonzero if $j < 0 < i$, in which case we use the anticommutation relation (4.51) to get

$$\langle a_i a_j^\dagger\rangle = \langle(\{a_i, a_j^\dagger\} - a_j^\dagger a_i)\rangle = \delta_{i+j,0}\langle 1\rangle - 0 = \delta_{i+j,0}. \tag{4.55}$$

In the spirit of the previous sections, we introduce the formal variable z and the generating functions

$$a(z) := \sum_{i\in\mathbb{Z}+\frac{1}{2}} a_i z^{-i-\frac{1}{2}} \quad \text{and} \quad a^\dagger(z) := \sum_{i\in\mathbb{Z}+\frac{1}{2}} a_i^\dagger z^{i-\frac{1}{2}}$$

in order to keep track of a large number of fermions at once. From the equations in (4.54) we obtain $\langle a(z)a(z')\rangle = \langle a^\dagger(z)a^\dagger(z')\rangle = 0$, whereas (4.55) gives

$$\begin{aligned}\langle a(z)a^\dagger(z')\rangle &= \sum_{i,j\in\mathbb{Z}+\frac{1}{2}} \langle a_i a_j^\dagger\rangle z^{-i-\frac{1}{2}} z'^{-j-\frac{1}{2}} = \sum_{j<0<i} \delta_{i+j,0} z^{-i-\frac{1}{2}} z'^{-j-\frac{1}{2}} \\ &= \sum_{i>0} z^{-i-\frac{1}{2}} z'^{i-\frac{1}{2}} = \sum_{n\geq 0} z^{-n-1} z'^{n} = \frac{1}{z-z'}.\end{aligned}$$

When considering the representation of a product of Fermions in $\mathcal{F}$ it is useful to have the annihilation operators on the right and the creation operators on the left. To justify this statement, let us first look at the Heisenberg algebra $\mathcal{B}$ and its left action on Bosonic Fock space $\mathbb{C}[p_*]$. Let $E = \sum_{n\geq 1} np_n \frac{\partial}{\partial p_n}$ be the Cauchy-Euler operator. Despite this being an infinite sum, its left action on $\mathbb{C}[p_*]$ is well-defined because $E \cdot f(p_*)$ is a finite sum for any polynomial $f(p_*)$. This is because all the annihilation operators $\frac{\partial}{\partial p_n} \in \mathcal{B}$ are on the right and all the creation operators $p_n \in \mathcal{B}$ are on the left. If we consider $E' = \sum_{n\geq 1} n \frac{\partial}{\partial p_n} p_n$ instead, the commutation relation $\left[\frac{\partial}{\partial p_n}, p_n\right] = 1$ gives $E' = \sum_{n\geq 1} n \left(1 + p_n \frac{\partial}{\partial p_n}\right)$, which gives rise to an infinite sum when acting on $1 \in \mathbb{C}[p_*]$, for example. For this reason, we introduce a *normal product* on the algebra $\mathcal{B}$, denoted by colons. It is defined inductively by

$$:1: \;=\; 1, \quad :b\frac{\partial}{\partial p_n}: \;=\; :b:\frac{\partial}{\partial p_n}, \quad :p_n b: \;=\; p_n :b:, \quad b \in \mathcal{B},$$

and by imposing the condition that creation and annihilation operators inside the colons commute with each other. Then, for example,

$$:\frac{\partial}{\partial p_n}p_n: \;=\; :p_n\frac{\partial}{\partial p_n}: \;=\; p_n\frac{\partial}{\partial p_n}.$$

We define the normal product in the Clifford algebra $\mathcal{A}$ in the same way, by placing the annihilation operators $\{a_i, a_i^\dagger\}_{i>0}$ on the right and the creation operators $\{a_i, a_i^\dagger\}_{i<0}$ on the left. This time we impose the condition that annihilation and creation operators *anti*commute with each other inside the colons. For example, if $i < 0 < j$ then $: a_i a_j^\dagger : \;=\; a_i a_j^\dagger$. If $j < 0 < i$ on the other hand, then

$$: a_i a_j^\dagger : \;=\; - : a_j^\dagger a_i : \;=\; -a_j^\dagger a_i \;=\; a_i a_j^\dagger - \{a_i, a_j^\dagger\}.$$

So for any $i, j \in \mathbb{Z} + \frac{1}{2}$ we can use the vacuum expectation value (4.55) to write

$$: a_i a_j^\dagger : \;=\; a_i a_j^\dagger - \langle a_i a_j^\dagger \rangle.$$

We define another generating series by

$$\sum_{n \in \mathbb{Z}} H_n z^{-n-1} \;=\; : a(z) a^\dagger(z) : \;=\; \sum_{i,j \in \mathbb{Z}+\frac{1}{2}} : a_i a_j^\dagger : z^{-i-j-1},$$

so that each coefficient is

$$H_n = \sum_{i \in \mathbb{Z}+\frac{1}{2}} : a_{-i} a_{i+n}^\dagger : \;. \tag{4.56}$$

From the previous discussion, the presence of normal products in this infinite sum ensures that the action of each H_n on $\mathcal{F}$ gives rise to well-defined finite sums. The following commutation relations hold:

$$[H_n, a_i] = a_{i+n}, \quad [H_n, a_i^\dagger] = -a_{i+n}^\dagger, \quad [H_n, H_m] = n\delta_{m+n,0}. \tag{4.57}$$

Let

$$H(p_*) := \sum_{n \geq 1} p_n H_n.$$

Using the commutation relations (4.57) between H_n and the fermions $a_i, a_i^\dagger$, we can compute the corresponding commutators between $H(p_*)$ and $a(z), a^\dagger(z)$:

$$\begin{aligned} [H(p_*), a(z)] &= \sum_{n \geq 1} \sum_{i \in \mathbb{Z}+\frac{1}{2}} p_n [H_n, a_i] z^{-i-\frac{1}{2}} = \sum_{n \geq 1} \sum_{i \in \mathbb{Z}+\frac{1}{2}} p_n a_{i+n} z^{-i-\frac{1}{2}} \\ &= \sum_{n \geq 1} p_n z^n \sum_{i \in \mathbb{Z}+\frac{1}{2}} a_i z^{-i-\frac{1}{2}} = \xi(p_*, z) a(z), \end{aligned}$$

where $\xi(p_*, z) = \sum_{n\geq 1} p_n z^n$ was introduced in Section 4.3. Hence

$$\begin{aligned} e^{H(p_*)}a(z) &= \sum_{k\geq 0}\frac{1}{k!}H(p_*)^k a(z) = \sum_{k\geq 0}\frac{1}{k!}a(z)(H(p_*)+\xi(p_*,z))^k \\ &= a(z)e^{H(p_*)+\xi(p_*,z)} = a(z)e^{H(p_*)}e^{\xi(p_*,z)}. \end{aligned} \tag{4.58}$$

A similar procedure yields

$$[H(p_*), a^\dagger(z)] = -\xi(p_*,z)a^\dagger(z) \quad \text{and} \quad e^{H(p_*)}a^\dagger(z) = a^\dagger(z)e^{-\xi(p_*,z)}. \tag{4.59}$$

We can now relate Fermionic Fock space $\mathcal{F}$ and Bosonic Fock space $\mathbb{C}[p_*][z,z^{-1}]$ [MJD00, Theorem 5.1].

Theorem 4.4.1. *The map* $\Phi : \mathcal{F} \to \mathbb{C}[p_*][z,z^{-1}]$ *given by*

$$\Phi(|u\rangle) := \sum_{\ell\in\mathbb{Z}} z^\ell \langle \ell | e^{H(p_*)} | u\rangle.$$

is an isomorphism of vector spaces.

We remind that for $\ell \in \mathbb{Z}$, the vector $\langle \ell | \in \mathcal{F}^*$ is the charge ℓ Maya diagram (4.53) obtained by shifting the vacuum state ℓ steps to the right. The full statement of the Boson-Fermion correspondence [MJD00, Theorem 5.2] also realizes the action of the Clifford algebra $\mathcal{A}$ on $\mathcal{F}$ in terms of the action of the Heisenberg algebra $\mathcal{B}$ on $\mathbb{C}[p_*][z,z^{-1}]$, but in our case we only need the linear isomorphism Φ. From the expression for H_n in equation (4.56) one readily sees that $H_n\mathcal{F}^\ell \subset \mathcal{F}^\ell$ for every charge ℓ subspace $\mathcal{F}^\ell \subset \mathcal{F}$, so $\Phi|_{\mathcal{F}^\ell}$ gives an isomorphism $\mathcal{F}^\ell \xrightarrow{\cong} z^\ell\,\mathbb{C}[p_*]$.

In particular we have a linear isomorphism between the charge zero Fock spaces: $\mathcal{F}^0 \xrightarrow{\cong} C[p_*]$. We briefly present an alternative way of describing the former, which is more commonly used in modern literature [Bur22; Kaz08; KL07; Lan10; Wan21]. Charge zero Fermionic Fock space $\mathcal{F}^0$ is the *semi-infinite wedge product* of Laurent series $\bigwedge^{\infty/2}\mathbb{C}[z,z^{-1}]$. It is the space spanned by the following vectors indexed by generalized partitions (see Definition 4.1.1):

$$v_\mu := z^{1-\mu_1}\wedge z^{2-\mu_2}\wedge z^{3-\mu_3}\wedge\cdots, \quad \mu_1\geq\mu_2\geq\cdots\geq 0. \tag{4.60}$$

The vacuum state is the vector corresponding to the partition of zero: $|\text{vac}\rangle = v_{(0,0,\dots)} = z^1\wedge z^2\wedge z^3\wedge\cdots$. The correspondence between charge zero Maya diagrams and semi-infinite wedge products is

$$\begin{array}{ccc} |\lambda_1, \lambda_2, \lambda_3, \dots\rangle & \longmapsto & z^{\lambda_1+\frac{1}{2}} \wedge z^{\lambda_2+\frac{1}{2}} \wedge z^{\lambda_3+\frac{1}{2}} \wedge \cdots, \\ |\frac{1}{2}-\mu_1, \frac{3}{2}-\mu_2, \frac{5}{2}-\mu_3, \dots\rangle & \longleftarrow & z^{1-\mu_1} \wedge z^{2-\mu_2} \wedge z^{3-\mu_3} \wedge \cdots, \end{array} \tag{4.61}$$

in other words $\lambda_j + \frac{1}{2} = j - \mu_j$. The condition $\lambda_{j+1} = \lambda_j + 1$ for j large enough translates to $\mu_{j+1} = \mu_j = 0$, and the action of the Clifford algebra $\mathcal{A}$ is given by composing (4.61) with the previously defined actions. All the previous statements about $\mathcal{F}$ in terms of Maya diagrams can be readily translated into statements about semi-infinite wedge products. In particular, the linear isomorphism $\Phi : \mathcal{F}^0 \to \mathbb{C}[p_*]$ sends the basis elements (4.60) to the Schur functions introduced in Section 3.4:

$$v_\mu \longmapsto s_\mu(p_*).$$

Therefore the disconnected Hurwitz potential $H^\bullet(\beta; p_*) \in \mathbb{C}[[p_*]][[\beta]]$ is mapped to

$$\sum_\mu s_\mu(1, 0, 0, \dots) e^{f_2(\mu)\beta} v_\mu \in \overline{\mathcal{F}^0}[[\beta]],$$

where $\overline{\mathcal{F}^0}$ denotes the completion of $\mathcal{F}^0$.

Now that we have a correspondence between $\mathcal{F}^0$ and $\mathbb{C}[p_*]$, we explain the Fermionic analogues of the Hirota bilinear identity and the KP wave function. The essential question that we would like to answer is: to what subset of $\mathcal{F}^0$ is the set of tau-functions in $\mathbb{C}[p_*]$ mapped to under the Boson-Fermion isomorphism? Firstly, we generalize expressions such as $H_n = \sum_{i \in \mathbb{Z}+\frac{1}{2}} : a_{-i} a^\dagger_{i+n} :$ to include general charge-preserving operators involving quadratic expressions in Fermions:

$$X_A = \sum_{i,j \in \mathbb{Z}+\frac{1}{2}} c_{ij} : a_{-i} a^\dagger_j : ,$$

where $A = (c_{ij})$ denotes an infinite matrix. As before, we have expressed such operators with the normal product in order for their action on $\mathcal{F}^0$ to be well-defined. We impose a further condition on A, namely

$$\text{there exists } N > 0 \text{ such that } c_{ij} = 0 \text{ whenever } |i - j| > N. \tag{4.62}$$

The infinite-dimensional Lie algebra $\mathfrak{gl}(\infty)$ is defined to be the vector space generated by operators X_A satisfying (4.62). For example, each H_n belongs to $\mathfrak{gl}(\infty)$ since its matrix coefficients are $c_{ij} = \delta_{i+n,j}$. Condition (4.62) ensures that the commutator bracket of $\mathfrak{gl}(\infty)$ only involves finite sums, and hence is well-defined. The infinite-dimensional Lie group corresponding to $\mathfrak{gl}(\infty)$ is

$$GL(\infty) = \{e^{X_1}e^{X_2}\cdots e^{X_k} : k \geq 0, X_i \in \mathfrak{gl}(\infty)\}.$$

Since each $X_i \in \mathfrak{gl}(\infty)$ preserves charge, every element in $GL(\infty)$ does so too. Thus the representation of $\mathcal{A}$ in $\mathcal{F}$ restricts to a representation of $GL(\infty)$ in $\mathcal{F}^0$. Denote the orbit of the vacuum state $|\text{vac}\rangle$ in this representation by $GL(\infty)|\text{vac}\rangle \subset \mathcal{F}^0$. The following theorem answers the question: to what subset of $\mathcal{F}^0$ is the set of tau-functions in $\mathbb{C}[p_*]$ mapped to under the Boson-Fermion isomorphism?

Theorem 4.4.2. *The set of tau-functions in* $\mathbb{C}[p_*]$ *corresponds to* $GL(\infty)|vac\rangle$ *under the isomorphism* $\Phi : \mathcal{F}^0 \to \mathbb{C}[p_*]$.

To prove this, first consider any $|u\rangle \in \mathcal{F}^0$ and denote $\Phi(|u\rangle)$ by $f(p_*)$. Let $\varepsilon(z^{-1})$ be shorthand notation for $\left(\frac{1}{z}, \frac{1}{2z^2}, \frac{1}{3z^3}, \dots\right)$. Using Wick's theorem, one can show [MJD00, Lemma 5.3] that

$$\langle 1|a(z) = \langle \text{vac}|e^{-H(\varepsilon(z^{-1}))}. \tag{4.63}$$

Using this we compute the image of $a(z)|u\rangle$ in $\mathbb{C}[p_*]$:

$$\begin{aligned}
\Phi(a(z)|u\rangle) &= \langle 1|e^{H(p_*)}a(z)|u\rangle \overset{(4.58)}{=} \langle 1|a(z)e^{H(p_*)}|u\rangle e^{\xi(p_*,z)} \\
&\overset{(4.63)}{=} \langle \text{vac}|e^{-H(\varepsilon(z^{-1}))}e^{H(p_*)}|u\rangle e^{\xi(p_*,z)} \\
&= \langle \text{vac}|e^{H(p_*-\varepsilon(z^{-1}))}|u\rangle e^{\xi(p_*,z)} = f(p_* - \varepsilon(z^{-1}))e^{\xi(p_*,z)} \\
&= f\left(p_1 - \frac{1}{z}, p_2 - \frac{1}{2z^2}, p_3 - \frac{1}{3z^3}, \dots\right) e^{\xi(p_*,z)}, \\
&\overset{(4.34)}{=} G_z(f)e^{\xi(p_*,z)} = G_z(\Phi(|u\rangle))e^{\xi(p_*,z)}.
\end{aligned} \tag{4.64}$$

Using the similar result $\langle -1|a^\dagger(z) = \langle \text{vac}|e^{H(\varepsilon(z^{-1}))}$ and equation (4.59) one also shows that

$$\Phi(a^\dagger(z)|u\rangle) = G_z^\dagger(\Phi(|u\rangle))e^{-\xi(p_*,z)}. \tag{4.65}$$

For the set of independent variables p'_*, denote the isomorphism $\mathcal{F}^0 \to \mathbb{C}[p'_*]$ by Φ'. Then $\Phi(|u\rangle) \in \mathbb{C}[p_*]$ satisfies the Hirota bilinear identity

$$\text{res}_z \left(G_z(\Phi(|u\rangle)) \cdot G_z^\dagger(\Phi'(|u\rangle))e^{\xi(p_*,z)-\xi(p'_*,z)}\right) = 0$$

if and only if

$$
\begin{aligned}
&\operatorname{res}_z \Big(\Phi(a(z)|u\rangle) \cdot \Phi'(a^\dagger(z)|u\rangle)\Big) = 0 \\
\iff &\operatorname{res}_z \left(\sum_{i,j \in \mathbb{Z}+\frac{1}{2}} \Phi(a_j|u\rangle) \cdot \Phi'(a_i^\dagger|u\rangle) z^{-i-j-1} \right) = 0 \\
\iff &\sum_{i \in \mathbb{Z}+\frac{1}{2}} \Phi(a_{-i}|u\rangle) \cdot \Phi'(a_i^\dagger|u\rangle) = 0 \\
\iff &(\Phi \otimes \Phi') \left(\sum_{i \in \mathbb{Z}+\frac{1}{2}} a_{-i}|u\rangle \otimes a_i^\dagger|u\rangle \right) = 0 \\
\iff &\sum_{i \in \mathbb{Z}+\frac{1}{2}} a_{-i}|u\rangle \otimes a_i^\dagger|u\rangle = 0. \qquad (4.66)
\end{aligned}
$$

We call this last equation the *Fermionic bilinear identity*, since it is the equivalent of the Hirota bilinear identity in charge zero Fermionic Fock space $\mathcal{F}^0$. By Corollary 4.3.8, the polynomial $\Phi(|u\rangle) \in \mathbb{C}[p_*]$ is a tau-function of the KP hierarchy (up to an additive constant) if and only if $|u\rangle$ satisfies the Fermionic bilinear identity (4.66). Therefore the statement of Theorem 4.4.2 is a direct corollary of the next theorem, which is preceded by the following small lemma.

Lemma 4.4.3. *Let*

$$|u\rangle := a_{i_1} \cdots a_{i_r} a_{j_1}^\dagger \cdots a_{j_r}^\dagger |\mathrm{vac}\rangle \in \mathcal{F}^0, \quad i_1 < \cdots < i_r < 0,\ j_1 < \cdots < j_r < 0,$$

be a generic charge zero Maya diagram. Then one can 'replace' the white and black stones arising from the action of a_{i_1} and $a_{j_1}^\dagger$ as follows:

$$(1 - b_{i_1,j_1}^\dagger b_{i_1,j_1})|u\rangle = (-1)^{r-1} a_{i_2} \cdots a_{i_r} a_{j_2}^\dagger \cdots a_{j_r}^\dagger |\mathrm{vac}\rangle,$$

where $b_{i_1,j_1} := a_{i_1} + a_{-j_1}$ and $b_{i_1,j_1}^\dagger := a_{-i_1}^\dagger + a_{j_1}^\dagger$.

Proof. We perform the proof for $r = 1$, the general case following analogously. Let $|u\rangle = a_i a_j^\dagger |\mathrm{vac}\rangle$ for $i, j < 0$. Then using the anticommutation relations (4.51):

$$b_{ij}|u\rangle = a_i a_i a_j^\dagger |\mathrm{vac}\rangle + a_{-j} a_i a_j^\dagger |\mathrm{vac}\rangle = -a_i a_{-j} a_j^\dagger |\mathrm{vac}\rangle = -a_i |\mathrm{vac}\rangle + a_i a_j^\dagger \underbrace{a_{-j}|\mathrm{vac}\rangle}_{=0}.$$

Thus

$$b^\dagger_{ij} b_{ij}|u\rangle = -a^\dagger_{-i} a_i|\text{vac}\rangle - a^\dagger_j a_i|\text{vac}\rangle = -|\text{vac}\rangle + a_i \underbrace{a^\dagger_{-i}|\text{vac}\rangle}_{=0} + a_i a^\dagger_j|\text{vac}\rangle = -|\text{vac}\rangle + |u\rangle.$$

Rearranging yields $(1 - b^\dagger_{ij} b_{ij})|u\rangle = |\text{vac}\rangle$. □

As previously mentioned, this next theorem gives a direct proof of Theorem 4.4.2. A similar proof can be found in [KR87, Proposition 7.2].

Theorem 4.4.4. *Let $|u\rangle \in \mathcal{F}^0$ be nonzero. Then $|u\rangle$ lies in $GL(\infty)|\text{vac}\rangle$ if and only if it satisfies the Fermionic bilinear identity*

$$\sum_{i \in \mathbb{Z}+\frac{1}{2}} a_{-i}|u\rangle \otimes a^\dagger_i|u\rangle = 0.$$

Proof. Suppose $|u\rangle = g|\text{vac}\rangle$ for some $g \in GL(\infty)$. If $g = 1$, then $|u\rangle = |\text{vac}\rangle$ clearly satisfies the Fermionic bilinear identity because either $a_{-i}|\text{vac}\rangle = 0$ or $a^\dagger_i|\text{vac}\rangle = 0$ for every $i \in \mathbb{Z} + \frac{1}{2}$. If $g = e^{X_A}$ with $X_A = \sum_{i,j} c_{ij} : a_{-i} a^\dagger_j :$, consider first the readily computed commutation relations

$$[X_A, a_{-i}] = \sum_{j \in \mathbb{Z}+\frac{1}{2}} c_{ji}\, a_{-j} \quad \text{and} \quad [X_A, a^\dagger_i] = \sum_{j \in \mathbb{Z}+\frac{1}{2}} (-c_{ij})\, a^\dagger_j.$$

Hence the transformation matrices corresponding to $\{a_{-i}\}_i$ and $\{a^\dagger_i\}_i$ are the contragredient of one another, so

$$\sum_{i \in \mathbb{Z}+\frac{1}{2}} [X_A, a_{-i}]|\text{vac}\rangle \otimes a^\dagger_i|\text{vac}\rangle + \sum_{i \in \mathbb{Z}+\frac{1}{2}} a_{-i} \otimes [X_A, a^\dagger_i]|\text{vac}\rangle = 0.$$

We write this more suggestively as

$$\begin{aligned} &\sum_{i \in \mathbb{Z}+\frac{1}{2}} \Big(X_A a_{-i}|\text{vac}\rangle \otimes a^\dagger_i|\text{vac}\rangle + a_{-i}|\text{vac}\rangle \otimes X_A a^\dagger_i|\text{vac}\rangle \Big) \\ = &\sum_{i \in \mathbb{Z}+\frac{1}{2}} \Big(a_{-i} X_A|\text{vac}\rangle \otimes a^\dagger_i|\text{vac}\rangle + a_{-i}|\text{vac}\rangle \otimes a^\dagger_i X_A|\text{vac}\rangle \Big). \end{aligned}$$

Hence if we consider the two power series in t given by

$$f(t) = \sum_{i\in\mathbb{Z}+\frac{1}{2}} e^{tX_A} a_{-i}|\mathrm{vac}\rangle \otimes e^{tX_A} a_i^\dagger|\mathrm{vac}\rangle,$$
$$g(t) = \sum_{i\in\mathbb{Z}+\frac{1}{2}} a_{-i} e^{tX_A}|\mathrm{vac}\rangle \otimes a_i^\dagger e^{tX_A}|\mathrm{vac}\rangle,$$

the result above implies that $\frac{d^n}{dt^n} f(t)\Big|_{t=0} = \frac{d^n}{dt^n} g(t)\Big|_{t=0}$ for all $n \geq 0$. Therefore $f(1) = g(1)$, or in other words

$$\sum_{i\in\mathbb{Z}+\frac{1}{2}} e^{X_A} a_{-i}|\mathrm{vac}\rangle \otimes e^{X_A} a_i^\dagger|\mathrm{vac}\rangle = \sum_{i\in\mathbb{Z}+\frac{1}{2}} a_{-i} e^{X_A}|\mathrm{vac}\rangle \otimes a_i^\dagger e^{X_A}|\mathrm{vac}\rangle \tag{4.67}$$
$$= \sum_{i\in\mathbb{Z}+\frac{1}{2}} a_{-i}|u\rangle \otimes a_i^\dagger|u\rangle.$$

But the first expression in (4.67) vanishes for the same reason as before, so $|u\rangle$ satisfies the Fermionic bilinear identity. Applying this reasoning repeatedly, one can show that (4.67) holds for any $g = e^{X_1}e^{X_2}\cdots e^{X_k} \in GL(\infty)$ instead of just e^{X_A}, which shows that any $|u\rangle = g|\mathrm{vac}\rangle$ satisfies the Fermionic bilinear identity. Conversely, suppose $|u\rangle \in \mathcal{F}^0$ satisfies $\sum_{i\in\mathbb{Z}+\frac{1}{2}} a_{-i}|u\rangle \otimes a_i^\dagger|u\rangle = 0$. The vector $|u\rangle$ is a linear combination of charge zero Maya diagrams

$$a_{i_1}\cdots a_{i_r} a_{j_1}^\dagger \cdots a_{j_r}^\dagger|\mathrm{vac}\rangle, \quad i_1 < \cdots < i_r < 0,\ j_1 < \cdots < j_r < 0.$$

The fermions b_{ij} and $b_{ij}^\dagger$ $(i, j < 0)$ from Lemma 4.4.3 satisfy $\{b_{ij}, b_{ij}^\dagger\} = 2$ and $(b_{ij})^2 = (b_{ij}^\dagger)^2 = 0$, so $(b_{ij}^\dagger b_{ij})^k = 2^{k-1} b_{ij}^\dagger b_{ij}$ for $k \geq 1$. Hence

$$GL(\infty) \ni e^{i\pi b_{ij}^\dagger b_{ij}/2} = 1 - b_{ij}^\dagger b_{ij}.$$

Thus by Lemma 4.4.3 we can transform $|u\rangle$ by a suitable element $g \in GL(\infty)$ to reduce it to the form

$$g|u\rangle = |\mathrm{vac}\rangle + \sum_{i,j<0} d_{ij} a_i a_j^\dagger|\mathrm{vac}\rangle + \cdots,$$

where $d_{ij} \in \mathbb{C}$ and the remaining terms involve $|\mathrm{vac}\rangle$ being acted on by four or more fermions. Next, acting with $g' = \exp\left(-\sum_{i,j<0} d_{ij} a_i a_j^\dagger\right) \in GL(\infty)$ yields

$$|u'\rangle := g'g|u\rangle = |\mathrm{vac}\rangle + \sum_{i_1,i_2,j_1,j_2<0} d_{i_1,i_2,j_1,j_2} a_{i_1} a_{i_2} a_{j_1}^\dagger a_{j_2}^\dagger|\mathrm{vac}\rangle + \cdots.$$

As we have shown in the first part of this proof in equation (4.67), the Fermionic bilinear identity (4.66) is invariant under the action of $GL(\infty)$, so that our initial assumption about $|u\rangle$ implies

$$\sum_{i\in\mathbb{Z}+\frac{1}{2}} a_{-i}|u'\rangle \otimes a_i^\dagger|u'\rangle = 0. \tag{4.68}$$

Since the terms in $|u'\rangle$ after $|\mathrm{vac}\rangle$ involve multiplication by at least four fermions we have $a_{-i}|u'\rangle \neq 0$ for $-i < 0$ and $a_i^\dagger|u'\rangle \neq 0$ for $i < 0$. Then (4.68) implies $a_i^\dagger|u'\rangle = 0$ for $-i < 0$ and $a_{-i}|u'\rangle = 0$ for $i < 0$. Thus $|u'\rangle$ is annihilated by all the annihilation operators, and so it must be a scalar multiple of $|\mathrm{vac}\rangle$. We therefore conclude that $|u\rangle = (g'g)^{-1}|u'\rangle$ lies in $GL(\infty)|\mathrm{vac}\rangle$. □

We have completed the proof of Theorem 4.4.2, which claimed that the tau-functions of the KP hierarchy are in bijective correspondence with elements in the orbit of the vacuum state. We denote the tau-function corresponding to $g|\mathrm{vac}\rangle \in GL(\infty)|\mathrm{vac}\rangle$ by

$$\tau_g(p_*) := \Phi(g|\mathrm{vac}\rangle) = \langle\mathrm{vac}|e^{H(p_*)}g|\mathrm{vac}\rangle.$$

Moreover, thanks to equations (4.64) and (4.65), the wave function $\psi_g(p_*, z) = \frac{G_z(\tau_g)}{\tau_g}e^{\xi(p_*,z)}$ and adjoint wave function $\psi^\dagger(p_*, z) = \frac{G_z^\dagger(\tau_g)}{\tau_g}e^{-\xi(p_*,z)}$ of $\tau_g(p_*, z)$ are

$$\psi_g(p_*, z) = \frac{\Phi(a(z)g|\mathrm{vac}\rangle)}{\Phi(g|\mathrm{vac}\rangle)} = \frac{\langle 1|e^{H(p_*)}a(z)g|\mathrm{vac}\rangle}{\langle\mathrm{vac}|e^{H(p_*)}g|\mathrm{vac}\rangle},$$
$$\psi_g^\dagger(p_*, z) = \frac{\Phi(a^\dagger(z)g|\mathrm{vac}\rangle)}{\Phi(g|\mathrm{vac}\rangle)} = \frac{\langle -1|e^{H(p_*)}a^\dagger(z)g|\mathrm{vac}\rangle}{\langle\mathrm{vac}|e^{H(p_*)}g|\mathrm{vac}\rangle}.$$

Our next goal is to understand the inclusion of the set of tau-functions in $\mathbb{C}[p_*]$, which we now know to be equivalent to the inclusion of $GL(\infty)|\mathrm{vac}\rangle$ in $\mathcal{F}^0$, as an embedding of a certain Grassmannian into a projectivized wedge space. This is called the Plücker embedding, and the equations that define it (the Plücker relations) will give us a relatively straightforward criterion to determine whether a function is a tau-function of the KP hierarchy.

4.5 Plücker embedding

We start by describing the simplest nontrivial instance of the Plücker embedding. Let $G(2,4)$ denote the set of 2-dimensional subspaces of the $\mathbb{C}$-vector space $V = \mathbb{C}^4$. Fix a basis $\{v_1, v_2, v_3, v_4\}$ of $\mathbb{C}^4$. Every element W of $G(2,4)$ is determined by a choice of basis $\{w_1, w_2\}$, or equivalently by a choice of a 4×2 matrix $(w_1\, w_2)$ of full rank. This choice is not unique, since the set of all possible bases of W is given by the $GL(2)$ orbit of $\{w_1, w_2\}$:

$$\left\{ aw_1 + bw_2, cw_1 + dw_2 : \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in GL(2) \right\}.$$

Thus we identify $G(2,4)$ with $V(2,4)/GL(2)$, where $V(2,4)$ is the set of 4×2 matrices of full rank and the action of $GL(2)$ is by right matrix multiplication. For an element $(w_1\, w_2) \in G(2,4)$, consider the wedge product $w_1 \wedge w_2 \in \bigwedge^2 \mathbb{C}^4$. A different representative $(w_1\, w_2)g, g \in GL(2)$, of the same equivalence class will give rise to the wedge product $\det(g) w_1 \wedge w_2$, so the assignment $G(2,4) \to \bigwedge^2 \mathbb{C}^4$ is only well-defined up to a scalar constant. Hence projectivizing the target space gives rise to a well-defined map

$$G(2,4) \longrightarrow \mathbb{P}\left(\bigwedge^2 \mathbb{C}^4\right).$$

This is the *Plücker embedding.* It maps elements of $G(2,4)$ to decomposable forms in $\mathbb{P}\left(\bigwedge^2 \mathbb{C}^4\right)$, and conversely every decomposable form lies in its image. The space $G(2,4)$ has a structure of a complex manifold of dimension $\dim V(2,4)/GL(2) = \dim V(2,4) - \dim GL(2) = 4$. The basis $\{v_1, v_2, v_3, v_4\}$ of $\mathbb{C}^4$ induces a basis $\{v_1 \wedge v_2, v_1 \wedge v_3, v_1 \wedge v_4, v_2 \wedge v_3, v_2 \wedge v_4, v_3 \wedge v_4\}$ of $\bigwedge^2 \mathbb{C}^4$, consisting of $\binom{4}{2} = 6$ elements. Hence $\mathbb{P}\left(\bigwedge^2 \mathbb{C}^4\right)$ is a complex manifold of dimension $\binom{4}{2} - 1 = 5$. It is therefore natural to ask what the defining equation of the Plücker embedding is, or equivalently the equation that determines whether a sum of wedge products is decomposable. To answer this consider a 4×2 matrix with columns w_1 and w_2, with $w_k = \sum_{i=1}^4 x_{ik} v_i$:

$$(w_1\, w_2) = \begin{pmatrix} x_{11} & x_{12} \\ x_{21} & x_{22} \\ x_{31} & x_{32} \\ x_{41} & x_{42} \end{pmatrix}.$$

This matrix determines an element of $G(2,4)$ if and only if it is of full rank, or equivalently if at least one of the six 2×2 subdeterminants

$$y_{ij} := x_{i1}x_{j2} - x_{i2}x_{j1}, \quad i < j, \tag{4.69}$$

is nonzero. These subdeterminants appear as coefficients in the wedge product

$$w_1 \wedge w_2 = \left(\sum_{i=1}^4 x_{i1} v_i\right) \wedge \left(\sum_{j=1}^4 x_{j2} v_j\right) = \sum_{i<j} y_{ij} v_i \wedge v_j.$$

Thus the linear coordinates of the image of $(w_1\, w_2) \in G(2,4)$ under the Plücker embedding are $y_{ij} = x_{i1}x_{j2} - x_{i2}x_{j1}$, which we call the Plücker coordinates of an element of $G(2,4)$. They are quadratic expressions in $\{x_{ik}\}$, and they satisfy the quadratic relation

$$y_{12}y_{34} - y_{13}y_{24} + y_{14}y_{23} = 0. \tag{4.70}$$

Conversely if $\sum_{i<j} Y_{ij} v_i \wedge v_j$ is any element of $\mathbb{P}\left(\bigwedge^2 \mathbb{C}^4\right)$ whose linear coordinates Y_{ij} satisfy (4.70), then Y_{ij} must be of the form (4.69), so $\sum_{i<j} Y_{ij} v_i \wedge v_j$ lies in the image of the Plücker embedding. Hence the relation above uniquely describes the Plücker embedding, and is called the *Plücker relation.*

The story for general Grassmannians is similar. Let $n \geq 0$ and $k \leq n$, and let $G(k,n)$ be the Grassmannian manifold of k-dimensional subspaces W of $V = \mathbb{C}^n$. Every subspace is determined by an $n \times k$ matrix of full rank up to a right $GL(k)$-action, so $G(k,n) \cong V(k,n)/GL(k)$. Taking the wedge product of basis elements up to a scalar factor defines the Plücker embedding

$$G(k,n) \longrightarrow \mathbb{P}\left(\bigwedge^k \mathbb{C}^n\right). \tag{4.71}$$

The dimension of $G(k,n)$ is $\dim V(k,n) - \dim GL(k) = nk - k^2 = k(n-k)$, while the dimension of $\mathbb{P}\left(\bigwedge^k \mathbb{C}^n\right)$ is the generally much larger number $\binom{n}{k} - 1$. Given a basis $\{v_1, \ldots, v_n\}$ of $\mathbb{C}^n$, the linear coordinate in $\bigwedge^k \mathbb{C}^n$ corresponding to the basis element $v_{i_1} \wedge \cdots \wedge v_{i_k}$ $(1 \leq i_1 < \cdots < i_k \leq n)$ is denoted by $Y_{i_1,\ldots,i_k}$. The relations that determine the Plücker embedding are once again linked to the $k \times k$ subdeterminants $y_{i_1,\ldots,i_k}$ of a matrix in $V(k,n)$ [MJD00, Theorems 8.1-8.2].

Theorem 4.5.1. *Let* $(Y_{i_1,\ldots,i_k})$ *be linear coordinates of an element of* $\mathbb{P}\left(\bigwedge^k \mathbb{C}^n\right)$*. Then this element lies in the image of the Plücker embedding* $G(k,n) \to \mathbb{P}\left(\bigwedge^k \mathbb{C}^n\right)$ *if and only if, for every choice of distinct indices* $1 \leq i_1, \ldots, i_{n-1}, j_1, \ldots, j_{n+1} \leq n$*, we have*

$$\sum_{\ell=1}^{k+1} (-1)^{\ell-1} Y_{i_1,\ldots,i_{n-1},j_\ell} Y_{j_1,\ldots,j_{\ell-1},j_{\ell+1},\ldots,j_{n+1}} = 0. \tag{4.72}$$

The equations above are the Plücker relations*.*

Notice that these relations are once again quadratic, just like the Hirota and Fermionic bilinear identities (4.45), (4.66).

We now give a brief overview of how one may understand the inclusion $GL(\infty)|\mathrm{vac}\rangle \subset \mathcal{F}^0$ as an infinite-dimensional analogue of the Plücker embedding (4.71), following [MJD00, Chapter 9] and [JM83, §1-§2]. First, for arbitrary $n \geq 1$, denote by $\mathcal{A}_n$ the subalgebra of the Clifford algebra $\mathcal{A}$ generated by the fermions $a_i, a_i^\dagger$ with indices $|i| < n$. Let

$$V_n := \bigoplus_{|i|<n} \mathbb{C}a_i, \quad V_n^\dagger := \bigoplus_{|i|<n} \mathbb{C}a_i^\dagger, \quad W_n := V_n \oplus V_n^\dagger$$

be the vector spaces generated by the individual fermions. Notice that the dimension of V_n and $V_n^\dagger$ is $2n$. Define the finite-dimensional Fermionic Fock space $\mathcal{F}_n$ in the same way as in the previous section.

Let $\mathcal{A}_n^\times \subset \mathcal{A}_n$ denote the subset of invertible elements. This is nonzero, since for example $(a_{-\frac{1}{2}} + a_{\frac{1}{2}}^\dagger)^{-1} = a_{-\frac{1}{2}} + a_{\frac{1}{2}}^\dagger$. For $g \in \mathcal{A}_n^\times$, define $T_g : \mathcal{A}_n \to \mathcal{A}_n$ by $a \mapsto gag^{-1}$. The *Clifford group* is

$$G(W_n) = \{g \in \mathcal{A}_n^\times : T_g(W_n) \subset W_n\}.$$

The anticommutator bracket $\{\cdot,\cdot\}$ defines a nondegenerate symmetric bilinear form on W_n. Denote the orthogonal subgroup of $GL(W_n)$ with respect to $\{\cdot,\cdot\}$ by $O(W_n)$. A straightforward check confirms that $g \in G(W_n) \implies T_g \in O(W_n)$. The converse is also true.

Lemma 4.5.2. *Any $T \in O(W_n)$ is of the form T_g for some $g \in G(W_n)$.*

Proof. An element $g \in W_n$ is invertible if and only if $\{g, g\} \neq 0$, in which case $g^{-1} = \frac{2g}{\{g,g\}}$. For such a g we have

$$T_g(a) = gag^{-1} = (-ag + \{a, g\})g^{-1} = -a + 2\Big\{a, \frac{g}{\sqrt{\{g,g\}}}\Big\}\frac{g}{\sqrt{\{g,g\}}}$$

for all $a \in \mathcal{A}_n$, so $-T_g$ is a reflection through the plane perpendicular to $\frac{g}{\sqrt{\{g,g\}}}$. The lemma follows because $O(W_n)$ is generated by reflections. $\square$

Lemma 4.5.3. *Let $g, g' \in G(W_n)$. Then $T_g = T_{g'} \iff g = cg'$ for some $c \in \mathbb{C}^\times$.*

Proof. The $\Longleftarrow$ direction follows by definition of T_g. Since $T_g \circ T_{g'^{-1}} = T_{gg'^{-1}}$ we can take $g' = 1$ without loss of generality. If $T_g = \mathrm{id}$ then g belongs to the centre of $\mathcal{A}_n$. Since the algebra $\mathcal{A}_n$ is isomorphic to the algebra of $2^{2n} \times 2^{2n}$ matrices [Bou59][Section 9.4], its centre is $\mathbb{C}$. $\square$

Define a subgroup of the Clifford group $G(W_n)$ by

$$G_n = \{g \in \mathcal{A}^\times : T_g(V_n) \subset V_n,\ T_g(V_n^\dagger) \subset V_n^\dagger\}.$$

As $n \to \infty$ this corresponds to the infinite-dimensional Lie group $GL(\infty)$ defined in the previous section. The reason for this is that given a Fermionic operator of the form $X_A = \sum_{i,j \in \mathbb{Z}+\frac{1}{2}} c_{ij} : a_{-i} a_j^\dagger :$, the commutation relations

$$[X_A, a_{-i}] = \sum_{j \in \mathbb{Z}+\frac{1}{2}} c_{ji}\, a_{-j} \quad \text{and} \quad [X_A, a_i^\dagger] = \sum_{j \in \mathbb{Z}+\frac{1}{2}} (-c_{ij})\, a_j^\dagger$$

imply that $e^{X_A} V_n e^{-X_A} \subset V_n$ and $e^{X_A} V_n^\dagger e^{-X_A} \subset V_n^\dagger$ for any $n \geq 1$, where we use

$$e^{X_A} a e^{-X_A} = 1 + [X_A, a] + \frac{1}{2!}[X_A, [X_A, a]] + \frac{1}{3!}[X_A, [X_A, [X_A, a]]] + \cdots .$$

Proposition 4.5.4. *The map $G_n \to GL(V_n), g \mapsto T_g$ is a surjective group homomorphism with kernel $\mathbb{C}$.*

Proof. Let $T \in GL(V_n)$, written as $Ta_{-i} = \sum_{|j|<n} c_{ji} a_{-j}$. Define $T^\dagger \in GL(V_n^\dagger)$ by $T^\dagger a_i^\dagger = \sum_{|j|<n} c_{ij}^{-1} a_j^\dagger$. Then

$$\{Ta_{-i}, T^\dagger a_j^\dagger\} = \sum_{|k|,|\ell|<n} c_{ki} c_{j\ell}^{-1} \{a_{-k}, a_\ell^\dagger\} = \sum_{|k|<n} c_{ki} c_{jk}^{-1} = \delta_{ji} = \{a_{-i}, a_j^\dagger\}.$$

Thus $T \oplus T^\dagger \in GL(W_n)$ lies in $O(W_n)$. By Lemma 4.5.2 we have $T \oplus T^\dagger = T_g$ for some $g \in G(W_n)$. By construction g lies in the subgroup G_n, and moreover $T = T_g|_{V_n}$. The fact that the kernel is $\mathbb{C}$ is a direct consequence of Lemma 4.5.3. $\square$

Now for $|u\rangle \in \mathcal{F}_n$, let

$$V_n(|u\rangle) := \{a \in V_n : a|u\rangle = 0\}$$

be the subspace of V_n consisting of elements that annihilate $|u\rangle$. Then, for example, $V_n(|\text{vac}\rangle) = \bigoplus_{0<i<n} \mathbb{C}a_i$ and

$$V_n(a_{i_1} \cdots a_{i_r} a_{j_1}^\dagger \cdots a_{j_r}^\dagger |\text{vac}\rangle) = \bigoplus_{\substack{0<i<n \\ i \neq -i_1, \ldots, -i_r}} \mathbb{C}a_i \oplus \bigoplus_{\ell=1}^{r} \mathbb{C}a_{j_\ell}$$

are both n-dimensional subspaces of the $2n$-dimensional space V_n, in other words elements of the Grassmannian $G(n, 2n)$. For $g \in G_n$ and $a \in V_n$ we can write $ag = gT_g{}^{-1}(a)$, so $a \in V_n(g|\text{vac}\rangle) \iff gT_g{}^{-1}(a)|\text{vac}\rangle = 0 \iff T_g{}^{-1}(a) \in V_n(|\text{vac}\rangle)$. Hence $V_n(g|\text{vac}\rangle) = T_g V_n(|\text{vac}\rangle)$, which means that $V_n(g|\text{vac}\rangle)$ is an n-dimensional subspace of V_n for every $g \in G_n$.

Proposition 4.5.5. *Let $g, g' \in G_n$. Then*

$$V_n(g|\mathrm{vac}\rangle) = V_n(g'|\mathrm{vac}\rangle) \iff g = cg' \textit{ for some } c \in \mathbb{C}^\times.$$

Proof. The $\Longleftarrow$ direction is clear. Since $V_n(g|\mathrm{vac}\rangle) = T_g V_n(|\mathrm{vac}\rangle)$ we can assume that $g' = 1$. Then suppose $V_n(g|\mathrm{vac}\rangle) = V_n(|\mathrm{vac}\rangle)$. Firstly, this implies that $a_i g|\mathrm{vac}\rangle = 0$ for all $0 < i < n$. Next, define the subspace $V_n^\dagger(g|\mathrm{vac}\rangle) \subset V_n^\dagger$ in the same way as before. Since $g \in G_n$, by the reasoning in Proposition 4.5.4 the matrix (c_{ij}) corresponding to $T_g|_{V_n}$ is the inverse transpose of the one for $T_g|_{V_n^\dagger}$. Hence we also have $T_g V_n^\dagger(|\mathrm{vac}\rangle) = V_n^\dagger(|\mathrm{vac}\rangle)$, and since the former subspace is equal to $V_n^\dagger(g|\mathrm{vac}\rangle)$, this implies that $a_i^\dagger g|\mathrm{vac}\rangle = 0$ for all $0 < i < n$. We have concluded that $g|\mathrm{vac}\rangle$ is annihilated by all the annihilation operators $\{a_i, a_i^\dagger\}_{0<i<n}$, so it must be a scalar multiple of $|\mathrm{vac}\rangle$. $\square$

The consequence of Propositions 4.5.4 and Propositions 4.5.5 is the following.

Corollary 4.5.6. *There is a bijective correspondence*

$$\begin{aligned} G_n|\mathrm{vac}\rangle/\mathbb{C}^\times &\longrightarrow G(n, 2n), \\ |u\rangle &\longmapsto V_n(|u\rangle). \end{aligned}$$

Proof. The injectivity is due to Proposition 4.5.5. Let W be an n-dimensional subspace of V_n, in other words an element of $G(n, 2n)$. Let $T \in GL(V_n)$ be such that $T(W) = V_n(|\mathrm{vac}\rangle)$. Then by Proposition 4.5.4 $T = T_{g^{-1}}$ for some $g \in G_n$, so $W = T_g V_n(|\mathrm{vac}\rangle) = V_n(g|\mathrm{vac}\rangle)$. $\square$

Taking $n \to \infty$ gives a correspondence between elements of the vacuum orbit $GL(\infty)|\mathrm{vac}\rangle \subset \mathcal{F}^0$ up to a scalar multiple, and "half-infinite-dimensional subspaces". The Grassmannian $G(\infty/2, \infty)$ of half-infinite-dimensional subspaces is called the *Sato Grassmannian.* The ambient space of $G(\infty/2, \infty)$ from the Plücker embedding is the semi-infinite wedge product of Laurent series $\bigwedge^{\infty/2} \mathbb{C}[z, z^{-1}]$, which is the equivalent characterization of charge zero Fermionic Fock space $\mathcal{F}^0$ given in Section 4.4. The image of the Sato Grassmannian in $\bigwedge^{\infty/2} \mathbb{C}[z, z^{-1}]$ consists of decomposable elements

$$\varphi_1(z) \wedge \varphi_2(z) \wedge \cdots$$

where $\varphi_i(z) = z^i +$ (lower order terms) for all $i > n$ for some $n \geq 1$.

Now that we have obtained our goal of understanding tau-functions of the KP hierarchy as elements of the Sato Grassmannian, we summarize the main points of Sections 4.2-4.5 in the following diagram.

$$\begin{tikzcd}
\left\{\begin{matrix}\text{Solutions } L = \partial_x + \sum_{i\geq 1} u_i(p_*)\partial_x^{-i} \\ \in \mathcal{PO}_p \text{ of the KP hierarchy}\end{matrix}\right\} \arrow[d, leftrightarrow, "\substack{\text{Example 4.2.3,} \\ \text{Proposition 4.3.6}}"'] & \\
\left\{\begin{matrix}\text{Tau-functions } \tau \in \mathbb{C}[p_*] \\ \text{of the KP hierarchy}\end{matrix}\right\} \arrow[r, leftrightarrow, "\text{Corollary 4.3.8}"] \arrow[d, leftrightarrow, "\substack{\text{Theorem 4.4.1,} \\ \text{Theorem 4.4.2}}"'] & \left\{\begin{matrix}\text{Polynomials } \tau \in \mathbb{C}[p_*] \text{ satisfying the} \\ \text{Hirota bilinear identity (4.45)}\end{matrix}\right\} \arrow[d, leftrightarrow, "\text{Theorem 4.4.1}"] \\
GL(\infty)|\mathrm{vac}\rangle \subset \mathcal{F}^0 \arrow[r, leftrightarrow, "\text{Theorem 4.4.4}"] \arrow[d, leftrightarrow, "\text{Corollary 4.5.6}"'] & \left\{\begin{matrix}\text{Vectors } |u\rangle \in \mathcal{F}^0 \text{ satisfying the} \\ \text{Fermionic bilinear identity (4.66)}\end{matrix}\right\} \\
G(\infty/2,\infty) \subset \bigwedge^{\infty/2}\mathbb{C}[z,z^{-1}] \arrow[r, leftrightarrow, "\text{Theorem 4.5.1}"] & \left\{\begin{matrix}\text{Decomposable elements of } \bigwedge^{\infty/2}\mathbb{C}[z,z^{-1}] \\ \text{satisfying the Plücker relations (4.72)}\end{matrix}\right\}
\end{tikzcd}$$

4.6 Hurwitz potential revisited

From the results of Section 3, the connected and disconnected Hurwitz potentials $H(\beta; p_*)$ and $H^\bullet(\beta; p_*)$ are related by $H^\bullet = e^H - 1$, and the disconnected potential is

$$H^\bullet(\beta; p_*) = \sum_\mu s_\mu(1,0,0,\dots)s_\mu(p_*)e^{f_2(\mu)\beta},$$

where the sum runs over all partitions μ and $s_\mu(p_*)$ are the Schur functions, and

$$f_2(\mu) = \frac{1}{2}\sum_{i=1}^{\ell(\mu)} \mu_i(\mu_i - 2i + 1).$$

The image of $H^\bullet(\beta; p_*)$ under the Boson-Fermion isomorphism is the following element in the formal completion of $\bigwedge^{\infty/2}\mathbb{C}[z, z^{-1}]$:

$$\sum_\mu s_\mu(1,0,0,\dots)v_\mu\, e^{f_2(\mu)\beta} \tag{4.73}$$

where $v_\mu = z^{1-\mu_1} \wedge z^{2-\mu_2} \wedge v^{3-\mu_3} \wedge \cdots$. By the results of Section 4.5, to show that $H^\bullet(\beta; p_*)$ is a tau-function of the KP hierarchy we show that (4.73) is of the form $\varphi_1(z) \wedge \varphi_2(z) \wedge \varphi_3(z) \wedge \cdots$, where $\varphi_i(z) = z^i + (\text{lower order terms})$ for i large enough. Following [KL07], we do this for $\beta = 0$ first.

For $i \geq 1$, let

$$\varphi_i(z) = z^i e^{z^{-1}} := \sum_{k \geq 0} \frac{1}{k!} z^{i-k} = \sum_{j \leq i} \frac{1}{(i-j)!} z^j.$$

Then

$$\begin{aligned}
\varphi_1(z) \wedge \varphi_2(z) \wedge \cdots &= \sum_{\substack{k_1, k_2, \cdots \geq 0 \\ i - k_i \neq i' - k_{i'}}} \left(\prod_{i \geq 1} \frac{1}{k_i!} \right) z^{1-k_1} \wedge z^{2-k_2} \wedge \cdots \\
&= \sum_{\substack{j_1, j_2, \cdots \in \mathbb{Z} \\ j_i \neq j_{i'}, j_i \leq i}} \left(\prod_{i \geq 1} \frac{1}{(i-j_i)!} \right) z^{j_1} \wedge z^{j_2} \wedge \cdots \qquad (j_i := i - k_i).
\end{aligned}$$

Now notice that for a fixed partition μ, the wedge product $z^{j_1} \wedge z^{j_2} \wedge \cdots$ is equal to $\pm v_\mu = \pm z^{1-\mu_1} \wedge z^{2-\mu_2} \wedge \cdots$ if and only if $(j_1, \ldots, j_{\ell(\mu)}) = (1 - \mu_1, \ldots, \ell(\mu) - \mu_{\ell(\mu)})$ up to reordering of the left-hand sequence, and $j_i = 0$ for $i > \ell(\mu)$. Let $\mathfrak{J}_\mu$ be the set of sequences $\mathrm{j} = (j_1, j_2, \ldots)$ satisfying these two conditions. Moreover, for $\mathrm{j} \in \mathfrak{J}_\mu$ let $\sigma_{\mathrm{j}} \in S_{\ell(\mu)}$ be the permutation such that $j_i = \sigma_{\mathrm{j}}(i) - \mu_{\sigma_{\mathrm{j}}(i)}$ for every $1 \leq i \leq \ell(\mu)$. Then the coefficient of v_μ in $\varphi_1(z) \wedge \varphi_2(z) \wedge \cdots$ is

$$\begin{aligned}
&\sum_{\mathrm{j} \in \mathfrak{J}_\mu} \operatorname{sgn}(\sigma_{\mathrm{j}}) \prod_{i=1}^{\ell(\mu)} \frac{1}{(i - j_i)!} = \sum_{\mathrm{j} \in \mathfrak{J}_\mu} \operatorname{sgn}(\sigma_{\mathrm{j}}) \prod_{i=1}^{\ell(\mu)} \frac{1}{(\mu_{\sigma_{\mathrm{j}}(i)} - \sigma_{\mathrm{j}}(i) + i)!} \\
&= \sum_{\sigma \in S_{\ell(\mu)}} \operatorname{sgn}(\sigma) \prod_{i=1}^{\ell(\mu)} \frac{1}{(\mu_{\sigma(i)} - \sigma(i) + i)!} = \det \left(\frac{1}{(\mu_j - j + i)!} \right) \\
&\overset{(3.23)}{=} \det \Big(s_{\mu_j - j + i}(1, 0, 0, \ldots) \Big) \overset{(3.24)}{=} s_\mu(1, 0, 0, \ldots).
\end{aligned}$$

This is precisely the coefficient of v_μ in the Fermionic version of $H^\bullet(0; p_*)$ in equation (4.73), so we are done for $\beta = 0$.

For general β, let us first consider the action of the diagonal endomorphisms of $\mathbb{C}[z, z^{-1}]$ on $\bigwedge^{\infty/2} \mathbb{C}[z, z^{-1}]$. A diagonal endomorphism $\operatorname{diag}(\ldots, a_{-1}, a_0, a_1, \ldots)$ of $\mathbb{C}[z, z^{-1}]$ with respect to the basis $\{z^i : i \in \mathbb{Z}\}$ sends $z^i \mapsto a_i z^i$. The obvious induced action on $\bigwedge^{\infty/2} \mathbb{C}[z, z^{-1}]$ would be

$$v_\mu \longmapsto \left(\prod_{i \geq 1} a_{i - \mu_i} \right) v_\mu,$$

but this is ill-defined as it gives rise to an infinite product. To get around this problem, we can exploit the fact that $\frac{a_{i-\mu_i}}{a_i} = 1$ for $i > \ell(\mu)$ to normalize the above expression so that $|\mathrm{vac}\rangle = v_{(0,0,\dots)}$ is mapped to $|\mathrm{vac}\rangle$ [Lan10]:

$$v_\mu \longmapsto \left(\prod_{i\geq 1} \frac{a_{i-\mu_i}}{a_i}\right) v_\mu = \left(\prod_{i=1}^{\ell(\mu)} \frac{a_{i-\mu_i}}{a_i}\right) v_\mu. \tag{4.74}$$

Now consider the diagonal matrix $\mathrm{diag}\big(e^{i(i-1)\beta/2}\big)_{i\in\mathbb{Z}}$. The product arising in (4.74) is

$$\prod_{i=1}^{\ell(\mu)} \frac{e^{(i-\mu_i)(i-\mu_i-1)\beta}/2}{e^{i(i-1)\beta/2}} = \prod_{i=1}^{\ell(\mu)} e^{\mu_i(\mu_i-2i+1)\beta/2} = e^{\sum_{i=1}^{\ell(\mu)} \mu_i(\mu_i-2i+1)\beta/2} = e^{f_2(\mu)\beta},$$

so the action of $\mathrm{diag}\big(e^{i(i-1)\beta/2}\big)_{i\in\mathbb{Z}}$ maps $\varphi_1(z)\wedge\varphi_2(z)\wedge\cdots = \sum_\mu s_\mu(1,0,0,\dots)v_\mu$ to $\sum_\mu s_\mu(1,0,0,\dots)v_\mu\, e^{f_2(\mu)\beta}$, which is the Fermionic counterpart of $H^\bullet(\beta;p_*)$. This is once again equal to a decomposable element $\varphi_1^\beta(z)\wedge\varphi_2^\beta(z)\wedge\cdots$ of $\bigwedge^{\infty/2}\mathbb{C}[z,z^{-1}]$, where

$$\varphi_i^\beta(z) = \sum_{j\leq i} \frac{1}{(i-j)!} e^{(j(j-1)-i(i-1))\beta/2}\, z^j.$$

We conclude:

Theorem 4.6.1. *The disconnected Hurwitz potential $H^\bullet(\beta;p_*)$ is a tau-function of the KP hierarchy.*

5 The Witten-Kontsevich theorem

5.1 Completing Kazarian and Lando's proof

In this section, we re-scale our independent variables $t_* = (t_0, t_1, t_2 \dots)$ and $p_* = (p_1, p_2, p_3, \dots)$ as follows:

$$t_i \longmapsto (2i+1)t_i, \quad p_j \longmapsto jp_j.$$

This is in order to be consistent with [Wit90; KL07]. We now finish outlining the proof of the Witten-Kontsevich theorem (Theorem 2.6.3). What we have so far is a generating series for the intersection numbers $\langle \tau_{d_1} \cdots \tau_{d_n} \rangle = \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n}$ of the form

$$F(t_*) = \sum_{(d_1,\dots,d_n)} \langle \tau_{d_1} \cdots \tau_{d_n} \rangle \frac{t_{d_1} \cdots t_{d_n}}{|\mathrm{Aut}(d_1, \dots, d_n)|},$$

which satisfies the string and dilaton equations

$$\frac{\partial F}{\partial t_0} = \frac{t_0^2}{2} + \sum_{i=0}^{\infty} t_{i+1} \frac{\partial F}{\partial t_i}, \qquad \frac{\partial F}{\partial t_1} = \frac{1}{3} \sum_{i=0}^{\infty} (2i+1) t_i \frac{\partial F}{\partial t_i} + \frac{1}{24}. \tag{5.1}$$

Witten's conjecture states that F is a tau-function of the KdV hierarchy, or equivalently that $U = \frac{\partial^2 F}{\partial t_0^2}$ is a solution of this hierarchy.

Next, we have the connected and disconnected simple Hurwitz potentials

$$H(\beta; p_*) = \sum_{g \geq 0, \mu \in \mathcal{P}} H_g^b(\mu) p_\mu \frac{\beta^b}{b!},$$
$$H^\bullet(\beta; p_*) = \sum_{g \geq 0, \mu \in \mathcal{P}} H_g^{\bullet, b}(\mu) p_\mu \frac{\beta^b}{b!},$$

where $p_\mu = p_{\mu_1} \cdots p_{\mu_n}$. They are related by the exponentiation relation $H^\bullet = e^H - 1$ from Proposition 3.3.1. We proved that $H^\bullet$ is a tau-function of the KP hierarchy in Theorem 4.6.1, and so $H = \log H^\bullet$ satisfies the first KP equation (see after Proposition 4.3.6)

$$\frac{\partial^2 H}{\partial p_2^2} = \frac{\partial^2 H}{\partial p_1 \partial p_3} - \frac{1}{2} \left(\frac{\partial^2 H}{\partial p_1^2} \right)^2 - \frac{1}{12} \frac{\partial^4 H}{\partial p_1^4}. \tag{5.2}$$

Moreover, the ELSV formula

$$H_g^b(\mu) = b! \prod_{i=1}^{n} \frac{\mu_i^{\mu_i}}{\mu_i!} \int_{\overline{\mathcal{M}}_{g,n}} \frac{1 - \lambda_1 + \lambda_2 - \cdots + (-1)^g \lambda_g}{(1 - \mu_1 \psi_1) \cdots (1 - \mu_n \psi_n)} \tag{5.3}$$

gave us a way of expressing intersections of psi-classes in terms of simple Hurwitz numbers in Proposition 3.5.3:

$$\langle \tau_{d_1} \cdots \tau_{d_n} \rangle = \sum_{\mu_1=1}^{d_1+1} \cdots \sum_{\mu_n=1}^{d_n+1} \left(\frac{1}{b!} \prod_{i=1}^{n} \frac{(-1)^{d_i-\mu_i+1}}{(d_i-\mu_i+1)!\mu_i^{\mu_i-1}} \right) H_g^b(\mu_1, \ldots, \mu_n), \tag{5.4}$$

where $b = 2g-2+n+|\mu|$ and $g = (d_1+\cdots+d_n-n+3)/3$ in each term of the sum. Since the ELSV formula involves integrating over the moduli space of curves, it is only valid for $2g-2+n>0$. With this mind, we separate the generating series $H(\beta; p_*)$ into unstable and stable parts by writing $H = H_{0;1} + H_{0;2} + H_{\mathrm{st}}$. The first two parts correspond to the terms with $(g,n) = (0,1)$ and $(g,n) = (0,2)$ in the sum, where the ELSV formula is not valid. The first one is

$$H_{0;1}(\beta; p_*) = \sum_{\mu=1}^{\infty} H_0^{\mu-1}(\mu)\, p_\mu \frac{\beta^{\mu-1}}{(\mu-1)!} = \sum_{\mu=1}^{\infty} \frac{\mu^{\mu-2}}{\mu!} p_\mu \beta^{\mu-1},$$

where we used $H_0^{\mu-1}(\mu) = \frac{\mu^{\mu-2}}{\mu}$, which was first sketched in [Hur91] and later proved in [Dén59]. The simple Hurwitz numbers in $H_{0;2}$ are computed in [Arn96, §4]:

$$H_0^{\mu_1+\mu_2}(\mu_1,\mu_2) = \begin{cases} \frac{\mu_1^{\mu_1}\mu_2^{\mu_2}(\mu_1+\mu_2-1)!}{\mu_1!\mu_2!}, & \mu_1 \neq \mu_2 \text{ or } \mu_1 = \mu_2 = 1, \\ \frac{1}{2}\frac{\mu_1^{\mu_1}\mu_2^{\mu_2}(\mu_1+\mu_2-1)!}{\mu_1!\mu_2!}, & \mu_1 = \mu_2 > 1, \end{cases}$$

Therefore

$$\begin{aligned} H_{0;2}(\beta; p_*) &= \sum_{\mu_1,\mu_2=1}^{\infty} H_0^{\mu_1+\mu_2}(\mu_1,\mu_2)\, p_{\mu_1} p_{\mu_2} \frac{\beta^{\mu_1+\mu_2}}{(\mu_1+\mu_2)!} \\ &= \frac{1}{2} p_1^2 \beta^2 + \frac{1}{2} \sum_{\substack{\mu_1,\mu_2=1 \\ (\mu_1,\mu_2)\neq(1,1)}}^{\infty} \frac{\mu_1^{\mu_1}\mu_2^{\mu_2}}{\mu_1!\mu_2!(\mu_1+\mu_2)} p_{\mu_1} p_{\mu_2} \beta^{\mu_1+\mu_2}. \end{aligned}$$

Therefore the KP equation (5.2) for H implies that $H_{\mathrm{st}} = H - H_{0;1} - H_{0;2}$ satisfies the following PDE:

$$\frac{\partial^2 H_{\mathrm{st}}}{\partial p_2^2} = \frac{\partial^2 H_{\mathrm{st}}}{\partial p_1 \partial p_3} - \frac{1}{2}\left(\frac{\partial^2 H_{\mathrm{st}}}{\partial p_1^2}\right)^2 - \frac{1}{12}\frac{\partial^4 H_{\mathrm{st}}}{\partial p_1^4} - \frac{1}{2}\beta^2 \frac{\partial^2 H_{\mathrm{st}}}{\partial p_1^2}. \tag{5.5}$$

Inspired by equation (5.4), introduce the change of independent variables given by

$$p_\mu = \sum_{d=\mu-1}^{\infty} \frac{(-1)^{d-\mu+1}}{(d-\mu+1)!\mu^{\mu-1}} \beta^{-\mu-\frac{2d+1}{3}}\, t_d, \tag{5.6}$$

and denote by $G_{\mathrm{st}}(\beta; t_*)$ the result of this coordinate change in $H_{\mathrm{st}}(\beta; p_*)$. This is a power series in $t_* = (t_0, t_1, \dots)$ whose coefficients are formal Laurent series in $\beta^{2/3}$. To see this, notice that the power of β in the monomial $p_\mu \beta^b = p_{\mu_1} \cdots p_{\mu_n} \beta^b$ when expressed in the t_* variables is

$$\begin{aligned} b + \sum_{i=1}^{n} \left(-\mu_i - \frac{2d_i + 1}{3} \right) &= 2g - 2 + n - \frac{1}{3} n - \frac{2}{3} \sum_{i=1}^{n} d_i \\ &= \frac{2}{3} \left(3g - 3 + n - \sum_{i=1}^{n} d_i \right), \end{aligned} \tag{5.7}$$

where we used $b = 2g - 2 + n + \sum_{i=1}^{n} \mu_i$. We can actually say more, namely that $G_{\mathrm{st}}(\beta; t_*)$ only contains nonnegative powers of $\beta^{2/3}$. To prove this, write the coefficient of $p_\mu \beta^b$ using the ELSV formula (5.3):

$$\begin{aligned} \frac{1}{b!} H_g^b(\mu) &= \prod_{i=1}^{n} \frac{\mu_i^{\mu_i}}{\mu_i!} \int_{\overline{\mathcal{M}}_{g,n}} \frac{1 - \lambda_1 + \lambda_2 - \cdots + (-1)^g \lambda_g}{(1 - \mu_1 \psi_1) \cdots (1 - \mu_n \psi_n)} \\ &= \prod_{i=1}^{n} \frac{\mu_i^{\mu_i}}{\mu_i!} \sum_{j=0}^{g} \sum_{d_1=1}^{\infty} \cdots \sum_{d_n=1}^{\infty} (-1)^j \int_{\overline{\mathcal{M}}_{g,n}} \lambda_j (\mu_1 \psi_1)^{d_1} \cdots (\mu_n \psi_n)^{d_n}. \end{aligned} \tag{5.8}$$

Integrating over $\overline{\mathcal{M}}_{g,n}$ only picks out monomials satisfying $j + \sum_{i=1}^{n} d_i = \dim \overline{\mathcal{M}}_{g,n} = 3g - 3 + n$, so the power of β given in (5.7) is $2j/3 \geq 0$.

Proposition 5.1.1. *The free term in β in $G_{st}(\beta; t_*)$ coincides with the generating series for intersections of ψ-classes:*

$$G_{st}(0; t_*) = F(t_*).$$

Proof. Since $G_{\mathrm{st}}(\beta; t_*)$ only contains nonnegative powers of $\beta^{2/3}$, the expression $G_{\mathrm{st}}(0; t_*)$ is well-defined. Collect the terms in $H_{\mathrm{st}}(\beta; p_*)$ corresponding to a given partition as follows:

$$H_{\mathrm{st}}(\beta; p_*) = \sum_{\mu} H_\mu(\beta)\, p_\mu, \quad \text{where} \quad H_\mu(\beta) := \sum_{g \geq 0} H_g^b(\mu) \frac{\beta^b}{b!}.$$

Writing $b = 2g - 2 + n + |\mu|$ as $b = \left(|\mu| + \frac{1}{3} n \right) + \frac{2}{3}(3g - 3 + n)$ and using equation (5.8) we obtain

$$\begin{aligned} H_\mu(\beta) &= C \sum_{g \geq 0} \sum_{j=0}^{g} \sum_{d_1=1}^{\infty} \cdots \sum_{d_n=1}^{\infty} (-1)^j \int_{\overline{\mathcal{M}}_{g,n}} \lambda_j (\mu_1 \psi_1)^{d_1} \cdots (\mu_n \psi_n)^{d_n} \, \beta^{\frac{2}{3}(j + \sum_{i=1}^{n} d_i)} \\ &= C \sum_{g \geq 0} \int_{\overline{\mathcal{M}}_{g,n}} \frac{1 - \beta^{2/3} \lambda_1 + \beta^{4/3} \lambda_2 - \cdots + (-1)^g \beta^{2g/3} \lambda_g}{(1 - \mu_1 \beta^{2/3} \psi_1) \cdots (1 - \mu_n \beta^{2/3} \psi_n)}, \end{aligned} \tag{5.9}$$

where $C = \beta^{|\mu|+\frac{1}{3}n} \prod_{i=1}^n \frac{\mu_i^{\mu_i}}{\mu_i!}$. Now collect the terms in $G_{\mathrm{st}}(\beta; t_*)$ in a similar way:

$$G_{\mathrm{st}}(\beta; t_*) = \sum_{d=(d_1,\ldots,d_n)} G_d(\beta) \frac{t_{d_1} \cdots t_{d_n}}{|\mathrm{Aut}(d_1, \ldots, d_n)|}.$$

To find $G_d(\beta)$ for any $d = (d_1, \ldots, d_n)$, we first apply the coordinate transformation (5.6) to $H_{\mathrm{st}}(\beta; p_*)$, using the notation $c_\mu^d = \frac{(-1)^{d-\mu+1}}{(d-\mu+1)!(\mu-1)!}$ from Lemma 3.5.2:

$$\begin{aligned}
G_{\mathrm{st}}(\beta; t_*) = H_{\mathrm{st}}(\beta; p_*(t_*)) &= \sum_{n\geq 0} \sum_{\mu_1=1}^{\infty} \cdots \sum_{\mu_n=1}^{\infty} H_\mu(\beta)\, p_{\mu_1} \cdots p_{\mu_n} \\
&= \sum_{n\geq 0} \sum_{\mu_1=1}^{\infty} \cdots \sum_{\mu_n=1}^{\infty} \sum_{d_1=\mu_1-1}^{\infty} \cdots \sum_{d_n=\mu_n-1}^{\infty} \left(\prod_{i=1}^n c_{\mu_i}^{d_i} \frac{\mu_i!}{\mu_i^{\mu_i}} \beta^{-\mu_i - \frac{2d_i+1}{3}} \right) H_\mu(\beta)\, t_{d_1} \cdots t_{d_n} \\
&= \sum_{n\geq 0} \sum_{d_1=1}^{\infty} \cdots \sum_{d_n=1}^{\infty} \underbrace{\sum_{\mu_1=1}^{d_1+1} \cdots \sum_{\mu_n=1}^{d_n+1} C^{-1} \left(\prod_{i=1}^n c_{\mu_i}^{d_i} \right) \beta^{-\frac{2}{3}\sum_i d_i} H_\mu(\beta)}_{=|\mathrm{Aut}(d_1,\ldots,d_n)| G_d(\beta)}\, t_{d_1} \cdots t_{d_n}
\end{aligned}$$

Combining this with (5.9) we obtain

$$\begin{aligned}
G_d(\beta) &= \beta^{-\frac{2}{3}\sum_i d_i} \sum_{g'\geq 0} \sum_{\mu_1=1}^{d_1+1} \cdots \sum_{\mu_n=1}^{d_n+1} \left(\prod_{i=1}^n c_{\mu_i}^{d_i} \right) \times \\
&\qquad \int_{\overline{\mathcal{M}}_{g',n}} \frac{1 - \beta^{2/3}\lambda_1 + \beta^{4/3}\lambda_2 - \cdots + (-1)^{g'} \beta^{2g'/3} \lambda_{g'}}{(1 - \mu_1 \beta^{2/3} \psi_1) \cdots (1 - \mu_n \beta^{2/3} \psi_n)} \\
&= \beta^{-\frac{2}{3}\sum_i d_i} \sum_{g'\geq 0} \int_{\overline{\mathcal{M}}_{g',n}} \left(\sum_{\mu_1=1}^{d_1+1} \frac{c_{\mu_1}^{d_1}}{1 - \mu_1 \beta^{2/3} \psi_1} \right) \cdots \left(\sum_{\mu_n=1}^{d_n+1} \frac{c_{\mu_n}^{d_n}}{1 - \mu_n \beta^{2/3} \psi_n} \right) \times \\
&\qquad \left(1 - \beta^{2/3}\lambda_1 + \beta^{4/3}\lambda_2 - \cdots + (-1)^{g'} \beta^{2g'/3} \lambda_{g'} \right) \\
&= \beta^{-\frac{2}{3}\sum_i d_i} \sum_{g'\geq 0} \int_{\overline{\mathcal{M}}_{g',n}} \left((\beta^{2/3}\psi_1)^{d_1} + \mathcal{O}(d_1+1) \right) \cdots \left((\beta^{2/3}\psi_n)^{d_n} + \mathcal{O}(d_n+1) \right) \times \\
&\qquad \left(1 - \beta^{2/3}\lambda_1 + \beta^{4/3}\lambda_2 - \cdots + (-1)^{g'} \beta^{2g'/3} \lambda_{g'} \right) \\
&= \int_{\overline{\mathcal{M}}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n} + \mathcal{O}(\beta^{2/3}),
\end{aligned}$$

where we used Lemma 3.5.2 in the third equality. We conclude that $G_d(0) = \langle \tau_{d_1} \cdots \tau_{d_n} \rangle$, and hence $G_{\mathrm{st}}(0; t_*) = F(t_*)$. □

The change of independent variables (5.6) induces in the following change of partial derivatives:

$$\begin{aligned}
\frac{\partial}{\partial p_1} &= \beta^{4/3}\frac{\partial}{\partial t_0},\\
\frac{\partial}{\partial p_2} &= 2\beta^{9/3}\frac{\partial}{\partial t_1} + 2\beta^{7/3}\frac{\partial}{\partial t_0},\\
\frac{\partial}{\partial p_3} &= 9\beta^{14/3}\frac{\partial}{\partial t_2} + 9\beta^{12/3}\frac{\partial}{\partial t_1} + \frac{9}{2}\beta^{10/3}\frac{\partial}{\partial t_0}.
\end{aligned}$$

Substituting this in first KP equation (5.5) for $H_{\mathrm{st}}(\beta; p_*)$, we see that the term containing $\beta^{14/3}$ cancels and so we can divide by $\beta^{16/3}$ to obtain

$$\frac{\partial^2 G_{\mathrm{st}}}{\partial t_0 \partial t_1} - \frac{1}{2}\left(\frac{\partial^2 G_{\mathrm{st}}}{\partial t_0^2}\right)^2 - \frac{1}{12}\frac{\partial^4 G_{\mathrm{st}}}{\partial t_0^4} + \beta^{2/3}\left(9\frac{\partial^2 G_{\mathrm{st}}}{\partial t_0 \partial t_2} - 4\frac{\partial^2 G_{\mathrm{st}}}{\partial t_1^2}\right) = 0.$$

Taking $\beta = 0$ and applying the proposition above we get

$$\frac{\partial^2 F}{\partial t_0 \partial t_1} - \frac{1}{2}\left(\frac{\partial^2 F}{\partial t_0^2}\right)^2 - \frac{1}{12}\frac{\partial^4 F}{\partial t_0^4} = 0,$$

which upon differentiation by t_0 yields the KdV equation for $U = \frac{\partial^2 F}{\partial t_0^2}$. The string and dilaton equations (5.1) for F become

$$\frac{\partial U}{\partial t_0} = 1 + \sum_{i=0}^{\infty} t_{i+1}\frac{\partial U}{\partial t_i}, \qquad \frac{\partial U}{\partial t_1} = \frac{2}{3}U + \frac{1}{3}\sum_{i=0}^{\infty}(2i+1)t_i\frac{\partial U}{\partial t_i}$$

when differentiated by t_0 twice. These give sufficient initial conditions to ensure that if U is a solution to the KdV equation, then it is a solution of the entire KdV hierarchy (see for example [LZ04, Remark 4.7.3]). This concludes the proof of the Witten-Kontsevich theorem.

In short, Kazarian and Lando's idea uses the ELSV formula as a bridge between intersections of ψ-classes on $\overline{\mathcal{M}}_{g,n}$ and Hurwitz numbers, in order to deduce properties of $F(t_*)$ from previously known properties of $H(\beta; p_*)$. An interesting question to ask is whether their method can be understood in terms of a reduction from the KP to the KdV hierarchy which we outlined in Example 4.2.5, and subsequently translated in a statement about tau-functions in Corollary 4.3.7. There does not seem to be an obvious way to do this. One of the reasons for this is that the transformation from p_* to t_* given by (5.6) does not provide one with a set of independent variables such that $F(t_*)$ is independent of the even-indexed ones.

One could also ask if the subsequent equations of the KP hierarchy for H give rise to higher equations of the KdV hierarchy for F. As we mentioned at the end of Section 4.3, the PDEs of the KP hierarchy are indexed by two-part partitions $(\mu_1, \mu_2) \in \mathcal{P}_d$ of $d \geq 4$ that do not contain a 1. The partition $(2,2)$ of $d = 4$ gives rise to the first KP equation which we have written above, while the partitions $(3,2)$ for $d = 5$, and $(4,2)$ and $(3,3)$ for $d = 6$ yield [Kaz08, Section 4]:

$$
\begin{aligned}
H_{3,2} &= H_{4,1} - \frac{1}{6} H_{2,1^3} - H_{1^2} H_{2,1}, \\
H_{4,2} &= H_{5,1} - \frac{1}{4} H_{3,1^3} - H_{1^2} H_{3,1} - \frac{1}{2} (H_{2,1})^2 + \frac{1}{8} (H_{1^3})^2 + \frac{1}{12} H_{1^2} H_{1^4} + \frac{1}{120} H_{1^6}, \\
H_{3,3} &= H_{5,1} - \frac{1}{3} H_{3,1^3} - H_{1^2} H_{3,1} - (H_{2,1})^2 + \frac{1}{4} (H_{1^3})^2 + \frac{1}{3} (H_{1^2})^3 + \frac{1}{3} H_{1^2} H_{1^4} + \frac{1}{45} H_{1^6}.
\end{aligned}
$$

We have used the short-hand notation

$$
H_{\mu_1^{n_1}, \ldots, \mu_\ell^{n_\ell}} = \frac{\partial^{n_1\mu_1 + \cdots + n_\ell \mu_\ell} H}{\partial p_{\mu_1}^{n_1} \cdots \partial p_{\mu_\ell}^{n_\ell}}.
$$

The corresponding PDEs for $H' := H_{\mathrm{st}} = H - H_{0;1} - H_{0;2}$ are

$$
\begin{aligned}
H'_{3,2} =& H'_{4,1} - \frac{1}{6} H'_{2,1^3} - H'_{1^2} H'_{2,1} - \frac{1}{2} \beta^2 H'_{2,1} - \frac{2}{3} \beta^3 H'_{1^2}, \\
H'_{4,2} =& H'_{5,1} - \frac{1}{4} H'_{3,1^3} - H'_{1^2} H'_{3,1} - \frac{1}{2} \left(H'_{2,1} \right)^2 + \frac{1}{8} (H'_{1^3})^2 + \frac{1}{12} H'_{1^2} H'_{1^4} + \frac{1}{120} H'_{1^6} \\
& - \beta^2 \left(\frac{1}{2} H'_{3,1} - \frac{1}{24} H'_{1^4} \right) - \frac{2}{3} \beta^3 H'_{2,1} - \frac{9}{8} \beta^4 H'_{2,1}, \\
H'_{3,3} =& H'_{5,1} - \frac{1}{3} H'_{3,1^3} - H'_{1^2} H'_{3,1} - \left(H'_{2,1} \right)^2 + \frac{1}{3} (H'_{1^2})^3 + \frac{1}{4} (H'_{1^3})^2 + \frac{1}{3} H'_{1^2} H'_{1^4} + \frac{1}{45} H'_{1^6} \\
& - \beta^2 \left(\frac{1}{2} H'_{3,1} - \frac{1}{2} (H'_{1^2})^2 - \frac{1}{6} H'_{1^4} \right) - \frac{4}{3} \beta^3 H'_{2,1} - \frac{7}{8} \beta^4 H'_{1^2}.
\end{aligned}
$$

Applying the changes in partial derivatives to the PDEs above like in the last section, with the use of the next two changes in partial derivatives

$$
\begin{aligned}
\frac{\partial}{\partial p_4} &= 64 \beta^{19/3} \frac{\partial}{\partial t_3} + 64 \beta^{17/3} \frac{\partial}{\partial t_2} + 32 \beta^{15/3} \frac{\partial}{\partial t_1} + \frac{32}{3} \beta^{13/3} \frac{\partial}{\partial t_0}, \\
\frac{\partial}{\partial p_5} &= 625 \beta^{24/3} \frac{\partial}{\partial t_4} + 625 \beta^{22/3} \frac{\partial}{\partial t_3} + \frac{625}{2} \beta^{20/3} \frac{\partial}{\partial t_2} + \frac{625}{6} \beta^{18/3} \frac{\partial}{\partial t_1} + \frac{625}{24} \beta^{16/3} \frac{\partial}{\partial t_0},
\end{aligned}
$$

all lead to the same outcome: the term containing the lowest power of $\beta^{2/3}$ (which is the $\beta^{17/3}$ term for $d = 5$ and the $\beta^{20/3}$ term for $d = 6$) vanishes, and dividing by the next highest power of $\beta^{2/3}$ leaves one with a PDE of the form

$$
\frac{\partial^2 G_{\mathrm{st}}}{\partial t_0 \partial t_1} - \frac{1}{2} \left(\frac{\partial^2 G_{\mathrm{st}}}{\partial t_0^2} \right)^2 - \frac{1}{12} \frac{\partial^4 G_{\mathrm{st}}}{\partial t_0^4} + \beta^{2/3} (\cdots) + \beta^{4/3} (\cdots) + \beta^{6/3} (\cdots) = 0.
$$

Once again, $\beta = 0$ yields the KdV equation for $U = \frac{\partial^2 F}{\partial t_0^2}$, rather than the higher

KdV equations. Thus our question has a negative answer, which is not too surprising considering that only the first KdV equation is needed to generate the entire KdV hierarchy when one has the string and dilaton equations.

As a conclusion to this thesis, we introduce a modern research area that has been heavily influenced by the Witten-Kontsevich theorem.

5.2 Beyond Witten's conjecture

Our discussion so far has centred around the moduli space of stable curves $\overline{\mathcal{M}}_{g,n}$, which is but a particular instance of a more general construction in the field of Gromov-Witten theory. We give a short introduction of the main objects in this field and how they relate to our work so far, with more details to be found in [FP97; CK99; Koc01; KV07; Ros14], for example.

Let X be a smooth projective variety over $\mathbb{C}$, and β an element in $A_1(X)$. Consider the tuple $(C, x_1, \ldots, x_n, \mu)$, where C is a genus g curve with n marked points $x_1, \ldots, x_n$, and $\mu : C \to X$ is a morphism satisfying $\mu_*([C]) = \beta$. An isomorphism between two such tuples $(C, x_1, \ldots, x_n, \mu)$ and $(C', x'_1, \ldots, x'_n, \mu')$ consists of an isomorphism $\tau : C \to C'$ sending x_i to x'_i, such that $\mu' \circ \tau = \mu$. Denote by $\mathcal{M}_{g,n}(X, \beta)$ the set of isomorphism classes of such tuples. Of course, in order to obtain a well-behaved space one may want to impose some additional conditions on β, depending on the nature of X. Our original moduli space of curves $\mathcal{M}_{g,n}$ is recovered when X is a point and $\beta = 0$.

There is a compactification $\overline{\mathcal{M}}_{g,n}(X, \beta)$ of the $\mathcal{M}_{g,n}(X, \beta)$, whose objects $(C, x_1, \ldots, x_n, \mu)$ consist of a pointed stable curve $(C, x_1, \ldots, x_n) \in \overline{\mathcal{M}}_{g,n}$ together with a *stable* map $\mu : C \to X$, which means that for every irreducible component $C' \subset C$ one must have:

- If $C' \cong \mathbb{P}^1$ and C' is mapped to a point by μ, then C' must contain at least three special points (we remind that a special point is either a marked point or a nodal point);
- If C' has genus 1 and is mapped to a point by μ, then C' must contain at least one special point.

The space $\overline{\mathcal{M}}_{g,n}(X, \beta)$ is a compact Deligne-Mumford stack, although in general not smooth [BM96]. Once again, our original moduli space of stable curves $\overline{\mathcal{M}}_{g,n}$ is recovered when X is a point and $\beta = 0$.

As before, there is morphism that forgets the last marked point

$$\pi : \overline{\mathcal{M}}_{g,n+1}(X,\beta) \to \overline{\mathcal{M}}_{g,n}(X,\beta)$$

which can be identified with the universal curve $\overline{\mathcal{C}}_{g,n}(X,\beta) \to \overline{\mathcal{M}}_{g,n}(X,\beta)$, and plays an important role in the theory. There are new morphisms that we did not have before. The first one forgets the stable map, while the second ones evaluates the stable map at the i^{th} marked point:

$$\begin{array}{rccc} p: & \overline{\mathcal{M}}_{g,n}(X,\beta) & \longrightarrow & \overline{\mathcal{M}}_{g,n}, \\ & (C, x_1, \dots, x_n, \mu) & \longmapsto & (C, x_1, \dots, x_n), \\ \mathrm{ev}_i: & \overline{\mathcal{M}}_{g,n}(X,\beta) & \longrightarrow & X, \\ & (C, x_1, \dots, x_n, \mu) & \longmapsto & \mu(x_i). \end{array}$$

Given generic subvarieties $X_1, \dots, X_n$ of X, their corresponding homology classes have Poincaré duals $\gamma_i \in H^*(X)$, and the intersection product

$$\mathrm{ev}_1^*\gamma_1 \cdot \mathrm{ev}_2^*\gamma_2 \cdots \mathrm{ev}_n^*\gamma_n \in H^*(\overline{\mathcal{M}}_{g,n}(X,\beta)) \tag{5.10}$$

represents those maps $\mu : C \to X$ such that $\mu(x_i) \in X_i$ for each $i = 1, \dots, n$ (we still take coefficients in $\mathbb{Q}$ in the cohomology ring). Since the location of the marked points on C varies over the moduli space, we can interpret the cohomology class above as representing the collection of morphisms $\mu : C \to X$ such that $\mu(C)$ intersects X_i for each $i = 1, \dots, n$. If (5.10) is a top degree class in $H^*(\overline{\mathcal{M}}_{g,n}(X,\beta))$, then the number of such morphisms should be finite and is given by

$$\int_{\overline{\mathcal{M}}_{g,n}(X,\beta)} \mathrm{ev}_1^*\gamma_1 \cdots \mathrm{ev}_n^*\gamma_n.$$

Unfortunately, the expression above does not make sense in general because the fundamental class of $\overline{\mathcal{M}}_{g,n}(X,\beta)$ is not always well-defined. When $\overline{\mathcal{M}}_{g,n}(X,\beta)$ happens to be smooth, the Grothendieck-Riemann-Roch theorem implies that its dimension is

$$\mathrm{vdim}\, \overline{\mathcal{M}}_{g,n}(X,\beta) := (3 - \dim X)(g-1) + c_1(X) \cap \beta + n.$$

This number is called the *virtual dimension* of $\overline{\mathcal{M}}_{g,n}(X,\beta)$. To get around the problem of not always having a fundamental class over which one can integrate, a *virtual fundamental class* $[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}} \in H_{2\,\mathrm{vdim}\,\overline{\mathcal{M}}_{g,n}(X,\beta)}(\overline{\mathcal{M}}_{g,n}(X,\beta))$ can be constructed [BF97; LT98] with similar properties to an ordinary fundamental class.

With this in mind, one defines the *primary Gromov-Witten invariants* to be

$$\langle \gamma_1 \cdots \gamma_n \rangle_{g,\beta}^X := \int_{[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}}} \mathrm{ev}_1^* \gamma_1 \cdots \mathrm{ev}_n^* \gamma_n.$$

Since we would like these invariants to not depend explicitly on the cycle β, introducing the Novikov ring [Get99, §1] allows us to write

$$\langle \gamma_1 \cdots \gamma_n \rangle_g^X = \sum_{\beta \in H_2^+(X,\mathbb{Z})} q^\beta \langle \tau_{d_1}(\gamma_1) \cdots \tau_{d_n}(\gamma_n) \rangle_{g,\beta}^X.$$

We will not comment on the Novikov ring, or on the meaning of $H_2^+(X,\mathbb{Z})$ and q^β. Now choose a basis $\gamma_1, \dots, \gamma_r$ of $H^2(X)$ and let $\gamma_0 = 1 \in H^0(X)$ be the Poincaré dual of the fundamental class $[X]$. The genus g Gromov-Witten potential is the generating series in $t_* = (t_0, t_1, \dots)$ and q given by

$$\Phi_g^X(q;t_*) := \sum_{(n_0,n_1,\dots)} \langle \gamma_0^{n_0} \gamma_1^{n_1} \cdots \rangle_g^X \prod_{i=0}^\infty \frac{t_{n_i}}{n_i!},$$

and the total Gromov-Witten potential is

$$Z^X(q;\hbar;t_*) := \exp\left(\sum_{g \geq 0} \hbar^{g-1} \Phi_g^X \right).$$

Analogously to what we did for $\overline{\mathcal{M}}_{g,n}$, one can construct ψ-classes on $\overline{\mathcal{M}}_{g,n}(X,\beta)$. Given the n sections of the universal curve $s_i : \overline{\mathcal{M}}_{g,n}(X,\beta) \to \overline{\mathcal{M}}_{g,n+1}(X,\beta)$ corresponding to the marked points, the pullback of the relative dualizing sheaf $\mathbb{L} \to \overline{\mathcal{M}}_{g,n+1}(X,\beta)$ by each s_i gives line bundles $\mathbb{L}_i \to \overline{\mathcal{M}}_{g,n}(X,\beta)$ whose fibre at $(C, x_1, \dots, x_n, \mu)$ is $T^*_{x_i}C$. The ψ-classes are $\psi_i = c_1(\mathbb{L}_i) \in H^2(\overline{\mathcal{M}}_{g,n}(X,\beta))$. The *descendant Gromov-Witten invariants* are

$$\langle \tau_{d_1}(\gamma_1) \cdots \tau_{d_n}(\gamma_n) \rangle_{g,\beta}^X = \int_{[\overline{\mathcal{M}}_{g,n}(X,\beta)]^{\mathrm{vir}}} \psi_1^{d_1} \cdots \psi_n^{d_n} \, \mathrm{ev}_1^* \gamma_1 \cdots \mathrm{ev}_n^* \gamma_n.$$

The genus g Gromov-Witten potential and the total Gromov-Witten potential are defined in a similar way:

$$D^X(q;\hbar;t_*) := \exp\left(\sum_{g \geq 0} \hbar^{g-1} F_g^X \right), \text{ where}$$
$$F_g^X(q;t_*) := \sum_{n \geq 0} \frac{1}{n!} \sum_{\substack{(d_1,\dots,d_n) \\ (a_1,\dots,a_n)}} \langle \tau_{d_1}(\gamma_{a_1}) \cdots \tau_{d_n}(\gamma_{a_n}) \rangle_g^X \, t_{d_1}^{a_1} \cdots t_{d_n}^{a_n}.$$

Thus the Witten-Kontsevich theorem concerns the descendant Gromov-Witten potential for $X = \{*\}$ (in this thesis, we took $\hbar = 1$ as is customary in the literature regarding the Witten-Kontsevich theorem). A generalization of Witten's conjecture to an arbitrary variety X is known as the Virasoro conjecture [Get99]. The statement is as follows: in [EHX97], the authors constructed a sequence $(L_k)_{k\geq -1}$ of differential operators satisfying

$$[L_k, L_\ell] = (k - \ell)L_{k+\ell}. \tag{5.11}$$

When $X = \{*\}$, the PDEs $L_{-1}D^X = 0$ and $L_0 D^X = 0$ correspond to the string and dilaton equations for $F(t_*)$.

Conjecture 5.2.1 (Virasoro)**.** *If X is a smooth projective variety over $\mathbb{C}$ then*

$$L_k D^X = 0 \textit{ for all } k \geq -1.$$

References

[Arn89] V. I. Arnold. *Mathematical Methods of Classical Mechanics.* Trans. Russian by K. Vogtmann and A. Weinstein. 2nd ed. Grad. Texts. in Math. 60. Springer, 1989.

[Arn96] V. I. Arnold. "Topological classification of trigonometric polynomials and combinatorics of graphs with an equal number of vertices and edges". In: *Functional Analysis and Its Applications* 30.1 (1996), pp. 1–14.

[Beh02] K. Behrend. *Cohomology of Stacks.* Lectures given at the School and Conference on Intersection Theory and Moduli, ICTP, Trieste. 2002.

[BF97] K. Behrend and B. Fantechi. "The intrinsic normal cone". In: *Invent. Math.* 128.1 (1997), pp. 45–88.

[BM96] K. Behrend and Y. Manin. "Stacks of stable maps and Gromov-Witten invariants". In: *Duke Math. J.* 85 (1996), pp. 1–60.

[Bou59] N. Bourbaki. *Algèbre: Chapitre 9.* Éléments de mathématique. Springer-Verlag, 1959.

[Bre93] G. E. Bredon. *Topology and Geometry.* Grad. Texts in Math. 139. Springer-Verlag, 1993.

[BRZ21] A. Buryak, P. Rossi, and D. Zvonkine. *Moduli spaces of residueless meromorphic differentials and the KP hierarchy.* arXiv:2110.01419. 2021.

[Bur22] A. Buryak. *Integrable systems as systems of PDEs with an infinite dimensional algebra of symmetries.* Extended lecture notes for course taught at the HSE University. 2022.

[Cav16] R. Cavalieri. "Hurwitz theory and the double ramification cycle". In: *Jpn. J. Math.* 11 (2016), pp. 305–331.

[Cho56] W. Chow. "On equivalence classes of cycles in an algebraic variety". In: *Ann. of Math.* 64.3 (1956), pp. 450–479.

[CK99] D. A. Cox and S. Katz. *Mirror symmetry and algebraic geometry.* Mathematical Surveys and Monographs 68. Amer. Math. Soc., 1999.

[CM16] R. Cavalieri and E. Miles. *Riemann Surfaces and Algebraic curves: A First Course in Hurwitz Theory.* London Mathematical Society Student Texts 87. Cambridge University Press, 2016.

[Dén59] J. Dénes. "The representation of a permutation as the product of a minimal number of transpositions and its connection with the theory of graphs". In: *Publ. Math. Ins. Hungar. Acad. Sci.* 4 (1959), pp. 63–70.

[Die08] T. tom Dieck. *Algebraic Topology.* EMS Textbooks in Mathematics. European Math. Soc., 2008.

[DM69] P. Deligne and D. Mumford. "The irreducibility of the space of curves of given genus". In: *Publications Mathématiques de l'IHÉS* 36 (1969), pp. 75–109.

[Dub+16] B. Dubrovin, S. Liu, D. Yang, and Y. Zhang. "Hodge integrals and tau-symmetric integrable hierarchies of Hamiltonian evolutionary PDEs". In: *Adv. in Math.* 293 (2016), pp. 382–435.

[Dun+15] P. Dunin-Barkowski, M. Kazarian, N. Orantin, S. Shadrin, and L. Spitz. "Polynomiality of Hurwitz numbers, Bouchard–Mariño conjecture, and a new proof of the ELSV formula". In: *Adv. in Math.* 279 (2015), pp. 67–103.

[EH16] D. Eisenbud and J. Harris. *3264 and All That: A Second Course in Algebraic Geometry.* Cambridge University Press, 2016.

[EHX97] T. Eguchi, K. Hori, and C. Xiong. "Quantum cohomology and Virasoro algebra". In: *Physics Letters B* 402 (1997), pp. 71–80.

[Eke+01] T. Ekedahl, S. Lando, M. Shapiro, and A. Vainshtein. "Hurwitz numbers and intersections on moduli spaces of curves". In: *Invent. Math.* 146.2 (2001), pp. 297–327.

[Eke+99] T. Ekedahl, S. Lando, M. Shapiro, and A. Vainshtein. "On Hurwitz numbers and Hodge integrals". In: *C. R. Acad. Sci. Paris* 328 (1999), pp. 1175–1180.

[Fab99] C. Faber. "A Conjectural Description of the Tautological Ring of the Moduli Space of Curves". In: *Moduli of Curves and Abelian Varieties.* Vol. 33. Aspects of Mathematics. Vieweg+Teubner Verlag, 1999, pp. 109–130.

[FH04] W. Fulton and J. Harris. *Representation Theory: A First Course.* Graduate Texts in Mathematics: Readings in Mathematics 129. Springer, 2004.

[FP02] B. Fantechi and R. Pandharipande. "Stable maps and branch divisors". In: *Compositio Math.* 130.3 (2002), pp. 345–364.

[FP03] C. Faber and R. Pandharipande. "Relative maps and tautological classes". In: *J. Eur. Math. Soc.* 7 (2003), pp. 13–49.

[FP97] W. Fulton and Pandharipande. *Notes on Stable Maps And Quantum Cohomology.* arXiv:alg-geom/9608011v2. Notes from a jointly taught class at the University of Chicago and lectures by the first author in Santa Cruz. 1997.

[Get99] E. Getzler. *The Virasoro conjecture for Gromov-Witten invariants.* arXiv:math/9812026. 1999.

[GJ08] I. P. Goulden and D. M. Jackson. "The KP hierarchy, branched covers, and triangulations". In: *Adv. in Math.* 219 (2008), pp. 932–951.

[GJ97] I. P. Goulden and D. M. Jackson. "Transitive factorisations into transpositions and holomorphic mappings on the sphere". In: *Proc. Amer. Math. Soc.* 125.1 (1997), pp. 51–60.

[GJV00] I. P. Goulden, D. M. Jackson, and A. Vainshtein. "The number of ramified coverings of the sphere by the torus and surfaces of higher genera". In: *Ann. of Comb.* 4 (2000), pp. 27–46.

[GP01] T. Graber and R. Pandharipande. "Constructions of nontautological classes of moduli spaces of curves". In: *Michigan Mathematical Journal* 51 (2001), pp. 93–110.

[GV03] T. Graber and R. Vakil. "Hodge integrals and Hurwitz numbers via virtual localization". In: *Compositio Math.* 135.1 (2003), pp. 25–36.

[GV05] T. Graber and R. Vakil. "Relative virtual localization and vanishing of tautological classes on moduli spaces of curves". In: *Duke Math. J.* 130.1 (2005), pp. 1–37.

[Har77] R. Hartshorne. *Algebraic geometry.* Grad. Texts in Math. 52. Springer, 1977.

[Hat01] A. Hatcher. *Algebraic Topology.* Cambridge University Press, 2001.

[HM98] J. Harris and I. Morrison. *Moduli of curves.* Grad. Texts in Math. 187. Springer-Verlag, 1998.

[Hor+03] K. Hori, S. Katz, A. Klemm, R. Pandharipande, R. Thomas, C. Vafa, R. Vakil, and E. Zaslow. *Mirror Symmetry.* Vol. 1. Clay Math. Monographs. Amer. Math. Soc., 2003.

[Hur91] A. Hurwitz. "Ueber Riemann'sche Flächen mit gegebenen Verzweigungspunkten". In: *Math. Ann.* 39.1 (1891), pp. 1–60.

[JM83] M. Jimbo and T. Miwa. "Solitons and Infinite Dimensional Lie Algebras". In: *Publ. RIMS, Kyoto Univ.* 19 (1983), pp. 943–1001.

[Kaz08] M. Kazarian. "KP hierarchy for Hodge integrals". In: *Adv. Math.* 221 (2008), pp. 1–21.

[KL06] Y. Kim and K. Liu. *A simple proof of Witten conjecture through localization.* math/0508384. 2006.

[KL07] M. E. Kazarian and S. K. Lando. "An algebro-geometric proof of Witten's conjecture". In: *J. Amer. Math. Soc.* 20 (2007), pp. 1079–1089.

[KL14] B. Kleiner and J. Lott. "Geometrization of three-dimensional orbifolds via Ricci flow". In: *Astérisque* 365 (2014), pp. 101–177.

[KM81] M. Kashiwara and T. Miwa. "The τ Function of the Kadomtsev-Petviashvili Equation". In: *Proc. Japan. Acad.* 57 (1981). Transformation Groups for Soliton Equations I, pp. 342–347.

[Knu83] F. F. Knudsen. "The projectivity of the moduli space of curves, II; III". In: *Math. Scand.* 52.2 (1983), pp. 161–199, 200–212.

[Koc01] J. Kock. *Notes on psi classes.* Notes from a mini-course in Belo Horizonte and a seminar at the Royal Institute of Technology in Stockholm. 2001.

[Kon92] M. Kontsevich. "Intersection theory on the moduli space of curves and the matrix airy function". In: *Commun. Math. Phys.* 147 (1992), pp. 1–23.

[KP70] B. B. Kadomtsev and V. I. Petviashvili. "On the Stability of Solitary Waves in Weakly Dispersing Media". In: *Soviet Physics Doklady* 15 (1970).

[KR87] V. G. Kac and A. K. Raina. *Bombay lectures on highest weight representations of infinite-dimensional Lie algebras.* Advanced Ser. in Math. Phys. 2. World Scientific Publishing, 1987.

[KV07] J. Kock and I. Vainsencher. *An Invitation to Quantum Cohomology: Kontsevich's Formula for Rational Plane Curves.* Vol. 249. Prog. in Math. Birkhauser Boston, 2007.

[Lan10] S. K. Lando. "Hurwitz numbers: on the edge between combinatorics and geometry". In: *Proc. Int. Cong. of Math.* Hyderabad, India, 2010, pp. 2444–2470.

[Liu10] C. C. M. Liu. *Lectures on the ELSV formula.* arXiv:1004.0853. 2010.

[Loo95] E. Looijenga. "On the tautological ring of $\mathcal{M}_g$". In: *Invent. Math.* 121 (1995), pp. 411–419.

[LT98] J. Li and G. Tian. “Virtual moduli cycles and Gromov-Witten invariants of algebraic varieties”. In: *J. Amer. Math. Soc.* 11 (1998), pp. 119–174.

[LZ04] S. K. Lando and A. K. Zvonkin. *Graphs on Surfaces and Their Applications*. Encyclopaedia of Mathematical Sciences 141. Springer-Verlag, 2004.

[Mir07] M. Mirzakhani. “Weil-Petersson volumes and intersection theory on the moduli space of curves”. In: *J. Amer. Math. Soc.* 20.1 (2007), pp. 1–23.

[Mir95] R. Miranda. *Algebraic Curves and Riemann Surfaces*. Grad. Stud. in Math. 5. Amer. Math. Soc., 1995.

[MJD00] T. Miwa, M. Jimbo, and E. Date. *Solitons: Differential equations, symmetries and infinite dimensional algebras*. Trans. Japanese by M. Reid. Cambridge University Press, 2000.

[MR99] J. E. Marsden and T. S. Ratiu. *Introduction to Mechanics and Symmetry. A Basic Exposition of Classical Mechanical Systems*. 2nd ed. Texts in Applied Mathematics 17. Springer New York, 1999.

[Mum83] D. Mumford. “Towards an Enumerative Geometry of the Moduli Space of Curves”. In: *Arithmetic and Geometry*. Vol. 36. Prog. in Math. 2. 1983.

[Mvo22] Y. Mvondo-She. “Integrable hierarchies, Hurwitz numbers and a branch point field in critical topologically massive gravity”. In: *SciPost Phys.* 12 (2022).

[Oko00] A. Okounkov. “Toda equations for Hurwitz numbers”. In: *Math. Res. Lett.* 7 (2000), pp. 447–453.

[OP01] A. Okounkov and R. Pandharipande. *Gromov-Witten theory, Hurwitz numbers, and Matrix models, I*. math.AG/0101147. 2001.

[Pan02] R. Pandharipande. “Three Questions in Gromov-Witten Theory”. In: *Proc. Int. Cong. of Math.* Vol. 2. Beijing, 2002, pp. 503–512.

[Ric22] A. Ricolfi. *An Invitation to Modern Enumerative Geometry*. Vol. 3. SISSA Springer Series. Springer, 2022.

[Ros14] S. Rose. *Introduction to Gromov-Witten Theory*. math.AG/1407.1260. 2014.

[Sag91] B. E. Sagan. *The Symmetric Group: Representations, Combinatorial Algorithms, and Symmetric Functions*. Wadsworth & Brooks/Cole Mathematics Series. Wadsworth, Inc., 1991.

[Sch20] J. Schmitt. *The moduli space of curves*. Lecture notes for a summer course at the University of Bonn. 2020.

[Sil09] J. H. Silverman. *The Arithmetic of Elliptic Curves*. 2nd ed. Grad. Texts in Math. 106. Springer, 2009.

[SMJ78] M. Sato, T. Miwa, and M. Jimbo. “Holonomic Quantum Fields I”. In: *Publ. Res. Inst. Math. Sci.* 14.1 (1978), pp. 223–267.

[Spi99] M. Spivak. *A Comprehensive Introduction to Differential Geometry*. 3rd ed. Vol. 1. Publish or Perish, Inc., 1999.

[Vak03] R. Vakil. “The moduli space of curves and its tautological ring”. In: *Notices of the Amer. Math. Soc.* 50.6 (2003), pp. 647–658.

[Vak08] R. Vakil. “The Moduli Space of Curves and Gromov-Witten Theory”. In: *Enumerative Invariants in Algebraic Geometry and String Theory*. Springer Berlin Heidelberg, 2008, pp. 143–198.

[Wan21] Z. Wang. “On affine coordinates for the tau-function for open intersection numbers”. In: *Nuc. Phys. B* 972 (2021).

[Wit90] E. Witten. “Two-dimensional gravity and intersection theory on moduli space”. In: *Surveys in differential geometry* 1 (1990), pp. 243–310.

[Zvo14] D. Zvonkine. *An introduction to moduli spaces of curves and their intersection theory*. Based on three lectures given at the Journées mathématiques de Glanon in July 2006. 2014.