



HAL
open science

Polynomial and topological tau functions of the Drinfeld–Sokolov hierarchies

Ann Du Crest de Villeneuve

► **To cite this version:**

Ann Du Crest de Villeneuve. Polynomial and topological tau functions of the Drinfeld–Sokolov hierarchies. Exactly Solvable and Integrable Systems [nlin.SI]. Université d'Angers, 2018. English. NNT : . tel-02022215v1

HAL Id: tel-02022215

<https://hal.science/tel-02022215v1>

Submitted on 17 Feb 2019 (v1), last revised 17 Apr 2019 (v2)

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

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

THÈSE

en vue de l'obtention du grade de

DOCTEURICE ÈS MATHÉMATIQUES

présentée à l'Université d'Angers par

Ann DU CREST DE VILLENEUVE

Fonctions tau polynomiales et topologique des hiérarchies de Drinfeld–Sokolov

soutenue le 13 décembre 2018 devant le jury composé de

Guido CARLET	Professeur à l'Université de Dijon	Président du jury
Tamara GRAVA	Professeure associée à la SISSA, Italie	Rapporteuse
Pol VANHAECKE	Professeur à l'Université de Poitiers	Rapporteur
Mattia CAFASSO	Maître de conférences à l'Université d'Angers	Directeur de thèse
Paolo ROSSI	Professeur associé à l'Université de Padoue, Italie	Codirecteur de thèse
Vladimir ROUBTSOV	Professeur à l'Université d'Angers	Codirecteur de thèse

préparée au Larema UMR CNRS 6093

PhD Thesis

Ann du Crest de Villeneuve

Polynomial and topological tau functions
of the Drinfeld–Sokolov hierarchies

*Fonctions tau polynomiales et topologique
des hiérarchies de Drinfeld–Sokolov*

Contents

Remerciements	9
Organization of the thesis	13
1 Introduction en français	15
1.1 Aperçu historique de la théorie et de nos résultats	15
1.2 La hiérarchie de Korteweg–de Vries : un exemple canonique	21
1.2.1 L'équation de Korteweg–de Vries	21
1.2.2 La hiérarchie de Korteweg–de Vries	25
1.2.3 Fonctions tau polynomiales de la hiérarchie KdV et les polynômes d'Adler–Moser	27
1.3 Fonctions tau polynomiales des hiérarchies de Drinfeld–Sokolov	28
1.3.1 Les hiérarchies de Drinfeld–Sokolov	29
1.3.2 Fonctions tau comme déterminants de Toeplitz	32
1.3.3 Fonctions tau polynomiales et applications	34
1.4 Hiérarchie de type D_4 et l'équivalence DR/DZ forte	37
1.4.1 Théories cohomologiques des champs et les hiérarchies de Dubrovin–Zhang	37
1.4.2 Les hiérarchies de double ramifications	41
1.4.3 Hiérarchie de type D_4 et l'équivalence DR/DZ forte	43
2 Introduction in English	47
2.1 Historical overview of the theory and of our results	47
2.2 The Korteweg–de Vries hierarchy: a canonical example	52
2.2.1 The Korteweg–de Vries equation	52
2.2.2 The Korteweg–de Vries hierarchy	57
2.2.3 Polynomial tau functions of the KdV hierarchy and the Adler–Moser polynomials	59
2.3 Polynomial tau functions of the Drinfeld–Sokolov hierarchies	60
2.3.1 The Drinfeld–Sokolov hierarchies	60
2.3.2 Tau functions as Toeplitz determinants	64
2.3.3 Polynomial tau functions and applications	65
2.4 Hierarchy of type D_4 and the strong DR/DZ equivalence	68

2.4.1	Cohomological field theories and the Dubrovin–Zhang hierarchies . . .	68
2.4.2	The double ramification hierarchies	72
2.4.3	Hierarchy of type D_4 and the strong DR/DZ equivalence	74
3	Polynomial tau functions of the KdV hierarchy	79
3.1	Pseudo-differential operators	79
3.2	The Kadomtsev–Petviashvili hierarchy	90
3.2.1	Lax pairs of the KP hierarchy	90
3.2.2	Equivalent forms of the KP hierarchy	92
3.3	The bilinear identity and tau functions	95
3.3.1	The bilinear identity	95
3.3.2	The tau functions of the KP hierarchy	99
3.4	The Korteweg–de Vries hierarchy	102
3.4.1	Lax pairs and other forms of the KdV hierarchy	102
3.4.2	The tau functions of the KdV hierarchy	104
3.5	The polynomial tau functions of KdV and the Adler–Moser polynomials . . .	105
3.5.1	The Adler–Moser Polynomials	106
3.5.2	A change of variables to the polynomial tau functions KdV	109
4	Polynomial tau functions of the DS hierarchies and Schur polynomials	115
4.1	Semisimple Lie algebras: a brief review	115
4.2	Untwisted affine Kac–Moody algebras	123
4.2.1	Definition and basic properties	123
4.2.2	Gradations and principal Heisenberg subalgebra	128
4.3	The Drinfeld–Sokolov hierarchies	134
4.3.1	Matrix Lax equations of the DS hierarchies	134
4.3.2	Integrability of the DS hierarchies	145
4.3.3	Example: From the A_1 -DS hierarchy to the KdV equation	149
4.3.4	The tau functions of the DS hierarchies	151
4.4	Tau functions as limits of block Toeplitz determinants	159
4.4.1	The Grassmannian approach and Toeplitz determinants	160
4.4.2	Relation with the tau functions of the DS hierarchies	170
4.5	Polynomial tau functions of the DS hierarchies and Schur polynomials	175
4.5.1	Young diagrams expansion of the tau function	175
4.5.2	Polynomial tau functions and bilinear equations	181
4.5.3	List of generalized Schur polynomials	188
5	Hierarchy of type D_4 and the strong DR/DZ equivalence	189
5.1	Infinite dimensional Hamiltonian structures	189
5.2	The double ramification hierarchy of type D_4	203
5.2.1	Cohomological field theories, FJRW theory	203
5.2.2	The double ramification hierarchies	213
5.2.3	Strong Double Ramification/Dubrovin–Zhang conjecture	217
5.2.4	Hamiltonian structure of the DR hierarchy of type D_4	220

5.2.5	Details on the computation of the Hamiltonian density $g_{1,1}$	225
5.3	The Drinfeld–Sokolov hierarchy of type D_4	228
5.3.1	Scalar Lax pair and bi-Hamiltonian structure of the DS hierarchy of type D_n	231
5.3.2	First Hamiltonian structure of the DS hierarchy of type D_4 and strong DR/DZ equivalence	236
A	Matrix realizations of classical Kac–Moody algebras	i
A.1	Classical simple Lie algebras	i
A.1.1	Examples: Rank 2 simple Lie algebras	iv
A.2	Classical untwisted affine Kac–Moody algebras	vi

Remerciements

Mes premiers remerciements, il va de soi, vont à Mattia d'avoir accepté de diriger ma thèse, ainsi qu'à Paolo et Volodya d'avoir accepté de me co-encadrer. Avoir trois encadrants de thèse m'a permis de ne jamais me sentir seulx,¹ d'avoir toujours eu quelqu'un vers qui me tourner mais aussi d'avoir à porter de main une variété de points de vue. Je les remercie tous les trois chaleureusement.

Puis je voudrais remercier Tamara Grava et Pol Vanhaecke d'avoir accepté de rapporter ma thèse ainsi que Guido Carlet d'avoir accepté de présider le jury de ma thèse. La rigueur dont ont fait preuve Tamara et Pol à la lecture de ma thèse me donne l'envie de persévérer. Je voudrais aussi remercier Étienne qui a encadré mon stage de Master et m'a conseillé de contacter Mattia, Paolo et Volodya, ainsi que de m'avoir accompagnés au début de ma thèse.

Mais ma thèse ne serait que peu sans les collaborations auxquelles j'ai eu la chance de participer. Je remercie Paolo de m'avoir lancés sur un sujet aussi passionnant et de m'avoir fait confiance dans notre collaboration. Je remercie Mattia de m'avoir proposé l'autre collaboration de ma thèse avec Di Yang de l'USTC en Chine. En fin de compte, j'ai eu la chance de travailler scientifiquement avec trois personnes sur deux sujets intimement liés et pourtant suffisamment distants pour m'offrir des perspectives complémentaires sur un domaine aussi vaste de la physique mathématique. Durant ces collaborations, j'ai eu le plaisir d'être accompagnés comme néophyte que j'étais. Pour autant, j'ai toujours eu le sentiment que mon avis était valorisé.

Durant ma thèse, j'ai eu la chance de travailler à l'Université de Dijon, à la SISSA à Trieste ainsi qu'au CRM, à l'Université de Montréal et à l'Université Concordia. Je remercie ces quatre universités de m'avoir accueillis dans les meilleures conditions. Il me faut aussi remercier les laboratoires Larema d'Angers et l'IMB à Dijon, l'école doctorale MathSTIC ainsi que l'action Marie Skłodowska-Curie IPaDEGAN d'avoir financé mes déplacements. J'ai rencontré lors de ces visites de nombreuses personnes que je remercie toutes et tous. Je tiens à remercier particulièrement John Harnad et Marco Bertola de m'avoir encadrés à Montréal. Et je remercie par ailleurs toutes les personnes extraordinaires que j'ai eu la chance de rencontrer à Montréal sans qui ma vie aujourd'hui serait vraisemblablement bien différente, en particulier, Anna, Grace, Jasper, Jeanne, Kellyann, Lau, Maggie, Ngozi, Ryan, Tyson, Yuan et plus que quiconque, B. Banjo Sullivan et Bobby.

¹J'utilise ici et là des terminaisons non standards et non genrées. Je me le permets, ce sont après tout mes remerciements. C'est assez simple, le x est muet.

Je remercie l'Université d'Angers et le laboratoire Larema de m'avoir accepté en thèse puis en ATER pour ma quatrième année. Par ailleurs, il me paraît évident que j'avais besoin d'une quatrième année pour rédiger ma thèse proprement. J'ai eu la chance de faire partie des dernières personnes à avoir obtenu un ATER pour la fin de thèse. J'espère que les politiques universitaires reconnaîtront suffisamment tôt l'erreur qui consisterait à ne (presque) plus offrir de positions ATER dans un futur proche. Forcer à finir une thèse de maths en trois ans, aujourd'hui en France, ce n'est pas seulement atteindre à la qualité des thèses préparées en France. C'est aussi atteindre à la santé des doctorant-es pour qui la thèse est souvent une expérience extrêmement stressante et difficile. Vis-à-vis des standards de la recherche à l'étranger, demander aux doctorant-es de finir leur thèse en trois ans c'est faire peser sur leurs épaules une pression qui me semble inutile et néfaste scientifiquement et mentalement. Gardons en mémoire que l'anxiété et la dépression chroniques touchent particulièrement les doctorant-es.

Sur une autre note, je tiens à remercier Frédéric Mangolte, Mohammed El Amrani, Michel Granger et Caroline Chalumeau de m'avoir aidé scientifiquement et professionnellement. Je remercie du fond du cœur Alexandra Le Petitcorps, simplement pour tout. Par ailleurs, je voudrais aussi remercier tous les élèves que j'ai eu le très grand plaisir d'avoir durant ces trois années et je n'ai qu'un seul souhait, que ce plaisir a été, dans une certaine mesure, réciproque.

Puis je voudrais remercier la faune doctorale du Larema. Ceux² d'avant : Benjamin, Delphine, Clément, Johan, Le, Mohamed, Viet-Anh ; et ceux d'après : Alexis, Axel, David, Jérôme, Marine, Marius, Sofia, sans oublier Yvain et Gauvain. J'espère que toutes et tous ont apprécié autant que moi parler de maths, d'anarchie et autres hooliganismes. Il me faut cependant remercier particulièrement Johan, qui m'a accueilli chaleureusement dans notre open-space I104. Puis ceux qui ont partagé avec moi cet espace et bien plus : Axel, Jérôme, Viet-Anh et Sofia.

Enfin, il me faut remercier ma famille, qui m'a soutenu durant ces trois années, en particulier ma mère et mon père. Ainsi que mes amis à Paris, à Lyon et ceux dispersés à travers le monde. Hugo et Antoine sans qui, malgré la distance, ma vie serait plus calme, ainsi que Lola, Jean-Baptiste et Richard que j'aime du fond du cœur ; il me faudrait une thèse pour leur exprimer ma gratitude. Mélie et Quentin dont la saveur unique et la joie qu'elle me procure me force à les visiter régulièrement et à qui je souhaite le plus grand des courages pour leurs fins de thèse respectives. Adèle, Èlise, Juliette, Quentin et Yann pour leur hospitalité ainsi que leur radicalité qui toujours m'inspirent. Puis j'aimerais avoir une pensée particulière pour Aaron et Marine qui, sur des plans parallèles et éloignés, m'ont apporté un soutien nécessaire et sincère pour lequel j'espère simplement avoir rendu la pareille. Enfin, je remercie Aria pour avoir été là pour moi, avoir partagé ma vie de thèse et ma vie tout court.

Puisqu'il s'agit d'une thèse de maths, je me dois de remercier particulièrement Sylvain. Sans lui, je n'aurais probablement pas préparé le concours pour entrer à l'ENS de Lyon ; sans lui, je n'aurais peut-être pas préparé l'agrégation ; sans lui, je n'aimerais pas autant les maths tout simplement. Les interactions avec Sylvain qui jalonnèrent ma thèse participèrent distinctement au plaisir que j'ai eu à la préparer. Il me faut aussi remercier toutes les personnes qui m'ont accompagné durant ma thèse que je n'ai pas la place de citer.

²Un doux mélange inclusif entre *celles* et *ceux*.

Mais en réalité, je souhaiterais remercier Mattia, ainsi que Paolo et Volodya différemment. Bien que, comme beaucoup, je n'ai eu l'expérience que d'une seule thèse, il me semble que Mattia m'a encadré avec brio, avec bienveillance, exigence et respect. Mattia m'a lancé sur un sujet passionnant et a su me transmettre sa propre passion. Il a toujours pris le temps d'écouter mes questions ainsi que mes questionnements. Il a su faire preuve de rigueur et exigence lorsque j'en avais besoin. En retour, lorsque je n'allais pas bien, comme c'est souvent arrivé durant ma thèse, non seulement je n'ai jamais eu peur d'aller lui en parler, mais en plus Mattia a fait preuve d'une compréhension et d'une compassion exemplaire, acceptant que je me retire le temps que ma santé mentale et physique s'améliore. Quelque soit la situation, Mattia me demandait d'abord si j'allais bien et ensuite où j'en étais.

Ma thèse fut parsemée d'événements personnels dont il serait absurde ne pas parler ici. Au début de ma troisième année de thèse, j'ai come out en tant que personne transgenre non binaire, à ma famille, à mon travail. Mattia, Paolo et Volodya ont fait preuve d'un respect à mon égard tel que ces quelques lignes ne me permettent pas d'exprimer la gratitude que j'ai à leur égard. Je me suis sentie en sécurité et pour cela je les remercie du fond du cœur. Plus que quiconque, Mattia a su établir un équilibre entre une relation professionnelle fructueuse et une relation humaine, amicale vertueuse. Mattia m'a encouragé et accompagné mathématiquement et humainement avec un enthousiasme et une rigueur qui forcent le respect. Je ne saurais le remercier suffisamment.

Organization of the thesis

Here we describe how the thesis is organized. Sections containing new results are preceded by a bold arrow \rightarrow . The other sections contain material that can be found in the literature and are given here so as to describe the context of our results.

Chapter 1 Polynomial tau functions of the Korteweg–de Vries hierarchy

Section 3.1: We detail the theory of pseudo–differential operators.

Section 3.2: We detail the theory of the Kadomtsev–Petviashvili hierarchy.

Section 3.3: We detail the theory of the bilinear identity (and its equivalence with the KP hierarchy) and of the tau function of the KP hierarchy.

Section 3.4: We detail the theory of the KdV hierarchy and its tau function.

- \rightarrow Section 3.5: We exhibit the change of variables that transforms the tau functions of the KdV hierarchy into the Adler–Moser polynomials using mainly KP theory; a work published [[Duc17](#)].

Chapter 2 Polynomial tau functions of the Drinfeld–Sokolov hierarchies and Schur polynomials

Section 4.1: We recall the theory of semisimple Lie algebras.

Section 4.2: We recall the theory of untwisted affine Kac–Moody algebras.

Section 4.3: We detail the theory of the Drinfeld–Sokolov hierarchies and Wu’s construction of a tau function [[Wu17](#)].

Section 4.4: We describe the construction of tau functions as Toeplitz determinants by Cafasso and Wu [[CW15a](#), [CW15b](#)].

- \rightarrow Section 4.5: We expand the tau function of the DS hierarchies as a sum over partitions of integers; through this we define generalized Schur polynomials. As an application, we give a criterion for the polynomiality of tau functions; we compute examples of polynomial tau functions and give applications to detecting bilinear equations of the DS hierarchies; a work published in [[CDY18](#)].

Chapter 3 Hamiltonian structure of the hierarchy of type D_4 and the strong DR/DZ equivalence

Section 5.1: We recall the theory of infinite-dim. Hamiltonian structures.

Section 5.2: We describe the construction of the double ramification (DR) hierarchies [Bur15a]; we recall the construction of the Dubrovin–Zhang (DZ) hierarchies [DZ08]; we explain the strong DR/DZ equivalence conjecture [BDGR18].

- Section 5.2.4: We detail the Hamiltonian structure of the double ramification hierarchy of type D_4 (for itself and in view of confirming the strong DR/DZ equivalence).

Section 5.3: We describe the theory of pseudo-differential operators of the second type and Liu–Wu–Zhang construction of scalar Lax pair and bi-Hamiltonian structure for the Drinfeld–Sokolov hierarchies of type D_n [LWZ10]

- Section 5.3.2: We detail the first Hamiltonian structure of the Drinfeld–Sokolov hierarchy of type D_4 and use the results to confirm the strong DR/DZ equivalence for this case; a work accepted for publication in [DR19].

Chapter 1

Introduction *en français*

1.1 Aperçu historique de la théorie et de nos résultats

Mise à part quelques équations, la section 1.1 est plutôt historique et devrait être lisible par les non spécialistes.

La présente thèse traite du calcul et des applications des fonctions tau des hiérarchies d'équations de Drinfeld–Sokolov. Les hiérarchies de Drinfeld–Sokolov sont des suites infinies d'équations aux dérivées partielles dont les flots commutent deux-à-deux. Pour cette raison, elles forment un système intégrable infini-dimensionnel. Elles furent introduites par V. G. Drinfeld et V. V. Sokolov dans leur article fondateur [DS84]. Dans ce dernier, ils montrèrent comment associer ces hiérarchies à n'importe quelle algèbre de Kac–Moody affine [Kac68, Moo68] \mathfrak{g} . Cette dernière, nous la réalisons comme l'extension centrale de l'algèbre de boucle d'une algèbre de Lie semisimple \mathfrak{g} , i.e.¹

$$\mathfrak{g} = \mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c, \quad (1.1)$$

avec le crochet de Lie approprié faisant de c un élément central ; voir (1.28).

Donnons la forme générale de ces systèmes d'équations. Une solution d'une hiérarchie de Drinfeld–Sokolov est la donnée de fonctions u^1, \dots, u^N qui dépendent d'une infinité de variables $\{t_i \mid i \geq 1\}$ avec $x = t_1$; le système d'équation de la hiérarchie a alors la forme

$$\{u_{t_i}^\alpha = A_i^\alpha(u^1, \dots, u^N) \mid \alpha \in \{1, \dots, N\}, i \geq 1\}, \quad (1.2)$$

¹Notons que l'algèbre de Kac–Moody affine associée à l'algèbre de Lie finie-dimensionnelle semisimple \mathfrak{g} est en réalité une extension co-centrale de l'algèbre \mathfrak{g} de (1.1) (voir [Kac90] §7). Cependant, les hiérarchies de Drinfeld–Sokolov n'impliquent pas l'extension co-centrale ; appeler \mathfrak{g} un algèbre affine est assez standard.

où $u_{t_i}^\alpha = \partial_{t_i} u^\alpha$ et où A_i^α est polynomial en les u^β et leurs dérivées par rapport à $x : u_x^\beta, u_{xx}^\beta$, etc. Par exemple, la célèbre équation de Korteweg–de Vries [Bou77, KdV95] est de cette forme :

$$u_t = \frac{1}{4}u_{xxx} + \frac{3}{2}uu_x. \quad (1.3)$$

Les hiérarchies de Drinfeld–Sokolov sont des suites d'équations ayant des propriétés similaires à celle de l'équation de Korteweg–de Vries et, comme nous le verrons, l'équation (1.3) joue un rôle crucial dans la théorie. Nous insistons sur le fait que bien que nous étudions des équations aux dérivées partielles, nous nous intéressons uniquement aux solutions algébriques de ces équations : polynômes, séries entières ou fractions rationnelles.

Dans la prochaine section de cette introduction, nous développerons les principaux concepts sur lesquels cette thèse repose et expliquerons nos résultats : [Duc17, CDY18, DR19]. Mais avant cela, nous faisons une brève histoire des hiérarchies de Drinfeld–Sokolov et des fonctions tau et espérons ainsi incorporer nos résultats à ce dense domaine de la physique mathématique.

Tout d'abord, dans le cas où \mathfrak{g} est l'algèbre affine associée à la l'algèbre de Lie simple $\mathfrak{sl}(n, \mathbb{C})$ (matrices complexes $n \times n$ de trace nulle), la hiérarchie de Drinfeld–Sokolov coïncide avec la n -ième hiérarchie de Gelfand–Dicke [GD76], ces dernières étant beaucoup plus simples ! La n -ième hiérarchie de Gelfand–Dicke admet la représentation suivante : considérons l'opérateur différentiel $L = \partial_x^n + u^{n-2}\partial_x^{n-2} + \dots + u^1\partial_x + u^0$ où les u^i sont des fonctions lisses des paramètres $\{t_i \mid i \geq 1, i \neq 0 \pmod n\}$ avec $t_1 = x$. Alors la n -ième hiérarchie de Gelfand–Dicke consiste en la famille d'équations²

$$\left\{ \frac{\partial L}{\partial t_i} = \left[(L^{i/n})_+, L \right] \mid i \geq 1, i \neq 0 \pmod n \right\}, \quad (1.4)$$

où $L^{1/n}$ la racine n -ième de L par rapport à la composition des opérateurs différentiels.³

En particulier, pour $\mathfrak{sl}(2, \mathbb{C})$ (fixant $n = 2$ dans (1.4)), on retrouve la hiérarchie de Korteweg–de Vries [Lax68, Miu68, GKM68] (dont le flot selon t_3 redonne l'équation de Korteweg–de Vries (1.3)). Les équations (1.4) sont appelées des *paires de Lax scalaires*. Pour la plupart, les hiérarchies de Drinfeld–Sokolov admettent une paire de Lax scalaire ainsi qu'une représentation bi-Hamiltonienne [DS84], les rendant intégrables de nombreuses façons. Cependant, dans le cas général, elles sont construites génériquement en tant que hiérarchies d'*équations de Lax matricielles* (ou équations de courbure nulle) que nous introduisons dans la section suivante. Notons que la commutativité des flots des hiérarchies de Drinfeld–Sokolov résulte essentiellement de l'existence de sous-algèbres abéliennes infinie-dimensionnelles de l'algèbre de Kac–Moody \mathfrak{g} .

Le développement de la notion de fonction tau. La notion de fonction tau émergea au début des années 1980 avec les travaux de l'école de Kyoto menée par Sato ainsi que Date, Jimbo,

²Le fait que nous ne considérons que les flots par rapport aux variables t_i avec $i \neq 0 \pmod n$ provient simplement du fait que pour $i = jn$, le flot est nul puisque $(L^{jn/n})_+ = L^j$ commute avec L .

³La définition de la racine n -ième nécessite l'introduction des opérateurs pseudo-différentiels ; voir §1.2.2.

Kashiwara, Miwa [JMS78, KM81, DKM81, DJKM81a] pour n'en citer que quelques uns. Originellement, la fonction tau est définie pour la hiérarchie de Kadomtsev–Petviashvili [Sat81] qui contient toutes les hiérarchies de Gelfand–Dickey.⁴ Une solution de la hiérarchie de Kadomtsev–Petviashvili est la donnée d'une infinité de fonctions ; la fonction tau est une seule fonction qui contient toute l'information de la solution. De plus, la fonction tau satisfait des équations différentielles partielles de nature différente de celles de la hiérarchie en soi. À savoir, les équations sur la fonction tau sont bilinéaires et intimement liées aux relations de Plücker. La hiérarchie de Kadomtsev–Petviashvili (KP) et sa fonction tau furent rapidement généralisées aux hiérarchies BKP et CKP, des hiérarchies de type KP liées aux algèbres de Lie de type orthogonale (B_n, D_n) et symplectique (C_n) [DJKM81b]. Pour les algèbres de Lie exceptionnelles, des hiérarchies de type KP furent définies par Kac et Wakimoto [KW89]. Nous retiendrons que la fonction tau est équivalente à la solution de la hiérarchie, mais que les deux donnent deux approches très différentes à la construction et la résolution de la hiérarchie.⁵

Dans le même temps, Edward Witten proposa sa célèbre conjecture de 1991 reliant une certaine fonction tau de la hiérarchie de Korteweg–de Vries à la fonction génératrice des nombres d'intersection sur l'espace de module des courbes stables $\overline{\mathcal{M}}_{g,n}$ [Wit91]. La conjecture de Witten, prouvée par Kontsevich l'année suivante [Kon92], initia une toute nouvelle façon d'approcher les hiérarchies intégrables que nous discutons ci-dessous.

Suite aux travaux de Drinfeld–Sokolov d'une part et de l'école de Kyoto et de Kac–Wakimoto d'autre part, plusieurs façons de définir les fonctions tau furent conçus pour les hiérarchies de Drinfeld–Sokolov. L'une d'elles, initiées par Kac [Kac90] et Hollowood et Miramontes [HM93, Mir99], fait usage des représentations de plus haut poids des algèbres de Kac–Moody affines, dans cette approche, on retrouve les équations bilinéaires. Une autre approche est de construire une certaine famille de densités Hamiltoniennes dites *tau symétriques* et qui, essentiellement par le lemme de Poincaré, donnent une définition naturelle de la fonction tau [DZ08].⁶

Dans la présente thèse, nous nous intéressons à deux classes de fonctions tau : les fonctions tau polynomiales et la fonction tau topologique. Les fonctions tau polynomiales forment une vaste classe de fonctions tau bien que les plus simples. Pour le cas particulier de la hiérarchie de Korteweg–de Vries, elles sont classifiées par les polynômes d'Adler–Moser [AM78]. Les zéros de ces polynômes évoluent selon un système intégrable dit de Calogero–Moser [AMcKM77] ; généralisé à certaines solutions de la hiérarchie de Kadomtsev–Petviashvili par Wilson [Wil98]. Ainsi, les fonctions tau polynomiales des hiérarchies de Drinfeld–Sokolov devraient avoir des propriétés similaires qu'il nous faut encore découvrir. D'autre part, la fonction tau topologique d'une hiérarchie de Drinfeld–Sokolov donnée est l'objet principale sur lequel se construisent les ponts entre les hiérarchies intégrables et la géométrie algébrique. Nous y reviendrons ci-dessous ; discutons en premier lieu les fonctions tau polynomiales.

⁴La hiérarchie de Kadomtsev–Petviashvili contient toutes les hiérarchies de Gelfand–Dickey dans le sens où toute solution d'une hiérarchie de Gelfand–Dickey est aussi solution de la hiérarchie de Kadomtsev–Petviashvili (l'inverse étant faux).

⁵Il doit être noté que la même école de Kyoto inventa, durant la même période, la théorie des fonctions tau pour les équations de Painlevé et les problèmes isomonodromiques [Miw81, JMW81]. Mais ce n'est pas la théorie que nous discutons ici.

⁶Pour la clarifier la chronologie : cet article fut originellement soumis à l'arXiv le 23 août 2001.

Fonctions tau polynomiales des hiérarchies de Drinfeld–Sokolov (traité au chapitre 4).

En 2012 Chao-Zhong Wu, inspiré par [HM93], proposa une construction universelle de densités Hamiltoniennes tau symétriques pour n’importe quelle hiérarchie de Drinfeld–Sokolov associée à une algèbre affine \mathfrak{g} [Wu17],⁷ donnant ainsi une fonction tau générique pour les hiérarchies de Drinfeld–Sokolov. Dans [CW15a, CW15b], Cafasso et Wu exprimèrent la fonction tau de Wu comme la limite de déterminants de Toeplitz *via* une certaine représentation fidèle de trace nulle de l’algèbre de Lie semisimple sous-jacente [DS84] (voir aussi l’appendice A). Cela généralisa en partie les premiers travaux de Cafasso [Caf08] sur les hiérarchies de Gelfand–Dickey.

La construction de Cafasso–Wu associe une fonction tau à tout élément $X \in \mathfrak{g}_-$ de l’algèbre affine \mathfrak{g} n’ayant que des puissances négatives du paramètre de boucle. Dans [CDY18], à partir des résultats de Cafasso et Wu, nous avons montré comment développer la fonction tau de Wu comme une somme sur toutes les partitions d’entiers (théorème 4.5.2) ; la somme fait intervenir des polynômes qui généralisent les polynômes de Schur [Sch01] (aussi [Mac95]). Conséquemment, nous avons obtenu un critère facile pour la polynomialité de la fonction tau : si $X \in \mathfrak{g}_-$ est tel que sa représentation matricielle est nilpotente (par rapport au produit usuel de matrices), alors la fonction tau associée est polynomiale.⁸ Nous avons donné quelques exemples non triviaux de fonctions tau polynomiales et décrit des applications aux équations bilinéaires des hiérarchies de Drinfeld–Sokolov. C’est l’objet du chapitre 4. Cela constituera un futur projet que d’étudier les propriétés de ces polynômes de Schur généralisés ainsi que de ces fonctions tau polynomiales.

Fonctions tau polynomiales de la hiérarchie KdV (traité au chapitre 3). Pendant que nous travaillons sur les fonctions tau polynomiales des hiérarchies de Drinfeld–Sokolov, nous nous intéressons au cas spécifique de la hiérarchie de Korteweg–de Vries (KdV). Cette dernière, comme nous l’avons décrit, est un cas particulier des hiérarchies de Drinfeld–Sokolov pour l’algèbre affine associée à l’algèbre de Lie simple $\mathfrak{sl}(2, \mathbb{C})$. De plus, toute solution de la hiérarchie de Korteweg–de Vries est aussi une solution de la hiérarchie de Kadomtsev–Petviashvili ; la réduction étant de mettre toutes les variables paires t_{2i} à zéro. Comme nous l’avons dit, les fonctions tau polynomiales de la hiérarchie de Korteweg–de Vries furent étudiées en profondeur et classifiées par Adler, McKean et Moser [AMcKM77, AM78]. De plus, comme cela fut compris d’abord par Sato [Sat81] (voir aussi [Hir04]), chaque fonction tau polynomiale de la hiérarchie de Korteweg–de Vries peut être écrite en terme des polynômes de Schur.

Le principal résultat de [AM78] est l’existence d’un unique changement des variables t_i qui transforme les fonctions tau polynomiales de la hiérarchie de Korteweg–de Vries en une suite de polynômes qui satisfait une simple récurrence différentielle, ce sont les polynômes d’Adler–Moser. Dans [Duc17], nous avons pu exhiber ce changement de variables (théorème 3.5.4 ; la preuve utilise essentiellement la théorie de la hiérarchie de Kadomtsev–Petviashvili. C’est l’objet du chapitre 3.

⁷Pour clarifier la chronologie : cet article fut originellement soumis à l’arXiv le 26 mars 2012.

⁸Dans certain cas, il se peut qu’en réalité une certaine puissance entière τ^a de la fonction tau soit polynomiale (dans les cas que nous avons calculés, $a = 1$ ou 2). Mais, de façon surprenante, dans tous les exemples que nous avons calculés, la fonction tau elle-même était polynomiale ; i.e. τ était un carré parfait.

Récemment, Kac et van de Leur contribuèrent de façon significative [KvdL18] en montrant que *toutes* les fonctions tau polynomiales de la hiérarchie de Kadomtsev–Petviashvili (incluant ainsi les hiérarchies de Gelfand–Dickey) peut être exprimée simplement comme un certain polynôme de Schur modulo un changement de variable.⁹ Ils calculèrent aussi les fonctions tau polynomiales de la hiérarchie de Kadomtsev–Petviashvili modifiée, mais nous n’en parlerons pas ici.

La conjecture de Witten et la fonction tau topologique. Nous revenons à présent à la conjecture de Witten et ses conséquences. La conjecture de Witten stipule que la série génératrice des nombres d’intersection sur l’espace de module des courbes stable $\overline{\mathcal{M}}_{g,n}$ est en réalité une fonction tau de la hiérarchie de Korteweg–de Vries [Wit91]. Cette conjecture fut prouvée par Maxim Kontsevich [Kon92] et l’on appelle aujourd’hui cette fonction tau la fonction tau de Witten–Kontsevich. Depuis lors, les fonctions tau devinrent un objet essentiel connectant la géométrie complexe, la géométrie algébrique et les systèmes intégrables dans un domaine actif de la physique mathématique parfois appelé la *symétrie miroir intégrable*. Peu de temps après, la conjecture de Witten fut généralisée pour connecter les fonctions tau des hiérarchies de Gelfand–Dickey aux classes r -spin de Witten [Wit93]. En 2003 Dubrovin et Zhang montrèrent que la fonction génératrice des invariants de Gromov–Witten de la droite complexe projective \mathbb{P}^1 est une fonction tau de la hiérarchie de Toda étendue [DZ03] (qui n’est *pas* une hiérarchie de Drinfeld–Sokolov).

Les théories cohomologiques des champs (CohFT pour *cohomological field theories*), introduites par Kontsevich et Manin [KM94], sont des classes de cohomologie sur l’espace de module des courbes stables définies en vue d’unifier les classes r -spin de Witten ainsi que les théories de Gromov–Witten. Au début des années 2010, Fan, Jarvis et Ruan définirent les CohFT de Fan–Jarvis–Ruan–Witten en vue de généraliser les classes r -spin de Witten à n’importe quelle singularité complexe [FJR11, FJR13].

D’autre part, dans [DZ08], Dubrovin et Zhang montrèrent comment associer à n’importe quelle CohFT semisimple¹⁰ une hiérarchie qui coïncident, dans certains cas,¹¹ à la hiérarchie de Drinfeld–Sokolov correspondante [GM05, FGM10]. Dans la construction de Dubrovin–Zhang, la hiérarchie est définie de sorte que la série génératrice de la CohFT semisimple est une fonction tau. Cette fonction tau particulière est appelée la *fonction topologique* ; par exemple, la fonction tau de Witten–Kontsevich est la fonction tau topologique de la hiérarchie de Korteweg–de Vries. Depuis lors, la symétrie miroir intégrable devint une question de tisser des liens entre les CohFT et les hiérarchies intégrables. Bien que certains exemples marchent très bien,¹² (voir [LRZ15] pour un clair exposé) il ne semble pas y avoir de schéma clair pour

⁹Ce résultat diffère des travaux originaux de Sato [Sat81] en ce qu’ici, on n’a plus besoin de considérer des combinaisons linéaires de polynômes de Schur.

¹⁰La semisimplicité signifie qu’une certaine algèbre associée est semisimple, i.e. somme direct d’algèbres simples.

¹¹Ces cas sont précisément ceux pour lesquels la hiérarchie de Drinfeld–Sokolov est construite à partir de l’algèbre affine associée à une algèbre de Lie simple de type $A_{n \geq 1}$, $D_{n \geq 4}$ ou E_6, E_7, E_8 (la classification ADE). Dans les cas restant, les hiérarchies de Dubrovin–Zhang et de Drinfeld–Sokolov sont aussi équivalentes dans le sens d’une transformation quasi-Miura (i.e. rationnelle).

¹²Parmi ceux-là, les exemples de correspondance « CohFT / Hiérarchie int. » que nous avons cités jusque là : Nombres d’intersection sur $\overline{\mathcal{M}}_{g,n}$ / Hiérarchie KdV ; Classe r -spin de Witten / Hiérarchies Gelfand–Dickey ; Théorie

l'instant.

L'un des buts de la construction de Dubrovin–Zhang était de classer les hiérarchies intégrables [DZ08, DLZ08]. L'idée était d'utiliser la théorie de l'intersection sur l'espace de module des courbes stables¹³ (et leurs généralisations : théorie de Gromov–Witten, théorie de Fan–Jarvis–Ruan–Witten) afin de définir des hiérarchies intégrables puis prouver que celles-ci sont équivalentes à des hiérarchies bien connues (telles que les hiérarchies de Drinfeld–Sokolov mais aussi d'autres). La relation d'équivalence entre les hiérarchies intégrables s'appelle une transformation de Miura : voir §1.4.2.

Les hiérarchies de double ramification (traité au chapitre 5). Récemment, Alexander Buryak, inspiré par la théorie des champs symplectique [EGH00], a défini une nouvelle classe de hiérarchies appelées les *hiérarchies de double ramification* [Bur15a]. Elles sont construites à partir des cycles de double ramification sur l'espace de module des courbes introduits par Graber et Vakil [GV05] (d'où le nom) et peuvent être définies pour n'importe quelle CohFT (pas seulement les CohFT semisimples, comme c'est le cas pour les hiérarchies de Dubrovin–Zhang). Dans ce même article, guidé par des exemples qu'il calcula, Buryak conjectura que dans le cas d'une CohFT semisimple, la hiérarchie de double ramification est équivalente à la hiérarchie de Dubrovin–Zhang correspondante *via* une transformation de Miura. En particulier, pour la CohFT associée à la théorie FJRW d'une singularité complexe simple de type ADE, la hiérarchie de double ramification, celle de Dubrovin–Zhang et celle de Drinfeld–Sokolov devraient coïncider selon la conjecture.

Les hiérarchies de double ramification possèdent plusieurs propriétés qui leur donnent un intérêt particulier. (1) Elles sont définies pour n'importe quelle CohFT. (2) Elles sont construites en tant que hiérarchies Hamiltoniennes pour lesquelles le crochet de Poisson est très simple (appelée le crochet standard) ; en particulier, (3) une famille de densités Hamiltoniennes tau symétrique est aisément obtenue [BDGR18]. (4) Leurs densités Hamiltoniennes satisfont des relations de récurrence fortes jusque lors inconnues [BR16b]. (5) Elles admettent une quantification relativement simple [BR16a, BDGR16], mais nous n'en parlerons pas ici. De plus, l'équivalence conjecturale avec les hiérarchies de Dubrovin–Zhang a été augmentée en une transformation de Miura uniquement déterminée, d'où la *conjecture DR/DZ forte* 5.2.23 [BDGR18]. Jusque là, l'équivalence DR/DZ forte a été confirmée dans plusieurs cas incluant les hiérarchies de Drinfeld–Sokolov de type A_1 à A_5 , la hiérarchie de Toda étendue et d'autres (voir 5.2.23).

Dans un papier accepté pour publication [DR19], nous avons confirmé l'équivalence DR/DZ forte pour la hiérarchie de Drinfeld–Sokolov associée à l'algèbre affine de l'algèbre de Lie simple de type D_4 . La confirmation de la conjecture implique le calcul de la première structure Hamiltonienne de la hiérarchie de Drinfeld–Sokolov de type D_4 . Pour cela, nous avons utilisé la paire de Lax scalaire de la hiérarchie de type D_n obtenue par Liu, Wu et Zhang [LWZ10]. Ces calculs, à notre connaissance, sont nouveaux vis-à-vis de la littérature.

de Gromov–Witten de \mathbb{P}^1 / Hiérarchie de Toda étendue.

¹³La construction de Dubrovin–Zhang s'applique en réalité à des objets fortement reliés appelés des variétés de Frobenius, mais nous ne nous concentrerons pas sur cette approche.

1.2 La hiérarchie de Korteweg–de Vries : un exemple canonique

Dans cette section, nous présentons certains des aspect fondamentaux de l'équation et de la hiérarchie de Korteweg–de Vries (KdV). Cela servira d'illustration à de nombreuses propriétés des hiérarchies intégrables. Le chapitre 3 est dédié aux fonctions tau polynomiales de la hiérarchie KdV ; nous décrivons les résultats obtenus dans [Duc17] à la fin de la section §1.2.

1.2.1 L'équation de Korteweg–de Vries

Mise à part ses applications en géométrie, l'équation KdV provient de mécanique des fluides. Elle décrit l'évolution de la surface de l'eau dans un canal, ce dernier modélisant des eaux peu profondes s'écoulant essentiellement dans une seule direction. Si l'on ignore la direction transverse, on peut décrire la hauteur de la surface de l'eau par une fonction scalaire $u(x, t)$ où t représente le temps et x représente la direction dans laquelle l'eau s'écoule. Alors u doit satisfaire l'équation KdV :

$$u_t = \frac{1}{4}u_{xxx} + \frac{3}{2}uu_x, \quad (1.5)$$

où $u_t = \partial_t u$ et $u_x = \partial_x u$, etc. L'équation ci-dessus est en réalité une version normalisée de KdV où les constantes physiques ont été éliminées. L'équation KdV fut en premier théorisée par Boussinesq [Bou77] puis proposée dans sa version actuelle par Korteweg et de Vries [KdV95]. Pour la dérivation physique de l'équation (et celle de l'équation KP), nous référons à l'introduction de [AC91] par Ablowitz et Clarkson.

Notons qu'en renormalisant u , x et t , on peut mettre des coefficients arbitraires devant chaque monôme u_t , u_{xxx} et uu_x ; le choix 1, 1/4 et 3/2 deviendrait clair à la section §1.2.2. L'équation KdV peut être écrite de diverses façons, permettant ainsi des connections avec de nombreux domaines *a priori* déconnectés. Ici nous décrivons quatre représentations : (a) la paire de Lax scalaire, (b) la paire de Lax matricielle, (c) la représentation bi-Hamiltonienne et (d) la représentation bilinéaire.

(a) La paire de Lax scalaire. Considérons l'opérateur $L = \partial_x^2 + u$ souvent appelé l'opérateur de Schrödinger.¹⁴ Alors on peut réécrire l'équation (1.5) comme

$$\frac{\partial L}{\partial t} = [P, L], \quad P = \partial_x^3 + \frac{3}{2}u\partial_x + \frac{3}{4}u_x, \quad (1.6)$$

¹⁴En effet, si u est vu comme un potentiel en mécanique quantique, alors l'équation de Schrödinger normalisée sur la fonction d'onde ψ devient $i\hbar\psi_t = -\hbar^2\psi_{xx} + u\psi$.

où $[L, P] = L \circ P - P \circ L$ est le commutateur entre opérateurs différentiels. La composition d'opérateurs différentiels est donnée par la règle de Leibniz, définie sur les monômes par¹⁵

$$a\partial_x^n \circ b\partial_x^p := \sum_{k=0}^n \binom{n}{k} ab_{(k)}\partial_x^{n+p-k}, \quad (1.7)$$

où $b_{(k)} := \partial_x^k(b)$, et étendue par linéarité. Nous omettrons le signe \circ lorsqu'il n'y a pas d'ambiguïté. La dérivation ∂_x est définie sur les opérateurs différentiels comme agissant sur les coefficients : $\partial_t(ad_x^n) := a_t\partial_x^n$. En particulier, l'action de ∂_t sur les opérateurs satisfait la règle de Leibniz. Ainsi, $\partial_t L = u_t$, mais c'est uniquement dû à la forme spécifique L . La paire (L, P) est ce qu'on appelle une *paire de Lax scalaire* pour l'équation KdV. Cette représentation est due à Peter D. Lax lui-même [Lax68].

Les équations (1.5) et (1.6) sont de nature différente : l'équation (1.5) est une EDP non linéaire en u , tandis que l'équation (1.6) est une EDO linéaire en L et les deux côtés de l'équation sont des dérivations (des opérateurs de Leibniz) sur les opérateurs différentiels. La formule pour l'opérateur P , à ce stade, semble plutôt *ad hoc*, mais elle entrera dans un élégant schéma avec la hiérarchie KdV. Un intérêt particulier de la paire de Lax scalaire est qu'elle permet de généraliser l'équation et la hiérarchie KdV à la hiérarchie de Kadomtsev–Petviashvili (voir §3.2).

(b) La paire de Lax matricielle. Considérons à présent l'opérateur différentiel matricielle d'ordre 1 suivant :

$$\mathcal{L} = \partial_x + \begin{pmatrix} 0 & \lambda - u \\ 1 & 0 \end{pmatrix}. \quad (1.8)$$

Ci-dessus, λ est une nouvelle indéterminée qu'on appelle le *paramètre spectral*. Toute identité doit être valable pour toute valeur complexe du paramètre λ . Considérons à présent la matrice suivante :

$$\mathcal{M} = \begin{pmatrix} -\frac{1}{4}u_x & \lambda^2 - \frac{1}{2}u\lambda - \frac{1}{2}(u)^2 - \frac{1}{4}u_{xx} \\ \lambda + \frac{1}{2}u & +\frac{1}{4}u_x \end{pmatrix}. \quad (1.9)$$

Alors

$$\frac{\partial \mathcal{L}}{\partial t} - [\mathcal{L}, \mathcal{M}] = \begin{pmatrix} 0 & -u_t + \frac{1}{4}u_{xxx} + \frac{3}{2}uu_x \\ 0 & 0 \end{pmatrix}.$$

Dans l'équation ci-dessus, $[\mathcal{L}, \mathcal{M}] := \partial_x \mathcal{M} + [\mathcal{N}, \mathcal{M}]$ où $\mathcal{N} = \mathcal{L} - \partial_x$ et où la dérivation $\partial_x \mathcal{M}$ est définie comme agissant sur les composantes de la matrice. Ainsi, $\partial_t \mathcal{L} = [\mathcal{L}, \mathcal{M}]$ si et seulement si $u_t = \frac{1}{4}u_{xxx} + \frac{3}{2}uu_x$, i.e. si et seulement si l'équation KdV (1.5) est satisfaite. Autrement dit, l'équation KdV (1.5) peut être réécrite comme

$$\frac{\partial \mathcal{L}}{\partial t} = [\mathcal{L}, \mathcal{M}]. \quad (1.10)$$

¹⁵Pour les non spécialistes, la composition des opérateurs différentiels donnée à l'équation (1.7) vient simplement de l'action sur des fonctions test. Par exemple, $\partial_x(f\partial_x(\psi)) = \partial_x(f\psi_x) = f_x\psi_x + f\psi_{xx}$, ce qui est en accord avec la formule $\partial_x \circ f\partial_x$ donnée par l'équation (1.7).

C'est équivalent à l'équation de courbure nulle

$$\partial_t \mathcal{N} - \partial_x \mathcal{M} - [\mathcal{N}, \mathcal{M}] = 0.$$

La paire $(\mathcal{L}, \mathcal{M})$ est appelée une *paire de Lax matricielle* pour l'équation KdV. Cette représentation est due à Ablowitz, Kaup, Newell et Segur [AKNS74].

Une fois de plus, les équations (1.6) sont (1.10) de nature différente ; cette dernière n'implique que des opérateurs différentiels d'ordre 1 et permet ainsi la représentation par courbure nulle. Un intérêt particulier est que les hiérarchies de Drinfeld–Sokolov sont construites génériquement comme des équations de Lax matricielles.

(c) La représentation bi-Hamiltonienne. On introduit l'anneau $\mathcal{A} = \mathbb{C}[u, u_x, u_{xx}, \dots]$ dont nous appelons les éléments des *polynômes différentiels* en u . L'opérateur ∂_x agissant sur \mathcal{A} s'identifie alors avec l'opérateur $\partial_x = \sum_{k \geq 0} u_{(k+1)} \partial_{u_{(k)}}$. Par ailleurs, on définit l'espace quotient suivant,¹⁶

$$\mathcal{F} := \mathcal{A} / \text{Im}(\partial_x),$$

dont nous appelons les éléments des *fonctionnelles locales*. Pour tout polynôme différentiel $f \in \mathcal{A}$, on note sa classe $\bar{f} \in \mathcal{F}$ par $\bar{f} = \int f$.¹⁷ En tant que conséquence immédiate, on obtient une formule d'intégration par partie : $\int f \partial_x^k g = - \int g \partial_x^k f$, ou, plus généralement, pour $k \geq 1$,

$$\int f \partial_x^k g = (-1)^k \int g \partial_x^k f. \quad (1.11)$$

Puis l'on définit sur \mathcal{A} l'opérateur $\delta/\delta u$, appelé la *dérivée variationnelle* (aussi appelé l'opérateur d'Euler–Lagrange), par

$$\begin{aligned} \frac{\delta f}{\delta u} &:= \sum_{k \geq 0} (-1)^k \partial_x^k \frac{\partial f}{\partial u_{(k)}} \\ &= \frac{\partial f}{\partial u} - \partial_x \frac{\partial f}{\partial u_x} + \partial_x^2 \frac{\partial f}{\partial u_{xx}} - \dots \end{aligned}$$

Alors pour tout f ne contenant pas de terme constant, $\delta f/\delta u = 0$ si et seulement si $f \in \text{Im}(\partial_x)$ (voir la proposition 5.1.1). En particulier, $\delta/\delta u$ est bien défini sur \mathcal{F} . On définit ensuite deux applications, $\{-, -\}_1, \{-, -\}_2 : \mathcal{F} \otimes \mathcal{F} \rightarrow \mathcal{F}$ by¹⁸

$$\{\bar{f}, \bar{g}\}_1 := \int \frac{\delta f}{\delta u} \partial_x \frac{\delta g}{\delta u}, \quad \{\bar{f}, \bar{g}\}_2 := \int \frac{\delta f}{\delta u} \left(\frac{1}{4} \partial_x^3 + u \partial_x + \frac{1}{2} u_x \right) \frac{\delta g}{\delta u}. \quad (1.12)$$

¹⁶Notons que l'espace quotient $\mathcal{F} = \mathcal{A}/\text{Im}(\partial_x)$ est seulement un espace vectoriel et non un anneau car l'opérateur ∂_x est certes linéaire, mais n'est pas un morphisme anneau (il satisfait la règle de Leibniz).

¹⁷Nous insistons sur le fait que le symbole intégrale ici n'est vraiment qu'une notation. Cela dit, par construction $\int \partial_x f = 0$, et cela peut être réalisé par les fonctions sur le cercle S^1 . Ainsi, on parle parfois de solutions périodiques de l'équation KdV (voir [Lax75]).

¹⁸Dans l'équation (1.12), les polynômes différentiels $f, g \in \mathcal{A}$ sont n'importe quels représentants des classes $\bar{f}, \bar{g} \in \mathcal{F}$ (i.e. $f = \bar{f} + \text{Im}(\partial_x)$). C'est bien défini puisque $\delta/\delta u$ s'annule sur $\text{Im}(\partial_x)$.

Ces applications sont antisymétriques et satisfont l'identité de Jacobi ; on appelle ce type d'applications des *structures de Poisson locales*.¹⁹ De plus, toute combinaison linéaire $\lambda\{-, -\}_1 + \mu\{-, -\}_2$ est encore une structure de Poisson locale ; on dit que les structures de Poisson sont compatibles. On définit aussi deux applications auxiliaires, notées pareillement $\{-, -\}_a : \mathcal{A} \otimes \mathcal{F} \rightarrow \mathcal{A}$, avec $a = 1, 2$, par

$$\{f, \bar{g}\}_1 := \sum_{k \geq 0} \frac{\partial f}{\partial u^{(k)}} \partial_x^{k+1} \frac{\delta g}{\delta u}, \quad \{f, \bar{g}\}_2 := \sum_{k \geq 0} \frac{\partial f}{\partial u^{(k)}} \partial_x^k \left(\frac{1}{4} \partial_x^3 + u \partial_x + \frac{1}{2} u_x \right) \frac{\delta g}{\delta u}. \quad (1.13)$$

Par construction, et grâce à l'équation (1.11), ces applications vérifient $\int \{f, \bar{g}\}_a = \int \{\bar{f}, g\}_a$, pour $a = 1, 2$. Ainsi, il est aisé de vérifier que l'équation KdV (1.3) est équivalente aux deux équations suivantes :

$$u_t = \{u, \bar{h}_1\}_1 = \partial_x \frac{\delta h_1}{\delta u}, \quad \bar{h}_1 = \int \frac{1}{4} u^3 + \frac{1}{8} u u_{xx}, \quad (1.14)$$

$$u_t = \{u, \bar{h}_2\}_2 = \left(\frac{1}{4} \partial_x^3 + u \partial_x + \frac{1}{2} u_x \right) \frac{\delta h_2}{\delta u}, \quad \bar{h}_2 = \int \frac{4}{8} u^2. \quad (1.15)$$

Les fonctionnelles locales \bar{h}_1 et \bar{h}_2 sont appelées des *fonctionnelles Hamiltoniennes* et les équations ci-dessus sont appelées des *représentations Hamiltoniennes* de l'équation KdV. Le fait que KdV admette deux représentations Hamiltoniennes avec deux structures de Poisson locales compatibles est ce qu'on appelle une *représentation bi-Hamiltonienne*.

La représentation (1.14) fut trouvée simultanément en 1971 par Gardner [Gar71] et Fadeev et Zakharov [FZ71] (bien que Fadeev et Zakharov l'attribuent à Gardner) ; on l'appelle la première représentation Hamiltonienne ou bien celle de Gardner–Zakharov–Fadeev. La représentation (1.15) est due à Magri [Mag78] et est appelée la seconde représentation ou bien celle de Magri.²⁰

Les représentations bi-Hamiltoniennes furent intensément étudiées par Dubrovin et Zhang (e.g. [DZ08, DLZ08]), en particulier dans le but de classifier les hiérarchies intégrables. De plus, les hiérarchies de double ramification, que nous étudions au chapitre 5, apparaissent naturellement dans une représentation Hamiltonienne.

(d) La représentation bilinéaire. Finalement, nous décrivons la représentation bilinéaire, ou de Hirota, de l'équation KdV. Contrairement aux représentations précédente, celle bilinéaire agit sur la fonction tau τ et non sur la solutions u . Comme mentionné précédemment, il y a diverses façons de définir la fonctions tau. Mais dans le cas de l'équation KdV, cela revient

¹⁹Puisque \mathcal{F} n'est pas un anneau, on ne peut pas écrire la règle de Leibniz pour les structures de Poisson locale. Ainsi, $(\mathcal{F}, \{-, -\}_a)$, pour $a = 1, 2$, est une algèbre de Lie, mais pas une algèbre de Poisson. Cela dit, pour toute fonctionnelle locale $\bar{g} \in \mathcal{F}$, les applications $f \mapsto \{f, \bar{g}\}_a$ de l'équation (1.13) satisfont la règle de Leibniz.

²⁰Magri en réalité contribua bien plus en étant le premier à découvrir comment reconstruire toute la hiérarchie KdV en utilisant la structure bi-Hamiltonienne. Ce faisant, une étape fondamentale était achevée dans la compréhension des techniques de reconstruction pour les hiérarchies intégrables. Cependant, nous ne le discuterons pas dans ce document.

toujours à définir τ par la relation²¹

$$u = -2 \frac{\partial^2 \log \tau}{\partial x^2}. \quad (1.16)$$

En particulier, à la donnée d'une solution u , la fonction tau τ n'est pas unique. On introduit à présent les dérivées bilinéaire, ou de Hirota, D_{t_i} avec $i \geq 1$. On pose²²

$$\exp \left(\sum_{i \geq 1} h_i D_{t_i} \right) (f, g) = f(\mathbf{t} + \mathbf{h}) \cdot g(\mathbf{t} - \mathbf{h}), \quad (1.17)$$

où chaque côté de l'équation est développé en les puissances de h_i et identifiés en tant que séries en les h_i . Par exemple, en posant $x = t_1$ et $t = t_3$ (les indices n'importent pas ici), on obtient

$$D_x(f, g) = \frac{\partial f}{\partial x} g - f \frac{\partial g}{\partial x}, \quad D_x D_t(f, g) = \frac{\partial^2 f}{\partial x \partial t} g + f \frac{\partial^2 g}{\partial x \partial t} - \frac{\partial f}{\partial x} \frac{\partial g}{\partial t} - \frac{\partial f}{\partial t} \frac{\partial g}{\partial x}.$$

Ainsi, c'est un peu long à vérifier, mais l'équation KdV (1.5), en substituant $u = -2\partial_x^2 \log \tau$, est équivalente à l'équation suivante,

$$(D_x^4 - 4D_x D_t)(\tau, \tau) = 0. \quad (1.18)$$

L'équation ci-dessus est appelée une équation de Hirota, d'après Ryogo Hirota qui les introduisit dans [Hir76]. Cette équation est relativement compacte compte tenu du nombre de termes que les dérivées logarithmiques font apparaître !

D'aucuns se demanderaient quel est l'intérêt de ce type d'équations. C'est en réalité moins clair pour la hiérarchie KdV que pour d'autres hiérarchies plus compliquées. En particulier, lorsqu'une solution de la hiérarchie consiste en plusieurs fonctions u^1, \dots, u^N (et pas seulement u), il y a toujours une seule fonction τ associée à la solution. Ainsi, la représentation bilinéaire condense toute la hiérarchie sur une seule fonction. De plus, il se peut que la fonction tau soit plus simple à calculer que la solution elle-même. C'est par exemple l'approche des chapitres 4 et 4, bien qu'ils n'impliquent pas directement les équations bilinéaires.

1.2.2 La hiérarchie de Korteweg–de Vries

La hiérarchie KdV fut introduite en 1968 simultanément par Lax [Lax68] et Gardner–Kruskal–Miura–Zabusky [Miu68, GKM68, GKMZ70].²³ Elle consiste en une suite d'équations commutant deux-à-deux (expliqué ci-dessous) et qui contient l'équation KdV (1.3), incluant ainsi l'équation KdV dans un système d'EDP infinie-dimensionnel. Dans ce qui suit, nous montrons comment construire la hiérarchie KdV en partant de (A) la paire de Lax scalaire, due

²¹Le facteur (-2) est expliqué à la proposition 3.4.5 ; il ne peut être éliminé puisque KdV est non linéaire.

²²Notons que lorsqu'on développe l'identité (1.17) le symbole $D_{t_{i_1}} \cdots D_{t_{i_n}}$ est un opérateur en soi et non un produit d'opérateurs.

²³Elle fut en réalité découverte pour la première fois par Andrew Lenard durant l'été 1967, bien qu'il n'est fameusement jamais écrit de papier sur le sujet. Voir [PS05] pour les détails de cette découverte et comment cela fut reconnu par Lax, Miura et al.

à Lax [Lax68] ; puis nous discuterons (B) sa paire de Lax matricielle, (C) sa représentation bi-Hamiltonienne et (D) sa représentation bilinéaire. Sa paire de Lax matricielle est la plus technique ; nous la discuterons en réalité à la section 1.3 avec le cas général des hiérarchies de Drinfeld–Sokolov.

(A) Les paires de Lax scalaire. Pour les paires de Lax scalaires, il suffit de considérer la racine carrée $L^{1/2}$ de l'opérateur de Schrödinger $L = \partial_x^2 + u$. Pour cela, nous commençons par étendre l'algèbre des opérateurs différentiels en une algèbre

$$\mathcal{D} = \left\{ \sum_{i=-\infty}^n a^i \partial^i \mid a^i \in \mathbb{C}[[x]], n \in \mathbb{Z} \right\}, \quad (1.19)$$

équipée du même produit défini à l'équation (1.7), mais où les coefficients binomiaux ont été généralisés à $\binom{n}{k}$ avec $n < 0$ en posant $\binom{n}{k} := n(n-1)\cdots(n-k+1)/k!$. Nous appelons les éléments de \mathcal{D} des *opérateurs pseudo-différentiels*. Dans cette algèbre, il existe alors un opérateur $L^{1/2} = \partial_x + \sum_{k<0} p^k \partial_x^k$, avec des fonctions $p^k \in \mathcal{A}$, satisfaisant $(L^{1/2})^2 = L$ (voir théorème 3.1.8). Alors la hiérarchie KdV peut être formulée de la façon suivante,

$$\left\{ \frac{\partial L}{\partial t_i} = \left[(L^{i/2})_+, L \right] \mid i > 0 \text{ odd} \right\}, \quad (1.20)$$

où $(-)_+$ désigne la projection sur le sous-espace de \mathcal{D} contenant seulement des puissances positives ou nulles de ∂_x (i.e. les opérateurs différentiels). En particulier, $(L^{3/2})_+ = P = \partial_x^3 + \frac{3}{2}u\partial_x + \frac{3}{4}u_x$, voir l'équation (1.6). Cette formulation de Lax scalaire de la hiérarchie KdV se généralise aisément à la n -ième hiérarchie de Gelfand–Dickey où l'opérateur $L = \partial_x^2 + u$ est remplacé par l'opérateur $L = \partial_x^n + u^{n-2}\partial_x^{n-2} + \cdots + u^0$.

Ainsi, les flots définis par le système (1.20) commutent deux-à-deux dans le sens où, pour tout entiers positifs i, j ,

$$\frac{\partial}{\partial t_i} \left[(L^{j/2})_+, L \right] = \frac{\partial}{\partial t_j} \left[(L^{i/2})_+, L \right]. \quad (1.21)$$

C'est ce qu'on appelle la *condition d'intégrabilité* de la hiérarchie KdV. La preuve est classique et peut être trouvée par exemple dans [DJM00] ; nous donnons une preuve tout à fait similaire pour la hiérarchie KP au théorème 3.2.1. Gardant à l'esprit que $L = \partial_x^2 + u$, il s'en suit de l'équation (1.21) que $\partial_{t_i} \partial_{t_j} u = \partial_{t_j} \partial_{t_i} u$. C'est évidemment une condition nécessaire à l'existence de solutions au moins C^1 , sans parler de solutions lisses. Notons que la fonction u doit à présent dépendre d'une infinité de variables $\{t_i \mid i > 0 \text{ odd}\}$. Nous supposons en général que

$$u \in \mathbb{C}[[t_1, t_3, t_5, \dots]].$$

Comme nous l'avons dit, nous traiterons la paire de Lax matricielle à la section 1.3 avec les hiérarchies de Drinfeld–Sokolov. Traitons à présent la représentation bi-Hamiltonienne.

(C) La représentation bi-Hamiltonienne. Soit un opérateur pseudo-différentiel $A \in \mathcal{D}$, on note $\text{res}(A)$ son coefficient en ∂_x^{-1} . Alors la représentation bi-Hamiltonienne est donnée par les deux mêmes structures de Poisson locales de l'équation (1.12), et par les densités Hamiltoniennes données par les résidus des puissances fractionnaires de $L = \partial_x^2 + u$, à savoir,

$$\begin{aligned} u_{t_i} &= \{u, \bar{h}_{1,i}\}_1 = \partial_x \frac{\delta h_{1,i}}{\delta u}, & \bar{h}_{1,i} &= \frac{4}{i+2} \int \text{res } L^{\frac{i+2}{2}}, \\ u_{t_i} &= \{u, \bar{h}_{2,i}\}_2 = \left(\frac{1}{4} \partial_x^3 + u \partial_x + \frac{1}{2} u_x \right) \frac{\delta h_{2,i}}{\delta u}, & \bar{h}_{2,i} &= \frac{4}{i} \int \text{res } L^{\frac{i}{2}}, \end{aligned}$$

pour $i > 0$ impair. En particulier, en comparant avec les équations (1.14) et (1.15), $\bar{h}_{1,3} = \bar{h}_1$ et $\bar{h}_{2,3} = \bar{h}_2$. De plus, les structures de Poisson locales peuvent être reformulées via l'opérateur L comme suit. Pour toute fonctionnelle locale $\bar{f} \in \mathcal{F}$, on note $\frac{\delta \bar{f}}{\delta L} = \partial_x^{-1} \circ \frac{\delta \bar{f}}{\delta u} \in \mathcal{A}$,²⁴ et soient $X = \frac{\delta \bar{f}}{\delta L}$ et $Y = \frac{\delta \bar{g}}{\delta L}$, alors

$$\{\bar{f}, \bar{g}\}_1 = \int \text{res} [XLY - LXY], \quad \{\bar{f}, \bar{g}\}_2 = \int \text{res} [(LX)_+ LY - (XL)_+ YL]. \quad (1.22)$$

(D) La représentation bilinéaire. Finalement, pour la représentation bilinéaire, on doit introduire les *polynômes de Schur élémentaires*, notés $p_k(\mathbf{t})$ avec $k \in \mathbb{Z}_{>0}$, définis par la relation

$$\exp \left(\sum_{i \in \mathbb{Z}_{>0}} t_i z^i \right) = \sum_{k \in \mathbb{Z}_{>0}} p_k(\mathbf{t}) z^k. \quad (1.23)$$

par exemple, $p_1 = t_1$, $p_2 = \frac{1}{2} t_1^2 + t_2$, $p_3 = \frac{1}{6} t_1^3 + t_2 t_1 + t_3$. Notons à présent \tilde{p}_k le polynôme p_k où toutes les variables t_{2i} ont été mises à 0. Alors la hiérarchie KdV (1.20), une fois effectuée la substitution $u = -2\partial_x^2 \log \tau$, est équivalente aux deux suites d'équations bilinéaires suivantes,²⁵

$$\left\{ \left\{ \begin{aligned} (\tilde{p}_{k+1}([D]) - \frac{1}{2} D_x D_{t_k})(\tau, \tau) &= 0, \\ (D_x D_{t_k} - D_x \tilde{p}_k([D]))(\tau, \tau) &= 0, \end{aligned} \right. \middle| k > 0 \text{ odd} \right\}, \quad (1.24)$$

où $[D] := (D_x, \frac{1}{3} D_{t_3}, \frac{1}{5} D_{t_5}, \dots)$. Les équations ci-dessus peuvent être déduites des résultats de [DJKM83] (voir aussi [ACvM12], lemme 4.1). En particulier, pour $k = 3$, $\tilde{p}_{k+1}([D]) - \frac{1}{2} D_x D_{t_k} = \frac{1}{24} (D_x^4 - 4 D_x D_{t_3})$ et l'on retrouve la forme bilinéaire (1.18) de l'équation KdV.

1.2.3 Fonctions tau polynomiales de la hiérarchie KdV et les polynômes d'Adler–Moser

Ici nous décrivons le résultat obtenu dans [Duc17] sur une transformation explicite des polynômes d'Adler–Moser vers les fonctions tau polynomiales de la hiérarchie KdV ; c'est l'objet du chapitre

²⁴La notation $\frac{\delta \bar{f}}{\delta L}$ peut sembler artificielle à ce stade, mais lorsque l'on considère les hiérarchies de Drinfeld–Sokolov pour une algèbre de Lie simple arbitraire (KdV étant associée à $\mathfrak{sl}(2, \mathbb{C})$), cela apparaît clairement.

²⁵Une forme bien plus compacte des équations (1.24) est donnée par la célèbre identité bilinéaire (cf. §3.3), cependant elle est moins explicite que celle donnée ici.

3, §3.5. Dans [AM78], Adler et Moser introduisirent une suite de polynômes θ_n , pour $n \geq 0$, dépendant de paramètres $\mathbf{r} = \{r_i \mid i \in \mathbb{Z}_{>0}^{\text{odd}}\}$, $r_1 = x$, définis par la récurrence différentielle suivante : $\theta_0 = 1$; $\theta_1 = x$; puis

$$\frac{\partial \theta_{n+1}}{\partial x} \theta_{n-1} - \theta_{n+1} \frac{\partial \theta_{n-1}}{\partial x} = (2n-1)(\theta_n)^2. \quad (1.25)$$

Lors du calcul de θ_n , la constante d'intégration choisie est r_{2n-1} . Par exemple, $\theta_2 = x^3 + r_3$ et $\theta_3 = x^6 + 5r_3x^3 + r_5x - 5r_3^2$. Ces polynômes sont désormais connus sous le nom de polynômes d'Adler–Moser. Dans ce même article, ils prouvèrent l'existence d'un unique changement de variables $\mathbf{r} \mapsto \mathbf{t}$ transformant le polynôme d'Adler–Moser $\theta_n(\mathbf{r})$ en une fonction tau $\tau_n(\mathbf{t})$ de la hiérarchie KdV dans le sens de l'équation (1.16). De plus, d'après des résultats de [AMcKM77], ils prouvèrent que *via* ce changement de variables, l'on retrouve *toutes* les fonctions tau polynomiales ainsi que *toutes* les solutions rationnelles de KdV, classifiant ainsi les solutions rationnelles. Cependant, cet unique changement de variables n'était pas connu. Dans [Duc17], nous avons prouvé le théorème suivant.

Théorème 1.2.1 ([Duc17]). *Le changement de variables suivant transforme les polynômes d'Adler–Moser en les fonctions tau polynomiales de la hiérarchie KdV : $r_1 = t_1 = x$, et*

$$\sum_{i \geq 2} \frac{r_{2i-1}}{\alpha_{2i-1}} z^{2i-1} = \tanh \left(\sum_{i \geq 2} t_{2i-1} z^{2i-1} \right). \quad (1.26)$$

où $\alpha_{2i-1} = (-1)^{i-1} 3^{2^2} \dots (2i-3)^2 (2i-1)$.

Voir théorème 3.5.4 dans le document. On peut interpréter le théorème 1.2.1 comme une renormalisation des variables, plus le choix différent, non trivial, d'une constante d'intégration pour la récurrence (1.25). Les premiers polynômes de la récurrence sont les suivants.

$$\begin{aligned} \tau_0 &= 1, & \tau_2 &= x^3 - 3t_3, \\ \tau_1 &= x, & \tau_3 &= x^6 - 15t_3x^3 - 45t_3^2 + 45t_5x. \end{aligned}$$

Ce sont des fonctions tau polynomiales bien connues, voir e.g. [Hir04] où elles sont calculées par des méthodes similaires *via* les polynômes de Schur élémentaires (voir équation (1.23)).

La preuve du théorème 1.2.1 est donnée à la section 3.5 ; elle repose essentiellement sur la théorie KP que nous développons à la section 3.2. Cela constituera un futur projet que de généraliser les polynômes d'Adler–Moser ainsi que les résultats de [AM78] et le théorème ci-dessus aux hiérarchies de Drinfeld–Sokolov en général (i.e. pour une algèbre de Lie simple arbitraire).

1.3 Fonctions tau polynomiales des hiérarchies de Drinfeld–Sokolov

Ici nous décrivons les résultats obtenus dans [CDY18] sur les fonctions tau polynomiales des hiérarchies de Drinfeld–Sokolov et des applications ; c'est l'objet du chapitre 4, §4.5. Nous

commençons par décrire la construction des hiérarchies de Drinfeld–Sokolov dans l’approche la plus directe, bien que technique : la paire de Lax matricielle. Dans cette approche, elles généralisent largement la représentation (1.10) de l’équation KdV. Mais tout d’abord, il nous faut introduire du matériel concernant les algèbres de Lie semisimples et affines.

1.3.1 Les hiérarchies de Drinfeld–Sokolov

Point de départ : une algèbre de Lie semisimple. Nous démarrons avec une algèbre de Lie finie-dimensionnelle semisimple (voir §4.1), i.e. une algèbre de Lie qui est une somme directe d’algèbres de Lie simple (i.e. qui ne contiennent pas d’idéaux non triviaux). Ces algèbres se décomposent ainsi,

$$\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+,$$

où \mathfrak{h} est une sous-algèbre abélienne maximale dont les éléments $H \in \mathfrak{h}$ sont tels que les applications adjointes $\text{ad}_H = [H, -] : \mathfrak{g} \rightarrow \mathfrak{g}$ sont simultanément diagonalisables. La sous-algèbre \mathfrak{h} est appelée une *sous-algèbre de Cartan* de \mathfrak{g} et sa dimension $r = \dim \mathfrak{h}$ est appelée le *rang* de \mathfrak{g} ; le rang est unique. Un résultat classique est que \mathfrak{g} est complètement déterminée (à isomorphisme près) par une unique matrice $(C_{ij}) \in \text{Mat}(r, \mathbb{Z})$, appelée sa *matrice de Cartan*, à laquelle on peut associer un ensemble de générateurs $\{H_i, E_i, F_i \mid 1 \leq i \leq r\}$ de \mathfrak{g} satisfaisant les relations suivantes :

$$[E_i, F_j] = \delta_{ij} H_i, \quad [H_i, E_j] = C_{ij} E_i, \quad [H_i, F_j] = -C_{ij} F_j, \quad [H_i, H_j] = 0,$$

Alors \mathfrak{h} est générée par les H_i , tandis que \mathfrak{n}_+ (resp. \mathfrak{n}_-) est générée par les E_i ’s (resp. the F_i ’s). De tels générateurs s’appellent des *générateurs de Weyl*. On a deux éléments spéciaux $E_\theta \in \mathfrak{n}_+$ et $E_{-\theta} \in \mathfrak{n}_-$ qui satisfont les relations suivantes : pour tout $i \in \{1, \dots, r\}$,

$$[E_\theta, E_i] = 0, \quad [E_{-\theta}, F_i] = 0.$$

Par exemple, l’algèbre de Lie simple $\mathfrak{sl}(2, \mathbb{C})$ (matrices complexes 2×2 de trace nulle) est de rang 1 et est générée par les matrices suivantes,

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Ces matrices satisfont les relations $[H, E] = 2E$, $[H, F] = -2F$ et $[E, F] = H$ de sorte que la matrice de Cartan de $\mathfrak{sl}(2, \mathbb{C})$ est $C = (2) \in \text{Mat}(1, \mathbb{Z})$. De plus, dans ce cas particulier, $E_\theta = E$ et $E_{-\theta} = F$.

L’algèbre affine associée. À l’algèbre finie-dimensionnelle semisimple \mathfrak{g} on associe une algèbre affine (non twistée)

$$\mathfrak{g} := \mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c, \quad (1.27)$$

qui est infinie-dimensionnelle. On équipe cette algèbre du crochet de Lie suivante : soient $X, Y \in \mathfrak{g}, p, q \in \mathbb{Z}$ et $a, b \in \mathbb{C}$,

$$[X\lambda^p + ac, Y\lambda^q + bc] = [X, Y]\lambda^{p+q} + p\delta_{p+q,0}(X, Y)_0 \cdot c, \quad (1.28)$$

où $(-, -)_0$ est une certaine forme bilinéaire symétrique non dégénérée sur \mathfrak{g} . Clairement, c est un élément central de \mathfrak{g} (i.e. c commute avec tous les autres éléments de \mathfrak{g}), faisant de $\mathbb{C}c$ un idéal non trivial de \mathfrak{g} . Si l'algèbre de Lie semisimple \mathfrak{g} est de type X_n , on note $X_n^{(1)}$ l'algèbre affine associée \mathfrak{g} . Les algèbres affines furent intensément étudiées par V. Kac, qui les introduisit parallèlement à R. Moody [Kac68, Moo68]. Les détails sont donnés dans §4.2.

L'algèbre affine \mathfrak{g} est générée par les éléments suivants :

$$\begin{aligned} e_i &= E_i, & f_i &= F_i, & \check{\alpha}_i &= H_i, \\ e_0 &= E_{-\theta}\lambda, & f_0 &= E_\theta\lambda^{-1}, & \check{\alpha}_0 &= H_0 + c, \end{aligned}$$

où $H_0 = [E_\theta, E_{-\theta}]$, avec $i \in \{1, \dots, r\}$ ($r = \dim \mathfrak{h}$ étant le rang de \mathfrak{g}). Puis on définit deux gradations importantes : la *gradation principale* et la *gradation homogène*, comme suit,

$$\begin{aligned} \text{Principal:} & \quad \deg e_i = 1, & \deg f_i &= -1, & \deg \check{\alpha}_i &= 0, & \mathfrak{g} &= \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}^k, \\ \text{Homogeneous:} & \quad \deg e_i = \delta_{i,0}, & \deg f_i &= -\delta_{i,0}, & \deg \check{\alpha}_i &= 0, & \mathfrak{g} &= \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_k. \end{aligned}$$

La gradation homogène, comme son nom l'indique, revient à poser $\deg \lambda = 1$ et $\deg \mathfrak{g} = \deg c = 0$. On note \mathfrak{g}^+ (resp. \mathfrak{g}_+) le sous-espace de des éléments de degré principal (resp. homogène) ≥ 0 ; on pose $\mathfrak{g}^- = \mathfrak{g} - \mathfrak{g}^+$ et $\mathfrak{g}_- = \mathfrak{g} - \mathfrak{g}_+$.

On introduit à présent l'*élément cyclique principal affine*

$$\Lambda = e_0 + e_1 + \dots + e_r = E_{-\theta}\lambda + E_1 + \dots + E_r. \quad (1.29)$$

En particulier, $\Lambda \in \mathfrak{g}^1$ (degré principal égal à 1). Un résultat fondamental (voir [Kac90]) sur les algèbres affines est la décomposition de \mathfrak{g} comme

$$\mathfrak{g} = \mathfrak{s} + \text{Im}(\text{ad}_\Lambda), \quad \mathfrak{s} \cap \text{Im}(\text{ad}_\Lambda) = \mathbb{C}c.$$

où $\mathfrak{s} = \text{Ker}(\text{ad}_\Lambda)$. On note $\mathfrak{s}^k = \mathfrak{s} \cap \mathfrak{g}^k$ le sous-espace de degré principal k . On appelle *exposants affines* de \mathfrak{g} les entiers $k \in \mathbb{Z}$ tels que $\dim \mathfrak{s}^k \neq 0$ et on note $m_k = \dim \mathfrak{s}^k$ la multiplicité de k . On note

$$\mathbb{E} = \left\{ k \in \mathbb{Z} \mid \dim \mathfrak{s}^k \neq 0 \right\}$$

le multi-ensemble des exposants affines de \mathfrak{g} , c'est-à-dire que l'élément $k \in \mathbb{E}$ a multiplicité m_k .

On note $\Lambda_k, \Lambda'_k, \Lambda''_k, \dots$ (m_k fois) les éléments engendrant \mathfrak{s}^k linéairement. Ces générateurs Λ_k (on omet les primes pour l'instant) peuvent être choisis de sorte qu'ils satisfassent les relations de Heisenberg

$$[\Lambda_i, \Lambda_j] = i\delta_{i+j,0} \cdot c. \quad (1.30)$$

On appelle \mathfrak{s} la sous-algèbre de Heisenberg principale de \mathfrak{g} . On note $\nu \in \mathbb{C} \setminus \{0\}$ le nombre complexe tel que $\Lambda_1 = \nu\Lambda$.

Par exemple, pour l'algèbre affine $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C}) \otimes \mathbb{C}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c$ l'élément cyclique principal affine est

$$\Lambda = F\lambda + E = \begin{pmatrix} 0 & \lambda \\ 1 & 0 \end{pmatrix}.$$

Les exposants affines de \mathfrak{g} sont exactement les nombres impairs et ont tous multiplicité 1, i.e. $\mathbb{E} = \mathbb{Z}^{\text{odd}}$. Les générateurs de la sous-algèbre de Heisenberg principale sont simplement les $\Lambda_{2i+1} = \Lambda^{2i+1} = \lambda^i \Lambda$ pour $i \in \mathbb{Z}$.

Les hiérarchies de Drinfeld–Sokolov. Nous avons à présent tout le matériel nécessaire à la description des hiérarchies de Drinfeld–Sokolov. Premièrement, on note $\mathbb{E}^+ = \mathbb{E} \cap \mathbb{Z}_{>0}$ l'ensemble des exposants affines positifs de \mathfrak{g} . On note alors $\mathbf{t} = \{t_i \mid i \in \mathbb{E}^+\}$, avec $t_1 = \nu^{-1}x$, et l'on considère l'algèbre suivante,

$$\mathfrak{g} \otimes \mathbb{C}[[\mathbf{t}]] = \left\{ \sum_{m \geq 0} \sum_{i_1, \dots, i_m \in \mathbb{E}^+} \sum_{k_1, \dots, k_m \geq 1} \xi_{i_1, \dots, i_m}^{k_1, \dots, k_m} \cdot t_{i_1}^{k_1} \cdots t_{i_m}^{k_m} \mid \xi_{i_1, \dots, i_m}^{k_1, \dots, k_m} \in \mathfrak{g} \right\}.$$

L'algèbre ci-dessus (et ses sous-algèbres) jouera le rôle d'« espace de fonctions » de la théorie. Le crochet de Lie est défini de façon similaire à celui sur $\hat{\mathfrak{g}} \otimes \mathbb{C}[\lambda, \lambda^{-1}]$, i.e. on pose $[\xi t_i, \eta t_j] = [\xi, \eta] t_i t_j$ puis on étend par multiplicativité. On considère à présent l'opérateur matricielle d'ordre 1 suivant,

$$\mathcal{L} = \partial_x + \Lambda + Q, \quad Q \in \hat{\mathfrak{b}}[[\mathbf{t}]],$$

où $\hat{\mathfrak{b}} = \hat{\mathfrak{h}} \oplus \hat{\mathfrak{n}}_+$ (typiquement, pour $\mathfrak{sl}(n, \mathbb{C})$, la sous-algèbre $\hat{\mathfrak{b}}$ consiste en les matrices triangulaires inférieures de trace nulle). Nous appelons de telles opérateurs des *connexions* (bien qu'il n'y ait pas de variété ni de fibré pour ainsi dire) et nous notons $\text{op}_{\mathfrak{g}}$ l'ensemble de ces connexions.²⁶ On considère ensuite le groupe multiplicatif suivant :

$$N := \{ \exp(\text{ad}_S) \mid S \in \hat{\mathfrak{n}}[[\mathbf{t}]] \}$$

et le faisons agir sur $\text{op}_{\mathfrak{g}}$ par

$$e^{\text{ad}_S} \mathcal{L} = \partial_x + e^{\text{ad}_S}(\Lambda + Q) - \nabla_S S, \quad \nabla_S S := \sum_{k \geq 0} \frac{1}{(k+1)!} (\text{ad}_S)^k \partial_x(S).$$

Les raisons pour définir cette action et le fait qu'elle est bien définie sont expliqués à la section §4.3.1. Alors il n'est pas difficile de montrer, qu'en effet, pour tout $\mathcal{L} \in \text{op}_{\mathfrak{g}}$, on a $e^{\text{ad}_S} \mathcal{L} \in \text{op}_{\mathfrak{g}}$. Les hiérarchies de Drinfeld–Sokolov sont définies sur les classes d'équivalences de ces

²⁶L'équation (1.8) donne un exemple de ces connexions \mathcal{L} avec $\hat{\mathfrak{g}} = \mathfrak{sl}(2, \mathbb{C})$, bien que nous ayons même choisi $Q \in \hat{\mathfrak{n}}_+[[\mathbf{t}]]$ puisque pour $\mathfrak{sl}(2, \mathbb{C})$, on a $\hat{\mathfrak{n}}_+ = \mathbb{C}E$.

connexions $\mathcal{L} \in \text{op}_{\mathfrak{g}}$, plutôt que sur les connexions elles-mêmes. À présent nous devons étendre notre espace de fonctions de la façon suivante :

$$\widehat{\mathfrak{g}} := \mathfrak{g}^+ \oplus \widehat{\mathfrak{g}}^-, \quad \mathfrak{g}^+ = \mathfrak{g}^{\geq 0} = \bigoplus_{k \geq 0} \mathfrak{g}^k, \quad \widehat{\mathfrak{g}}^- = \prod_{k < 0} \mathfrak{g}^k.$$

De la même façon, on définit naturellement $\widehat{\mathfrak{g}}[[\mathbf{t}]]$ et ses sous-algèbres. De telles définitions sont implicites pour \mathfrak{g}_+ et $\widehat{\mathfrak{g}}_-$.

Un résultat fondamental qui fut prouvé pour la première fois par Drinfeld et Sokolov [DS84] (ici nous donnons une version un peu différente due à Wu [Wu17]) est l'existence d'une unique fonction $U \in \widehat{\mathfrak{g}}^-[[\mathbf{t}]]$ telle que la connexion $\mathcal{L}_0 := e^{-\text{ad}_U} \mathcal{L}$ prenne la forme

$$\mathcal{L}_0 = \partial_x + \Lambda + H, \quad H \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]], \quad (1.31)$$

et telle que $(e^{\text{ad}_U} \xi)_c = 0$ pour tout $\xi \in \mathfrak{s}^+$, où $(-)_c$ désigne la composante le long de $\mathbb{C}c$ dans la décomposition (1.27) de \mathfrak{g} . Les composantes des fonctions U et H sont calculables récursivement. Le point clé ici est que, grâce aux relations de Heisenberg (1.30), l'algèbre $\widehat{\mathfrak{s}}^-[[\mathbf{t}]]$ est abélienne ! On peut à présent formuler la paire de Lax matricielle des hiérarchies de Drinfeld–Sokolov. Soit une connexion $\mathcal{L} \in \text{op}_{\mathfrak{g}}$, la hiérarchie de Drinfeld–Sokolov de type \mathfrak{g} (ou la hiérarchie \mathfrak{g} -DS) consiste en le système d'équations suivant :

$$\left\{ \frac{\partial \mathcal{L}}{\partial t_i} = [(e^{\text{ad}_U} \Lambda_i)_+, \mathcal{L}] \mid i \in \mathbb{E}^+ \right\}. \quad (1.32)$$

Comme suggéré précédemment, si \mathcal{L} est solution de la hiérarchie, alors il en est de même pour $e^{\text{ad}_S} \mathcal{L}$ pour tout $S \in \mathfrak{n}[[\mathbf{t}]]$. Si la multiplicité de l'exposant $i \in \mathbb{E}^+$ est plus grande que 1, alors on note t_i, t'_i, t''_i, \dots les variables associées. C'est essentiellement de l'abélianité de $\widehat{\mathfrak{s}}^-[[\mathbf{t}]]$ que résulte la commutativité des flots de la hiérarchie, et donc son intégrabilité. Le membre de droite de ces équations sont des matrices dont les composantes sont des polynômes différentiels en les composantes de la matrice Q . Finalement, on montre aisément que les composantes de la fonction H sont des quantités conservées dans le sens où, pour tout $j \in \mathbb{E}^+$,

$$\int \frac{\partial H}{\partial t_j} = 0.$$

Par exemple, pour la hiérarchie KdV (1.20) est équivalente au système (1.32), avec les données suivantes :

$$\mathcal{L} = \partial_x + \Lambda + Q, \quad \Lambda = \begin{pmatrix} 0 & \lambda \\ 1 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & -u \\ 0 & 0 \end{pmatrix}, \quad \Lambda_{2i+1} = \lambda^i \Lambda = \begin{pmatrix} 0 & \lambda^{i+1} \\ \lambda^i & 0 \end{pmatrix}.$$

1.3.2 Fonctions tau comme déterminants de Toeplitz

Ici nous expliquons la construction d'une fonction tau générique pour toute hiérarchie de Drinfeld–Sokolov par C.-Z. Wu [Wu17]. Puis nous expliquons comment réaliser ces fonctions tau comme la limite de déterminants de Toeplitz par Cafasso et Wu [CW15a, CW15b]. Finalement, nous décrivons les résultats obtenus dans [CDY18] sur le calcul et les applications des fonctions tau polynomiales à partir des travaux de Cafasso et Wu.

Fonctions tau des hiérarchies de Drinfeld–Sokolov. Dans [Wu17], Wu définit une fonction tau générique pour n'importe quelle hiérarchie de Drinfeld–Sokolov. Tout d'abord nous introduisons une fonction $\Omega = \sum_{i \in \mathbb{E}^+} \frac{\Omega^i}{i} \Lambda_{-i}$ définie par $\partial_x \Omega = -H$, où $H \in \widehat{\mathfrak{s}^-}[[\mathbf{t}]]$ est définie à l'équation (1.31). Ainsi les fonctions Ω^j ne sont pas des polynômes différentiels en les composantes de Q , cependant les fonctions $\partial_{t_i} \Omega^j$ le sont et satisfont la relation de fermeture $\partial_{t_i} \Omega^j = \partial_{t_j} \Omega^i$. En posant $H = \sum_{i \in \mathbb{E}^+} -i^{-1} \tilde{h}^i \Lambda_{-i}$, on obtient

$$\frac{\partial \tilde{h}^i}{\partial t_j} = \frac{\partial \tilde{h}^j}{\partial t_i}. \quad (1.33)$$

L'équation ci-dessus est appelée la condition de *symétrie tau* pour les densités conservées \tilde{h}^j . Par le lemme de Poincaré, il s'en suit que l'on peut définir une fonction $\tau \in \mathbb{C}[[\mathbf{t}]]$ telle que²⁷

$$\frac{\partial \log \tau}{\partial x \partial t_j} = \tilde{h}^j. \quad (1.34)$$

On appelle τ la *fonction tau*. La fonction tau de Wu est générique dans le sens où sa construction marche pour toute algèbre affine (les définitions précédentes étaient quelque peu au cas par cas). Maintenant en considérant la série formelle $\Theta = e^U e^\Omega$ (nous montrons qu'elle est bien définie), on peut réexprimer la fonction tau par la relation

$$\frac{\partial \log \tau}{\partial t_j} = (\Theta \Lambda_j \Theta^{-1})_c \quad (1.35)$$

Fonctions tau comme déterminants de Toeplitz. Dans [CW15a, CW15b], Cafasso et Wu montrèrent comment exprimer la fonction tau τ associée à la fonction Θ comme limite de *déterminants de Toeplitz*. Tout d'abord, nous fixons une représentation fidèle de trace nulle²⁸

$$\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(n)$$

comme à l'appendice A. Les matrices de Toeplitz sont des matrices de taille « $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ » constantes le long des diagonales, i.e. de la forme

$$\mathbf{T}(\gamma) = (\gamma_{I-J})_{I, J \geq 0} = \begin{pmatrix} \gamma_0 & \gamma_{-1} & \gamma_{-2} & & \\ \gamma_1 & \gamma_0 & \gamma_{-1} & \ddots & \\ \gamma_2 & \gamma_1 & \gamma_0 & \ddots & \\ & \ddots & \ddots & \ddots & \end{pmatrix},$$

où les γ_I sont certains coefficients que, dans notre cas, nous choisissons dans $\mathfrak{gl}(n)[[\mathbf{t}]]$. On rassemble ces éléments dans une série formelle $\gamma = \sum_{I \in \mathbb{Z}} \gamma_I \lambda^I \in \mathfrak{gl}(n)[[\lambda, \lambda^{-1}, \mathbf{t}]]$ (même λ que

²⁷L'équation (1.34) est équivalente au fait de poser $\partial_{t_j} \log \tau = -\Omega^j$. Nous insistons sur le fait que puisque les fonctions Ω^j ne sont pas des polynômes différentiels, cela n'implique pas que les \tilde{h}^j soient des dérivées exactes de polynômes différentiels. En particulier, l'équation (1.34) n'implique pas que $\int \tilde{h}^j = 0$.

²⁸Par $\mathfrak{gl}(n)$ on entend $\text{Mat}(n, \mathbb{C})$ (matrices $n \times n$ complexes) équipée du crochet de Lie définie par le commutateur.

précédemment). Bien entendu, ce n'est pas clair que le déterminant de $\mathbf{T}(\gamma)$ peut être défini et c'est un aspect technique de cette approche. Pour ce faire, on considère la matrice de Toeplitz tronquée $\mathbf{T}_N(\gamma) = (\gamma_{I-J})$ avec $I, J \in \{1, \dots, N\}$.

À la section 4.4 nous expliquons le résultat suivant de [CW15a, CW15b]. Considérons un élément quelconque $X \in \mathfrak{g}_-$ et la fonction suivante que nous appelons la *fonction Xi matricielle*,

$$\Xi(\mathbf{t}; \lambda) := \sum_{j \in \mathbb{E}^+} t_j \Lambda_j \in \mathfrak{s}^+ [[\mathbf{t}]].$$

Alors (1) la fonction $J_X(\mathbf{t}) = e^{\Xi(\mathbf{t}; \lambda)} e^{-X}$ est une série formelle bien définie de $\mathfrak{gl}(n)[[\lambda, \lambda^{-1}, \mathbf{t}]]$ ²⁹ ; (2) la limite $\lim_{N \rightarrow \infty} \det \mathbf{T}_N(J_X(\mathbf{t}))$ est une série formelle bien définie de $\mathbb{C}[[\mathbf{t}]]$ et (3) la fonction

$$\tau_X(\mathbf{t}) := \left[\lim_{N \rightarrow \infty} \det \mathbf{T}_N(J_X(\mathbf{t})) \right]^\kappa \quad (1.36)$$

est une fonction tau de la hiérarchie g-DS au sens de l'équation (1.34). Ici κ est un nombre associé à la représentation $\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(n)$ par $(Y, Z)_0 = \kappa \cdot \text{tr}(\pi(Y)\pi(Z))$ pour tous $Y, Z \in \mathfrak{g}$ (avec $(-, -)_0$ la forme bilinéaire standard sur \mathfrak{g}). Typiquement, pour les algèbres de Lie de type A_n et C_n , on a $\kappa = 1$ et pour les types B_n et D_n on a $\kappa = 1/2$.

Plus précisément, la fonction Θ de l'équation (1.34) à laquelle la fonction tau $\tau_X(\mathbf{t})$ est associée est celle telle que $\Theta|_{\mathbf{t}=0} = e^{-X}$. Il est facile de montrer que pour tout $X \in \mathfrak{g}_-$, il existe une telle fonction Θ ; voir le lemme 4.4.2.

1.3.3 Fonctions tau polynomiales et applications

Nous expliquons à présent les résultats de [CDY18] sur les fonctions tau de Drinfeld–Sokolov et applications ; voir chapitre 4, §4.5. On choisit un élément $X \in \mathfrak{g}_-$ et la fonction tau associée $\tau_X(\mathbf{t})$ de l'équation (1.36). Notre premier résultat est une expansion en série entière de $\tau_X(\mathbf{t})$ sur les partitions d'entiers (ou, de façon équivalente, sur les diagrammes de Young) ; on note \mathbb{Y} l'ensemble des partitions d'entiers. Pour ce faire, il nous faut introduire la *matrice de Laurent* $\mathbf{L}(\gamma)$, qui est aussi constante le long des diagonales, mais de taille « $\mathbb{Z} \times \mathbb{Z}$, » à savoir

$$\mathbf{L}(\gamma) = (\gamma_{I-J})_{I, J \in \mathbb{Z}} = \begin{pmatrix} \ddots & \ddots & \ddots & & & & \\ \ddots & \gamma_0 & \gamma_{-1} & \gamma_{-2} & & & \\ \ddots & \gamma_1 & \gamma_0 & \gamma_{-1} & \ddots & & \\ & \gamma_2 & \gamma_1 & \gamma_0 & \ddots & & \\ & & \ddots & \ddots & \ddots & & \end{pmatrix}.$$

²⁹Les problèmes de convergences sont traités de façon purement algébrique *via* la topologie induite par la valuation. C'est expliqué au début de la section 4.4

On considère donc les matrices de Laurent $r := \mathbf{L}(e^X)$ et $s(\mathbf{t}) := \mathbf{L}(e^{\Xi(\mathbf{t};\lambda)})$ ainsi que les déterminants suivants,³⁰

$$r_\mu := \det(r_{i-\mu_i-1; j-1})_{i,j \in [1, \ell(\mu)]} \in \mathbb{C}, \quad s_\mu(\mathbf{t}) := \det(s_{i-1; j-\mu_j-1})_{i,j \in [1, \ell(\mu)]} \in \mathbb{C}[\mathbf{t}], \quad (1.37)$$

où $\mu \in \mathbb{Y}$ est n'importe quelle partition de longueur $\ell(\mu)$. On appelle r_μ les *coordonnées de Plücker* de e^X et on appelle $s_\mu(\mathbf{t})$ les *polynômes de Schur généralisés* de type (\mathfrak{g}, π) . Dans le cas où $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$ et π est la représentation fondamentale, les polynômes $s_\mu(\mathbf{t})$ coïncident avec les polynômes de Schur classiques [Mac95], une fois que l'on a posé $t_k = 0$ pour tout $k = 0 \pmod n$. Nous avons le prouvé le théorème suivant. C'est une application de la formule de Cauchy-Binet et une généralisation des travaux fondateurs de Sato sur les fonctions tau de la hiérarchie KP [Sat81].

Théorème 1.3.1 ([CDY18]). *Soit $X \in \mathfrak{g}_-$, la fonction tau de l'équation (1.36) admet l'expansion suivante en terme de partitions d'entiers :*

$$\tau_X(\mathbf{t})^{\frac{1}{\kappa}} = \sum_{\mu \in \mathbb{Y}} s_\mu(\mathbf{t}) \cdot r_\mu = 1 + s_{\square} r_{\square} + s_{\square\square} r_{\square\square} + s_{\square\square\square} r_{\square\square\square} + s_{\square\square\square\square} r_{\square\square\square\square} + \dots \quad (1.38)$$

Voir théorème 4.5.2 dans le document. À partir de l'équation (1.38) on trouve un critère facile pour la polynomialité de la fonction tau comme suit. Si la donnée initiale $X \in \mathfrak{g}_-$ est telle que $\pi(X)$ est une matrice nilpotente, alors les composantes de la matrice de Laurent $r = \mathbf{L}(e^X)$ s'annulent lorsque l'on s'éloigne suffisamment de la diagonale principale. Conséquemment, pour une partition μ suffisamment longue, les coordonnées de Plücker r_μ de X finissent par s'annuler. Cela prouve l'assertion suivante,

$$\forall X \in \mathfrak{g}_-, \pi(X) \text{ est une matrice nilpotente} \implies \tau_X(\mathbf{t})^{\frac{1}{\kappa}} \text{ est polynomiale.}$$

À la section §4.5.2 nous donnons plusieurs exemples de fonctions tau polynomiales pour les hiérarchies de Drinfeld–Sokolov associées aux algèbres de Lie simples de type A_1, A_2, B_2 et D_4 . Pour le cas A_1 – i.e. $\mathfrak{sl}(2, \mathbb{C})$, i.e. la hiérarchie KdV – toutes les fonctions tau que nous avons calculées étaient, à un changement de variable près, les polynômes d'Adler–Moser que nous avons discutés à la section 1.2.3 de cette introduction. Bien que ça n'ait pas été encore prouvé, nous pensons que notre algorithme recouvre *tous* les polynômes d'Adler–Moser. Cela pourrait être prouvé en comparant les degrés des polynômes. Pour les autres cas A_2, B_2, D_4 , les fonctions tau polynomiales deviennent rapidement non triviales.

Cela constituera un futur projet que d'étudier les propriétés de ces polynômes. Notamment, en les voyant comme des polynômes en x avec paramètres $\mathbf{t}^* = \{t_i \mid i \neq 1\}$, l'on devrait étudier comment les zéros de ces polynômes évoluent par rapport à \mathbf{t}^* . Comme noté à la section §1.1, dans le cas A_1 , les zéros des polynômes d'Adler–Moser évoluent en formant un système de type Calogero–Moser [AMcKM77, Wil98]. C'est encore à comprendre pour les hiérarchies de Drinfeld–Sokolov en général.

³⁰Il est important de noter que, dans l'équation (1.37), les indices i, j représentent les vraies composantes complexes des matrices de Laurent r et $s(\mathbf{t})$. Ils ne représentent pas les indices de block qui étaient noté avec des indices capitaux I, J jusque là.

Fonctions tau polynomiales et équations bilinéaire. Les fonctions polynomiales ont le mérite d'être des expressions finies et sont ainsi bien plus simples à manipuler dans les calculs pratiques. Comme application, nous pouvons détecter les équations bilinéaires (i.e. similaires à celles de l'équation (1.24)) que les fonctions tau de la hiérarchie doivent satisfaire. La forme bilinéaire des hiérarchies de Drinfeld–Sokolov est connue explicitement pour toutes les hiérarchies de Gelfand–Dickey (i.e. pour $\mathfrak{sl}(n, \mathbb{C})$), voir équation (4.97). Cependant, à notre connaissance, de telles équations bilinéaires explicites pour toute hiérarchie de Drinfeld–Sokolov ne sont pas données dans la littérature. Cependant, elles doivent être déductibles des résultats de la série de papier fondatrice « *Transformation groups for soliton equations* » [KM81]–[DJKM83].

Nous avons donc injecté les fonctions tau polynomiales que nous avons calculées dans des ansatz d'équations bilinéaires et avons requis que ces équations s'annulent. Pour des fonctions tau polynomiales de degré suffisamment haut, cela détermina les coefficients des dites équations bilinéaires. Bien entendu, cela ne signifie pas que toutes les fonctions tau de la hiérarchie doivent satisfaire ces équations. Cependant, nous pensons que cela donne de bonnes indications pour identifier ces équations.

Dans ce qui suit, $\text{HDS}_{X_n}^{[\text{deg} \leq g]}$ désigne l'espace vectoriel des équations bilinéaires de degré $\leq g$ que les fonctions de la hiérarchie DS de type X_n (A_n ou B_n , etc.) satisfont. Le degré des dérivées de Hirota (introduites à l'équation (1.17)) est défini par $\text{deg } D_i = i$ où $D_i = D_{t_i}$. Nous avons obtenu les propositions suivantes ; premièrement pour le cas B_2 .

Proposition 1.3.2. *On a les estimations de dimension suivantes,*

$$\dim_{\mathbb{C}} \text{HDS}_{B_2}^{[\text{deg} \leq 4]} = 0, \quad \dim_{\mathbb{C}} \text{HDS}_{B_2}^{[\text{deg} \leq 6]} \leq 1, \quad \dim_{\mathbb{C}} \text{HDS}_{B_2}^{[\text{deg} \leq 8]} \leq 2.$$

De plus, les éléments de $\text{HDS}_{B_2}^{[\text{deg} \leq 8]}$ sont combinaisons linéaires des équations suivantes,

$$\begin{aligned} (D_1^6 - 5D_1^3 D_3 - 5D_3^2 + 9D_1 D_5)(\tau, \tau) &= 0, \\ (D_1^8 + 7D_1^5 D_3 - 35D_1^2 D_3^2 - 21D_1^3 D_5 - 42D_3 D_5 + 90D_1 D_7)(\tau, \tau) &= 0. \end{aligned}$$

Les équations ci-dessus peuvent être trouvées dans [DKM81] comme étant les deux premières équations bilinéaires non triviales de la hiérarchie BKP. Cette dernière contient toutes les hiérarchies de type B_n et D_n (les types orthogonaux spéciaux). Dans la même façon, la hiérarchie KP contient toutes les hiérarchies de type A_n .

Nous avons une proposition similaire pour la hiérarchie de type D_4 . Notons que pour le cas D_4 , on a deux variables associées à l'exposant 3 : t_3 et $t_{3'}$. C'est dû au fait que la sous-algèbre de Heisenberg \mathfrak{s} de type D_4 possède un sous-espace dimension 2 pour tout exposant dans $3\mathbb{Z}$.

Proposition 1.3.3. *On a les estimations de dimension suivantes,*

$$\dim_{\mathbb{C}} \text{HDS}_{D_4}^{[\text{deg} \leq 4]} = 0, \quad \dim_{\mathbb{C}} \text{HDS}_{D_4}^{[\text{deg} \leq 6]} \leq 3.$$

De plus, les éléments de $\text{HDS}_{D_4}^{[\text{deg} \leq 6]}$ sont combinaisons linéaires des équations suivantes,

$$\begin{aligned} (2D_1^3 D_{3'} + 4D_3 D_{3'} - 3D_{3'}^2)(\tau, \tau) &= 0, \\ (D_1^3 D_3 - D_1^3 D_{3'} + D_3 D_{3'} - D_3^2)(\tau, \tau) &= 0, \\ (D_1^6 + 9D_1 D_5 - 10D_1^3 D_3 + 5D_1^3 D_{3'} - 5D_3 D_{3'})(\tau, \tau) &= 0. \end{aligned}$$

1.4 Hiérarchie de type D_4 et l'équivalence DR/DZ forte

Finalement, nous décrivons les résultats concernant la confirmation de l'équivalence Double Ramification / Dubrovin–Zhang forte (DR/DZ) ; un travail accepté pour publication [DR19]. C'est l'objet du chapitre 5. Tout d'abord, nous expliquons brièvement le concept de théorie cohomologique des champs (CohFT, pour *cohomological field theory*), l'exemple de la théorie de Fan–Jarvis–Ruan–Witten (FJRW) et la construction de Dubrovin–Zhang.

1.4.1 Théories cohomologiques des champs et les hiérarchies de Dubrovin–Zhang

Espaces de module des courbes stables. Essentiellement, l'espace de module des courbes stables $\overline{\mathcal{M}}_{g,n}$ est l'espace qui paramétrise les classes d'isomorphisme de courbes complexes C , compactes, possiblement nodales, de genre g , avec la donnée de n points $x_1, \dots, x_n \in C$ appelés des *marques*. Les marques ne peuvent être des nœuds et doivent être distinctes. Cela permet de dire que deux courbes marquées $(C, (x_1, \dots, x_n))$ et $(C', (x'_1, \dots, x'_n))$ sont isomorphes s'il existe un biholomorphisme $\phi : C \rightarrow C'$ tel que $\phi(x_i) = x'_i$. On dit qu'une telle courbe est stable si son groupe d'automorphisme est fini. L'hypothèse de stabilité est nécessaire afin d'obtenir un espace de module bien défini ; elle donne la condition suivante,

$$2g - 3 + n \geq 0.$$

On définit alors l'espace de module des courbes stables de genre g avec n marques comme l'espace quotient suivant,³¹

$$\overline{\mathcal{M}}_{g,n} := \{(C, (x_1, \dots, x_n))\} / \sim$$

C'est une orbifold complexe compact, i.e. grossièrement, un espace topologique qui est localement isomorphe à un espace affine Euclidien quotienté par l'action d'un groupe fini. En tant qu'orbifold, elle a la dimension complexe suivante,

$$\dim_{\mathbb{C}} \overline{\mathcal{M}}_{g,n} = 3g - 3 + n.$$

Cette orbifold vient avec plusieurs applications et fibrés importants. Les détails peuvent être trouvés à la section §5.2.1 ; nous référons à [Zvo12]. Tout d'abord, le groupe symétrique S^n agit sur $\overline{\mathcal{M}}_{g,n}$ en permutant les points marqués. Puis on a trois applications naturelles :

- ◊ L'application d'oubli, $\pi : \overline{\mathcal{M}}_{g,n+1} \rightarrow \overline{\mathcal{M}}_{g,n}$, qui oublie le dernier point marqué ;
- ◊ L'application de recollement séparante, $\sigma : \overline{\mathcal{M}}_{g_1, n_1+1} \times \overline{\mathcal{M}}_{g_2, n_2+1} \rightarrow \overline{\mathcal{M}}_{g_1+g_2, n_1+n_2}$, qui recolle deux courbes sur leur dernier point marqué ;
- ◊ L'application de recollement non séparante, $\tau : \overline{\mathcal{M}}_{g, n+2} \rightarrow \overline{\mathcal{M}}_{g+1, n}$, qui recolle une courbe sur elle-même en identifiant ses deux derniers points marqués.

³¹La bar sur $\overline{\mathcal{M}}_{g,n}$ vient du fait qu'il s'agit en réalité de la compactification de l'espace $\mathcal{M}_{g,n}$ qui paramétrise les mêmes courbes mais non nodales. Ce dernier n'est pas compact malheureusement.

Soit un point $x \in \overline{\mathcal{M}}_{g,n}$, i.e. la classe d'isomorphisme d'une courbe, disons C . Alors il peut être montré que $\pi^{-1}(x) \in \overline{\mathcal{M}}_{g,n+1}$ est en réalité isomorphe à la courbe C elle-même (avec l'action de son groupe d'automorphisme). On peut alors définir des sections $s_i : \overline{\mathcal{M}}_{g,n} \rightarrow \overline{\mathcal{M}}_{g,n+1}$ en posant $s_i(x) = \tilde{x}_i \in \pi^{-1}(x)$, où \tilde{x}_i est l' i -ème marque sur la courbe $\pi^{-1}(x)$.

Puis on les fibrés suivants. Les fibrés en droite $L_i \rightarrow \overline{\mathcal{M}}_{g,n}$, pour $i \in \{1, \dots, n\}$, sont construits de telle façon que les fibres $L_i|_x$ en un point $x \in \overline{\mathcal{M}}_{g,n}$ satisfont $L_i|_x \simeq T_{\tilde{x}_i}^* \pi^{-1}(x)$, où $\tilde{x}_i = s_i(x)$. Ces fibrés sont appelés les *fibrés tautologiques*. Le fibré $\mathbb{H} \rightarrow \overline{\mathcal{M}}_{g,n}$ est défini de sorte que $\mathbb{H}|_x \simeq T^* \pi^{-1}(x)$. Ce fibré est appelé le *fibré de Hodge* et est de rang g . On peut considérer les classes de Chern de ces fibrés :

$$\psi_i := c_1(L_i) \in H^2(\overline{\mathcal{M}}_{g,n}, \mathbb{Q}), \quad \lambda_j := c_j(\mathbb{H}) \in H^{2j}(\overline{\mathcal{M}}_{g,n}, \mathbb{Q}).$$

Les classes ψ_1, \dots, ψ_n sont appelées de *classes psi* et les classes $\lambda_1, \dots, \lambda_g$ sont appelées les *classes lambda*.

Théories cohomologiques des champs. Les théories cohomologiques des champs (CohFT) furent introduites par Kontsevich et Manin [KM94]. Une CohFT consiste en la donnée d'un espace vectoriel complexe V , avec un vecteur spécial e_1 , équipé d'une métrique $\eta = (\eta^{\alpha\beta})$, ainsi que d'applications

$$c_{g,n} : V^{\otimes n} \rightarrow H^*(\overline{\mathcal{M}}_{g,n}, \mathbb{C})$$

qui satisfont les propriétés suivantes (on choisit une base (e_1, e_2, \dots, e_N) et utilisons la notation d'Einstein) :

- (i) $c_{g,n}$ est équivariante par rapport à la permutation simultanée des copies de V dans $V^{\otimes n}$ et celle des marques sur les courbes ;
- (ii) $c_{0,3}(e_1, e_\alpha, e_\beta) = \eta^{\alpha\beta}$;
- (iii) $\pi^* c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = c_{g,n+1}(e_{\alpha_1}, \dots, e_{\alpha_n}, e_1)$;
- (iv) $\sigma^* c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = c_{g_1, n_1+1}(e_1, \dots, e_{n_1}, e_\mu) \eta^{\mu\nu} c_{g_2, n_2+1}(e_\nu, e_1, \dots, e_{n_2})$;
- (v) $\tau^* c_{g+1, n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = \eta^{\mu\nu} c_{g, n+2}(e_\mu, e_\nu, e_{\alpha_1}, \dots, e_{\alpha_n})$.

In fine, on demande à ce que les classes $c_{g,n}$ « se comportent naturellement » par rapport aux applications π, σ et τ , considérant la métrique $\eta^{\alpha\beta}$.

Soit une CohFT $\{c_{g,n} \mid g, n \geq 0\}$, on définit les *nombre d'intersection* (ou *corrélateurs*) associés à la CohFT par

$$\langle \tau_{d_1}(e_{\alpha_1}) \cdots \tau_{d_n}(e_{\alpha_n}) \rangle_g := \int_{\overline{\mathcal{M}}_{g,n}} c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) \cdot \psi_1^{d_1} \cdots \psi_n^{d_n} \in \mathbb{C}.$$

avec $\alpha_i \in \{1, \dots, N\}$ et $d_j \geq 0$. Ces nombres s'annulent si le degré cohomologique de l'intégrande n'est pas égal à $3g - 3 + n$. On définit alors le *potentiel* F de la CohFT comme la série génératrice des nombres d'intersection, c'est-à-dire,

$$F_g(t^*) := \sum_{\substack{n \geq 0 \\ 2g-2+n > 0}} \frac{1}{n!} \sum_{d_1, \dots, d_n \geq 0} \langle \tau_{d_1}(e_{\alpha_1}) \cdots \tau_{d_n}(e_{\alpha_n}) \rangle \cdot t_{d_1}^{\alpha_1} \cdots t_{d_n}^{\alpha_n}. \quad (1.39)$$

et $F(t_*^*; \varepsilon) = \sum_{g \geq 0} \varepsilon^{2g} F_g(t_*^*)$. Cela introduit des variables $t_*^* := \{t_{d_j}^{\alpha_i} \mid \alpha_i \in \{1, \dots, N\}, d_j \geq 0\}$ parfois appelées les *variables topologiques*.

Par exemple, la *CohFT triviale* a pour espace vectoriel $V = \mathbb{C}$, métrique $\eta = 1$ et classes $c_{g,n} = 1$. Elle joue un rôle important car elle est reliée à la hiérarchie KdV [Wit91] comme nous le verrons ci-dessous. Dans la même veine, la *CohFT de Hodge* a les mêmes caractéristiques, sauf que les classes sont données par $c_{g,n} = \Lambda(s) := 1 + \sum_{j=1}^g \lambda^j s^j$. Cette dernière est reliée à la hiérarchie d'« onde longue intermédiaire » [Bur15a].

Un autre exemple important et la théorie de Gromov–Witten d'une variété projective lisse X équipée d'une classe d'homologie $\beta \in H_2(X, \mathbb{Z})$. Elle est construite à partir de l'espace de module $\overline{\mathcal{M}}_{g,n}(X, \beta)$ des courbes de genre g , avec n marques $(C, (x_i))$, avec une application stable $f : C \rightarrow X$ telle que $f_*[C] = \beta$. Nous rappelons la construction à l'exemple 5.2.3. L'espace vectoriel de la CohFT est donné par $V = H^*(X, \mathbb{C})$ et avec la métrique $\eta(v, w) = \int_{\mathcal{P}(v)} w$, où $\mathcal{P} : H^*(X, \mathbb{C}) \rightarrow H_*(X, \mathbb{C})$ est la dualité de Poincaré. Les classes sont définies par

$$c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_N}) = p_* \text{ev}^*(e_{\alpha_1}, \dots, e_{\alpha_N}), \quad (1.40)$$

où $p : \overline{\mathcal{M}}_{g,n}(X, \beta) \rightarrow \overline{\mathcal{M}}_{g,n}$ envoie $(C, (x_i), f)$ sur $(C, (x_i))$ (oublie l'application f) et où $\text{ev} : \overline{\mathcal{M}}_{g,n}(X, \beta) \rightarrow X^n$ envoie $(C, (x_i), f)$ sur $(f(x_i))$ (évalue l'application f). La définition formelle est relativement complexe et implique la classe fondamentale virtuelle $[\overline{\mathcal{M}}_{g,n}(X, \beta)]^{\text{vir}}$ introduite par Behrend et Fantechi [BF97].

Théorie de Fan–Jarvis–Ruan–Witten. Un dernier exemple important pour la suite est la théorie FJRW d'une paire (W, G) . Ici $W \in \mathbb{C}[z_1, \dots, z_m]$ est un polynôme quasi-homogène avec poids q_1, \dots, q_n possédant une singularité isolée à l'origine. Et G est un groupe de matrices diagonales $\gamma = (\gamma_1, \dots, \gamma_m)$ qui laissent le polynôme W invariant, et qui contient l'élément $J = (e^{2i\pi q_1}, \dots, e^{2i\pi q_m})$.

L'espace vectoriel est défini par $V = \bigoplus_{\gamma \in G} (\mathcal{Q}_{W_\gamma} \otimes dz_\gamma)^G$, où $W_\gamma = W|_{\text{Fix}(\gamma)}$ et \mathcal{Q}_{W_γ} est son anneau Jacobien et où $dz^\gamma = \bigwedge_{z_i \in \text{Fix}(\gamma)} dz_i$. On donne une brève revue de la théorie FJRW associée et de la CohFT associée à la section §5.2.1, voir [FJR11, FJR13]. Puisque les polynômes quasi-homogènes possédant une singularité isolée sont classifiées par les diagrammes de Dynkin de type *ADE*, on a une correspondance entre les théories FJRW et les algèbres de Lie semisimples du type correspondant.

Les hiérarchies de Dubrovin–Zhang. Les hiérarchies de Dubrovin–Zhang (DZ) sont des hiérarchies d'EDP du même type que celui des hiérarchies de Drinfeld–Sokolov décrites à la section §1.3.1. Elles sont définie pour tout CohFT semisimple. Cette dernière condition signifie que le potentiel F de la CohFT de l'équation (1.39) doit être tel que l'algèbre complexe N -dimensionnelle définie par les constantes de structure

$$f_{\beta\gamma}^\alpha = \eta^{\alpha\mu} \left. \frac{\partial^3 F_0}{\partial t_0^\mu \partial t_0^\gamma \partial t_0^\beta} \right|_{t_{p>0}^* = 0}$$

est semisimple. On note

$$\Omega_{\alpha,p;\beta,q}(t_*^*; \varepsilon) = \frac{\partial^2 F}{\partial t_p^\alpha \partial t_q^\beta} = \sum_{g \geq 0} \Omega_{\alpha,p;\beta,q}^{[2g]} \varepsilon^{2g}.$$

Puis l'on définit des densités Hamiltoniennes et une structure de Poisson locale par

$$h_{\alpha,p}^{\text{DZ},(0)}(v^*) := \Omega_{\alpha,p+1;1,0}^{[0]}(t_0^* = v^*, 0, 0, \dots), \quad K_{\text{DZ},(0)}^{\alpha\beta} := \eta^{\alpha\beta} \partial_x.$$

Ces densités Hamiltoniennes commutent vis-à-vis du crochet ci-dessus [DZ03]. Cela forme la hiérarchie DZ en genre 0, ou la *hiérarchie principale* de la CohFT. Pour construire la hiérarchie en genre supérieur, on utilise la *solution topologique* ; c'est la solution v^α de la hiérarchie DZ déterminée par la donnée initiale

$$v^\alpha(x, t_*^* = 0) = \delta^{\alpha,1} x.$$

Alors il existe des fonctions rationnelles P_p^α telles que [Get02, BPS12]

$$F_g(t_0^*, t_1^*, t_2^*, \dots) = F_g(P_0^*(v_0^*, \dots, v_{3g-2}^*), \dots, P_{3g-2}^*(v_0^*, \dots, v_{3g-2}^*), 0, 0, \dots) |_{x=0}.$$

Considérons à présent le changement de coordonnées

$$\varphi(v^\alpha) = w^\alpha := v^\alpha + \sum_{g \geq 1} \varepsilon^{2g} \frac{\partial^2 F_g(v_0^*, \dots, v_{3g-2}^*)}{\partial t_0^\alpha \partial x}. \quad (1.41)$$

La hiérarchie DZ en genre supérieur est alors obtenue en appliquant la transformation inverse φ^{-1} à la hiérarchie en genre 0. C'est-à-dire,

$$h_{\alpha,p}^{\text{DZ}}(w_*^*) := h_{\alpha,p}^{\text{DZ},(0)} \circ \varphi^{-1}(w_*^*), \quad K_{\text{DZ}}^{\alpha\beta}(w_*^*) = ((\varphi^{-1})^* K_{\text{DZ},(0)})^{\alpha\beta}(w_*^*),$$

où $(\varphi^{-1})^* K_{\text{DZ},(0)}$ est l'opérateur différentiel transformé ; la transformation est donnée à l'équation (5.18). Ces données forment un système Hamiltonien au sens de l'équation (1.14), i.e.

$$\frac{\partial w^\alpha}{\partial t_p^\alpha} = \{w^\alpha, \bar{h}_{\alpha,p}^{\text{DZ}}\}_{\text{DZ}} = \sum_{\beta=1}^N K_{\text{DZ}}^{\alpha\beta} \frac{\delta h_{\alpha,p}^{\text{DZ}}}{\delta u^\beta}. \quad (1.42)$$

Les Hamiltoniens $h_{\alpha,p}^{\text{DZ}}$ appartiennent à l'anneau

$$\widehat{\mathcal{A}} := \mathbb{C}[[u^\alpha, \varepsilon]][[u_{k>0}^\alpha]]$$

des *polynômes différentiels étendus*. De façon similaire, on note $\widehat{\mathcal{F}} = \widehat{\mathcal{A}}/\text{Im}(\partial_x)$. On définit une gradation sur cet anneau en posant $\deg u_k^\alpha = k$ et $\deg \varepsilon = -1$ de sorte que ε compte le nombre de dérivées dans une monôme de degré total 0. Le paramètre ε est souvent appelé le *paramètre de dispersion* et joue un rôle important dans ce qui suit. Il découle immédiatement de la définition que les Hamiltoniens $h_{\alpha,p}^{\text{DZ}}$ sont tau symétriques dans le sens de [DZ08], i.e.

$$\frac{\partial h_{\alpha,p-1}^{\text{DZ}}}{\partial t_q^\beta} = \frac{\partial h_{\beta,q-1}^{\text{DZ}}}{\partial t_p^\alpha}.$$

L'équation ci-dessus est à comparer avec l'équation (1.33).

Par construction, le potentiel de la CohFT est une fonction tau de la hiérarchie DZ associée à la solution topologique. Il a été prouvé depuis que la hiérarchie DZ construite à partir de la théorie FJRW d'une singularité simple de type ADE coïncide (à certaines considérations près) avec la hiérarchie de Drinfeld–Sokolov associée à l'algèbre de Lie simple correspondante [GM05, FGM10]. Voir [LRZ15] pour un clair exposé sur le sujet. Cette correspondance³² entre CohFT et hiérarchies intégrables nous l'appellerons *correspondance de Witten–Kontsevich généralisée*.

Le premier exemple de telle correspondance (avant même la définition des CohFT et des hiérarchies DZ) est le célèbre théorème de Witten–Kontsevich [Wit91, Kon92] qui stipule que le potentiel de la CohFT triviale est la fonction tau topologique de la hiérarchie KdV (1.20) (qui coïncide avec la hiérarchie DZ associée à la CohFT triviale). Plus précisément, considérons le potentiel $F_{\text{WK}}(t_*)$ de la CohFT triviale :

$$F_{\text{WK}}(t_*) = \sum_{g \geq 0} \sum_{\substack{n \geq 0 \\ 2g - 2 + n > 0}} \frac{1}{n!} \sum_{d_1, \dots, d_n \geq 0} t_{d_1} \cdots t_{d_n} \int_{\mathcal{M}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n}.$$

Puis effectuons les substitutions $F_{\text{WK}} = -\frac{1}{2} \log(\tau_{\text{WK}})$ puis $t_i \mapsto (2i + 1)!! t_{2i+1}$ où $(2i + 1)!! = 1 \cdot 3 \cdot 5 \cdots (2i + 1)$. Alors la fonction qui en résulte τ_{WK} est un fonction tau de la hiérarchie KdV au sens où $v = -2\partial_x^2 \log(\tau_{\text{WK}})$ est solution de (1.20) (dans la substitution, $t_0 \mapsto t_1$ et on pose $t_1 = x$).³³ De plus, v satisfait $v(x, t_{i>1} = 0) = x$.

1.4.2 Les hiérarchies de double ramifications

On explique à présent la construction des hiérarchies de double ramification (DR) introduites par Buryak [Bur15a]. Comme mentionné précédemment, il est conjecturé que ces dernières sont équivalentes aux hiérarchies DZ. Mais elles sont, dans une certaine mesure, plus facile à construire et satisfont de fortes relations de récurrence jusque lors inconnues. De plus, elles sont définie pour toute CohFT, pas seulement celle semisimples. Dans le cas de la théorie FJRW d'une singularité simple de type ADE, la conjecture DR/DZ forte stipule que les hiérarchies DR, DZ et DS coïncident. Ce fait a été vérifié pour les hiérarchies de type A_1 à A_5 (et d'autre, voir ci-dessous) et nous l'avons confirmé dans le cas D_4 dans une prépublication à paraître décrite au chapitre 5.

Structure Hamiltonienne. Tout d'abord, la structure de Poisson locale est simplement donnée par l'opérateur $\eta^{\alpha\beta} \partial_x$, où $\eta^{\alpha\beta}$ est la métrique de la CohFT, c'est-à-dire que pour deux fonc-

³²Il s'agit bien d'une correspondance « CohFT / Hiérarchie int. » dans le sens où le potentiel de la CohFT semisimple est une fonction tau de la hiérarchie DZ, mais permet aussi de construire la structure de Poisson locale ainsi que toutes les densités Hamiltoniennes. En retour, la fonction tau topologique, vue comme le potentiel, contient par construction toute les classes de cohomologie et donc, toute l'information de la CohFT.

³³Pour des raisons apparentes dans cet exemple, on appelle souvent F la fonction tau.

tionnelles locales $\bar{f}, \bar{g} \in \widehat{\mathcal{F}}$,

$$\{\bar{f}, \bar{g}\} = \int \sum_{\alpha, \beta=1}^N \frac{\delta f}{\delta u^\alpha} \eta^{\alpha\beta} \partial_x \frac{\delta f}{\delta u^\beta}.$$

Pour les densités Hamiltoniennes, on introduit les *cycles de double ramification*. Soient des entiers a_1, \dots, a_n tels que $\sum a_i = 0$ et soit un entier $g \geq 0$. Le cycle de double ramification $\text{DR}_g(a_1, \dots, a_n)$ est, heuristiquement, le Poincaré dual de la classe d'homologie du locus des courbes complexes marquées $(C, (x_i))$ telles que $\sum a_i x_i$ soit le diviseur des zéros et pôles de quelque fonction méromorphe. On considère x_i comme étant un pôle si $a_i > 0$ et un zéro sinon. Une définition plus précise est donnée à la section §5.2.2.

Puis l'on définit les nombres d'intersection suivants : soient $\alpha \in \{1, \dots, N\}$ et $g, d \geq 0$,

$$P_{\alpha, d; \underline{\alpha}}^{g; \underline{a}} := \int_{\text{DR}_g(-\sum a_i, \underline{a})} \lambda_g \psi_1^d c_{g, n+1}(e_\alpha, e_{\alpha_1}, \dots, e_{\alpha_n}),$$

où $\underline{\alpha} = (\alpha_1, \dots, \alpha_n) \subset \{1, \dots, N\}$ et $\underline{a} = (a_1, \dots, a_n) \subset \mathbb{Z}$. La classe λ_g est la classe de Chern $c_g(\mathbb{H})$ de plus haut degré du fibré de Hodge. Il peut être montré que ces nombres sont polynomiaux en les a_i et homogènes de degré $2g$ de sorte que l'on peut les écrire comme

$$P_{\alpha, d; \underline{\alpha}}^{g; \underline{a}} = \sum_{\substack{b_1, \dots, b_n \geq 0 \\ \sum b_j = 2g}} \widetilde{P}_{\alpha, d; \underline{\alpha}}^{g; \underline{b}} a_1^{b_1} \cdots a_n^{b_n} \in \mathbb{C}_{2g}[a_1, \dots, a_n],$$

On peut à présent définir les densités Hamiltoniennes :

$$g_{\alpha, d} := \sum_{\substack{g, n \geq 0 \\ 2g-2+n > 0}} \frac{\varepsilon^{2g}}{n!} \sum_{\substack{b_1, \dots, b_n \geq 0 \\ \sum b_j = 2g}} \widetilde{P}_{\alpha, d; \underline{\alpha}}^{g; \underline{b}} \prod_{j=1}^n u_{b_j}^{\alpha_j}. \quad (1.43)$$

Comme précédemment, la donnée de la structure de Poisson locale $\eta^{\alpha\beta} \partial_x$ et des densités Hamiltoniennes $g_{\alpha, d}$ forment un système Hamiltonien au sens de l'équation (1.42). Les densités $g_{\alpha, d}$ satisfont la relation de récurrence suivante : tout d'abord, on pose $g_{\alpha, -1} := \eta_{\alpha\mu} u^\mu$, où $\eta_{\alpha\mu}$ est la matrice inverse de $\eta^{\alpha\mu}$, puis l'on a [BR16b]

$$\partial_x(D-1)g_{\alpha, p+1} = \{g_{\alpha, p}, \bar{g}_{1,1}\} = \sum_{\beta, \gamma=1}^N \sum_{k \geq 0} \frac{\partial g_{\alpha, p}}{\partial u_k^\beta} \eta^{\beta\gamma} \partial_x^{k+1} \frac{\delta g_{1,1}}{\delta u^\gamma},$$

où $D = \varepsilon \partial_\varepsilon + \sum_{s \geq 0} u_s^\alpha \partial_{u_s^\alpha}$. Il peut être aisément montré que cela permet effectivement de calculer les densités récursivement à partir de la seule donnée $\bar{g}_{1,1}$. Nous insistons sur le fait que cette récurrence précisément est nouvelle vis-à-vis de la théorie des hiérarchies intégrables.

D'autre part, ces densités Hamiltoniennes ne sont pas tau symétriques. Pour obtenir des densités tau symétriques, il suffit de considérer les polynômes différentiels

$$h_{\alpha, p} := \frac{\delta g_{\alpha, p}}{\delta u^1}.$$

Par rapport à la structure de Poisson $\eta^{\alpha\beta} \partial_x$, ces polynômes sont des densités Hamiltoniennes tau symétriques.

La conjecture d'équivalence DR/DZ forte. Afin d'énoncer la conjecture, il nous faut définir la notion de *transformation de Miura* qui joue le rôle de « changement de coordonnées » dans ce contexte. Une transformation de Miura est toute application inversible $\varphi : \widehat{\mathcal{A}} \rightarrow \widehat{\mathcal{A}}$, $v^\alpha := \varphi(u^\alpha) \in \mathcal{A}^{[0]}$, telle que

$$\det \left[\frac{\partial v^\alpha}{\partial u^\beta} \Big|_{\varepsilon=0} \right]_{\alpha, \beta \in [1, N]} \neq 0.$$

On choisit à présent des coordonnées particulières pour la hiérarchie que nous appelons *normales* suivant [DZ08]. À la donnée d'une structure tau $h_{\alpha, p}$, elles sont définies par

$$\tilde{u}^\alpha := \eta^{\alpha\beta} h_{\alpha, -1}.$$

L'application $\varphi : u^\alpha \mapsto \tilde{u}^\alpha$ est une transformation de Miura. On note

$$h_{\alpha, p}^{\text{DR}} := h_{\alpha, p} \circ \varphi, \quad K_{\text{DR}}^{\alpha\beta} := \varphi^*(\eta^{\alpha\beta} \partial_x) \quad (1.44)$$

les densités Hamiltoniennes et la structure de Poisson écrits dans les coordonnées normales. On peut à présent énoncer la conjecture d'équivalence DR/DZ forte. Plus de détails sont donnés à la conjecture 5.2.23.

Conjecture 1.4.1 (Équivalence DR/DZ forte [BDGR18]). *Pour toute CohFT semisimple il existe un unique polynôme différentiel $f \in \widehat{\mathcal{A}}^{[-2]}$ tel que les hiérarchies DR et DZ coïncident après la transformation de Miura définie par*

$$\tilde{u}^\alpha = w^\alpha + \eta^{\alpha\beta} \partial_x \{f, \bar{h}_{\beta, 0}^{\text{DZ}}\}_{\text{DZ}},$$

En particulier, pour les CohFT définies par la théorie FJRW d'une singularité simple de type ADE, la conjecture stipule que les hiérarchie DR, DZ et DS coïncident. L'équivalence DR/DZ forte a été confirmée dans un certain nombre de cas [Bur15a, BG16, BDGR18] : la CohFT triviale ; la CohFT de Hodge totale ; les hiérarchies de type A_3, A_4, A_5 ; la théorie de Gromov–Witten de \mathbb{P}^1 ; jusqu'à l'ordre 5 pour toutes les CohFT de rang 1 ; jusqu'à l'ordre 2 pour toutes les CohFT semisimple.

1.4.3 Hiérarchie de type D_4 et l'équivalence DR/DZ forte

Finalement, on explique comment nous avons confirmé l'équivalence DR/DZ forte dans le cas de la hiérarchie de type D_4 . Cela revient à prouver que les hiérarchies DR et DS de type D_4 coïncident. Pour ce faire, nous avons utilisé le théorème suivant de Buryak et Guéré [BG16] : soit une CohFT semisimple, et supposons qu'il existe une transformation de Miura $\varphi : u^\alpha \mapsto \tilde{u}^\alpha(u_*^*; \varepsilon)$, telle que

$$\frac{\partial \tilde{u}^\alpha}{\partial u^1} = \delta^{\alpha, 1}, \quad \varphi^*(\eta^{\alpha\beta} \partial_x) = K_{\text{DZ}}^{\alpha\beta}, \quad \overline{g_{1,1} \circ \varphi} = \bar{h}_{1,1}^{\text{DZ}}, \quad (1.45)$$

où $\varphi^*(\eta^{\alpha\beta} \partial_x)$ est donnée par (5.18). Alors la hiérarchie DR, sous la transformation φ , coïncide avec la hiérarchie DZ. En calculant la hiérarchie DR associée à la théorie FJRW de type D_4 ,

on trouve que la transformation de Miura depuis les coordonnées u^α vers les coordonnées normales \tilde{u}^α est donnée par

$$\begin{cases} \tilde{u}^1 = u^1 + (-\frac{1}{36}(u_1^3)^2 - \frac{1}{36}u^3u_2^3)\varepsilon^2 - \frac{1}{45}u_4^3\varepsilon^4 \\ \tilde{u}^2 = u^2 \\ \tilde{u}^3 = u^3 \\ \tilde{u}^4 = u^4 \end{cases} \quad (1.46)$$

C'est en accord avec la première condition de l'équation (1.45). Les densités Hamiltoniennes $g_{\alpha,0}$ et $g_{1,1}$ dont données à la section §5.2.4.

Il reste donc à expliquer comment nous avons calculer la hiérarchie de Drinfeld–Sokolov de type D_4 . Pour ce faire, nous avons utilisé la paire de Lax scalaire et la structure Hamiltonienne trouvée par Liu, Wu et Zhang [LWZ10]. En effet, les hiérarchies DS sont construites génériquement comme des hiérarchies de paires de Lax matricielles. Mais précisément pour le cas des hiérarchies de type D_n , Drinfeld et Sokolov ne donnèrent pas une paire de Lax scalaire pour toutes les équations de la hiérarchie. Afin d'énoncer le problème, il nous faut discuter l'algèbre affine de type $D_4^{(1)}$.

La paire de Lax matricielle de la hiérarchie DS de type D_4 . Ce que nous discutons ici se généralise aux algèbres affines de type $D_n^{(1)}$. Rappelons que l'algèbre affine \mathfrak{g} de type $D_4^{(1)}$ est l'algèbre affine associée à l'algèbre simple D_4 , cette dernière étant isomorphe à l'algèbre $\mathfrak{o}(8)$ (matrices complexes 8×8 antisymétriques – non hermitiennes). I.e. $D_4^{(1)}$ est isomorphe à

$$\mathfrak{g} = \mathfrak{o}(8) \otimes \mathbb{C}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c.$$

Cette algèbre affine a pour ensembles d'exposants positifs $\mathbb{E}^+ = \mathbb{Z}_{>0}^{\text{odd}} \cup (3\mathbb{Z}_{>0}^{\text{odd}})' = \{1, 3, 3', 5, 7, 9, 9', \dots\}$, ce qui veut dire que tout exposant de la forme $3k$ avec k impair, a multiplicité 2. On note Λ_k (resp. Γ_k), avec k impair, le générateur de la sous-algèbre de Heisenberg associé à l'exposant $3k$ (resp. $(3k)'$). De façon similaire, on note t_k (resp. \hat{t}_k) la variable associée à Λ_k (resp. Γ_k). Les Λ_k et Γ_k sont donnés aux équations (5.54) et (5.55). Cela donne les équations de Lax matricielles suivantes [DS84]

$$\frac{\partial \mathcal{L}}{\partial t_k} = [(e^{\text{ad}_U} \Lambda_k)_+, \mathcal{L}], \quad (1.47)$$

$$\frac{\partial \mathcal{L}}{\partial \hat{t}_k} = [(e^{\text{ad}_U} \Gamma_k)_+, \mathcal{L}], \quad (1.48)$$

pour $k \in \mathbb{Z}_{>0}^{\text{odd}}$, où $\mathcal{L} = \partial_x + \Lambda + Q$, avec $Q \in \mathring{\mathfrak{b}}[[t]]$ et où Λ est l'élément cyclique principal donné à l'équation (1.29).

La paire de Lax scalaire la structure Hamiltonienne. Dans [DS84], Drinfeld et Sokolov décrivent la structure Hamiltonienne par rapport aux variables t_k , mais pas \hat{t}_k (il fallait une nouvelle algèbre d'opérateurs). La structure de Poisson locale, à la façon de l'équation (1.22),

est définie comme suit [DS84]. Premièrement, l'opérateur de Lax scalaire est donné par

$$L = \partial_x^6 + \partial_x^{-1} \sum_{\mu=1}^3 \left(s^\mu \partial_x^{2\mu-1} + \partial_x^{2\mu-1} s^\mu \right) + \partial_x^{-1} \varrho \partial_x^{-1} \varrho, \quad s^4 = (\varrho)^2, \quad (1.49)$$

Puis en notant $L' = \partial_x \circ L$, la première structure de Poisson est donnée par

$$\{\bar{f}, \bar{g}\}_1 = \int \text{res } X [(\partial_x Y_+ L')_- - (L' Y_+ \partial_x)_- - (\partial_x Y_- L')_+ + (L' Y_- \partial_x)_+] dx, \quad (1.50)$$

pour tout $\bar{f}, \bar{g} \in \widehat{\mathcal{F}}$,³⁴ où $X = \frac{\delta f}{\delta L'}$, $Y = \frac{\delta g}{\delta L'}$ sont donnés à l'équation (5.66). Cela permet de définir la paire de Lax scalaire ainsi que la structure Hamiltonienne³⁵ de l'équation (1.47) par [DS84]

$$\frac{\partial L}{\partial t_k} = \left[(L^{k/6})_+, L \right], \quad \iff \quad \frac{\partial s^\alpha}{\partial t_k} = \{s^\alpha, H_{k+6}\}_1, \quad H_k = \frac{6}{k} \int \text{res } L^{k/6}.$$

Cette dernière est relativement similaire à celle de la hiérarchie KdV (1.20), (1.14). Pour les flots par rapport aux variables \hat{t}_k , il n'est pas possible d'utiliser la racine 6-ième $L^{1/6}$.

Ce que firent Liu, Wu et Zhang [LWZ10] c'est qu'ils construisirent une nouvelle algèbre $\widetilde{\mathcal{D}}$ qui intersecte \mathcal{D} (équation (1.19)) mais qui ne la contient pas. Essentiellement, ils utilisèrent la topologie sur $\widehat{\mathcal{A}}$ induite par la gradation $\widehat{\mathcal{A}} = \bigoplus_{k \in \mathbb{Z}} \widehat{\mathcal{A}}_k$ définie par $\deg s_k^\alpha = k$, pour définir une structure d'anneau (avec le même produit que (1.7)) sur l'espace

$$\widetilde{\mathcal{D}} = \left\{ \sum_{i \geq m} \sum_{j \geq m-i} a^{i,j} \partial_x^i \mid m \in \mathbb{Z}, a^{i,j} \in \widehat{\mathcal{A}}_j \right\}.$$

L'algèbre qui en résulte $\widetilde{\mathcal{D}}$, dont nous appelons les éléments des *opérateurs pseudo-différentiels du second type*, contient des séries avec une infinité de termes de degrés positifs et négatifs.

Le résultat est qu'il existe un unique opérateur $L^{1/2} \in \widetilde{\mathcal{D}}$ tel que $(L^{1/2})^2 = L$ satisfaisant certaines contraintes (avec L donné à l'équation (1.49)). Cela permet de définir la représentation de Lax scalaire de l'équation (1.48) [LWZ10] par

$$\frac{\partial L}{\partial \hat{t}_k} = \left[(L^{k/2})_+, L \right]. \quad \iff \quad \frac{\partial s^\alpha}{\partial \hat{t}_k} = \{s^\alpha, \hat{H}_{k+2}\}_1, \quad \hat{H}_k = \frac{2}{k} \int \text{res } L^{k/2}.$$

Calculs explicites de la structure Hamiltonienne et comparaison avec la hiérarchie DR. Finalement, nous donnons les calculs explicites effectués à la section §5.3.2 sur la hiérarchie DS de type D_4 et l'équivalence entre la hiérarchie DS et DR confirmant ainsi la conjecture forte dans le cas D_4 . Pour ce faire, nous devons substituer les variables t_k et \hat{t}_k , avec $k \in \mathbb{Z}_{>0}^{\text{odd}}$,

³⁴Notons qu'ici, $\widehat{\mathcal{F}} = \widehat{\mathcal{A}}/\text{Im}(\partial_x)$ où $\widehat{\mathcal{A}} = \mathbb{C}[[s_0^\alpha; \varepsilon]][[s_{k>0}^\alpha]]$.

³⁵Notons qu'ici on ne s'intéresse qu'à l'une des deux structures Hamiltoniennes. C'est parce que ce que nous voulons c'est comparer les structures Hamiltoniennes de deux hiérarchies *a priori* différentes. La structure bi-Hamiltonienne a de nombreux intérêts, mais nous n'en avons pas besoin ici.

par les variables topologiques $t_{\alpha,p}$, avec $\alpha \in \{1, \dots, 4\}$ et $p \geq 0$. De façon similaire, nous devons définir des densités Hamiltoniennes tau symétriques $h_{\alpha,p}$ au sens de l'équation (1.33). On a, pour $\mu \in \{1, 2, 3\}$ et $p \geq 0$,

$$\begin{aligned} t_{\mu,p} &= \Gamma_{\mu,p} \cdot t_{6p+2\alpha-1}, & h_{\mu,p-1} &= \Gamma_{\mu,p}^{-1} \cdot \text{res } L^{(6p+2\alpha-1)/6}, \\ t_{4,p} &= \Gamma_{4,p} \cdot \hat{t}_{2p+1}, & h_{4,p-1} &= \Gamma_{4,p}^{-1} \cdot \text{res } L^{(2p+1)/2}. \end{aligned}$$

Les facteurs $\Gamma_{\alpha,p}$ sont donnés à l'équation (5.68). Il reste à substituer les fonctions s^α par les coordonnées normales $\tilde{u}^\alpha = \eta^{\alpha\beta} h_{\beta,-1}$ de la hiérarchie DS de type D_4 (on note volontairement ces coordonnées comme celles de la hiérarchie DR). On note $h_{\alpha,p}^{\text{DS}}$ les densités Hamiltoniennes tau symétriques écrites dans les coordonnées normales \tilde{u}^α . De façon similaire, dans les coordonnées normales, la première structure de Poisson (1.50) prend la forme suivante

$$(K_{\text{DS}}^{\alpha\beta}) = \begin{pmatrix} \left(\frac{1}{6} \tilde{u}_2^3 \partial_x + \frac{1}{2} \tilde{u}_1^3 \partial_x^2 + \frac{1}{3} \tilde{u}^3 \partial_x^3 \right) \varepsilon^2 + \frac{4}{15} \partial_x^5 \varepsilon^4 & 0 & 6\partial_x & 0 \\ 0 & 6\partial_x & 0 & 0 \\ 6\partial_x & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\partial_x \end{pmatrix},$$

où $(K_{\text{DS}}^{\alpha\beta})$ est l'opérateur différentiel matriciel tel que $\{\bar{f}, \bar{g}\}_1 = \int \frac{\delta f}{\delta \tilde{u}^\alpha} K_{\text{DS}}^{\alpha\beta} \frac{\delta g}{\delta \tilde{u}^\beta}$. En tenant compte du théorème de Buryak et Guéré (voir le début de la section §1.4.3), on a à présent tout le matériel nécessaire pour énoncer le théorème principal de cette section. Sa preuve (i.e. les calculs) est donnée à la section §5.3.2.

Théorème 1.4.2 (Équivalence DR/DZ forte pour le cas D_4 [DR19]). *L'équivalence DR/DZ forte est vérifiée pour la hiérarchie de type D_4 via la transformation de Miura (1.46). Plus précisément, dans les coordonnées normales \tilde{u}^α ,*

- (i) *Les structures de Poisson locales sont les mêmes : $K_{\text{DR}}^{\alpha\beta} = K_{\text{DS}}^{\alpha\beta}$;*
 - (ii) *Les densités Hamiltoniennes $h_{1,1}^{\text{DR}}$ et $h_{1,1}^{\text{DS}}$ diffèrent d'une dérivée totale, i.e. $\bar{h}_{1,1}^{\text{DR}} = \bar{h}_{1,1}^{\text{DS}}$.*
- De plus, les densités Hamiltoniennes au niveau 0 coïncident $h_{\alpha,0}^{\text{DR}} = h_{\alpha,0}^{\text{DS}}$ pour $\alpha \in \{1, \dots, 4\}$.*

Chapter 2

Introduction *in English*

2.1 Historical overview of the theory and of our results

Apart from a few equations, section 2.1 is rather historical and shall be readable by non specialists.

The present thesis deals with the computation and applications of tau functions of the Drinfeld–Sokolov hierarchies of equations. The Drinfeld–Sokolov hierarchies are infinite sequences of partial differential equations which flows commute pairwise. For that reason they form an infinite-dimensional integrable system. They were introduced in 1984 by V. G. Drinfeld and V. V. Sokolov in their seminal article [DS84]. In the latter, they showed how to associate these hierarchies to any affine Kac–Moody algebra \mathfrak{g} [Kac68, Moo68]. The latter we realize as the central extension of the loop algebra of a semisimple Lie algebra \mathfrak{g} , i.e.¹

$$\mathfrak{g} = \mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c, \quad (2.1)$$

with the appropriate Lie bracket making c into a central element; see (2.28).

Let us give a simple picture of these systems of equations. A solution of a Drinfeld–Sokolov hierarchy is made of functions u^1, \dots, u^N which depend on infinitely many variables $\{t_i \mid i \geq 1\}$ with $x = t_1$; the system of equations of the hierarchy then has the form

$$\{u_{t_i}^\alpha = A_i^\alpha(u^1, \dots, u^N) \mid \alpha \in \{1, \dots, N\}, i \geq 1\}, \quad (2.2)$$

where $u_{t_i}^\alpha = \partial_{t_i} u^\alpha$ and where A_i^α is polynomial in the u^β 's and their derivatives w.r.t. x : u_x^β, u_{xx}^β ,

¹Notice that the affine Kac–Moody algebra associated to the finite-dimensional semisimple Lie algebra \mathfrak{g} is actually a co-central extension of the algebra \mathfrak{g} in (2.1) (see [Kac90] §7). However, the Drinfeld–Sokolov hierarchies do not involve the co-central extension; calling \mathfrak{g} an affine algebra is somehow customary.

etc. For example, the celebrated Korteweg–de Vries equation [Bou77, KdV95] is of this form:

$$u_t = \frac{1}{4}u_{xxx} + \frac{3}{2}uu_x. \quad (2.3)$$

The Drinfeld–Sokolov hierarchies are sequences of equations that share many features with the Korteweg–de Vries equation and, as we will see, Equation (2.3) plays a pivotal role in the theory. We shall emphasize that although we study partial differential equations, we are only looking for algebraic solutions of these equations: polynomials, power series or rational functions.

In the next sections of this introduction, we will develop the main concepts that this thesis is based on and explain our results: [Duc17, CDY18, DR19]. But before that, let us give a brief history of Drinfeld–Sokolov hierarchies and tau functions and how our works fit into this dense area of mathematical physics.

First of all, in the case where \mathfrak{g} is the affine algebra associated to the simple Lie algebra $\mathfrak{sl}(n, \mathbb{C})$ ($n \times n$ complex traceless matrices), the Drinfeld–Sokolov hierarchy coincides with the n -th Gelfand–Dickey hierarchy [GD76], the latter being much simpler! The n -th Gelfand–Dickey hierarchy admits the following representation: first consider the differential operator $L = \partial_x^n + u^{n-2}\partial_x^{n-2} + \dots + u^1\partial_x + u^0$ where the u^i 's are smooth functions of parameters $\{t_i \mid i \geq 1, i \neq 0 \pmod n\}$ with $t_1 = x$. Then the n -th Gelfand–Dickey hierarchy consists of the family of equations²

$$\left\{ \frac{\partial L}{\partial t_i} = \left[(L^{i/n})_+, L \right] \mid i \geq 1, i \neq 0 \pmod n \right\}, \quad (2.4)$$

where $L^{1/n}$ is the n -th root of L with respect to the composition of differential operators.³ In particular, for $\mathfrak{sl}(2, \mathbb{C})$ (specializing $n = 2$ in (2.4)) it comes down to the Korteweg–de Vries hierarchy [Lax68, Miu68, GKM68] (whose flow along t_3 reads the Korteweg–de Vries equation (2.3)). Equations (2.4) are called *scalar Lax pairs*. For the most part, the Drinfeld–Sokolov hierarchies admit a scalar Lax pair along with a bi-Hamiltonian representation [DS84] making them integrable in many ways. However, in the general case, they are generically constructed as hierarchies of *matrix Lax equations* (or zero-curvature equations) which we introduce in the next section. Let us note that the commutation of the flows of the Drinfeld–Sokolov hierarchies essentially results from the existence of infinite-dimensional abelian subalgebras of the affine Kac–Moody algebra \mathfrak{g} .

The development of the notion of tau function. The notion of tau function on the other hand emerged in the early 1980's with the works of the Kyoto school led by Sato along with Date, Jimbo, Kashiwara, Miwa [JMS78, KM81, DKM81, DJKM81a] to cite a few. Originally, the tau function is defined for the Kadomtsev–Petviashvili hierarchy [Sat81] which contains all the Gelfand–Dickey hierarchies.⁴ A solution of the Kadomtsev–Petviashvili hierarchy is the data

²The fact that we only consider flows w.r.t. variables t_i with $i \neq 0 \pmod n$ simply comes from the fact that for $i = jn$, the flow is null since then $(L^{j/n})_+ = L^j$ commutes with L .

³The definition of the n -th root necessitates the introduction of pseudo-differential operators; see §2.2.2.

⁴The Kadomtsev–Petviashvili hierarchy contains all the Gelfand–Dickey hierarchies in the sense that any solution of any of the Gelfand–Dickey hierarchies is also a solution of the Kadomtsev–Petviashvili hierarchy (the converse being wrong).

of infinitely many functions; the tau function is one single function that contains all the information of a solution. Moreover, the tau function satisfies partial differential equations that are different in nature to that of the hierarchy itself. Namely, the equations on the tau function are bilinear and deeply connected to Plücker relations. The Kadomtsev–Petviashvili (KP) hierarchy and its tau functions were soon generalized to the so-called BKP and CKP hierarchies which are connected to Lie algebras of orthogonal (B_n, D_n) and symplectic (C_n) type [DJKM81b]. For exceptional Lie algebras, KP-like hierarchies were defined by Kac and Wakimoto [KW89]. We shall retain that a tau function is equivalent to a solution of the hierarchy, yet that the two give very different approaches to constructing or solving the hierarchy.⁵

In the meantime, Edward Witten proposed his famous 1991 conjecture connecting a certain tau function of the Korteweg–de Vries hierarchy to the generating series of the intersection numbers on the moduli spaces of stable curves $\overline{\mathcal{M}}_{g,n}$ [Wit91]. Witten’s conjecture, proven by Maxim Kontsevich [Kon92] the year after, initiated a brand new way of understanding integrable hierarchies which we discuss below.

Subsequently to the works of Drinfeld–Sokolov on the one hand and of the Kyoto school and Kac–Wakimoto on the other, several ways of defining tau functions were designed for Drinfeld–Sokolov hierarchies. One of them, initiated by Kac [Kac90] and Hollowood and Miramontes [HM93, Mir99], makes use of highest weight representations of affine Kac–Moody algebras, in this approach, it connects to the bilinear equations. A different approach is to construct a specific family of Hamiltonian densities called *tau symmetric* and which, essentially via Poincaré lemma, give a natural definition of the tau function [DZ08].⁶

In the present thesis, we are interested in two classes of tau functions: the polynomial tau functions and the topological tau function. Polynomial tau functions of a given Drinfeld–Sokolov hierarchy form a vast class of tau functions yet the simplest ones. For the specific case of the Korteweg–de Vries hierarchy, they are known to be classified by the Adler–Moser polynomials [AM78]. The zeros of these polynomials evolve through an integrable system called of Calogero–Moser type [AMcKM77]; generalized to certain solutions of the Kadomtsev–Petviashvili hierarchy by Wilson in [Wil98]. Hence polynomial tau functions of the Drinfeld–Sokolov hierarchies are expected to present analogous features yet to be discovered. On the other hand, *the* topological tau function of a given Drinfeld–Sokolov hierarchy is the main object upon which bridges from integrable hierarchies and geometry are built. We will get to it below; let us first discuss polynomial tau functions.

Polynomial tau functions of the Drinfeld–Sokolov hierarchies (treated in Chapter 4).

In 2012 Chao–Zhong Wu, inspired by [HM93], proposed a universal construction of tau symmetric Hamiltonian densities for any of the Drinfeld–Sokolov hierarchies associated to an affine Kac–Moody algebra \mathfrak{g} [Wu17].⁷ This gives a generic tau function for the Drinfeld–Sokolov hierarchies. In [CW15a, CW15b], Cafasso and Wu expressed Wu’s tau function as the limit of Toeplitz determinants through a certain faithful traceless representation of the underlying sem-

⁵It should be noted that the same Kyoto school invented, around the same time, the theory of tau functions for Painlevé equations and isomonodromic problems [Miw81, JMW81]. But this not the theory we discuss here.

⁶To clarify the chronology: this article was originally submitted to the arXiv on August 23 2001.

⁷To clarify the chronology: this article was originally submitted to the arXiv on March 26 2012.

simple Lie algebra [DS84] (see also Appendix A). This generalized part of the early work of Cafasso [Caf08] on the Gelfand–Dickey hierarchies.

Cafasso–Wu construction associates a tau function to any element $X \in \mathfrak{g}_-$ of the affine algebra \mathfrak{g} that has only negative powers of the loop parameter. In [CDY18], using the results of Cafasso and Wu, we showed how to expand Wu’s tau function as a sum over all partitions of integers (Theorem 4.5.2); the summands involve polynomials that generalize Schur polynomials [Sch01] (also [Mac95]). As a consequence, we obtained an easy criterion for polynomiality of the tau function: if $X \in \mathfrak{g}_-$ is such that its matrix representation is nilpotent (w.r.t. to the usual matrix power), then the associated tau function is polynomial.⁸ We gave several non trivial examples of polynomial tau functions and described applications to the bilinear equations of the Drinfeld–Sokolov hierarchies. This is the object of Chapter 4. It will constitute a future project to study the properties of these generalized Schur polynomials and of these polynomial tau functions.

Polynomial tau functions of the Korteweg–de Vries hierarchy (treated in Chapter 3).

While working on polynomial tau functions of the Drinfeld–Sokolov hierarchies, we became interested in the specific case of the Korteweg–de Vries hierarchy. The latter, as we described, is a special case of the Drinfeld–Sokolov hierarchies for the affine Lie algebra associated to the simple Lie algebra $\mathfrak{sl}(2, \mathbb{C})$. Moreover, every solution of the Korteweg–de Vries hierarchy is also a solution of the Kadomtsev–Petviashvili hierarchy; the reduction being to set all evenly indexed variables t_{2i} to zero. As we have said, the polynomial tau functions of the Korteweg–de Vries hierarchy have been much studied and classified by Adler, McKean and Moser [AMcKM77, AM78]. Moreover, as it was first understood by Sato [Sat81] (see also [Hir04]), each polynomial tau function of the Korteweg–de Vries hierarchy can be written in terms of Schur polynomials.

The main result of [AM78] is that there exists a change of the variables t_i which transforms the polynomial tau functions of the Korteweg–de Vries hierarchy into a sequence of polynomials which satisfies a much simple differential recursion, the so-called Adler–Moser polynomials. In [Duc17] we were able to exhibit this change of variables (Theorem 3.5.4); the proof uses essentially the Kadomtsev–Petviashvili theory. This is the object of Chapter 3.

Recently, Kac and van de Leur made significant progress in the field [KvdL18] by showing that *all* the polynomial tau functions of the Kadomtsev–Petviashvili hierarchy (thus including the Gelfand–Dickey hierarchies) can be expressed surprisingly simply as a Schur polynomial modulo a certain change of variables.⁹ They also worked out the polynomial tau functions of the modified Kadomtsev–Petviashvili hierarchy, but we will not discuss it here.

Witten’s conjecture and the topological tau function. We now come back to Witten’s conjecture and its consequences. Witten’s conjecture states that the generating series of the

⁸In certain cases, it can actually be that some integer power τ^a of the tau function τ is polynomial (in the cases we computed, $a = 1$ or 2). But interestingly enough, in all cases we computed, the tau function itself was polynomial; i.e. when $a = 2$, τ was a perfect square. We have not proven this fact and only observed it on examples.

⁹This result differs from the original work of Sato [Sat81] for Kac and van de Leur they did not have to consider linear combinations of Schur polynomials.

intersection numbers on the moduli spaces of stable curves $\overline{\mathcal{M}}_{g,n}$ is actually a tau function of the Korteweg–de Vries hierarchy [Wit91]. The conjecture was proven by Maxim Kontsevich [Kon92] and this tau function is now called the Witten–Kontsevich tau function. Since then, tau functions became an essential object that connects complex geometry, algebraic geometry and integrable systems in an active area of mathematical physics sometimes referred to as *integrable mirror symmetry*. Soon after, Witten’s conjecture was generalized to connect tau functions of Gelfand–Dickey hierarchies and Witten’s r -spin classes [Wit93]. In 2003 Dubrovin and Zhang showed that the generating function of the Gromov–Witten invariants of the complex projective line \mathbb{P}^1 is a tau function of the extended Toda hierarchy [DZ03] (which is *not* a Drinfeld–Sokolov hierarchy).

Cohomological field theories (CohFT), introduced by Kontsevich and Manin in [KM94], are cohomology classes on the moduli spaces of stable curves that are defined in view of unifying Witten’s r -spin classes as well as Gromov–Witten theories. In the early 2010’s, Fan, Jarvis and Ruan defined the so-called Fan–Jarvis–Ruan–Witten CohFTs in view of generalizing Witten’s r -spin classes to arbitrary simple complex singularities [FJR11, FJR13].

On the other hand, in [DZ08], Dubrovin and Zhang showed how to associate to any semisimple¹⁰ CohFT an integrable hierarchy that was shown to coincide, in some cases,¹¹ with the Drinfeld–Sokolov hierarchies [GM05, FGM10]. In Dubrovin–Zhang construction, the hierarchy is designed such that the generating series of the semisimple CohFT it is based on is a tau function. This special tau function we call the *topological tau function*; for example, the Witten–Kontsevich tau function is the topological tau function of the Korteweg–de Vries hierarchy. Since then, integrable mirror symmetry became a matter of building bonds between CohFTs and integrable hierarchies. Although some examples work very well,¹² (see [LRZ15] for a clear expositions) there seem to be no universal pattern so far.

One of the goals of Dubrovin–Zhang construction was to classify integrable hierarchies [DZ08, DLZ08]. The idea was to use the intersection theory of the moduli spaces of stable curves¹³ (and generalizations: Gromov–Witten theory, Fan–Jarvis–Ruan–Witten theory) to define integrable hierarchies and then prove that they are equivalent to well-known hierarchies (such as the Drinfeld–Sokolov hierarchies but also others). The equivalence relation between integrable hierarchies is called a Miura transform [Miu68]; see §2.4.2.

The double ramification hierarchies (treated in Chapter 5). Recently, Alexander Buryak, inspired by ideas from symplectic field theory [EGH00], defined a new kind of integrable hierarchies he called the *double ramification hierarchies* [Bur15a]. They are constructed via the

¹⁰The semisimplicity means that an associated algebra is semisimple, i.e. the direct sum of simple algebras.

¹¹These cases are those where the Drinfeld–Sokolov hierarchy is built from the affine Lie algebra associated to a simple Lie algebra of type $A_{n \geq 1}$, $D_{n \geq 4}$ or E_6, E_7, E_8 (the *ADE* classification). In the remaining cases, the Dubrovin–Zhang and Drinfeld–Sokolov hierarchies are still equivalent in the sense of quasi-Miura (i.e. rational) transformations.

¹²Among them, the examples of mirror symmetry “CohFT / Int. hierarchy” we have cited so far: Intersection numbers of $\overline{\mathcal{M}}_{g,n}$ / Korteweg–de Vries hierarchy; Witten’s r -spin classes / Gelfand–Dickey hierarchies; Gromov–Witten theory of \mathbb{P}^1 / Extended Toda hierarchy.

¹³Dubrovin–Zhang construction actually applies to very much related objects called Frobenius manifolds, but we will not focus on this approach.

double ramification cycles on the moduli spaces of stable curves introduced by Graber and Vakil [GV05] (hence the name) and can be defined for arbitrary CohFTs (not only semisimple ones, as is in the Dubrovin–Zhang hierarchies). In that same article, guided by examples he worked out, Buryak conjectured that, in the semisimple case, the double ramification hierarchies are equivalent to the Dubrovin–Zhang hierarchies via a Miura transform. In particular, for the CohFT associated to the Fan–Jarvis–Ruan–Witten theory of a simple singularity of ADE type, the double ramification, Dubrovin–Zhang and Drinfeld–Sokolov hierarchies are conjectured to coincide. Since then, the double ramification hierarchies have been intensively studied by, for example but not only, A. Buryak, J. Guéré [BG16], P. Rossi [BR16a, BR16b, BGR17] and B. Dubrovin [BDGR16, BDGR18].

The double ramification hierarchies have several features that make them of particular interest. (1) They are defined for arbitrary CohFTs. (2) They are constructed as Hamiltonian equations for which the Poisson bracket is very simple (called the standard bracket); in particular, (3) tau symmetric densities are easily obtained [BDGR18]. (4) Their Hamiltonian densities satisfy strong recursion relations that are new to integrable hierarchies [BR16b]. (5) They admit a rather simple quantization [BR16a, BDGR16], but we will not discuss it here. Moreover, the conjectural equivalence with the Dubrovin–Zhang hierarchies has been enhanced into a uniquely determined Miura transform, hence *the strong DR/DZ conjecture 5.2.23* [BDGR18]. So far, the strong DR/DZ equivalence has been confirmed in several cases including the Drinfeld–Sokolov hierarchies of type A_1 to A_5 , the extended Toda hierarchy and others (see 5.2.23).

In [DR19], we confirmed the strong DR/DZ equivalence for the Drinfeld–Sokolov hierarchy associated to the affine algebra of the simple Lie algebra of type D_4 . This is the object of Chapter 5. The confirmation of the conjecture involves the computation of the first Hamiltonian structure of the Drinfeld–Sokolov hierarchy of type D_4 . To do so, we used the scalar Lax pair representation of the hierarchies of type D_n obtained by Liu, Wu and Zhang [LWZ10]. These computations, to the best of our knowledge, are new to the literature.

2.2 The Korteweg–de Vries hierarchy: a canonical example

In this section, we present some of the important features of the Korteweg–de Vries equation (KdV for short) and the KdV hierarchy. This will serve as an illustration of many of the properties of integrable hierarchies. Chapter 3 is dedicated to the polynomial tau functions of the KdV hierarchy; we describe the results we obtained in [Duc17] at the end of §2.2.

2.2.1 The Korteweg–de Vries equation

Aside from its application in geometry, the KdV equation arose from fluid mechanics. It describes the evolution of the surface of the water in a canal, the latter modelling shallow water flowing essentially in one direction. If one ignores the transverse direction, we can describe the height of the surface of the water by a scalar function $u(t, x)$ where t represents time and x

represents the one direction in which the water flows. Then u must satisfy the KdV equation:

$$u_t = \frac{1}{4}u_{xxx} + \frac{3}{2}uu_x, \quad (2.5)$$

where $u_t = \partial_t u$ and $u_x = \partial_x u$, etc. The above equation is actually a normalized version of KdV where physical constants have been ruled out. The KdV equation was first theorized by Boussinesq in 1877 [Bou77] and then proposed in its current form by Korteweg and de Vries in 1895 [KdV95]. For the physical derivation of the KdV equation (and that of the KP equation), we refer to the introduction of [AC91] by Ablowitz and Clarkson.

Notice that by rescaling u , x and t , we can put arbitrary coefficients in front of each monomial u_t , u_{xxx} and uu_x ; the choice 1, 1/4 and 3/2 will become clear in §2.2.2. The KdV equation can be written in many ways, thereby allowing connections to many a priori disconnected areas. Here we show four representations that we will use throughout the document: (a) the scalar Lax pair, (b) the matrix Lax pair, (c) the bi-Hamiltonian representation and (d) the bilinear representation.

(a) The scalar Lax pair. Consider the differential operator $L = \partial_x^2 + u$ often referred to as the Schrödinger operator.¹⁴ Then we can rewrite Equation (2.5) as

$$\frac{\partial L}{\partial t} = [P, L], \quad P = \partial_x^3 + \frac{3}{2}u\partial_x + \frac{3}{4}u_x, \quad (2.6)$$

where $[L, P] = L \circ P - P \circ L$ is the commutator between differential operators. The composition of operators is given by Leibniz's rule, defined on monomials by¹⁵

$$a\partial_x^n \circ b\partial_x^p := \sum_{k=0}^n \binom{n}{k} ab_{(k)}\partial_x^{n+p-k}, \quad (2.7)$$

where $b_{(k)} := \partial_x^k(b)$, and extended by linearity. We will omit the \circ sign when there is no ambiguity. The derivative ∂_t is defined to act on operators by differentiating each coefficient: $\partial_t(a\partial_x^n) := a_t\partial_x^n$. In particular, the action of ∂_t on operators satisfies Leibniz's rule. Therefore, $\partial_t L = u_t$, but this is due to the specific form L . The pair (L, P) is called a *scalar Lax pair* for the KdV equation. This representation is due to Peter D. Lax himself [Lax68].

Equations (2.5) and (2.6) are different in nature: Equation (2.5) is a PDE which is nonlinear in u , while Equation (2.6) is a linear ODE in L and both sides of the equation are derivations (Leibniz operators) on differential operators. The formula for the operator P , at this stage, seems rather ad hoc, but it will fit into a very clear and simple pattern with the KdV hierarchy. A particular interest of the scalar Lax pair is that it easily generalizes to the Kadomtsev–Petviashvili hierarchy (see Section 3.2).

¹⁴Indeed, if u is seen as a potential in quantum mechanics, then the normalized Schrödinger equation on a wave function ψ reads $i\hbar\psi_t = -\hbar^2\psi_{xx} + u\psi$.

¹⁵For the unaccustomed reader, the composition of differential operator given in Equation (2.7) simply comes from the action on test functions. For example, $\partial_x(f\partial_x(\psi)) = \partial_x(f\psi_x) = f_x\psi_x + f\psi_{xx}$, which agrees with the formula for $\partial_x \circ f\partial_x$ given by Equation (2.7).

(b) The matrix Lax pair. Now consider the following order 1 matrix-valued differential operator:

$$\mathcal{L} = \partial_x + \begin{pmatrix} 0 & \lambda - u \\ 1 & 0 \end{pmatrix}. \quad (2.8)$$

In the above, λ is a new indeterminate which we call the *spectral parameter*. Any identity should hold for any complex value of λ . Now consider the following matrix:

$$\mathcal{M} = \begin{pmatrix} -\frac{1}{4}u_x & \lambda^2 - \frac{1}{2}u\lambda - \frac{1}{2}(u)^2 - \frac{1}{4}u_{xx} \\ \lambda + \frac{1}{2}u & +\frac{1}{4}u_x \end{pmatrix}. \quad (2.9)$$

Then

$$\frac{\partial \mathcal{L}}{\partial t} - [\mathcal{L}, \mathcal{M}] = \begin{pmatrix} 0 & -u_t + \frac{1}{4}u_{xxx} + \frac{3}{2}uu_x \\ 0 & 0 \end{pmatrix}.$$

In the above equation, $[\mathcal{L}, \mathcal{M}] := \partial_x \mathcal{M} + [\mathcal{N}, \mathcal{M}]$ where $\mathcal{N} = \mathcal{L} - \partial_x$ and where the derivative $\partial_x \mathcal{M}$ is defined component-wise. Then $\partial_t \mathcal{L} = [\mathcal{L}, \mathcal{M}]$ if and only if $u_t = \frac{1}{4}u_{xxx} + \frac{3}{2}uu_x$, i.e. if and only if the KdV equation (2.5) holds. In other words, the KdV equation (2.5) can be rewritten as

$$\frac{\partial \mathcal{L}}{\partial t} = [\mathcal{L}, \mathcal{M}]. \quad (2.10)$$

It is equivalent to the zero-curvature equation

$$\partial_t \mathcal{N} - \partial_x \mathcal{M} - [\mathcal{N}, \mathcal{M}] = 0.$$

The pair $(\mathcal{L}, \mathcal{M})$ is called a *matrix Lax pair* for the KdV equation. This representation is due to Ablowitz, Kaup, Newell and Segur [AKNS74].

Once again, Equations (2.6) and (2.10) are different in nature; the latter only involves order 1 differential operators and therefore allows this zero-curvature representation. A particular interest is that the Drinfeld–Sokolov hierarchies are generically constructed as matrix Lax pair equations.

(c) The bi-Hamiltonian representation. We now introduce the ring $\mathcal{A} = \mathbb{C}[u, u_x, u_{xx}, \dots]$ whose elements we call *differential polynomials* in u . The operator ∂_x on \mathcal{A} identifies with the operator $\partial_x = \sum_{k \geq 0} u_{(k+1)} \partial_{u_{(k)}}$. Besides, we define the following quotient space,¹⁶

$$\mathcal{F} := \mathcal{A} / \text{Im}(\partial_x),$$

whose elements we call *local functionals*. Given a differential polynomial $f \in \mathcal{A}$ we denote its class $\bar{f} \in \mathcal{F}$ by $\bar{f} = \int f$.¹⁷ As an immediate consequence, we get an integration by parts

¹⁶Notice that the quotient space \mathcal{F} is only a vector space and not a ring since the operator ∂_x is linear yet not a homomorphism of algebras (it satisfies Leibniz's rule).

¹⁷We shall emphasize that the integral sign really is purely notational. Nonetheless, by construction $\int \partial_x f = 0$, and this can be realized by defining solutions on the circle S^1 . Hence we sometimes talk about periodic solutions of the KdV equation (see e.g. [Lax75]).

formula: $\int f \partial_x g = - \int g \partial_x f$, or, more generally for $k \geq 1$,

$$\int f \partial_x^k g = (-1)^k \int g \partial_x^k f. \quad (2.11)$$

From this we define on \mathcal{A} the operator $\delta/\delta u$, called the *variational derivative* (also called the Euler–Lagrange operator), by

$$\begin{aligned} \frac{\delta f}{\delta u} &:= \sum_{k \geq 0} (-1)^k \partial_x^k \frac{\partial f}{\partial u^{(k)}} \\ &= \frac{\partial f}{\partial u} - \partial_x \frac{\partial f}{\partial u_x} + \partial_x^2 \frac{\partial f}{\partial u_{xx}} - \dots \end{aligned}$$

Then for any f that contains no constant term, $\delta f/\delta u = 0$ if and only if $f \in \text{Im}(\partial_x)$ (see Proposition 5.1.1). In particular, $\delta/\delta u$ is well defined on \mathcal{F} . We then define two maps, $\{-, -\}_1, \{-, -\}_2 : \mathcal{F} \otimes \mathcal{F} \rightarrow \mathcal{F}$ by¹⁸

$$\{\bar{f}, \bar{g}\}_1 := \int \frac{\delta f}{\delta u} \partial_x \frac{\delta g}{\delta u}, \quad \{\bar{f}, \bar{g}\}_2 := \int \frac{\delta f}{\delta u} \left(\frac{1}{4} \partial_x^3 + u \partial_x + \frac{1}{2} u_x \right) \frac{\delta g}{\delta u}. \quad (2.12)$$

These maps are skew symmetric and satisfy Jacobi's identity; we call such maps *local Poisson structures*.¹⁹ Moreover, any linear combination $\lambda \{-, -\}_1 + \mu \{-, -\}_2$ is still a local Poisson structure; we say that the Poisson structures are compatible. We also define two auxiliary maps, denoted similarly, $\{-, -\}_a : \mathcal{A} \otimes \mathcal{F} \rightarrow \mathcal{A}$, with $a = 1, 2$, by

$$\{f, \bar{g}\}_1 := \sum_{k \geq 0} \frac{\partial f}{\partial u^{(k)}} \partial_x^{k+1} \frac{\delta g}{\delta u}, \quad \{f, \bar{g}\}_2 := \sum_{k \geq 0} \frac{\partial f}{\partial u^{(k)}} \partial_x^k \left(\frac{1}{4} \partial_x^3 + u \partial_x + \frac{1}{2} u_x \right) \frac{\delta g}{\delta u}. \quad (2.13)$$

By construction, and thanks to Equation (2.11), these maps satisfy $\int \{f, \bar{g}\}_a = \{\bar{f}, \bar{g}\}_a$, for $a = 1, 2$. Then it can be easily checked that the KdV equation (2.3) is equivalent to the following equations:

$$u_t = \{u, \bar{h}_1\}_1 = \partial_x \frac{\delta h_1}{\delta u}, \quad \bar{h}_1 = \int \frac{1}{4} u^3 + \frac{1}{8} u u_{xx}, \quad (2.14)$$

$$u_t = \{u, \bar{h}_2\}_2 = \left(\frac{1}{4} \partial_x^3 + u \partial_x + \frac{1}{2} u_x \right) \frac{\delta h_2}{\delta u}, \quad \bar{h}_2 = \int \frac{4}{8} u^2. \quad (2.15)$$

The local functionals \bar{h}_1 and \bar{h}_2 are called *Hamiltonian functionals* and the above equations are called *Hamiltonian representations* of the KdV equation. The fact that KdV admits two Hamiltonian representations with two compatible local Poisson structures is what we call a *bi-Hamiltonian representation*.

¹⁸In Equation (2.12), the differential polynomials $f, g \in \mathcal{A}$ are any representative of the classes $\bar{f}, \bar{g} \in \mathcal{F}$ (i.e. $\bar{f} = f + \text{Im}(\partial_x)$). This is well defined since $\delta/\delta u$ vanishes on $\text{Im}(\partial_x)$.

¹⁹Since \mathcal{F} is not a ring, we cannot state Leibniz's rule for local Poisson structures. Therefore, $(\mathcal{F}, \{-, -\}_a)$, for $a = 1, 2$, is a Lie algebra, but not a Poisson algebra. Nonetheless, for any local functional $\bar{g} \in \mathcal{F}$, the maps $f \mapsto \{f, \bar{g}\}_a$ of Equation (2.13) satisfy Leibniz's rule w.r.t. f .

Representation (2.14) was found simultaneously in 1971 by Gardner [Gar71] and Fadeev and Zakharov [FZ71] (although Fadeev and Zakharov attribute it to Gardner); it is called the first- or the Gardner–Zakharov–Fadeev Hamiltonian structure of KdV. Representation (2.15) is due to Magri [Mag78] and is called the second- or Magri’s Hamiltonian structure.²⁰

Bi-Hamiltonian representations were much studied by Dubrovin and Zhang (e.g. [DZ08, DLZ08]), especially in view of classifying integrable hierarchies. Moreover, the double ramification hierarchies, which we study in Chapter 5, naturally arise a Hamiltonian systems.

(d) The bilinear representation. Finally we describe the bilinear, or Hirota’s, representation of the KdV equation. Contrary to the previous representations, the bilinear one acts on the tau function τ rather than on the solution u . As mentioned above, there are different ways of defining the tau function, but in the case of the KdV equation, it always comes down to defining τ by the relation²¹

$$u = -2 \frac{\partial^2 \log \tau}{\partial x^2}. \quad (2.16)$$

In particular, given a solution u , tau function τ is not unique. We now introduce the bilinear, or Hirota’s, derivatives D_{t_i} , with $i \geq 1$. These derivatives are defined by the following relation: consider two smooth functions f, g of parameters $\mathbf{t} := \{t_i \mid i \geq 1\}$, we set²²

$$\exp \left(\sum_{i \geq 1} h_i D_{t_i} \right) (f, g) = f(\mathbf{t} + \mathbf{h}) \cdot g(\mathbf{t} - \mathbf{h}), \quad (2.17)$$

where both sides shall be expanded in powers of the h_i ’s and identified as series in the h_i ’s. For example, by setting $x = t_1$ and $t = t_3$ (the indices do not matter here), we get

$$D_x(f, g) = \frac{\partial f}{\partial x} g - f \frac{\partial g}{\partial x}, \quad D_x D_t(f, g) = \frac{\partial^2 f}{\partial x \partial t} g + f \frac{\partial^2 g}{\partial x \partial t} - \frac{\partial f}{\partial x} \frac{\partial g}{\partial t} - \frac{\partial f}{\partial t} \frac{\partial g}{\partial x}.$$

Then it is a bit lengthy to check, but the KdV equation (2.5), when substituting $u = -2 \partial_x^2 \log \tau$, is equivalent to the following equation,

$$(D_x^4 - 4D_x D_t)(\tau, \tau) = 0. \quad (2.18)$$

Equation (2.18) is called a Hirota equation, after Ryogo Hirota who introduced these derivatives [Hir76]. It is fairly compact considering the number of terms the logarithmic derivatives make appear.

One may naturally ask the interest of such equations. It is actually less clear for the KdV hierarchy than for other, more complicated hierarchies. In particular, when a solution of the

²⁰Magri actually brought much more to the field as he first discovered how to recover the entire hierarchy using the bi-Hamiltonian structure. This was an important step in understanding the richness of reconstruction techniques for integrable hierarchies. However, we will not use it in this document.

²¹The (-2) factor is explained in Proposition 3.4.5; it cannot be ruled out since the KdV equation is nonlinear.

²²Notice importantly that when expanding identity (2.17) the symbol $D_{t_{i_1}} \cdots D_{t_{i_n}}$ is an operator in itself and not a product of operators.

hierarchy consists of several functions u^1, \dots, u^N (and not just u), there is still one single tau function τ . Therefore, the bilinear representation condenses the whole hierarchy on one single function. Moreover, it can be the case that the tau function is easier to compute than the solution itself. This is, for instance, the approach of Chapters 3 and 4, although they do not involve directly the bilinear equations.

2.2.2 The Korteweg–de Vries hierarchy

The KdV hierarchy was introduced in 1968 simultaneously by Lax [Lax68] and Gardner–Kruskal–Miura–Zabusky [Miu68, GKM68, GKMZ70].²³ It consists of a sequence of equations that commute pairwise (explained below) and which contains the KdV equation (2.3), thus embedding the KdV equation within an infinite-dimensional system of PDEs. In what follows, we show how to construct the KdV hierarchy starting with (A) its scalar Lax pair, due to Lax [Lax68]; then we will discuss (B) its matrix Lax pair, (C) its bi-Hamiltonian representation and (D) its bilinear representation. The matrix Lax pair is the most technical one; we will actually discuss it in Section 2.3 along with the case of an arbitrary Drinfeld–Sokolov hierarchy.

(A) The scalar Lax pairs. For the scalar Lax pair, it suffices to consider the square root $L^{1/2}$ of the Schrödinger operator $L = \partial_x^2 + u$. For that, we first extend the algebra of differential operators into a larger algebra,

$$\mathcal{D} = \left\{ \sum_{i=-\infty}^n a^i \partial^i \mid a^i \in \mathbb{C}[[x]], n \in \mathbb{Z} \right\}, \quad (2.19)$$

equipped with the same product defined in Equation (2.7), but where the binomial coefficients have been generalized to $\binom{n}{k}$ with $n < 0$ by setting $\binom{n}{k} := n(n-1) \cdots (n-k+1)/k!$. The elements of \mathcal{D} we call *pseudo-differential operators*. In this larger algebra, there exists an operator $L^{1/2} = \partial_x + \sum_{k < 0} p^k \partial_x^k$, with functions $p^k \in \mathcal{A}$, which satisfies $(L^{1/2})^2 = L$ (see Theorem 3.1.8). Then the KdV hierarchy can be formulated as

$$\left\{ \frac{\partial L}{\partial t_i} = \left[(L^{i/2})_+, L \right] \mid i > 0 \text{ odd} \right\}, \quad (2.20)$$

where $(-)_+$ denotes the projection onto the subspace of \mathcal{D} containing only nonnegative powers of ∂_x (i.e. differential operators). In particular, $(L^{3/2})_+ = P = \partial_x^3 + \frac{3}{2}u\partial_x + \frac{3}{4}u_x$, see Equation (2.6). This scalar Lax formulation of the KdV hierarchy easily generalizes to the n -th Gelfand–Dickey hierarchy where the operator $L = \partial_x^2 + u$ is replaced by the operator $L = \partial_x^n + u^{n-2}\partial_x^{n-2} + \cdots + u^0$.

So, the flows defined by the system (2.20) commute pairwise in the sense that for any odd positive integers i, j ,

$$\frac{\partial}{\partial t_i} \left[(L^{j/2})_+, L \right] = \frac{\partial}{\partial t_j} \left[(L^{i/2})_+, L \right]. \quad (2.21)$$

²³It was actually first discovered by Andrew Lenard during the summer 1967, although he famously wrote no paper on the subject. See [PS05] for the details of that discovery and how it was acknowledged by Lax, Miura et al.

This is the *integrability condition* of the KdV hierarchy. The proof is classic and can be found for example in [DJM00]; we give a very similar proof for the KP hierarchy in Theorem 3.2.1. Keeping in mind that $L = \partial_x^2 + u$, it follows from Equation (2.21) that $\partial_{t_i} \partial_{t_j} u = \partial_{t_j} \partial_{t_i} u$. This is of course a necessary condition for the existence of at least C^1 solutions, let alone smooth solutions. Notice that the function u must now depend on infinitely many variables $\{t_i \mid i > 0 \text{ odd}\}$. We will generally assume that

$$u \in \mathbb{C}[[t_1, t_3, t_5, \dots]].$$

As we said, we will treat the matrix Lax pair in Section 2.3 along with the case of an arbitrary Drinfeld–Sokolov hierarchy. Let us now treat the bi-Hamiltonian representation.

(C) The bi-Hamiltonian representation. Given a pseudo-differential operator $A \in \mathcal{D}$, we denote by $\text{res}(A)$ its coefficient in ∂_x^{-1} . Then the bi-Hamiltonian representation is given by the two same local Poisson structures of Equation (2.12), and by the Hamiltonian densities given by the residues of the fractional powers of $L = \partial_x^2 + u$, namely,

$$\begin{aligned} u_{t_i} = \{u, \bar{h}_{1,i}\}_1 &= \partial_x \frac{\delta h_{1,i}}{\delta u}, & \bar{h}_{1,i} &= \frac{4}{i+2} \int \text{res } L^{\frac{i+2}{2}}, \\ u_{t_i} = \{u, \bar{h}_{2,i}\}_2 &= \left(\frac{1}{4} \partial_x^3 + u \partial_x + \frac{1}{2} u_x \right) \frac{\delta h_{2,i}}{\delta u}, & \bar{h}_{2,i} &= \frac{4}{i} \int \text{res } L^{\frac{i}{2}}, \end{aligned}$$

for $i > 0$ odd. In particular, comparing with Equations (2.14) and (2.15), $\bar{h}_{1,3} = \bar{h}_1$ and $\bar{h}_{2,3} = \bar{h}_2$. Moreover, the local Poisson structures can be reformulated using the operator L as follows. For any local functional $\bar{f} \in \mathcal{F}$, denote $\frac{\delta \bar{f}}{\delta L} = \partial_x^{-1} \circ \frac{\delta \bar{f}}{\delta u} \in \mathcal{A}$,²⁴ and let $X = \frac{\delta \bar{f}}{\delta L}$ and $Y = \frac{\delta \bar{g}}{\delta L}$, then

$$\{\bar{f}, \bar{g}\}_1 = \int \text{res} [XLY - LXY], \quad \{\bar{f}, \bar{g}\}_2 = \int \text{res} [(LX)_+ LY - (XL)_+ YL]. \quad (2.22)$$

(D) The bilinear representation. Finally, for the bilinear representation, we need to introduce the so-called *elementary Schur polynomials*, denoted $p_k(\mathbf{t})$ with $k \in \mathbb{Z}_{>0}$, defined by the relation

$$\exp \left(\sum_{i \in \mathbb{Z}_{>0}} t_i z^i \right) = \sum_{k \in \mathbb{Z}_{>0}} p_k(\mathbf{t}) z^k. \quad (2.23)$$

For instance, $p_1 = t_1$, $p_2 = \frac{1}{2} t_1^2 + t_2$, $p_3 = \frac{1}{6} t_1^3 + t_2 t_1 + t_3$. Now denote \tilde{p}_k the polynomial p_k where all even variables t_{2i} are set to 0. Then the KdV hierarchy (2.20), once we substitute $u = -2\partial_x^2 \log \tau$, is equivalent to the following two sequences of bilinear equations,²⁵

$$\left\{ \left\{ \begin{aligned} (\tilde{p}_{k+1}([D]) - \frac{1}{2} D_x D_{t_k})(\tau, \tau) &= 0, \\ (D_x D_{t_k} - D_x \tilde{p}_k([D]))(\tau, \tau) &= 0, \end{aligned} \right. \middle| k > 0 \text{ odd} \right\}, \quad (2.24)$$

²⁴The notation $\frac{\delta \bar{f}}{\delta L}$ seems a bit artificial here, but when one considers Drinfeld–Sokolov hierarchies for arbitrary simple Lie algebras (KdV being associated to $\mathfrak{sl}(2, \mathbb{C})$), it becomes clearer.

²⁵A much more compact form of Equations (2.24) is given by the celebrated bilinear identity (cf. §3.3), however it is less explicit than the one given here.

where $[D] := (D_x, \frac{1}{3}D_{t_3}, \frac{1}{5}D_{t_5}, \dots)$. The above can be deduced from results of [DJKM83] (see also [ACvM12], Lemma 4.1). In particular, for $k = 3$, $\tilde{p}_{k+1}([D]) - \frac{1}{2}D_x D_{t_k} = \frac{1}{24}(D_x^4 - 4D_x D_{t_3})$ so that we recover the bilinear form (2.18) of the KdV equation.

2.2.3 Polynomial tau functions of the KdV hierarchy and the Adler–Moser polynomials

Here we describe the result obtained in [Duc17] on the explicit transformation from the Adler–Moser polynomials to the polynomial tau functions of the KdV hierarchy; it is the object of Chapter 3, §3.5. In [AM78], Adler and Moser introduced a sequence of polynomials θ_n , for $n \geq 0$, depending on parameters $\mathbf{r} = \{r_i \mid i \in \mathbb{Z}_{>0}^{\text{odd}}\}$, $r_1 = x$, defined by the following differential recursion: $\theta_0 = 1$; $\theta_1 = x$; then

$$\frac{\partial \theta_{n+1}}{\partial x} \theta_{n-1} - \theta_{n+1} \frac{\partial \theta_{n-1}}{\partial x} = (2n-1)(\theta_n)^2. \quad (2.25)$$

When computing θ_n , the integration constant is set to r_{2n-1} . For example, $\theta_2 = x^3 + r_3$ and $\theta_3 = x^6 + 5r_3x^3 + r_5x - 5r_3^2$. These polynomials are now known as the Adler–Moser polynomials. In that same article, they proved that there exists a unique change of variables $\mathbf{r} \mapsto \mathbf{t}$ which transforms the Adler–Moser polynomial $\theta_n(\mathbf{r})$ into a tau function $\tau_n(\mathbf{t})$ of the KdV hierarchy in the sense of Equation (2.16). Moreover, from results of [AMcKM77], they proved that through this change of variables, one recovers *all* the polynomial tau functions of KdV as well as *all* the rational solutions of KdV, thus classifying the rational solutions. However, this unique change of variables was not known. In [Duc17], we proved the following theorem.

Theorem 2.2.1 ([Duc17]). *The following change of variables transforms the Adler–Moser polynomials into the polynomial tau functions of the KdV hierarchy: $r_1 = t_1 = x$, and*

$$\sum_{i \geq 2} \frac{r_{2i-1}}{\alpha_{2i-1}} z^{2i-1} = \tanh \left(\sum_{i \geq 2} t_{2i-1} z^{2i-1} \right). \quad (2.26)$$

where $\alpha_{2i-1} = (-1)^{i-1} 3^2 5^2 \dots (2i-3)^2 (2i-1)$.

See Theorem 3.5.4 in the document. We can interpret Theorem 2.2.1 as rescaling the variables, plus choosing a different, nontrivial integration constant for recursion (2.25). The first few tau functions read the following.

$$\begin{aligned} \tau_0 &= 1, & \tau_2 &= x^3 - 3t_3, \\ \tau_1 &= x, & \tau_3 &= x^6 - 15t_3x^3 - 45t_3^2 + 45t_5x. \end{aligned}$$

These are well-known tau functions of the KdV hierarchy, see e.g. [Hir04] where they are computed using similar methods and the elementary Schur polynomials (see Equation (2.23)).

The proof of Theorem 2.2.1 is given in Section 3.5; it relies essentially on KP theory which we develop in Section 3.2. It would constitute a further project to generalize the Adler–Moser polynomials as well as the result of [AM78] and the above theorem to arbitrary Drinfeld–Sokolov hierarchies (i.e. for arbitrary simple Lie algebras).

2.3 Polynomial tau functions of the Drinfeld–Sokolov hierarchies

Here we describe the results obtained in [CDY18] on polynomial tau functions of the Drinfeld–Sokolov hierarchies and applications; it is the object of Chapter 4, §4.5. We start by reviewing the construction of the Drinfeld–Sokolov hierarchies in the most straightforward, although technical way: the matrix Lax pair. In that way, they largely generalize representation (2.10) of the KdV equation. But first we need to introduce some material on semisimple and affine Lie algebras.

2.3.1 The Drinfeld–Sokolov hierarchies

Starting point: A semisimple Lie algebra. We start with a finite-dimensional semisimple Lie algebra \mathfrak{g} (see §4.1), i.e. a Lie algebra which is a direct sum of simple Lie algebras (i.e. algebras containing no nontrivial ideals). These algebras decompose nicely in the following way:

$$\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+,$$

where \mathfrak{h} is a maximal abelian subalgebra whose elements $H \in \mathfrak{h}$ are such that all the adjoint maps $\text{ad}_H = [H, -] : \mathfrak{g} \rightarrow \mathfrak{g}$ are simultaneously diagonalizable. The subalgebra \mathfrak{h} is called a *Cartan subalgebra* of \mathfrak{g} and its dimension $r = \dim \mathfrak{h}$ is called the *rank* of \mathfrak{g} ; the rank is unique. A classical result is that \mathfrak{g} is completely determined (up to isomorphism) by a unique matrix $(C_{ij}) \in \text{Mat}(r, \mathbb{Z})$, called its *Cartan matrix*, to which we can associate a set of generators $\{H_i, E_i, F_i \mid 1 \leq i \leq r\}$ of \mathfrak{g} satisfying the following relations:

$$[E_i, F_j] = \delta_{ij} H_i, \quad [H_i, E_j] = C_{ij} E_j, \quad [H_i, F_j] = -C_{ij} F_j, \quad [H_i, H_j] = 0,$$

Then \mathfrak{h} is generated by the H_i 's, while \mathfrak{n}_+ (resp. \mathfrak{n}_-) is generated by the E_i 's (resp. the F_i 's). Such generators are called *Weyl generators*. There are two special elements $E_\theta \in \mathfrak{n}_+$ and $E_{-\theta} \in \mathfrak{n}_-$ which satisfy the following relations: for all $i \in \{1, \dots, r\}$,

$$[E_\theta, E_i] = 0, \quad [E_{-\theta}, F_i] = 0.$$

For example, the simple Lie algebra $\mathfrak{sl}(2, \mathbb{C})$ (2×2 complex traceless matrices) has rank 1 and is generated by the following matrices,

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

They satisfy the relations $[H, E] = 2E$, $[H, F] = -2F$ and $[E, F] = H$ so that the Cartan matrix of $\mathfrak{sl}(2, \mathbb{C})$ is $C = (2) \in \text{Mat}(1, \mathbb{Z})$. Moreover, in this particular example, $E_\theta = E$ and $E_{-\theta} = F$.

The associated affine algebra. To the finite-dimensional semisimple Lie algebra \mathfrak{g} we associate the so-called (untwisted) affine algebra

$$\mathfrak{g} := \mathring{\mathfrak{g}} \otimes \mathbb{C}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c, \quad (2.27)$$

which is infinite-dimensional. We equip this algebras with the following Lie bracket: given $X, Y \in \mathring{\mathfrak{g}}$, $p, q \in \mathbb{Z}$ and $a, b \in \mathbb{C}$,

$$[X\lambda^p + ac, Y\lambda^q + bc] = [X, Y]\lambda^{p+q} + p\delta_{p+q,0}(X, Y)_0 \cdot c, \quad (2.28)$$

where $(-, -)_0$ denotes a certain nondegenerate symmetric bilinear form on $\mathring{\mathfrak{g}}$. Clearly, c belongs to the center of the algebra \mathfrak{g} (i.e. c commutes with all elements of \mathfrak{g}), making $\mathbb{C}c$ into a nontrivial ideal of \mathfrak{g} . If the semisimple Lie algebra \mathfrak{g} has type X_n , we will denote by $X_n^{(1)}$ the associated affine algebra \mathfrak{g} . Affine algebras were much studied especially by V. Kac, who introduced them simultaneously with R. Moody [Kac68, Moo68]. The details are explained in §4.2.

The affine algebra \mathfrak{g} is generated by the following elements:

$$\begin{aligned} e_i &= E_i, & f_i &= F_i, & \check{\alpha}_i &= H_i, \\ e_0 &= E_{-\theta}\lambda, & f_0 &= E_{\theta}\lambda^{-1}, & \check{\alpha}_0 &= H_0 + c, \end{aligned}$$

where $H_0 = [E_{\theta}, E_{-\theta}]$, with $i \in \{1, \dots, r\}$ ($r = \dim \mathfrak{h}$) being the rank of $\mathring{\mathfrak{g}}$. From these we define two important gradations on \mathfrak{g} , the *principal gradation* and the *homogeneous gradation*, as follows:

$$\begin{aligned} \text{Principal:} & \quad \deg e_i = 1, & \deg f_i &= -1, & \deg \check{\alpha}_i &= 0, & \mathfrak{g} &= \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}^k, \\ \text{Homogeneous:} & \quad \deg e_i = \delta_{i,0}, & \deg f_i &= -\delta_{i,0}, & \deg \check{\alpha}_i &= 0, & \mathfrak{g} &= \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_k. \end{aligned}$$

The homogeneous gradation, as the name suggests, amounts to setting $\deg \lambda = 1$ and $\deg \mathring{\mathfrak{g}} = \deg c = 0$. We denote by \mathfrak{g}^+ (resp. \mathfrak{g}_+) the subspace of elements of principal (resp. homogeneous) degree ≥ 0 ; we set $\mathfrak{g}^- = \mathfrak{g} - \mathfrak{g}^+$ and $\mathfrak{g}_- = \mathfrak{g} - \mathfrak{g}_+$.

We now introduce the so-called *affine principal cyclic element*

$$\Lambda = e_0 + e_1 + \dots + e_r = E_{-\theta}\lambda + E_1 + \dots + E_r. \quad (2.29)$$

In particular, $\Lambda \in \mathfrak{g}^1$ (principal degree 1). A fundamental result (see [Kac90]) on affine algebras is that \mathfrak{g} decomposes as follows,

$$\mathfrak{g} = \mathfrak{s} + \text{Im}(\text{ad}_{\Lambda}), \quad \mathfrak{s} \cap \text{Im}(\text{ad}_{\Lambda}) = \mathbb{C}c.$$

where $\mathfrak{s} = \text{Ker}(\text{ad}_{\Lambda})$. We denote by $\mathfrak{s}^k = \mathfrak{s} \cap \mathfrak{g}^k$ the subspace of principal degree k . We call the *affine exponents* of \mathfrak{g} the integers $k \in \mathbb{Z}$ such that $\dim \mathfrak{s}^k \neq 0$ and denote $m_k = \dim \mathfrak{s}^k$ called the *multiplicity* of exponent k . We denote by

$$\mathbb{E} = \left\{ k \in \mathbb{Z} \mid \dim \mathfrak{s}^k \neq 0 \right\}$$

the multiset of affine exponents of \mathfrak{g} , meaning that elements $k \in \mathbb{E}$ have multiplicity m_k .

We denote $\Lambda_k, \Lambda'_k, \Lambda''_k, \dots$ (as many as m_k) the spanning elements of \mathfrak{s}^k . These generators Λ_k (we omit the primes for now) can be chosen in such a way that they satisfy the Heisenberg relations

$$[\Lambda_i, \Lambda_j] = i\delta_{i+j,0} \cdot c. \quad (2.30)$$

We call \mathfrak{s} the principal Heisenberg subalgebra of \mathfrak{g} . We denote $\nu \in \mathbb{C} \setminus \{0\}$ the complex number such that $\Lambda_1 = \nu\Lambda$.

For example, for the affine algebra $\mathfrak{g} = \mathfrak{sl}(2, \mathbb{C}) \otimes \mathbb{C}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c$ the principal affine cyclic element reads

$$\Lambda = F\lambda + E = \begin{pmatrix} 0 & \lambda \\ 1 & 0 \end{pmatrix}.$$

The affine exponents of \mathfrak{g} are exactly the odd integers and all have multiplicity 1, i.e. $\mathbb{E} = \mathbb{Z}^{\text{odd}}$. The generators of the principal Heisenberg subalgebra simply read $\Lambda_{2i+1} = \Lambda^{2i+1} = \lambda^i \Lambda$ for $i \in \mathbb{Z}$.

The Drinfeld–Sokolov hierarchies. We now have all the material we need to describe the Drinfeld–Sokolov hierarchies. First, we denote by $\mathbb{E}^+ = \mathbb{E} \cap \mathbb{Z}_{>0}$ the set of positive exponents of \mathfrak{g} . Then we denote $\mathbf{t} := \{t_i \mid i \in \mathbb{E}^+\}$, with $t_1 = \nu^{-1}x$, and we consider the following algebra,

$$\mathfrak{g} \otimes \mathbb{C}[[\mathbf{t}]] = \left\{ \sum_{m \geq 0} \sum_{i_1, \dots, i_m \in \mathbb{E}^+} \sum_{k_1, \dots, k_m \geq 1} \zeta_{i_1, \dots, i_m}^{k_1, \dots, k_m} \cdot t_{i_1}^{k_1} \dots t_{i_m}^{k_m} \mid \zeta_{i_1, \dots, i_m}^{k_1, \dots, k_m} \in \mathfrak{g} \right\}.$$

The above algebra (and its subalgebras) will serve as the “function space” of the theory. The Lie bracket on it is defined similarly to that on $\mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}]$, i.e. we set $[\xi t_i, \eta t_j] = [\xi, \eta] t_i t_j$ and then extend by multiplicativity. Now consider the following degree 1 matrix-valued operator:

$$\mathcal{L} = \partial_x + \Lambda + Q, \quad Q \in \mathring{\mathfrak{b}}[[\mathbf{t}]],$$

where $\mathring{\mathfrak{b}} = \mathring{\mathfrak{h}} \oplus \mathring{\mathfrak{n}}_+$ (typically, for $\mathfrak{sl}(n, \mathbb{C})$, the subalgebra $\mathring{\mathfrak{b}}$ consists of lower triangular traceless matrices). Such operators we call *connections* (even though there is no manifold or fibre bundle per say) and denote the set of such connections $\mathcal{L} \in \text{op}_{\mathfrak{g}}$.²⁶ We then consider the following multiplicative group:

$$N := \{ \exp(\text{ad}_S) \mid S \in \mathring{\mathfrak{n}}[[\mathbf{t}]] \}$$

and make it act on $\text{op}_{\mathfrak{g}}$ via

$$e^{\text{ad}_S} \mathcal{L} = \partial_x + e^{\text{ad}_S}(\Lambda + Q) - \nabla_S S, \quad \nabla_S S := \sum_{k \geq 0} \frac{1}{(k+1)!} (\text{ad}_S)^k \partial_x(S).$$

²⁶Equation (2.8) gives an example of the connection \mathcal{L} with $\mathring{\mathfrak{g}} = \mathfrak{sl}(2, \mathbb{C})$, although for this one we even chose $Q \in \mathring{\mathfrak{n}}_+[[\mathbf{t}]]$ since for $\mathfrak{sl}(2, \mathbb{C})$, we have $\mathring{\mathfrak{n}}_+ = \mathbb{C}E$.

The reasons for defining this specific action and the well-definedness of the latter are explained in §4.3.1. Then it is not hard to show that, indeed, for any $\mathcal{L} \in \text{op}_{\mathfrak{g}}$, we have $e^{\text{ads}} \mathcal{L} \in \text{op}_{\mathfrak{g}}$. The Drinfeld–Sokolov hierarchies are defined on equivalence classes of connections $\mathcal{L} \in \text{op}_{\mathfrak{g}}$, rather than on the connection itself. Now we need to extend our function space in the following way:

$$\widehat{\mathfrak{g}} := \mathfrak{g}^+ \oplus \widehat{\mathfrak{g}}^-, \quad \mathfrak{g}^+ = \mathfrak{g}^{\geq 0} = \bigoplus_{k \geq 0} \mathfrak{g}^k, \quad \widehat{\mathfrak{g}}^- = \prod_{k < 0} \mathfrak{g}^k.$$

Through this we naturally consider $\widehat{\mathfrak{g}}[[\mathbf{t}]]$ and its subalgebras. We also imply similar definitions for \mathfrak{g}_+ and $\widehat{\mathfrak{g}}_-$.

A fundamental result that was proven by Drinfeld and Sokolov [DS84] (here we give a slightly different version due to Wu [Wu17]) is that there exists a unique function $U \in \widehat{\mathfrak{g}}^-[[\mathbf{t}]]$ such that the connection $\mathcal{L}_0 := e^{-\text{ad}U} \mathcal{L}$ takes the form

$$\mathcal{L}_0 = \partial_x + \Lambda + H, \quad H \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]], \quad (2.31)$$

and such that $(e^{\text{ad}U} \xi)_c = 0$ for all $\xi \in \mathfrak{s}^+$, where $(-)_c$ denotes the component along $\mathbb{C}c$ in the decomposition (2.27) of \mathfrak{g} . The components of the functions U and H are computed recursively. The key point here is that, thanks to the Heisenberg relations (2.30), the algebra $\widehat{\mathfrak{s}}^-[[\mathbf{t}]]$ is abelian! We can now formulate the matrix Lax representation of the Drinfeld–Sokolov hierarchies. Given a connection $\mathcal{L} \in \text{op}_{\mathfrak{g}}$, the Drinfeld–Sokolov hierarchy of type \mathfrak{g} (or the \mathfrak{g} -DS hierarchy) consists of the following system of equations:

$$\left\{ \frac{\partial \mathcal{L}}{\partial t_i} = [(e^{\text{ad}U} \Lambda_i)_+, \mathcal{L}] \mid i \in \mathbb{E}^+ \right\}. \quad (2.32)$$

As suggested above, if \mathcal{L} is a solution of the hierarchy, then so does $e^{\text{ads}} \mathcal{L}$ for any $S \in \mathfrak{n}[[\mathbf{t}]]$. If the multiplicity of exponent $i \in \mathbb{E}^+$ is greater than 1, then we denote t_i, t'_i, t''_i, \dots the associated variables. It is essentially the abelianity of $\widehat{\mathfrak{s}}^-[[\mathbf{t}]]$ that makes the Drinfeld–Sokolov equations commute pairwise and thus integrable. The right-hand side of these equations are matrices whose components are differential polynomials in the components of the matrix Q . Hence these equations eventually take the form of Equation (2.2). Finally, we easily show that the components of H are conserved densities in the sense that for any $j \in \mathbb{E}^+$,

$$\int \frac{\partial H}{\partial t_j} = 0.$$

For example, the KdV hierarchy (2.20) is equivalent to the system (2.32), with the following data: $\mathbb{E}^+ = \mathbb{Z}_{>0}^{\text{odd}}$ and

$$\mathcal{L} = \partial_x + \Lambda + Q, \quad \Lambda = \begin{pmatrix} 0 & \lambda \\ 1 & 0 \end{pmatrix}, \quad Q = \begin{pmatrix} 0 & -u \\ 0 & 0 \end{pmatrix}, \quad \Lambda_{2i+1} = \lambda^i \Lambda = \begin{pmatrix} 0 & \lambda^{i+1} \\ \lambda^i & 0 \end{pmatrix}.$$

2.3.2 Tau functions as Toeplitz determinants

Here we explain the construction of a generic tau function for arbitrary Drinfeld–Sokolov hierarchies by C.-Z. Wu [Wu17]. Then we explain how to realize these tau functions as the limit of large size Toeplitz determinants by Cafasso and Wu [CW15a, CW15b]. Finally, we describe our results of [CDY18] on the computation and applications of polynomial tau functions based on the works of Cafasso and Wu.

Tau functions of the Drinfeld–Sokolov hierarchies. In [Wu17], C.-Z. Wu defined a generic tau function for an arbitrary \mathfrak{g} -DS hierarchy as follows. First we introduce a function $\Omega = \sum_{i \in \mathbb{E}^+} \frac{\Omega^i}{i} \Lambda_{-i}$ defined by $\partial_x \Omega = -H$, where $H \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]]$ is defined in Equation (2.31). Then the functions Ω^j are *not* differential polynomials in the components of Q , however the functions $\partial_{t_i} \Omega^j$ are and they satisfy the closure condition $\partial_{t_i} \Omega^j = \partial_{t_j} \Omega^i$. By setting $H = \sum_{i \in \mathbb{E}^+} -i^{-1} \tilde{h}^i \Lambda_{-i}$, we get

$$\frac{\partial \tilde{h}^i}{\partial t_j} = \frac{\partial \tilde{h}^j}{\partial t_i}. \quad (2.33)$$

The above equation is called the *tau symmetry* condition on the conserved densities \tilde{h}_i . From Poincaré lemma, it follows that we can define a function $\tau \in \mathbb{C}[[\mathbf{t}]]$ such that²⁷

$$\frac{\partial \log \tau}{\partial x \partial t_j} = \tilde{h}^j. \quad (2.34)$$

The function τ we call the *tau function*. Wu’s tau function is generic in the sense that its construction works for arbitrary affine algebras (previous definitions where somehow case-by-case). Now by considering the formal power series $\Theta = e^U e^\Omega$ (we show that is it well defined), we can redefine the tau function by the relation

$$\frac{\partial \log \tau}{\partial t_j} = (\Theta \Lambda_j \Theta^{-1})_c \quad (2.35)$$

Tau functions as Toeplitz determinants. In [CW15a, CW15b], Cafasso and Wu showed how to express the tau function τ associated to the function Θ as the limit of large size *Toeplitz determinants*. First of all, we fix a faithful traceless representation²⁸

$$\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(n)$$

²⁷Equation (2.34) is equivalent to setting $\partial_{t_j} \log \tau = -\Omega^j$. We shall emphasize that since the Ω^j ’s are *not* differential polynomials, it does *not* mean that the \tilde{h}^j ’s are exact derivatives of differential polynomials. In particular, Equation (2.34) does *not* imply that $\int \tilde{h}^j = 0$.

²⁸By $\mathfrak{gl}(n)$ we mean $\text{Mat}(n, \mathbb{C})$ ($n \times n$ complex matrices) equipped with Lie bracket defined by the commutator.

as in Appendix A. Toeplitz matrices are infinite matrices of size “ $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ ” that are constant along diagonals, i.e. of the form

$$\mathbf{T}(\gamma) = (\gamma_{I-J})_{I,J \geq 0} = \begin{pmatrix} \gamma_0 & \gamma_{-1} & \gamma_{-2} & & \\ \gamma_1 & \gamma_0 & \gamma_{-1} & \ddots & \\ \gamma_2 & \gamma_1 & \gamma_0 & \ddots & \\ & \ddots & \ddots & \ddots & \end{pmatrix},$$

where the γ_I 's are certain coefficients that, in our case, we set to be elements of $\mathfrak{gl}(n)[[\mathbf{t}]]$. We gather these coefficients in a power series $\gamma = \sum_{I \in \mathbb{Z}} \gamma_I \lambda^I \in \mathfrak{gl}(n)[[\lambda, \lambda^{-1}, \mathbf{t}]]$ (same λ as before). Of course it is not clear that the determinants of $\mathbf{T}(\gamma)$ can be defined and it is a technical aspect of this approach. To do so, we consider the truncated Toeplitz matrices $\mathbf{T}_N(\gamma) = (\gamma_{I-J})$ with $I, J \in \{1, \dots, N\}$.

In §4.4 we explain the following result of [CW15a, CW15b]. Consider an arbitrary element $X \in \mathfrak{g}_-$ and the following function that we call the *matrix Xi function*,

$$\Xi(\mathbf{t}; \lambda) := \sum_{j \in \mathbb{E}^+} t_j \Lambda_j \in \mathfrak{s}^+ [[\mathbf{t}]].$$

Then (1) the function $J_X(\mathbf{t}) = e^{\Xi(\mathbf{t}; \lambda)} e^X$ is a well defined power series in $\mathfrak{gl}(n)[[\lambda, \lambda^{-1}, \mathbf{t}]]$;²⁹ (2) the limit $\lim_{N \rightarrow \infty} \det \mathbf{T}_N(J_X(\mathbf{t}))$ is a well defined power series in $\mathbb{C}[[\mathbf{t}]]$ and (3) the function

$$\tau_X(\mathbf{t}) := \left[\lim_{N \rightarrow \infty} \det \mathbf{T}_N(J_X(\mathbf{t})) \right]^\kappa \quad (2.36)$$

is a tau function of the \mathfrak{g} -DS hierarchy in the sense of Equation (2.34). Here κ is a number associated to the representation $\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(n)$ by $(Y, Z)_0 = \kappa \cdot \text{tr}(\pi(Y)\pi(Z))$ for all $Y, Z \in \mathfrak{g}$ (with $(-, -)_0$ the standard bilinear form on \mathfrak{g}). Typically, for simple Lie algebras of type A_n and C_n , we have $\kappa = 1$ and for types B_n and D_n we have $\kappa = 1/2$.

More precisely, the function Θ of Equation (2.34) to which the tau function $\tau_X(\mathbf{t})$ is associated is the one such that $\Theta|_{\mathbf{t}=0} = e^X$. It can be easily shown that for any $X \in \mathfrak{g}_-$, there exists such a function Θ ; see Lemma 4.4.2.

2.3.3 Polynomial tau functions and applications

We now explain the results of [CDY18] on polynomial tau functions of the Drinfeld–Sokolov hierarchies and applications; see Chapter 4, §4.5. We choose an element $X \in \mathfrak{g}_-$ and the associated tau function $\tau_X(\mathbf{t})$ of Equation (2.36). Our first point is a series expansion of $\tau_X(\mathbf{t})$ over integer partitions (or equivalently, over Young diagrams); we denote by \mathbb{Y} the set of integer partitions. To do so, we need to introduce the *Laurent matrix* $\mathbf{L}(\gamma)$, which is also a diagonally

²⁹All issues of convergence are treated in a purely algebraic way using the valuation topology. This is explained at the beginning of §4.4

constant matrix, yet of size “ $\mathbb{Z} \times \mathbb{Z}$,” namely,

$$\mathbf{L}(\gamma) = (\gamma_{I-J})_{I,J \in \mathbb{Z}} = \begin{pmatrix} \ddots & \ddots & \ddots & & & & \\ & \ddots & \gamma_0 & \gamma_{-1} & \gamma_{-2} & & \\ & \ddots & \gamma_1 & \gamma_0 & \gamma_{-1} & \ddots & \\ & & \gamma_2 & \gamma_1 & \gamma_0 & \ddots & \\ & & & \ddots & \ddots & \ddots & \\ & & & & \ddots & \ddots & \ddots \end{pmatrix}.$$

So, consider the Laurent matrices $r := \mathbf{L}(e^X)$ and $s(\mathbf{t}) := \mathbf{L}(e^{\Xi(\mathbf{t}; \lambda)})$ along with the following determinants³⁰

$$r_\mu := \det(r_{i-\mu_i-1; j-1})_{i,j \in [1, \ell(\mu)]} \in \mathbb{C}, \quad s_\mu(\mathbf{t}) := \det(s_{i-1; j-\mu_j-1})_{i,j \in [1, \ell(\mu)]} \in \mathbb{C}[\mathbf{t}], \quad (2.37)$$

where $\mu \in \mathbb{Y}$ is an integer partition of length $\ell(\mu)$. We call r_μ the *Plücker coordinates* of e^X and we call $s_\mu(\mathbf{t})$ the *generalized Schur polynomials* of type (\mathfrak{g}, π) . In the case where $\mathfrak{g} = \mathfrak{sl}(n, \mathbb{C})$ and π is the fundamental representation, the polynomials $s_\mu(\mathbf{t})$ coincide with the traditional Schur polynomials [Mac95], up to setting $t_k = 0$ whenever $k \equiv 0 \pmod n$. We proved the following theorem. It shall be seen as an application of the well-known Cauchy–Binet formula as well as a generalization of Sato’s seminal result on the tau functions of the KP hierarchy [Sat81].

Theorem 2.3.1 ([CDY18]). *For any $X \in \mathfrak{g}_{<0}$, the tau function of Equation (2.36) admits the following expansion in terms of integer partitions:*

$$\tau_X(\mathbf{t})^{\frac{1}{\kappa}} = \sum_{\mu \in \mathbb{Y}} s_\mu(\mathbf{t}) \cdot r_\mu = 1 + s_{\square} r_{\square} + s_{\square\square} r_{\square\square} + s_{\square\square\square} r_{\square\square\square} + \cdots \quad (2.38)$$

See Theorem 4.5.2 in the document. From Equation (2.38) we find an easy criterion for polynomiality of the tau function as follows. If the initial data $X \in \mathfrak{g}_-$ is such that $\pi(X)$ is a nilpotent matrix, then the matrix components of $r = \mathbf{L}(e^X)$ vanish as we move further away from the main diagonal. Consequently, for partitions μ of a higher length, the Plücker coordinates r_μ of X eventually vanish. This proves that

$$\forall X \in \mathfrak{g}_-, \pi(X) \text{ is a nilpotent matrix} \implies \tau_X(\mathbf{t})^{\frac{1}{\kappa}} \text{ is polynomial.}$$

In §4.5.2 we give several examples of polynomial tau functions for the Drinfeld–Sokolov hierarchies associated to the simple Lie algebras of type A_1, A_2, B_2 and D_4 . For the A_1 case—i.e. $\mathfrak{sl}(2, \mathbb{C})$, i.e. the KdV hierarchy—all the tau functions we computed were, up to change of variables, the Adler–Moser polynomials which we discussed in the section 2.2.3 of this introduction. Although it has not been proven yet, we believe that our algorithm recovers *all* the Adler–Moser polynomials (up to the right change of variables). This shall be showed using the degree of these polynomials. For the remaining cases A_2, B_2, D_4 , the tau functions become rapidly fairly nontrivial.

³⁰Very much importantly, in Equation (2.37), the indices i, j refer to the actual complex entries of the Laurent matrices r and $s(\mathbf{t})$. They do not refer to the matrix entries which were denoted by capital indices I, J so far.

It will constitute a future project to study the properties of these polynomials. In particular, seeing these polynomials as polynomials in x with parameters $\mathbf{t}^* = \{t_i \mid i \neq 1\}$, it should be investigated how the zeros of these polynomials evolve w.r.t. \mathbf{t}^* . As noted in §2.1, in the A_1 case, the zeros of the Adler–Moser polynomials evolve through time so as to form a system of Calogero–Moser type [AMcKM77, Wil98]. This is yet to be understood for arbitrary Drinfeld–Sokolov hierarchies.

Polynomial tau functions and bilinear equations. Besides being the simplest tau functions, the polynomial tau functions have the merit of being a finite expression and are therefore much easier to use in practical computations. As an application, we can detect possible bilinear equation (i.e. similar to Equation (2.24)) that the polynomial tau functions of the hierarchies must satisfy. The bilinear form of the Drinfeld–Sokolov hierarchies are known explicitly for all the Gelfand–Dickey hierarchies (i.e. for $\mathfrak{sl}(n, \mathbb{C})$), see Equation (4.97). However, to the best of our knowledge, such explicit bilinear equations for arbitrary Drinfeld–Sokolov hierarchies are not pointed in the literature. Although they must be deducible from results of the seminal series of paper “Transformation groups for soliton equations” [KM81]–[DJKM83].

So, we plugged the polynomial tau functions we computed in ansatz of bilinear equations and ask for the equations to return 0. With polynomial tau functions of sufficiently high degree, this determined the coefficients of said bilinear equations. Of course, it does not mean that all tau functions must satisfy the bilinear equations we found. However, we believe it gives a good hint on who these equations must be.

In what follows, $\text{HDS}_{X_n}^{[\text{deg} \leq g]}$ denotes the vector space of bilinear equations of degree $\leq g$ that the tau functions of the DS hierarchy of type X_n (A_n or B_n , etc.) satisfy. The degree of Hirota derivatives (introduced in Equation (2.17)) is simply set as $\text{deg } D_i = i$ where $D_i = D_{t_i}$. We proved the following propositions; first for the B_2 case.

Proposition 2.3.2. *The following dimension estimates hold true*

$$\dim_{\mathbb{C}} \text{HDS}_{B_2}^{[\text{deg} \leq 4]} = 0, \quad \dim_{\mathbb{C}} \text{HDS}_{B_2}^{[\text{deg} \leq 6]} \leq 1, \quad \dim_{\mathbb{C}} \text{HDS}_{B_2}^{[\text{deg} \leq 8]} \leq 2.$$

Moreover, the only possible elements in $\text{HDS}_{B_2}^{[\text{deg} \leq 8]}$ are linear combinations of the following,

$$\begin{aligned} (D_1^6 - 5D_1^3 D_3 - 5D_3^2 + 9D_1 D_5)(\tau, \tau) &= 0, \\ (D_1^8 + 7D_1^5 D_3 - 35D_1^2 D_3^2 - 21D_1^3 D_5 - 42D_3 D_5 + 90D_1 D_7)(\tau, \tau) &= 0. \end{aligned}$$

The above equations can be found in [DKM81] as the two first nontrivial equations of the bilinear formulation of the BKP hierarchy. The latter contains all the hierarchies of type B_n and D_n (the special orthogonal types). In a similar way the KP hierarchy contains all the hierarchies of type A_n .

We also have the following proposition for the DS hierarchy of type D_4 . Notice that for the D_4 case, we have two variables associated to the exponent 3: t_3 and t'_3 . This is due to the fact that the Heisenberg subalgebra \mathfrak{s} of type D_n has subspaces of multiplicity 2 for all exponents in $3\mathbb{Z}^{\text{odd}}$.

Proposition 2.3.3. *The following dimension estimates hold true*

$$\dim_{\mathbb{C}} \text{HDS}_{D_4}^{[\text{deg} \leq 4]} = 0, \quad \dim_{\mathbb{C}} \text{HDS}_{D_4}^{[\text{deg} \leq 6]} \leq 3.$$

Moreover, the only possible elements in $\text{HDS}_{D_4}^{[\text{deg} \leq 6]}$ are linear combinations of

$$\begin{aligned} (2D_1^3 D_{3'} + 4D_3 D_{3'} - 3D_3^2)(\tau, \tau) &= 0, \\ (D_1^3 D_3 - D_1^3 D_{3'} + D_3 D_{3'} - D_3^2)(\tau, \tau) &= 0, \\ (D_1^6 + 9D_1 D_5 - 10D_1^3 D_3 + 5D_1^3 D_{3'} - 5D_3 D_{3'})(\tau, \tau) &= 0. \end{aligned}$$

2.4 Hierarchy of type D_4 and the strong DR/DZ equivalence

Finally, we describe our results on the confirmation of the strong Double Ramification / Dubrovin–Zhang equivalence (DR/DZ); a work accepted for publication [DR19]. It is the object of Chapter 5 we will briefly explain the concept of cohomological field theories (CohFT), the example of Fan–Jarvis–Ruan–Witten theory (FJRW) and the construction of the Dubrovin–Zhang hierarchies.

2.4.1 Cohomological field theories and the Dubrovin–Zhang hierarchies

Moduli spaces of stables curves. Essentially speaking, the moduli space $\overline{\mathcal{M}}_{g,n}$ is the space that parameterizes isomorphism classes of compact, possibly nodal, complex curves C with genus g together with n points on it $x_1, \dots, x_n \in C$ which we call *markings*. Markings cannot be nodes though and must be distinct from each other. That allows to define two such genus g marked curves $(C, (x_1, \dots, x_n))$ and $(C', (x'_1, \dots, x'_n))$ to be isomorphic if there exists a biholomorphism $\phi : C \rightarrow C'$ such that $\phi(x_i) = x'_i$. We say that such a curve is *stable* if its group of automorphism is finite. The stability assumption is necessary in order to have a well defined moduli space; it translates into the following condition,

$$2g - 3 + n \geq 0.$$

We define the moduli space of stable genus g curves with n marked points as the following quotient space,³¹

$$\overline{\mathcal{M}}_{g,n} := \{(C, (x_1, \dots, x_n))\} / \sim$$

It is a compact complex orbifold, i.e., roughly speaking, a topological space that is locally homeomorphic an affine Euclidean vector space quotiented by the action of a finite group. As an orbifold, it has a well defined complex dimension

$$\dim_{\mathbb{C}} \overline{\mathcal{M}}_{g,n} = 3g - 3 + n.$$

³¹The bar on $\overline{\mathcal{M}}_{g,n}$ comes from the fact that it is actually the compactification of another space, $\mathcal{M}_{g,n}$, which parameterizes the same curves yet not nodal. The latter is not compact unfortunately.

This orbifold comes with several important maps and bundles. The details can be found in §5.2.1; we refer to [Zvo12]. First of all, the symmetric group S^n act on $\overline{\mathcal{M}}_{g,n}$ by permuting the marked points. Then we have three natural maps:

- ◊ The forgetful map, $\pi : \overline{\mathcal{M}}_{g,n+1} \rightarrow \overline{\mathcal{M}}_{g,n}$, which forgets the last marked point;
- ◊ The separating attaching map, $\sigma : \overline{\mathcal{M}}_{g_1,n_1+1} \times \overline{\mathcal{M}}_{g_2,n_2+1} \rightarrow \overline{\mathcal{M}}_{g_1+g_2,n_1+n_2}$, which glues two curves on their last marked point;
- ◊ The non-separating attaching map, $\tau : \overline{\mathcal{M}}_{g,n+2} \rightarrow \overline{\mathcal{M}}_{g+1,n}$, which glues a curve on itself by identifying its two last marked points.

Let a point $x \in \overline{\mathcal{M}}_{g,n}$, i.e., the isomorphism class of a curve, say C . Then it can be shown that $\pi^{-1}(x) \in \overline{\mathcal{M}}_{g,n+1}$ is actually isomorphic to the curve C itself (together with the action of its own automorphism group). We can then define sections $s_i : \overline{\mathcal{M}}_{g,n} \rightarrow \overline{\mathcal{M}}_{g,n+1}$ by setting $s_i(x) = \tilde{x}_i \in \pi^{-1}(x)$, where \tilde{x}_i is the i -th marked point on the curve $\pi^{-1}(x)$.

Then we have important fibre bundles. The line bundles $L_i \rightarrow \overline{\mathcal{M}}_{g,n}$, for $i \in \{1, \dots, n\}$, are constructed in such a way that the fibres $L_i|_x$ at a point $x \in \overline{\mathcal{M}}_{g,n}$ satisfy $L_i|_x \simeq T_{\tilde{x}_i}^* \pi^{-1}(x)$, where $\tilde{x}_i = s_i(x)$. These bundles L_1, \dots, L_n are called the *tautological bundles*. The bundle $\mathbb{H} \rightarrow \overline{\mathcal{M}}_{g,n}$ is defined in such a way that $\mathbb{H}|_x \simeq T^* \pi^{-1}(x)$. This bundle \mathbb{H} is called the *Hodge bundle* and is of rank g . We can then consider the Chern classes of these bundles:

$$\psi_i := c_1(L_i) \in H^2(\overline{\mathcal{M}}_{g,n}, \mathbb{Q}), \quad \lambda_j := c_j(\mathbb{H}) \in H^{2j}(\overline{\mathcal{M}}_{g,n}, \mathbb{Q}).$$

The classes ψ_1, \dots, ψ_n we call the *psi classes* and the classes $\lambda_1, \dots, \lambda_g$ we call the *lambda classes*.

Cohomological field theory. Using the material above, we can define cohomological field theories (CohFTs) which were introduced by Kontsevich and Manin [KM94]. A CohFT consists of the data of a finite-dimensional complex vector space V , with a special vector e_1 , and equipped with a metric $\eta = (\eta^{\alpha\beta})$, together with maps

$$c_{g,n} : V^{\otimes n} \rightarrow H^*(\overline{\mathcal{M}}_{g,n}, \mathbb{C})$$

which satisfy the following properties (we choose a base (e_1, e_2, \dots, e_N) and we use Einstein notation):

- (i) $c_{g,n}$ is equivariant w.r.t. to simultaneous permutation of copies of V in $V^{\otimes n}$ and of the marked points on the curves;
- (ii) $c_{0,3}(e_1, e_\alpha, e_\beta) = \eta^{\alpha\beta}$;
- (iii) $\pi^* c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = c_{g,n+1}(e_{\alpha_1}, \dots, e_{\alpha_n}, e_1)$;
- (iv) $\sigma^* c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = c_{g_1,n_1+1}(e_1, \dots, e_{n_1}, e_\mu) \eta^{\mu\nu} c_{g_2,n_2+1}(e_\nu, e_1, \dots, e_{n_2})$;
- (v) $\tau^* c_{g+1,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = \eta^{\mu\nu} c_{g,n+2}(e_\mu, e_\nu, e_{\alpha_1}, \dots, e_{\alpha_n})$.

All in all, we ask for the classes $c_{g,n}$ to “behave naturally” with respect to the maps π, σ and τ , considering the metric $\eta^{\alpha\beta}$.

Given a CohFT $\{c_{g,n} \mid g, n \geq 0\}$, we define the *intersection numbers* (or *correlators*) associated to the CohFT by

$$\langle \tau_{d_1}(e_{\alpha_1}) \cdots \tau_{d_n}(e_{\alpha_n}) \rangle_g := \int_{\overline{\mathcal{M}}_{g,n}} c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) \cdot \psi_1^{d_1} \cdots \psi_n^{d_n} \in \mathbb{C}.$$

with $\alpha_i \in \{1, \dots, N\}$ and $d_j \geq 0$. These numbers vanish if the cohomological degree of the integrand differs from $3g - 3 + n$. We then define the *potential* F of a CohFT as the generating series of the intersection numbers, namely,

$$F_g(t_*^*) := \sum_{\substack{n \geq 0 \\ 2g-2+n > 0}} \frac{1}{n!} \sum_{d_1, \dots, d_n \geq 0} \langle \tau_{d_1}(e_{\alpha_1}) \cdots \tau_{d_n}(e_{\alpha_n}) \rangle \cdot t_{d_1}^{\alpha_1} \cdots t_{d_n}^{\alpha_n}. \quad (2.39)$$

and $F(t_*^*; \varepsilon) = \sum_{g \geq 0} \varepsilon^{2g} F_g(t_*^*)$. That introduces variables $t_*^* := \{t_{d_j}^{\alpha_i} \mid \alpha_i \in \{1, \dots, N\}, d_j \geq 0\}$ which we sometimes call the *topological variables*.

For example, the *trivial CohFT* has vector space $V = \mathbb{C}$, metric $\eta = 1$ and classes $c_{g,n} = 1$. It plays an important role for it is connected to the KdV hierarchy [Wit91] as we will see below. In the same fashion, the *Hodge CohFT* has the same data except that the classes are $c_{g,n} = \Lambda(s) := 1 + \sum_{j=1}^g \lambda^j s^j$. The latter is connected to the “intermediate long wave” hierarchy (ILW) [Bur15a]

Another important example is the Gromov–Witten theory of a smooth projective variety X with a given homology class $\beta \in H_2(X, \mathbb{Z})$. They are constructed via the moduli spaces $\overline{\mathcal{M}}_{g,n}(X, \beta)$ of genus g , n marked curves $(C, (x_i))$ together with a stable map $f : C \rightarrow X$ such that $f_*[C] = \beta$. We recall its construction in Example 5.2.3. The vector space of the CohFT is $V = H^*(X, \mathbb{C})$ and the metric is given by $\eta(v, w) = \int_{\mathcal{P}(v)} w$, where $\mathcal{P} : H^*(X, \mathbb{C}) \rightarrow H_*(X, \mathbb{C})$ is the Poincaré duality. The classes are defined by

$$c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = p_* \text{ev}^*(e_{\alpha_1}, \dots, e_{\alpha_n}), \quad (2.40)$$

where $p : \overline{\mathcal{M}}_{g,n}(X, \beta) \rightarrow \overline{\mathcal{M}}_{g,n}$ maps $(C, (x_i), f)$ to $(C, (x_i))$ (forgets the map f) and where $\text{ev} : \overline{\mathcal{M}}_{g,n}(X, \beta) \rightarrow X^n$ maps $(C, (x_i), f)$ to $(f(x_i))$ (evaluates the map). The proper definition is complex and involves the virtual fundamental class $[\overline{\mathcal{M}}_{g,n}(X, \beta)]^{\text{vir}}$ introduced by Behrend and Fantechi [BF97].

Fan–Jarvis–Ruan–Witten theory. One last example which is of interest here is the FJRW theory of a pair (W, G) . Here $W \in \mathbb{C}[z_1, \dots, z_m]$ is a quasi-homogeneous polynomial with weights q_1, \dots, q_m which has an isolated singularity at the origin. And G is group of diagonal matrices $\gamma = (\gamma_1, \dots, \gamma_m)$ which leave the polynomial W invariant, and which contains the element $J = (e^{2i\pi q_1}, \dots, e^{2i\pi q_m})$.

The vector space is defined by $V = \bigoplus_{\gamma \in G} (\mathcal{Q}_{W_\gamma} \otimes dz_\gamma)^G$, where $W_\gamma = W|_{\text{Fix}(\gamma)}$ and \mathcal{Q}_{W_γ} is its Jacobian ring and where $dz^\gamma = \bigwedge_{z_i \in \text{Fix}(\gamma)} dz_i$. We give a brief review of FJRW theory and the associated CohFT in §5.2.1, see [FJR11, FJR13]. Since quasi-homogeneous polynomials with isolated singularities are classified by the Dynkin diagrams of type *ADE*, there is a correspondence between FJRW theories and semisimple Lie algebras of the corresponding type.

The Dubrovin–Zhang hierarchies. The Dubrovin–Zhang hierarchies (DZ for short) are hierarchies of PDEs of the same type as those of Drinfeld and Sokolov described in §2.3.1. They are defined for any semisimple CohFT. The latter means that the potential F of Equation

(2.39) must be such that the N -dimensional complex algebra defined by the structure constants

$$f_{\beta\gamma}^\alpha = \eta^{\alpha\mu} \frac{\partial^3 F_0}{\partial t_0^\mu \partial t_0^\gamma \partial t_0^\beta} \Big|_{t_{p>0}^* = 0}$$

is semisimple. Let us denote

$$\Omega_{\alpha,p;\beta,q}(t_*^*; \varepsilon) = \frac{\partial^2 F}{\partial t_p^\alpha \partial t_q^\beta} = \sum_{g \geq 0} \Omega_{\alpha,p;\beta,q}^{[2g]} \varepsilon^{2g}.$$

Then we define Hamiltonian densities and local Poisson structure:

$$h_{\alpha,p}^{\text{DZ},(0)}(v^*) := \Omega_{\alpha,p+1;1,0}^{[0]}(t_0^* = v^*, 0, 0, \dots), \quad K_{\text{DZ},(0)}^{\alpha\beta} := \eta^{\alpha\beta} \partial_x.$$

These Hamiltonian densities commute w.r.t. to the above bracket [DZ08]. They form the so-called genus 0 Dubrovin–Zhang hierarchy, or the *principal hierarchy* of the CohFT. To construct the full genus DZ hierarchy, we make use of the so-called *topological solution*; it is the one solution v^α of the DZ hierarchy that is determined by the initial data

$$v^\alpha(x, t_*^* = 0) = \delta^{\alpha,1} x.$$

Then there exist rational functions P_p^α such that [Get02, BPS12]

$$F_g(t_0^*, t_1^*, t_2^*, \dots) = F_g(P_0^*(v_0^*, \dots, v_{3g-2}^*), \dots, P_{3g-2}^*(v_0^*, \dots, v_{3g-2}^*), 0, 0, \dots) \Big|_{x=0}.$$

Now consider the change of coordinates

$$\varphi(v^\alpha) = w^\alpha := v^\alpha + \sum_{g \geq 1} \varepsilon^{2g} \frac{\partial^2 F_g(v_0^*, \dots, v_{3g-2}^*)}{\partial t_0^\alpha \partial x}. \quad (2.41)$$

The full genera DZ hierarchy is then obtained by applying the inverse transform φ^{-1} to the genus 0 hierarchy. Namely,

$$h_{\alpha,p}^{\text{DZ}}(w_*^*) := h_{\alpha,p}^{\text{DZ},(0)} \circ \varphi^{-1}(w_*^*), \quad K_{\text{DZ}}^{\alpha\beta}(w_*^*) = ((\varphi^{-1})^* K_{\text{DZ},(0)})^{\alpha\beta}(w_*^*),$$

where $(\varphi^{-1})^* K_{\text{DZ},(0)}$ is the transformed differential operator; the transformation is given in Equation (5.18). These data form a Hamiltonian system in the sense of Equation (2.14), i.e.

$$\frac{\partial w^\alpha}{\partial t_p^\alpha} = \{w^\alpha, \bar{h}_{\alpha,p}^{\text{DZ}}\}_{\text{DZ}} = \sum_{\beta=1}^N K_{\text{DZ}}^{\alpha\beta} \frac{\delta h_{\alpha,p}^{\text{DZ}}}{\delta u^\beta}. \quad (2.42)$$

The Hamiltonians $h_{\alpha,p}^{\text{DZ}}$ lie in the ring

$$\widehat{\mathcal{A}} := \mathbb{C}[[u^\alpha, \varepsilon]][u_{k>0}^\alpha]$$

of *extended differential polynomials*. Similarly, $\widehat{\mathcal{F}} = \widehat{\mathcal{A}}/\text{Im}(\partial_x)$. We define a gradation on it by setting $\deg u_k^\alpha = k$ and $\deg \varepsilon = -1$ so that ε counts the number of derivatives in a monomial of total degree 0. Parameter ε is often referred to as the *dispersion parameter* and plays an important role in what follows. It is immediate from the definition that the Hamiltonians $h_{\alpha,p}^{\text{DZ}}$ are tau symmetric in the sense of [DZ08], i.e.

$$\frac{\partial h_{\alpha,p-1}^{\text{DZ}}}{\partial t_q^\beta} = \frac{\partial h_{\beta,q-1}^{\text{DZ}}}{\partial t_p^\alpha}.$$

The above equation is to be compared with (2.33).

By construction, the potential of the CohFT is a tau function of the DZ hierarchy and is associated to the topological solution. It has been proven over the years that the DZ hierarchy constructed from the FJRW theory of a simple singularity of type *ADE* coincides (up to certain non trivial subtleties) with the Drinfeld–Sokolov hierarchy associated to the corresponding simple Lie algebra of the same type [GM05, FGM10]. See [LRZ15] for a clear exposé on that matter. This correspondence³² between CohFTs and integrable hierarchies we call *generalized Witten–Kontsevich correspondence*.

The very first example of such correspondence (prior to the definition of CohFTs and of the DZ hierarchy) is the celebrated Witten–Kontsevich theorem [Wit91, Kon92] which states (in our framework) that the potential of the trivial CohFT is the topological tau function of the KdV hierarchy (2.20) (which coincides with the DZ hierarchy associated the trivial CohFT). More precisely, consider the potential $F_{\text{WK}}(t_*)$ of the trivial CohFT:

$$F_{\text{WK}}(t_*) = \sum_{g \geq 0} \sum_{\substack{n \geq 0 \\ 2g-2+n > 0}} \frac{1}{n!} \sum_{d_1, \dots, d_n \geq 0} t_{d_1} \cdots t_{d_n} \int_{\mathcal{M}_{g,n}} \psi_1^{d_1} \cdots \psi_n^{d_n}.$$

Now substitute $F_{\text{WK}} = -\frac{1}{2} \log(\tau_{\text{WK}})$ and then $t_i \mapsto (2i+1)!! t_{2i+1}$ where $(2i+1)!! = 1 \cdot 3 \cdot 5 \cdots (2i+1)$. Then the resulting function τ_{WK} is a tau function of the KdV hierarchy in the sense that $v = -2\partial_x^2 \log(\tau_{\text{WK}})$ solves (2.20) (in the substitution, $t_0 \mapsto t_1$ and we set $t_1 = x$).³³ Moreover, v satisfies $v(x, t_{i>1} = 0) = x$.

2.4.2 The double ramification hierarchies

We now explain how to construct Buryak’s double ramification hierarchies (DR) introduced in [Bur15a]. The latter, as mentioned previously, are conjectured to be equivalent to the DZ hierarchies. Yet they are, in some extent, easier to construct, and they satisfy strong recursion relations that are new to the theory of integrable hierarchies. Moreover, they can be defined for arbitrary CohFTs, not only semisimple ones. In the case of the FJRW theory of a simple singularity of type *ADE*, it is conjectured that the DR, DZ and DS hierarchies coincide. A fact

³²It is a correspondence “CohFT / Int. hierarchy” in the sense that the potential of the semisimple CohFT is a tau function of the DZ hierarchy, but also allows to construct the local Poisson structure as well as all the Hamiltonian densities. On the other hand, the topological tau function, seen as the potential, contains by construction all the cohomology classes and therefore, all the information of the CohFT.

³³For reasons that appear clear with this example, we often call the potential F itself the tau function.

verified for hierarchies of type A_1, \dots, A_5 (and others, see below) and which we confirmed for the case of D_4 in a forthcoming preprint and described in Chapter 5.

Hamiltonian structure. First of all, the local Poisson structure is simply given by the operator $\eta^{\alpha\beta} \partial_x$, where $\eta^{\alpha\beta}$ is the metric of the CohFT, meaning that for two local functionals $\bar{f}, \bar{g} \in \widehat{\mathcal{F}}$,

$$\{\bar{f}, \bar{g}\} = \int \sum_{\alpha, \beta=1}^N \frac{\delta f}{\delta u^\alpha} \eta^{\alpha\beta} \partial_x \frac{\delta g}{\delta u^\beta}.$$

For the Hamiltonian densities we introduce the *double ramification cycles*. Let integers a_1, \dots, a_n be such that $\sum a_i = 0$. Given $g \geq 0$, the double ramification cycle $\text{DR}_g(a_1, \dots, a_n) \in H^{2g}(\overline{\mathcal{M}}_{g,n}, \mathbb{Q})$ is, heuristically, the Poincaré dual of the homology class of the locus of marked complex curves $(C, (x_i))$ such that $\sum a_i x_i$ is the divisor of the zeros and poles of some meromorphic function. We consider x_i a pole if $a_i > 0$ and a zero otherwise. A more precise definition is given in §5.2.2.

Form this, we define the following intersection numbers: let $\alpha \in \{1, \dots, N\}$ and let $g, d \geq 0$,

$$P_{\alpha, \underline{d}; \underline{\alpha}}^{g; \underline{a}} := \int_{\text{DR}_g(-\sum a_i, \underline{a})} \lambda_g \psi_1^d c_{g, n+1}(e_\alpha, e_{\alpha_1}, \dots, e_{\alpha_n}),$$

where $\underline{\alpha} = (\alpha_1, \dots, \alpha_n) \subset \{1, \dots, N\}$ and $\underline{a} = (a_1, \dots, a_n) \subset \mathbb{Z}$. The class λ_g is the Chern class $c_g(\mathbb{H})$ of highest degree of the Hodge bundle. It can be shown that these numbers are polynomial in the a_i 's and homogeneous of degree $2g$ so that they can be written as

$$P_{\alpha, \underline{d}; \underline{\alpha}}^{g; \underline{a}} = \sum_{\substack{b_1, \dots, b_n \geq 0 \\ \sum b_j = 2g}} \widetilde{P}_{\alpha, \underline{d}; \underline{\alpha}}^{g; \underline{b}} a_1^{b_1} \dots a_n^{b_n} \in \mathbb{C}_{2g}[a_1, \dots, a_n],$$

We can now define the Hamiltonian densities:

$$g_{\alpha, d} := \sum_{\substack{g, n \geq 0 \\ 2g - 2 + n > 0}} \frac{\varepsilon^{2g}}{n!} \sum_{\substack{b_1, \dots, b_n \geq 0 \\ \sum b_j = 2g}} \widetilde{P}_{\alpha, \underline{d}; \underline{\alpha}}^{g; \underline{b}} \prod_{j=1}^n u_{b_j}^{\alpha_j}. \quad (2.43)$$

Again, the data of the local Poisson structure $\eta^{\alpha\beta} \partial_x$ and of the Hamiltonian densities $g_{\alpha, d}$ form a Hamiltonian system in the sense of Equation (2.42). The densities $g_{\alpha, d}$ satisfy the following recursion relation: First we set $g_{\alpha, -1} := \eta_{\alpha\mu} u^\mu$, where $\eta_{\alpha\mu}$ is the inverse matrix of $\eta^{\alpha\mu}$, then we have [BR16b]

$$\partial_x(D-1)g_{\alpha, p+1} = \{g_{\alpha, p}, \bar{g}_{1,1}\} = \sum_{\beta, \gamma=1}^N \sum_{k \geq 0} \frac{\partial g_{\alpha, p}}{\partial u_k^\beta} \eta^{\beta\gamma} \partial_x^{k+1} \frac{\delta g_{1,1}}{\delta u^\gamma},$$

where $D = \varepsilon \partial_\varepsilon + \sum_{s \geq 0} u_s^\alpha \partial_{u_s^\alpha}$. It can be easily shown that it indeed allows to compute the densities recursively from the data of $\bar{g}_{1,1}$. It shall be emphasized that this specific form of recursion is new to the theory on integrable hierarchies.

On the other hand, these Hamiltonian densities are not tau symmetric. To obtain tau symmetric densities, it suffices to consider differential polynomials

$$h_{\alpha,p} := \frac{\delta g_{\alpha,p}}{\delta u^1}.$$

With respect to the Poisson structure $\eta^{\alpha\beta} \partial_x$, they are tau symmetric Hamiltonian densities.

The strong DR/DZ equivalence conjecture. In order to state the DR/DZ conjecture, we must define *Miura transforms* which will play the role of “change of coordinates” in this context. A Miura transform is any invertible map $\varphi : \widehat{\mathcal{A}} \rightarrow \widehat{\mathcal{A}}$, $v^\alpha := \varphi(u^\alpha) \in \mathcal{A}^{[0]}$, which is such that

$$\det \left[\frac{\partial v^\alpha}{\partial u^\beta} \Big|_{\varepsilon=0} \right]_{\alpha, \beta \in [1, N]} \neq 0.$$

We now choose special coordinates for the hierarchy that we call *normal* following [DZ08]. Given the tau structure $h_{\alpha,p}$, they are defined as

$$\tilde{u}^\alpha := \eta^{\alpha\beta} h_{\alpha,-1}.$$

The map $\varphi : u^\alpha \mapsto \tilde{u}^\alpha$ is a Miura transform. We denote

$$h_{\alpha,p}^{\text{DR}} := h_{\alpha,p} \circ \varphi, \quad K_{\text{DR}}^{\alpha\beta} := \varphi^*(\eta^{\alpha\beta} \partial_x) \quad (2.44)$$

the transformed Hamiltonian densities and local Poisson structure written in normal coordinates. We can now state the strong DR/DZ equivalence conjecture. More details are given in Conjecture 5.2.23.

Conjecture 2.4.1 (Strong DR/DZ equivalence [BDGR18]). *For any semisimple CohFT there exists a unique differential polynomial $f \in \widehat{\mathcal{A}}^{[-2]}$ such that the DR and DZ hierarchies coincide up to the normal Miura transform defined by*

$$\tilde{u}^\alpha = w^\alpha + \eta^{\alpha\beta} \partial_x \{f, \bar{h}_{\beta,0}^{\text{DZ}}\}_{\text{DZ}},$$

In particular, for the CohFT consisting of the FJRW theory of a simple singularity of type *ADE*, the conjecture states that the DR, DZ and DS hierarchies coincide. The strong DR/DZ equivalence has been confirmed in a number of cases [Bur15a, BG16, BDGR18]: The trivial CohFT; the full Hodge class CohFT; the hierarchies of type A_3, A_4, A_5 ; the Gromov–Witten theory of \mathbb{P}^1 ; up to genus 5 for any CohFT of rank 1; up to genus 2 for any semisimple CohFT.

2.4.3 Hierarchy of type D_4 and the strong DR/DZ equivalence

Finally, we explain how we confirmed the strong DR/DZ equivalence in the case of the hierarchy of type D_4 . It amounts to proving that the DR and DS hierarchies of type D_4 coincide. We will compare them in normal coordinates. To do so, we used the following useful theorem of

Buryak and Guéré [BG16]: Given a semisimple CohFT, suppose there exists a Miura transform $\varphi : u^\alpha \mapsto \tilde{u}^\alpha(u_\bullet^*, \varepsilon)$, such that

$$\frac{\partial \tilde{u}^\alpha}{\partial u^1} = \delta^{\alpha,1}, \quad \varphi^*(\eta^{\alpha\beta} \partial_x) = K_{\text{DZ}}^{\alpha\beta}, \quad \overline{g_{1,1} \circ \varphi} = \overline{h_{1,1}^{\text{DZ}}}, \quad (2.45)$$

where $\varphi^*(\eta^{\alpha\beta} \partial_x)$ is given in (5.18). Then the DR hierarchy, under the Miura transform φ , coincides with the DZ hierarchy. When computing the DR hierarchy associated to the FJRW CohFT of type D_4 , we find that the Miura transform from the coordinates u^α to the normal coordinates \tilde{u}^α reads

$$\begin{cases} \tilde{u}^1 = u^1 + (-\frac{1}{36}(u_1^3)^2 - \frac{1}{36}u_1^3 u_2^3)\varepsilon^2 - \frac{1}{45}u_4^3 \varepsilon^4 \\ \tilde{u}^2 = u^2 \\ \tilde{u}^3 = u^3 \\ \tilde{u}^4 = u^4 \end{cases} \quad (2.46)$$

This agrees with the first condition of Equation (2.45). The Hamiltonian densities $g_{\alpha,0}$ and $g_{1,1}$ are given in §5.2.4.

So, it remains to explain how to compute the Drinfeld–Sokolov hierarchy of type D_4 . To do so, we will use the scalar Lax pair, in the sense of Equation (3.34), and the Hamiltonian structure found by Liu, Wu and Zhang [LWZ10]. Indeed, the DS hierarchies are constructed generically as hierarchies of matrix Lax pair, but specifically for the case of the hierarchies of type D_n , Drinfeld and Sokolov did not give a full picture of the (bi-)Hamiltonian structure. However they constructed the local Poisson structure. In order to state the problem, it is necessary to discuss the affine algebra of type $D_4^{(1)}$.

The matrix Lax pair of the DS hierarchy of type D_4 . What we discuss here adapts to the affine algebra of type $D_n^{(1)}$ in general. Recall that the affine algebra \mathfrak{g} of type $D_4^{(1)}$ is the affine algebra associated to the simple algebra D_4 , the latter being isomorphic to the algebra $\mathfrak{o}(8)$ (8×8 antisymmetric—not antihermitian—complex matrices). That is, $D_4^{(1)}$ is isomorphic to

$$\mathfrak{g} = \mathfrak{o}(8) \otimes \mathbb{C}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c.$$

This affine algebra has set of positive exponents $\mathbb{E}^+ = \mathbb{Z}_{>0}^{\text{odd}} \cup (3\mathbb{Z}_{>0}^{\text{odd}})' = \{1, 3, 3', 5, 7, 9, 9', \dots\}$, meaning that all exponent of the form $3k$, with k odd, has multiplicity 2. We denote Λ_k (resp. Γ_k), with k odd, the generator of the Heisenberg subalgebra associated to the exponent $3k$ (resp. $(3k)'$). Similarly, we denote t_k (resp. \hat{t}_k) the variable associated to Λ_k (resp. to Γ_k). The Λ_k 's and Γ_k 's are given in Equation (5.54) and (5.55). This gives the matrix Lax equations [DS84]

$$\frac{\partial \mathcal{L}}{\partial t_k} = [(e^{\text{ad}_U} \Lambda_k)_+, \mathcal{L}], \quad (2.47)$$

$$\frac{\partial \mathcal{L}}{\partial \hat{t}_k} = [(e^{\text{ad}_U} \Gamma_k)_+, \mathcal{L}], \quad (2.48)$$

for $k \in \mathbb{Z}_{>0}^{\text{odd}}$, where $\mathcal{L} = \partial_x + \Lambda + Q$, with $Q \in \mathfrak{b}[[\mathbf{t}]]$ and where Λ is the principal cyclic elements given in Equation (2.29).

The scalar Lax pair and Hamiltonian structure. In their original paper [DS84], Drinfeld and Sokolov described the Hamiltonian structure of the flows w.r.t. the variables t_k , but not \hat{t}_k (it needed a new algebra of operator). The local Poisson structure, in the fashion of Equation (2.22), is given as follows [DS84]. First, the scalar Lax operator reads

$$L = \partial_x^6 + \partial_x^{-1} \sum_{\mu=1}^3 \left(s^\mu \partial_x^{2\mu-1} + \partial_x^{2\mu-1} s^\mu \right) + \partial_x^{-1} \varrho \partial_x^{-1} \varrho, \quad s^4 = (\varrho)^2, \quad (2.49)$$

Then by denoting $L' = \partial_x \circ L$, the first Poisson structure reads

$$\{\bar{f}, \bar{g}\}_1 = \int \text{res } X [(\partial_x Y_+ L')_- - (L' Y_+ \partial_x)_- - (\partial_x Y_- L')_+ + (L' Y_- \partial_x)_+] dx, \quad (2.50)$$

for any $\bar{f}, \bar{g} \in \widehat{\mathcal{F}}$,³⁴ where $X = \frac{\delta f}{\delta L}$, $Y = \frac{\delta g}{\delta L}$ are given in Equation (5.66). This allows to define a scalar Lax pair along with a Hamiltonian representation³⁵ of Equation (2.47) by [DS84]

$$\frac{\partial L}{\partial t_k} = \left[(L^{k/6})_+, L \right], \quad \iff \quad \frac{\partial s^\alpha}{\partial t_k} = \{s^\alpha, H_{k+6}\}_1, \quad H_k = \frac{6}{k} \int \text{res } L^{k/6}.$$

The latter, as we can see, is fairly similar to that of the KdV hierarchy (2.20), (2.14). Now for the flows w.r.t the variables \hat{t}_k , it not possible to use the 6-th root $L^{1/6}$.

What Liu, Wu and Zhang [LWZ10] did, is that they constructed a new algebra $\widetilde{\mathcal{D}}$ which intersects \mathcal{D} (Equation (2.19)) but does not contain it. Essentially, they used the topology on $\widehat{\mathcal{A}}$ induced by the gradation $\widehat{\mathcal{A}} = \bigoplus_{k \in \mathbb{Z}} \widehat{\mathcal{A}}_k$ defined by $\deg s_k^\alpha = k$, to defined a ring structure (with the same product as (2.7)) on the space

$$\widetilde{\mathcal{D}} = \left\{ \sum_{i \geq m} \sum_{j \geq m-i} a^{i,j} \partial_x^i \mid m \in \mathbb{Z}, a^{i,j} \in \widehat{\mathcal{A}}_j \right\}.$$

The resulting algebra $\widetilde{\mathcal{D}}$, whose elements are called *pseudo-differential operators (PDOs) of the second type*, allows for infinitely many terms of positive *and* negative degrees.

The result is that there exists a unique operator $L^{1/2} \in \widetilde{\mathcal{D}}$ such that $(L^{1/2})^2 = L$ satisfying some constraints (with L given in (2.49)). It allows to define a scalar Lax pair representation of Equation (2.48) [LWZ10] as

$$\frac{\partial L}{\partial \hat{t}_k} = \left[(L^{k/2})_+, L \right]. \quad \iff \quad \frac{\partial s^\alpha}{\partial \hat{t}_k} = \{s^\alpha, \hat{H}_{k+2}\}_1, \quad \hat{H}_k = \frac{2}{k} \int \text{res } L^{k/2}.$$

Explicit computation of the Hamiltonian structure and comparison with the DR hierarchy. Finally, we give the explicit computations performed in §5.3.2 on the DS hierarchy of type D_4 and the result that they do coincide with the DR hierarchy of type D_4 , thus confirming

³⁴Note that here, $\widehat{\mathcal{F}} = \widehat{\mathcal{A}}/\text{Im}(\partial_x)$ where $\widehat{\mathcal{A}} = \mathbb{C}[[s_0^\alpha; \varepsilon]][s_{k>0}^\alpha]$.

³⁵Notice that here we only need to compute one of the two Hamiltonian representations of the hierarchy. This is because what we want is to compare the Hamiltonian structures of two a priori distinct hierarchies. The bi-Hamiltonian property has many interest, but we do not need it here.

the strong DR/DZ equivalence. To do so, we need to substitute the variables t_k and \hat{t}_k , with $k \in \mathbb{Z}_{>0}^{\text{odd}}$, for the topological variables $t_{\alpha,p}$, with $\alpha \in \{1, \dots, 4\}$ and $p \geq 0$. Similarly, we need to define tau symmetric Hamiltonian densities $h_{\alpha,p}$ in the sense of Equation (2.33). We have, for $\mu \in \{1, 2, 3\}$ and $p \geq 0$,

$$\begin{aligned} t_{\mu,p} &= \Gamma_{\mu,p} \cdot t_{6p+2\alpha-1}, & h_{\mu,p-1} &= \Gamma_{\mu,p}^{-1} \cdot \text{res } L^{(6p+2\alpha-1)/6}, \\ t_{4,p} &= \Gamma_{4,p} \cdot \hat{t}_{2p+1}, & h_{4,p-1} &= \Gamma_{4,p}^{-1} \cdot \text{res } L^{(2p+1)/2}. \end{aligned}$$

The factors $\Gamma_{\alpha,p}$ are given in (5.68). It remains to substitute the functions s^α for the normal coordinates $\tilde{u}^\alpha = \eta^{\alpha\beta} h_{\beta,-1}$ of the DS hierarchy of type D_4 (we purposefully denote them like the normal coordinates of the DR hierarchy). We denote by $h_{\alpha,p}^{\text{DS}}$ the tau symmetric Hamiltonian densities $h_{\alpha,p}$ written in normal coordinates \tilde{u}^α . Similarly, in normal coordinates, the first Poisson structure (2.50) takes the following form,

$$(K_{\text{DS}}^{\alpha\beta}) = \begin{pmatrix} \left(\frac{1}{6} \tilde{u}_2^3 \partial_x + \frac{1}{2} \tilde{u}_1^3 \partial_x^2 + \frac{1}{3} \tilde{u}^3 \partial_x^3 \right) \varepsilon^2 + \frac{4}{15} \partial_x^5 \varepsilon^4 & 0 & 6\partial_x & 0 \\ 0 & 6\partial_x & 0 & 0 \\ 6\partial_x & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\partial_x \end{pmatrix},$$

where $(K_{\text{DS}}^{\alpha\beta})$ is the matrix-valued differential operator such that $\{\bar{f}, \bar{g}\}_1 = \int \frac{\delta f}{\delta \tilde{u}^\alpha} K_{\text{DS}}^{\alpha\beta} \frac{\delta g}{\delta \tilde{u}^\beta}$. Considering Buryak and Guéré's theorem (see the beginning of §2.4.3), we now have all the material needed to state the main theorem of this section. Its proof (that is, the computations) is given in §5.3.2.

Theorem 2.4.2 (Strong DR/DZ for D_4 [DR19]). *The strong DR/DZ equivalence holds true for the hierarchy of type D_4 under the normal Miura transform (2.46). More precisely, in the normal coordinates \tilde{u}^α ,*

(i) *The local Poisson structures are the same: $K_{\text{DR}}^{\alpha\beta} = K_{\text{DS}}^{\alpha\beta}$;*

(ii) *The Hamiltonian densities $h_{1,1}^{\text{DR}}$ and $h_{1,1}^{\text{DS}}$ differ by a total x -derivative, i.e. $\bar{h}_{1,1}^{\text{DR}} = \bar{h}_{1,1}^{\text{DS}}$.*

Moreover, the level 0 Hamiltonian densities coincide: $h_{\alpha,0}^{\text{DR}} = h_{\alpha,0}^{\text{DS}}$ for $\alpha \in \{1, \dots, 4\}$.

Chapter 3

Polynomial tau functions of the Korteweg–de Vries hierarchies

3.1 Pseudo-differential operators

In this section we define the ring of algebraic pseudo-differential operators and give some of their essential properties that we will use through out the document. We refer to the classic books [DJM00], [BBT03], [Dic03] for more details on pseudo-differential operators and integrable hierarchies.

Definition and basic properties

Definition 3.1.1 (Pseudo-differential operators). The algebra \mathcal{D} of pseudo-differential operators is the following set of formal power series in an indeterminate ∂ over the ring of complex-valued formal power series,

$$\mathcal{D} = \left\{ \sum_{i=-\infty}^n a^i \partial^i \mid a^i \in \mathbb{C}[[x]], n \in \mathbb{Z} \right\},$$

equipped with the non commutative product

$$a\partial^n \circ b\partial^m = \sum_{k \geq 0} \binom{n}{k} a b_{(k)} \partial^{n+m-k}, \quad \binom{n}{k} := \frac{(n-k+1) \cdots (n-1) \cdot n}{k!}, \quad (3.1)$$

where $b_{(k)}$ (sometimes just b_k) denotes the k -th derivative w.r.t. x of the function b and $b_{(0)} = b$. When there is no ambiguity, we will simply call them *operators* and omit the \circ notation for the product.¹

¹The notation a^i for i -th coefficient of X is not the most standard. But in the sequel, and especially in Chapter 5, we will denote $u_k^\alpha := \partial_x^k u^\alpha$ (as is customary in [DZ08, DLZ08, BR16b] to cite a few). Therefore, in what follows, it is the subscript that will denote the order of differentiation.

For our purpose, the algebra \mathcal{D} is defined over $\mathbb{C}[[x]]$. However, we can define pseudo-differential operators over any commutative integral domain R equipped with a derivation δ that we denote $\text{DO}(R, \delta)$. For instance, it is often defined over $C^\infty(\mathbb{R}, \mathbb{C})$ with the derivation being the usual derivation w.r.t. the variable along \mathbb{R} . Our approach is purely algebraic and involves only formal power series so that, in our case,

$$\mathcal{D} = \text{DO}(\mathbb{C}[[x]], \partial = \partial_x).$$

The algebra of \mathcal{D} is sometimes denoted $\mathcal{D} = R((\partial^{-1}))$ (where $R = \mathbb{C}[[x]]$). This notation is intended as a non-commutative version of the ring $R((T^{-1}))$ of formal Laurent series in T^{-1} , i.e. formal power series in T and T^{-1} with only finitely many nonzero coefficients of positive degree.

Equation (3.1) naturally generalizes the binomial coefficients $\binom{n}{k}$ for negative values of n . And if $n < 0$, then the sequence $\binom{n}{k}$ diverges. They still satisfy the recursion relation

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}, \quad n \in \mathbb{Z}, k \geq 0,$$

which extends Pascal's triangle as shown in Figure 3.1. The formula for binomial coefficients of Equation (3.1) readily implies the relation

$$\binom{n}{k} = (-1)^k \binom{k-n-1}{k}, \quad n \geq 0, k \in \{0, \dots, n\}, \quad (3.2)$$

(so that $k - n - 1 < 0$). This relation can be easily seen in Figure 3.1 as the identification of coefficients (up to a sign) within the k -th column modulo a shift $n \mapsto k - n - 1$.

$n \setminus k$	0	1	2	3	4
-4	1	-4	10	-20	35
-3	1	-3	6	-10	15
-2	1	-2	3	-4	5
-1	1	-1	1	-1	1
0	1				
1	1	1			
2	1	2	1		
3	1	3	3	1	
4	1	4	6	4	1

Figure 3.1: Generalized binomial coefficients in an extended Pascal's "triangle"

Definition 3.1.2. Given a pseudo-differential operator of order n , $A = \sum_{k \leq n} a^k \partial^k$ and an integer $m \in \mathbb{Z}$, we call the m -truncated part of A the operator

$$A_{\geq m} = \sum_{k=m}^n a^k \partial^k.$$

In the same fashion, we call respectively the *positive part* and *negative part* of A the operators

$$A_+ = A_{\geq 0}, \quad A_- = A - A_+.$$

Finally, we call the *residue* of A its coefficient of degree -1 in ∂ , i.e.,

$$\text{res}_\partial A = a^{-1}$$

(we denote simply $\text{res } A$ when there is no ambiguity).

The positive part A_+ of A is also called its *differential part* while its negative part A_- is also called its *integral part* (sometimes its *purely pseudo-differential part*). We call respectively differential operators and integral operators the elements of the following sets:

$$\mathcal{D}_+ = \left\{ \sum_{k=0}^n a^k \partial^k, n \geq 0 \right\}, \quad \mathcal{D}_- = \left\{ \sum_{k=-\infty}^n a^k \partial^k, n < 0 \right\}.$$

The product of pseudo-differential operators stabilizes these sets, that is, $\mathcal{D}_+ \times \mathcal{D}_+ \xrightarrow{\circ} \mathcal{D}_+$ and $\mathcal{D}_- \times \mathcal{D}_- \xrightarrow{\circ} \mathcal{D}_-$. That makes \mathcal{D}_+ and \mathcal{D}_- into subalgebras of \mathcal{D} .

Order and valuation

Similarly to the case of usual (commutative) formal Laurent series, we can define the *order* and the *valuation* of a pseudo-differential operator as follows.

Definition 3.1.3 (Order, valuation). Let $A = \sum a^k \partial^k$ be a nonzero pseudo-differential operator

- (i) We call *order* (or *degree*) of A the highest integer n such that $a^n \neq 0$, and denote it $n = \text{deg}(A)$. The order of the 0 operator is set as $\text{deg}(0) = -\infty$.
- (ii) We call *valuation* of A the least integer v such that $a^v \neq 0$, and denote it $v = \text{val}(A)$. If there is no such integer, then we set $\text{val}(A) = -\infty$. The valuation of the 0 operator is set as $\text{val}(0) = +\infty$;
- (iii) We say that A is *monic* if $a^n = 1$.

We will often use the “big O” notation as follows: For any integer $m \in \mathbb{Z}$, we say that A is a “big O” of ∂^m , and we write

$$A = O(\partial^m) \Leftrightarrow \text{deg } A \leq m.$$

In order to state the next proposition, it is necessary to set the following arithmetic rules for infinity. As usual, $+\infty$ (respectively $-\infty$) is set to be greater (respectively less) than any real number. Besides, we set $+\infty + \infty = +\infty$ and $-\infty - \infty = -\infty$. With these rules, and by integrality of the ring $\mathbb{C}[[x]]$, we get the following proposition.

Proposition 3.1.4. For any pseudo-differential operators A, B ,

$$\text{deg}(A + B) \leq \max(\text{deg}(A), \text{deg}(B)), \quad \text{val}(A + B) \geq \min(\text{val}(A), \text{val}(B)), \quad (3.3)$$

$$\text{deg}(A \circ B) = \text{deg}(A) + \text{deg}(B). \quad (3.4)$$

Remark 3.1.5. We cannot get a formula of the type $\text{val}(A \circ B) = \text{val}(A) + \text{val}(B)$ (which holds true for nonzero elements in the ring $R((T^{-1}))$, where R is a commutative integral domain). For example, $\partial \circ x = x\partial + 1$. So $\text{val}(\partial \circ x) = 0$, while $\text{val}(\partial) + \text{val}(x) = 1 + 0 = 1$.

When we do explicit computations of products of pseudo-differential operators, we often have to limit ourselves to computing some truncated part of said product. It is especially the case when we use computer programs. On the other hand, if one wants to compute one specific coefficient of a product, say $\text{res } AB$, then the computation must involve only finitely many coefficients of A and B . Therefore, it is helpful to know precisely “how much” we need to truncate the input operators A and B to compute $\text{res } AB$. The following proposition gives the optimal truncation needed to compute a product of operators down to some valuation.

Proposition 3.1.6 (Optimal valuations). *Let A, B be two pseudo-differential operators:*

$$A = \sum_{p \leq n_a} a^p \partial^p, \quad B = \sum_{p \leq n_b} b^p \partial^p,$$

i.e., A has order n_a and B of order n_b . Then for any integer $v \leq n_a + n_b$, the following truncated operators coincide:

$$(A \circ B)_{\geq v} = (A_{\geq v-n_b} \circ B_{\geq v-n_a})_{\geq v}.$$

Moreover, these valuations are optimal in the sense that for any $w_a \leq v - n_b$ and $w_b \leq v - n_a$, we get the same result:

$$(A_{\geq w_a} \circ B_{\geq w_b})_{\geq v} = (A_{\geq v-n_b} \circ B_{\geq v-n_a})_{\geq v}.$$

The way to interpret this proposition is that the coefficients of the operator $A \circ B$ down to some degree v depend only on $a^{n_a}, \dots, a^{v-n_b}$ and $b^{n_b}, \dots, b^{v-n_a}$.

Proof. The proof is straightforward. The product of A and B reads

$$A \circ B = \sum_{p \leq n_a} \sum_{q \leq n_b} \sum_{k \geq 0} \binom{p}{k} a^p b_{(k)}^q \partial^{p+q-k}.$$

So the coefficient of degree v of $A \circ B$ reads

$$\sum_{p \leq n_a} \sum_{q \leq n_b} \sum_{k \geq 0} \binom{p}{k} a^p b_{(k)}^q \delta_{p+q-k, v}.$$

We want to identify, in the above sum, the contribution in a^p and b^p with the least value of p . Clearly, this contribution only happens when $k = 0$, that is, in the sum

$$\sum_{p \leq n_a} \sum_{q \leq n_b} a^p b^q \delta_{p+q, v} = a^{n_a} b^{v-n_a} + a^{n_a-1} b^{v-n_a+1} + \dots + a^{v-n_b} b^{n_b}.$$

Therefore, the above sum only involves the terms $a^{n_a}, \dots, a^{v-n_b}$ and $b^{n_b}, \dots, b^{v-n_a}$. \square

As a direct consequence of the above proposition, we get the following corollary for raising an operator to some power, down to some valuation, with optimal valuations of the factors. This formula can be helpful when programming with pseudo-differential operators, since it gives an optimal knowledge of what is needed to compute.

Corollary 3.1.7. *Let A be a pseudo-differential operator of order $m \in \mathbb{Z}$.*

(i) *For any integers $q \geq 1$ and $v \leq qm$, the following formula uses optimal valuations:*

$$(A^q)_{\geq v} = \left(A_{\geq w} \circ (\cdots \circ (A_{\geq w} \circ A_{\geq w})_{\geq w+m} \cdots)_{\geq v-m} \right)_{\geq v}, \quad (3.5)$$

where $w = v - (q - 1)m = m + v - qm$.

(ii) *As a consequence, if we denote*

$$A = \sum_{k \leq m} a^k \partial^k, \quad A^q = \sum_{k \leq qm} a^{q \cdot k} \partial^k,$$

then the coefficient $a^{q \cdot v}$ is a differential polynomial in the coefficients a^m, \dots, a^{m+v-qm} .

In the above proposition, we called a *differential polynomial* in a^m, \dots, a^{m+v-qm} a polynomial in these functions and their derivatives. More generally, given functions u^1, \dots, u^n we call *differential polynomials* in u^1, \dots, u^n the elements of the polynomial algebra

$$\mathcal{A}(u^1, \dots, u^n) := \mathbb{C} \left[u_{(k)}^\alpha \mid 1 \leq \alpha \leq n, k \geq 0 \right].$$

The n -th root of an operator

In the study of integrable hierarchies, a key features of pseudo-differential operators is the existence of a n -th root for any monic operator of order $n \geq 1$. This will allow us to express the hierarchy in terms of the fractional powers of some pseudo-differential operator.

Theorem 3.1.8 (n -th Root). *Let A be a monic pseudo-differential operator of order $n \geq 1$, that is, of the form $A = \partial^n + \sum_{k < n} a^k \partial^k$.*

(i) *There exists a unique monic operator of order 1,*

$$P = \partial + \sum_{k < 1} p^k \partial^k,$$

such that $P^n = A$, called the n -th root of A .

(ii) *The computation of coefficients $p^0, \dots, p^k, k \leq 0$ amounts to equating the operators*

$$A_{\geq k+n-1} = \left(P_{\geq k} \circ (\cdots \circ (P_{\geq k} \circ P_{\geq k})_{\geq k+1} \cdots)_{\geq k+n-2} \right)_{\geq k+n-1}, \quad (3.6)$$

and p^k is a differential polynomial in $a^{n-1}, \dots, a^{k+n-1}$ of the form

$$p^k = \frac{1}{n} a^{k+n-1} + R(a^{n-1}, \dots, a^{k+n-2}).$$

Remark 3.1.9. If we relax the condition that P must be monic, then we have n distinct roots of the form $P = e^{2ik\pi/n}\partial + O(1)$, with $k \in \{0, \dots, n\}$. If moreover the operator A is not monic, i.e. $A = a^n \partial^n + O(\partial^{n-1})$, with $a^n \notin \{0, 1\}$, then there are n distinct roots of the form $P = (a^n)^{1/n} e^{2ik\pi/n} \partial + O(1)$.

Proof. We prove (ii). Denote $v = k + n - 1$ and assume that the operator P such that $P^n = A$ exists. Thanks to Corollary 3.1.7, we know that a^k is a differential polynomial in p^0, \dots, p^{1+k-n} (substituting $m = 1$ and $q = n$ in p^{m+v-qn} in 3.1.7(ii)), which implies that a^v is a differential polynomial in p^0, \dots, p^k . So if we want to compute a^v in terms of p^0, \dots, p^k , we only need to equate the two truncated operators

$$((P_{\geq k})^n)_{\geq v} = A_{\geq v}. \quad (3.7)$$

Then applying Equation (3.5), one finds Equation (3.6). On the other hand, developing Equation 3.7, we get

$$\begin{aligned} (P_{\geq k})^n &= \sum_{\substack{m_1, \dots, m_n \\ \in \{k, \dots, 1\}}} p^{m_1} \partial^{m_1} \circ \dots \circ p^{m_{n-1}} \partial^{m_{n-1}} \circ p^{m_n} \partial^{m_n} \\ &= \sum_{\substack{m_1, \dots, m_n \\ \in \{k, \dots, 1\}}} \sum_{k_1, \dots, k_{n-1} \geq 0} \binom{m_1}{k_1} \cdots \binom{m_{n-1}}{k_{n-1}} p^{m_1} \left(\dots \left(p^{m_{n-1}} p_{(k_{n-1})}^{m_n} \right)_{(k_{n-2})} \cdots \right)_{(k_1)} \partial^{\sum m_i - \sum k_i} \end{aligned}$$

Now if we select the term of order v , that is, for which $v = \sum m_i - \sum k_i$, we obtain that

$$a^v = \sum_{\substack{m_1, \dots, m_n \\ \in \{k, \dots, 1\}}} \sum_{\substack{k_1, \dots, k_{n-1} \geq 0 \\ v = \sum m_i - \sum k_i}} \binom{m_1}{k_1} \cdots \binom{m_{n-1}}{k_{n-1}} p^{m_1} \left(\dots \left(p^{m_{n-1}} p_{(k_{n-1})}^{m_n} \right)_{(k_{n-2})} \cdots \right)_{(k_1)}. \quad (3.8)$$

In the above equation, for the condition $v = \sum m_i - \sum k_i$ under the sum to be fulfilled, and since the k_i 's are nonnegative, it is necessary that

$$v - \sum_{i=1}^n m_i \leq 0. \quad (3.9)$$

Now we want to extract the coefficient p^k from the big sum in Equation (3.8). To do so, let $i_0 \in \{1, \dots, n\}$ be such that $m_{i_0} = k$. Then all the other m_i 's must equal 1 since then, by definition of v ,

$$v - \sum_{i=1}^n m_i = v - (n - 1 + k) = 0,$$

which saturates Equation (3.9). As a saturation, it fixes all the k_i 's to 0, and we have n possible choices for $m_{i_0} = k$. Therefore,

$$p^k = \frac{1}{n} a^v - \frac{1}{n} \sum_{\substack{m_1, \dots, m_n \\ \in \{k+1, \dots, 1\}}} \sum_{\substack{k_1, \dots, k_{n-1} \geq 0 \\ v = \sum m_i - \sum k_i}} \binom{m_1}{k_1} \cdots \binom{m_{n-1}}{k_{n-1}} p^{m_1} \left(\dots \left(p^{m_{n-1}} p_{(k_{n-1})}^{m_n} \right)_{(k_{n-2})} \cdots \right)_{(k_1)}.$$

(In the first sum, the interval $\{k, \dots, 1\}$ has been changed into $\{k+1, \dots, 1\}$.) Consequently, by recursion on $k \leq 0$, the coefficient p^k is a differential polynomial in $a^{n-1}, \dots, a^{k+n-1}$ of the form $p^k = \frac{1}{n}a^{k+n-1} + R(a^{n-1}, \dots, a^{k+n-2})$. \square

Example 3.1.10. The square root of the so-called Schrödinger operator $L = \partial^2 + u$ reads

$$(\partial^2 + u)^{1/2} = \partial + \frac{1}{2}u\partial^{-1} - \frac{1}{4}u_{(1)}\partial^{-2} + \frac{1}{8}(u_{(2)} - (u)^2)\partial^{-3} + \dots \quad (3.10)$$

Multiplicative inverse

Another important property of pseudo-differential operators that is essential in the study of integrable hierarchies is the existence of a unique multiplicative inverse, as long as the leading coefficient of said operator is invertible within the ring $\mathbb{C}[[x]]$. The following theorem makes a more precise statement.

Theorem 3.1.11 (Multiplicative inverse). *Let A be a monic operator of order $n \in \mathbb{Z}$, that is, of the form $A = \partial^n + \sum_{k < n} a^k \partial^k$.*

(i) *There exists a unique monic operator of order $-n$ that we denote*

$$A^{-1} = \partial^{-n} + \sum_{k < -n} \tilde{a}^k \partial^k,$$

such that $AA^{-1} = A^{-1}A = 1$.

(ii) *The computation of coefficients $\tilde{a}^{-n-1}, \dots, \tilde{a}^k$, with $k < -n$, amounts to equating the operators*

$$(A_{\geq 2n+k} \circ (A^{-1})_{\geq k})_{\geq k+n} = 1,$$

and \tilde{a}^k is a differential polynomial in a^{n-1}, \dots, a^{2n+k} of the form

$$\tilde{a}^k = -a^{2n+k} + J(a^n, \dots, a^{2n+k-1}).$$

The first few terms of A^{-1} read

$$A^{-1} = 1 - a^{n-1}\partial^{-1} - \left(a^{n-2} - (a^{n-1})^2 - na_{(1)}^{n-1}\right)\partial^{-2} - \dots$$

Remark 3.1.12. If the operator A is not monic, i.e. if $A = a^n \partial^n + \mathcal{O}(\partial^{n-1})$, with $a^n \notin \{0, 1\}$, then by denoting $A = a^n B$ (with B monic), we easily see that A 's inverse is $A^{-1} = (a^n)^{-1} B^{-1}$.

Proof. We prove point (ii). Let $a^n = \tilde{a}^{-n} = 1$ and let $k < -n$ (so that $2n+k < n$). Let us assume that there exists such an operator A^{-1} . Then by Proposition 3.1.6, we know that

$$(A \circ A^{-1})_{\geq k+n} = (A_{\geq 2n+k} \circ (A^{-1})_{\geq k})_{\geq k+n}.$$

Consequently, in order to compute coefficient \tilde{a}^k , it is sufficient to equate the truncated operators above. We get

$$\begin{aligned} A_{\geq 2n+k} \circ (A^{-1})_{\geq k} &= \sum_{m=2n+k}^n \sum_{m'=k}^{-n} a^m \partial^m \circ \tilde{a}^{m'} \partial^{m'} \\ &= \sum_{m=2n+k}^n \sum_{m'=k}^{-n} \sum_{\ell \geq 0} \binom{m}{\ell} a^m \tilde{a}_{(\ell)}^{m'} \partial^{m+m'-\ell}. \end{aligned}$$

Clearly, the equation for $k = -n$ simply reads $1 = 1$. So set $k < -n$. If we select the term of degree $n+k$, we obtain

$$\begin{aligned} 0 &= \sum_{m=2n+k}^n \sum_{m'=k}^{-n} \sum_{\ell \geq 0} \binom{m}{\ell} a^m \tilde{a}_{(\ell)}^{m'} \delta_{m+m'-\ell, n+k} \\ &= \sum_{m=2n+k}^n \sum_{m'=k}^{-n} \binom{m}{m+m'-n-k} a^m \tilde{a}_{(m+m'-n-k)}^{m'} \mathbf{1}_{m+m'-n-k \geq 0}, \end{aligned}$$

where the symbol $\mathbf{1}_{m+m'-n-k \geq 0}$ means that we select only the terms for which $m+m'-n-k \geq 0$. If one saturates this inequality, i.e. imposes $m+m' = n+k$, then the least admissible value of m' is $m' = k$, which imposes $m = n$. Consequently, the above expression contains the term $a^n \tilde{a}^k = \tilde{a}^k$, and that is the only term involving \tilde{a}^k . In other words,

$$\tilde{a}^k = - \sum_{m=2n+k}^n \sum_{m'=k+1}^{-n} \binom{m}{m+m'-n-k} a^m \tilde{a}_{(m+m'-n-k)}^{m'} \mathbf{1}_{m+m'-n-k \geq 0}.$$

By recursion on $k < -n$, we get that \tilde{a}^k is a differential polynomial in a^{n-1}, \dots, a^{2n+k} .

Finally, in the above equation, if $m = 2n+k$, then the inequality $m+m'-n-k \geq 0$ saturates and imposes $m' = -n$. Consequently,

$$\tilde{a}^k = -a^{2n+k} + J(a^n, \dots, a^{2n+k-1}),$$

where J is a differential polynomial that does not depend on a^{2n+k} . One can check that A and A^{-1} commute. \square

Formal adjoint

Proposition 3.1.13 (Formal adjoint). *The algebra \mathcal{D} has a natural involution $(-)^* : \mathcal{D} \rightarrow \mathcal{D}$, called the formal adjoint, defined on monomials by*

$$(a\partial^k)^* = (-1)^k \partial^k \circ a, \quad (3.11)$$

and extended on \mathcal{D} by linearity. It is an anti-homomorphism, i.e. for any $A, B \in \mathcal{D}$,

$$(AB)^* = B^*A^*.$$

Notice that since $(A^*)^{-1}A^* = 1$ and $(A^{-1})^*A^* = (AA^{-1})^* = 1$, then

$$(A^*)^{-1} = (A^{-1})^*.$$

Proof. We prove the anti-homomorphism property on monomials. We have,

$$(a\partial^n \circ b\partial^p)^* = \sum_{k \geq 0} \binom{n}{k} (ab_k \partial^{n+p-k})^* = (-1)^{n+p} \sum_{k \geq 0} (-1)^k \binom{n}{k} \partial^{n+p-k} \circ (ab_k).$$

On the other hand,

$$(b\partial^p)^* \circ (a\partial^n)^* = (-1)^{n+p} \partial^p \circ b\partial^n \circ a,$$

so that it amounts to proving that

$$\partial^p \circ b\partial^n \circ a = \sum_{k \geq 0} (-1)^k \binom{n}{k} \partial^{n+p-k} \circ (ab_k). \quad (3.12)$$

Now, using the relation $\partial^n \circ b = \sum_{k \geq 0} \binom{n}{k} b_k \partial^{n-k}$, we can write

$$\partial^p \circ b\partial^n \circ a = \partial^{n+p} \circ (ab) - \sum_{k \geq 1} \binom{n}{k} \partial^p \circ b_k \partial^{n-k} \circ a.$$

Using the same argument, we get that

$$\partial^p \circ b\partial^n \circ a = \partial^{n+p} \circ (ab) - \binom{n}{1} \partial^{n+p-1} \circ (ab_1) + \sum_{k \geq 2} \binom{n}{k} \partial^p \circ b_k \partial^{n-k} \circ a.$$

Using the same argument recursively, we prove Equation (3.12). \square

Lie algebra structure

Here we discuss the structure of Lie algebra with which \mathcal{D} is naturally equipped. For more details on Lie algebras in general, see Section 4.1. As an associative algebra, \mathcal{D} is naturally equipped with a structure of Lie algebra by defining the bracket, for any operators $A, B \in \mathcal{D}$,

$$[A, B] = A \circ B - B \circ A.$$

Because this bracket is the commutator, it immediately satisfies the skew-symmetry property, $[A, B] + [B, A] = 0$, along with Jacobi's identity:

$$[A, [B, C]] + [B, [C, A]] + [C, [A, B]] = 0. \quad (3.13)$$

It is worth noticing, by a direct check, that the commutator $[A, B]$ contains no term of degree $\deg A + \deg B$, so that

$$\deg [A, B] \leq \deg A + \deg B - 1. \quad (3.14)$$

As in the case of linear Lie algebras, the following proposition holds.

Proposition 3.1.14. For any operator $A \in \mathcal{D}$, the adjoint mapping $\text{ad}_A : B \mapsto [A, B]$ is a derivation of the algebra \mathcal{D} for the product \circ , i.e. for any $B, C \in \mathcal{D}$,

$$[A, BC] = B[A, C] + [A, B]C.$$

As a direct consequence of Proposition 3.1.14, and similarly to the case of linear Lie algebras, we get the following corollary by recursion on $q \geq 1$.

Corollary 3.1.15. For any operators $A, B \in \mathcal{D}$ and any integer $q \geq 1$, the following equation holds,

$$[A^q, B] = A^{q-1}[A, B] + A^{q-2}[A, B]A + \cdots + [A, B]A^{q-1}.$$

In turns, Corollary 3.1.15 implies the following proposition which is necessary in order to inject the Gelfand–Dickey hierarchies, a fortiori the KdV hierarchy, into the KP hierarchy.

Corollary 3.1.16. Let $A, B \in \mathcal{D}$ be two operators. There exists an integer $q \geq 1$ such that $[A^q, B] = 0$ if and only if $[A, B] = 0$.

Proof. Let us assume that $[A, B] \neq 0$ and prove that it implies that $[A^q, B] \neq 0$ for any $q \geq 1$. We denote $A = \sum_{k \leq n} a^k \partial^k$ and $[A, B] = \sum_{k \leq m} c^k \partial^k$, where $a^n \neq 0$ and $c^m \neq 0$. Then

$$[A^q, B] = A^{q-1}[A, B] + A^{q-2}[A, B]A + \cdots + [A, B]A^{q-1},$$

has leading term $q(a^n)^{q-1}c^m$, which cannot vanish. Conversely, if $[A, B] = 0$, then for any $q \geq 1$, $[A^q, B] = 0$. \square

Local functionals

Recall that we set our functions to be formal power series in some indeterminate x . The derivation $\partial = \partial_x$ on $\mathbb{C}[[x]]$ defines a quotient space

$$\overline{\mathbb{C}[[x]]} := \mathbb{C}[[x]] / (\text{Im}(\partial) \oplus \mathbb{C}1),$$

whose elements are called *local functionals*. We denote by $\int : \mathbb{C}[[x]] \rightarrow \overline{\mathbb{C}[[x]]}$ the canonical projection onto local functionals; the class of an element $f \in \mathbb{C}[[x]]$ is denoted alternatively by

$$\bar{f} = \int f = \int f(x)dx \in \overline{\mathbb{C}[[x]]}.$$

Notice that $\overline{\mathbb{C}[[x]]}$ is not a ring for these equivalence classes do not respect the product in $\mathbb{C}[[x]]$. As a direct consequence, we get a “integration by parts” formula: $\int f g_{(1)} = -\int g f_{(1)}$, or more generally,

$$\int f g_{(k)} = (-1)^k \int g f_{(k)}. \quad (3.15)$$

The space of local functionals plays a pivotal role in the Hamiltonian structure of integrable hierarchies. A first step towards it is the Adler trace, introduced by Mark Adler in [Adl79].

Proposition 3.1.17 (Adler trace). *The Lie algebra \mathcal{D} is naturally equipped with a linear form $\langle - \rangle : \mathcal{D} \rightarrow \mathbb{C}[[x]]$, called the Adler trace, defined by, for any $A \in \mathcal{D}$,*

$$\langle A \rangle = \int \operatorname{res} A \in \overline{\mathbb{C}[[x]]},$$

which satisfies the fundamental trace property $\langle [A, B] \rangle = 0$. The Adler trace produces a non-degenerate symmetric bilinear form (a scalar product):

$$(A, B) = \langle AB \rangle,$$

that is ad-invariant, i.e. $(A, [B, C]) = ([A, B], C)$. This scalar product defines an isomorphism of vector spaces $(\mathcal{D}_+)^ \simeq \mathcal{D}_-$.*

Proof. We start by proving the trace property, which we can prove on monomials by linearity. Clearly, if both $A, B \in \mathcal{D}_+$ or both $A, B \in \mathcal{D}_-$, then $\langle AB \rangle = 0$. Let us assume that $A = a\partial^n$, $n \geq 0$ and $B = b\partial^p$, $p \leq -1$ and $n + p \geq -1$, then

$$\begin{aligned} \langle AB \rangle &= \int \operatorname{res} a\partial^n b\partial^p \\ &= \int \operatorname{res} \sum_{k \geq 0} \binom{n}{k} ab_{(k)} \partial^{n+p-k} \\ &= \int \binom{n}{n+p+1} ab_{(n+p+1)}. \end{aligned}$$

On the other hand,

$$\begin{aligned} \langle BA \rangle &= \int \operatorname{res} b\partial^p a\partial^n \\ &= \int \binom{p}{n+p+1} ba_{(n+p+1)} \\ &= \int (-1)^{n+p+1} \binom{p}{n+p+1} ab_{(n+p+1)}, \end{aligned}$$

where we have used the integration part property (5.2). Now thanks to the symmetry property of binomial coefficients (3.2), the two expressions under the integral sign coincide.

The ad-invariance property follows from the trace property. The non-degeneracy follows from the fact that a nonzero monomial $a\partial^n$ is non-degenerately paired with any nonzero monomial $\partial^{-n-1}b$ (as long as $\bar{ab} \notin \operatorname{Im}\partial$):

$$\langle (a\partial^n) \circ (\partial^{-n-1}b) \rangle = \int \operatorname{res} a\partial^{-1}b = \int ab$$

The above proves the duality $(\mathcal{D}_+)^* \simeq \mathcal{D}_-$. □

3.2 The Kadomtsev–Petviashvili hierarchy

In this section we define the celebrated Kadomtsev–Petviashvili (KP) hierarchy, which first appeared in [Sat81]. Even though the main goal of this chapter is to discuss the polynomial tau functions of KdV (rather than KP), we will see that KP will be our main tool in proving our statements. Moreover, once the main theorems on KP are proved, then most of the job is done. The KdV hierarchy then appears as a particular case of the KP hierarchy.

Let us first have a preliminary discussion. As we will see, a solution to the equations of the KP hierarchy is expressed as a single pseudo-differential operator $Q \in \mathcal{D}$ of the form

$$Q = \partial + \sum_{k < 0} q^k \partial^k = \partial + q^{-1} \partial^{-1} + q^{-2} \partial^{-2} + \dots \quad (3.16)$$

In order to be able to express the KP hierarchy, we ask the functions q^i to depend on infinitely many variables

$$\mathbf{t} := (x; t_1, t_2, t_3, \dots).$$

To avoid problems of convergence, we restrict our space of functions to setting

$$q^i \in \mathbb{C}[[\mathbf{t}]] := \mathbb{C}[[x, t_1, t_2, t_3, \dots]].$$

Thereby, we've extended our space of functions; in other words, from now on,

$$\mathcal{D} = \text{DO}(\mathbb{C}[[\mathbf{t}]]), \partial = \partial_x).$$

This change does not affect any of the results given above.

We naturally define the action of ∂_{t_i} on pseudo-differential operators by setting, for any operator $A = \sum_{k \leq n} a^k \partial^k \in \mathcal{D}$,

$$\frac{\partial A}{\partial t_i} := \sum_{k \leq n} \frac{\partial a^k}{\partial t_i} \partial^k.$$

This action automatically satisfies Leibniz's rule

$$\frac{\partial}{\partial t_i}(A \circ B) = \frac{\partial A}{\partial t_i} \circ B + A \circ \frac{\partial B}{\partial t_i}.$$

In particular, the action of ∂_{t_i} commutes with the composition by ∂ :

$$\frac{\partial}{\partial t_i}(\partial \circ A) = \partial \circ \frac{\partial A}{\partial t_i}.$$

3.2.1 Lax pairs of the KP hierarchy

Our first theorem gives a definition of the KP hierarchy in its pseudo-differential operators formulation, one of many ways to define the hierarchy.

Theorem 3.2.1 (KP hierarchy [Sat81]). *Let $Q \in \mathcal{D}$ be an operator of the form*

$$Q = \partial + q^{-1}\partial^{-1} + q^{-2}\partial^{-2} + \cdots. \quad (3.17)$$

Then the following flows commute pairwise, producing the so-called KP hierarchy:

$$\frac{\partial Q}{\partial t_i} = [(Q^i)_+, Q], \quad i \in \mathbb{Z}_{\geq 1}. \quad (3.18)$$

The operator Q is often called the *Lax operator* of the KP hierarchy. Equations (3.18) are said to be equations of *Lax type* and $(Q, (Q^i)_+)$ is called a Lax pair.

Proof. (of Theorem 3.2.1). We give the same proof as in [DJM00], pp. 14–15 (for the KdV case). First of all since $[Q^i, Q] = 0$, we have

$$[(Q^i)_+, Q] = -[(Q^i)_-, Q].$$

We easily check that $\deg [(Q^i)_-, Q] = \deg \partial_{t_i} Q = -1$, so that the equations of the hierarchy are well defined. The flows of Equation (3.18) commute if for any $i, j \geq 1$, $\partial_{t_i} \partial_{t_j} Q = \partial_{t_j} \partial_{t_i} Q$. Let us write

$$\partial_{t_i} \partial_{t_j} Q = \partial_{t_i} [(Q^j)_+, Q] = [\partial_{t_i} (Q^j)_+, Q] + [(Q^j)_+, \partial_{t_i} Q].$$

First, it is easy to see that

$$\partial_{t_i} (Q^j)_+ = (\partial_{t_i} Q^j)_+ = [(Q^i)_+, Q^j]_+.$$

Indeed, prior to projecting onto the subspace \mathcal{D}_+ , we have

$$\begin{aligned} \partial_{t_i} Q^j &= (\partial_{t_i} Q) Q^{j-1} + Q (\partial_{t_i} Q) Q^{j-2} + \cdots + Q^{j-1} (\partial_{t_i} Q) \\ &= [(Q^i)_+, Q^j]. \end{aligned}$$

Therefore,

$$\partial_{t_i} \partial_{t_j} Q = \left[[(Q^i)_+, Q^j]_+, Q \right] + \left[(Q^j)_+, [(Q^i)_+, Q] \right]. \quad (3.19)$$

On the other hand, using the fact that for any $A, B \in \mathcal{D}$, $[A_+, B_+]_+ = [A_+, B_+]$ and $[A_-, B_-]_+ = 0$, we see that

$$\begin{aligned} [(Q^i)_+, Q^j]_+ &= [(Q^i)_+, (Q^j)_+]_+ + [(Q^i)_+, (Q^j)_-]_+ \\ &= [(Q^i)_+, (Q^j)_+] + [Q^i, (Q^j)_-]_+ \\ &= [(Q^i)_+, (Q^j)_+] - [Q^i, (Q^j)_+]_+. \end{aligned}$$

Therefore,

$$\begin{aligned} \partial_{t_i} \partial_{t_j} Q &= \left[[(Q^i)_+, (Q^j)_+]_+, Q \right] - \left[[Q^i, (Q^j)_-]_+, Q \right] + \left[(Q^j)_+, [(Q^i)_+, Q] \right] \\ &= \left[[(Q^i)_+, (Q^j)_+]_+, Q \right] + \left[Q, [(Q^j)_+, (Q^i)_+] \right] + \left[(Q^j)_+, [(Q^i)_+, Q] \right]. \end{aligned}$$

Now using Jacobi's identity (3.13) on the two rightmost terms of the above equation, we get

$$\partial_{t_i} \partial_{t_j} Q = \left[[(Q^j)_+, Q^i]_+, Q \right] + \left[(Q^i)_+, [(Q^j)_+, Q] \right].$$

According to Equation (3.19), we just proved that $\partial_{t_i} \partial_{t_j} Q = \partial_{t_j} \partial_{t_i} Q$. \square

The flow along t_1 straightforwardly reads $\partial_{t_1} Q = [\partial, Q] = \partial(Q)$, so that $\partial_{t_1} = \partial$. As a convention, we set

$$t_1 := x.$$

The first two flows along t_2 gives the following equations

$$\begin{aligned} \partial_{t_2} q^{-1} &= q_{(2)}^{-1} + 2q_{(1)}^{-2}, \\ \partial_{t_2} q^{-2} &= q_{(2)}^{-2} + 2q_{(1)}^{-3} + 2q^{-1} q_{(1)}^{-1}. \end{aligned}$$

Similarly, the flows along t_3 gives

$$\partial_{t_3} q^{-1} = q_{(3)}^{-1} + 3q_{(2)}^{-2} + 3q_{(1)}^{-3} + 6q^{-1} q_{(1)}^{-1}.$$

Eliminating q^{-2} and q^{-3} plus renaming $v = -2q^{-1}$, $t_2 = y$ and $t_3 = t$, we find

$$\partial_y^2 v = \frac{4}{3} \partial_x \left(\partial_t v + \frac{9}{4} v \partial_x v - \frac{3}{4} \partial_x^3 v \right).$$

The above equation is called the *KP equation* (if we also rename $t_2 = y$ and $t_3 = t$), introduced by Kadomtsev and Petviashvili in [KP70]; hence the *KP hierarchy*. In the next proposition, we give infinitely many conserved quantities expressed via the Adler trace.

Proposition 3.2.2. *The local functionals $H_j = \langle Q^j \rangle$ are conserved quantities of the KP hierarchy in the sense that for any $i \geq 1$,*

$$\frac{\partial H_j}{\partial t_i} = \int \text{res} \frac{\partial Q^j}{\partial t_i} = 0.$$

Proof. Since $\partial = \partial_x = \partial_{t_1}$ commute with the other ∂_{t_i} 's, we can permute ∂ and the integral sign so that, by linearity of the residue map, $\partial_{t_i} H_j = \int \text{res} \partial_{t_i} Q^j$. Then by Equation (3.18),

$$\frac{\partial H_j}{\partial t_i} = \int \text{res} [(Q^i)_+, Q^j].$$

Yet by Proposition 3.1.17, the above vanishes. \square

3.2.2 Equivalent forms of the KP hierarchy

The KP hierarchy admits multiple formulations: the isospectral deformation system, Sato's equation, the bilinear identity. But all these rely on the existence of the so-called "dressing operator." The idea is to "dress up" the operator ∂ with the operator Q via an operator M : $Q = M \partial M^{-1}$. Then finding the operator M amounts to finding Q .

Proposition 3.2.3 (Dressing operator). *Let Q be an operator of the form of the KP Lax operator, that is, $Q = \partial + \sum_{k < 0} q^k \partial^k$. There exists an operator M , called the dressing operator, of the form*

$$M = 1 + \sum_{k < 0} m^k \partial^k, \quad \text{such that } Q = M \partial M^{-1}. \quad (3.20)$$

Moreover, the operator M is determined by Q up to a transformation of the form $M \mapsto M \circ C$, with $C = 1 + \sum_{k < 0} c^k \partial^k$, where the functions c^k do not depend on x (i.e. $\partial(c^k) = 0$).

Proof. First, since $\deg M = 0$, then $\deg M^{-1} = 0$ too. Therefore, $\deg M \partial M^{-1} = 1 = \deg Q$, so that the equation $Q = M \partial M^{-1}$ is meaningful. Second, if $N = MC$, then

$$N \partial N^{-1} = MC \partial C^{-1} M^{-1} = MC (\partial(C^{-1}) + C^{-1}) \partial M^{-1} = M \partial M^{-1},$$

where $\partial(C^{-1})$ stands for the operator which coefficient of degree i is the derivative of the coefficient of degree i of C^{-1} (which are constant too). Consequently, transformation $M \mapsto MC$ leaves Q invariant. To prove that Q determines M (up to $M \mapsto MC$), it suffices to equate $M \partial = QM$. Then by methods similar to the proofs of Theorems 3.1.8 and 3.1.11, we find that the coefficients m^k are determined recursively. \square

Now that we have the existence of the dressing operator, we can translate the flows of the KP hierarchy on Q into the flows on M . The latter is called Sato's equation.

Theorem 3.2.4 (Sato's equation). *Let Q be an operator of the form of Equation 3.16 and let M be an operator such that $Q = M \partial M^{-1}$. Then the operator Q satisfies the KP hierarchy (3.18) if and only if the operator M satisfies the following hierarchy of equations, called Sato's equations*

$$\frac{\partial M}{\partial t_i} = -(M \partial^i M^{-1})_- M, \quad i \geq 1. \quad (3.21)$$

Proof. Let us start from Equation (3.21) on M and prove that, equivalently, $Q = M \partial M^{-1}$ satisfies Equation (3.18). First, from Equation (3.21) it is not hard to see that $\partial_{t_i} M^{-1} = M^{-1} (M \partial^i M^{-1})_-$. Therefore, we can write the flows of M and M^{-1} as

$$\frac{\partial M}{\partial t_i} = -(Q^i)_- M, \quad \frac{\partial M^{-1}}{\partial t_i} = M^{-1} (Q^i)_-.$$

Then, by definition of M ,

$$\begin{aligned} \frac{\partial Q}{\partial t_i} &= \frac{\partial M}{\partial t_i} \partial M^{-1} + M \partial \frac{\partial M^{-1}}{\partial t_i} \\ &= -(Q^i)_- M \partial M^{-1} + M \partial M^{-1} (Q^i)_- \\ &= [Q, (Q^i)_-]. \end{aligned}$$

Clearly, these equations follow from one another by equivalences. \square

Our next important theorem is the equivalence between the KP hierarchy and what is called the *linear system of isospectral deformation* of KP (or simply its *linear system*). But before stating the theorem, we need to define an action of pseudo-differential on some space of functions as follows. First we introduce the so-called *xi function*:

$$\xi(\mathbf{t}; \lambda) = \sum_{i \geq 1} t_i \lambda^i \quad (3.22)$$

Then, for any integer $k \in \mathbb{Z}$, we set $\partial^k e^{\xi(\mathbf{t}; \lambda)} := \lambda^k e^{\xi(\mathbf{t}; \lambda)}$, which naturally extends the usual definition for $k \geq 0$. We then extend this definition on \mathcal{D} and get that, for $A = \sum_{k \leq n} a^k \partial^k \in \mathcal{D}$,

$$A e^{\xi(\mathbf{t}; \lambda)} = \left(\sum_{k \leq n} a^k \lambda^k \right) e^{\xi(\mathbf{t}; \lambda)}.$$

Given this definition, we can define the following “function space”:

$$\mathcal{W} = \left\{ A e^{\xi(\mathbf{t}; \lambda)} \mid A \in \mathcal{D} \right\} \subset \mathbb{C}[[\mathbf{t}, \lambda, \lambda^{-1}]].$$

This space is, of course, completely ad hoc. It is also clearly in one-to-one correspondence with \mathcal{D} . We can define an action of \mathcal{D} on \mathcal{W} by

$$A \left(B e^{\xi(\mathbf{t}; \lambda)} \right) = A \circ B \left(e^{\xi(\mathbf{t}; \lambda)} \right).$$

We can now state the following theorem.

Theorem 3.2.5 (KP linear system). *The KP hierarchy (3.18) on the operator Q holds if and only if there exists a function $w \in \mathcal{W}$ such that the following hierarchy of equations is satisfied,*

$$\begin{cases} Qw = \lambda w; \\ \partial_{t_i} w = (Q^i)_+ w; \end{cases} \quad i \geq 1, \lambda \in \mathbb{C}. \quad (3.23)$$

This system is called the linear system of KP; the complex number λ is called the spectral parameter; the function w is called the wave function associated to λ .

Proof. Suppose there exists a function $w \in \mathcal{W}$ such that (3.23) holds. Then

$$\begin{aligned} \frac{\partial Q}{\partial t_i} w + Q \frac{\partial w}{\partial t_i} &= \lambda \frac{\partial w}{\partial t_i} \\ \frac{\partial Q}{\partial t_i} w + Q(Q^i)_+ &= (Q^i)_+ Q w \\ \left(\frac{\partial Q}{\partial t_i} + [Q, (Q^i)_+] \right) w &= 0. \end{aligned}$$

The above equation holds for arbitrary values of $\lambda \in \mathbb{C}$, which means that the operator $\partial_{t_i} Q + [Q, (Q^i)_+]$ has infinitely many independent solutions (one $w(\lambda)$ for each value of λ), while being independent of λ . Therefore, it vanishes, i.e

$$\frac{\partial Q}{\partial t_i} + [Q, (Q^i)_+] = 0.$$

We now prove the converse. Let an operator $M = 1 + \sum_{k<0} m^k \partial^k$ be such that $Q = M\partial M^{-1}$. We will prove that the function $w = M e^{\xi(t;\lambda)}$ is a wave function of the linear system of Kp. Notice that by construction of the xi function,

$$\partial_{t_i} e^{\xi(t;\lambda)} = \partial^i e^{\xi(t;\lambda)} = \lambda^i e^{\xi(t;\lambda)}.$$

Now we easily see that

$$Qw = M\partial e^{\xi(t;\lambda)} = \lambda w,$$

so that the function w is a eigenvector with eigenvalue λ . Now considering Sato's equation (3.21) on the dressing operator M , we get that

$$\begin{aligned} \frac{\partial w}{\partial t_i} &= \frac{\partial M}{\partial t_i} e^{\xi(t;\lambda)} + M \frac{\partial e^{\xi(t;\lambda)}}{\partial t_i} \\ &= -(Q^i)_- M - e^{\xi(t;\lambda)} + M \partial^i e^{\xi(t;\lambda)} \\ &= -(Q^i)_- w + M \partial^i M^{-1} w \\ &= -(Q^i)_- w + Q^i w = (Q^i)_+ w. \end{aligned}$$

Therefore, there exist an eigenfunction w such that the linear system of KP holds. \square

3.3 The bilinear identity and tau functions

3.3.1 The bilinear identity

In this section, we follow the approach of [BBT03]. In the proof of Theorem 3.2.5, we saw an important fact: Let Q satisfy the KP hierarchy, then any dressing operator $M = 1 + \sum_{k<0} m^k \partial^k$, i.e. such that $Q = M\partial M^{-1}$, provides a wave function of the linear system of KP by setting

$$w(t; \lambda) := M e^{\xi(t;\lambda)} = \left(1 + \frac{m^{-1}}{\lambda} + \frac{m^{-2}}{\lambda^2} + \dots \right) e^{\xi(t;\lambda)}, \quad (3.24)$$

On the other hand, since the KP hierarchy is equivalent to the *existence* of a wave function, we can chose the wave function to be of the form of Equation (3.51) without loss of generality. We also define the *adjoint wave function* by

$$w^*(t; \lambda) := (M^*)^{-1} e^{-\xi(t;\lambda)}.$$

We get the following lemma.

Lemma 3.3.1. *The adjoint wave function obeys the following hierarchy of equation:*

$$\begin{cases} Q^* w = \lambda w^*; \\ \partial_{t_i} w^* = -(Q^i)_+^* w^*; \end{cases} \quad i \in \mathbb{Z}. \quad (3.25)$$

Proof. By properties of the formal adjoint, we have that $Q^* = -(M^{-1})^* \partial M^*$ (since $\partial^* = -\partial$), so that

$$Q^* w^* = -(M^{-1})^* \partial e^{-\xi(\mathbf{t}; \lambda)} = \lambda w^*.$$

On the other hand, since $(M^{-1})^* = (M^*)^{-1}$, we have

$$\begin{aligned} \partial_{t_i} w^* &= (\partial_{t_i} M^{-1})^* e^{-\xi(\mathbf{t}; \lambda)} + (M^*)^{-1} \partial_{t_i} e^{-\xi(\mathbf{t}; \lambda)} \\ &= (M^{-1} (Q^i)_-)^* e^{-\xi(\mathbf{t}; \lambda)} + (M^*)^{-1} \partial^i e^{-\xi(\mathbf{t}; \lambda)} \\ &= (Q^i)_-^* (M^*)^{-1} e^{-\xi(\mathbf{t}; \lambda)} - (Q^i)^* (M^*)^{-1} e^{-\xi(\mathbf{t}; \lambda)} = -(Q^i)_+^* w^*. \end{aligned}$$

□

We can now state the celebrated bilinear identity, first introduced by R. Hirota's in his lecture notes [Hir76]. This theorem expresses an equivalence between the KP hierarchy and a bilinear formulation of the latter.

Theorem 3.3.2 (Bilinear identity). *Let Q be an operator of the form (3.16) and let M be a dressing operator for Q . Let w be the function*

$$w(\mathbf{t}; \lambda) = M e^{\xi(\mathbf{t}; \lambda)} \in \mathcal{W}.$$

Then w is a wave function of the linear system of KP (3.23) (i.e., Q satisfies the KP hierarchy (3.18)) if and only if w satisfies the following hierarchy of bilinear equations:

$$\oint \frac{d\lambda}{2i\pi} \left(\partial_{t_1}^{j_1} \cdots \partial_{t_m}^{j_m} w(\mathbf{t}; \lambda) \right) \cdot w^*(\mathbf{t}'; \lambda) = 0, \quad (3.26)$$

for any $m \geq 0$ and $j_1, \dots, j_m \geq 0$. Moreover, it is equivalent to the identity

$$\oint \frac{d\lambda}{2i\pi} w(\mathbf{t}; \lambda) \cdot w^*(\mathbf{t}'; \lambda) = 0, \quad \forall \mathbf{t}, \mathbf{t}'.$$

In the above expressions, the contour integral denotes the residue around $\lambda = \infty$, i.e. the coefficient in λ^{-1} .

Remark 3.3.3. Everything we state is purely formal. In this fashion, it does not really make sense to say “ $\forall \mathbf{t}, \mathbf{t}'$.” What we actually mean by that is that the second form of the bilinear identity is an equality holding within the ring $\mathbb{C}[[\mathbf{t}, \mathbf{t}', \lambda, \lambda^{-1}]]$.

We proceed by a series of lemmas.

Lemma 3.3.4. *For any operators $A, B \in \mathcal{D}$,*

$$\oint \frac{d\lambda}{2i\pi} A(e^{\lambda x}) \cdot B(e^{-\lambda x}) = \text{res}_\partial (AB^*).$$

Proof. Let us denote $A = \sum_{k \leq n_a} a^k \partial^k$ and $B = \sum_{k \leq n_b} b^k \partial^k$. The left hand-side of the above equality reads

$$\oint \frac{d\lambda}{2i\pi} A(e^{\lambda x}) \cdot B(e^{-\lambda x}) = \operatorname{res}_\lambda \sum_{k \leq n_a} a^k \lambda^k \sum_{\ell \leq n_b} b^\ell (-1)^\ell \lambda^\ell = \sum_{\ell \leq n_b} (-1)^\ell a^{-\ell-1} b^\ell.$$

The left-hand side reads

$$\operatorname{res}_\partial (AB^*) = \operatorname{res}_\partial \sum_{k \leq n_a} a^k \partial^k \circ \sum_{\ell \leq n_b} (-1)^\ell \partial^\ell b^\ell = \sum_{\ell \leq n_b} (-1)^\ell a^{-\ell-1} b^\ell.$$

□

Lemma 3.3.5. *If Q satisfies the KP hierarchy (3.18) (i.e. if w satisfies (3.23)), then w satisfies the bilinear identity (3.26).*

Proof. Since w satisfies (3.23), we have $\partial_{t_i} w = (Q^i)_+ w$. And since $(Q^i)_+$ is a polynomial in ∂ , any term of the form $\partial_{t_1}^{j_1} \cdots \partial_{t_m}^{j_m} w$ eventually reads a polynomial in the $\partial^k(w)$'s. Consequently, it is sufficient to prove the bilinear identity for integrand $\partial^k(w) \cdot w^*$, with $k \geq 0$. Yet because $w = Me^{\xi(t;\lambda)}$ and because M only acts on $t_1 = x$, we can write

$$\begin{aligned} \oint \frac{d\lambda}{2i\pi} \partial^k(w) \cdot w^* &= \oint \frac{d\lambda}{2i\pi} \partial^k \left(Me^{\xi(t;\lambda)} \right) \cdot (M^*)^{-1} e^{-\xi(t;\lambda)} \\ &= \oint \frac{d\lambda}{2i\pi} \left((\partial^k M) e^{\lambda x} \right) \cdot \left((M^*)^{-1} e^{-\lambda x} \right) \\ &= \operatorname{res}_\partial (\partial^k M M^{-1}) = \operatorname{res}_\partial \partial^k = 0, \end{aligned}$$

where we have used Lemma 3.3.4. The compact form of the bilinear identity (the one with $\forall \mathbf{t}, \mathbf{t}'$) is obtained by a Taylor expansion around $\mathbf{t} = \mathbf{t}'$, i.e by replacing $\mathbf{t} = \mathbf{t}' + \Pi^\vee$ and expanding the integrand. □

For the final step, we prove a statement that is a bit stronger than the implication that is converse to Lemma 3.3.5.

Proposition 3.3.6. *Let $w, \tilde{w} \in \mathcal{W}$ be two formal power series of the form*

$$w(\mathbf{t}; \lambda) = Me^{\xi(\mathbf{t}; \lambda)}, \quad \tilde{w}(\mathbf{t}; \lambda) = \tilde{M}e^{-\xi(\mathbf{t}; \lambda)},$$

where the operators $M, \tilde{M} \in \mathcal{D}$ have the form

$$M = 1 + \sum_{k < 0} m^k \partial^k, \quad \tilde{M} = 1 + \sum_{k < 0} \tilde{m}^k \partial^k.$$

If the functions w and \tilde{w} satisfy the equation

$$\oint \frac{d\lambda}{2i\pi} w(\mathbf{t}; \lambda) \cdot \tilde{w}(\mathbf{t}'; \lambda) = 0, \quad \forall \mathbf{t}, \mathbf{t}', \quad (3.27)$$

then $\tilde{M} = (M^*)^{-1}$ (so $\tilde{w} = w^*$) and w is a wave function of the linear system of KP (3.23) (equivalently, the operator $Q = M\partial M^{-1}$ satisfies the KP hierarchy (3.18)).

Clearly, the above proposition, along with Lemma 3.3.5, implies Theorem 3.3.2.

Proof. Let us first prove that if w and \tilde{w} satisfy Equation (3.27), then $\tilde{M} = (M^*)^{-1}$. By Lemma 3.3.4, we see that for any $k \geq 0$,

$$\text{res}_\partial \partial^k M N^* = \oint \frac{d\lambda}{2i\pi} \left(\partial^k M e^{\xi(t;\lambda)} \right) \cdot \left(\tilde{M} e^{-\xi(t;\lambda)} \right) = \oint \frac{d\lambda}{2i\pi} \partial^k(w) \cdot \tilde{w} = 0,$$

where we have used Equation (3.27) in the last step. On the other hand, $\tilde{M} = 1 + O(\partial^{-1})$ and this readily implies that $\tilde{M}^* = 1 + O(\partial^{-1})$ as well. Consequently,

$$M \tilde{M}^* = 1 + A, \quad A \in \mathcal{D}_-,$$

so that $\text{res}_\partial \partial^k A = 0$ for any $k \geq 0$. Therefore, $A = 0$, which means that $\tilde{M} = (M^*)^{-1}$.

We now prove that M satisfies Sato's equation (which is one equivalent form of the KP hierarchy, cf. Theorem 3.2.4), i.e. that

$$\frac{\partial M}{\partial t_i} = -(Q^i)_- M,$$

where $Q = M \partial M^{-1}$. First, we observe that the flows of w and e^ξ are related by

$$\left(\frac{\partial M}{\partial t_i} + (Q^i)_- M \right) e^{\xi(t;\lambda)} = \left(\frac{\partial}{\partial t_i} - (Q^i)_+ \right) w.$$

Indeed, using $\partial_{t_i} e^\xi = \partial^i e^\xi$ and $M \partial^i = Q^i M$, we get

$$\begin{aligned} \left(\frac{\partial M}{\partial t_i} + (Q^i)_- M \right) e^{\xi(t;\lambda)} &= \frac{\partial w}{\partial t_i} - M \frac{\partial e^{\xi(t;\lambda)}}{\partial t_i} + (Q^i)_- w \\ &= \frac{\partial w}{\partial t_i} - M \partial^i e^{\xi(t;\lambda)} + (Q^i)_- w \\ &= \frac{\partial w}{\partial t_i} - Q^i w + (Q^i)_- w = \left(\frac{\partial}{\partial t_i} - (Q^i)_+ \right) w. \end{aligned}$$

Since $(Q^i)_+$ is a differential operator in $\partial = \partial_{t_1}$ and using Equation (3.27), we have, for $k \geq 0$,

$$\begin{aligned} 0 &= \oint \frac{d\lambda}{2i\pi} \left(\partial^k (\partial_{t_i} - (Q^i)_+) \right) w \cdot \tilde{w} \\ &= \oint \frac{d\lambda}{2i\pi} \left(\partial^k (\partial_{t_i} M + (Q^i)_- M) e^{\xi(t;\lambda)} \right) \cdot \tilde{M} e^{-\xi(t;\lambda)}. \end{aligned}$$

By Lemma 3.3.4, this is equivalent to

$$0 = \text{res}_\partial \partial^k (\partial_{t_i} M + (Q^i)_- M) \circ M^{-1},$$

where we have used that $\tilde{M}^* = M^{-1}$. The above holding for any $k \geq 0$, we get that $\partial_{t_i} M + (Q^i)_- M = 0$, which is what was to be shown. \square

3.3.2 The tau functions of the KP hierarchy

Now that we have the bilinear identity, we can prove the essential property of the existence of a tau function for any solution of the KP hierarchy. This statement is formulated in the following theorem. We still follow the approach of [BBT03] (pp. 352–354).

Theorem 3.3.7 (Tau function (KP)). *Let $M = 1 + \sum_{k < 0} m^k \partial^k$ be an operator satisfying Sato's equation (3.21) and define functions $w, w^* \in \mathcal{W}$ by*

$$w(\mathbf{t}; \lambda) = M e^{\xi(\mathbf{t}; \lambda)}, \quad w^*(\mathbf{t}; \lambda) = (M^*)^{-1} e^{-\xi(\mathbf{t}; \lambda)}.$$

If w and w^ satisfy the bilinear identity (3.26), then there exists a (non unique, nonzero) function $\tau \in \mathbb{C}[[\mathbf{t}]]$, called the tau function, such that*

$$w(\mathbf{t}; \lambda) = \frac{\tau(\mathbf{t} - [\lambda^{-1}])}{\tau(\mathbf{t})} e^{\xi(\mathbf{t}; \lambda)}, \quad w^*(\mathbf{t}; \lambda) = \frac{\tau(\mathbf{t} + [\lambda^{-1}])}{\tau(\mathbf{t})} e^{-\xi(\mathbf{t}; \lambda)},$$

where $[\lambda^{-1}] := (\frac{1}{\lambda}, \frac{1}{2\lambda^2}, \frac{1}{3\lambda^3}, \dots)$ and $\xi(\mathbf{t}; \lambda) = \sum_{i \geq 1} t_i \lambda^i$.

Remark 3.3.8. Notice that since $\tau(\mathbf{t})$ does not depend on λ , it does not contribute to the residue. Therefore, the bilinear identity can be reformulated as

$$\oint \frac{d\lambda}{2i\pi} \tau(\mathbf{t} - [\lambda^{-1}]) \cdot \tau(\mathbf{t}' + [\lambda^{-1}]) \exp \left[\sum_{i \geq 1} (t_i - t'_i) \lambda^i \right] = 0, \quad \forall \mathbf{t}, \mathbf{t}'.$$

To prove Theorem 3.3.7, we will proceed with two lemmas. For these lemmas, it will be convenient to introduce a function $\widehat{m} \in \mathcal{W}$ defined by

$$w(\mathbf{t}; \lambda) = M e^{\xi(\mathbf{t}; \lambda)} =: \widehat{m}(\mathbf{t}; \lambda) \times e^{\xi(\mathbf{t}; \lambda)}$$

By construction, if $M = 1 + \sum_{k < 0} m^k \partial^k$, then

$$\widehat{m}(\mathbf{t}; \lambda) = 1 + \sum_{k < 0} \frac{m^k}{\lambda^k}.$$

Accordingly, we define a function \widehat{m}^* by

$$w^*(\mathbf{t}; \lambda) = (M^*)^{-1} e^{-\xi(\mathbf{t}; \lambda)} =: \widehat{m}^*(\mathbf{t}; \lambda) \times e^{-\xi(\mathbf{t}; \lambda)}$$

Lemma 3.3.9. *If w, w^* satisfy the bilinear identity, then the function \widehat{m} satisfies the functional equation*

$$\frac{\widehat{m}(\mathbf{t} - [\mu^{-1}]; \nu)}{\widehat{m}(\mathbf{t}; \mu)} = \frac{\widehat{m}(\mathbf{t} - [\nu^{-1}]; \mu)}{\widehat{m}(\mathbf{t}; \nu)} \quad (3.28)$$

Proof. We start by making two simple observations. First, for any function $f(\lambda) = 1 + \sum_{k < 0} f^k \lambda^k$, using the geometric series, we have

$$\oint \frac{d\lambda}{2i\pi} \frac{f(\lambda)}{1 - \lambda/\mu} = \mu (f(\mu) - 1). \quad (3.29)$$

Second, using the power series expansion of the logarithm, i.e. $\log(1+x) = \sum_{k \geq 1} (-1)^{k+1} x^k k^{-1}$, we see that

$$e^{-\xi(\mathbf{t}-[\mu^{-1}]; \lambda)} = \frac{e^{-\xi(\mathbf{t}; \lambda)}}{1 - \lambda/\mu}. \quad (3.30)$$

Now, applying the bilinear identity to \mathbf{t} and $\mathbf{t}' = \mathbf{t} - [\mu^{-1}]$ and using Equations (3.29) and (3.30), we get

$$\begin{aligned} 0 &= \oint \frac{d\lambda}{2i\pi} w(\mathbf{t}; \lambda) \cdot w^*(\mathbf{t} - [\mu^{-1}]; \lambda) \\ &= \oint \frac{d\lambda}{2i\pi} \widehat{m}(\mathbf{t}; \lambda) e^{\xi(\mathbf{t}; \lambda)} \cdot \widehat{m}^*(\mathbf{t} - [\mu^{-1}]; \lambda) e^{-\xi(\mathbf{t}-[\mu^{-1}]; \lambda)} \\ &= \oint \frac{d\lambda}{2i\pi} \frac{\widehat{m}(\mathbf{t}; \lambda) \cdot \widehat{m}^*(\mathbf{t} - [\mu^{-1}]; \lambda)}{1 - \lambda/\mu} = \mu (\widehat{m}(\mathbf{t}; \mu) \cdot \widehat{m}^*(\mathbf{t} - [\mu^{-1}]; \mu) - 1). \end{aligned}$$

In other words,

$$\widehat{m}^*(\mathbf{t} - [\mu^{-1}]; \mu) = \frac{1}{\widehat{m}(\mathbf{t}; \mu)}. \quad (3.31)$$

Similarly, we want to apply the bilinear identity to \mathbf{t} and $\mathbf{t}' = \mathbf{t} - [\mu^{-1}] - [v^{-1}]$. Let us denote

$$g(\mathbf{t}; \lambda) := \widehat{m}(\mathbf{t}; \lambda) \cdot \widehat{m}^*(\mathbf{t} - [\mu^{-1}] - [v^{-1}]; \lambda),$$

then we get

$$\begin{aligned} 0 &= \oint \frac{d\lambda}{2i\pi} w(\mathbf{t}; \lambda) \cdot w^*(\mathbf{t} - [\mu^{-1}] - [v^{-1}]; \lambda) \\ &= \oint \frac{d\lambda}{2i\pi} \frac{g(\mathbf{t}; \lambda)}{(1 - \lambda/\mu)(1 - \lambda/v)} \\ &= \oint \frac{d\lambda}{2i\pi} \frac{v}{v - \mu} \frac{g(\mathbf{t}; \lambda)}{1 - \lambda/\mu} - \frac{\mu}{v - \mu} \frac{g(\mathbf{t}; \lambda)}{1 - \lambda/v} \\ &= \frac{\mu v}{v - \mu} (g(\mathbf{t}; \mu) - g(\mathbf{t}; v)). \end{aligned}$$

Consequently, $g(\mathbf{t}; \mu) = g(\mathbf{t}; v)$, i.e.

$$\widehat{m}(\mathbf{t}; \mu) \cdot \widehat{m}(\mathbf{t} - [\mu^{-1}] - [v^{-1}]; \mu) = \widehat{m}(\mathbf{t}; v) \cdot \widehat{m}(\mathbf{t} - [\mu^{-1}] - [v^{-1}]; v).$$

Using Equation (3.31), we get

$$\frac{\widehat{m}(\mathbf{t} - [\mu^{-1}]; v)}{\widehat{m}(\mathbf{t}; \mu)} = \frac{\widehat{m}(\mathbf{t} - [v^{-1}]; \mu)}{\widehat{m}(\mathbf{t}; v)}$$

□

Lemma 3.3.10. *The function \widehat{m} satisfies the functional equation (3.28) if and only if there exists a function $\tau \in \mathbb{C}[[\mathbf{t}]]$ such that*

$$\widehat{m}(\mathbf{t}; \lambda) = \frac{\tau(\mathbf{t} - [\lambda^{-1}])}{\tau(\mathbf{t})}$$

Proof. First of all, it is easy to check that \widehat{m} of this form satisfies Equation (3.28). For the converse statement, let us take the logarithm of Equation (3.28):

$$h(\mathbf{t} - [\mu^{-1}]; \nu) - h(\mathbf{t}; \nu) = h(\mathbf{t} - [\nu^{-1}]; \mu) - h(\mathbf{t}; \mu), \quad (3.32)$$

where $h = \log \widehat{m}$. The function h is well defined since $\widehat{m} = 1 + \mathcal{O}(\lambda^{-1})$. We introduce the generating function of time derivatives:

$$\nabla_\lambda := \sum_{i \geq 1} \lambda^{-i-1} \partial_{t_i}.$$

Notice that for any function ϕ , we automatically get $(\partial_\lambda - \nabla_\lambda)\phi(\mathbf{t} - [\lambda^{-1}]) = 0$. Applying $(\partial_\nu - \nabla_\nu)$ to Equation (3.32) we get

$$(\partial_\nu - \nabla_\nu)h(\mathbf{t} - [\mu^{-1}]; \nu) - (\partial_\nu - \nabla_\nu)h(\mathbf{t}; \nu) = -\nabla_\nu h(\mathbf{t}; \mu).$$

Let us denote $(\partial_\nu - \nabla_\nu)h(\mathbf{t}; \nu) =: \sum_{i \geq 1} \gamma^i(\mathbf{t})\nu^{-i-1}$. Then expanding in powers of ν , the above equation reads

$$\gamma^i(\mathbf{t} - [\mu^{-1}]) - \gamma^i(\mathbf{t}) = \partial_{t_i} h(\mathbf{t}; \mu). \quad (3.33)$$

We denote $\Gamma^{ij} = \partial_{t_i} \gamma^j - \partial_{t_j} \gamma^i$. Then by symmetry of the second partial derivatives, we can eliminate the right-hand side in Equation (3.33) and get

$$\Gamma^{ij}(\mathbf{t} - [\mu^{-1}]) - \Gamma^{ij}(\mathbf{t}) = 0.$$

Now, expanding in powers of μ , we get $\sum_{i \geq 1} \mu^{-i-1} \partial_{t_i} \Gamma^{ij} = \Gamma^{ij}$, so that Γ^{ij} is actually independent of all the variables t_i , $i \geq 1$.

On the other hand, Γ^{ij} is by construction a polynomial in the $\partial_{t_k} m^\ell$'s. Using the KP hierarchy, we can express $\partial_{t_k} m^\ell$ as a polynomial in the $\partial^p m^\ell$'s so that Γ^{ij} is a polynomial in the independent variables $m_p^\ell = \partial^p m^\ell$. Because these variables are independent and because Γ^{ij} is constant, it reduces to its constant term.

Now notice that the function $\widehat{m} \equiv 1$ solves Equation (3.28). For this particular solution, $h = \log \widehat{m} = 0$, so that $\Gamma^{ij} = 0$ as well. Therefore, $\Gamma^{ij} = 0$ for any solution (since it is constant). By Poincaré lemma, it means that there exists a function $F \in \mathbb{C}[[\mathbf{t}]]$ such that $\gamma^i = \partial_{t_i} F$; we denote $\tau = \exp F$. Inserting this in Equation (3.33) we get that

$$\log \tau(\mathbf{t} - [\mu^{-1}]) - \log \tau(\mathbf{t}) = \log \widehat{m}(\mathbf{t}; \mu) + C,$$

where $C \in \mathbb{C}$ is constant, i.e.

$$\widehat{m}(\mathbf{t}; \mu) = \tilde{C} \cdot \frac{\tau(\mathbf{t} - [\mu^{-1}])}{\tau(\mathbf{t})},$$

where $\tilde{C} = e^{-C}$. The constant \tilde{C} can be absorbed into τ (which is defined, in particular, up to multiplication by a constant) to get the result. The equation on w^* follows directly from that on w . \square

The two lemmas above prove Theorem 3.3.7.

3.4 The Korteweg–de Vries hierarchy

3.4.1 Lax pairs and other forms of the KdV hierarchy

As we already said, the KdV hierarchy can be seen a reduction of the KP hierarchy by imposing $L = Q^2 \in \mathcal{D}_+$, i.e. $Q_-^2 = 0$. Let us first state the KdV hierarchy.

Theorem 3.4.1 (KdV hierarchy [Lax68]). *Let $L = \partial^2 + u \in \mathcal{D}_+$ be a differential operator. Then the following flows commute pairwise, producing the so-called KdV hierarchy:*

$$\frac{\partial L}{\partial t_i} = \left[(L^{i/2})_+, L \right], \quad i \in \mathbb{Z}_{\geq 1}^{\text{odd}}. \quad (3.34)$$

The first flows of the KdV hierarchy read $u_{t_1} = u_1$ (where $u_{t_i} = \partial_{t_i} u$ and $u_1 = u_x, u_2 = u_{xx}$, etc.), then

$$\begin{aligned} u_{t_3} &= \frac{1}{4}u_3 + \frac{3}{2}uu_1, \\ u_{t_5} &= \frac{1}{16}u_5 + \frac{5}{8}uu_3 + \frac{5}{4}u_1u_2 + \frac{15}{8}u^2u_1, \\ u_{t_7} &= \frac{1}{64}u_7 + \frac{7}{32}uu_5 + \frac{21}{32}u_1u_4 + \frac{35}{32}u_2u_3 + \frac{35}{32}(u)^2u_3 + \frac{35}{8}uu_1u_2 + \frac{35}{32}(u_1)^3 + \frac{35}{16}(u)^3u_1. \end{aligned}$$

The flow with respect to t_3 reads the so-called KdV equation, first theorized by J. Boussinesq in 1877 [Bou77] and then proposed in its current acceptance by D. J. Korteweg and G. de Vries in 1895 [KdV95].

First of all, Theorem 3.1.8 ensures the existence of the square root of L . Second, as in the KP hierarchy and because $[L^{i/2}, L^{j/2}] = 0$ for any $i, j \geq 1$, we can rewrite the equations as

$$\frac{\partial L}{\partial t_i} = \left[L, (L^{i/2})_- \right], \quad i \in \mathbb{Z}_{\geq 1}^{\text{odd}}.$$

Then one can easily show that these equations are well defined.

Just like for KP, there is a straightforward proof of Theorem 3.4.1 (see proof of Theorem 3.2.1). However, we will take a different approach to the problem and show that we can deduce the commutativity of the KdV flows from that of the KP flows. To do so, we start with an important lemma.

Lemma 3.4.2. *Let P be a pseudo-differential operator of the form $P = \partial + \sum_{k < 0} p^k \partial^k$. Then for any integer $i \in \mathbb{Z}_{\geq 1}^{\text{odd}}$, the following equivalence holds:*

$$\frac{\partial P^2}{\partial t_i} + [(P^i)_-, P^2] = 0 \quad \text{if and only if} \quad \frac{\partial P}{\partial t_i} + [(P^i)_-, P] = 0.$$

Proof. The proof is entirely similar to that of Corollary 3.1.16. Let us denote

$$A_i = [(P^i)_-, P] = \sum_{k < 0} a^{i,k} \partial^k.$$

Suppose that $\partial_{t_i} P + A_i \neq 0$ and denote m_i its order, i.e. $\partial_{t_i} p^{m_i} + a^{i,m_i} \neq 0$. Then by Leibniz's rule, we can write

$$\begin{aligned} \frac{\partial P^2}{\partial t_i} + [(P^i)_-, P^2] &= P \frac{\partial P}{\partial t_i} + \frac{\partial P}{\partial t_i} P + P A_i + A_i P \\ &= 2 (\partial_{t_i} p^{m_i} + a^{i,m_i}) \partial^{m_i+1} + O(\partial^{m_i}). \end{aligned}$$

Therefore, $\partial_{t_i} P^2 + [(P^i)_-, P^2] \neq 0$ as well. The converse is obvious. \square

Remark 3.4.3. With exactly the same arguments, we could have shown the following statement (that we use for the Gelfand–Dickey hierarchies): For any $n \geq 1$,

$$\frac{\partial P^n}{\partial t_i} + [(P^i)_-, P^n] = 0 \quad \text{if and only if} \quad \frac{\partial P}{\partial t_i} + [(P^i)_-, P] = 0.$$

Proof. (of Theorem 3.4.1). Thanks to Lemma 3.4.2, the Lax equations of the KdV hierarchy can be interpreted as the Lax equations of the KP hierarchy on the operator $P = L^{1/2}$. Then Theorem 3.2.1 ensures that the flows commute. \square

In the proof above, we implicitly proved that any operator $P = \partial + \sum_{k < 0} p^k \partial^k$ such that $(P^2)_- = 0$ and which satisfies the KP hierarchy, provides a solution to the KdV hierarchy. In the same fashion, we can define the linear system of KdV as follows.

Theorem 3.4.4 (linear system of KdV). *The KdV hierarchy (3.34) on the operator $L = \partial^2 + u$ holds if and only if there exists a function $w \in \mathcal{W}$ such that the following hierarchy of equations is satisfied,*

$$\begin{cases} Lw = \lambda^2 w; \\ \partial_{t_i} w = (L^{i/2})_+ w; \end{cases} \quad i \in \mathbb{Z}_{\geq 1}^{\text{odd}}, \lambda \in \mathbb{C}. \quad (3.35)$$

As in the KP case, one can explicitly construct a wave function w that satisfies the linear system of KdV as follows. Let $M = \partial + \sum_{k < 0} m^k \partial^k$ be an operator such that $P = M \partial M^{-1}$ where $P = L^{1/2}$. Then

$$w(\mathbf{t}; \lambda) = M e^{\xi(\mathbf{t}; \lambda)}$$

provides a wave function that solves the system (3.35). Here $\xi(\mathbf{t}; \lambda) = \sum_{i \geq 1} t_i \lambda^i$ as well, but we could have gotten rid of the even variables and still get a valid wave function.

As w is also a wave function for the linear system of KP, we still have the existence of a tau function $\tau \in \mathbb{C}[[\mathbf{t}]]$ such that

$$w(\mathbf{t}; \lambda) = \frac{\tau(\mathbf{t} - [\lambda^{-1}])}{\tau(\mathbf{t})} e^{\xi(\mathbf{t}; \lambda)}.$$

That defines the tau functions of the KdV hierarchy. In the next section, we show that the constraint $(P^2)_- = 0$ of the KdV hierarchy implies much simpler formulae for the tau function.

3.4.2 The tau functions of the KdV hierarchy

Proposition 3.4.5 (Tau function (KdV)). *Let $L = \partial^2 + u$ be a solution of the KdV hierarchy. Let τ be the tau function associated to the square root operator $P = L^{1/2}$ (seen as a solution to the KP hierarchy). Then τ is determined by the constraint*

$$u = -2 \frac{\partial^2}{\partial x^2} \log \tau = 2 \frac{(\tau_{(1)})^2 - \tau \tau_{(2)}}{(\tau)^2}. \quad (3.36)$$

Proof. Let us recall how the tau function of KP is defined starting from the operator P . First of all, P goes like this:

$$P = \partial - \frac{1}{2} u \partial^{-1} + \mathcal{O}(\partial^{-2})$$

(Equation (3.10)). Since $P = \partial + \mathcal{O}(\partial^{-1})$, there exists an operator $M = 1 + \mathcal{O}(\partial^{-1})$ (the dressing operator) such that $P = M \partial M^{-1}$ (Proposition 3.2.3). It is easy to show that

$$M = 1 - m^{-1} \partial^{-1} - \left(m^{-2} + (m^{-1})^2 \right) \partial^{-2} + \mathcal{O}(\partial^{-3}), \quad \text{such that } \frac{\partial m^{-1}}{\partial x} = -\frac{1}{2} u,$$

From the operator M , we define the function

$$w(\mathbf{t}; \lambda) = M e^{\xi(\mathbf{t}; \lambda)} = \left(1 + m^{-1} \lambda^{-1} + \mathcal{O}(\lambda^{-2}) \right) e^{\xi(\mathbf{t}; \lambda)} \in \mathcal{W},$$

where $\xi(\mathbf{t}; \lambda) = \sum_{i \geq 1} t_i \lambda^i$, which produces a solution to the linear system of KP (3.23), the latter being equivalent to the KP hierarchy (Theorem 3.2.5). Then the fact that the wave function satisfies Equation (3.23) implies the existence of the tau function $\tau \in \mathbb{C}[[\mathbf{t}]]$ such that

$$w(\mathbf{t}; \lambda) \times e^{-\xi(\mathbf{t}; \lambda)} = \frac{\tau(\mathbf{t} - [\lambda^{-1}])}{\tau(\mathbf{t})} = 1 + \frac{1}{\tau} \frac{\partial \tau}{\partial x} \lambda^{-1} + \mathcal{O}(\lambda^{-2}).$$

Comparing the three equations above, we get that

$$u = -2 \frac{\partial m^{-1}}{\partial x} = -2 \frac{\partial}{\partial x} \left(\frac{1}{\tau} \frac{\partial \tau}{\partial x} \right) = -2 \frac{\partial^2}{\partial x^2} \log \tau.$$

□

As we saw, given an operator Q satisfying the KP hierarchy, one only has to impose the constraint $(Q^2)_- = 0$ to produce a solution to the KdV hierarchy. It remains to know how that works at the level of tau functions. The next proposition is essential.

Proposition 3.4.6. *Let τ be a tau function of the KP hierarchy. If $\partial_{t_2} \tau = 0$, then the function $u = -2 \partial_x^2 \log \tau$ is a solution of the KdV hierarchy.*

Remark 3.4.7. The converse is wrong: A tau function τ of KP can be a tau function of KdV and depend on t_2 (or other even variables). Indeed, since the tau function is defined—in the KdV case—by $u = -2 \partial_x^2 \log \tau$, then any transformation of the form

$$\tau(\mathbf{t}) \mapsto \tau(\mathbf{t}) \times \exp [A(\mathbf{t})x + B(\mathbf{t})], \quad \frac{\partial A}{\partial x} = \frac{\partial B}{\partial x} = 0,$$

defines the same function u , and therefore produces the same operator $L = \partial^2 + u$ that satisfies the KdV hierarchy.

Notice that, in practice, we often compute tau functions of KdV such that they do not depend on even variables. In particular, they do not depend t_2 . Then it suffices to prove that they satisfy the KP hierarchy, since Proposition 3.4.6 ensures that they will automatically be tau function of KdV.

As we will see in Section 3.5, it can actually be easier to prove that a tau function satisfies the KP hierarchy, rather than the KdV hierarchy.

Proof of Proposition 3.4.6. We use the isospectral formulation of KP and consider the wave function w associated to τ via the relation

$$w(\mathbf{t}; \lambda) = \frac{\tau(\mathbf{t} - [\lambda^{-1}])}{\tau(\mathbf{t})} e^{\xi(\mathbf{t}; \lambda)}.$$

Since $\partial_{t_2} \tau = 0$, we have $\partial_{t_2} w = \lambda^2 w$. By Theorem 3.2.5, w satisfies the system of equations

$$\begin{cases} Pw = \lambda w; \\ \partial_{t_i} w = (P^i)_+ w; \end{cases} \quad i \in \mathbb{Z}_{\geq 1}.$$

In particular, $(P^2)_- w = 0$; indeed,

$$0 = \partial_{t_2} w - (P^2)_+ w = \lambda^2 w - (P^2)_+ w = (P^2)_- w.$$

Since P is independent of λ , yet $(P^2)_- w = 0$ for arbitrary values of $\lambda \in \mathbb{C}$, it follows that $(P^2)_- = 0$. Equivalently $(P^2)_+ = P^2 \in \mathcal{D}_+$ and the operator $L = P^2 =: \partial^2 + u$ satisfies the linear system system

$$\begin{cases} Lw = \lambda^2 w; \\ \partial_{t_i} w = (L^{i/2})_+ w; \end{cases} \quad i \in \mathbb{Z}_{\geq 1}^{\text{odd}}, \quad (3.37)$$

which is equivalent to the KdV hierarchy. □

3.5 The polynomial tau functions of KdV and the Adler–Moser polynomials

As described in the former section, to any solution u of the KdV hierarchy, we associate a tau function defined by

$$u = -2 \frac{\partial^2}{\partial x^2} \log \tau = 2 \frac{(\tau_{(1)})^2 - \tau \tau_{(2)}}{(\tau)^2}$$

If τ is polynomial then u is clearly rational. On the other hand, it was proven by Airault, McKean and Moser [AMcKM77] that there are denumerably many rational solutions of KdV, each one being the orbit of the function $u_n = n(n + 1)/x^2$ under the flows of the hierarchy.

In [AM78], the authors constructed the Adler–Moser polynomials $\theta_n(x = r_1, r_3, \dots, r_{2n-1})$ for $n \geq 0$, defined by the recursion

$$\frac{\partial \theta_{n+1}}{\partial x} \theta_{n-1} - \theta_{n+1} \frac{\partial \theta_{n-1}}{\partial x} = (2n-1)(\theta_n)^2.$$

An important result of [AM78] (cf. Theorem 3.5.1) is that there exists a unique change of variables that transforms the Adler–Moser polynomials into polynomial tau functions of KdV and that we recover all rational solutions of KdV. But we did not know what this change of variables was.

In this section, we show that the following change of variables transforms the Adler–Moser polynomials into the polynomial tau functions of KdV: $r_1 = t_1 = x$, and

$$\sum_{i \geq 2} \frac{r_{2i-1}}{\alpha_{2i-1}} z^{2i-1} = \tanh \left(\sum_{i \geq 2} t_{2i-1} z^{2i-1} \right).$$

where $\alpha_{2i-1} = (-1)^{i-1} 3^2 5^2 \dots (2i-3)^2 (2i-1)$. To do so, we apply this change of variables to the Adler–Moser polynomials to get some polynomials $\tau_n(t_1, t_3, t_5, \dots)$. Then by seeing them as functions $\tau_n(t_1, t_2, t_3, \dots)$ of even and odd times we show that they are tau functions of the Kadomstev–Petviashvili hierarchy (KP). Then by Proposition 3.4.6, since the τ_n 's actually depend only on odd times, they are indeed tau functions of KdV.

It is well known how to compute the polynomial tau functions of KdV without using the Adler–Moser polynomials. For instance, in [Hir04] R. Hirota constructs a family of tau functions of KP in terms of Wronskians of the elementary Schur polynomials, which can be reduced to recover the polynomial tau functions of KdV. But the Adler–Moser polynomials reveal a recursive structure in the space of rational solutions of KdV. It would be interesting to investigate how to generalize this to the Drinfeld–Sokolov hierarchies.

3.5.1 The Adler–Moser Polynomials

Let $\mathbf{r} = (x = r_1, r_3, r_5, \dots)$ be a set of oddly-indexed variables.² We save the usual variables $\mathbf{t} = (x = t_1, t_3, t_5, \dots)$ for the tau functions of KdV. The Adler–Moser polynomials form a sequence

$$\theta_n(r_1, r_3, \dots, r_{2n-1}), \quad n \geq 0,$$

defined by the following recursion [AM78]: $\theta_0 = 1$, $\theta_1 = x$ and for $n \geq 1$,

$$\frac{\partial \theta_{n+1}}{\partial x} \theta_{n-1} - \theta_{n+1} \frac{\partial \theta_{n-1}}{\partial x} = (2n+1)(\theta_n)^2. \quad (3.38)$$

This recursion leaves an integration constant that is chosen to be r_{2n-1} when computing θ_n . We can check that θ_n has degree $d_n = \frac{1}{2}n(n+1)$ in x and r_{2n-1} is actually the coefficient of x^{d_n-2}

²In their original paper [AM78] the authors use the variables τ_i instead. But since then, the letter τ has been used rather for the tau functions. Moreover, we use odd indices so that we can later complete the set into the variables of KP.

in this polynomial. The first few polynomials read

$$\begin{aligned}
\theta_0 &= 1; \\
\theta_1 &= x; \\
\theta_2 &= x^3 + r_3; \\
\theta_3 &= x^6 + 5r_3x^3 + r_5x - 5r_3^2; \\
\theta_4 &= x^{10} + 15r_3x^7 + 7r_5x^5 - 35r_3r_5x^2 + 175r_3^3x - \frac{7}{3}r_5^2 + r_7x^3 + r_3r_7; \\
\theta_5 &= x^{15} + 35r_3x^{12} + 28r_5x^{10} + 175r_3^2x^9 + 9r_7x^8 - 105r_3r_5x^7 + r_9x^6 - 147r_5^2x^5 - 63r_3r_7x^5 \\
&\quad + 3675r_3^2r_5x^4 - 18375r_3^4x^3 - 21r_5r_7x^3 + 5r_3r_9x^3 + 735r_3r_5^2x^2 - 315r_3^2r_7x^2 - \frac{9}{5}r_7^2x + r_5r_9x \\
&\quad - 49r_5^3 + 42r_3r_5r_7 - 5r_3^2r_9.
\end{aligned}$$

In that same article ([AM78], pp. 17–18), the authors state the following theorem.

Theorem 3.5.1 (Adler–Moser, 1978). *There exists a unique change of variables $\mathbf{r} \mapsto \mathbf{t}$ that transforms the Adler–Moser polynomials $\theta_n(\mathbf{r})$ into the polynomial tau functions $\tau_n(\mathbf{t})$ of KdV. That is, the rational functions $u_n = -2\partial^2 \log \tau_n$ define operators $L_n = \partial^2 + u_n$ that satisfy the Lax system of KdV:*

$$\frac{\partial L_n}{\partial t_{2i-1}} = \left[\left(L_n^{\frac{2i-1}{2}} \right)_+, L_n \right].$$

We aim to prove that the desired change of variables is given by: $r_1 = t_1 = x$, and

$$\sum_{i \geq 2} \frac{r_{2i-1}}{\alpha_{2i-1}} z^{2i-1} = \tanh \left(\sum_{i \geq 2} t_{2i-1} z^{2i-1} \right).$$

where $\alpha_{2i-1} = (-1)^{i-1} 3^{2i-2} \dots (2i-3)^2 (2i-1)$. The latter coefficients α_{2i-1} where already given in Adler and Moser’s article. Notice that this change of variables amounts to simply changing the choice of the integration constant in the differential recursion (3.38). To prove the statement, let us first recall and prove some results stated in [AM78]. These lemmas relate the Adler–Moser polynomials to a Wronskian representation through a multiplicative factor and a simple rescaling of the variables.

Let $\mathbf{s} = (x = s_1, s_3, s_5, \dots)$ be another set of variables and let $\psi^j(\mathbf{s})$, $j \geq 0$ be functions defined by:

$$\sum_{j \geq 1} \psi^j z^{2j-1} = \sinh(xz) + \cosh(xz) \sum_{i \geq 2} s_{2i-1} z^{2i-1}, \quad (3.39)$$

and $\psi^0 = 0$. It readily implies the recursion $\psi_{(2)}^j = \psi^{j-1}$, where as before $\psi_{(i)}^j = \partial_x^i \psi^j$. The first few terms read

$$\begin{aligned}
\psi^0 &= 0; \\
\psi^1 &= 1; \\
\psi^2 &= \frac{1}{6}x^3 + s_3^3; \\
\psi^3 &= \frac{1}{120}x^5 + \frac{1}{2}x^2s_3^3 + s_5^5; \\
\psi^4 &= \frac{1}{5040}x^7 + \frac{1}{24}x^4s_3^3 + \frac{1}{2}x^2s_5^5 + s_7^7; \\
\psi^5 &= \frac{1}{362880}x^9 + \frac{1}{720}x^6s_3^3 + \frac{1}{24}x^4s_5^5 + \frac{1}{2}x^2s_7^7 + s_9^9.
\end{aligned}$$

We now define the Wronskians of these functions:

$$W_n := \text{Wr}(\psi^1, \dots, \psi^n) = \det \left(\psi_{(i-1)}^j \right)_{i,j=1, \dots, n}.$$

Lemma 3.5.2. *The Wronskians W_n of the functions ψ^j , satisfy the recursion*

$$\frac{\partial W_{n+1}}{\partial x} W_{n-1} - W_{n+1} \frac{\partial W_{n-1}}{\partial x} = (W_n)^2. \quad (3.40)$$

Proof. For any smooth function $\chi(x)$ with respect to x , denote $W_n(\chi) := \text{Wr}(\psi^1, \dots, \psi^n, \chi)$. Then one has the identity:

$$\frac{\partial W_n(\chi)}{\partial x} W_{n+1} - W_n(\chi) \frac{\partial W_{n+1}}{\partial x} = W_{n+1}(\chi) W_n. \quad (3.41)$$

This can be proven by noticing that the left-hand side is a linear differential operator of order $n + 1$ which vanishes for $\chi = \psi^1, \dots, \psi^n$, as well as for ψ^{n+1} . Yet the functions ψ^j being linearly independent, the left-hand side must be proportional to $W_{n+1}(\chi)$. Comparing the highest coefficient, one obtains Equation (3.41). Now thanks to the relation $\psi_{(2)}^j = \psi^{j-1}$ and $\psi^1 = x$, we can compute that for $\chi = 1$,

$$W_n(1) = (-1)^n W_{n-1}. \quad (3.42)$$

Indeed, first we have $W_n(1) = \text{Wr}(\psi^1, \dots, \psi^n, 1) = (-1)^n \text{Wr}(1, \psi^1, \dots, \psi^n)$. Then writing down the determinants one obtains that

$$\text{Wr}(1, \psi^1, \dots, \psi^n) = \begin{vmatrix} 1 & 0 & 0 & \cdots & 0 \\ x & 1 & 0 & \cdots & 0 \\ \psi^2 & \psi_{(1)}^2 & \psi_{(2)}^2 & \cdots & \psi_{(n)}^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \psi^n & \psi_{(1)}^n & \psi_{(2)}^n & \cdots & \psi_{(n)}^n \end{vmatrix} = \begin{vmatrix} \psi_{(2)}^2 & \cdots & \psi_{(n)}^2 \\ \vdots & \ddots & \vdots \\ \psi_{(2)}^n & \cdots & \psi_{(n)}^n \end{vmatrix} = \begin{vmatrix} \psi^1 & \cdots & \psi_{(n-2)}^1 \\ \vdots & \ddots & \vdots \\ \psi^{n-1} & \cdots & \psi_{(n-2)}^{n-1} \end{vmatrix} = W_n,$$

hence Equation (3.42) holds true. Then setting $\chi = 1$ in Equation (3.41), one finds Equation (3.40). \square

Now since $W_0 = \theta_0 = 1$ and $W_1 = \theta_1 = x$, the two sequences of polynomials differ only by a multiplicative factor that can be computed:

$$\theta_n(\mathbf{r}) = \mu_n W_n(\mathbf{s}), \quad \mu_n = \prod_{j=1}^k (2k - 2j + 1)^j. \quad (3.43)$$

Lemma 3.5.3. *The parameters of the Adler–Moser polynomials $\theta_n(\mathbf{r})$ and those of the Wronskians $W_n(\mathbf{s})$ are related via a rescaling:*

$$s_{2i-1} = \frac{r_{2i-1}}{\alpha_{2i-1}}, \quad \alpha_{2i-1} = (-1)^{i-1} 3^2 5^2 \cdots (2i-3)^2 (2i-1) \quad (3.44)$$

Proof. It all has to do with the choice of the normalization in the recursions (3.38) and (3.40). For θ_n , the choice is such that r_{2n-1} is the coefficient of x^{dn-2} . Moreover if θ_n is a solution of (3.38), so is $\theta_n + c\theta_{n-2}$ so that the normalization can be expressed as

$$\theta_n = \dot{\theta}_n + r_{2n-1}\theta_{n-2},$$

where $\dot{\theta}_n := \theta_n |_{r_{2n-1} = 0}$. Similarly, define $\dot{W}_n := W_n |_{s_{2n-1} = 0}$. Since $\psi^n = \dot{\psi}^n + s_{2n-1}$, then by Equation (3.42),

$$W_n = \dot{W}_n + s_{2n-1} \text{Wr}(\psi^1, \dots, \psi^n, 1) = \dot{W}_n + (-1)^{k-1} s_{2n-1} W_{n-2}.$$

Comparing the last two equations and using Equation (3.43) one obtains the result. \square

3.5.2 A change of variables to the polynomial tau functions KdV

The aim of this section is to prove the following theorem.

Theorem 3.5.4 ([Duc17]). *The following change of variables transforms the Adler–Moser polynomials into the polynomial tau functions of KdV: $r_1 = t_1 = x$, and*

$$\sum_{i \geq 2} \frac{r_{2i-1}}{\alpha_{2i-1}} z^{2i-1} = \tanh \left(\sum_{i \geq 2} t_{2i-1} z^{2i-1} \right). \quad (3.45)$$

where $\alpha_{2i-1} = (-1)^{i-1} 3^2 5^2 \dots (2i-3)^2 (2i-1)$.

As a matter of fact, this change of variables does not affect the first few variables except for a rescaling. Here are the first variables r_{2i-1} in terms of the t_{2i-1} 's: $r_1 = t_1 = x$, and then

$$\begin{aligned} r_3 &= -3 \cdot t_3, & r_7 &= -1575 \cdot t_7, & r_{11} &= -9823275 \cdot (t_{11} - t_3^2 t_5), \\ r_5 &= 45 \cdot t_5, & r_9 &= 99255 \cdot (t_9 - \frac{1}{3} t_3^3), & r_{13} &= 1404728325 \cdot (t_{13} - t_3^2 t_7 - t_3 t_5^2). \end{aligned}$$

We voluntarily factorized by the constants α_{2i-1} ; in turns, the variable r_{2i-1} in terms of the s_{2i-1} 's is the expression within the parentheses, i.e. $r_1 = s_1 = x$ and then

$$\begin{aligned} r_3 &= s_3, & r_7 &= s_7, & r_{11} &= s_{11} - s_3^2 s_5, \\ r_5 &= s_5, & r_9 &= s_9 - \frac{1}{3} s_3^3, & r_{13} &= s_{13} - s_3^2 s_7 - s_3 s_5^2. \end{aligned}$$

Notice that the change of variables is homogeneous for the grading $\deg r_i = \deg t_i = \deg s_i = i$, as proven by Adler and Moser [AM78]. Under this change of variables, we obtain the following polynomial tau functions of KdV:

$$\begin{aligned} \tau_0 &= 1, \\ \tau_1 &= x, \\ \tau_2 &= x^3 - 3t_3, \\ \tau_3 &= x^6 - 15t_3x^3 - 45t_3^2 + 45t_5x, \\ \tau_4 &= x^{10} - 45t_3x^7 + 315t_5x^5 + 4725t_3t_5x^2 - 4725t_3^3x - 4725t_5^2 - 1475t_7x^3 + 4725t_3t_7. \end{aligned}$$

Then the polynomials τ_n/μ_n (cf. Equation (3.43)) correspond to the tau functions computed via the Wronskians of the elementary Schur polynomials following Hirota [Hir04].

To prove Theorem 3.5.4, we introduce another sequence of functions defined by the generating series.

$$\sum_{j \geq 1} \phi^j z^{2j-1} = \sinh(xz) + \cosh(xz) \tanh \left(\sum_{i \geq 2} t_{2i-1} z^{2i-1} \right). \quad (3.46)$$

It amounts to applying the change of variables of Equation (3.45) to the functions ψ^j of Equation (3.39). The first few terms read

$$\begin{aligned} \phi^0 &= 0; \\ \phi^1 &= 1; \\ \phi^2 &= \frac{1}{6}x^3 + s_3^3; \\ \phi^3 &= \frac{1}{120}x^5 + \frac{1}{2}x^2s_3^3 + s_5^5; \\ \phi^4 &= \frac{1}{5040}x^7 + \frac{1}{24}x^4s_3^3 + \frac{1}{2}x^2s_5^5 + s_7^7; \\ \phi^5 &= \frac{1}{362880}x^9 + \frac{1}{720}t_3^3x^6 + \frac{1}{24}t_5^5x^4 + \frac{1}{2}t_7^7x^2 - \frac{1}{19683}t_3^{27} + \frac{1}{729}t_9t_3^{24} - \frac{4}{243}t_9^2t_3^{21} + \frac{28}{243}t_9^3t_3^{18} - \frac{14}{27}t_9^4t_3^{15} \\ &\quad + \frac{14}{9}t_9^5t_3^{12} - \frac{28}{9}t_9^6t_3^9 + 4t_9^7t_3^6 - 3t_9^8t_3^3 + t_9^9. \end{aligned}$$

With these functions, we define another sequence of Wronskians:

$$\tau_n := \text{Wr}(\phi^1, \dots, \phi^n), \quad (3.47)$$

In what follows, we prove that these Wronskians are tau functions of the KP hierarchy. Yet because they depend only on odd times, they are tau functions of KdV (cf. Proposition 3.4.6), which proves Theorem 3.5.4. We use the same approach as in [IZ92] (also as in [Caf08]). First we need the following lemma.

Lemma 3.5.5. *The functions ϕ^j satisfy the following relation for any integers $i, j \geq 1$:*

$$\phi_{(2i-1)}^j - \frac{\partial \phi^j}{\partial t_{2i-1}} = \sum_{k=1}^{j-i-1} \phi^k a^{j-i-k+1}, \quad (3.48)$$

where $\phi_{(2i-1)}^j = \partial_x^{2i-1} \phi^j$. Here the a^j 's are functions that do not depend on x and are defined by

$$\sum_{j \geq 2} a^j z^{2j-1} = \tanh \left(\sum_{i \geq 2} t_{2i-1} z^{2i-1} \right).$$

Proof. It is a direct calculation: differentiating Equation (3.46) and using a Cauchy product, one obtains

$$\sum_{j \geq 1} \left(\phi_{(2i-1)}^j - \frac{\partial \phi^j}{\partial t_{2i-1}} \right) z^{2j-1} = \sum_{j \geq 1} z^{2j-1} \sum_{k=0}^{j-i} \phi^k a^{j-i-k+1}.$$

Then, noticing that $\phi^0 = 0$ and $a^0 = a^1 = 0$, one gets the correct boundaries in the last sum. \square

Remark 3.5.6. This relation is to be compared with the one satisfied by the elementary Schur polynomials defined by

$$\exp\left(\sum_{i \geq 1} t_i z^i\right) = \sum_{j \geq 1} p^j z^j. \tag{3.49}$$

These polynomials satisfy the relation $p_{(i)}^j = \partial_{t_i} p^j$. Now define \tilde{p}^j to be p^{2j-1} where all even times are set to 0, i.e. $\tilde{p}^j = p^{2j-1} |_{\{t_{2i} = 0, \forall i \geq 1\}}$. They satisfy the relation

$$\tilde{p}_{(2i-1)}^j = \frac{\partial \tilde{p}^j}{\partial t_{2i-1}}.$$

The last relation allows to prove that their Wronskians $\text{Wr}(\tilde{p}^1, \dots, \tilde{p}^n)$ are the tau functions of KdV the same way we prove that the Wronskians of the ϕ^j 's are (see for instance [I792]). In particular the Wronskians of the \tilde{p}^j 's and those of the ϕ^j 's coincide.

Now let us introduce all the variables ($x = t_1, t_2, t_3, \dots$) of the KP hierarchy (that is, odd and even). To prove that the τ_n 's are tau functions of KP we use Sato's equation, which is equivalent to the KP hierarchy (see Theorem 3.2.4). We state Sato's equation in an equivalent form in the following proposition.

Proposition 3.5.7. *Let Δ_n be the following differential operator*

$$\Delta_n(\chi) := \frac{\text{Wr}(\chi, \phi^1, \dots, \phi^n)}{\text{Wr}(\phi^1, \dots, \phi^n)} = \frac{1}{\tau_n} \text{Wr}(\chi, \phi^1, \dots, \phi^n),$$

for any differentiable function χ with respect to x . The following equation holds for any $i \geq 1$:

$$\frac{\partial \Delta_n}{\partial t_i} = (\Delta_n \partial^i \Delta_n^{-1})_+ \Delta_n - \Delta_n \partial^i. \tag{3.50}$$

Then from Equation (3.50) we readily get Sato's equation as follows. Set $M_n = \Delta_n \partial^{-n}$, which clearly is a monic operator of order 1 (i.e. $M_n = \partial + \mathcal{O}(1)$), then

$$\frac{\partial M_n}{\partial t_i} = -(M_n \partial^i M_n^{-1})_- M_n.$$

Proof. It is sufficient to prove the equality when acting on ϕ^1, \dots, ϕ^n which are n linearly independent functions. Yet these functions are solutions of the equation $\Delta_n(\phi^j) = 0$, so it amounts to proving that

$$\frac{\partial \Delta_n}{\partial t_i}(\phi^j) + \Delta_n(\phi_i^j) = 0.$$

If $i = 2\ell$ is even, then we only have to prove that $\Delta_n(\phi_{2\ell}^j) = 0$. Yet $\phi_2^j = \phi^{j-1}$ so that $\phi_{2\ell}^j = \phi^{j-\ell}$, or 0 if $j \leq \ell$. Eventually, it amounts to $\Delta_n(\phi^{j-\ell}) = 0$ which holds true. If $i = 2\ell - 1$ is odd, by Lemma 3.5.5, it amounts to proving that

$$\frac{\partial \Delta_n}{\partial t_{2\ell-1}}(\phi^j) + \Delta_n \frac{\partial}{\partial t_{2\ell-1}}(\phi^j) + \sum_{k=1}^{j-\ell-1} \Delta_n(\phi^k a^{j-\ell-k+1}) = 0.$$

Yet the functions a^j do not depend on x , so the last sum vanishes. And for the other terms it amounts to

$$\frac{\partial \Delta_n}{\partial t_{2\ell-1}}(\phi^j) + \Delta_n \frac{\partial}{\partial t_{2\ell-1}}(\phi^j) = \frac{\partial}{\partial t_{2\ell-1}} \Delta_n(\phi^j) = 0.$$

□

Finally, the following proposition states that the dressing operator M_n is indeed related to the Wronskians τ_n via the wave function, i.e. that τ_n is a tau function of KP.

Proposition 3.5.8. *The dressing operator M_n and the Wronskian τ_n satisfy the wave function relation*

$$M_n e^{\xi(\mathbf{t}; \lambda)} = \frac{\tau_n(\mathbf{t} - [\lambda^{-1}])}{\tau_n(\mathbf{t})} e^{\xi(\mathbf{t}; \lambda)}, \quad (3.51)$$

where $\xi(\mathbf{t}; \lambda) = \sum_{i \geq 1} t_i \lambda^i$ and $[\lambda^{-1}] := (\frac{1}{\lambda}, \frac{1}{3\lambda^3}, \frac{1}{5\lambda^5}, \dots)$.

To prove that, we need the following lemma.

Lemma 3.5.9. *The shift of the functions ϕ^j reads the following triangular relation for any $j \geq 1$,*

$$\phi^j(\mathbf{t} - [\lambda^{-1}]) = \phi^j(\mathbf{t}) - \lambda^{-1} \phi_1^j(\mathbf{t}) + \sum_{i=1}^{j-1} (\phi^i(\mathbf{t}) - \lambda^{-1} \phi_1^i(\mathbf{t})) b^{j-i}(\mathbf{t}), \quad (3.52)$$

where the b^j 's are functions that do not depend on x and are defined by

$$\sum_{j \geq 0} b^j z^{2j} = \operatorname{sech}(z\lambda^{-1}) \left[1 - z\lambda^{-1} \tanh(z\lambda^{-1}) - z\lambda^{-1} \tanh(\eta) + \tanh(z\lambda^{-1}) \tanh(\eta) \right]^{-1}, \quad (3.53)$$

with $b^0 = 1$, and³

$$\eta(\mathbf{t}; z) := \sum_{i \geq 2} t_{2i-1} z^{2i-1}.$$

Here $\operatorname{sech} = 1/\cosh$ and the exponent -1 stands for the multiplicative inverse of formal power series.

Proof. Using the fact that

$$\tanh^{-1}(z\lambda^{-1}) = z\lambda^{-1} + \sum_{i \geq 2} \frac{z^{2i-1}}{(2i-1)\lambda^{2i-1}}$$

for $z\lambda^{-1} \in (-1, 1)$, and applying the sum formulae of hyperbolic functions, one obtains that

$$\sum_{j \geq 1} \phi^j(\mathbf{t} - [\lambda^{-1}]) z^{2j-1} = \frac{\operatorname{sech}(z\lambda^{-1}) \sum_{j \geq 1} (\phi^j - \lambda^{-1} \phi_1^j) z^{2j-1}}{1 - z\lambda^{-1} \tanh(z\lambda^{-1}) - z\lambda^{-1} \tanh(\eta) + \tanh(z\lambda^{-1}) \tanh(\eta)}.$$

Moreover, the above denominator has its constant term equal to 1 and only even powers, so is its multiplicative inverse. Therefore, the series $\sum_{j \geq 0} b^j z^{2j}$ in Equation (3.53) is well defined and has indeed $b^0 = 1$, hence the triangular relation (3.52). □

³We emphasize the fact that $\eta(\mathbf{t}; \lambda)$ does not depend on $t_1 = x$, contrary to $\xi(\mathbf{t}; \lambda)$.

Remark 3.5.10. As in Lemma (3.5.5), this relation is to be compared with the one satisfied by the elementary Schur polynomials, namely,

$$p^j(\mathbf{t} - [\lambda^{-1}]) = p^j(\mathbf{t}) - \lambda^{-1}p^{j-1}(\mathbf{t}).$$

We can now prove Proposition 3.5.8 which concludes the proof of Theorem 3.5.4.

Proof of Proposition 3.5.8. We prove an equivalent equation which only uses differential operator:

$$\Delta_n e^{\xi(\mathbf{t}; \lambda)} = \lambda^n e^{\xi(\mathbf{t}; \lambda)} \frac{\tau_n(\mathbf{t} - [\lambda^{-1}])}{\tau_n(\mathbf{t})}.$$

Thanks to Lemma 3.5.9, we can rewrite the right-hand side as

$$\frac{\lambda^n e^{\xi(\mathbf{t}; \lambda)}}{\tau_n} \begin{vmatrix} \phi^1 - \lambda^{-1}\phi_1^1 & \cdots & \phi_{n-1}^1 - \lambda^{-1}\phi_n^1 \\ \vdots & & \vdots \\ \phi^n - \lambda^{-1}\phi_1^n + \sum_{i=1}^{j-1} (\phi^i - \lambda^{-1}\phi_1^i) b^{j-i} & \cdots & \phi_{n-1}^n - \lambda^{-1}\phi_n^n + \sum_{i=1}^{j-1} (\phi_{n-1}^i - \lambda^{-1}\phi_n^i) b^{j-i} \end{vmatrix}.$$

On the other hand, the left-hand side reads

$$\Delta_n e^{\xi(\mathbf{t}; \lambda)} = \frac{1}{\tau_n} \begin{vmatrix} e^{\xi(\mathbf{t}; \lambda)} & \lambda e^{\xi(\mathbf{t}; \lambda)} & \cdots & \lambda^n e^{\xi(\mathbf{t}; \lambda)} \\ \phi^1 & \phi_1^1 & \cdots & \phi_n^1 \\ \vdots & \vdots & & \vdots \\ \phi^n & \phi_1^n & \cdots & \phi_n^n \end{vmatrix}.$$

Using operations on rows and columns, we can easily check that these two expressions are equal. \square

Chapter 4

Polynomial tau functions of the Drinfeld–Sokolov hierarchies and generalized Schur polynomials

4.1 Semisimple Lie algebras: a brief review

In this section we briefly review the theory of semisimple Lie algebra. In the sequel, our starting point will often be the data of a semisimple Lie algebra together with a choice of a set simple roots and Weyl generators. We review these concepts. As stated above, for proofs of the statements below, we refer to the books [Hum72], [Car05], [EW06] and to the unpublished quality notes [Igu11].

Abelian and semisimple Lie algebras

Let \mathfrak{g} be a complex Lie algebra,¹ that is, an algebra over the field of complex numbers \mathbb{C} whose product, denoted $(X, Y) \mapsto [X, Y]$ satisfies the following properties: for any $X, Y, Z \in \mathfrak{g}$,

- ◊ Skew symmetry: $[X, Y] + [Y, X] = 0$;
- ◊ Jacobi's identity: $[X, [Y, Z]] + [Y, [Z, X]] + [Z, [X, Y]] = 0$.

Notice that in what follows, all Lie algebras are implicitly considered over complex numbers.

We say that \mathfrak{g} is *abelian* if its product is trivial, i.e. if $[X, Y] = 0$ for any $X, Y \in \mathfrak{g}$. We say that \mathfrak{g} is *simple* if it is finite dimensional,² non abelian and if it has no ideal other than $\{0\}$ and \mathfrak{g} itself. We say that \mathfrak{g} is *semisimple* if it is a direct sum of simple Lie algebra; in particular, \mathfrak{g} is finite-dimensional. If \mathfrak{g} is a semisimple Lie algebra which decomposes into $\mathfrak{g} = \mathfrak{g}_1 \oplus \cdots \oplus \mathfrak{g}_m$ with the \mathfrak{g}_i 's simple, then each \mathfrak{g}_i is a simple ideal of \mathfrak{g} .

We will sometimes use the term *finite Lie algebra* to mean *complex semisimple Lie algebra*; this terminology will become clear later with the introduction of *affine Lie algebras*.

¹We denote \mathfrak{g} the semisimple Lie algebra; the notation \mathfrak{g} will be used for the associated affine Lie algebra, as is done in [Kac90, Wu17].

²For some authors, the simplicity condition does not require finite-dimensionality. In our context, this requirement is rather customary and convenient.

Any algebra A can be naturally equipped with a structure of Lie algebra by defining the bracket $[X, Y] := XY - YX$ (the commutator) for any $X, Y \in A$. In particular, we denote by $\mathfrak{gl}(n)$ the algebra $\text{Mat}(n)$ of complex square matrices of size n equipped with the bracket defined by the commutator. A finite-dimensional Lie algebra is called *linear* if it is a Lie subalgebra of $\mathfrak{gl}(n)$ for some n .

Example 4.1.1 (Algebra $\mathfrak{sl}(2)$). The simplest example of a simple Lie algebra is the algebra $\mathfrak{sl}(2)$ made of three generators $\{H, E, F\}$ which satisfy the following relation,

$$[E, F] = H, \quad [H, E] = 2E, \quad [H, F] = -2F. \quad (4.1)$$

This algebra can be realized with complex square matrices of size 2 as follows:

$$H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad E = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Automorphisms and the adjoint representation

A *morphism* from a Lie algebra \mathfrak{g}_1 to another Lie algebra \mathfrak{g}_2 is a map $\phi : \mathfrak{g}_1 \rightarrow \mathfrak{g}_2$ such that $\phi([X, Y]) = [\phi(X), \phi(Y)]$ for any $X, Y \in \mathfrak{g}_1$. An *automorphism* of a Lie algebra \mathfrak{g} is a morphism $\phi : \mathfrak{g} \rightarrow \mathfrak{g}$ that is bijective (then its inverse is automatically a morphism too). The group of automorphisms (under composition) of a Lie algebra \mathfrak{g} is denoted $\text{Aut}(\mathfrak{g})$.

To any element $X \in \mathfrak{g}$ we associate a linear endomorphism $\text{ad}_X : \mathfrak{g} \rightarrow \mathfrak{g}$ defined by $\text{ad}_X(Y) = [X, Y]$ and called the *adjoint endomorphism* of X . The adjoint map is not a morphism but a *derivation*, i.e. for any $X, Y, Z \in \mathfrak{g}$,

$$\text{ad}_X[Y, Z] = [\text{ad}_X(Y), Z] + [Y, \text{ad}_X(Z)],$$

the above being equivalent to Jacobi's identity. The above equation can also be rewritten equivalently as

$$\text{ad}_{[X, Y]}(Z) = [\text{ad}_X, \text{ad}_Y](Z),$$

where $[\text{ad}_X, \text{ad}_Y] = \text{ad}_X \circ \text{ad}_Y - \text{ad}_Y \circ \text{ad}_X$ (the commutator within the algebra $\text{End}(\mathfrak{g})$ of endomorphisms of \mathfrak{g}). This allows to define the so-called *adjoint representation*

$$\begin{aligned} \text{ad} : \mathfrak{g} &\longrightarrow \text{End}(\mathfrak{g}) \\ X &\longmapsto \text{ad}_X. \end{aligned}$$

To sum up, Jacobi's identity is equivalent to the fact that the adjoint endomorphism $\text{ad}_X : \mathfrak{g} \rightarrow \mathfrak{g}$, associated to any element $X \in \mathfrak{g}$, is a derivation; it is also equivalent to the fact that the adjoint map $\text{ad} : \mathfrak{g} \rightarrow \text{End}(\mathfrak{g})$ is a morphism of algebras (i.e. defines a representation of \mathfrak{g} on \mathfrak{g} itself as a vector space).

Killing form

The adjoint representation $\text{ad} : \mathfrak{g} \rightarrow \text{End}(\mathfrak{g})$ defines a symmetric bilinear form $\kappa : \mathfrak{g} \otimes \mathfrak{g} \rightarrow \mathbb{C}$, called the Killing form

$$\kappa(X, Y) = (X, Y) := \text{tr}(\text{ad}_X \circ \text{ad}_Y). \quad (4.2)$$

Since the trace depends only on the endomorphism, the Killing form is well defined. Cartan's second criterion ([EW06], p.82) states that a complex Lie algebra is semisimple if and only if its Killing form is nondegenerate. Moreover, the Killing form is said to be *invariant* (*ad-invariant* or *associative*) in the sense that for any $X, Y, Z \in \mathfrak{g}$,

$$([X, Y], Z) = (X, [Y, Z]).$$

Using Schur's lemma ([Hum72], p. 26), one shows that any invariant scalar product (i.e. a nondegenerate symmetric bilinear form) on \mathfrak{g} is necessarily a scalar multiple of the Killing form ([Hum72], p.118).

Inner automorphisms

We say that an element $X \in \mathfrak{g}$ is *ad-semisimple* if its adjoint endomorphism $\text{ad}_X : \mathfrak{g} \rightarrow \mathfrak{g}$ is semisimple (by which we mean that it is diagonalizable). Accordingly, we say that $X \in \mathfrak{g}$ is *ad-nilpotent* if ad_X is nilpotent. For any ad-nilpotent element $X \in \mathfrak{g}$ we can define a linear endomorphism of \mathfrak{g}

$$\exp(\text{ad}_X) := \sum_{k \geq 0} \frac{1}{k!} (\text{ad}_X)^k,$$

often denoted e^{ad_X} . By expanding the exponential, it can be easily seen that $\exp(\text{ad}_X) : \mathfrak{g} \rightarrow \mathfrak{g}$ is actually an automorphism of algebras with inverse $\exp(-\text{ad}_X)$ ([Car05], p. 26). Such an automorphism is called *inner*. The set of all inner automorphisms then forms a subgroup of $\text{Aut}(\mathfrak{g})$ that we denote

$$\text{Inn}(\mathfrak{g}) = \{ \exp(\text{ad}_X) \mid X \in \mathfrak{g} \text{ ad-nilpotent} \}.$$

The composition of two inner automorphisms is given by the celebrated Baker–Campbell–Hausdorff (BCH) formula [Cam97, Bak02, Hau06]: for $X, Y \in \mathfrak{g}$ ad-nilpotent,

$$\exp(\text{ad}_X) \exp(\text{ad}_Y) = \exp(\text{ad}_{X \bullet Y}),$$

where

$$X \bullet Y = \sum_{n \geq 1} \frac{(-1)^{n-1}}{n} \sum_{i=1}^n \sum_{r_i + s_i > 0} \frac{[X^{r_1} Y^{s_1} \dots X^{r_n} Y^{s_n}]}{\left(\sum_{j=1}^n r_j + s_j \right) \cdot \prod_{k=1}^n r_k! s_k!}, \quad (4.3)$$

where

$$[X^{r_1} Y^{s_1} \cdots X^{r_n} Y^{s_n}] = \underbrace{[X, \cdots]}_{r_1}, \underbrace{[X, [Y, \cdots]}_{s_1}, \underbrace{[Y, \cdots]}_{r_n}, \underbrace{[X, [Y, \cdots]}_{s_n} Y \cdots].$$

Equation (4.3) is due to Dynkin [Dyn47] (see [Ser92] §7 as well) and clearly shows that the map $(X, Y) \rightarrow X \bullet Y$ is a homomorphism. The first few terms read

$$X \bullet Y = X + Y + \frac{1}{2}[X, Y] + \frac{1}{12}([X, [X, Y]] + [Y, [Y, X]]) + \cdots$$

Cartan subalgebras and root system

A subalgebra \mathfrak{h} of a semisimple Lie algebra \mathfrak{g} is called a *Cartan subalgebra* if \mathfrak{h}

- ◊ Is abelian;
- ◊ Contains only semisimple elements;
- ◊ Is maximal with respect to the dimension.

Any semisimple Lie algebra \mathfrak{g} contains a non-unique Cartan subalgebra. Moreover, for any two Cartan subalgebras \mathfrak{h}_1 and \mathfrak{h}_2 , there exists an ad-nilpotent element $X \in \mathfrak{g}$ such that $\mathfrak{h}_2 = e^{\text{ad}_X}(\mathfrak{h}_1)$ (e.g. [Car05], p. 34); we say that \mathfrak{h}_1 and \mathfrak{h}_2 are *conjugate* to one another. In particular, all Cartan subalgebras have the same dimension r ; we call r the *rank* of \mathfrak{g} and denote

$$\text{rk } \mathfrak{g} = \dim \mathfrak{h} = r.$$

Let \mathfrak{h} be a Cartan subalgebra of a semisimple Lie algebra \mathfrak{g} . We call a *root* of \mathfrak{g} any nonzero linear functional $\alpha \in \mathfrak{h}^*$ such that there exists a nonzero $X \in \mathfrak{g}$ satisfying the eigenvalue equation

$$[H, X] = \alpha(H)X, \quad \forall H \in \mathfrak{h}$$

(i.e. we co-diagonalize the elements of \mathfrak{h}). We denote by Δ the set of all roots and call it the *root system*. Since \mathfrak{h} is abelian and all its elements are ad-semisimple, then the ad_H 's, for $H \in \mathfrak{h}$, are simultaneously diagonalizable. This provides the so-called *root space decomposition* of \mathfrak{g} :

$$\mathfrak{g} = \mathfrak{h} \oplus \bigoplus_{\alpha \in \Delta} \mathfrak{g}_\alpha, \quad \mathfrak{g}_\alpha = \left\{ X \in \mathfrak{g} \mid \forall H \in \mathfrak{h}, [H, X] = \alpha(H)X \right\}.$$

We also denote $\mathfrak{g}_0 = \mathfrak{h}$. Thanks to Jacobi's identity, for any $\alpha, \beta \in \Delta$ and $X \in \mathfrak{g}_\alpha, Y \in \mathfrak{g}_\beta$, we have $[X, Y] \in \mathfrak{g}_{\alpha+\beta}$; in other words, the root space decomposition induces a gradation:

$$[\mathfrak{g}_\alpha, \mathfrak{g}_\beta] \subset \mathfrak{g}_{\alpha+\beta}. \quad (4.4)$$

As a consequence, any $X \in \mathfrak{g}_\alpha$ is ad-nilpotent. Each eigenspace \mathfrak{g}_α is a one-dimensional complex vector space ([EW06], p. 100). The Killing form κ satisfies $(\mathfrak{g}_\alpha, \mathfrak{g}_\beta) = 0$ for any $\alpha, \beta \in \Delta$ such that $\alpha + \beta \neq 0$. In particular, κ pairs \mathfrak{g}_α and $\mathfrak{g}_{-\alpha}$ nondegenerately.

Similarly, κ is also nondegenerate on $\mathfrak{h} = \mathfrak{g}_0$. Hence, we are provided with a natural isomorphism $\mathfrak{h}^* \rightarrow \mathfrak{h} : \phi \mapsto T_\phi$ such that $\phi(H) = (T_\phi, H)$ for any $H \in \mathfrak{h}$. The latter engenders a scalar product on \mathfrak{h}^* defined by, for any $\phi, \psi \in \mathfrak{h}^*$,

$$(\phi, \psi) := (T_\phi, T_\psi). \quad (4.5)$$

Using the the scalar product (4.5), we define a non symmetric pairing as follows, for any roots $\alpha, \beta \in \Delta$,

$$\langle \beta, \alpha \rangle := \frac{2(\beta, \alpha)}{(\alpha, \alpha)}.$$

It automatically satisfies $\langle \alpha, \alpha \rangle = 2$. Among other properties, the roots of a semisimple Lie algebra satisfy the following.

- (i) The set Δ spans \mathfrak{h}^* ;
- (ii) If $\alpha \in \Delta$, then $-\alpha \in \Delta$ and no other multiple of α does;
- (iii) For any $\alpha, \beta \in \Delta$, the quantity $\langle \beta, \alpha \rangle$ is an integer;
- (iv) The set Δ is stable by the following reflections:

$$\sigma_\alpha(\beta) = \beta - \langle \beta, \alpha \rangle \alpha \in \Delta.$$

If one replaces \mathfrak{h}^* by a generic vector space E , then properties (i) to (iv) define what is called an *abstract root system* on E . As a matter of fact, the classification of root systems classifies simple Lie algebras [Hum72, Car05].

Simple roots, Cartan matrix

From the root system Δ we can always extract a (non unique) set of *simple roots* (or a *basis*) denoted $\Pi = \{\alpha_1, \dots, \alpha_r\} \subset \Delta$, satisfying the following properties:

- ◊ The set Π spans \mathfrak{h}^* ;
- ◊ For any $\alpha \in \Delta$, there exist $k_1, \dots, k_r \in \mathbb{Z}$ such that

$$\alpha = k_1 \alpha_1 + \dots + k_r \alpha_r,$$

such that the k_i 's all have the same sign (or are null).

(Note that if we relax the assumption that Π has r elements, the two properties above allow to prove it.) From the set of simple roots Π we naturally define a partition $\Delta = \Delta_+ \sqcup \Delta_-$, where Δ_+ (resp. Δ_-) is made of all roots which decompose in Π with only nonnegative coefficients (resp. nonpositive); obviously $\Pi \subset \Delta_+$.

If α, β are two *distinct* simple roots, then their pairing $\langle \beta, \alpha \rangle$ is heavily constrained. Indeed, $\langle \beta, \alpha \rangle \in \{0, -1, -2, -3\}$; more precisely, up to exchanging α and β ,

- ◊ Either $\langle \beta, \alpha \rangle = 0$ and then $\langle \alpha, \beta \rangle = 0$ as well;
- ◊ Or $\langle \beta, \alpha \rangle = -1$ and then $\langle \alpha, \beta \rangle \in \{-1, -2, -3\}$.

By numbering the simple roots $\Pi = \{\alpha_1, \dots, \alpha_r\}$ we can define, for any semisimple Lie algebra, a matrix $(C_{ij}) \in \text{Mat}(r, \mathbb{Z})$, called the *Cartan matrix*, by

$$C_{ij} = \langle \alpha_j, \alpha_i \rangle.$$

(Notice that the indices are purposefully put in the “reverse” order.) The constraints on the pairings $\langle \beta, \alpha \rangle$ induce similar constraints on the Cartan matrix as follows. First, $C_{ii} = 2$, then if $i \neq j$, then $C_{ij} \in \{0, -1, -2, -3\}$; up to exchanging i and j ,

- ◊ Either $C_{ij} = 0$ and then $C_{ji} = 0$ as well;
- ◊ Or $C_{ij} = -1$ and then $C_{ji} \in \{-1, -2, -3\}$.

The Cartan matrix determines the semisimple Lie algebra \mathfrak{g} up to isomorphism.

Weyl generators, Chevalley involution

Let \mathfrak{g} be a rank- r semisimple Lie algebra with Cartan matrix (C_{ij}) and a set of simple roots $\Pi = \{\alpha_1, \dots, \alpha_r\}$. A set of *Weyl generators* is a subset $\{E_i, F_i, H_i \mid 1 \leq i \leq r\}$ of \mathfrak{g} which generates the algebra \mathfrak{g} and whose elements satisfy the following relations: for any $1 \leq i, j \leq r$,

$$[E_i, F_j] = \delta_{ij} H_i, \quad [H_i, E_j] = C_{ij} E_i, \quad [H_i, F_j] = -C_{ij} F_j, \quad [H_i, H_j] = 0, \quad (4.6)$$

where δ_{ij} is the Kronecker symbol. In particular the H_i 's generate a Cartan subalgebra that we still denote \mathfrak{h} . Any semisimple Lie algebra admits a set of Weyl generators. For any two sets of Weyl generators $\{E_i, F_i, H_i\}$ and $\{\tilde{E}_i, \tilde{F}_i, \tilde{H}_i\}$, there exist an inner automorphism e^{ad_X} and a permutation σ of $\{1, \dots, r\}$ such that for any $1 \leq i \leq r$,

$$\tilde{E}_i = e^{\text{ad}_X} E_{\sigma(i)}, \quad \tilde{F}_i = e^{\text{ad}_X} F_{\sigma(i)}, \quad \tilde{H}_i = e^{\text{ad}_X} H_{\sigma(i)}.$$

Besides the relations of Equation (4.6), the Weyl generators satisfy the so-called Chevalley-Serre relations:

$$(\text{ad}_{E_i})^{1-C_{ij}} E_j = 0, \quad (\text{ad}_{F_i})^{1-C_{ij}} F_j = 0. \quad (4.7)$$

Relations of Equations (4.6) and (4.7) constitute a complete set of relations in \mathfrak{g} . With any set of Weyl generators we associate the so-called *Chevalley involution* $\Omega : \mathfrak{g} \rightarrow \mathfrak{g}$ (i.e. such that $\Omega^2 = \text{id}$) defined by

$$\Omega(H_i) = -H_i, \quad \Omega(E_i) = -F_i, \quad \Omega(F_i) = -E_i.$$

From Equation 4.6, we easily see that the *derived algebra*, namely $\mathfrak{g}' = [\mathfrak{g}, \mathfrak{g}]$ of \mathfrak{g} , is \mathfrak{g} itself:

$$\mathfrak{g}' := [\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}. \quad (4.8)$$

This is specific to semisimple algebra and does not necessarily happen in other types of Lie algebras.

Cartan, Borel and nilpotent subalgebras

As we have seen, for any root $\alpha \in \Delta$ (so $\alpha \neq 0$), each element of \mathfrak{g}_α is ad-nilpotent. For that reason, we call the *principal nilpotent subalgebra* the subalgebra \mathfrak{n}_+ (resp. \mathfrak{n}_-) generated by the E_i 's (resp. the F_i 's) for $1 \leq i \leq r$. To that we add the so-called *Borel subalgebra* \mathfrak{b}_+ (resp. \mathfrak{b}_-) generated by the H_i 's and the E_i 's (resp. the H_i 's and the F_i 's), along with the Cartan subalgebra

\mathfrak{h} generated by the H_i 's. To sum up, given a set of Weyl generators $\{E_i, F_i, H_i \mid 1 \leq i \leq r\}$ come five important subalgebras:

$$\begin{aligned}\mathfrak{n}_+ &= \langle E_1, \dots, E_r \rangle, & \mathfrak{b}_+ &= \langle E_1, \dots, E_r, H_1, \dots, H_r \rangle, & \mathfrak{h} &= \langle H_1, \dots, H_r \rangle, \\ \mathfrak{n}_- &= \langle F_1, \dots, F_r \rangle, & \mathfrak{b}_- &= \langle F_1, \dots, F_r, H_1, \dots, H_r \rangle.\end{aligned}$$

Considering the partition $\Delta = \Delta_+ \sqcup \Delta_-$ between positive and negative roots (w.r.t. the set of simple roots Π),

$$\mathfrak{n}_+ = \bigoplus_{\alpha \in \Delta_+} \mathfrak{g}_\alpha, \quad \mathfrak{n}_- = \bigoplus_{\alpha \in \Delta_-} \mathfrak{g}_\alpha,$$

along with $\mathfrak{b}_+ = \mathfrak{h} \oplus \mathfrak{n}_+$ and $\mathfrak{b}_- = \mathfrak{h} \oplus \mathfrak{n}_-$. In the sequel, we will often denote $\mathfrak{n} := \mathfrak{n}_-$ and $\mathfrak{b} := \mathfrak{b}_-$ when there is no ambiguity. In the end, the semisimple Lie algebra \mathfrak{g} decomposes as

$$\mathfrak{g} = \mathfrak{n}_- \oplus \mathfrak{h} \oplus \mathfrak{n}_+.$$

Highest root, normalized invariant bilinear form

Let $\alpha \in \Delta$ be a root which we decompose as $\alpha = k_1\alpha_1 + \dots + k_r\alpha_r$ in the basis $\Pi = \{\alpha_1, \dots, \alpha_r\}$. We call the *height* of α the integer $\text{ht}(\alpha) = k_1 + \dots + k_r$. There exists a unique root ([Car05], p. 251),

$$\theta = a_1\alpha_1 + \dots + a_r\alpha_r \in \Delta_+ \quad (4.9)$$

called the *highest root*, such that for any root $\alpha = k_1\alpha_1 + \dots + k_r\alpha_r$, we have $k_i \leq a_i$. The integer

$$h = 1 + a_1 + \dots + a_r \quad (4.10)$$

is called the *Coxeter number* of \mathfrak{g} (the +1 becomes clear later).

As stated at the end of the section on the Killing form, any invariant scalar product is a scalar multiple of the Killing form $\kappa : \mathfrak{g} \otimes \mathfrak{g} \rightarrow \mathbb{C}$. So let

$$\begin{aligned}b_0 : \mathfrak{g} \otimes \mathfrak{g} &\longrightarrow \mathbb{C} \\ X \otimes Y &\longmapsto (X, Y)_0\end{aligned}$$

be an invariant nondegenerate symmetric bilinear form on \mathfrak{g} and let $\mu \in \mathbb{C}$ be such that $b_0 = \mu \cdot \kappa$; in particular, the restriction of b_0 on the Cartan subalgebra \mathfrak{h} is also nondegenerate. The form b_0 then induces an isomorphism $\mathfrak{h}^* \rightarrow \mathfrak{h} : \phi \rightarrow T_{0,\phi}$ defined by $(T_{0,\phi}, H)_0 = \phi(H)$, for all $\phi \in \mathfrak{h}^*$ and $H \in \mathfrak{h}$. As usual, this naturally defines a nondegenerate symmetric bilinear form on \mathfrak{h}^* that we still denote $b_0 : \mathfrak{h}^* \otimes \mathfrak{h}^* \rightarrow \mathbb{C}$, defined by

$$(\phi, \psi)_0 := (T_{0,\phi}, T_{0,\psi})_0$$

for all $\phi, \psi \in \mathfrak{h}^*$. We fix the constant $\mu \in \mathbb{C}$ by imposing the condition

$$(\theta, \theta)_0 = 2.$$

The invariant nondegenerate symmetric bilinear form $b_0 : \mathfrak{h}^* \otimes \mathfrak{h}^* \rightarrow \mathbb{C}$ is called the *normalized bilinear form* on \mathfrak{h}^* (see for instance [Kac90], p.85 §6.4).

Root space \mathbb{Z} -gradation

Using the Coxeter number, we can refine the root space decomposition into a \mathbb{Z} -gradation as follows. Given the Weyl generators, we set

$$\deg E_i = 1, \quad \deg F_i = -1, \quad \deg H_i = 0.$$

Since the highest root is a sum $h-1$ simple roots (where h is the Coxeter number, see Equation 4.10) we get the following \mathbb{Z} -gradation:

$$\mathring{\mathfrak{g}} = \bigoplus_{k \in \mathbb{Z}} \mathring{\mathfrak{g}}^k, \quad \mathring{\mathfrak{g}}^k = \{0\} \text{ if } k \notin \{1-h, \dots, h-1\}. \quad (4.11)$$

Note that, again, $\mathring{\mathfrak{g}}^0 = \mathring{\mathfrak{h}}$. This gradation amounts to setting

$$\mathring{\mathfrak{g}}^1 = \bigoplus_{\alpha_i \in \Pi} \mathring{\mathfrak{g}}_{\alpha_i} = \text{Span}\{E_1, \dots, E_r\}, \quad \mathring{\mathfrak{g}}^{-1} = \bigoplus_{\alpha_i \in \Pi} \mathring{\mathfrak{g}}_{-\alpha_i} = \text{Span}\{F_1, \dots, F_r\}.$$

In particular, $\dim \mathfrak{h} = \dim \mathring{\mathfrak{g}}^1 = \dim \mathring{\mathfrak{g}}^{-1} = \text{rk } \mathring{\mathfrak{g}} = r$. It is easy to see from Equation (4.6) that the remaining subspaces have the following form: for any $k \in \{1, \dots, h-1\}$,

$$\begin{aligned} \mathring{\mathfrak{g}}^k &= \text{Span} \left\{ [E_{i_1}, [E_{i_2}, \dots, [E_{i_{k-1}}, E_{i_k}] \dots]] \mid 1 \leq i_1, \dots, i_k \leq r \right\}. \\ \mathring{\mathfrak{g}}^{-k} &= \text{Span} \left\{ [F_{i_1}, [F_{i_2}, \dots, [F_{i_{k-1}}, F_{i_k}] \dots]] \mid 1 \leq i_1, \dots, i_k \leq r \right\}. \end{aligned}$$

Simple co-roots and associated Weyl generators

Starting from the set of simple roots Π , there is a constructive way of building a set of Weyl generators as follows. For any simple root $\alpha_i \in \Pi$, we have defined the element $T_{0, \alpha_i} \in \mathring{\mathfrak{h}}$ which is such that $\alpha_i(H) = (T_{0, \alpha_i}, H)_0$ for any $H \in \mathfrak{h}$. We then define the so-called *simple coroots* by

$$H_i := \frac{2T_{0, \alpha_i}}{(\alpha_i, \alpha_i)_0}. \quad (4.12)$$

Since b_0 is proportional to κ , we automatically get $\alpha_j(H_i) = 2(\alpha_j, \alpha_i)/(\alpha_i, \alpha_i) = C_{ij}$. Now let elements $X_j \in \mathring{\mathfrak{g}}_{\alpha_j}$ and $X_{-j} \in \mathring{\mathfrak{g}}_{-\alpha_j}$, where $j \in \{1, \dots, r\}$, then we also get $[H_i, X_j] = C_{ij}X$ and $[H_i, X_{-j}] = -C_{ij}X_{-j}$. It remains to choose $E_i \in \mathring{\mathfrak{g}}_{\alpha_i}$ and $F_j \in \mathring{\mathfrak{g}}_{-\alpha_j}$ such that $[E_i, F_j] = \delta_{ij}H_i$. First, for any $H' \in \mathfrak{h}$

$$(H', [X_i, X_{-j}])_0 = ([H', X_i], X_{-j})_0 = \alpha_j(H') (X_i, X_{-j})_0 = (H', T_{0, \alpha_i})_0 (X_i, X_{-j})_0.$$

Since b_0 is nondegenerate on $\mathring{\mathfrak{h}}$, it follows that $[X_i, X_{-j}] = (X_i, X_{-j})_0 T_{0, \alpha_i}$. Moreover, $(X_i, X_{-j})_0 = 0$ unless $i = j$, and if so, b_0 pairs nondegenerately $\mathring{\mathfrak{g}}_{\alpha_i}$ and $\mathring{\mathfrak{g}}_{-\alpha_i}$ so that we can choose X_{-i} such that $(X_i, X_{-i})_0 = 1$. Therefore, we obtain $[X_i, X_{-i}] = T_{0, \alpha_i} \neq 0$ and $[X_i, X_{-j}] = 0$ if $i \neq j$. Then it suffices to set $E_i = X_i$ and

$$F_j = \frac{2X_{-j}}{(\alpha_j, \alpha_j)_0}$$

to get the relation $[E_i, F_j] = \delta_{ij}H_i$. The family $\{E_i, F_i, H_i \mid 1 \leq i \leq r\}$ is then a set of Weyl generators.

Notice that since $C_{ii} = \langle \alpha_i, \alpha_i \rangle = 2$, each triplet $\{E_i, F_i, H_i\}$ forms a subalgebra of \mathfrak{g} with relations $[E_i, F_i] = H_i$, $[H_i, E_i] = 2E_i$ and $[H_i, F_i] = -2F_i$, i.e. isomorphic to $\mathfrak{sl}(2)$ (see Example 4.1.1). This exact construction also applies to any root $\alpha \in \Delta$ (not necessarily a simple root).

Remark 4.1.2. Since the simple roots $\{\alpha_1, \dots, \alpha_r\}$ form a basis of the vector space \mathfrak{h}^* , then the simple co-roots $\{H_1, \dots, H_r\}$ form a basis of \mathfrak{h} . On the other hand, by definition of the roots and of the Cartan matrix, $C_{ij}E_j = [H_i, E_j] = \alpha_j(H_i)E_j$. In other words, the vector $(C_{1j}, \dots, C_{rj}) \in \mathbb{Z}^r$ reads the coordinates of the root α_j in the basis $\Pi^\vee = \{H_1, \dots, H_r\}$; we rewrite this statement as follows:

$$(C_{ij})_{i,j \in [1,r]} = \text{Mat}_{\Pi^\vee}(\alpha_1, \dots, \alpha_r).$$

In particular, the matrix (C_{ij}) is invertible.

Dynkin diagrams

Dynkin diagrams are a way to encode the entries of the Cartan matrix in a graph with simple, double or triple edges, and possibly a direction on the edges. Since the Cartan matrix determines the semisimple Lie algebra \mathfrak{g} up to isomorphism, so does the Dynkin diagram.

We start by numbering the simple roots $\Pi = \{\alpha_1, \dots, \alpha_r\}$, which defines the Cartan matrix $(C_{ij}) \in \text{Mat}(r, \mathbb{Z})$. The Dynkin diagram of \mathfrak{g} has r vertices, each one labelled by a simple root. Earlier we described all the different cases that can arise for the computation of $C_{ij} = \langle \alpha_j, \alpha_i \rangle$. According to these cases, we draw an edge (or not) between vertices α_i and α_j and add a direction on the edge as follows:

- ◊ $\alpha_i \bullet \bullet \alpha_j$ if $C_{ij} = 0$ (which implies that $C_{ji} = 0$ as well);
- ◊ $\alpha_i \bullet \rightarrow \alpha_j$ if $C_{ij} = -1$ and $C_{ji} = -1$ as well;
- ◊ $\alpha_i \bullet \rightleftarrows \alpha_j$ if $C_{ij} = -1$ and $C_{ji} = -2$;
- ◊ $\alpha_i \bullet \rightleftarrows \alpha_j$ if $C_{ij} = -1$ and $C_{ji} = -3$.

The Dynkin diagram of a simple Lie algebra is a connected graph. The Dynkin diagram of a semisimple Lie algebra is a graph which connected components are the Dynkin diagrams of its simple ideals. A case by case study of simple Lie algebra shows that all Dynkin diagrams (of simple Lie algebras) are exhausted in Table 4.1.

4.2 Untwisted affine Kac–Moody algebras

In this section we review some of the properties of loop algebras and affine algebras that we will use in the sequel. We refer to the books [Kac90] and [Car05] for more details and proofs.

4.2.1 Definition and basic properties

The finite-dimensional data

In accordance with the previous section, we need the following material. Let \mathfrak{g} be a semisimple Lie algebra of rank r with Cartan matrix $C \in \text{Mat}(r, \mathbb{Z})$. Let Δ be the root system and $\Pi =$

Table 4.1: Finite Dynkin diagrams

Type	Lie algebra	Dynkin diagram	Type	Dynkin diagram
$A_n, n \geq 1$	$\mathfrak{sl}(n+1)$		E_6	
$B_n, n \geq 3$	$\mathfrak{o}(2n+1)$		E_7	
$C_n, n \geq 2$	$\mathfrak{sp}(2n)$		E_8	
$D_n, n \geq 4$	$\mathfrak{o}(2n)$		F_4	
			G_2	

$\{\alpha_1, \dots, \alpha_r\}$ a set of simple roots. Let $\theta \in \Delta$ be the highest root w.r.t. Π and let b_0 be the normalized bilinear form, i.e. such that $(\theta, \theta)_0 = 2$. We choose Weyl generators $\{E_i, F_i, H_i \mid 1 \leq i \leq r\}$ with H_i is the simple co-root associated to α_i (4.12) and E_i, F_i chosen accordingly. Let $\Omega : \mathfrak{g} \rightarrow \mathfrak{g}$ be the Chevalley involution associated to our Weyl generators. As usual, we denote by $\mathfrak{h}, \mathfrak{b} = \mathfrak{b}_-$ and $\mathfrak{n} = \mathfrak{n}_-$ the Cartan, negative Borel and negative principal nilpotent subalgebras respectively.

On top of that, we need two specific vectors: $E_\theta \in \mathfrak{g}_\theta$ and $E_{-\theta} \in \mathfrak{g}_{-\theta}$. Since both \mathfrak{g}_θ and $\mathfrak{g}_{-\theta}$ are one-dimensional, we need two constraints to fix E_θ and $E_{-\theta}$. We impose

$$\Omega(E_\theta) = -E_{-\theta}, \quad (E_\theta, E_{-\theta})_0 = 1.$$

Notice that since θ is the highest root, we automatically get

$$[E_\theta, \mathfrak{n}_+] = 0, \quad [E_{-\theta}, \mathfrak{n}] = 0. \quad (4.13)$$

Finally, we set $\alpha_0 := -\theta$ (not a simple root) and define $H_0 := T_{0, \alpha_0}$ ³, the co-root of α_0 w.r.t. b_0 . Then using the invariance property of b_0 and the normalization $(E_\theta, E_{-\theta})_0 = 1$, we easily show that

$$H_0 = [E_{-\theta}, E_\theta].$$

We decompose H_0 in the basis Π^\vee of \mathfrak{h} and α_0 in the basis Π of \mathfrak{h}^* (similarly to Equation (4.9)) to get the following relations,

$$\begin{aligned} \alpha_0 + a_1 \alpha_1 + \dots + a_r \alpha_r &= 0, \\ H_0 + a_1^\vee H_1 + \dots + a_r^\vee H_r &= 0. \end{aligned}$$

(Notice that $a_i \neq \theta(H_i)$ and $a_i^\vee \neq \alpha_i(T_{0, \theta})$ since neither the basis Π nor Π^\vee is orthogonal.) The integer

$$h^\vee = 1 + a_1^\vee + \dots + a_r^\vee$$

³Notice that $H_0 = 2T_{0, -\theta}/(\theta|\theta)_0$ since $(\theta|\theta)_0 = 2$

we call the *dual Coxeter number* of \mathfrak{g} . Notice that since $H_0 = T_{0,\alpha_0}$ and $\alpha_0 = -\theta$, we have

$$\alpha_0(H_0) = -\theta(T_{0,-\theta}) = (T_{0,-\theta}, T_{0,-\theta})_0 = (\theta|\theta)_0 = 2, \quad (4.14)$$

which also implies $\theta(T_{0,\theta}) = 2$. The integers (a_0, a_1, \dots, a_r) , with $a_0 = 1$ (resp. $(a_0^\vee, a_1^\vee, \dots, a_r^\vee)$, with $a_0^\vee = 1$), are called the *Kac labels* (resp. the *dual Kac labels*).

Structure of loop and affine algebras

Given all these data, we call the *loop algebra* of \mathfrak{g} the vector space

$$\bar{\mathfrak{g}} = \mathfrak{g}[\lambda, \lambda^{-1}] := \mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}] = \left\{ \sum_{k=m}^n X_k \lambda^k \mid m, n \in \mathbb{Z}, X_k \in \mathfrak{g} \right\},$$

i.e. Laurent polynomials over \mathfrak{g} . This vector space is naturally equipped with a Lie bracket defined by

$$[X\lambda^p, Y\lambda^q] = [X, Y]\lambda^{p+q},$$

for any $X, Y \in \mathfrak{g}$ and $p, q \in \mathbb{Z}$. We extend the invariant scalar product $b_0 : \mathfrak{g} \rightarrow \mathfrak{g}$ on $\bar{\mathfrak{g}}$ by setting

$$(X\lambda^p, Y\lambda^q)_0 := \delta_{p+q,0}(X, Y)_0,$$

which produces an invariant nondegenerate symmetric bilinear form on $\bar{\mathfrak{g}}$. It is clearly isotropic since $(X\lambda^p, X\lambda^p)_0 = 0$.

We then proceed to a central extension of the loop algebra and call it the *affine algebra* associated to \mathfrak{g} ,⁴ namely,

$$\mathfrak{g} = \bar{\mathfrak{g}} \oplus \mathbb{C}c,$$

with the following Lie bracket: for any $X, Y \in \mathfrak{g}$, $p, q \in \mathbb{Z}$ and $a, b \in \mathbb{C}$,

$$[X\lambda^p + ac, Y\lambda^q + bc] = [X, Y]\lambda^{p+q} + p\delta_{p+q,0}(X, Y)_0c. \quad (4.15)$$

Clearly, c belongs to the center of the algebra \mathfrak{g} (i.e. c commutes with all elements of \mathfrak{g}), making $\mathbb{C}c$ into an ideal of \mathfrak{g} . Therefore, we have a natural surjective homomorphism

$$\pi : \mathfrak{g} \rightarrow \mathfrak{g}/\mathbb{C}c \simeq \bar{\mathfrak{g}},$$

hence the central extension. We again extend b_0 on \mathfrak{g} to produce an invariant nondegenerate bilinear form on \mathfrak{g} defined by

$$(X\lambda^p + aK, Y\lambda^q + bK)_0 := \delta_{p+q,0}(X, Y)_0.$$

Remark 4.2.1. The affine algebras that arise through the construction we described above are often called *untwisted* affine algebras. But since we will not consider the twisted cases, we simply call them *affine algebras* if there is no ambiguity.

⁴Notice that the actual affine algebra is a co-central extension $\tilde{\mathfrak{g}} = \mathfrak{g}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c \oplus \mathbb{C}d$, i.e. such that its derived algebra reads $\tilde{\mathfrak{g}}' = \mathfrak{g}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c$ [Kac90]. In other words, here we call *affine algebra* what is actually its derived algebra. This denomination is convenient here since the co-central element d would play no role.

Weyl generators and affine Cartan matrix

Starting from the Weyl generators $\{E_i, F_i, H_i \mid 1 \leq i \leq r\}$, we define the following elements of the affine algebra \mathfrak{g} : for all $1 \leq i \leq r$,

$$e_i = E_i, \quad f_i = F_i, \quad \check{\alpha}_i = H_i, \quad (4.16)$$

$$e_0 = E_{-\theta}\lambda, \quad f_0 = E_{\theta}\lambda^{-1}, \quad \check{\alpha}_0 = H_0 + c, \quad (4.17)$$

where we recall that $H_0 = [E_{-\theta}, E_{\theta}]$. Then there exists a matrix $(A_{ij}) \in \text{Mat}(r+1, \mathbb{Z})$, called the *affine Cartan matrix*⁵ of \mathfrak{g} , such that for any $0 \leq i \leq r$,

$$[e_i, f_i] = \delta_{ij}\check{\alpha}_i, \quad [\check{\alpha}_i, e_i] = A_{ij}e_j, \quad [\check{\alpha}_i, f_i] = -A_{ij}f_j, \quad [\check{\alpha}_i, \check{\alpha}_j] = 0, \quad (4.18)$$

Notice that by Equation 4.15, $[e_0, f_0] = [E_{-\theta}, E_{\theta}] + (E_{-\theta}, E_{\theta})_0 K = \check{\alpha}_0$, as stated above. Moreover, since $[E_{\theta}, \mathfrak{n}_+] = [E_{-\theta}, \mathfrak{n}] = 0$, we have that

$$[f_0, \mathfrak{n}_+] = 0, \quad [e_0, \mathfrak{n}] = 0, \quad (4.19)$$

(recall that $\mathfrak{n} = \mathfrak{n}_-$).

It is well known [Kac90] that the family $\{e_i, f_i, \check{\alpha}_i \mid 0 \leq i \leq r\}$ generates the affine algebra \mathfrak{g} ; the same family—except for $\check{\alpha}_0$ that we change for H_0 —generates the loop algebra $\bar{\mathfrak{g}}$. These vectors also satisfy the so-called affine Chevalley-Serre relations: for any $0 \leq i \leq 0$,

$$(\text{ad}_{e_i})^{1-A_{ij}} e_j = 0, \quad (\text{ad}_{f_i})^{1-A_{ij}} f_j = 0. \quad (4.20)$$

Relations of Equations (4.18) and (4.20) form a complete set of relations in \mathfrak{g} . Finally, we can extend the the Chevalley involution $\Omega : \mathfrak{g} \rightarrow \mathfrak{g}$ by

$$\Omega(\check{\alpha}_i) = -\check{\alpha}_i, \quad \Omega(e_i) = -f_i, \quad \Omega(f_i) = -e_i.$$

Affine Cartan subalgebra and Kac labels

By Equation 4.18, the vectors $\check{\alpha}_0, \dots, \check{\alpha}_r$ are all ad-semisimple (i.e. each $\text{ad}_{\check{\alpha}_i}$ is diagonalizable) and commute pairwise. We denote

$$\mathfrak{h} = \text{Span}\{\check{\alpha}_0, \dots, \check{\alpha}_r\} = \mathfrak{h}^{\circ} \oplus \mathbb{C}c,$$

and call it the *affine Cartan subalgebra* of the affine algebra \mathfrak{g} . Since $\check{\alpha}_0 = H_0 + c$ and since c is linearly independent from the loop algebra, then the family $\{\check{\alpha}_0, \dots, \check{\alpha}_r\}$ also forms a basis of \mathfrak{h} .

In Remark 4.1.2, we explained that the vector (C_{1j}, \dots, C_{rj}) , of the Cartan matrix C of the semisimple algebra \mathfrak{g} , actually reads the components of the simple root α_j in the basis $\Pi^{\vee} = \{H_1, \dots, H_r\}$ of the Cartan subalgebra \mathfrak{h}° . In particular, we saw that C is invertible.

⁵Again, following Remark 4.2.1, affine Cartan matrices are actually more general than the ones that arise in the above construction; here we only discuss the case of *untwisted* affine Cartan matrices.

Now since for $1 \leq i, j \leq r$, we have $[\check{\alpha}_i, e_j] = [H_i, E_j] = C_{ij}e_j$, it follows that $A_{ij} = C_{ij}$ for all $1 \leq i, j \leq r$. In particular, the matrix C is an invertible $r \times r$ submatrix of the matrix A . Second, $A_{00} = 2$, as shown below,

$$[\check{\alpha}_0, e_0] = [H_0, E_0]\lambda = \alpha_0(H_0)E_0\lambda = 2e_0,$$

where we have used the fact that c commutes with all elements, along with $\alpha_0(H_0) = 2$ (Equation 4.14). Finally, since $\alpha_0 = -\theta$, then for any $1 \leq i \leq r$,

$$[\check{\alpha}_i, e_0] = [H_i, E_0] = -\theta(H_i)E_0 = -\sum_{j=1}^r a_j \alpha_j(H_i) e_0,$$

so that $A_{i0} = -\sum_{j=1}^r a_j C_{ij}$ for all $1 \leq i \leq r$. Similarly, we prove that $A_{0i} = -\sum_{j=1}^r a_i^\vee C_{ij}$. To sum up, the untwisted affine Cartan matrix A has the form

$$A = \begin{pmatrix} 2 & A_{01} & \cdots & A_{0r} \\ A_{10} & C_{11} & \cdots & C_{1r} \\ \vdots & \vdots & & \vdots \\ A_{r0} & C_{r1} & \cdots & C_{rr} \end{pmatrix}, \quad A_{i0} = -\sum_{j=1}^r a_j C_{ij}, \quad A_{0i} = -\sum_{j=1}^r a_i^\vee C_{ij}.$$

where, again, $C_{ij} = \alpha_j(H_i)$ and $C_{ii} = 2$ in particular. From this we show that $\dim \text{Ker} A = 1$; more precisely we show that

$$A \begin{pmatrix} a_0 \\ \vdots \\ a_r \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}, \quad (a_0^\vee \cdots a_r^\vee) A = (0 \cdots 0). \quad (4.21)$$

We prove the equation on the left. It amounts to proving that for any $0 \leq i \leq r$, we have $\sum_{j=0}^r A_{ij} a_j = 0$. If $i \neq 0$, the identity follows directly from Equation (4.21). If $i = 0$, then it amounts to

$$\begin{aligned} \sum_{j=0}^r A_{0j} a_j &= a_0 A_{00} + \sum_{j=1}^r a_j A_{0j} = 2 - \sum_{j,c=1}^r a_j C_{cj} a_c^\vee = 2 - \sum_{j,c=1}^r a_j \alpha_j(H_c) a_c^\vee \\ &= 2 - \alpha_0(H_0) = 0. \end{aligned}$$

where we have used Equation (4.14).

Remark 4.2.2. Now we have seen to different types of Cartan matrix: the matrix C which is associated to the semisimple Lie algebra \mathfrak{g} and the matrix A , associated to the affine algebra \mathfrak{g} . Cartan matrices can actually be generalized to any matrix with integer entries such that the diagonal entries equal 2, the off-diagonal entries are nonpositive, and such that if an off-diagonal entry is 0, then the transposed entry is 0 too. These generalized Cartan matrices can be classified such that they correspond either to a semisimple Lie algebra (in which case they are called *finite*), an affine algebra (in which case they are called *affine*) or to none of those (in which case they are called *indefinite*). That is the reason why we call semisimple complex Lie algebras *finite algebras*.

Untwisted affine Cartan matrices and Dynkin diagrams

If the underlying semisimple Lie algebra \mathfrak{g} is actually *simple* and has Dynkin diagram of type X_n ($X = A, B, C, \dots$), then the Dynkin diagram of the untwisted affine algebra \mathfrak{g} is denoted $X_n^{(1)}$. The untwisted affine Dynkin diagram determines the Cartan matrix as in the case of semisimple Lie algebras. The Cartan matrices of these algebras are exhausted in Table 4.2.1. The white dot denotes the vertex labelled by the root $\alpha_0 = -\theta$. Type $A_1^{(1)}$ is treated separately.

Table 4.2: Untwisted affine Dynkin diagrams (or extended Dynkin diagrams)

Type	Affine algebra	Dynkin diagram	Type	Dynkin diagram
$A_1^{(1)}$	$\widehat{\mathfrak{sl}}(2)$		$E_6^{(1)}$	
$A_n^{(1)}, n \geq 2$	$\widehat{\mathfrak{sl}}(n+1)$		$\tilde{E}_7^{(1)}$	
$B_n^{(1)}, n \geq 3$	$\widehat{\mathfrak{v}}(2n+1)$		$E_8^{(1)}$	
$C_n^{(1)}, n \geq 2$	$\widehat{\mathfrak{sp}}(2n)$		$F_4^{(1)}$	
$D_n^{(1)}, n \geq 4$	$\widehat{\mathfrak{v}}(2n)$		$G_2^{(1)}$	

4.2.2 Gradations and principal Heisenberg subalgebra

Principal and homogeneous gradations

For any vector $\mathbf{s} = (s_0, s_1, \dots, s_r) \in \mathbb{Z}_{\geq 0}^{r+1}$ we can associate a gradation on the affine algebra \mathfrak{g} as follows. We set $\deg_{\mathbf{s}} e_i = s_i$, $\deg_{\mathbf{s}} f_i = -s_i$ and $\deg_{\mathbf{s}} \check{\alpha}_i = 0$ for all $0 \leq i \leq r$. We denote by $\mathfrak{g}_{k,\mathbf{s}}$ the space of elements of \mathfrak{g} of degree $k \in \mathbb{Z}$ and set

$$[\mathfrak{g}_{j,\mathbf{s}}, \mathfrak{g}_{k,\mathbf{s}}] \subset \mathfrak{g}_{j+k,\mathbf{s}}.$$

Then we have the decomposition

$$\mathfrak{g} = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_{k,\mathbf{s}}.$$

We also define, for any $m \in \mathbb{Z}$,

$$\mathfrak{g}_{\geq m,\mathbf{s}} = \bigoplus_{k \geq m} \mathfrak{g}_{k,\mathbf{s}}, \quad \mathfrak{g}_{< m,\mathbf{s}} = \bigoplus_{k < m} \mathfrak{g}_{k,\mathbf{s}}.$$

Two gradations will be of particular interests: the *principal gradation*, which correspond to the vector $\mathbf{s}^1 = (1, 1, \dots, 1)$ and the *homogeneous gradation*, which correspond to the vector $\mathbf{s}^0 = (1, 0, \dots, 0)$. When there is no ambiguity, we will denote $\mathfrak{g}^k = \mathfrak{g}_{k,\mathbf{s}^1}$ (resp. $\mathfrak{g}_k = \mathfrak{g}_{k,\mathbf{s}^0}$) the

degree- c subspace under the principal (resp. homogeneous) gradation. For the sake of clarity, these two gradations are defined as follows,

$$\begin{array}{llll} \text{Principal:} & \deg e_i = 1, & \deg f_i = -1, & \deg \check{\alpha}_i = 0, & \mathfrak{g} = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}^k, \\ \text{Homogeneous:} & \deg e_i = \delta_{i,0}, & \deg f_i = -\delta_{i,0}, & \deg \check{\alpha}_i = 0, & \mathfrak{g} = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g}_k. \end{array}$$

It follows that

$$\mathfrak{g}^0 = \mathfrak{h} = \mathfrak{h} \oplus \mathbb{C}c, \quad \mathfrak{g}_0 = \mathfrak{g} \oplus \mathbb{C}c.$$

Clearly, both the principal and homogeneous gradations transfer to the loop algebra via the quotient by the ideal $\mathbb{C}c$. We denote $\bar{\mathfrak{g}} = \bigoplus_{k \in \mathbb{Z}} \bar{\mathfrak{g}}^k$ (for the principal gradation) and $\bar{\mathfrak{g}} = \bigoplus_{k \in \mathbb{Z}} \bar{\mathfrak{g}}_k$ (for the homogeneous gradation).

Notice that that \mathbb{Z} -gradation of \mathfrak{g} induced by the root space decomposition, i.e $\deg E_i = 1$, $\deg F_i = -1$ and $\deg H_i = 0$, is preserved by the inclusion map $\mathfrak{g} \hookrightarrow \mathfrak{g}$. Similarly to the \mathbb{Z} -gradation of \mathfrak{g} induced by the root space decomposition, the principal gradation of \mathfrak{g} amounts to setting

$$\mathfrak{g}^1 = \text{Span}\{e_0, e_1, \dots, e_r\}, \quad \mathfrak{g}^{-1} = \text{Span}\{f_0, f_1, \dots, f_r\}.$$

In particular, $\dim \mathfrak{h} = \dim \mathfrak{g}^1 = \dim \mathfrak{g}^{-1} = r + 1$. It easy to see from Equation (4.18) that the remaining subspaces have the following form: for any $k > 1$,

$$\mathfrak{g}^k = \text{Span} \{ [e_{i_1}, [e_{i_2}, \dots, [e_{i_{k-1}}, e_{i_k}] \dots]] \mid 0 \leq i_1, \dots, i_k \leq r \}. \quad (4.22)$$

$$\mathfrak{g}^{-k} = \text{Span} \{ [f_{i_1}, [f_{i_2}, \dots, [f_{i_{k-1}}, f_{i_k}] \dots]] \mid 0 \leq i_1, \dots, i_k \leq r \}. \quad (4.23)$$

On the other hand, the homogeneous gradation amounts to setting $\mathfrak{g}_1 = \mathbb{C}e_0$ and $\mathfrak{g}_{-1} = \mathbb{C}f_0$. The homogeneous subspace of degree $k > 1$ is generated by arbitrary long commutators $[e_{i_1}, [e_{i_2}, \dots, [e_{i_{k-1}}, e_{i_M}] \dots]]$, with $M \geq k$, such that the index 0 occurs exactly k times among the i_1, \dots, i_M . A similar assertion holds for \mathfrak{g}_{-k} . Consequently, $\mathfrak{g}_k \subset \mathfrak{g} \otimes \mathbb{C}\lambda^k$, and since the homogeneous gradation produces a decomposition of \mathfrak{g} , it follows that for any $k \in \mathbb{Z}$,

$$\mathfrak{g}_k = \mathfrak{g} \otimes \mathbb{C}\lambda^k,$$

hence the name *homogeneous gradation*. In particular,

$$\mathfrak{g}_{>0} = \lambda \mathfrak{g}[\lambda], \quad \mathfrak{g}_{<0} = \lambda^{-1} \mathfrak{g}[\lambda^{-1}].$$

Moreover, it is clear from the discussion above that for any $k \geq 1$,

$$\mathfrak{g}_k \subset \bigoplus_{j \geq 1} \mathfrak{g}^j, \quad \mathfrak{g}_{-k} \subset \bigoplus_{j \geq 1} \mathfrak{g}^{-j}.$$

As a consequence, $\mathfrak{g}_{>0} \subset \mathfrak{g}^{>0}$ and $\mathfrak{g}_{<0} \subset \mathfrak{g}^{>0}$.

Finite exponents and principal Cartan subalgebra

The exponents can be defined both in the case of the finite algebras (i.e. semisimple Lie algebra) and affine algebras. The first are called *finite exponents* and the latter *affine exponents*. As we will see, the affine exponents are related to the finite exponents via an arithmetical modular relation. We start with the finite exponents.

We still use the same set of Weyl generators $\{E_i, F_i, H_i \mid 1 \leq i \leq r\}$ where the H_i 's are the simple co-roots and the E_i 's and F_i 's are chosen accordingly. We define the so-called *finite principal cyclic element*:

$$J = E_0 + E_1 + \cdots + E_r.$$

We call the *principal Cartan subalgebra* of \mathfrak{g} the subalgebra $\text{Ker}(\text{ad}_J)$. We then decompose $\text{Ker}(\text{ad}_J)$ through the gradation of Equation (4.11) that is induced by the simple roots:

$$\text{Ker}(\text{ad}_J) = \bigoplus_{c=1-h}^{h-1} \text{Ker}(\text{ad}_J)^c, \quad \text{Ker}(\text{ad}_J)^k = \text{Ker}(\text{ad}_J) \cap \mathfrak{g}^k.$$

It is a classical result that [Kos59] that $\text{Ker}(\text{ad}_J)$ is a Cartan subalgebra of \mathfrak{g} , hence the name. In particular,

$$\dim \text{Ker}(\text{ad}_J) = \text{rk } \mathfrak{g} = r.$$

Then an integer $k \in \{0, \dots, h-1\}$ is said to be a finite exponent of \mathfrak{g} if $\text{Ker}(\text{ad}_J)^{k[h-1]} \neq \{0\}$; if so, we say that k has multiplicity $m_k = \dim \text{Ker}(\text{ad}_J)^{k[h-1]}$. (Here, $k[h-1]$ means $k \bmod h-1$.) We gather the finite exponents into an ordered multi-set that we denote

$$\mathbb{F}(\mathfrak{g}) = \{k_1^{m_1}, \dots, k_{h-1}^{m_{h-1}}\} = \left\{ k^{m_k} \mid 1 \leq k \leq h-1, m_k = \dim \text{Ker}(\text{ad}_J)^{k[h-1]} \right\}.$$

A multi-set means that each element has a multiplicity; a null multiplicity means that the element does not belong to the set. The fact that $\mathbb{F}(\mathfrak{g})$ is ordered simply means that $k_i < k_j$ if $i < j$. Since $\dim \text{Ker}(\text{ad}_J) = r$, we must have $m_1 + \cdots + m_{h-1} = r$. On the other hand, one can show that $m_1 = m_{h-1} = 1$. When there is no ambiguity, we will simply denote $\mathbb{F} = \mathbb{F}(\mathfrak{g})$.

Example 4.2.3 (Finite exponents of $\mathfrak{sl}(n)$). We start with the algebra $\mathfrak{g} = \mathfrak{sl}(3)$. We can present $\mathfrak{sl}(3)$ via the following system of generators:

$$\begin{aligned} E_1 &= \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & F_1 &= \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, & H_1 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \\ E_2 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, & F_2 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, & H_2 &= \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}. \end{aligned} \tag{4.24}$$

It is not hard to see that $\mathfrak{sl}(3)$ has Coxeter number equal to 3. Therefore, we can define the elements $E_0 \in \mathfrak{g}^{-2}$, $F_0 \in \mathfrak{g}^2$ and $H_0 \in \mathfrak{g}^0$ by $E_0 = [F_1, F_2]$, $F_0 = [E_1, E_2]$ and $H_0 = [E_0, F_0]$, which

reads

$$E_0 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad F_0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad H_0 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Consequently, the finite principal cyclic element, which is given by $J = E_0 + E_1 + E_2$, is the matrix

$$J = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

Then a direct calculation shows that the kernel of ad_J is

$$\text{Ker}(\text{ad}_J) = \mathbb{C}(E_0 + E_1 + E_2) \oplus \mathbb{C}(F_0 + F_1 + F_2)$$

Yet $E_0 + E_1 + E_2 = J \in \mathfrak{sl}(3)^{1[3]}$ and $F_0 + F_1 + F_2 \in \mathfrak{sl}(3)^{2[3]}$ so that the finite exponents of $\mathfrak{sl}(3)$ are (where all multiplicity are equal to 1)

$$\mathbb{F}(\mathfrak{sl}(3)) = \{1, 2\}.$$

Similarly, we could have found that the finite exponents of $\mathfrak{sl}(2)$ read

$$\mathbb{F}(\mathfrak{sl}(2)) = \{1\}.$$

One can easily generalize these results to find that the finite principal cyclic element of $\mathfrak{sl}(n)$ is given by $J = E_{1,n} + \sum_{i=1}^{n-1} E_{i+1,i}$, that is,

$$J = \begin{pmatrix} 0 & \cdots & & 0 & 1 \\ 1 & & & & 0 \\ 0 & 1 & & & \vdots \\ \vdots & & \ddots & & \\ 0 & \cdots & & 0 & 1 & 0 \end{pmatrix},$$

and that the set of finite exponents (where all multiplicity are equal to 1) is

$$\mathbb{F}(\mathfrak{sl}(n)) = \{1, 2, \dots, n-1\}. \quad (4.25)$$

Affine exponents and principal Heisenberg subalgebra

The affine exponents are constructed in a similar way. First, we define the *affine principal cyclic element* of the loop algebra

$$\bar{\Lambda} = e_0 + e_1 + \cdots + e_r \in \bar{\mathfrak{g}}.$$

We denote $\bar{\mathfrak{s}}$ its kernel, that is,

$$\bar{\mathfrak{s}} := \text{Ker}(\text{ad}_{\bar{\Lambda}} : \bar{\mathfrak{g}} \rightarrow \bar{\mathfrak{g}}).$$

It is proven in [Kac90] that $\bar{\mathfrak{s}}$ is an abelian algebra. We decompose $\bar{\mathfrak{s}} = \bigoplus_{k \in \mathbb{Z}} \bar{\mathfrak{s}}^k$ through the principal gradation. Then we say that an integer $j \in \mathbb{Z}$ is an affine exponent of \mathfrak{g} if $\bar{\mathfrak{s}}^j \neq \{0\}$, and if so, with multiplicity $m_j = \dim \bar{\mathfrak{s}}^j$. We gather the affine exponents into a multi-set that we denote

$$\mathbb{E}(\mathfrak{g}) = \{j^{m_j} \mid j \in \mathbb{Z}, m_j = \dim \bar{\mathfrak{s}}^j\}.$$

We also denote $\mathbb{E}^+(\mathfrak{g}) = \mathbb{E}(\mathfrak{g}) \cap \mathbb{Z}_{>0}$ the set of *positive exponents* (0 is never an exponent). The affine exponents are related to the finite exponents via the following relation ([Kac90], p. 298 Proposition 14.3):

$$\mathbb{E}(\mathfrak{g}) = \{j + ph \mid j \in \mathbb{F}(\mathring{\mathfrak{g}}), p \in \mathbb{Z}\},$$

the latter we often write in the condensed form $\mathbb{E}(\mathfrak{g}) = \mathbb{F}(\mathring{\mathfrak{g}}) + h\mathbb{Z}$. We will denote $\mathbb{E} = \mathbb{E}(\mathfrak{g})$ when there is no ambiguity.

Following [Kac90], we can lift the abelian algebra $\bar{\mathfrak{s}} \subset \bar{\mathfrak{g}}$ in \mathfrak{g} as follows. There exists a subalgebra $\mathfrak{s} \subset \mathfrak{g}$ such that, through the principal gradation, $\dim \mathfrak{s}^k = \dim \bar{\mathfrak{s}}^k$ for any exponent $k \in \mathbb{E}(\mathfrak{g})$, and which admits a set of generators $\Lambda_j, j \in \mathbb{E}(\mathfrak{g})$, satisfying the Heisenberg commutation relations:

$$[\Lambda_i, \Lambda_j] = i\delta_{i+j,0} \cdot c, \quad \forall i, j \in \mathbb{E}(\mathfrak{g}). \quad (4.26)$$

Moreover, $\mathfrak{s}/\mathbb{C}c = \bar{\mathfrak{s}}$. We call \mathfrak{s} the principal Heisenberg subalgebra of \mathfrak{g} . Notice that, in particular,

$$[\Lambda_i, \Lambda_j] = 0, \quad \forall i, j \in \mathbb{E}^+(\mathfrak{g}), \quad (4.27)$$

which hold for $i, j \in \mathbb{E}(\mathfrak{g}) \cap \mathbb{Z}_{<0}$ as well, making $\mathfrak{s}^{>0}$ and $\mathfrak{s}^{<0}$ into abelian subalgebras of \mathfrak{s} . The condition (4.26) on the generators Λ_j leaves yet one degree of freedom; we normalize them via the condition

$$(\Lambda_i, \Lambda_j)_0 = h\delta_{i+j,0}, \quad (4.28)$$

where h is the Coxeter number of $\mathring{\mathfrak{g}}$. We denote $\nu \in \mathbb{C}$ the complex number such that $\Lambda_1 = \nu\Lambda$.

Still in [Kac90], it is proven that $\bar{\Lambda} \in \bar{\mathfrak{g}}$ is ad-semisimple, i.e.

$$\bar{\mathfrak{g}} = \bar{\mathfrak{s}} \oplus \text{Im}(\text{ad}_{\bar{\Lambda}}). \quad (4.29)$$

Now, on the level of \mathfrak{g} , since c is central, then $c \in \text{Ker}(\text{ad}_{\Lambda}) = \mathfrak{s}$. On the other hand, from Equation (4.26) we see that $c = \nu^{-1}\text{ad}_{\Lambda}(\Lambda_{-1})$, where $\Lambda_1 = \nu\Lambda$, so that $c \in \text{Im}(\text{ad}_{\Lambda})$. Since $\bar{\mathfrak{g}} = \text{Ker}(\text{ad}_{\bar{\Lambda}}) \oplus \text{Im}(\text{ad}_{\bar{\Lambda}})$ and $\mathfrak{g} = \bar{\mathfrak{g}} \oplus \mathbb{C}c$, it follows that

$$\mathfrak{g} = \mathfrak{s} + \text{Im}(\text{ad}_{\Lambda}), \quad \mathfrak{s} \cap \text{Im}(\text{ad}_{\Lambda}) = \mathbb{C}c. \quad (4.30)$$

Equation (4.30) is a key fact in the building of the Drinfeld–Sokolov hierarchies.

Example 4.2.4 (Affine exponents of $\mathfrak{sl}(n)$). Very similar computations to those of Example 4.2.3 allow to compute the affine exponents of $\mathfrak{sl}(n)$. Again, we treat the case of $\mathfrak{sl}(3)$, the other cases are computed in the same fashion. Recall that

$$e_0 = E_0\lambda, \quad e_1 = E_1, \quad e_2 = E_2.$$

Then using the same Weyl generators of $\mathfrak{sl}(3)$ as in Equation (4.24), the (projection of the) affine principal cyclic element $\Lambda = e_0 + e_1 + e_2$ reads

$$\Lambda = \begin{pmatrix} 0 & 0 & \lambda \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

In order to compute $\mathfrak{s} = \text{Ker}(\text{ad}_\Lambda)$, it is convenient to first compute the powers of Λ :

$$\Lambda^2 = \begin{pmatrix} 0 & \lambda & 0 \\ 0 & 0 & \lambda \\ 1 & 0 & 0 \end{pmatrix}, \quad \Lambda^3 = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix},$$

from which we get that $\Lambda^{3k+1} = \lambda^k \Lambda$ and $\Lambda^{3k+2} = \lambda^k \Lambda^2$ for any $k \in \mathbb{Z}$. Knowing that, a direct computation shows that the affine principal Heisenberg subalgebra decomposes as

$$\mathfrak{s} = \bigoplus_{k \in \mathbb{Z}} \mathbb{C} \Lambda^{3k+1} \oplus \mathbb{C} \Lambda^{3k+2} = \left\{ \begin{pmatrix} 0 & \lambda c_2 & \lambda c_1 \\ c_1 & 0 & \lambda c_2 \\ c_2 & c_1 & 0 \end{pmatrix} \mid c_1, c_2 \in \mathbb{C}[\lambda, \lambda^{-1}] \right\}.$$

Yet w.r.t. the principal gradation, $\Lambda = e_0 + e_1 + e_2 \in \widehat{\mathfrak{sl}}(3)^1$, so that the set of exponents reads

$$\mathbb{E}(\mathfrak{sl}(3)) = \bigcup_{k \in \mathbb{Z}} \{3k + 1, 3k + 2\} = \mathbb{Z} - 3\mathbb{Z}.$$

With similar computations we can show that

$$\mathbb{E}(\mathfrak{sl}(2)) = \bigcup_{k \in \mathbb{Z}} \{2k + 1\} = \mathbb{Z} - 2\mathbb{Z} = \mathbb{Z}^{\text{odd}}.$$

One can easily generalize these results to find that the affine principal cyclic element of $\widehat{\mathfrak{sl}}(n)$ is given by

$$\Lambda = \begin{pmatrix} 0 & \cdots & 0 & \lambda \\ 1 & & & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix},$$

and that the set of affine exponents (where all multiplicity are equal to 1) is

$$\mathbb{E}(\mathfrak{sl}(n)) = \bigcup_{k \in \mathbb{Z}} \{nk + 1, nk + 2, \dots, nk + n - 1\} = \mathbb{Z} - n\mathbb{Z}. \quad (4.31)$$

4.3 The Drinfeld–Sokolov hierarchies

In their seminal paper [DS84], Drinfeld–Sokolov built the so-called Drinfeld–Sokolov hierarchies based on the loop algebra $\bar{\mathfrak{g}} = \mathring{\mathfrak{g}}[\lambda, \lambda^{-1}]$. Hollowood and Miramontes [HM93] showed that one can build the exact same hierarchies of equations based on the affine algebra $\mathfrak{g} = \bar{\mathfrak{g}} \oplus \mathbb{C}c$ and that this approach allows to define a notion of dressing operator. In the present section, we follow the approach of C.-Z. Wu [Wu17] who, inspired by [HM93], produced a systematic definition of tau symmetric Hamiltonian densities (in the sense of [DZ08])—and therefore of a tau function—for the Drinfeld–Sokolov hierarchy associated to any affine algebra \mathfrak{g} .

We shall notice that the original construction of Drinfeld and Sokolov makes use the principal and homogeneous gradations, respectively defined by the vectors $\mathbf{s}^1 = (1, \dots, 1)$ and $\mathbf{s}^0 = (1, 0, \dots, 0) = (\delta_{0,j})_{j \in [1,r]}$ (see Section 4.2.2), but also applies if the homogeneous gradation is replaced by any vector of the form $(\delta_{i,j})_{j \in [1,r]}$. In this thesis, we are only interested in those defined by the homogeneous gradation.

On the other hand, de Groot, Hollowood and Miramontes generalized the DS hierarchies as follows. Take two vectors $\mathbf{s} = (s_0, \dots, s_r)$ and $\mathbf{s}' = (s'_0, \dots, s'_r)$ of relatively coprime non-negative integers. We say that $\mathbf{s} \geq \mathbf{s}'$ if $s_i = 0$ implies $s'_i = 0$ (e.g. $\mathbf{s}^1 \geq \mathbf{s}^0$). Then de Groot–Hollowood–Miramontes showed, in particular, how to adapt the Drinfeld–Sokolov procedure when one replaces \mathbf{s}^1 by \mathbf{s} and \mathbf{s}^0 by \mathbf{s}' [GHM92, HM93]. They also generalized the construction to the case when the principal cyclic element $\Lambda = e_0 + \dots + e_r$ is replaced by any constant element of \mathfrak{g} which is semisimple; and further generalizations [GHM92]. This gives rise to the generalized Drinfeld–Sokolov hierarchies. Notice that the work of de Groot–Hollowood–Miramontes applies to untwisted affine algebras.

Finally, Kac and Wakimoto also generalized the DS hierarchies in the following sense [KW89]. For each affine Lie algebra \mathfrak{g} together with an integrable highest weight representation V of \mathfrak{g} and a vertex operator construction R of V , Kac and Wakimoto formulated a hierarchy of soliton equations. These equations can be written down in terms of Hirota bilinear equations and their super analogue [KW89]. When \mathfrak{g} is the untwisted affinization of a simply laced finite Lie algebra, the Kac–Wakimoto hierarchy coincides with the corresponding generalized Drinfeld–Sokolov hierarchy defined by Groot, Hollowood and Miramontes [GHM92, HM93]. In particular, if the highest weight representation is the basic one, and the vertex operator realization is constructed from the principal Heisenberg subalgebra \mathfrak{s} of \mathfrak{g} , then the Kac–Wakimoto hierarchy is equivalent to the Drinfeld–Sokolov hierarchy associated to \mathfrak{g} for the homogeneous gradation.

4.3.1 Matrix Lax equations of the DS hierarchies

Lie algebra of \mathfrak{g} -valued formal power series

Let $\mathring{\mathfrak{g}}$ be a semisimple Lie algebra, $\bar{\mathfrak{g}} = \mathring{\mathfrak{g}}[\lambda, \lambda^{-1}]$ be the associated loop algebra and $\mathfrak{g} = \bar{\mathfrak{g}} \oplus \mathbb{C}c$ be the affine algebra with the material described in the previous section. We consider the vector

space

$$\mathfrak{g}[[x]] := \mathfrak{g} \otimes \mathbb{C}[[x]] = \left\{ \sum_{k \geq 0} \xi^{[k]} x^k \mid \xi^{[k]} \in \mathfrak{g} \right\}$$

of \mathfrak{g} -valued formal power series in x . The elements of this vector space will be what we call our *functions*. We equip this space with the \mathbb{Z} -gradation in powers of x :

$$\mathfrak{g}[[x]] = \bigoplus_{k \in \mathbb{Z}} \mathfrak{g} \otimes \mathbb{C}x^k;$$

subspaces of degree $k < 0$ are null. This vector space admits a natural bracket

$$\left[\sum_{k \geq 0} \xi^{[j]} x^j, \sum_{k \geq 0} \eta^{[k]} x^k \right] = \sum_{j, k \geq 0} [\xi^{[j]}, \eta^{[k]}] x^{j+k}. \quad (4.32)$$

where the bracket $[\xi^{[j]}, \eta^{[k]}]$ in \mathfrak{g} is given by Equation (4.15) (Equation (4.32) is actually the same kind of bracket as the one on the loop algebra). The bracket (4.32) is well defined for each term of degree c in $\mathfrak{g} \otimes \mathbb{C}x^k$ is clearly a finite sum.

Next we consider the multiplicative group

$$G = \{ \exp(\text{ad}_T) \mid T \in \mathfrak{g}[[x]] \text{ ad-nilpotent} \}.$$

This group acts in on $\mathfrak{g}[[x]]$ in a similar way the group of inner automorphisms of \mathfrak{g} acts on \mathfrak{g} : for $T \in G$ and $\xi \in \mathfrak{g}[[x]]$,

$$e^{\text{ad}_T} \xi = \sum_{k, n \geq 0} \frac{1}{n!} \left[\sum_{j \geq 0} \text{ad} \left(T^{[j]} x^j \right) \right]^n \xi^{[k]} x^k,$$

where as before, $\text{ad} \left(T^{[j]} x^j \right) \xi^{[k]} x^k = [T^{[j]}, \xi^{[k]}] x^{j+k}$. This action is well defined since, again, each term of a given degree involves only finitely many terms.

Gauge equivalence classes of connections

The initial data of the *Drinfeld–Sokolov hierarchy of \mathfrak{g} -type* (or simply the \mathfrak{g} -DS hierarchy) consists of a connection on the algebra $\mathfrak{g}[[x]]$ of the specific form

$$\mathcal{L} = \partial + \Lambda + Q, \quad Q \in \mathring{\mathfrak{b}}[[x]], \quad (4.33)$$

where $\partial = \partial_x$ and $\Lambda = e_0 + e_1 + \cdots + e_r \in \mathfrak{g}$ is the affine principal cyclic element. This type of connection first appeared in [DS84] (p. 2010). The set of such connections we denote by

$$\text{op}_{\mathfrak{g}} = \left\{ \partial + \Lambda + Q \mid Q \in \mathring{\mathfrak{b}}[[x]] \right\}.$$

Remark 4.3.1. By a connection we simply mean an linear endomorphism of $\mathfrak{g}[[x]]$ which satisfies Leibniz's rule; i.e. a derivation. The action of \mathcal{L} on $\mathfrak{g}[[x]]$ that makes it into a connection we define by

$$\left[\mathcal{L}, \xi \otimes x^k \right] = \xi \otimes kx^{k-1} + [\Lambda + Q, \xi] \otimes x^k, \quad (4.34)$$

for any $\xi \in \mathfrak{g}$ and $k \geq 0$, where $[\Lambda + Q, \xi]$ is given by Equation (4.32). We call them connections in reference to connections on a vector bundle on a manifold. Here we use the terminology and some notations of [LVY18].

For $\mathcal{L} = \partial + \Lambda + Q \in \text{op}_{\mathfrak{g}}$ and $T \in \mathfrak{g}[[x]]$ ad-nilpotent, we define

$$e^{\text{ad}_T} \mathcal{L} := \partial + e^{\text{ad}_T}(\Lambda + Q) - \nabla_T T, \quad (4.35)$$

where

$$\nabla_T T = \frac{e^{\text{ad}_T} - 1}{\text{ad}_T} \partial T := \sum_{k \geq 0} \frac{1}{(k+1)!} (\text{ad}_T)^k \partial(T), \quad (4.36)$$

and where $\partial(T)$ is the derivative of T w.r.t. x component-wise (the first equality is simply notational). For the moment, $e^{\text{ad}_T} \mathcal{L}$ is considered simply as a derivation on $\mathfrak{g}[[x]]$ and may not belong to $\text{op}_{\mathfrak{g}}$ anymore.

Remark 4.3.2. Equations (4.35) and (4.36) become clear when we choose a matrix realization of \mathfrak{g} . In that case, e^T is well defined as a matrix exponential, and, by Leibniz's rule,

$$e^{\text{ad}_T} \partial = e^T \partial \circ e^{-T} = \partial + e^T \partial(e^{-T}).$$

The derivative of the matrix exponential is given by the formula

$$\partial e^A = e^A \frac{1 - e^{-\text{ad}_A}}{\text{ad}_A} \partial A,$$

for A any square matrix (see for instance [Ros02], Theorem 5 §1.2). Replacing A by $-T$, one finds that

$$e^{\text{ad}_T} \partial = \partial + \frac{1 - e^{\text{ad}_T}}{\text{ad}_T} \partial T = \partial - \nabla_T T, \quad (4.37)$$

hence Equation (4.36).

Since any element of \mathfrak{n} is ad-nilpotent and since \mathfrak{n} is a subalgebra of \mathfrak{g} , we can define the following subgroup N of G :

$$N = \{ \exp(\text{ad}_S) \mid S \in \mathfrak{n}[[x]] \} \subset G.$$

The following proposition holds.

Proposition 4.3.3 (Gauge classes). *The group N acts on $\text{op}_{\mathfrak{g}}$, i.e. for any connection $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ and any $S \in \mathfrak{n}[[x]]$, the derivation*

$$e^{\text{ads}} \mathcal{L} = \partial + e^{\text{ads}}(\Lambda + Q) - \nabla_S S$$

belongs to $\text{op}_{\mathfrak{g}}$. We say that \mathcal{L} and $e^{\text{ads}} \mathcal{L}$ are gauge equivalent, or that \mathcal{L} and $e^{\text{ads}} \mathcal{L}$ belong to the same gauge class denoted $[\mathcal{L}]$.

Proof. We need to prove that for any $S \in \mathfrak{n}[[x]]$ and any $\mathcal{L} = \partial + \Lambda + Q \in \text{op}_{\mathfrak{g}}$, there exists $\tilde{Q} \in \mathfrak{b}[[x]]$ such that

$$e^{\text{ads}} \mathcal{L} = \partial + \Lambda + \tilde{Q}.$$

From Equation (4.35), it amounts to finding $\tilde{Q} \in \mathfrak{b}[[x]]$ such that

$$\Lambda + \tilde{Q} = e^{\text{ads}}(\Lambda + Q) - \nabla_S S.$$

Developing the exponentials, we get

$$\tilde{Q} = \sum_{k \geq 1} \frac{1}{k!} (\text{ad}_S)^k \Lambda + \sum_{k \geq 0} \frac{1}{k!} (\text{ad}_S)^k Q - \sum_{k \geq 0} \frac{1}{(k+1)!} (\text{ad}_S)^k \partial S.$$

First, using $[e_0, \mathfrak{n}] = 0$ (4.19), we see that

$$\sum_{k \geq 1} \frac{1}{k!} (\text{ad}_S)^k \Lambda = \sum_{k \geq 1} \frac{1}{k!} (\text{ad}_S)^k (e_1 + \cdots + e_r).$$

Yet using the root space decomposition of $\mathfrak{g} = \bigoplus_k \mathfrak{g}^k$, we have $e_1, \dots, e_r \in \mathfrak{g}^1$ and $S \in \mathfrak{g}^{<0}[[x]]$, so that

$$\text{ad}_S(e_1 + \cdots + e_r) \in \mathfrak{b}[[x]],$$

where we used $\mathfrak{b} = \mathfrak{g}^{\leq 0}$. Similarly, since $Q \in \mathfrak{b}[[x]]$, it follows that $\text{ad}_S Q \in \mathfrak{b}[[x]]$ as well. Finally, $\partial S \in \mathfrak{n}[[x]]$ by linearity of ∂ , which implies that $\text{ad}_S \partial S \in \mathfrak{n}[[x]] \subset \mathfrak{b}[[x]]$. We conclude that $\tilde{Q} \in \mathfrak{b}[[x]]$. \square

Canonical coordinates system

Let us introduce an alternative definition to the finite exponents [Kos59]. Define the *finite principal nilpotent element*

$$I = J - E_0 = E_1 + \cdots + E_r.$$

Since $I \in \mathfrak{g}^1$, then for all $j \in \{1 - h, \dots, h - 1\}$, we have $\text{ad}_I \mathfrak{g}^{j-1} \subset \mathfrak{g}^j$. Moreover, for all $j \in \{1, \dots, h - 1\}$, the map $\text{ad}_I : \mathfrak{g}^{-j-1} \rightarrow \mathfrak{g}^{-j}$ is injective and an integer $j \in \{1, \dots, h - 1\}$ is a finite exponent of \mathfrak{g} if and only if the map $\text{ad}_I : \mathfrak{g}^{-j-1} \rightarrow \mathfrak{g}^{-j}$ is *not* surjective; if so with multiplicity $m_j = \dim \mathfrak{g}^{-j-1} - \dim \mathfrak{g}^{-j}$ [Kos59].

Therefore for any j a finite exponent of \mathfrak{g} , since $\text{ad}_I : \mathfrak{g}^{-j-1} \rightarrow \mathfrak{g}^{-j}$ is injective and not surjective, we can choose a subspace $\mathcal{V}^j \subset \mathfrak{g}^{-j}$ such that $\mathfrak{g}^{-j} = \mathcal{V}^j \oplus [I, \mathfrak{g}^{-j-1}]$. Now since $\mathfrak{b} = \mathfrak{g}^{\leq 0}$ and $\mathfrak{n} = \mathfrak{g}^{< 0}$, we get

$$\mathfrak{b} = \mathcal{V} \oplus [I, \mathfrak{n}], \quad \mathcal{V} = \bigoplus_{j \in \mathbb{F}(\mathfrak{g})} \mathcal{V}^j, \quad (4.38)$$

where $\mathbb{F}(\mathfrak{g})$ are the finite exponents of \mathfrak{g} . Since $\mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n}$ and since $\text{ad}_I : \mathfrak{n} \rightarrow \mathfrak{b}$ is injective, it follows that $\dim \mathcal{V} = \dim \mathfrak{b} - \dim \mathfrak{n} = \dim \mathfrak{h}$, that is

$$\dim \mathcal{V} = \text{rk } \mathfrak{g} = r.$$

Such a subspace $\mathcal{V} \subset \mathfrak{b}$ is called a *Drinfeld–Sokolov gauge* for reasons that become clear with the next proposition.

Proposition 4.3.4 (Canonical coordinates). *Let $\mathcal{V} \subset \mathfrak{b}$ be a gauge and let $\mathcal{L} = \partial + \Lambda + Q \in \text{op}_{\mathfrak{g}}$. There exists a unique function $S \in \mathfrak{n}[[x]]$ such that the connection $\tilde{\mathcal{L}} = e^{-\text{ad}_S} \mathcal{L}$ has the form*

$$\tilde{\mathcal{L}} = \partial + \Lambda + \tilde{Q}, \quad \tilde{Q} \in \mathcal{V}[[x]].$$

Moreover, the functions \tilde{Q} and S are differential polynomials in the components of Q . The functions $\{\tilde{Q}^j \mid j \in \mathbb{F}(\mathfrak{g})\}$ are called the canonical coordinates of \mathcal{L} .

Proof. We want to equate $\partial + \Lambda + Q = e^{\text{ad}_S}(\partial + \Lambda + \tilde{Q})$. Since $[e_0, \mathfrak{n}] = 0$, then $e^{\text{ad}_S} \Lambda = e_0 + e^{\text{ad}_S} I$, so that it amounts to solving

$$I + Q = e^{\text{ad}_S}(I + \tilde{Q}) - \nabla_S S.$$

Let us decompose \tilde{Q} and S following the gradation of \mathfrak{b} , i.e.

$$\begin{aligned} \tilde{Q} &= \tilde{Q}^0 + \tilde{Q}^1 + \cdots + \tilde{Q}^{h-1}, \\ S &= S^1 + \cdots + S^{h-1}, \end{aligned}$$

with $\tilde{Q}^i, S^i \in \mathfrak{g}^{-i}$, possibly $\tilde{Q}^i = 0$ (if i is not a finite exponent). It is not hard to see that by equating each term of the above equation in each component \mathfrak{g}^{-j} , we find that $\tilde{Q}^j + [I, S^{j+1}]$ can be expressed in terms of $\tilde{Q}^0, \dots, \tilde{Q}^{j-1}$ and S^1, \dots, S^j and Q^0, \dots, Q^{j-1} . For example, the first equation, in \mathfrak{g}^0 and \mathfrak{g}^{-1} , read

$$\begin{aligned} Q^0 &= \tilde{Q}^0 + [S^1, I], \\ Q^1 &= \tilde{Q}^1 + [S^2, I] + [S^1, \tilde{Q}^0] + \frac{1}{2}[S^1, [S^1, I]] - \partial S^1. \end{aligned}$$

Now since $\mathfrak{g}^{-j} = \mathcal{V}^j \oplus [I, \mathfrak{g}^{-j-1}]$, then \tilde{Q}^j and S^{j+1} are unique and can be found recursively as differential polynomials in Q^0, \dots, Q^j . \square

To sum up, the choice of a gauge $\mathcal{V} \subset \mathfrak{b}$ such that $\mathfrak{b} = \mathcal{V} \oplus [I, \mathfrak{n}]$ entirely determines the gauge class of a connection $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ through unique r canonical coordinates \tilde{Q} , associated to \mathcal{V} , indexed by the finite exponents of \mathfrak{g} .

Example 4.3.5 (with $\mathfrak{g} = \mathfrak{sl}(2)$). We treat the $\mathfrak{sl}(2)$ case; we recall the classical Weyl generators of this algebra:

$$E = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad F = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

The root space gradation assigns $\deg E = -\deg F = 1$ and $\deg H = 0$. The Cartan, negative Borel and principal nilpotent subalgebras are given by

$$\mathfrak{n} = \mathbb{C}F, \quad \mathfrak{h} = \mathbb{C}H, \quad \mathfrak{b} = \mathfrak{h} \oplus \mathfrak{n} = \mathbb{C}H \oplus \mathbb{C}F.$$

The principal nilpotent element I is simply $I = E$. Since $\mathfrak{n} = \mathbb{C}F$ and $[E, F] = H$, it follows that $\mathfrak{b} = \mathbb{C}F \oplus [I, \mathfrak{n}]$. Hence, we naturally choose the gauge $\mathcal{V} = \mathbb{C}F = \mathfrak{n}$.

We start a connection $\mathcal{L} = \partial + \Lambda + Q$, with $Q \in \mathfrak{b}[[x]]$ and we want to represent it in the form $\tilde{\mathcal{L}} = e^{\text{ads}} \mathcal{L} = \partial + \Lambda + \tilde{Q}$ with $\tilde{Q}, S \in \mathfrak{n}[[x]]$ (since $\mathcal{V} = \mathfrak{n}$). We decompose the functions Q, \tilde{Q} and S through the root space gradation:

$$Q = H \otimes q^0 + F \otimes q^1, \quad \tilde{Q} = F \otimes \tilde{q}^1, \quad S = F \otimes s^1,$$

where $q^0, q^1, \tilde{q}^1, s^1 \in \mathbb{C}[[x]]$. (In the proof of Proposition 4.3.4, $H \otimes q^0$ is what we denoted Q^0 , etc.) Since $I = E$, the two first equations read

$$\begin{aligned} H \otimes q^0 &= [F \otimes s^1, E], \\ F \otimes q^1 &= F \otimes \tilde{q}^1 + \frac{1}{2} [F \otimes q^0, [F \otimes q^0, E]] - F \otimes s_{(1)}^1, \end{aligned}$$

where, as before, $s_{(1)}^1 = \partial_x(s^1)$. The first equation yields $s^1 = -q^0$, so that the second equation becomes $F \otimes q^1 = F \otimes (\tilde{q}^1 - (q^0)^2 + q_{(1)}^0)$, which yields

$$\tilde{q}^1 = q^1 + (q^0)^2 - q_{(1)}^0.$$

In the end, we obtain

$$\mathcal{L} = \partial + \begin{pmatrix} q^0 & \lambda + q^1 \\ 1 & -q^0 \end{pmatrix} = e^{-\text{ads}} \left(\partial + \begin{pmatrix} 0 & \lambda + \tilde{q}^1 \\ 1 & 0 \end{pmatrix} \right), \quad s^1 = -q^0, \quad \tilde{q}^1 = q^1 + (q^0)^2 - q_{(1)}^0.$$

Hamiltonian expansion

Let \mathcal{X} be any subspace of \mathfrak{g} and $\mathcal{X} = \bigoplus_{k \in \mathbb{Z}} \mathcal{X}^k$, with $\mathcal{X}^k \subset \mathfrak{g}^k$, be its decomposition through the principal gradation of \mathfrak{g} . We define

$$\widehat{\mathcal{X}} := \mathcal{X}^+ \oplus \widehat{\mathcal{X}}^-, \quad \mathcal{X}^+ = \bigoplus_{k \geq 0} \mathcal{X}^k, \quad \widehat{\mathcal{X}}^- = \prod_{k < 0} \mathcal{X}^k.$$

In particular, $\mathcal{X}^+ = \mathcal{X}^{\geq 0}$. We define \mathcal{X}_+ and $\widehat{\mathcal{X}}_-$ similarly for the homogeneous gradation. Elements of $\widehat{\mathcal{X}}$ consists of formal power series $\xi = \sum_{k \in \mathbb{Z}} \xi^k$, with $\xi^k \in \mathcal{X}^k$, with finitely many nonzero terms of positive degree. We define ξ^+ as the projection of ξ on \mathcal{X}^+ and $\xi^- = \xi - \xi^+$; we define ξ_+ and ξ_- similarly.

Remark 4.3.6. Importantly, $\mathcal{X}^+ = \mathcal{X}^{\geq 0}$, but $\widehat{\mathcal{X}}^- \neq \mathcal{X}^{< 0}$, since $\mathcal{X}^{< 0}$ is a direct sum, while $\widehat{\mathcal{X}}^-$ is a direct product. For example,

$$\widehat{\mathfrak{g}} = \mathring{\mathfrak{g}}((\lambda^{-1})) \oplus \mathbb{C}c,$$

i.e. a central extension of \mathfrak{g} -valued formal Laurent series in λ . The Lie bracket on \mathfrak{g} extends to $\widehat{\mathfrak{g}}$; similarly for the action of G on $\mathfrak{g}[[x]]$.

We recall that $\mathfrak{g} = \mathfrak{s} + \text{Im}(\text{ad}_\Lambda)$ and $\mathfrak{s} \cap \text{Im}(\text{ad}_\Lambda) = \mathbb{C}c$, where $\mathfrak{s} = \text{Ker}(\text{ad}_\Lambda)$ is the principal abelian subalgebra. These spaces decompose through the principal gradation:

$$\mathfrak{s} = \bigoplus_{k \in \mathbb{Z}} \mathfrak{s}^k, \quad \text{Im}(\text{ad}_\Lambda) = \bigoplus_{k \in \mathbb{Z}} \text{Im}(\text{ad}_\Lambda)^k,$$

where $\mathfrak{s}^k = \mathfrak{s} \cap \mathfrak{g}^k$ and $\text{Im}(\text{ad}_\Lambda)^k = \text{Im}(\text{ad}_\Lambda) \cap \mathfrak{g}^k$. They decompose through the homogeneous gradation as well.

Now, in comparison to the the group N of elements of the form e^{ad_S} with $S \in \mathfrak{n}[[x]]$, we want to define an action on $\mathfrak{g}[[x]]$ of elements of the form e^{ad_U} , with $U \in \widehat{\mathfrak{g}}^-[[x]]$. It is not obvious that one can define such an endomorphism e^{ad_U} , however, since U is assumed to be of negative degree, it is not hard to see that for any $A \in \mathfrak{g}[[x]]$, the expression $e^{\text{ad}_U} A$ contains finitely many terms for each degree w.r.t. the principal gradation. Hence, the action is well defined. The following theorem, which plays a major role in the building of the Drinfeld–Sokolov hierarchies, first appeared in [DS84] in its version based on $\bar{\mathfrak{g}}$; the following version is due to C.-Z. Wu.

Theorem 4.3.7 ([Wu17]). *Let $\mathcal{L} = \partial + \Lambda + Q \in \text{op}_{\mathfrak{g}}$ be a connection. There exists a unique function $U \in \widehat{\mathfrak{g}}^-[[x, t]]$ such that the following holds true.⁶*

(i) *The connection $\mathcal{L}_0 = e^{-\text{ad}_U} \mathcal{L}$ has the form⁷*

$$\mathcal{L}_0 = \partial + \Lambda + H, \quad H \in \widehat{\mathfrak{s}}^-[[x]] \quad (4.39)$$

(notice that $\mathcal{L}_0 \notin \text{op}_{\mathfrak{g}}$). *The coefficients $U^i, H^i \in \mathfrak{g}^{-i}[[x]]$ are differential polynomials in the component of Q .*

(ii) *For every $\xi \in \mathfrak{s}^+$, the following relation holds,*

$$(e^{\text{ad}_U} \xi)_c = 0. \quad (4.40)$$

Here, for $\eta \in \mathfrak{g}$, the complex number η_c denotes the component of η along the center $\mathbb{C}c$ in the decomposition $\mathfrak{g}^0 = \mathbb{C}\check{\alpha}_1 \oplus \cdots \oplus \mathbb{C}\check{\alpha}_r \oplus \mathbb{C}c$.

Vocabulary. The connection $\mathcal{L}_0 = e^{-\text{ad}_U} \mathcal{L}$ we call a *Hamiltonian expansion* of \mathcal{L} , for reasons that will become clear later (cf. Proposition 4.3.21).

⁶Note that in [DS84] the function U is not unique.

⁷We chose the convention $\mathcal{L}_0 = e^{-\text{ad}_U} \mathcal{L}$ instead of $\mathcal{L}_0 = e^{\text{ad}_U} \mathcal{L}$ (as is done in [DS84] for the latter) so that the Drinfeld–Sokolov equations take the form (4.43).

Proof. We first prove point (i); we proceed in the same fashion as the proof of Proposition 4.3.4. We first need to assume that

$$U \in \widehat{\mathfrak{g}}^- [[x]], \quad H \in (\mathfrak{g}^0 \oplus \widehat{\mathfrak{g}}^-) [[x]].$$

Equating $\mathcal{L}_0 = e^{-\text{ad}_U} \mathcal{L}$ amounts to

$$\Lambda + Q = e^{\text{ad}_U} (\Lambda + H) - \nabla_U U.$$

Denoting the components $H^i, U^i \in \mathfrak{g}^{-i} [[x]]$, one finds that $H^i + [U^{i+1}, \Lambda]$ can be expressed in terms of H^0, \dots, H^{i-1} and U^1, \dots, U^i . We then obtain $H^0 \in \mathfrak{g}^0 \in \mathbb{C} [[x]]$ and, for $i > 0$,

$$H^i \in \mathfrak{s}^{-i} [[x]], \quad [U^{i+1}, \Lambda] \in \text{Im}(\text{ad}_\Lambda)^{-i} [[x]]$$

But since $\widehat{\mathfrak{g}}^{-i} = \mathfrak{s}^{-i} + \text{Im}(\text{ad}_\Lambda)^{-i}$, then H^i and $[U^{i+1}, \Lambda]$ can be recursively determined and are differential polynomials in the coordinates of Q . Finally, $\mathfrak{s}^0 = \{0\}$ so that $H^0 = 0$ and $H \in \widehat{\mathfrak{s}}^- [[x]]$.

We now prove point (ii). Since only $[U^i, \Lambda]$ is determined, two cases must be distinguished. Either i is an affine exponent, in which case U^i is entirely determined; or i is not an exponent, in which case U^i is determined up to the addition of a multiple of Λ_{-i} . So let i be a positive exponent of \mathfrak{g} . Since $c \in \mathfrak{g}^0$, then for $\xi = \Lambda_i$, Equation (4.40) reads

$$0 = (e^{\text{ad}_U} \Lambda_i)_c = [U^i, \Lambda_i]_c + \frac{1}{2}[U^1, [U^{i-1}, \Lambda_i]]_c + \dots \quad (4.41)$$

where the remaining terms involve only the functions U^1, \dots, U^{i-1} . Let us assume that U^1, \dots, U^{i-1} are uniquely determined; choose U^i and consider the transformation $U^i \mapsto U^i + \mu_i \Lambda_{-i}$ (which does not affect Equation (4.56)). Since $[\Lambda_{-i}, \Lambda_i] = ic$, then Equation (4.41) gives

$$i\mu_i = -[U^i, \Lambda_i]_c - \frac{1}{2}[U^1, [U^{i-1}, \Lambda_i]]_c - \dots$$

which determines μ_i . Therefore, Equation (4.41) entirely determines the function U . Moreover, since the Λ_j 's form a basis of \mathfrak{s} , then Equation (4.40) is satisfied for any $\xi \in \mathfrak{s}^+$. \square

The next proposition tells us what happens to the functions U and H of Theorem 4.3.7 when the connection \mathcal{L} undergoes a gauge transformation.

Proposition 4.3.8 (Wu, 2016). *The function H of Theorem 4.3.7 is invariant with respect to gauge transformation. More precisely, let $U \in \widehat{\mathfrak{g}}^- [[x, t]]$ be the unique function such that $(e^{\text{ad}_U} \xi)_c = 0$ for any $\xi \in \mathfrak{s}^+$ and*

$$e^{-\text{ad}_U} \mathcal{L} = \partial + \Lambda + H,$$

where $H \in \widehat{\mathfrak{s}}^- [[x, t]]$. Then for any function $S \in \mathfrak{n} [[x]]$, there exists a unique function $\tilde{U} \in \widehat{\mathfrak{g}}^- [[x, t]]$ such that $(e^{\text{ad}_{\tilde{U}}} \xi)_c = 0$ for any $\xi \in \mathfrak{s}^+$ and

$$e^{-\text{ad}_{\tilde{U}}} \tilde{\mathcal{L}} = \partial + \Lambda + H. \quad (4.42)$$

Proof. We prove that the function $\tilde{U} \in \widehat{\mathfrak{g}}[[x, t]]$ such that $e^{\text{ad}_{\tilde{U}}} = e^{\text{ad}_S} e^{\text{ad}_U}$ satisfies the conditions of the proposition. First, by the homomorphism property of the BCH formula, and since $\mathfrak{n} \subset \mathfrak{g}^{<0}$, it follows that $\tilde{U} \in \widehat{\mathfrak{g}}^-[[x, t]]$. Second,

$$e^{\text{ad}_{\tilde{U}}} \tilde{\mathcal{L}} = e^{-\text{ad}_U} \mathcal{L} = \partial + \Lambda + H.$$

Third, recall that the bracket on \mathfrak{g} is given by

$$[X\lambda^p + ac, Y\lambda^q + bc] = [X, Y]\lambda^{p+q} + q\delta_{p+q,0}(X, Y)_0 \cdot c,$$

for $X, Y \in \mathfrak{g}$, $p, q \in \mathbb{Z}$ and $a, b \in \mathbb{C}$. So for $S \in \mathfrak{n}[[x]]$, we have $(\text{ad}_S(\mathfrak{g}))_c = 0$ and also $\text{ad}_S c = 0$. Therefore, ad_S commutes with the projection onto $\mathbb{C}c$. Consequently, for any $\xi \in \mathfrak{s}^+$,

$$(e^{\text{ad}_{\tilde{U}}} \xi)_c = e^{\text{ad}_S}(e^{\text{ad}_U} \xi)_c = 0.$$

Therefore, the function \tilde{U} satisfies the condition. \square

The matrix Lax pairs of the Drinfeld–Sokolov equations

From now on we will assume that our functions Q, U, H , etc. depend not only on x but also on an “evolution variable” t (or even multiple variables, depending on the dynamics considered). In other words, hereafter, Q, U, H , etc. $\in \widehat{\mathfrak{g}}[[x, t]]$ (or otherwise specified). That clearly does not affect the previous results and will allow us to produce a dynamics.

Let $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ be a connection and let $U \in \widehat{\mathfrak{g}}^-[[x, t]]$ be a function as in Theorem 4.3.7. For $\xi \in \mathfrak{s}^+$, we consider the following equations:⁸

$$\frac{\partial \mathcal{L}}{\partial t} = [(e^{\text{ad}_U} \xi)_+, \mathcal{L}], \quad (4.43)$$

$$\frac{\partial \mathcal{L}}{\partial t} = [(e^{\text{ad}_U} \xi)^+, \mathcal{L}]. \quad (4.44)$$

Notice that since Λ is constant, then $\partial_t \mathcal{L} = \partial_t Q$. Before proving that the above equations are consistent (well defined) or integrable, we give two important propositions. Proofs of these propositions can be found in [DS84] (Proposition 3.7).

Proposition 4.3.9 (Gauge symmetry). *Let $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ be a connection and let $\xi \in \mathfrak{s}^+$. Assume that \mathcal{L} satisfies Equation (4.43) with flow ∂_t . Then for any function $S \in \mathfrak{n}[[x, t]]$, the operator $\tilde{\mathcal{L}} = e^{\text{ad}_S} \mathcal{L}$ also satisfies Equation (4.43), i.e.*

$$\frac{\partial \tilde{\mathcal{L}}}{\partial t} = [(e^{\text{ad}_{\tilde{U}}} \xi)_+, \tilde{\mathcal{L}}],$$

where $e^{\text{ad}_{\tilde{U}}} = e^{\text{ad}_S} e^{\text{ad}_U}$.

Proposition 4.3.10. *If \mathcal{L} is a solution of Equation (4.43), then there exists a gauge transformation $\tilde{\mathcal{L}} = e^{\text{ad}_S} \mathcal{L}$, for some $S \in \mathfrak{n}[[x, t]]$, such that $\tilde{\mathcal{L}}$ is a solution of Equation (4.44).*

⁸The fact that we denote the flows of Equations (4.43) and (4.44) both by ∂_t is an anticipation of Proposition 4.3.10; the latter justifies that the notation is consistent.

Remark 4.3.11. An important consequence of Proposition 4.3.10 is that both Equations (4.43) and (4.44) are invariant under the transformation $\xi \mapsto \xi + \eta$ where $\eta \in \mathfrak{s}^-$. This is obvious when considering Equation (4.44), and it holds true for Equation (4.43) by Proposition 4.3.10. Therefore, we can assume that $\xi \in \mathfrak{s}^+$ without loss of generality.

With the last two propositions, we can consider Equation (4.43) modulo gauge transformations. We call the *Drinfeld–Sokolov equation of type \mathfrak{g}* (or \mathfrak{g} -DS equation) associated to $\xi \in \mathfrak{s}^+$ the Equation (4.43) modulo gauge transformation of \mathcal{L} . In analogy to the KP hierarchy (3.18), in order to define the Drinfeld–Sokolov hierarchies, we need to introduce infinitely many time variables, one for each affine exponent of \mathfrak{g} :

$$\mathbf{t} := \{t_j \mid j \in \mathbb{E}^+\}.$$

Definition 4.3.12 (Drinfeld–Sokolov hierarchy). Let $Q \in \mathfrak{b}[[\mathbf{t}]]$ be a function and a connection $\mathcal{L} = \partial + \Lambda + Q \in \text{op}_{\mathfrak{g}}$. We call the Drinfeld–Sokolov hierarchy of type \mathfrak{g} (or simply the \mathfrak{g} -DS hierarchy) the family of equations

$$\left\{ \partial_{t_j} \mathcal{L} = [(e^{\text{ad}_U} \Lambda_j)_+, \mathcal{L}] \mid j \in \mathbb{E}^+ \right\} \quad (4.45)$$

modulo gauge transformation of \mathcal{L} . We call $\partial_{t_j} \mathcal{L} = [(e^{\text{ad}_U} \Lambda_j)_+, \mathcal{L}]$ the j -th \mathfrak{g} -DS equation.

We make a few remarks.

Remark 4.3.13. The function $Q = 0$ is always a solution. Indeed, then the connection $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ has the form $\mathcal{L} = \partial + \Lambda$, and is already in the form $\mathcal{L}_0 = \partial + \Lambda + H$, with $H = 0 \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]]$. The function $U = 0$ also does the job of Theorem 4.3.7. The j -th \mathfrak{g} -DS equation then becomes $\partial_{t_j} \mathcal{L} = [\Lambda_j, \partial + \Lambda]$. Since $Q = 0$, then $\partial_{t_j} \mathcal{L} = 0$, and by constancy of Λ_j and commutativity of \mathfrak{s}^+ , it follows that $[\Lambda_j, \partial + \Lambda] = 0$ as well. Hence $Q = 0$ is a solution.

Remark 4.3.14. Similarly to the KP hierarchy (3.18), the first equation of the \mathfrak{g} -DS hierarchy, i.e. for $\xi = \Lambda_1 = \nu \Lambda$ (with associated flow ∂_{t_1}) simply reads

$$\frac{\partial Q}{\partial t_1} = \nu \frac{\partial Q}{\partial x} \quad (4.46)$$

Indeed, considering Equation (4.44), we clearly have $(e^{\text{ad}_U} \Lambda_1)^+ = \Lambda_1 + [U^1, \Lambda_1]$, while $[U^1, \Lambda] = Q$ (by construction of U). Therefore $[\mathcal{L}, (e^{\text{ad}_U} \Lambda_1)^+] = \nu[\partial + Q + \Lambda, Q + \Lambda] = \nu \partial_x Q$.

Remark 4.3.15. In most cases, for a simple Lie algebra \mathfrak{g} , the subspace \mathfrak{s}^j of principal degree j is one-dimensional. However, it might not be the case (e.g. in the D_n case with n even, cf. Section A.1, or for a nonsimple Lie algebra). If so, then we may distinguish several flows: t_j, t'_j, t''_j , etc., one for each generator of \mathfrak{s}^j .

Equivalence between Drinfeld–Sokolov and Wu’s constructions

Before continuing, let us have a few words on the equivalence between Drinfeld–Sokolov’s construction, which is based on the loop algebra $\bar{\mathfrak{g}} = \mathfrak{g}[\lambda, \lambda^{-1}]$, and the one of Wu, described above, which is based on the (derived) affine algebra $\mathfrak{g} = \bar{\mathfrak{g}} \oplus \mathbb{C}c$. In [DS84], the authors give the following theorem (in comparison to Wu’s theorem 4.3.7):

Theorem 4.3.16 ([DS84]). *Let $\mathcal{L} = \partial + \Lambda + Q \in \text{op}_{\widehat{\mathfrak{g}}}$ be a connection. There exists a non unique function $\overline{U} \in \widehat{\mathfrak{g}}^-[[x, t]]$ such that the connection $\mathcal{L}_0 = e^{-\text{ad}_{\overline{U}}}\mathcal{L}$ has the form*

$$\mathcal{L}_0 = \partial + \Lambda + \overline{H}, \quad \overline{H} \in \widehat{\mathfrak{s}}^-[[x]] \quad (4.47)$$

The coefficients $\overline{U}^i, \overline{H}^i \in \widehat{\mathfrak{g}}^{-i}[[x]]$ are differential polynomials in the component of Q . Moreover, the function \overline{U} is unique if we impose $\overline{U} \in (\widehat{\mathfrak{s}}^\perp)^-$, where the orthogonal is taken with respect to the standard bilinear form on \mathfrak{g} .

Then Drinfeld and Sokolov define the following equations:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial t} &= [(e^{\text{ad}_{\overline{U}}}\xi)_+, \mathcal{L}], \\ \frac{\partial \mathcal{L}}{\partial t} &= [(e^{\text{ad}_{\overline{U}}}\xi)^+, \mathcal{L}]. \end{aligned}$$

We prove that the two constructions define the exact same equations.

Proof. First of all, the function $U \in \widehat{\mathfrak{g}}^-[[x, t]]$ of Theorem 4.3.7 also belongs to $\widehat{\mathfrak{g}}^-[[x, t]]$, as a set, since $c \notin \mathfrak{g}^-$. Then since $(e^{\text{ad}_U}\xi)_c = 0$ for all $\xi \in \mathfrak{s}^+$, then in particular, $(e^{\text{ad}_U}\Lambda)_c = 0$; moreover, $(e^{\text{ad}_U}Q)_c = 0$ since $Q \in \mathfrak{g}_0$. It follows that the action of e^{ad_U} on \mathcal{L} within \mathfrak{g} is identical to its action within the quotient $\overline{\mathfrak{g}} = \mathfrak{g}/\mathbb{C}c$. Similarly, since as a set $\overline{\mathfrak{s}}^- = \mathfrak{s}^-$, it follows that the function $H \in \widehat{\mathfrak{s}}^-[[x, t]]$ also belongs to $\widehat{\mathfrak{s}}^-[[x]]$ as a set. Therefore, the function U is also a functions as in Theorem 4.3.22. Finally, since $(e^{\text{ad}_U}\xi)_c = 0$, it follows that the two constructions define the same equations. \square

Consistency of the Drinfeld–Sokolov equations

The first thing we need to prove is that the DS equations are consistent (well defined). Since $\mathcal{L} = \partial + \Lambda + Q$, it follows that $\partial_t \mathcal{L} = \partial_t Q \in \mathfrak{b}[[x, t]]$. Hence it is necessary, for Equation (4.43) to be consistent, that the following proposition holds.

Proposition 4.3.17 (Consistency). *Let $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ be a connection and let $\xi \in \mathfrak{s}^+$. Then*

$$[\mathcal{L}, (e^{\text{ad}_U}\xi)_+] \in \mathfrak{b}[[x, t]].$$

Proof. Here we adapt the proof given in [DS84] (§6.2 and Proposition 3.7)—where the construction is on the loop $\overline{\mathfrak{g}} \simeq \mathfrak{g}/\mathbb{C}c$ —to the construction of [Wu17]—based on the affine algebra \mathfrak{g} . We will proceed in several steps. First, let us prove that⁹

$$[\mathcal{L}, e^{\text{ad}_U}\widehat{\mathfrak{s}}] = \mathbb{C}[[x, t]]. \quad (4.48)$$

Let $M = e^{\text{ad}_U}\xi$, then $[\mathcal{L}, M] = e^{\text{ad}_U}[\mathcal{L}_0, \xi]$, and since ξ is constant, $[\mathcal{L}_0, \xi] = [\Lambda + H, \xi]$. The latter belongs to $\mathbb{C}[[x, t]]$ since all terms involved belong to the Heisenberg subalgebra. Since

⁹Compare (4.48) with Lemma 6.3 of [DS84], which states that $[\mathcal{L}, \widehat{\mathfrak{s}}] = \{0\}$.

c is central, it follows that $[\mathcal{L}, M] \in \mathbb{C}c[[x, t]]$ as well. We easily see that by adjusting ξ , we recover all elements of the center.

Next, we prove that

$$[\mathcal{L}, M_+] \in (\mathfrak{g}_0 \cap \mathfrak{g}^{\leq 0})[[x, t]]. \quad (4.49)$$

(Recall that \mathfrak{b} consists of elements of $\mathfrak{g}_0 \cap \mathfrak{g}^{\leq 0}$ that have no component in c .) By Equation (4.48), there exists $A \in \mathbb{C}[[x, t]]$ such that $[\mathcal{L}, M] = A \cdot c$, so that

$$[\mathcal{L}, M_+] = [M_-, \mathcal{L}] + A \cdot c.$$

To show Equation (4.49), we first prove that

$$[M_-, \mathcal{L}] \in (\mathfrak{g}^0 \oplus \widehat{\mathfrak{g}}^-)[[x, t]]. \quad (4.50)$$

On the one hand, $M_- \in \widehat{\mathfrak{g}}^-[[x, t]]$ and $\widehat{\mathfrak{g}}^- \subset \widehat{\mathfrak{g}}^-$, so that $\partial M_- \in \widehat{\mathfrak{g}}^-[[x, t]]$. On the other hand, $\Lambda \in \mathfrak{g}^1$ and $Q \in \mathfrak{b}[[x, t]]$, so $[\Lambda + Q, M_-] \in (\mathfrak{g}^0 \oplus \widehat{\mathfrak{g}}^-)[[x, t]]$. Hence Equation (4.50). With similar arguments, we can see that $[\mathcal{L}, M_+] \in \mathfrak{g}_+[[x, t]]$ and $[\mathcal{L}, M_-] \in (\mathfrak{g}_0 \oplus \widehat{\mathfrak{g}}^-)[[x, t]]$, so that

$$[\mathcal{L}, M_-] \in \mathfrak{g}_0[[x, t]]. \quad (4.51)$$

Equations (4.50) and (4.51) together prove (4.49).

From Equation (4.49), it is clear that what remains is to show that $[\mathcal{L}, M_+]_c = 0$, i.e.

$$(\partial M_+)_c + [\Lambda + Q, M_+]_c = 0. \quad (4.52)$$

By Theorem 4.3.7, $M_c = 0$, so $(\partial M_+)_c = 0$ as well. for the remaining part, let us recall the formula for the bracket in \mathfrak{g} : for any $X, Y \in \mathfrak{g}$, $p, q \in \mathbb{Z}$ and $a, b \in \mathbb{C}$,

$$[X\lambda^p + ac, Y\lambda^q + bc] = [X, Y]\lambda^{p+q} + q\delta_{p+q,0}(X, Y)_0 \cdot c.$$

Clearly, the bracket produces a nonzero component in c only if $p > 0$ and $q < 0$ (or the reverse). In other words, for any $\eta, \zeta \in \mathfrak{g}$, we have $[\eta, \zeta]_c \neq 0$ only if $\eta \in \mathfrak{g}_{>0}$ and $\zeta \in \mathfrak{g}_{<0}$ (or the reverse). Yet $\Lambda = E_{-\theta}\lambda + I \in \mathfrak{g}_1 \oplus \mathfrak{g}_0$ and $Q \in \mathfrak{g}_0[[x, t]]$, while $M_+ \in \mathfrak{g}_+[[x, t]]$. So $[\Lambda + Q, M_+]_c = 0$, which concludes the proof of Proposition 4.3.17. \square

4.3.2 Integrability of the DS hierarchies

Commutativity of the flows

In this section, we prove that the flows defined by the DS equations are integrable in the sense of the following theorem.

Theorem 4.3.18 (Drinfeld–Sokolov [DS84]). *Let $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ be a connection. Let $\xi, \eta \in \mathfrak{s}^+$ and consider the associated DS equations, namely,*

$$\frac{\partial \mathcal{L}}{\partial t} = [(e^{\text{ad}_U} \xi)_+, \mathcal{L}], \quad \frac{\partial \mathcal{L}}{\partial s} = [(e^{\text{ad}_U} \eta)_+, \mathcal{L}],$$

where $U \in \widehat{\mathfrak{g}}^-[[x, t]]$ is a function as in Theorem 4.3.7. Then the flows defined by the above equations commute pairwise, that is,

$$\frac{\partial^2 \mathcal{L}}{\partial t \partial s} = \frac{\partial^2 \mathcal{L}}{\partial s \partial t}.$$

We will proceed by a series of lemmas, as is done in [DS84]. Let us introduce the following vector space,

$$\mathcal{S}_{\mathcal{L}} = \{A \in \widehat{\mathfrak{g}}[[x, t]] \mid A_c = 0, \partial_t \mathcal{L} = [A, \mathcal{L}]\}.$$

Lemma 4.3.19. *Let $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ be a connection and let $U \in \widehat{\mathfrak{g}}^-$ be a function as in Theorem 4.3.7, then*

$$\mathcal{S}_{\mathcal{L}} \subset e^{\text{ad}_U} \widehat{\mathfrak{s}}[[x, t]].$$

Proof. Let $A \in \widehat{\mathfrak{g}}[[x, t]]$ be a function such that $\partial_t \mathcal{L}_0 = [A, \mathcal{L}_0]$, which is equivalent to

$$\frac{\partial H}{\partial t} + \frac{\partial A}{\partial x} = [A, \Lambda] + [A, H].$$

Let us decompose A and H through the principal gradation: $A = \sum_{i \geq m} A^i$ (for some $m \in \mathbb{Z}$) and $H = \sum_{i \geq 1} H^i$, with $A^i, H^i \in \mathfrak{g}^{-i}[[x, t]]$. Then since H contains only negative terms, we get the following system of equations:

$$\frac{\partial H^i}{\partial t} + \frac{\partial A^i}{\partial x} = [A^{i+1}, \Lambda] + \sum_{j=m}^{i-1} [A^j, H^{i-j}] \in \mathfrak{g}^{-i}.$$

The first equation, within \mathfrak{g}^{-m} , reads

$$\frac{\partial H^{m-1}}{\partial t} = [A^m, \Lambda].$$

possibly $H^{m-1} = 0$ if $m \leq 1$ or is not an exponent. If $m \leq 1$ or is not an exponent, then the equation simply becomes $[\Lambda, A^m] = 0$. If $m > 1$ and is an exponent, the above implies that $[A^m, \Lambda] \in (\mathfrak{s} \cap \text{Im}(\text{ad}_{\Lambda}))^-[[x, t]]$. Yet $(\mathfrak{s} \cap \text{Im}(\text{ad}_{\Lambda}))^- = \{0\}$ (since the center $\mathbb{C}c \in \mathfrak{g}^0$), so that $[A^m, \Lambda] = 0$. In both cases, $A^m \in \mathfrak{s}^{-m}[[x, t]]$. Similarly, if $A^m, A^{m+1}, \dots, A^i \in \mathfrak{s}[[x, t]]$, then the equation in \mathfrak{g}^{-i} becomes

$$\frac{\partial H^i}{\partial t} + \frac{\partial A^i}{\partial x} = [A^{i+1}, \Lambda],$$

If $i > 0$ and is an exponent, the above implies that $[A^{i+1}, \Lambda] = 0$. If $i \leq 0$ or is not an exponent, then the equation becomes $\partial_x A^i = [A^{i+1}, \Lambda]$, implying that $[A^{i+1}, \Lambda] \in (\mathfrak{s} \cap \text{Im}(\text{ad}_{\Lambda}))^+[[x, t]]$. Now, $(\mathfrak{s} \cap \text{Im}(\text{ad}_{\Lambda}))^+ = \mathbb{C}c$, yet $A_c = 0$, so that that $[A^{i+1}, \Lambda] = 0$. In both case, $A^{i+1} \in \mathfrak{s}^{-i-1}[[x, t]]$. By induction, $A \in \widehat{\mathfrak{s}}[[x, t]]$. \square

Lemma 4.3.20. *Let $\xi \in \mathfrak{s}^+$, let $U \in \widehat{\mathfrak{g}}^-[[x, t]]$ be a function as in Theorem 4.3.7 and let us denote $M = e^{\text{ad}_U} \xi$. For any $A \in \mathcal{S}_{\mathcal{L}}$ the following equation holds,*

$$\frac{\partial M}{\partial t} = [A, M].$$

Proof. It is given that $\partial_t \mathcal{L} = [A, \mathcal{L}]$, or equivalently, $[\partial_t - A, \mathcal{L}] = 0$. Since $\mathcal{L}_0 = e^{-\text{ad}_U} \mathcal{L}$, by applying $e^{-\text{ad}_U}$ to the latter, we find

$$[\partial_t - \tilde{A}, \mathcal{L}_0] = 0, \quad \tilde{A} = e^{-\text{ad}_U} A - \nabla_U U, \quad (4.53)$$

where we have used Equation (4.37) for the expression of \tilde{A} . In other words, $\tilde{A} \in \mathcal{S}_{\mathcal{L}_0}$. By Lemma 4.3.19 it follows that $\tilde{A} \in \widehat{\mathfrak{s}}[[x, t]]$.

On the other hand, we want to prove that $[\partial_t - A, M] = 0$. Since $M = e^{\text{ad}_U} \xi$, by applying $e^{-\text{ad}_U}$ to the latter, we find

$$[\partial_t - \tilde{A}, \xi] = 0.$$

Yet $\xi \in \mathfrak{s}^+$ and, in particular, is constant so that it reduces to $[\tilde{A}, \xi] = 0$. Now, by Remark 4.3.11, we can assume that $\tilde{A} \in \mathfrak{s}^+[[x, t]]$, so that $[\tilde{A}, \xi] = 0$ holds true by commutativity of \mathfrak{s}^+ .¹⁰ \square

We can now prove that the flows of the Drinfeld–Sokolov hierarchies commute pairwise, that is, are integrable. Still, the proof first appeared in [DS84].

Proof of Theorem 4.3.18. Let us denote $M = e^{\text{ad}_U} \xi$ and $N = e^{\text{ad}_U} \eta$. Then by definition,

$$\begin{aligned} \partial_s \partial_t \mathcal{L} &= \partial_s [M_+, \mathcal{L}] \\ &= [\partial_s M_+, \mathcal{L}] + [M_+, [N_+, \mathcal{L}]]. \end{aligned}$$

By Lemma 4.3.20, we know that $\partial_s M = [N_+, M]$, so that

$$\begin{aligned} \partial_s \partial_t \mathcal{L} &= [[N_+, M]_+, \mathcal{L}] + [M_+, [N_+, \mathcal{L}]] \\ &= [[N_+, M]_+, \mathcal{L}] - [N_+, [\mathcal{L}, M_+]] - [\mathcal{L}, [M_+, N_+]], \end{aligned} \quad (4.54)$$

where we have used Jacobi's identity to obtain the second equation. We can then use the antisymmetry of the bracket to rewrite the above as

$$\partial_s \partial_t \mathcal{L} = [[N_+, M]_+, \mathcal{L}] + [N_+, [M_+, \mathcal{L}]] + [[M_+, N_+], \mathcal{L}].$$

In the same way we obtained Equation (4.54), one finds that

$$\partial_t \partial_s \mathcal{L} = [[M_+, N]_+, \mathcal{L}] + [N_+, [M_+, \mathcal{L}]].$$

¹⁰Notice that since $[\tilde{A}, \xi] = 0$, we also have $[\partial_t + \tilde{A}, \xi] = 0$. So it seems as though we could have proved that $[\partial_t + A, M] = 0$ as well. Yet that is not true because the relation between A and \tilde{A} is precisely given by $\partial_t + \tilde{A} = e^{-\text{ad}_U} (\partial_t + A)$, so that \tilde{A} cannot be replaced by $-\tilde{A}$. The relation $[\partial_t + \tilde{A}, \xi] = 0$ holds true as well, but does not imply $[\partial_t + A, M] = 0$.

Subtracting the two equations, we obtain

$$\partial_s \partial_t \mathcal{L} - \partial_t \partial_s \mathcal{L} = [[N_+, M]_+ + [M_+, N_+] - [M_+, N]_+, \mathcal{L}]. \quad (4.55)$$

Yet since $[M, N] = e^{\text{ad}_U} [\xi, \eta] = 0$ (by commutativity of \mathfrak{s}^+), we have

$$\begin{aligned} [N_+, M]_+ &= [N - N_-, M]_+ \\ &= [N_-, M]_+ = [M_+, N_-]_+, \end{aligned}$$

where we have used the fact that $[M_-, N_-]_+ = 0$. Finally, we obtain that in Equation (4.55),

$$\begin{aligned} [N_+, M]_+ + [M_+, N_+] - [M_+, N]_+ &= [M_+, N_-]_+ + [M_+, N_+] - [M_+, N]_+ \\ &= -[M_+, N_+]_+ + [M_+, N_+] = 0. \end{aligned}$$

Consequently, the two derivatives commute. \square

Conserved densities

Similarly to the construction of local functionals in Section 3.1, we naturally define an analogue for our \mathfrak{g} -valued functions: we set

$$\overline{\mathfrak{g}[[x, t]]} := \mathfrak{g}[[x, t]] / (\text{Im}(\partial) \oplus \mathbb{C}1).$$

We call the elements of this quotient *local functionals* as well and alternatively denote

$$\bar{F} = \int F = \int F(x) dx \in \overline{\mathfrak{g}[[x, t]]}$$

the class of an element $F \in \mathfrak{g}[[x, t]]$. Again, if our functions depend on more variables than just x and t , the above definition adapts.

Proposition 4.3.21 (Conserved densities). *Let $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ and let $U \in \widehat{\mathfrak{g}}^-[[x, t]]$ and $H \in \widehat{\mathfrak{s}}^-[[x, t]]$ be functions as in Theorem 4.3.7. Let $H^i \in \mathfrak{s}^{-i}$ be such that $H = \sum_{i \geq 1} H^i$. Let $\xi \in \mathfrak{s}^+$ and assume that \mathcal{L} satisfies the associated DS equation (4.43):*

$$\frac{\partial \mathcal{L}}{\partial t} = [(e^{\text{ad}_U} \xi)_+, \mathcal{L}].$$

Then the H^i 's, $i \in \mathbb{E}^{<0}$, are conserved densities in the sense that

$$\frac{\partial}{\partial t} \int H^i = 0.$$

The above proposition justifies the denomination *Hamiltonian expansion* for the operator $\mathcal{L}_0 = e^{-\text{ad}_U} \mathcal{L}$, although we still need to prove that the conserved densities H^i commute with respect to a local Poisson bracket, see Section 5.1.

Proof. Let us denote $M = -e^{\text{ad}_U} \xi$ so that we start with $[\partial_t + M, \mathcal{L}] = 0$, i.e. $M \in \mathcal{S}_{\mathcal{L}}$. As in the proof of Lemma 4.3.20, it amounts to $\tilde{M} \in \mathcal{S}_{\mathcal{L}_0}$, where \tilde{M} is given by Equation (4.53) (the formula does not matter here). Yet since, by Lemma 4.3.20, $\mathcal{S}_{\mathcal{L}_0} \subset \widehat{\mathfrak{s}}[[x, t]]$, then

$$0 = [\partial_t + \tilde{M}, \mathcal{L}_0] = \frac{\partial H}{\partial t} - \frac{\partial \tilde{M}}{\partial x} - [\tilde{M}, \Lambda + H] = \frac{\partial H}{\partial t} - \frac{\partial \tilde{M}}{\partial x}.$$

Consequently, $\partial_t H = \partial_x \tilde{M}$, meaning that $\partial_t \bar{H} = 0$. \square

4.3.3 Example: From the A_1 -DS hierarchy to the KdV equation

In this section we show explicitly how to obtain the KdV equation from the first nontrivial equation of the DS hierarchy of type A_1 (i.e. $\mathfrak{sl}(2)$). However, this example is developed from the definition of the DS hierarchies given in the original paper [DS84], not the version given in [Wu17] (which is the one described in the previous section). We recall the following theorem (we drop the bar over symbols for this section specifically).

Theorem 4.3.22 ([DS84]). *Let $\mathcal{L} = \partial + \Lambda + Q \in \text{op}_{\mathfrak{g}}$. There exists a non unique function $U \in \widehat{\mathfrak{g}}^- [[x, t]]$ such that the connection $\mathcal{L}_0 = e^{-\text{ad}_U} \mathcal{L}$ has the form*

$$\mathcal{L}_0 = \partial + \Lambda + H, \quad H \in \widehat{\mathfrak{s}}^- [[x]] \quad (4.56)$$

The coefficients $U^i, H^i \in \mathfrak{g}^- [[x]]$ are differential polynomials in the component of Q . Moreover, the function U is unique if we impose $U \in (\widehat{\mathfrak{s}}^\perp)^- [[t]]$, where the orthogonal is taken with respect to the standard bilinear form on \mathfrak{g} .

Conveniently, we start with a connection $\mathcal{L} = \partial + \Lambda + Q \in \text{op}_{\mathfrak{g}}$ in its canonical form $\mathcal{L} = \widetilde{\mathcal{L}}$, i.e.

$$\mathcal{L} = \partial + \begin{pmatrix} 0 & \lambda + q \\ 1 & 0 \end{pmatrix}.$$

The principal Heisenberg subalgebra of $\widehat{\mathfrak{sl}}(2)$ is generated by the Λ^k 's with $k \in \mathbb{E}(\mathfrak{sl}(2)) = \mathbb{Z}^{\text{odd}}$. We are interested in the second equation of the hierarchy (the first describing the flow along t_1 as a translation in x), i.e for $\Lambda^3 = \lambda\Lambda$. More precisely, we want to show that the equation

$$\frac{\partial \mathcal{L}}{\partial t_3} = [(e^{\text{ad}_U} \Lambda^3)_+, \mathcal{L}], \quad \Lambda^3 = e_1 \lambda + f_1 \lambda^2 = \begin{pmatrix} 0 & \lambda^2 \\ \lambda & 0 \end{pmatrix},$$

reads nothing but the KdV equation on the function $q \in \mathbb{C} [[x, t_3]]$. The subscript $+$ denotes the projection onto the subspace $\mathfrak{g}_{\geq 0}$ of nonnegative homogeneous degree (i.e. with only nonnegative powers of λ).

Recall that, in the above equation, the $U \in (\mathfrak{s}^\perp)^- [[x, t_3]]$ is such that the connection defined by $\mathcal{L}_0 = e^{-\text{ad}_U} \mathcal{L}$ has the form

$$\mathcal{L}_0 = \partial + \Lambda + H, \quad H \in \widehat{\mathfrak{s}}^- [[x]].$$

Finding the components of U that will contribute

Since $U \in (\mathfrak{s}^\perp)^- [[x, t_3]]$ and $(\mathfrak{s}^\perp)^- = \prod_{k < 0} (\mathfrak{s}^\perp)^k$, the first thing we need to figure out is how many terms $U^i \in (\mathfrak{s}^\perp)^{-i}$ we need to compute. The easiest way to compute the components of U is through solving the equation $\mathcal{L}_0 = e^{-\text{ad}_U} \mathcal{L}$ for each degree w.r.t. the principal gradation (as in the proof of Theorem 4.3.7).

On the other hand, we need to compute $(e^{\text{ad}_U} \Lambda^3)_+$, where

$$\Lambda^3 = e_1 \lambda + f_1 \lambda^2 \in \bar{\mathfrak{g}}_1 \oplus \bar{\mathfrak{g}}_2,$$

so that we only to compute U down to the homogeneous degree -2 . Therefore, it is worth first finding the homogeneous subspaces $(\mathfrak{s}^\perp)_{-i}$, $i = 0, -1, -2$, and then decompose them through the principal gradation. To do so, we first notice that the Heisenberg generators $\Lambda^{-1} \in \mathfrak{s}^{-1}$ and $\Lambda^{-3} \in \mathfrak{s}^{-3}$ decompose through the homogeneous gradation as

$$\Lambda^{-1} = f_1 + e_1\lambda^{-1} \in \bar{\mathfrak{g}}_0 \oplus \bar{\mathfrak{g}}_{-1}, \quad \Lambda^{-3} = f_1\lambda^{-1} + e_1\lambda^{-2} \in \bar{\mathfrak{g}}_{-1} \oplus \bar{\mathfrak{g}}_{-2}$$

Therefore, it suffices to subtract the vectors $e_1\lambda^{-1}$ and $e_1\lambda^{-2}$ (for instance) from $\bar{\mathfrak{g}}_0 \oplus \bar{\mathfrak{g}}_{-1} \oplus \bar{\mathfrak{g}}_{-2}$ to get a decomposition of $(\mathfrak{s}^\perp)_0 \oplus (\mathfrak{s}^\perp)_{-1} \oplus (\mathfrak{s}^\perp)_{-2}$. In other words,

$$(\mathfrak{s}^\perp)_0 = \mathbb{C}\check{\alpha}_1 \oplus \mathbb{C}f_1, \quad (\mathfrak{s}^\perp)_{-1} = \mathbb{C}\check{\alpha}_1\lambda^{-1} \oplus \mathbb{C}f_1\lambda^{-1}, \quad (\mathfrak{s}^\perp)_{-2} = \mathbb{C}\check{\alpha}_1\lambda^{-2} \oplus \mathbb{C}f_1\lambda^{-2},$$

is one possible decomposition of $(\mathfrak{s}^\perp)_0 \oplus (\mathfrak{s}^\perp)_{-1} \oplus (\mathfrak{s}^\perp)_{-2}$. This tells us that we only need to compute the components of U along each of the vectors in the above decomposition. Now we need to decompose these spaces through the principal gradation. To do so, notice the following:

$$\begin{aligned} \check{\alpha}_1 &\in \bar{\mathfrak{g}}^0, & \check{\alpha}_1\lambda^{-1} &= [f_0, f_1] \in \bar{\mathfrak{g}}^{-2}, & \check{\alpha}_1\lambda^{-2} &= [f_0, f_1\lambda^{-1}] \in \bar{\mathfrak{g}}^{-4}, \\ f_1 &\in \bar{\mathfrak{g}}^{-1}, & f_1\lambda^{-1} &= \frac{1}{2} [f_1, \check{\alpha}_1\lambda^{-1}] \in \bar{\mathfrak{g}}^{-3}, & f_1\lambda^{-2} &= \frac{1}{2} [f_1, \check{\alpha}_1\lambda^{-2}] \in \bar{\mathfrak{g}}^{-5}. \end{aligned}$$

Consequently, there U has no component along $\check{\alpha}_1$ (since $U \in (\mathfrak{s}^\perp)^- [[x, t_3]]$). Moreover, $[f_1\lambda^{-1}, \Lambda^3] = -\check{\alpha}_1\lambda^{-1} \in \bar{\mathfrak{g}}_{-1}$, so that the component along $f_1\lambda^{-2}$ does not contribute to $(e^{\text{ad}_U} \Lambda^3)_+$.

To sum up, we can decompose the function $U \in (\mathfrak{s}^\perp)^- [[x, t_3]]$ through the principal gradation as follows,

$$U = f_1 \otimes u^1 + \check{\alpha}_1\lambda^{-1} \otimes u^2 + f_1\lambda^{-1} \otimes u^3 + \check{\alpha}_1\lambda^{-2} \otimes u^4 + O(\lambda^{-2}),$$

where the component proportional to $u^i \in \mathbb{C} [[x, t_3]]$ belongs to $(\mathfrak{s}^\perp)^{-i}$; the above are exactly the components that will contribute to $(e^{\text{ad}_U} \Lambda^3)_+$.

Computing the components of U

We can now compute the functions u^i via the equation $\mathcal{L}_0 = e^{-\text{ad}_U} \mathcal{L}$, i.e.

$$\Lambda + Q = e^{\text{ad}_U} (\Lambda + H) - \nabla_U U.$$

We decompose the function H in the basis $\{\Lambda^c \mid c \in \mathbb{Z}^{\text{odd}}\}$ of \mathfrak{s} , i.e.

$$H = \Lambda^{-1} \otimes h^1 + \Lambda^{-3} \otimes h^3 + \dots$$

Since $\Lambda \in \mathfrak{s}^1$ and $Q = f_1 \otimes q \in \bar{\mathfrak{g}}^{-1}$, the equation in principal degree 0 reads

$$0 = [f_1 \otimes u^1, \Lambda],$$

so that $f_1 \otimes u^1 \in \text{Ker}(\text{ad}_\Lambda) \cap \text{Im}(\text{ad}_\Lambda)$. Therefore, $u^1 = 0$. Then, the equation in principal degree -1 reads

$$\begin{aligned} f_1 \otimes q &= \Lambda^{-1} \otimes h^1 + [\check{\alpha}_1\lambda^{-1} \otimes u^2, \Lambda] \\ &= (f_1 + e_1\lambda^{-1}) \otimes h^1 + (-2f_1 + 2e_1\lambda^{-1}) \otimes u^2 \\ &= f_1 \otimes (h^1 - 2u^2) + e_1\lambda^{-1} \otimes (h^1 + 2u^2), \end{aligned}$$

from we get that $u^2 = -\frac{1}{4}q$ and also $h^1 = \frac{1}{2}q$. Keeping on with the computations, we find that the components of U we are interested in read

$$u^1 = 0, \quad u^2 = -\frac{1}{4}q, \quad u^3 = \frac{1}{4}q_{(1)}, \quad u^4 = \frac{1}{8}q_{(2)} - \frac{5}{16}(q)^2.$$

We sum up these in the matrix form of U :

$$U = \begin{pmatrix} -\frac{1}{4}q\lambda^{-1} + \left(\frac{1}{8}q_{(2)} - \frac{5}{16}(q)^2\right)\lambda^{-2} & \frac{1}{4}q_{(1)}\lambda^{-1} \\ 0 & -\left(\frac{1}{8}q_{(2)} - \frac{5}{16}(q)^2\right)\lambda^{-2} \end{pmatrix} + \mathcal{O}(\lambda^{-2})$$

Computing the KdV equation

Now that we have the matrix $U \in (\mathfrak{s})^- [[x, t_3]]$, we need to compute $(e^{\text{ad}_U} \Lambda^3)_+$. Since $U \in (\mathfrak{s}^\perp)_{\leq 1} [[x, t_3]]$ and $\Lambda^3 \in \bar{\mathfrak{g}}_1 \oplus \bar{\mathfrak{g}}_2$, we get that

$$(e^{\text{ad}_U} \Lambda^3)_+ = \Lambda^3 + [U, \Lambda^3] + \frac{1}{2} [U, [U, \Lambda^3]].$$

We find

$$(e^{\text{ad}_U} \Lambda^3)_+ = \begin{pmatrix} \frac{1}{4}q_{(1)} & \lambda^2 - \frac{1}{2}q\lambda + \frac{1}{4}q_{(2)} - \frac{1}{2}(q)^2 \\ \lambda - \frac{1}{2}q & -\frac{1}{4}q_{(1)} \end{pmatrix}$$

Together with the connection $\mathcal{L} = \partial + \Lambda + Q$, the above is a well known matrix Lax pair of the KdV equation; indeed,

$$\frac{\partial \mathcal{L}}{\partial t_3} - [(e^{\text{ad}_U} \Lambda^3)_+, \mathcal{L}] = \begin{pmatrix} 0 & \partial_{t_3} q - \frac{1}{4}q_{(3)} + \frac{3}{2}qq_{(1)} \\ 0 & 0 \end{pmatrix}.$$

So, the above equation holds if and only if $\partial_{t_3} q = \frac{1}{4}q^{(3)} - \frac{3}{2}qq_{(1)}$. Substituting $q = -u$, we get that the equation along t_3 of the A_1 -DS hierarchy holds i.f.f.

$$u_{t_3} = \frac{1}{4}u_{(3)} + \frac{3}{2}uu_{(1)},$$

which is the KdV equation.

4.3.4 The tau functions of the DS hierarchies

In this section we define the tau functions of the Drinfeld–Sokolov hierarchy in the sense of Dubrovin and Zhang [DZ08], that is, through tau symmetric Hamiltonian densities (we define these notions below). Then we express the tau function through a zero-curvature expression of the Drinfeld–Sokolov hierarchies. This universal construction, which applies to the DS hierarchy of any type \mathfrak{g} (for \mathfrak{g} an untwisted affine algebra) is due, again, to C.-Z. Wu [Wu17]. In the latter article, Wu actually defines a tau function even for twisted affine algebra, but we will not use that in the sequel.

Definition of the tau function

Let $\mathcal{L} = \partial + \Lambda + Q \in \text{op}_{\mathfrak{g}}$ and let functions $U \in \widehat{\mathfrak{g}}^-[[\mathbf{t}]]$ and $H \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]]$ be as in Theorem 4.3.7. Let us consider a function $\Omega \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]]$, which we decompose in the basis $\{\Lambda_j \mid j \in \mathbb{E}\}$ as

$$\Omega = \sum_{j \in \mathbb{E}^+} \frac{\Omega^j}{j} \Lambda_{-j},$$

and such that

$$\frac{\partial \Omega}{\partial x} = -H, \quad (4.57)$$

As a formal power series, the function Ω is well defined as a primitive of the function $-H$. However, Ω may no longer be a differential polynomial in the components q^1, \dots, q^r of Q (as is H). Nonetheless, we ask Equation (4.57) to hold within the ring of differential polynomials in the q^i 's. For that reason, the function Ω is defined up to the addition of some constant term in $\widehat{\mathfrak{s}}^-$; equivalently, its coefficients Ω^j are defined up to the addition of some constant complex number.

Lemma 4.3.23 ([Wu17]). *Each function $\partial_{t_i} \Omega^j$, for $i, j \in \mathbb{E}^+$, is a differential polynomial in the components of the function Q ; they satisfy the symmetry condition*

$$\frac{\partial \Omega^j}{\partial t_i} = \frac{\partial \Omega^i}{\partial t_j}. \quad (4.58)$$

Proof. Recall that $\mathcal{L} = e^{\text{ad}_U}(\partial + \Lambda + H)$. It can be verified straightforwardly that

$$\frac{\partial \mathcal{L}}{\partial t_j} = e^{\text{ad}_U} \frac{\partial H}{\partial t_j} + [\nabla_{t_j, U}(U), \mathcal{L}], \quad (4.59)$$

where, similarly to Equation (4.34),

$$\nabla_{t_j, U}(U) = \frac{e^{\text{ad}_U} - 1}{\text{ad}_U} \frac{\partial U}{\partial t_j} = \sum_{k \geq 0} \frac{1}{(k+1)!} (\text{ad}_U)^k \frac{\partial U}{\partial t_j}.$$

Now using $\partial_{t_j} \mathcal{L} = [(e^{\text{ad}_U} \Lambda_j)_+, \mathcal{L}]$, we can rewrite Equation (4.59) as

$$e^{\text{ad}_U} \frac{\partial H}{\partial t_j} + [\nabla_{t_j, U}(U) + (e^{\text{ad}_U} \Lambda_j)_-, \mathcal{L}] = + [e^{\text{ad}_U} \Lambda_j, \mathcal{L}]. \quad (4.60)$$

Now, notice that

$$[e^{\text{ad}_U} \Lambda_j, \mathcal{L}] = e^{\text{ad}_U} [\Lambda_j, \partial + \Lambda + H] = -e^{\text{ad}_U} \frac{\partial \Omega^j}{\partial x} \cdot c,$$

where we have used $[\Lambda_j, \Lambda_k] = j\delta_{j+k,0} \cdot c$ and $H = -\partial_x \Omega$ (4.57). Therefore, Equation (4.60) becomes

$$\frac{\partial H}{\partial t_j} + \left[e^{-\text{ad}_U} \left(\nabla_{t_j, U}(U) + (e^{\text{ad}_U} \Lambda_j)_- \right), \partial + \Lambda + H \right] = -\frac{\partial \Omega^j}{\partial x} \cdot c. \quad (4.61)$$

Now, since $\mathfrak{g} = \mathfrak{s} + \text{Im}(\text{ad}_\Lambda)$ with $\mathfrak{s} \cap \text{Im}(\text{ad}_\Lambda) \in \mathfrak{g}^0$, it follows that there exists two functions $G^{[j]} \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]]$ and $\tilde{G}^{[j]} \in (\text{Im}(\text{ad}_\Lambda))^-[[\mathbf{t}]]$ such that

$$e^{-\text{ad}_U} \left(\nabla_{t_j, U}(U) + (e^{\text{ad}_U} \Lambda_j)_- \right) = G^{[j]} + \tilde{G}^{[j]} \in \widehat{\mathfrak{g}}^-. \quad (4.62)$$

Equation (4.60) then splits into three parts:

$$\partial_{t_j} H - \partial_x G^{[j]} = 0, \quad (4.63)$$

$$[G^{[j]}, \Lambda]_c + \partial_x \Omega^j = 0, \quad (4.64)$$

$$[\tilde{G}^{[j]}, \Lambda + H] - \partial_x \tilde{G}^{[j]} = 0.$$

From the definition of Ω , Equation (4.63) reads $\partial_x \partial_{t_j} \Omega = -\partial_x G^{[j]}$, so that

$$\frac{\partial \Omega}{\partial t_j} = -G^{[j]}. \quad (4.65)$$

There would actually be an integration constant $K \in \mathbb{C}$, yet for the particular solution $Q = 0$ of the hierarchy, we have $U = H = 0$ so that Ω is constant and $G^{[j]} = 0$ as well by Equation (4.62) (since $(\Lambda_j)_- = 0$). Therefore, the integration constant vanishes. It follows from Equation (4.65) that

$$[\Lambda_i, G^{[j]}]_c = -[\Lambda_i, \partial_{t_j} \Omega] = -\frac{\partial \Omega^i}{\partial t_j}.$$

Since U is made of differential polynomials in Q , it follows that the $\partial_{t_j} \Omega^i$ is as well. Taking $i = 1$ in the above and using $\Lambda_1 = \nu \Lambda$ and (4.64), we find

$$\frac{\partial \Omega^1}{\partial t_j} = -\nu [\Lambda, G^{[j]}]_c = -\nu \frac{\partial \Omega^j}{\partial x}.$$

By symmetry of the second derivatives, we get

$$\frac{\partial^2 \Omega^j}{\partial x \partial t_i} = \frac{\partial^2 \Omega^i}{\partial x \partial t_j}.$$

Since the $\partial_{t_j} \Omega^i$'s are differential polynomials in Q , it follows that $\partial_{t_j} \Omega^i = \partial_{t_i} \Omega^j$ up to a constant. Yet, again, for $Q = 0$, the Ω^i 's vanish and so does the constant. Hence the result. \square

By Poincaré lemma, Equation (4.58) implies the existence of a function $\tau \in \mathbb{C}[[\mathbf{t}]]$ such that $\Omega^j = \partial_{t_j} \log \tau$. We can now define the tau function.

Definition 4.3.24 (Tau function). Let $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ and let $H \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]]$ be as in Theorem 4.3.7. Let Ω be such that $\partial_x \Omega + H = 0$, with $\Omega = \sum_{j>0} j^{-1} \Omega^j \Lambda_{-j}$. We call a tau function of the \mathfrak{g} -DS hierarchy any function $\tau \in \mathbb{C}[[\mathbf{t}]]$ such that

$$\frac{\partial \log \tau}{\partial t_j} = -\Omega^j. \quad (4.66)$$

We say that two tau functions $\tau, \tilde{\tau}$ coincide if their logarithm differ from a linear function of the variables t_j , i.e. if there exist $c_j \in \mathbb{C}$ such that

$$\log \tilde{\tau} = \log \tau + \sum_{j \in \mathbb{E}^+} c_j t_j.$$

Tau symmetric conserved densities

We now establish the link between the tau function and the coefficients $H^i \in \mathfrak{s}^{-i}[[\mathbf{t}]]$ of the function H as in Theorem 4.3.7. First recall that the functions H^i are differential polynomials in the q^i 's and are conserved densities in the sense that for any $i, j \in \mathbb{E}^+$, we have

$$\frac{\partial}{\partial t_j} \int H^i = 0$$

(Proposition 4.3.21). By the above we mean that there exists functions $F^i \in \mathfrak{s}^{-i}[[\mathbf{t}]]$ —which are also differential polynomial in the q^i 's—such that $\partial_{t_j} H^i = \partial F^i$ (with $\partial = \partial_x$). Therefore, in the sequel, we will consider the formal expression

$$\partial^{-1}(\partial_{t_j} H) \in \widehat{\mathfrak{s}}^- [[\mathbf{t}]],$$

as the only differential polynomial expression such $(\partial^{-1}(\partial_{t_j} H))^i = F^i$, with the integration constant fixed to 0.

Proposition 4.3.25. *The function H as in Theorem 4.3.7 is related to the tau function through the relation*

$$\frac{\partial^2 \log \tau}{\partial t_i \partial t_j} = j \frac{\partial^{-1}(\Lambda_j, \partial_{t_i} H)}{(\Lambda_j, \Lambda_{-j})}. \quad (4.67)$$

The above does not depend on the normalization factor h (in that case, the Coxeter number of \mathfrak{g}) in $(\Lambda_i, \Lambda_j)_0 = h\delta_{i+j,0}$, and can be rewritten as

$$\frac{\partial^2 \log \tau}{\partial t_i \partial t_j} = -\partial^{-1} [\partial_{t_i} H, \Lambda_j]_c.$$

It is proved straightforwardly by noticing that, by the definition $\partial_x \Omega = -H$, we have $[H, \Lambda_j] = \partial(\Omega^j) \cdot c$ and then using $\Omega^j = -\partial_{t_j} \log \tau$.

We naturally define the following differential polynomials:

$$\tilde{h}^j = -j \frac{(\Lambda_j, H)}{(\Lambda_j, \Lambda_{-j})} \in \mathcal{A}_q.$$

They are conserved densities as well and satisfy the tau symmetry property in the sense of the following proposition.

Proposition 4.3.26 (Tau symmetric densities). *The conserved densities \tilde{h}^j satisfy the tau symmetry property: for any $i, j \in \mathbb{E}^+$,*

$$\frac{\partial \tilde{h}^j}{\partial t_i} = \frac{\partial \tilde{h}^i}{\partial t_j}.$$

It suffices to notice that $\partial_{t_i} \tilde{h}^j = j(\Lambda_j, \Lambda_j)^{-1}(\Lambda_j, \partial_{t_i} \partial_x \Omega) = -\partial_x \partial_{t_i} \Omega^j$ and then use $\partial_{t_i} \Omega^j = \partial_{t_j} \Omega^i$ (see Equation (4.58)).

Zero-curvature representation and dressing of the connection

In this section, we construct a dressing operator for the connection \mathcal{L} as well as a zero-curvature representation of the \mathfrak{g} -DS hierarchy. We follow [Wu17], §3.3. Let functions $U \in \widehat{\mathfrak{g}}^-[[\mathbf{t}]]$ and $H \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]]$ be as in Theorem 4.3.7. Let $\Omega \in \widehat{\mathfrak{g}}^-[[\mathbf{t}]]$ be such that $\partial\Omega = -H$. We now introduce a new function, $T = U \bullet \Omega$, where \bullet denotes the Baker–Campbell–Hausdorff product (4.3), i.e. T is characterised by

$$e^{\text{ad}_T} = e^{\text{ad}_U} e^{\text{ad}_\Omega}.$$

Since both U and Ω have only negative degrees w.r.t. the principal gradation, $T \in \widehat{\mathfrak{g}}^-[[\mathbf{t}]]$ is well defined. In general, the function T may not be a differential polynomials in the q^i 's, for Ω may not as well. Notice that in [Wu17] (which we still follow here), the author introduces the power series $\Theta = e^U e^\Omega$ which belongs to the formal Kac–Moody group of \mathfrak{g} . However, everything can be formulated using only T which need not introducing the formal Kac–Moody group.

In the study of the KP hierarchy (Section 3.2) we have seen that there always exists an operator M such that the scalar Lax operator Q can be written as $Q = M\partial M^{-1}$ (Proposition 3.2.3); the operator M is what we call a *dressing operator* for the operator Q . The proposition below is an adaptation of this notion of dressing to the case of connections of $\text{op}_{\mathfrak{g}}$; the operator e^{ad_T} plays the role of of the dressing M . However we will not dress up the connection $\mathcal{L} \in \text{op}_{\mathfrak{g}}$, but the following connection:

$$\mathcal{L}_1 := \nu \mathcal{L} - \Omega^1 \cdot c,$$

where $\nu \in \mathbb{C}$ is such that $\Lambda_1 = \nu\Lambda$ and $\partial_{t_1} = \nu\partial$ as well.

Proposition 4.3.27 (Dressing operator). *Let $\mathcal{L} \in \text{op}_{\mathfrak{g}}$ and let the functions U, H be as in Theorem 4.3.7. Recall the function*

$$\Omega = \sum_{j \in \mathbb{E}^+} \frac{\Omega^j}{j} \Lambda_{-j} \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]]$$

such that $\partial\Omega = -H$. Let $T \in \widehat{\mathfrak{g}}^-[[\mathbf{t}]]$ be such that $e^{\text{ad}_T} = e^{\text{ad}_U} e^{\text{ad}_\Omega}$. Let $\mathcal{L}_1 = \nu \mathcal{L} - \Omega^1 \cdot c$, then

$$\mathcal{L}_1 = e^{\text{ad}_T} (\partial_{t_1} + \Lambda_1). \quad (4.68)$$

We call the automorphism e^{ad_T} a dressing operator for the connection \mathcal{L}_1 ; or we call T a dressing function.

As advertised for, this naturally agrees with a notion of dressing since, if we consider the formal series $\Theta = e^U e^\Omega$, the equation becomes $\mathcal{L}_1 = \Theta(\partial_{t_1} + \Lambda_1)\Theta^{-1}$.

Proof. First, via $\Lambda_1 = \nu\Lambda$ and $\partial_{t_1} = \nu\partial$, it amounts to proving that $\mathcal{L} = e^{\text{ad}_U} e^{\text{ad}_\Omega} (\partial + \Lambda + \nu^{-1}\Omega^1 \cdot c)$. In turns, from $\mathcal{L} = e^{\text{ad}_U} (\partial + \Lambda + H)$, it amounts to proving that

$$e^{-\text{ad}_\Omega} (\partial + \Lambda + H) = \partial + \Lambda + \nu^{-1}\Omega^1 \cdot c. \quad (4.69)$$

So we compute: using (4.37), $\partial\Omega = -H$ and the commutativity of $\widehat{\mathfrak{s}}^-$, we have

$$e^{-\text{ad}_\Omega} \partial = \partial + \sum_{k \geq 0} \frac{(-1)^k}{(k+1)!} (\text{ad}_\Omega)^k \partial\Omega = \partial - H.$$

On the other hand, using again the commutativity of $\widehat{\mathfrak{s}}^-$ along with the Heisenberg relations (4.26), we have

$$\begin{aligned} e^{-\text{ad}_\Omega}(\Lambda + H) &= \Lambda + H - [\Omega, \Lambda] \\ &= \Lambda + H - \nu^{-1} \Omega^1 [\Lambda_{-1}, \Lambda_1] \\ &= \Lambda + H + \nu^{-1} \Omega^1 \cdot c. \end{aligned}$$

Hence Equation (4.69), hence the proposition holds. \square

Let $S \in \mathfrak{n}[[\mathfrak{t}]]$ and consider the gauge transformation $\mathcal{L} \mapsto \tilde{\mathcal{L}} = e^{\text{ad}_S} \mathcal{L}$. Since $H = \tilde{H}$ is invariant under gauge transformation (and so is Ω) while $e^{\text{ad}_{\tilde{U}}} = e^{\text{ad}_S} e^{\text{ad}_U}$, then $e^{\text{ad}_{\tilde{T}}} = e^{\text{ad}_S} e^{\text{ad}_T}$. The following lemma holds.

Lemma 4.3.28. *There exists $S \in \mathfrak{n}[[\mathfrak{t}]]$ and $\tilde{T} \in \widehat{\mathfrak{g}}_-[[\mathfrak{t}]]$ such that $e^{\text{ad}_S} e^{\text{ad}_T} = e^{\text{ad}_{\tilde{T}}}$.*

Proof. The problem here is that if $T \in \widehat{\mathfrak{g}}^-[[\mathfrak{t}]]$, there can be some component $T_0 \in \mathfrak{g}_0[[\mathfrak{t}]]$ (namely, of the form $[e_{i_1}, [\dots, e_{i_k}]] \dots$ with $i_1, \dots, i_k \neq 0$). But if we find a gauge transformation $e^{\text{ad}_{\tilde{T}}} = e^{\text{ad}_S} e^{\text{ad}_T}$ such that $\tilde{T}_0 = 0$, then by definition $\tilde{T} \in \widehat{\mathfrak{g}}_-[[\mathfrak{t}]]$. From $e^{\text{ad}_S} e^{\text{ad}_T} = e^{\text{ad}_{S \bullet T}}$, since the BCH formula (4.3) is a homomorphism, then $(S \bullet T)_0 = S_0 \bullet T_0$. We easily see that choosing $S_0 = -T_0$, we get $S_0 \bullet T_0 = 0$. Now recall that $\mathfrak{n} = \mathfrak{g}_0 \cap \mathfrak{g}^{<0}$ and therefore $T_0 \in \mathfrak{n}[[\mathfrak{t}]]$ so that $S_0 = -T_0$ is consistent. That proves the lemma. \square

From now on we will assume that we chose a slice \mathcal{L} of the gauge class $[\mathcal{L}]$ such that the function T belongs to $\widehat{\mathfrak{g}}_-[[\mathfrak{t}]]$. We can rewrite the \mathfrak{g} -DS hierarchy using the dressing operator T as follows.

Lemma 4.3.29. *The \mathfrak{g} -DS hierarchy (4.43) is equivalent to the following system of equations,*

$$\left\{ \partial_{t_j} \mathcal{L}_1 = [(e^{\text{ad}_T} \Lambda_j)_+, \mathcal{L}_1] \mid j \in \mathbb{E}^+ \right\}. \quad (4.70)$$

Proof. First notice that $e^{\text{ad}_T} \Lambda_j = e^{\text{ad}_U} \Lambda_j - \Omega^j \cdot c$. Indeed, $[\Omega, \Lambda_j] = -\Omega^j \cdot c$ from which we get that

$$e^{\text{ad}_\Omega} \Lambda_j = \Lambda_j - \Omega^j \cdot c. \quad (4.71)$$

Therefore,

$$\begin{aligned} [\mathcal{L}_1, (e^{\text{ad}_T} \Lambda_j)_+] &= [\nu \mathcal{L} - \Omega^1 \cdot c, (e^{\text{ad}_U} \Lambda_j)_+ - \Omega^j \cdot c] \\ &= \nu [\mathcal{L}, (e^{\text{ad}_U} \Lambda_j)_+] - \nu [\mathcal{L}, \Omega^j \cdot c] \\ &= -\nu \partial_{t_j} \mathcal{L} - \nu \partial \Omega^j \cdot c. \end{aligned}$$

On the other hand,

$$\partial_{t_j} \mathcal{L}_1 = \nu \partial_{t_j} \mathcal{L} + \partial_{t_j} \Omega^1 \cdot c.$$

We conclude via $\partial_{t_j} \Omega^1 = \nu \partial \Omega^j$ (Equation (4.58)). \square

The following proposition, which we take from [Wu17] (Lemma 3.7), already appeared in [HM93, HMSG94, Mir99].

Proposition 4.3.30. *The dressing function $T \in \widehat{\mathfrak{g}}_-[[\mathfrak{t}]]$, which is such that $e^{\text{ad}_T} = e^{\text{ad}_U} e^{\text{ad}_\Omega}$, satisfies the following system of consistent equations,*

$$\left\{ \nabla_{t_j, T}(T) = -(e^{\text{ad}_T} \Lambda_j)_- \mid j \in \mathbb{B}^+ \right\}, \quad (4.72)$$

where

$$\nabla_{t_j, T}(T) = \sum_{k \geq 0} \frac{1}{(k+1)!} (\text{ad}_T)^k \frac{\partial T}{\partial t_j}.$$

Moreover, the system (4.72) is equivalent to the \mathfrak{g} -DS hierarchy (4.72) on the connection \mathcal{L} .

Remark 4.3.31. If we take $\Theta = e^T$ in the Kac–Moody group of \mathfrak{g} , then it is not hard to see from the relation $e^{\text{ad}_T} \partial_{t_j} = \partial_{t_j} - \nabla_{t_j, T}(T)$ (4.37) that $\nabla_{t_j, T}(T) = (\partial_{t_j} \Theta) \Theta^{-1}$, so that the equations of (4.72) become

$$\partial_{t_j} \Theta = -(\Theta \Lambda_j \Theta^{-1})_- \Theta,$$

as in [Wu17], Lemma 3.7. Compare with Sato’s equation (Theorem 3.2.4) on the dressing of the scalar Lax operator in the study of the KP hierarchy.

Proof. First, these equations are consistent since we chose T in $\widehat{\mathfrak{g}}_-[[\mathfrak{t}]]$ thanks to Lemma 4.3.28. Second, notice that Equation (4.70) is equivalent to

$$\frac{\partial \mathcal{L}_1}{\partial t_j} = [\mathcal{L}_1, (e^{\text{ad}_T} \Lambda_j)_-].$$

Indeed, it suffices to see that $[\mathcal{L}_1, e^{\text{ad}_T} \Lambda_j] = e^{\text{ad}_T} [\partial_{t_1} + \Lambda_1, \Lambda_j] = 0$ by Equation (4.68) and commutativity of \mathfrak{s}^+ . Substituting this into Equation (4.68), we get

$$\begin{aligned} 0 &= \partial_{t_j} \left(e^{\text{ad}_T} (\partial_{t_1} + \Lambda_1) \right) + [(e^{\text{ad}_T} \Lambda_j)_-, \mathcal{L}_1] \\ &= e^{\text{ad}_T} \left(\partial_{t_j} (\partial_{t_1} + \Lambda_1) \right) + [\nabla_{t_j, T}(T), \mathcal{L}_1] + [(e^{\text{ad}_T} \Lambda_j)_-, \mathcal{L}_1] \\ &= [\nabla_{t_j, T}(T) + (e^{\text{ad}_T} \Lambda_j)_-, \mathcal{L}_1]. \end{aligned}$$

Let us denote, for $j \in \mathbb{B}^+$,

$$\Delta^{[j]} = e^{-\text{ad}_T} \left(\nabla_{t_j, T}(T) + (e^{\text{ad}_T} \Lambda_j)_- \right).$$

Then $\Delta^{[j]} \in \widehat{\mathfrak{g}}_-[[\mathfrak{t}]]$, and satisfies

$$\left[\Delta^{[j]}, \partial_{t_1} + \Lambda_1 \right] = 0.$$

Since $\mathfrak{g}_{<0} = \mathfrak{s}_{<0} \oplus \text{Im}(\text{ad}_\Lambda)_{<0}$, it follows that

$$\Delta^{[j]} = \sum_{i \in \mathbb{B}^+} c_i \Lambda_{-i},$$

with $c_j \in \mathbb{C}$ constants. Now notice that for the special solution $Q = 0$, which implies $U = H = \Omega = 0$ and therefore $T = 0$ as well, we get $\Delta^{[j]} = 0$ (since $(\Lambda_j)_- = 0$). Therefore, $\Delta^{[j]} = 0$ identically. Moreover, we easily derive Equations (4.70) from (4.72). \square

We can now formulate the zero-curvature representation of the g-DS hierarchy.

Proposition 4.3.32 (Zero-curvature representation). *Let $j \in \mathbb{E}^+$ be a positive exponent and define the following connection,*

$$\mathcal{L}_j := e^{\text{ad}_T}(\partial_{t_j} + \Lambda_j).$$

They satisfy the zero-curvature equation:

$$[\mathcal{L}_i, \mathcal{L}_j] = 0. \quad (4.73)$$

Proof. The proof is immediate: by construction, for any $i, j \in \mathbb{E}^+$, we have $[\mathcal{L}_i, \mathcal{L}_j] = e^{\text{ad}_T}[\partial_{t_i} + \Lambda_i, \partial_{t_j} + \Lambda_j]$, and by commutativity of \mathfrak{s}^+ , this yields Equation (4.73). \square

Relation with the tau function

Using Proposition 4.3.30, we can rewrite the operator \mathcal{L}_j as

$$\mathcal{L}_j = \partial_{t_j} + \Lambda_j + Q^{[j]}, \quad Q^{[j]} = (e^{\text{ad}_T} \Lambda_j)_+ - \Lambda_j \in \mathfrak{g}_+[[\mathbf{t}]].$$

Now recall that the tau function of the g-DS hierarchy was defined as any function $\tau \in \mathbb{C}[[\mathbf{t}]]$ such that

$$\Omega^j = -\frac{\partial \log \tau}{\partial t_j},$$

where $\Omega = \sum_{j \in \mathbb{E}^+} j^{-1} \Omega^j \Lambda_{-j}$ is such that $\partial \Omega = -H$ and $H \in \widehat{\mathfrak{s}}^-[[\mathbf{t}]]$ is as in Theorem 4.3.7. From this, we get the following proposition which redefines the tau function in terms of the dressing function T .

Proposition 4.3.33. *The tau function $\tau \in \mathbb{C}[[\mathbf{t}]]$ of the g-DS hierarchy satisfies the following system of equations,*

$$\frac{\partial \log \tau}{\partial t_j} = (e^{\text{ad}_T} \Lambda_j)_c, \quad j \in \mathbb{E}^+, \quad (4.74)$$

where the subscript c denotes the component along $\mathbb{C}c$ in the decomposition $\mathfrak{h} = \mathbb{C}c \oplus \bigoplus_{i=1}^r \mathbb{C}\check{\alpha}_i$ of the affine Cartan subalgebra.

Proof. The proof is straightforward:

$$\begin{aligned} Q^{[j]} &= (e^{\text{ad}_T} \Lambda_j)_+ - \Lambda_j \\ &= (e^{\text{ad}_U} (\Lambda_j - \Omega^j \cdot c))_- - \Lambda_j \\ &= (e^{\text{ad}_U} \Lambda_j)_+ - \Omega^j \cdot c - \Lambda_j. \end{aligned}$$

Since $(e^{\text{ad}_U} \Lambda_j)_c = 0$ (4.40), we get that $(Q^{[j]})_c = -\Omega^j$. Yet by definition of $Q^{[j]}$, we also have $(Q^{[j]})_c = (e^{\text{ad}_T} \Lambda_j)_c$, hence the result. \square

4.4 Tau functions as limits of block Toeplitz determinants

In this section, we follow [CW15a, CW15b]. First of all, we fix an n -dimensional faithful and traceless matrix realization of \mathfrak{g} that we denote $\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(n, \mathbb{C})$. For the classical simple Lie algebras, we choose π as in Appendix A. This representation extends to $\mathfrak{g} = \mathfrak{g}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c$ by

$$\pi(X\lambda^p + a \cdot c) = \pi(X)\lambda^p + a \cdot c.$$

In the sequel, we will identify \mathfrak{g} or \mathfrak{g} with their representation. The goal of this section is to prove the following theorem, due to Cafasso and Wu:

Theorem 4.4.1 ([CW15a]). *Let $X \in \mathfrak{g}_-$ and let $\Xi(\mathbf{t}; \lambda) = \sum_{j \in \mathbb{B}^+} t_j \Lambda_j$. Then¹¹*

$$\tau_X := \left[\det \left(e^{\Xi(\mathbf{t}; \lambda)} e^X \right) \right]^\kappa,$$

is well defined (provided the context developed below) and is a tau function of the \mathfrak{g} -DS hierarchy.

Of course, an important part of the present section will be devoted to providing the context in which the determinant defined above is well defined. The number κ depends on the representation of $\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(n, \mathbb{C})$; it is given by

$$(X, Y)_0 = \kappa \cdot \text{tr}(\pi(X)\pi(Y)),$$

where $(X, Y)_0$ is the normalized invariant form (see Section 4.1). In the cases we will study and given the representation in Appendix A, $\kappa = 1$ for the A_n and C_n cases and $\kappa = 1/2$ for the B_n and D_n cases.

Before getting into the subject, we make an important observation. Since we have fixed a representation, given any $X \in \mathfrak{g}_-$, the power series

$$e^X = \sum_{n \geq 0} \frac{X^n}{n!} = \text{id} + \mathcal{O}(\lambda^{-1})$$

is well defined: each term in λ^k is a finite sum; the product of elements of \mathfrak{g} is defined by the matrix product. Similarly, e^X is well defined for all $X \in \mathfrak{g}_+$. Following Remark 4.3.31, we can then define the following function,

$$\Theta := e^U e^\Omega \in G_-[[\mathbf{t}]],$$

where $\partial_x \Omega = -H$ and where U, H are functions as in Theorem 4.3.7. The function Θ we call the *(matrix) dressing function* and is obviously equivalent to the (matrix) dressing operator $e^{\text{ad}_T} = e^{\text{ad}_U} e^{\text{ad}_\Omega}$. Then Equations (4.72), which are equivalent to the \mathfrak{g} -DS hierarchy, become

$$\frac{\partial \Theta}{\partial t_j} = -(\Theta \Lambda_j \Theta^{-1})_- \Theta. \quad (4.75)$$

¹¹Notice that in [CW15a, CW15b], $g = \exp(\sum t_j \Lambda_j)$ is replaced with $g^{-1} = \exp(-\sum t_j \Lambda_j)$. This is due to the fact that in [CW15a, CW15b], the authors use the convention to write the \mathfrak{g} -DS hierarchy as $\partial_{t_j} \mathcal{L} = [(e^{\text{ad}_U} \Lambda_j)_+, \mathcal{L}]$; while in this thesis, we use the convention $\partial_{t_j} \mathcal{L} = [\mathcal{L}, (e^{\text{ad}_U} \Lambda_j)_+]$ (as in [CDY18]), which amounts to replacing t_j with $-t_j$ and therefore g with g^{-1} .

Let us denote $\Theta|_{t=0}$ ($t_1 = vx = 0$ included) the initial data of a solution Θ of the \mathfrak{g} -DS hierarchy. The following lemma holds.

Lemma 4.4.2. *For any $X \in \mathfrak{g}_-$, there exists a solution Θ of the \mathfrak{g} -DS hierarchy such that*

$$\Theta|_{t=0} = e^X.$$

To show that, it suffices to write $e^X = e^Y \Theta|_{t=0}$ with $e^Y = e^X \Theta|_{t=0}^{-1}$. Then

$$\begin{aligned} (e^Y \Theta \Lambda_j \Theta^{-1} e^{-Y})_- e^Y \Theta &= (e^{\text{ad}_Y} \Theta \Lambda_j \Theta^{-1})_- e^Y \Theta \\ &= e^Y (\Theta \Lambda_j \Theta^{-1})_- \Theta, \end{aligned}$$

where in the second equation, we have used the fact that $Y \in \mathfrak{g}_-$. Hence, if Θ satisfies Equations (4.75), then so does $e^Y \Theta$. Yet $(e^Y \Theta)|_{t=0} = e^X$, so the lemma is proved. We can rephrase this result as *the group $\{e^X \mid X \in \mathfrak{g}_-\}$ acts transitively on the space of dressing operators.*

4.4.1 The Grassmannian approach and Toeplitz determinants

Preliminaries on topology and the formal loop group

In what follows, we want to define a structure such that something like $e^{\Xi(t;\lambda)} e^X$ converges for some topology. This structure we call the *formal loop group* associated to \mathfrak{g} . To do so, we will consider formal power series of the form

$$\gamma = \sum_{I \in \mathbb{Z}} \gamma_I \lambda^I, \quad \gamma_I = \sum_{A \geq 0} \gamma_{I,A} \zeta^A, \quad \gamma_{I,A} \in \mathfrak{gl}(n) [[\mathbf{t}]].$$

In other words, $\gamma \in \mathfrak{gl}(n) [[\lambda, \lambda^{-1}, \zeta, \mathbf{t}]]$. Here the problem is that the Cauchy product of two power series of this form may not always be well defined, some coefficients in λ^I may give infinite sums. To solve this problem, we define a topology given by the valuation distance.

For this section, let us denote by R the ring

$$R = \mathfrak{gl}(n) [[\lambda, \lambda^{-1}, \mathbf{t}]],$$

which is not a commutative ring nor is it an integral domain. The Cauchy product of $f, g \in R [[\zeta]]$ is given by

$$f \cdot g := \sum_{A \geq 0} \zeta^A \sum_{B=0}^A f_{A-B} \cdot g_B,$$

where $f_{A-B} \cdot g_B$ is defined by the usual matrix product. Similarly to the case of pseudo-differential operators (see Definition 3.1.3), the valuation of $f \in R [[\zeta]]$, denoted $\text{val}_\zeta f$, is defined as the smallest $V \geq 0$ such that $f_V \neq 0$, i.e.

$$\text{val}_\zeta \sum_{A \geq 0} f_A \zeta^A = \inf \{A \geq 0 \mid f_A \neq 0\}.$$

If we set $\text{val}_\zeta 0 = +\infty$, then $\text{val}_\zeta (f + g) \geq \min\{\text{val}_\zeta f, \text{val}_\zeta g\}$ and

$$\text{val}_\zeta (f \cdot g) \geq \text{val}_\zeta f + \text{val}_\zeta g$$

(we would have had an equality if R were an integral domain). Then the valuation distance can be defined by

$$\text{dist}(f, g) = 2^{-\text{val}(f-g)}$$

The topology induced by this distance is such that a sequence $f_N = \sum_{a \geq 0} f_{N,A} \zeta^A$ converges to $f = \sum_{a \geq 0} f_A \zeta^A$ if and only if for all $B \geq 0$, there exists $N \geq 0$ such that for all $A \leq B$, we have $f_{N,A} = f_A$. This construction makes $R[[\zeta]]$ into a *complete topological ring* (see e.g. [Bou88] §4).

Below we define the notion of *admissible series* in a way that generalizes to the matrix case the one given by Borodin and Okounkov in [BO00] in the scalar case. In [CW15b] a similar definition is given, although it makes use of the principal gradation rather than the homogeneous gradation; this applies well to the computation of the topological solution.

Definition 4.4.3 (Admissible series). Let $\gamma = \sum_{I \in \mathbb{Z}} \gamma_I \lambda^I$ be a formal power series with entries $\gamma_I \in \text{gl}(n)[[\mathbf{t}, \zeta]]$. We say that γ is admissible if for any $I \in \mathbb{Z}$, if the valuation of the entries satisfy

$$\text{val}_\zeta (\gamma_I) \geq |I|, \quad (4.76)$$

This definition seems arbitrary, but below we will see that it naturally applies to the computation of $\exp(\Xi(\mathbf{t}; \lambda)) \exp(X)$. The following elementary lemma holds.

Lemma 4.4.4. *If $\gamma, \tilde{\gamma}$ are two admissible series, then their Cauchy product (over \mathbb{Z}),*

$$\gamma \cdot \tilde{\gamma} := \sum_{I \in \mathbb{Z}} \lambda^I \sum_{J \in \mathbb{Z}} \gamma_{I-J} \tilde{\gamma}_J, \quad (4.77)$$

is well defined and is also an admissible series.

Proof. Let $\gamma, \tilde{\gamma}$ be two admissible series, then

$$\gamma \cdot \tilde{\gamma} = \sum_{I \in \mathbb{Z}} \lambda^I \sum_{J \in \mathbb{Z}} \gamma_{I-J} \tilde{\gamma}_J = \sum_{I \in \mathbb{Z}} \lambda^I \left(\gamma_I \tilde{\gamma}_0 + \sum_{J > 0} \gamma_{I-J} \tilde{\gamma}_J + \gamma_{I+J} \tilde{\gamma}_{-J} \right).$$

On the one hand,

$$\begin{aligned} \text{val}_\zeta (\gamma_{I-J} \tilde{\gamma}_J) &\geq |I - J| + |J| \\ &\geq 2|J| - |I|, \end{aligned}$$

and similarly $\text{val}_\zeta (\gamma_{I+J} \tilde{\gamma}_{-J}) \geq 2|J| - |I|$, so that both $\text{val}_\zeta (\gamma_{I-J} \tilde{\gamma}_J)$ and $\text{val}_\zeta (\gamma_{I+J} \tilde{\gamma}_{-J})$ are strictly increasing functions of $J > 0$. In particular, for any $A \geq 0$, only finitely many terms appear in the coefficient in ζ^A of $(\gamma \cdot \tilde{\gamma})_I$. This shows that the Cauchy product is well defined. On the other hand, $\text{val}_\zeta (\gamma_I \tilde{\gamma}_0) \geq |I|$, but also

$$\text{val}_\zeta (\gamma_{I-J} \tilde{\gamma}_J) \geq |I - J| + |J| \geq |I|,$$

and similarly $\text{val}_\zeta (\gamma_{I+J} \tilde{\gamma}_{-J}) \geq |I|$. This proves that $\gamma \cdot \tilde{\gamma}$ is an admissible series. \square

Clearly, the Cauchy product (4.77) is associative and the element $\text{id} \in \mathfrak{gl}(n)$ is the identity element for this law. This makes the set of all admissible series into a monoid. Naturally, we say that an admissible series $\gamma \in R[[\zeta]]$ is invertible if there exists an admissible series $\gamma^{-1} \in R[[\zeta]]$ such that

$$\gamma \cdot \gamma^{-1} = \gamma^{-1} \cdot \gamma = \text{id}. \quad (4.78)$$

This leads us to the following definition.

Definition 4.4.5 (Formal loop group). We call a (formal) *admissible loop* any power series $\gamma \in R[[\zeta]]$ which is invertible for the Cauchy product (4.77). We denote $\text{LG}(n)$, and call the (formal) *loop group*, the multiplicative group of admissible loops, i.e.

$$\text{LG}(n) = \{\gamma \in R[[\zeta]] \mid \gamma \text{ is admissible (4.76) and invertible (4.78)}\},$$

where $R = \mathfrak{gl}(n)[[\lambda, \lambda^{-1}, \mathbf{t}]]$.

We now explain how this relates to computing $e^{\Xi(\mathbf{t}; \lambda)} e^X$ for $X \in \mathfrak{g}_-$ and $\Xi(\mathbf{t}; \lambda) = \sum_{j \in \mathbb{B}^+} t_j \Lambda_j$. First, we define a map $\Phi : \mathfrak{g} \rightarrow \mathfrak{g}[[\zeta]]$, defined by

$$\Phi : X = \sum_{I \in \mathbb{Z}} X_I \lambda^I \longmapsto X_\zeta := \sum_{I \in \mathbb{Z}} X_I \lambda^I \zeta^{|I|}$$

(where the sum over \mathbb{Z} is finite). Now for any $X = \sum_{I \in \mathbb{Z}} X_I \lambda^I$ (where the sum is finite), we can define the exponential of X_ζ by

$$\exp(X_\zeta) = \sum_{k \geq 0} \frac{1}{k!} \left[\sum_{I \in \mathbb{Z}} X_I \lambda^I \zeta^{|I|} \right]^k.$$

In the above equation, the product $X_I X_J$ is defined as the matrix product within the representation we chose. Then $e^{X_\zeta} \in \text{LG}(n)$, i.e. is an admissible loop. Indeed, it suffices to write

$$\exp(X_\zeta) = \sum_{I \in \mathbb{Z}} \lambda^I \sum_{k \geq 0} \frac{1}{k!} \sum_{\substack{I_1, \dots, I_k \in \mathbb{Z} \\ I_1 + \dots + I_k = I}} \zeta^{|I_1| + \dots + |I_k|} X_{I_1} \dots X_{I_k}.$$

Since $|I_1| + \dots + |I_k| \geq |I|$, it follows that e^{X_ζ} is an admissible loop; its inverse is e^{-X_ζ} .

We now apply that same construction to

$$\Xi(\mathbf{t}; \lambda) = \sum_{j \in \mathbb{B}^+} t_j \Lambda_j.$$

The map $\Phi : \mathfrak{g} \rightarrow \mathfrak{g}[[\zeta]]$ naturally extends to a map $\Phi : \mathfrak{g}[[\mathbf{t}]] \rightarrow \mathfrak{g}[[\mathbf{t}]] [[\zeta]]$. The Λ_j 's are not homogeneous, although, since $\mathfrak{g}_- \cap \mathfrak{g}^+ = \emptyset$, we can decompose each Λ_j through the homogeneous gradation with only positive powers of λ ; we write $\Lambda_j = \sum \Lambda_{j,I} \lambda^I$. Then

$$\Xi(\mathbf{t}; \lambda) = \sum_{I \geq 0} F_I(\mathbf{t}) \lambda^I, \quad F_I(\mathbf{t}) = \sum_{j \in \mathbb{B}^+} t_j \Lambda_{j,I}.$$

We assume $F_I(\mathbf{t})$ to be a finite sum, meaning that for a fixed $I \geq 0$, there is only finitely many Λ_j 's such that $\text{val}_\lambda \Lambda_j \leq I$. This will be realized by the matrix representations we choose (see Appendix A). We define

$$\Xi_\zeta(\mathbf{t}; \lambda) := \sum_{I \geq 0} F_I(\mathbf{t}) \lambda^I \zeta^{|I|} = \Phi(\Xi(\mathbf{t}; \lambda)).$$

Then there is no obstruction to defining

$$\begin{aligned} \exp(\Xi_\zeta(\mathbf{t}; \lambda)) &= \sum_{k \geq 0} \frac{1}{k!} \left[\sum_{I \geq 0} F_I(\mathbf{t}) \lambda^I \zeta^{|I|} \right]^k \\ &= \sum_{I \in \mathbb{Z}} \lambda^I \sum_{k \geq 0} \frac{1}{k!} \sum_{\substack{I_1, \dots, I_k \geq 0 \\ I_1 + \dots + I_k = I}} \zeta^{|I_1| + \dots + |I_k|} F_{I_1}(\mathbf{t}) \cdots F_{I_k}(\mathbf{t}). \end{aligned}$$

It is not hard to see that the coefficient in λ^I is a well defined power series in \mathbf{t} . Therefore, $e^{\Xi_\zeta(\mathbf{t}; \lambda)}$ is an admissible loop with inverse $e^{-\Xi_\zeta(\mathbf{t}; \lambda)}$.

Now let $X \in \mathfrak{g}_-$, i.e. with only negative powers of λ . Let us denote $\gamma = e^{X_\zeta}$ and $g(\mathbf{t}) = e^{\Xi_\zeta(\mathbf{t}; \lambda)}$; these loops have the form

$$\gamma = \sum_{I \geq 0} \gamma_{-I} \lambda^{-I} = \text{id} + \mathcal{O}(\lambda^{-1}), \quad g(\mathbf{t}) = \sum_{I \geq 0} g_I \lambda^I = \text{id} + \mathcal{O}(\lambda),$$

with $\text{val}_\zeta \gamma_{-I} \geq I$ and $\text{val}_\zeta g_I \geq I$. The Cauchy product of $g(\mathbf{t})$ and γ can be written as

$$\begin{aligned} J_\gamma(\mathbf{t}) := g(\mathbf{t}) \cdot \gamma &= \sum_{I \in \mathbb{Z}} \lambda^I \sum_{J \geq \max\{0, -I\}} g_{I+J} \gamma_{-J} \\ &= \sum_{I \geq 0} \lambda^I (g_I \gamma_0 + g_{I+1} \gamma_{-1} + \cdots) + \sum_{I < 0} \lambda^I (g_0 \gamma_{-I} + g_1 \gamma_{-I-1} + \cdots). \end{aligned}$$

We easily see, once again, that this is indeed an admissible loop in the sense of Equation (4.76). Moreover, if we define

$$J_\gamma^+ = \sum_{I \geq 0} \gamma_{-I} \lambda^I = \text{id} + \mathcal{O}(\lambda), \quad J_\gamma^- = \sum_{I \geq 0} g_I \lambda^{-I} = \text{id} + \mathcal{O}(\lambda^{-1}),$$

then $J_\gamma(\mathbf{t})$ admits two factorization:

$$J_\gamma(\mathbf{t}) = g(\mathbf{t}) \cdot \gamma = J_\gamma^-(\mathbf{t}) \cdot J_\gamma^+(\mathbf{t}). \quad (4.79)$$

The above equation will be important in the sequel for the Borodin–Okounkov formula. In what follows, we will write $X = X_\zeta$ and $\Xi(\mathbf{t}; \lambda) = \Xi_\zeta(\mathbf{t}; \lambda)$ when there is no ambiguity.

Laurent block matrices and formal loops

Following the definition of the formal loop group, let \mathcal{H}^n be the subspace of formal power series over \mathbb{Z} in a parameter λ , say $v = \sum_{I \in \mathbb{Z}} v_I \lambda^I$, with entries $v_I \in \mathbb{C}^n [[\zeta]]$ such that $\text{val}_\zeta v_I \geq |I|$, i.e.

$$\mathcal{H}^n = \left\{ v = \sum_{I \in \mathbb{Z}} v_I \lambda^I \mid v_I \in \mathbb{C}^n [[\zeta]], \text{val}_\zeta v_I \geq |I| \right\} \subset \mathbb{C}^n [[\lambda, \lambda^{-1}, \zeta]].$$

The space \mathcal{H}^n is spanned by the standard family of vectors $\{e_\alpha \lambda^I \mid \alpha \in \{1, \dots, n\}, I \in \mathbb{Z}\}$, where e_α is the column vector with 1 at the α -th position and 0 elsewhere. With respect to this basis, a vector $v = \sum_{I \in \mathbb{Z}} v_I \lambda^I \in \mathcal{H}^n$ can be identified with its coordinates in the basis $\{\lambda^I \mid I \in \mathbb{Z}\}$:

$$v \sim \begin{pmatrix} \vdots \\ v_{-1} \\ v_0 \\ v_1 \\ \vdots \end{pmatrix} \quad v_I = \sum_{A \geq |I|} \zeta^A \sum_{\alpha=1}^n v_{I,A}^\alpha e_\alpha \in \mathbb{C}^n [[\zeta]].$$

We denote by \mathcal{H}_+ and \mathcal{H}_- the subspaces consisting of vectors with only nonnegative and negative powers of λ respectively, i.e.

$$\mathcal{H}_+ = \left\{ v = \sum_{I \geq 0} v_I \lambda^I \right\} \simeq \mathbb{C}^n [[\lambda]], \quad \mathcal{H}_- = \left\{ v = \sum_{I < 0} v_k \lambda^I \right\} \simeq \lambda^{-1} \mathbb{C}^n [[\lambda^{-1}]].$$

We denote

$$\text{pr}_+ : \mathcal{H} \rightarrow \mathcal{H}_+, \quad \text{pr}_- : \mathcal{H} \rightarrow \mathcal{H}_-$$

the projections on \mathcal{H}_+ and \mathcal{H}_- respectively.

Given an admissible loop $\gamma = \sum_{I \in \mathbb{Z}} \gamma_I \lambda^I \in \text{LG}(n)$ we can define an action of γ on v by a Cauchy product:

$$\gamma v = \sum_{I \in \mathbb{Z}} \lambda^I \sum_{J \in \mathbb{Z}} \gamma_{I-J} v_J \in \mathcal{H}^n \quad (4.80)$$

It follows from Equation (4.80) that γ admits an infinite-dimensional “ $\mathbb{Z} \times \mathbb{Z}$ ” matrix representation as

$$\mathbf{L}(\gamma) := (\gamma_{I-J})_{I, J \in \mathbb{Z}} = \begin{pmatrix} \ddots & \ddots & \ddots & & & & \\ \ddots & \gamma_0 & \gamma_{-1} & \gamma_{-2} & & & \\ \ddots & \gamma_1 & \gamma_0 & \gamma_{-1} & \ddots & & \\ & \gamma_2 & \gamma_1 & \gamma_0 & \ddots & & \\ & & \ddots & \ddots & \ddots & & \end{pmatrix}. \quad (4.81)$$

The matrix $L(\gamma)$ we call the *Laurent matrix* of γ . We have $\gamma v = L(\gamma)v$, where the second product is the standard product of matrices. It follows that

$$\gamma \cdot \tilde{\gamma} = L(\gamma) \cdot L(\tilde{\gamma}),$$

making the Laurent matrix map into a homomorphism

$$\mathbf{L} : \text{LG}(n) \rightarrow \left\{ (\gamma_{I-J})_{I,J \in \mathbb{Z}} \mid \gamma_K \in \mathbb{C}[[\zeta]], \text{val}_\zeta \gamma_K \geq |K| \right\}$$

Notations. We shall emphasize the important distinction, hereafter, between capital indices I, J and small indices i, j . Capital indices I, J denote “block indices” in the sense that the coefficient γ_I of a loop $\gamma = \sum_{I \in \mathbb{Z}} \gamma_I \lambda^I$ is a matrix in $\gamma_I \in \text{GL}(n)$. On the other hand, small indices i, j will denote “usual indices” of a matrix, e.g. $(\gamma_I)_{i,j} \in \mathbb{C}$. For the Laurent matrix of a loop, the relation between the block indices and the usual indices is given in the following relation:

$$\mathbf{L}(\gamma)_{In+\alpha, Jn+\beta} = (\gamma_{I-J})_{\alpha+1, \beta+1}, \quad I, J \in \mathbb{Z}, \alpha, \beta \in \{0, \dots, n-1\}.$$

For any admissible loop $\gamma \in \text{LG}(n)$ of the form

$$\gamma = \text{id} + O(\lambda^{-1}),$$

we associate the subspace $W_\gamma \subset \mathcal{H}^n$ defined by $W_\gamma = \gamma(\mathcal{H}_+^n)$. Then W_γ satisfies the periodic condition $\lambda W \subset W$. Moreover, it is not hard to see, thanks to the form of γ , that the restriction of the projection $\text{pr}_+|_{W_\gamma} : W_\gamma \rightarrow \mathcal{H}_+^n$ is bijective. Conversely, one associate such a loop $\gamma = \text{id} + O(\lambda^{-1})$ to any subspace $W \subset \mathcal{H}^n$ which satisfies the condition

$$\lambda W \subset W, \quad \text{pr}_+|_W : W \xrightarrow{\sim} \mathcal{H}_+^n.$$

The set of all such subspaces is what we call the “big cell” associate to \mathcal{H}^n ; we denote Gr^0 . The notion of big cell was introduced by [SW85], although the idea was already present in [Sat81].

Hereafter we will identify any admissible loop $\gamma \in \text{LG}(n)$ of the form $\gamma = \text{id} + O(\lambda^{-1})$ with the element $W_\gamma \in \text{Gr}^0$ such that $W_\gamma = \gamma(\mathcal{H}_+)$. We identify W_γ with the $\mathbb{Z}_{\geq 0} \times \mathbb{Z}$ matrix that represents the map $\gamma|_{\mathcal{H}^n} : \mathcal{H}^n \rightarrow W_\gamma$, namely:

$$W_\gamma \sim (\gamma_{I-J})_{I \in \mathbb{Z}, J \geq 0} = \begin{pmatrix} \vdots & \vdots & \vdots & \ddots \\ \text{id} & \gamma_{-1} & \gamma_{-2} & \cdots \\ 0 & \text{id} & \gamma_{-1} & \cdots \\ 0 & 0 & \text{id} & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

In what follows, we will drop the superscript n and simply denote $\mathcal{H} = \mathcal{H}^n$.

Toeplitz and Hankel block matrices

Besides the Laurent matrix of a loop $\gamma = \sum_{I \in \mathbb{Z}} \gamma_I \lambda^I$, we introduce the two following block matrices; all of them are $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ matrices:

$$\mathbf{T}(\gamma) = (\gamma_{I-J})_{I, J \geq 0}, \quad \mathbf{H}(\gamma) = (\gamma_{I+J+1})_{I, J \geq 0}, \quad \tilde{\mathbf{H}}(\gamma) = (\gamma_{-I-J-1})_{I, J \geq 0}.$$

The matrix $\mathbf{T}(\gamma)$ we call the block Toeplitz matrix associated to γ the matrices $\mathbf{H}(\gamma)$ and $\tilde{\mathbf{H}}(\gamma)$ we call the Hankel block matrices. These matrices can be visualized as follows:

$$\mathbf{T}(\gamma) = \begin{pmatrix} \gamma_0 & \gamma_{-1} & \gamma_{-2} & & \\ \gamma_1 & \gamma_0 & \gamma_{-1} & \ddots & \\ \gamma_2 & \gamma_1 & \gamma_0 & \ddots & \\ & \ddots & \ddots & \ddots & \end{pmatrix}, \quad \mathbf{H}(\gamma) = \begin{pmatrix} \gamma_1 & \gamma_2 & \gamma_3 & & \\ \gamma_2 & \gamma_3 & \gamma_4 & \ddots & \\ \gamma_3 & \gamma_4 & \gamma_5 & \ddots & \\ & \ddots & \ddots & \ddots & \end{pmatrix}.$$

(The matrix $\tilde{\mathbf{H}}(\gamma)$ is the same as $\mathbf{H}(\gamma)$ but with negative indices). Now introduce the following involution on \mathcal{H} :

$$\begin{aligned} \iota : \mathcal{H} &\longrightarrow \mathcal{H} \\ v(\lambda) &\longmapsto \lambda^{-1}v(\lambda^{-1}). \end{aligned}$$

It is clear that $\iota^2 = \text{id}$ and that the restrictions $\iota : \mathcal{H}_+ \rightarrow \mathcal{H}_-$ and $\iota : \mathcal{H}_- \rightarrow \mathcal{H}_+$ are bijective. Then it is not hard to show that the Toeplitz and Hankel matrices are the matrix representations of the following endomorphisms:

$$\mathbf{T}(\gamma) = \text{pr}_+ \cdot \gamma|_{\mathcal{H}_+}, \quad \mathbf{H}(\gamma) = \text{pr}_+ \cdot \gamma \cdot \iota|_{\mathcal{H}_+}, \quad \tilde{\mathbf{H}}(\gamma) = \iota \cdot \text{pr}_- \cdot \gamma|_{\mathcal{H}_+}.$$

Now the Toeplitz and Hankel matrices are not homomorphisms, contrary to the Laurent matrix, yet the following lemma holds [BS99].

Lemma 4.4.6. *Given two loops γ_1, γ_2 , the following identity holds,*

$$\mathbf{T}(\gamma_1) \cdot \mathbf{T}(\gamma_2) = \mathbf{T}(\gamma_1 \cdot \gamma_2) - \mathbf{H}(\gamma_1) \cdot \tilde{\mathbf{H}}(\gamma_2).$$

In particular, $\mathbf{T}(\gamma_1) \cdot \mathbf{T}(\gamma_2) = \mathbf{T}(\gamma_1 \cdot \gamma_2)$ whenever $\mathbf{T}(\gamma_1)$ is upper triangular or $\mathbf{T}(\gamma_2)$ is lower triangular.

Proof. It suffices to use the definition of these matrices:

$$\begin{aligned} \mathbf{T}(\gamma_1) \cdot \mathbf{T}(\gamma_2) &= \text{pr}_+ \cdot \gamma_1 \cdot \text{pr}_+ \cdot \gamma_2|_{\mathcal{H}_+} \\ &= \text{pr}_+ \cdot \gamma_1 \cdot (\text{id} - \text{pr}_-) \cdot \gamma_2|_{\mathcal{H}_+} \\ &= \text{pr}_+ \cdot \gamma_1 \cdot \gamma_2|_{\mathcal{H}_+} - \text{pr}_+ \cdot \gamma_1 \cdot \iota \cdot \iota \cdot \text{pr}_- \cdot \gamma_2|_{\mathcal{H}_+} \end{aligned}$$

□

The Sato–Segal–Wilson tau function

In this section, we give a definition of a tau function associated to a loop $\gamma = \text{id} + O(\lambda^{-1})$, or equivalently to a point $W_\gamma \in \text{Gr}^0$, called the Sato–Segal–Wilson tau function, following [CW15a]¹². It originally appeared in a certain form in [Sat81] and [SW85]. One of the aims is to identify this tau function with the limit of certain Toeplitz determinant and then with a tau function of the g-DS hierarchy. We define the following maps:

$$\Omega_+ = \text{pr}_+ \cdot \gamma|_{\mathcal{H}_+} : \mathcal{H}_+ \rightarrow \mathcal{H}_+, \quad \Omega_- = \text{pr}_- \cdot \gamma|_{\mathcal{H}_+} : \mathcal{H}_+ \rightarrow \mathcal{H}_-$$

along with

$$h_W = \Omega_- \cdot (\Omega_+)^{-1} = \text{pr}_-|_W \cdot (\text{pr}_+|_W)^{-1} : \mathcal{H}_+ \rightarrow \mathcal{H}_-$$

These maps are well defined since, by hypothesis, $\text{pr}_+|_W$ is bijective.

Recall the power series

$$g(\mathbf{t}) = \exp(\Xi(\mathbf{t}; \lambda)) = \text{id} + O(\lambda).$$

Consider the following submatrices of the Laurent matrix of g ,

$$a = (g_{I-J})_{I, J \geq 0}, \quad b = (g_{I-J})_{I \geq 0, J < 0}, \quad d = (g_{I-J})_{I, J < 0}.$$

They are the matrices of the following maps,

$$a = \text{pr}_+ \cdot g|_{\mathcal{H}_+}, \quad b = \text{pr}_+ \cdot g|_{\mathcal{H}_-}, \quad d = \text{pr}_- \cdot g|_{\mathcal{H}_-}.$$

Notice that actually $a = g|_{\mathcal{H}_+}$ since by the form of g , we see that $g(\mathcal{H}_+) \subset \mathcal{H}_+$. The latter also implies that the map $\text{pr}_- \cdot g|_{\mathcal{H}_+}$ vanishes identically. Consequently, the loop g admits the following block matrix representation,

$$g = \begin{pmatrix} a & b \\ 0 & d \end{pmatrix}.$$

Definition 4.4.7 (Sato–Segal–Wilson tau function [SW85]). Let $W_\gamma \in \text{Gr}^0$ be a subspace and let $g = \exp(\Xi(\mathbf{t}; \lambda))$. We call the Sato–Segal–Wilson the formal power series

$$\begin{aligned} \hat{\tau}_{W_\gamma} &= \det \left(\text{id} + a^{-1} \cdot b \cdot h_{W_\gamma} \right) \\ &= \det \left(\text{id} + b \cdot h_{W_\gamma} \cdot a^{-1} \right). \end{aligned}$$

For now, it is unclear why the determinant that defines $\hat{\tau}_{W_\gamma}(\mathbf{t})$ is well defined. This will be solved by the following lemma.

Lemma 4.4.8. For any $X \in \mathfrak{g}_-$ and associated loop $\gamma = e^X$, the following identity holds,

$$bh_{W_\gamma} a^{-1} = -\mathbf{H}(g^{-1}\gamma)\tilde{\mathbf{H}}(\gamma^{-1}g)$$

¹²We shall emphasize that the terminology *tau function* here anticipates the fact that it will indeed be identified with a tau function of the DS hierarchy, see Section 4.4.2.

Proof. First, we write

$$\begin{aligned} a^{-1} &= \text{pr}_+ g^{-1} |_{\mathcal{H}_+}, & b &= \text{pr}_+ g |_{\mathcal{H}_-}, \\ \Omega_+^{-1} &= \text{pr}_+ \gamma^{-1} |_{\mathcal{H}_+}, & \Omega_- &= \text{pr}_- \gamma |_{\mathcal{H}_+}. \end{aligned}$$

Therefore,

$$bh_{W_\gamma} a^{-1} = \text{pr}_+ g \text{pr}_- \gamma \text{pr}_+ \gamma^{-1} g^{-1} |_{\mathcal{H}_+}.$$

On the other hand, since $\iota^2 = \text{id}$, we have

$$\mathbf{H}(g\gamma) \tilde{\mathbf{H}}(\gamma^{-1} g^{-1}) = \text{pr}_+ g \gamma \text{pr}_- \gamma^{-1} g^{-1} |_{\mathcal{H}_+}.$$

Therefore, it suffices to prove that (hereafter we underline the summands for clarity)

$$\underline{\text{pr}_+ g \gamma \text{pr}_- \gamma^{-1} g^{-1}} + \underline{\text{pr}_+ g \text{pr}_- \gamma \text{pr}_+ \gamma^{-1} g^{-1}} |_{\mathcal{H}_+} = 0.$$

Then it suffices to use $\text{pr}_- = \text{id} - \text{pr}_+$ and the fact that $\Omega_+ = \text{pr}_+ \gamma |_{\mathcal{H}_+}$ and $\Omega_+^{-1} = \text{pr}_+ \gamma^{-1} |_{\mathcal{H}_+}$:

$$\begin{aligned} & \underline{\text{pr}_+ g \gamma \text{pr}_- \gamma^{-1} g^{-1}} + \underline{\text{pr}_+ g (\text{id} - \text{pr}_+) \gamma \text{pr}_+ \gamma^{-1} g^{-1}} |_{\mathcal{H}_+} \\ &= \underline{\text{pr}_+ g \gamma \text{pr}_- \gamma^{-1} g^{-1}} + \underline{\text{pr}_+ g \gamma \text{pr}_+ \gamma^{-1} g^{-1}} - \underline{\text{pr}_+ g \text{pr}_+ \gamma \text{pr}_+ \gamma^{-1} g^{-1}} |_{\mathcal{H}_+} \\ &= \underline{\text{pr}_+ g \gamma \gamma^{-1} g^{-1}} - \underline{\text{pr}_+ g \Omega_+ \Omega_+^{-1} g^{-1}} |_{\mathcal{H}_+} \\ &= \text{id} - \text{id} |_{\mathcal{H}_+} = 0. \end{aligned}$$

□

Now thanks to Lemma 4.4.8, the Sato–Segal–Wilson tau function can be written as

$$\hat{\tau}_{W_\gamma} = \det \left(1 - \mathbf{H}(g\gamma) \tilde{\mathbf{H}}(\gamma^{-1} g^{-1}) \right).$$

Notice that because the loops $g(t)$ and γ are admissible, the entries of $\mathbf{H}(g\gamma)_{I,J}$ and $\tilde{\mathbf{H}}(\gamma^{-1} g^{-1})_{I,J}$ have valuation greater than $|I + J + 1|$. In particular, both these Hankel matrices are Hilbert–Schmidt operators in the sense that the norm

$$\|\mathbf{H}(g\gamma)\|^2 = \sum_{I,J \geq 0} |\mathbf{H}(g\gamma)_{IJ}|^2$$

converges for the topology induced by the valuation (it suffices to see that the partial sums form a Cauchy sequence). Then the product of two Hilbert–Schmidt operators is of trace class and the determinant $\det(\text{id} + T)$ with T trace-class is a well defined Fredholm determinant. We refer to [Sim05] for more details.

The Borodin–Okounkov formula and a formal Szegő–Widom theorem

In this section, we identify the Sato–Segal–Wilson with the limit of the Toeplitz determinant of the loop $g\gamma$ by means of the Borodin–Okounkov formula and a formal version of the Szegő–Widom theorem; we still follow [CW15b].

Besides the infinite $\mathbb{Z}_{\geq 0} \times \mathbb{Z}_{\geq 0}$ Toeplitz matrix, we associate to a loop $\phi = \sum_{I \in \mathbb{Z}} \phi \lambda^I$ the following matrix:

$$\mathbf{T}_N(\phi) = (\phi_{I-J})_{I,J=0}^N = \begin{pmatrix} \phi_0 & \phi_{-1} & \cdots & \phi_{-N} \\ \phi_1 & \phi_0 & \cdots & \phi_{-N+1} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_N & \phi_{N-1} & \cdots & \phi_0 \end{pmatrix},$$

called the N -truncated block Toeplitz matrix. We denote its determinant

$$\mathbf{D}_N(\phi) := \det \mathbf{T}_N(\phi).$$

Moreover, we denote $p_N : \mathcal{H}_+ \rightarrow \mathcal{H}_+$ the projection on the subspace of power series of the form $\sum_{I=0}^N v_I \lambda^I$ and $q_N = \text{id} - p_N$. Finally, if ϕ is admissible, we denote

$$\mathbf{Z}(\phi) = \det \left(\text{id} - \mathbf{H}(\phi) \tilde{\mathbf{H}}(\phi^{-1}) \right),$$

which, as noted at the end of the previous section, is a well defined Fredholm determinant. We can now present the celebrated Borodin–Okounkov formula as it appeared in [BW00].

Theorem 4.4.9 (Borodin–Okounkov formula [BO00]). *Let $\phi \in \text{LG}(n)$ be an admissible loop such that $\phi = \phi_+ \phi_- = \psi_- \psi_+$ admits two distinct factorizations such that*

$$\psi_-(\infty) = \phi_-(\infty) = \text{id}, \quad \psi_+(0) = \text{id} + N_+, \quad \phi_+ = \text{id} + M_+,$$

where N_+ and M_+ are strictly lower triangular matrices. Then for any $N \geq 1$,

$$\mathbf{D}_N(\phi) = \mathbf{Z}(\phi) \det(\text{id} - \mathbf{K}_N(\phi)),$$

where $\mathbf{K}_N(\phi) = q_N \mathbf{K}(\phi) q_N$ and

$$\mathbf{K}(\phi) = \mathbf{H}(\phi_- \psi_+^{-1}) \tilde{\mathbf{H}}(\psi_-^{-1} \phi_+).$$

Notice that $\mathbf{K}_N(\phi)$ is an infinite matrix but its determinant is a well defined Fredholm determinant for the same reasons $\mathbf{Z}(\phi)$ is. We refer to [CW15b] (Theorem 2.3) for a proof of Theorem 4.4.9 which is adapted to the present case of formal matrix-valued power series; the proof follows that of [BW00]. As a corollary, we obtain a formal version of the Szegő–Widom theorem (see [Wid76] for example).

Corollary 4.4.10. *Under the conditions of Theorem 4.4.9, we have*

$$\lim_{N \rightarrow \infty} \mathbf{D}_N(\phi) = \mathbf{Z}(\phi) = \det(\text{id} - \mathbf{H}(\phi) \tilde{\mathbf{H}}(\phi^{-1})). \quad (4.82)$$

Proof. Using the Borodin–Okounkov formula,

$$\log Z(\phi) - \log \mathbf{D}_N(\phi) = \log (\det(\text{id} - \mathbf{K}_N(\phi))).$$

Then, it is not hard to see that the entries of $\mathbf{K}_N(\phi)$ have a valuation greater or equal to Nh , so that

$$\text{val}(\log Z(\phi) - \log \mathbf{D}_N(\phi)) \geq Nh,$$

which implies that Equation (4.82) holds. \square

Now to identify the Sato–Segal–Wilson tau function with the limit of the Toeplitz determinant of $J_\gamma = g^{-1}\gamma$, it suffices to show that J_γ satisfies the assumptions of the Borodin Okounkov theorem. First, as noted at the beginning of Section 4.4.1,

$$J_\gamma = g\gamma = J_\gamma^- J_\gamma^+$$

Therefore, J_γ admits two different factorizations, with $g, J_\gamma^+ \in \text{id} + \mathcal{O}(\lambda)$ and $\gamma, J_\gamma^- \in \text{id} + \mathcal{O}(\lambda^{-1})$. So J_γ satisfies the assumptions of Theorem 4.4.9.

Corollary 4.4.11 ([CW15b]). *For any loop $\gamma = \text{id} + \mathcal{O}(\lambda^{-1})$ and $J_\gamma(\mathbf{t}) = g(\mathbf{t}) \cdot \gamma$, the following identity holds,*

$$\hat{\tau}_{W_\gamma}(\mathbf{t}) = \lim_{N \rightarrow \infty} \mathbf{D}_N(J_\gamma(\mathbf{t})) =: \mathbf{D}_\infty(J_\gamma(\mathbf{t})).$$

4.4.2 Relation with the tau functions of the DS hierarchies

The matrix wave function

We start this section by defining the notion of matrix wave (or Baker-Akhiezer) function associated to a loop $\gamma = \text{id} + \mathcal{O}(\lambda^{-1})$ and the power series $g(\mathbf{t}) = \exp(\Xi(\mathbf{t}))$ as defined in [Dic03] (see Definition 12.1.3) and given again in [CW15b] (see Definition 2.9).

Definition 4.4.12. Let $\gamma = \text{id} + \mathcal{O}(\lambda^{-1})$ and the associated point of the big cell $\text{Gr}^0 \ni W_\gamma = \gamma(\mathcal{H}_+)$. We say that a loop $w(\mathbf{t}) \in \text{LG}(n)$ is a wave function (or Baker-Akhiezer) function associated to γ if

- (i) For any $v \in \mathcal{H}_+$, we have $w(\mathbf{t})v \in W_\gamma$;
- (ii) $\text{pr}_+ \cdot g(\mathbf{t}) \cdot w(\mathbf{t}) = \text{id}_{\mathcal{H}}$.

This definition is to be compared with the notion of wave function in KP theory and the bilinear identity developed in Section 3.3.

Lemma 4.4.13. *For any loop $\gamma = \text{id} + \mathcal{O}(\lambda^{-1})$, the following function is a wave function associated to γ :*

$$w(\mathbf{t}) = g^{-1}(\mathbf{t})J_\gamma^-(\mathbf{t}).$$

Proof. For point (i), it amounts to proving that for any $v \in \mathcal{H}_+$, we have $\gamma^{-1}w(\mathbf{t})v \in \mathcal{H}_+$. Indeed,

$$\begin{aligned}\gamma^{-1}w(\mathbf{t})v &= \gamma^{-1}g^{-1}(\mathbf{t})J_Y^-(\mathbf{t})v \\ &= (J_Y^-(\mathbf{t})J_Y^+(\mathbf{t}))^{-1}J_Y^-(\mathbf{t})v \\ &= (J_Y^+(\mathbf{t}))^{-1}v,\end{aligned}$$

the latter belongs to \mathcal{H}_+ since $(J_Y^+(\mathbf{t}))^{-1} = \text{id} + \mathcal{O}(\lambda)$. Point (ii) is immediate since

$$\text{pr}_+ \cdot g(\mathbf{t})w(\mathbf{t}) = \text{pr}_+ \cdot J_Y^{-1}(\mathbf{t}) = \text{id}.$$

□

Then in [CW15a], and in analytical context which adapts to our formal context, the authors related the Sato–Segal–Wilson tau function with the wave function via the following theorem (see Theorem 2.11 and Corollary 2.12 of [CW15a]).

Theorem 4.4.14 ([CW15a]). *Given a point $W_Y \in \text{Gr}^0$ and its associated wave function $w(\mathbf{t})$, the following identity holds, for any $j \in \mathbb{E}^+$,*

$$\frac{\partial \log \hat{\tau}_{W_Y}}{\partial t_j} = - \oint \frac{d\lambda}{2i\pi} \text{tr} \left(\frac{\partial w}{\partial \lambda} w^{-1} \Lambda_j \right), \quad (4.83)$$

where the contour integral is to be understood as the sum of all residues; the trace is to be understood as the trace of the representation we chose.

Theorem 4.4.14 can be understood as a generalization of 3.3.7 in KP theory. Indeed, in the latter the wave function is scalar and related to the tau function by

$$w(\mathbf{t}; \lambda) = \frac{\tau(\mathbf{t} - [\lambda^{-1}])}{\tau(\mathbf{t})} g^{-1}(\mathbf{t}), \quad g(\mathbf{t}) = \exp \left(\sum_{i \geq 1} t_i \lambda^i \right).$$

Let us verify that in this scalar case the KP wave function satisfies Equation (4.83). First, Equation (4.83) is equivalent to

$$\frac{\partial \log \hat{\tau}_{W_Y}}{\partial t_j} = - \oint \frac{d\lambda}{2i\pi} \text{tr} \left(\frac{\partial w}{\partial \lambda} w^{-1} \frac{\partial g}{\partial t_j} g^{-1} \right)$$

Then, the trace is the identity, so it reduces to

$$\begin{aligned}\oint \frac{d\lambda}{2i\pi} \frac{\partial w}{\partial \lambda} w^{-1} \frac{\partial g^{-1}}{\partial t_j} g &= - \oint \frac{d\lambda}{2i\pi} \frac{\partial \tau(\mathbf{t} - [\lambda^{-1}])}{\partial \lambda} \frac{1}{\tau(\mathbf{t} - [\lambda^{-1}])} \lambda^j \\ &= - \oint \sum_{i \geq 1} \frac{\partial \log \tau(\mathbf{t} - [\lambda^{-1}])}{\partial t_i} \lambda^{j-i-1} = - \frac{\partial \log \tau}{\partial t_j}\end{aligned}$$

Sato–Segal–Wilson vs. Drinfeld–Sokolov tau functions

We now make the link between the Sato–Segal–Wilson tau function and the Drinfeld–Sokolov tau function as defined by Wu [Wu17] in (4.74): given a dressing operator Θ ,

$$\frac{\partial \log \tau_\Theta}{\partial t_j} = (\Theta \Lambda_j \Theta^{-1})_c.$$

Recall that the dressing operator $\Theta = e^U e^\Omega = \text{id} + O(\lambda^{-1}) \in \widehat{\mathfrak{g}}^-[[\mathbf{t}]]$. The following lemma connects the wave function $w(\mathbf{t})$ and the dressing operator $\Theta(\mathbf{t})$ in the case where the initial data $\gamma = \text{id} + O(\lambda^{-1})$ is taken as $\gamma = \Theta^{-1}|_{t=0}$ [CW15a].

Lemma 4.4.15. *Let $\Theta \in \widehat{\mathfrak{g}}^-[[\mathbf{t}]]$ be an arbitrary solution of the \mathfrak{g} -DS hierarchy and let $\gamma = \Theta^{-1}|_{t=0} = \text{id} + O(\lambda^{-1})$. Then the corresponding wave function is given by*

$$w(\mathbf{t}) = g^{-1}(\mathbf{t})\Theta^{-1}(\mathbf{t}).$$

Moreover, if $\tau_{W_\gamma}(\mathbf{t})$ is the Sato–Segal–Wilson tau function associated to γ , then Equation (4.83) becomes the following,

$$\frac{\partial \log \hat{\tau}_{W_\gamma}}{\partial t_j} = \oint \frac{d\lambda}{2i\pi} \text{tr} \left(\frac{\partial \Theta}{\partial \lambda} \Lambda_j \Theta^{-1} \right) \quad (4.84)$$

Proof. First, we clearly have $\text{pr}_+ g(\mathbf{t})w(\mathbf{t}) = \text{pr}_+ \Theta^{-1}(\mathbf{t}) = \text{id}$. So it remains to prove that for any $v \in \mathcal{H}_+$, we have $\gamma^{-1}w(\mathbf{t})v \in \mathcal{H}_+$. Let us denote $\hat{w}(\mathbf{t}) = \gamma^{-1}w(\mathbf{t})$, then

$$\begin{aligned} \partial_{t_j} w &= -g^{-1} \Lambda_j \Theta^{-1} - g^{-1} \Theta^{-1} (\partial_{t_j} \Theta) \Theta^{-1} \\ &= -g^{-1} \Theta^{-1} (\Theta \Lambda_j \Theta^{-1}) + g^{-1} \Theta^{-1} (\Theta \Lambda_j \Theta^{-1})_- \\ &== -w(\Theta \Lambda_j \Theta^{-1})_+. \end{aligned} \quad (4.85)$$

Therefore, $\hat{w}^{-1} \partial_{t_j} \hat{w} = w^{-1} \partial_{t_j} w = -(\Theta \Lambda_j \Theta^{-1})_+ \in \mathfrak{g}_+[[\mathbf{t}]]$. Moreover the initial condition of $\hat{w}(\mathbf{t})$ reads

$$\hat{w}(\mathbf{t})|_{t=0} = \gamma^{-1} (g(\mathbf{t}) \Theta^{-1}(\mathbf{t}))|_{t=0} = \gamma^{-1} \gamma = 1, \quad (4.86)$$

since $g(\mathbf{t})|_{t=0} = e^0 = 1$ and $\gamma = \Theta^{-1}(\mathbf{t})|_{t=0}$. Now, Equations (4.85) and (4.86) together imply that $\hat{w}(\mathbf{t})$ contains only nonnegative powers of λ , hence the result. Then to prove Equation (4.84) it suffices to use the analyticity of g , the commutativity within the trace and an integration by part:

$$\begin{aligned} \frac{\partial \log \hat{\tau}_{W_\gamma}}{\partial t_j} &= - \oint \frac{d\lambda}{2i\pi} \text{tr} \left(\frac{\partial g^{-1} \Theta^{-1}}{\partial \lambda} \Theta g \Lambda_j \right) \\ &= - \oint \frac{d\lambda}{2i\pi} \text{tr} \left(\frac{\partial g^{-1}}{\partial \lambda} g \Lambda_j + g^{-1} \frac{\partial \Theta^{-1}}{\partial \lambda} \Theta g \Lambda_j \right) \\ &= - \oint \frac{d\lambda}{2i\pi} \text{tr} \left(\frac{\partial \Theta^{-1}}{\partial \lambda} \Theta \Lambda_j \right) \\ &= \oint \frac{d\lambda}{2i\pi} \text{tr} \left(\frac{\partial \Theta}{\partial \lambda} \Lambda_j \Theta^{-1} \right) \end{aligned}$$

□

We can now state the main theorem of this section.

Theorem 4.4.16 ([CW15a]). *Let Θ be a solution of the \mathfrak{g} -DS hierarchy and let τ_Θ be the associated tau function. Let $\gamma = \Theta^{-1}|_{t=0} = \text{id} + O(\lambda^{-1})$ and let $\hat{\tau}_{W_\gamma}$ be the associated Sato–Segal–Wilson tau function. Then*

$$\log \tau_\Theta = \kappa \log \hat{\tau}_{W_\gamma},$$

where κ is defined by $(Y, Z)_0 = \kappa \cdot \text{tr}(\pi(Y)\pi(Z))$ for all $Y, Z \in \mathfrak{g}$.

We proceed by two lemmas; we give the proofs of [CW15a].

Lemma 4.4.17. *For any $A, B \in \mathfrak{g}$, the following identity holds,*

$$[A, B]_c = \kappa \oint \frac{d\lambda}{2i\pi} \text{tr}((\partial_\lambda A)B).$$

Proof. It suffices to prove the identity on elements of the form $X\lambda^p, Y\lambda^q \in \mathfrak{g}$ with $X, Y \in \mathfrak{g}$ and $p, q \in \mathbb{Z}$. By the definition (4.15), we have

$$[X\lambda^p, Y\lambda^q]_c = \kappa p \delta_{p+q, 0} \text{tr}(XY).$$

On the other hand,

$$\oint \frac{d\lambda}{2i\pi} \text{tr}((\partial_\lambda X\lambda^p)Y\lambda^q) = \oint \frac{d\lambda}{2i\pi} p\lambda^{p+q-1} \text{tr}(XY) = p\delta_{p+q, 0} \text{tr}(XY).$$

□

Lemma 4.4.18. *For any $X \in \mathfrak{g}_-$ and $Y \in \mathfrak{g}$, the following identity holds,*

$$(e^X Y e^{-X})_c = \kappa \oint \frac{d\lambda}{2i\pi} \text{tr}((\partial_\lambda e^X) Y e^{-X}).$$

Proof. To simplify notations, we write $X' = \partial_\lambda X$ and $\langle X \rangle = \oint \frac{d\lambda}{2i\pi} \text{tr} X$. First, we have

$$\partial_\lambda e^X = \sum_{p \geq 0} \frac{1}{p!} \sum_{q=0}^{p-1} X^q X' X^{p-q-1}.$$

Let us denote $\langle (\partial_\lambda e^X) Y e^{-X} \rangle = \sum_{m \geq 1} R_m$ with

$$R_m = \left\langle \sum_{k=0}^{m-1} \frac{(-1)^k}{(m-k)!k!} \left[\sum_{j=0}^{m-k-1} X^j X' X^{m-k-j-1} \right] Y X^k \right\rangle.$$

We rewrite R_m as

$$\begin{aligned}
 R_m &= \left\langle \sum_{k=0}^{m-1} \frac{(-1)^k}{(m-k)!k!} X' \sum_{j=0}^{m-k-1} X^j Y X^k \right\rangle \\
 &= \frac{1}{m!} \left\langle X' \sum_{j=0}^{m-1} \sum_{k=0}^{m-j-1} (-1)^k \binom{m}{k} X^j Y X^{m-1-j} \right\rangle \\
 &= \frac{1}{m!} \left\langle X' \sum_{j=0}^{m-1} (-1)^{m-1-j} \binom{m-1}{m-1-j} X^j Y X^{m-1-j} \right\rangle \\
 &= \frac{1}{m!} \langle X' (\text{ad}_X)^{m-1} Y \rangle,
 \end{aligned}$$

where in the third equation we have used the recursion relation $\binom{m-1}{k} - \binom{m}{k+1} = -\binom{m-1}{k+1}$. By Lemma 4.4.17, it follows that

$$\begin{aligned}
 \langle (\partial_\lambda e^X) Y e^{-X} \rangle &= \frac{1}{\kappa} \sum_{m \geq 1} \frac{1}{m!} [X, (\text{ad}_X)^{m-1} Y]_c \\
 &= \frac{1}{\kappa} \sum_{m \geq 1} \frac{1}{m!} ((\text{ad}_X)^m Y)_c \\
 &= \frac{1}{\kappa} (e^{\text{ad}_X} Y)_c = \frac{1}{\kappa} (e^X Y e^{-X})_c
 \end{aligned}$$

□

Proof of Theorem 4.4.16. Now it suffices to use the definition (4.74) of Wu's tau function for the \mathfrak{g} -DS hierarchy, Lemma 4.4.18 and Equation (4.84), for all $j \in \mathbb{E}^+$,

$$\begin{aligned}
 \frac{\partial \log \tau_\Theta}{\partial t_j} &= (\Theta \Lambda_j \Theta^{-1})_c \\
 &= \oint \frac{d\lambda}{2i\pi} \text{tr} \left(\frac{\partial \Theta}{\partial \lambda} \Lambda_j \Theta^{-1} \right) \\
 &= \kappa \frac{\partial \log \hat{\tau}_{W_Y}}{\partial t_j}.
 \end{aligned}$$

Therefore, $\log \tau_\Theta$ and $\log \hat{\tau}_{W_Y}$ differ from a linear function of the variables t_j , which, in our framework, amounts to their coincidence. □

It remains to derive Theorem 4.4.1 (which we stated at the beginning of Section 4.4) from Theorem 4.4.16 (which we just proved). Let us first recall Theorem 4.4.1; notice that from now on we denote $\det = \mathbf{D}_\infty$.

Theorem 4.4.19 ([CW15a]). *Let $X \in \mathfrak{g}_-$ and let $\Xi(\mathbf{t}; \lambda) = \sum_{j \in \mathbb{E}^+} t_j \Lambda_j$. Then*

$$\tau_X := \left[\det \left(e^{\Xi(\mathbf{t}; \lambda)} e^X \right) \right]^\kappa,$$

is a tau function of the \mathfrak{g} -DS hierarchy.

First, since we chose $J_\gamma = g\gamma$, by Corollary 4.4.11, we have

$$\hat{\tau}_{W_\gamma} = \lim_{N \rightarrow \infty} \mathbf{D}_N \left(e^{\Xi(t;\lambda)} e^X \right) =: \det \left(e^{\Xi(t;\lambda)} e^X \right).$$

Moreover, $\gamma = \Theta^{-1}|_{t=0}$ and, as noted at the beginning of Section 4.4, for any $X \in \mathfrak{g}_-$ there exists Θ such that $\Theta|_{t=0} = e^{-X}$. Therefore, if we denote Θ_X the solution associated to the initial data $\Theta_X|_{t=0} = e^{-X}$, we can write

$$\log \tau_X := \log \tau_{\Theta_X} = \kappa \log \tau_{W_\gamma} = \kappa \det \left(e^{\Xi(t;\lambda)} e^X \right).$$

In particular,

$$\tau_X = \left[\det \left(e^{\Xi(t;\lambda)} e^X \right) \right]^\kappa.$$

4.5 Polynomial tau functions of the DS hierarchies and Schur polynomials

4.5.1 Young diagrams expansion of the tau function

Preliminaries on notations

We denote \mathbb{Y} the set of integer partitions ('Y' for Young diagrams; see below), i.e. the set of decreasing finite sequences of positive integers, i.e.

$$\mathbb{Y} = \{ \mu = (\mu_1, \dots, \mu_r) \subset \mathbb{Z}_{\geq 1} \mid r \geq 0, \mu_i \geq \mu_{i+1} \}$$

We denote $\ell(\mu) = r$ and call it the *length* of μ . If $\ell(\mu) = 0$ the partition is denoted $\mu = \emptyset$ and called the *empty partition*. We denote by

$$|\mu| = \mu_1 + \dots + \mu_{\ell(\mu)}$$

the *weight* of μ , i.e. the integer of which μ is a partition.

To any partition we associate a *Young diagram* as follows. Given a partition $\mu = (\mu_1 \geq \dots \geq \mu_{\ell(\mu)})$, its Young diagram is a table of squares, all flushed to the left, with $\ell(\mu)$ rows and where row i contains μ_i squares; see Equation (4.87) for an example.

Using the Young diagram of a partition, we define its *Frobenius notation* as follows. First, denote $d(\mu)$ is the number of squares in the main diagonal of the Young diagram of μ (i.e. the squares with coordinates (i, i) , $i = 1, \dots, d(\mu)$). Denote k_i the number of squares in row i that are strictly to the right of square (i, i) ; denote l_i is the number of squares in column i that are strictly below square (i, i) . The integers k_i, l_i and $d(\mu)$ determine μ ; we denote

$$\mu = (\mu_1, \dots, \mu_{\ell(\mu)}) = (k_1, \dots, k_{d(\mu)} \mid l_1, \dots, l_{d(\mu)})$$

For example, the partition $\mu = (5441)$ has length $\ell(\mu) = 4$, weight $|\mu| = 14$, and can be written in the Frobenius notation as $\mu = (421|310)$, as illustrated below.

$$v = (5441) \quad \leftrightarrow \quad \begin{array}{cccc} \blacksquare & \square & \square & \square \\ \square & \blacksquare & \square & \square \\ \square & \square & \blacksquare & \square \\ \square & \square & \square & \square \end{array} \begin{array}{l} (4) \\ (2) \\ (1) \\ (1) \end{array} \quad \leftrightarrow \quad v = (421|310) \quad (4.87)$$

Let $X \in \mathfrak{g}_-$ and recall $\Xi(\mathbf{t}; \lambda) = \sum_{j \in \mathbb{B}^+} t_j \Lambda_j$. Let $\gamma = e^X$ and $g(\mathbf{t}) = e^{\Xi(\mathbf{t}; \lambda)}$. Let their Laurent matrices be denoted by

$$r := \mathbf{L}(\gamma), \quad s(\mathbf{t}) := \mathbf{L}(g(\mathbf{t})).$$

We introduce the following determinants.

Definition 4.5.1. Let a partition $\mu \in \mathbb{Y}$. We call the *Plücker coordinates* of γ with respect to the representation $\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(n)$ the determinant

$$r_\mu := \det(r_{i-\mu_i-1; j-1})_{i, j \in [1, \ell(\mu)]} \in \mathbb{C}. \quad (4.88)$$

We call the *generalized Schur polynomials* of type (\mathfrak{g}, π) the determinant

$$s_\mu(\mathbf{t}) := \det(s_{i-1; j-\mu_j-1})_{i, j \in [1, \ell(\mu)]} \in \mathbb{C}[\mathbf{t}]. \quad (4.89)$$

We also set $r_\emptyset = s_\emptyset = 1$. For the adjoint representation $\pi(X) = \text{ad}_X \in \text{End}(\mathfrak{g})$, we call $s_\mu(\mathbf{t})$ the *intrinsic Schur polynomial* of type \mathfrak{g} .

Moreover, we define matrices $D_{IJ}(\mathbf{t})$ and Z_{IJ} , for $I, J \geq 0$, by

$$\frac{\text{id} - e^{\Xi(\mathbf{t}; \lambda)} e^{-\Xi(\mathbf{t}; \mu)}}{\lambda - \mu} = \sum_{I, J=0}^{\infty} D_{IJ}(\mathbf{t}) \lambda^{I+1} \mu^{J+1}, \quad (4.90)$$

$$\frac{\text{id} - e^{X(\lambda)} e^{-X(\mu)}}{\lambda - \mu} = \sum_{I, J=0}^{\infty} Z_{IJ} \lambda^{-I-1} \mu^{-J-1}. \quad (4.91)$$

The above equations are to be understood as formal expansion in the parameters λ, μ . Define $s_{(i|j)}, r_{(i|j)}$, for $i, j \geq 0$, via

$$\begin{aligned} (D_{IJ})_{ab} &= s_{(nI+a-1|nJ+n-b)}, \\ (Z_{IJ})_{ab} &= r_{(I+n-a|nJ+b-1)}, \end{aligned}$$

where $a, b \in \{1, \dots, n\}$. We call Z_{IJ} the *matrix-valued affine coordinates* and $r_{X(i|j)}$ the *affine coordinates*. These matrices were introduced in [BY17] for case of $\mathfrak{sl}(2)$.

Young diagram expansion

The aim of this section is to prove the following theorem.

Theorem 4.5.2 ([CDY18]). *The tau function $\tau_X(\mathbf{t}) = [\det(g(\mathbf{t}) \cdot \gamma)]^k$ of the \mathfrak{g} -DS hierarchy (cf. Theorem 4.4.1) admits the following Young diagrams expansion,*

$$\begin{aligned} \tau_X(\mathbf{t})^{\frac{1}{k}} &= \sum_{\mu \in \mathbb{Y}} s_\mu(\mathbf{t}) \cdot r_\mu & (4.92) \\ &= 1 + s_{\square} r_{\square} + s_{\begin{smallmatrix} \square \\ \square \end{smallmatrix}} r_{\begin{smallmatrix} \square \\ \square \end{smallmatrix}} + s_{\begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix}} r_{\begin{smallmatrix} \square & \square \\ \square & \square \end{smallmatrix}} + s_{\begin{smallmatrix} \square & \square \\ \square & \square \\ \square & \square \end{smallmatrix}} r_{\begin{smallmatrix} \square & \square \\ \square & \square \\ \square & \square \end{smallmatrix}} + \dots \end{aligned}$$

Moreover, if $\mu = (k_1, \dots, k_{d(\mu)} \mid l_1, \dots, l_{d(\mu)})$ is the Frobenius notation of μ and $l = l_1 + \dots + l_{d(\mu)}$, then $s_\mu(\mathbf{t})$ and r_μ have the following expressions

$$s_\mu(\mathbf{t}) = \det \left(s_{(k_i | l_j)} \right)_{i,j \in [1, d(\mu)]}, \quad r_\mu = (-1)^l \det \left(r_{(k_i | l_j)} \right)_{i,j \in [1, d(\mu)]}. \quad (4.93)$$

Remark 4.5.3. Notice that if we grade Young diagrams by setting $\deg \mu = |\mu| = \mu_1 + \dots + \mu_{\ell(\mu)}$ (i.e. the integer of which μ is a partition), then Equation (4.92) is a well defined power series for this grading: each term of a given degree is a finite sum of $s_\mu r_\mu$'s and each $s_\mu r_\mu$ is a finite determinant. Therefore, besides all that was said in the previous sections, Equation (4.92) gives an intrinsic formal expansion of the tau function.

Remark 4.5.4. In the A_n case, if one chooses the fundamental representation $A_n \simeq \mathfrak{sl}(n+1)$, then the Schur polynomials $s_\mu(\mathbf{t})$ are exactly the traditional Schur polynomials [Mac95] but with the constraint $t_{k(n+1)} = 0$.

Remark 4.5.5. Once Theorem 4.5.2 is proven, formula (4.93) is due to Giambelli type formula, see e.g. [BY17, EH11, Mac95].

To prove Theorem 4.5.2 we will proceed in several lemmas. The first one is a well known result, although we wanted to present a rigorous proof to the reader.

Lemma 4.5.6. *Let $\mathbb{Y}^{(m)}$ be the following subset of \mathbb{Y} ,*

$$\mathbb{Y}^{(m)} = \{ \mu = (\mu_1 \geq \mu_2 \geq \dots) \in \mathbb{Y} \mid \mu_1 \leq m, \ell(\mu) \leq m + 1 \}.$$

Let two rectangular matrices, S of size $(m+1) \times (2m+1)$ and R of size $(2m+1) \times (m+1)$, of the specific form

$$S = \left(\begin{array}{ccc|cccc} S_{0,-m} & \cdots & S_{0,-1} & 1 & 0 & \cdots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ S_{m,-m} & \cdots & S_{m,-1} & S_{m,0} & \cdots & S_{m,m-1} & 1 \end{array} \right), \quad R = \left(\begin{array}{ccc|ccc} R_{-m,0} & \cdots & R_{-m,m} \\ \vdots & & \vdots \\ R_{-1,0} & & R_{-1,m} \\ \hline 1 & \ddots & R_{0,m} \\ 0 & \ddots & \vdots \\ \vdots & \ddots & 1 & R_{m-1,m} \\ 0 & \cdots & 0 & 1 \end{array} \right).$$

For a partition $\mu \in \mathbb{Y}^{(m)}$, we define determinants R_μ and S_μ in the same fashion as in Equation (4.88). Then the following identity holds,

$$\det SR = \sum_{\mu \in \mathbb{Y}^{(m)}} S_\mu R_\mu.$$

Notice that because of the specific definition of the matrices R and S , the determinants R_μ and S_μ can only be defined for $\mu \in \mathbb{Y}^{(m)}$, hence the proposition. For example, if $m = 1$, one easily computes that

$$\det \begin{pmatrix} S_{0,-1} & 1 & 0 \\ S_{1,-1} & S_{1,0} & 1 \end{pmatrix} \cdot \begin{pmatrix} R_{-1,0} & R_{-1,1} \\ 1 & R_{0,1} \\ 0 & 1 \end{pmatrix} = S_{\emptyset} R_{\emptyset} + S_{\square} R_{\square} + S_{\square} R_{\square}.$$

Proof. It is a particular case of the well known Cauchy–Binet formula (see e.g. [Gan00]). The latter states that if A is an $n \times n'$ matrix and B is an $n' \times n$ matrix, then the determinant of AB is the sum of the product of the minors over all minors of size n ; in other words,

$$\det AB = \sum_{1 \leq i_1 < \dots < i_n \leq n} \begin{vmatrix} A_{1,i_1} & \dots & A_{1,i_n} \\ \vdots & & \vdots \\ A_{n,i_1} & \dots & A_{n,i_n} \end{vmatrix} \begin{vmatrix} B_{i_1,1} & \dots & B_{i_1,n} \\ \vdots & & \vdots \\ B_{i_n,1} & \dots & B_{i_n,n} \end{vmatrix}.$$

So let a sequence of $m + 1$ integers $-m \leq i_1 < i_2 < \dots < i_{m+1} \leq m$. We can assume without loss of generality that

$$(i_1, \dots, i_p, i_{p+1}, \dots, i_{m+1}) = (i_1, \dots, i_p, p, \dots, m), \quad (4.94)$$

for some $p \in \{0, \dots, m + 1\}$. If $p = 0$, the sequence reads $(0, \dots, m)$; if $p = m + 1$, the sequence is simply such that $i_{m+1} < m$. We define

$$\mu_k = k - i_k - 1 \quad k \in \{1, \dots, p\}$$

Then we automatically have $\mu_{k+1} \leq \mu_k$ from $i_{k+1} > i_k$ and it is not hard to see that $\mu_k \geq 0$ starting from $i_{p+1} = p$ and then applying a backward induction to $i_{p+1} > i_p > i_{p-1} > \dots$. Therefore, the μ_k 's form a partition $\mu = (\mu_1 \geq \dots \geq \mu_p)$. If $p = 0$, the partition is the empty partition $\mu = \emptyset$. Moreover, by construction we have $\mu_1 = -i_1 \leq m$ and $\ell(\mu) = p \leq m + 1$, so that $\mu \in \mathbb{Y}^{(m)}$. Reciprocally, any partition of $\mu \in \mathbb{Y}^{(m)}$ of length $\ell(\mu) = p \in \{0, \dots, m + 1\}$ determines a unique sequence (i_1, \dots, i_{m+1}) satisfying Equation (4.94). This explains the expansion over partitions of $\mathbb{Y}^{(m)}$.

It remains to prove that given a sequence $-m \leq i_1 < i_2 < \dots < i_{m+1} \leq m$ satisfying Equation (4.94) for some $0 \leq p \leq m + 1$, and given $\mu = (\mu_1 \geq \dots \geq \mu_p) \in \mathbb{Y}^{(m)}$ the partition it determines, the minors of R and S are respectively given by Equation (4.88). We prove it for the matrix S .

◊ If $p = m + 1$, it is immediate: the partition μ has length $\ell(\mu) = m + 1$ so that S_μ given in Equation (4.88) is indeed the minor of size $m + 1$ associated to the sequence (i_1, \dots, i_{m+1}) .

◊ If $p = 0$, the sequence is given by $(0, \dots, m)$ for which the minor is the determinant of a lower triangular matrix with 1's on the diagonal and therefore equals 1; accordingly, the associated partition is the empty partition $\mu = \emptyset$, for which we set $S_\emptyset = 1$.

◊ If $1 \leq p \leq m$, the minor of S associated to the sequence (i_1, \dots, i_{m+1}) has the form

$$\left| \begin{array}{ccc|ccc} S_{0,i_1} & \cdots & S_{0,i_p} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ S_{p-1,i_1} & \cdots & S_{p-1,i_p} & 0 & \cdots & 0 \\ \hline S_{p,i_1} & \cdots & S_{p,i_p} & 1 & 0 & \cdots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ & & & & & 1 & 0 \\ S_{m,i_1} & \cdots & S_{m,i_p} & S_{m,p} & \cdots & S_{m,m-1} & 1 \end{array} \right| = \left| \begin{array}{ccc} S_{0,i_1} & \cdots & S_{0,i_p} \\ \vdots & & \vdots \\ S_{p-1,i_1} & \cdots & S_{p-1,i_p} \end{array} \right|. \quad (4.95)$$

By definition, the right-hand side of the above equation is S_μ as defined in Equation (4.88); which concludes the proof. \square

We now come back to the Laurent matrices $r = \mathbf{L}(\gamma)$ and $s(\mathbf{t}) = \mathbf{L}(g(\mathbf{t}))$. For $m \geq 1$, we define the following finite-size rectangular matrices:

$$s_m(\mathbf{t}) = (s_{i,j}(\mathbf{t}) \mid i \in [0, m], j \in [-m, m]), \quad r_m = (r_{i,j} \mid i \in [-m, m], j \in [0, m]).$$

These matrices have entries in $\mathbb{C}[[\mathbf{t}, \zeta]]$. Since $\gamma = \text{id} + \mathcal{O}(\lambda^{-1})$ and $g(\mathbf{t}) = \text{id} + \mathcal{O}(\lambda)$, the matrices $s_m(\mathbf{t})$ and r_m have the form assumed in Lemma (4.5.6).

Lemma 4.5.7. *The sequence $(\det(s_m(\mathbf{t}) \cdot r_m))_{m \geq 1}$ converges w.r.t. the topology induced by the valuation in ζ . Moreover,*

$$\lim_{m \rightarrow \infty} \det(s_m(\mathbf{t}) \cdot r_m) = \sum_{\mu \in \mathbb{Y}} s_\mu(\mathbf{t}) \cdot r_\mu. \quad (4.96)$$

Proof. Let us denote $u_m := \det(s_m s_m(\mathbf{t}))$. Let $N = \lfloor m/n \rfloor$ (the floor integer part of m/n , where n is the dimension of the representation $\pi : \mathfrak{g} \rightarrow \mathfrak{gl}(n)$ we chose). Then by Lemma 4.5.6, we have

$$\sum_{\mu \in \mathbb{Y}} s_\mu(\mathbf{t}) \cdot r_\mu - u_m = \sum_{\mu \in \mathbb{Y} \setminus \mathbb{Y}^{(m)}} s_\mu(\mathbf{t}) \cdot r_\mu.$$

Now consider a partition $\mu \in \mathbb{Y} \setminus \mathbb{Y}^{(m)}$ with minimal length $\ell(\mu) = m + 2$, minimal $\mu_1 = m + 1$ and minimal weight $|\mu| = \mu_1 + \dots + \mu_{\ell(\mu)}$. This is realized by $\mu = (m + 1, 1, \dots, 1)$. The minor associated to μ reads

$$s_\mu = \left| \begin{array}{cccc|ccc} s_{0,-m-1} & 1 & 0 & \cdots & 0 \\ s_{1,-m-1} & s_{1,0} & 1 & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ s_{m+1,-m-1} & s_{m+1,0} & \cdots & s_{m+1,m} & 1 \end{array} \right| = s_{0,-m-1} (s_{1,0} \cdots s_{m+1,m-1}).$$

Therefore,

$$\text{val}_\zeta s_\mu \geq \text{val}_\zeta s_{0,-m-1} \geq |N|$$

since $N = \lfloor m/n \rfloor$. Clearly, $\lim_{m \rightarrow \infty} \text{val}_\zeta s_\mu = \infty$. Moreover, any partition in $\mathbb{Y} \setminus \mathbb{Y}^{(m)}$ contains terms of higher valuation in ζ , so that $\lim_{m \rightarrow \infty} \text{val}_\zeta \sum_{\mu \in \mathbb{Y}} s_\mu(\mathbf{t}) \cdot r_\mu - u_m = \infty$; this concludes the proof. \square

Now let $\phi = g(\mathbf{t}) \cdot \gamma$ and let $T_N(\phi)$ be the associated N -truncated Toeplitz block matrix, which we recall the definition below,

$$T_N(\phi) = (\phi_{I-J})_{I,J \in [0,N]}.$$

To finally prove Theorem 4.5.2, we make the following observation.

If $m = Nn$, then the matrices $s_{Nn}(\mathbf{t})$ and r_{Nn} admit a block decomposition:

$$s_{Nn}(\mathbf{t}) = \begin{pmatrix} g_N & \cdots & g_1 & \text{id} & 0 & \cdots & 0 \\ \vdots & & & \ddots & \ddots & \ddots & \vdots \\ g_{2N} & \cdots & & & g_1 & \text{id} \end{pmatrix} \quad r_{Nn} = \begin{pmatrix} \gamma_{-N} & \cdots & \gamma_{-2N} \\ \vdots & & \vdots \\ \gamma_{-1} & & \\ \text{id} & \ddots & \\ 0 & \ddots & \gamma_{-1} \\ \vdots & & \text{id} \\ 0 & \cdots & 0 \end{pmatrix}$$

Consequently, if we denote $\theta_{nN} = s_{nN}(\mathbf{t}) \cdot r_{nN}$, then in terms of blocks,

$$(\theta_{nN})_{I,J} = \sum_{K=-N}^N g_{I-K} \gamma_{K-J}, \quad (T_N(\phi))_{I,J} = \sum_{K \in \mathbb{Z}} g_{I-K} \gamma_{K-J}.$$

Therefore,

$$(T_N(\phi))_{I,J} - (\theta_{nN})_{I,J} = \sum_{K=N}^{+\infty} g_{I-K} \gamma_{K-J} + g_{I+K} \gamma_{-K-J}.$$

Yet $\text{val}_\zeta (g_{I-K} \gamma_{K-J}) \geq 2N - |I| - |J|$; this holds for $g_{I+K} \gamma_{-K-J}$ as well. Now this computation was done block-wise; we let the reader convince themselves that it implies that component-wise,

$$\lim_{m \rightarrow \infty} (\theta_m)_{i,j} = T(\phi)_{i,j}$$

for all $i, j \geq 0$. This concludes the proof of Theorem 4.5.2.

4.5.2 Polynomial tau functions and bilinear equations

Now consider the case where $X \in \mathfrak{g}_-$ is such that $\pi(X)$ is a nilpotent matrix. Then $\gamma = e^X$ admits a finite expansion in negative powers of λ . It follows that for sufficiently large partitions, the Plücker coordinates r_μ of X vanish. As a consequence, we get the following corollary to Theorem 4.5.2.

Corollary 4.5.8. *For any $X \in \mathfrak{g}_-$ such that $\pi(X)$ is a nilpotent matrix, the function $(\tau_X)^\frac{1}{\kappa}$ is polynomial.*

Let us condense the above statement in what will hold as the leitmotiv of this chapter's remainder:

$$\forall X \in \mathfrak{g}_-, \pi(X) \text{ is a nilpotent matrix} \implies \tau_X(\mathbf{t})^\frac{1}{\kappa} \text{ is polynomial.}$$

Although this has not been proven, there is strong indication that the above procedure, in the A_1 case, recovers all tau functions of KdV. For other cases though, we do not know if we would recover all polynomial tau function. In what follows we compute the first few polynomial tau functions of the DS hierarchy of type \mathfrak{g} for $\mathfrak{g} = A_1, A_2, B_2$ and D_4 .

We will use the particular tau functions to deduce possible bilinear equations of small degrees. Note that each Drinfeld–Sokolov hierarchy has infinitely many solutions. The usual question is to find particular solutions (and their tau-functions) to the DS hierarchy (e.g. to solve all PDEs in this hierarchy together). Here, as we mentioned above, we will also consider the reverse: deduce possible bilinear equations associated to a hierarchy from particular solutions.

Remark 4.5.9. If $\kappa = 1$, then the tau function itself is a polynomial. Interestingly enough, in the computations that we performed, even when $\kappa = 1/2$, we obtained some polynomial tau functions; in other words, the finite sum in (4.92) is a perfect square.

Bilinear derivatives

The bilinear derivatives are operators $D_{i_1} \cdots D_{i_k} : \mathbb{C}[[\mathbf{t}]] \otimes \mathbb{C}[[\mathbf{t}]] \rightarrow \mathbb{C}[[\mathbf{t}]]$, for $i_1, \dots, i_k \in \mathbb{E}^+$, defined by the following identity: for any $f, g \in \mathbb{C}[[\mathbf{t}]]$,¹³

$$\exp\left(\sum_{i \in \mathbb{E}^+} h_i D_i\right)(f, g) = f(\mathbf{t} + \mathbf{h})g(\mathbf{t} - \mathbf{h}).$$

For example,

$$\begin{aligned} D_i(f, g) &= \frac{\partial f}{\partial t_i} g - f \frac{\partial g}{\partial t_i}, \\ D_i D_j(f, g) &= \frac{\partial^2 f}{\partial t_i \partial t_j} g + f \frac{\partial^2 g}{\partial t_i \partial t_j} - \frac{\partial f}{\partial t_i} \frac{\partial g}{\partial t_j} - \frac{\partial f}{\partial t_j} \frac{\partial g}{\partial t_i}. \end{aligned}$$

The bilinear derivatives give a way to translate the \mathfrak{g} -DS hierarchy on the solution into equations on the tau function itself.

¹³Notice that each expression $D_{i_1} \cdots D_{i_k}$ is a bilinear operator in itself is not the product of D_{i_1} to D_{i_k} (this would actually not make sense).

Example 4.5.10. In the A_n case, the equations of the DS hierarchy, when translated onto the tau function τ , are equivalent to the following two chains of bilinear equations (these can be deduced from results of [DJKM83], they also appear in [ACvM12], Lemma 4.1):

$$\begin{aligned} (p_{k+1}([D]) - \frac{1}{2}D_1D_k)(\tau, \tau) &= 0, \\ (D_1D_k - \frac{1}{2}D_2D_{k-1} - D_1p_k([D]))(\tau, \tau) &= 0. \end{aligned} \quad (4.97)$$

In the above, $[D] := (D_1, \frac{1}{2}D_2, \frac{1}{3}D_3, \dots)$ while the p_k 's are the elementary Schur polynomials which can be defined by

$$\exp\left(\sum_{i \geq 1} t_i \lambda^i\right) = \sum_{k \geq 1} p_k(\mathbf{t}) \lambda^k.$$

They were already introduced in Chapter 3, Equation (3.49). Notice that since we reduced to the A_n case, all variables $t_{k(n+1)} = 0$ because $h = n + 1$ (the Coxeter number of $\hat{\mathfrak{g}}$).

The bilinear derivatives have a natural gradation given by $\deg D_{i_1} \cdots D_{i_k} = i_1 + \cdots + i_k$. Denote by $\text{HDS}_{\mathfrak{g}}$ the linear space of bilinear equations¹⁴ satisfied by the Drinfeld–Sokolov hierarchies of type \mathfrak{g} , which decomposes into homogeneous subspaces

$$\text{HDS}_{\mathfrak{g}} = \bigoplus_{i \geq 0} \text{HDS}_{\mathfrak{g}}^{[i]}.$$

The gradation allows us to list all possible bilinear equations up to certain degree.

The A_1 case

Let us chose the standard matrix realization $\hat{\mathfrak{g}} = \mathfrak{sl}(2; \mathbb{C})$. Consider the following two elements in $\lambda^{-1}\hat{\mathfrak{g}}[[\lambda^{-1}]]$

$$\frac{1}{\lambda}F = \frac{1}{\lambda} \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \frac{1}{\lambda}E = \frac{1}{\lambda} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}. \quad (4.98)$$

The associated polynomial tau functions are

$$\tau_1 = 1 + t_1, \quad \tau_2 = 1 + t_3 - \frac{t_1^3}{3} \quad (4.99)$$

respectively. Similarly, one computes polynomial tau functions corresponding to elements of the form $\lambda^{-k}F$, $\lambda^{-k}E$, $k \geq 2$. For example, for $k = 2$, we obtain

$$\tau_3 = 1 + 2t_3 - t_5t_1 + t_3^2 + \frac{t_1^3}{3} + \frac{1}{3}t_3t_1^3 - \frac{1}{45}t_1^6, \quad (4.100)$$

$$\tau_4 = 1 - t_3t_7 + 2t_5 + t_5^2 + t_3^3t_1 - t_3t_5t_1^2 - t_3t_1^2 + \frac{1}{3}t_7t_1^3 - \frac{t_1^5}{15} - \frac{1}{15}t_5t_1^5 + \frac{1}{105}t_3t_1^7 - \frac{t_1^{10}}{4725}, \quad (4.101)$$

corresponding to $\lambda^{-2}F$ and $\lambda^{-2}E$, respectively.

¹⁴The ‘‘H’’ in ‘‘HDS’’ stands for Ryugo Hirota who first introduced the bilinear derivative, also called Hirota derivatives [Hir76].

Remark 4.5.11. Up to a certain affine transformation of the variables, the above polynomials are the Adler–Moser polynomials studied in Chapter 1. This is no surprise since the Adler–Moser polynomials recover *all* the polynomial tau functions KdV. Although we do not have a proof of this statement, we believe that the procedure we use here allows to recover all the Adler–Moser polynomials. A first clue in this direction is by identifying the degree in t_1 (which characterizes each Adler–Moser polynomial).

Now consider all bilinear equations up to degree 4:

$$(\beta + \alpha_0 D_1^2 + \alpha_1 D_1^4 + \alpha_2 D_1 D_3)(\tau, \tau) = 0 \quad (4.102)$$

where $\beta, \alpha_0, \alpha_1, \alpha_2$ are complex constants. Requiring that τ_1, τ_2 satisfy the above ansatz (4.102), we find that up to a multiplicative constant there is only one possible choice of coefficients:

$$(D_1^4 - 4D_1 D_3)(\tau, \tau) = 0. \quad (4.103)$$

Similarly up to degree 6, we find out only two more possible linearly independent bilinear equations that are satisfied by $\tau_1, \tau_2, \tau_3, \tau_4$

$$(D_1^6 + 20D_1^3 D_3 - 96D_1 D_5)(\tau, \tau) = 0, \quad (4.104)$$

$$(D_1^3 D_3 + 2D_3^2 - 6D_1 D_5)(\tau, \tau) = 0, \quad (4.105)$$

which are identified to two of the well-known bilinear equations for the hierarchy of type A_1 (the KdV hierarchy) (see Example 4.5.10). Consequently, we have shown that

$$\dim_{\mathbb{C}} \text{HDS}_{A_1}^{[\text{deg} \leq 6]} \leq 3.$$

Moreover, (4.103)–(4.105) are the three only possible choices of homogeneous basis (up to constant factors) of $\text{HDS}_{\mathfrak{g}}^{[\text{deg} \leq 6]}$.

The A_2 case

We still chose the standard matrix realization $\mathfrak{g} = \mathfrak{sl}(3; \mathbb{C})$. Consider for example the following two elements in $\lambda^{-1} \mathfrak{g}[[\lambda^{-1}]]$:

$$X_1 = \frac{1}{\lambda} \begin{pmatrix} 0 & 0 & 0 \\ a_1 & 0 & 0 \\ a_2 & a_3 & 0 \end{pmatrix}, \quad X_2 = \frac{1}{\lambda} \begin{pmatrix} 0 & a_1 & a_2 \\ 0 & 0 & a_3 \\ 0 & 0 & 0 \end{pmatrix}, \quad (4.106)$$

where a_1, a_2, a_3 are arbitrary constants. The corresponding polynomial tau functions will be denoted by τ_1, τ_2 , respectively. We have

$$\begin{aligned} \tau_1 = & 1 + a_2 t_1 + \frac{1}{2} a_1 t_1^2 - \frac{1}{2} a_3 t_1^2 + \frac{1}{8} a_1 a_3 t_1^4 - \frac{1}{160} a_1^2 a_3 t_1^6 + \frac{1}{160} a_1 a_3^2 t_1^6 - \frac{1}{1792} a_1^2 a_3^2 t_1^8 + a_1 t_2 + a_3 t_2 \\ & + \frac{1}{16} a_1^2 a_3 t_1^4 t_2 + \frac{1}{16} a_1 a_3^2 t_1^4 t_2 + \frac{3}{2} a_1 a_3 t_2^2 - \frac{1}{8} a_1^2 a_3 t_1^2 t_2^2 + \frac{1}{8} a_1 a_3^2 t_1^2 t_2^2 + \frac{1}{32} a_1^2 a_3^2 t_1^4 t_2^2 + \frac{1}{4} a_1^2 a_3 t_2^3 \\ & + \frac{1}{4} a_1 a_3^2 t_2^3 + \frac{1}{16} a_1^2 a_3^2 t_2^4 - \frac{1}{4} a_1^2 a_3 t_1^2 t_4 - \frac{1}{4} a_1 a_3^2 t_1^2 t_4 - \frac{1}{2} a_1^2 a_3 t_2 t_4 + \frac{1}{2} a_1 a_3^2 t_2 t_4 - \frac{1}{4} a_1^2 a_3^2 t_1^2 t_2 t_4 \\ & - \frac{1}{4} a_1^2 a_3^2 t_2^2 + \frac{1}{2} a_1^2 a_3 t_1 t_5 - \frac{1}{2} a_1 a_3^2 t_1 t_5 + \frac{1}{4} a_1^2 a_3^2 t_1 t_7, \end{aligned}$$

$$\begin{aligned}
\tau_2 = & 1 - \frac{1}{8}a_1t_1^4 + \frac{1}{8}a_3t_1^4 + \frac{1}{20}a_2t_1^5 + \frac{1}{640}a_1a_3t_1^8 - \frac{1}{358400}a_1^2a_3t_1^{12} + \frac{1}{358400}a_1a_3^2t_1^{12} - \frac{1}{90112000}a_1^2a_3^2t_1^{16} \\
& - \frac{1}{2}a_1t_1^2t_2 - \frac{1}{2}a_3t_1^2t_2 - \frac{1}{12800}a_1^2a_3t_1^{10}t_2 - \frac{1}{12800}a_1a_3^2t_1^{10}t_2 + \frac{1}{2}a_1t_2^2 - \frac{1}{2}a_3t_2^2 - a_2t_1t_2^2 + \frac{1}{16}a_1a_3t_1^4t_2^2 \\
& - \frac{13}{17920}a_1^2a_3t_1^8t_2^2 + \frac{13}{17920}a_1a_3^2t_1^8t_2^2 + \frac{3}{1126400}a_1^2a_3^2t_1^{12}t_2^2 - \frac{1}{320}a_1^2a_3t_1^6t_2^3 - \frac{1}{320}a_1a_3^2t_1^6t_2^3 - \frac{3}{8}a_1a_3t_2^4 \\
& - \frac{1}{128}a_1^2a_3t_1^4t_2^4 + \frac{1}{128}a_1a_3^2t_1^4t_2^4 - \frac{1}{10240}a_1^2a_3^2t_1^8t_2^4 - \frac{1}{32}a_1^2a_3t_1^2t_2^5 - \frac{1}{32}a_1a_3^2t_1^2t_2^5 - \frac{1}{32}a_1^2a_3t_2^6 + \frac{1}{32}a_1a_3^2t_2^6 \\
& + \frac{1}{256}a_1^2a_3^2t_1^6t_2^6 + \frac{1}{256}a_1^2a_3^2t_2^8 + a_1t_4 + a_3t_4 + \frac{1}{1280}a_1^2a_3t_1^8t_4 + \frac{1}{1280}a_1a_3^2t_1^8t_4 - \frac{3}{2}a_1a_3t_1^2t_2t_4 \\
& + \frac{1}{160}a_1^2a_3t_1^6t_2t_4 - \frac{1}{160}a_1a_3^2t_1^6t_2t_4 - \frac{1}{12800}a_1^2a_3^2t_1^{10}t_2t_4 + \frac{1}{32}a_1^2a_3t_1^4t_2^2t_4 + \frac{1}{32}a_1a_3^2t_1^4t_2^2t_4 - \frac{1}{8}a_1^2a_3t_1^2t_2^3t_4 \\
& + \frac{1}{8}a_1a_3^2t_1^2t_2^3t_4 - \frac{1}{320}a_1^2a_3^2t_1^6t_2^3t_4 - \frac{3}{16}a_1^2a_3t_2^4t_4 - \frac{3}{16}a_1a_3^2t_2^4t_4 - \frac{1}{32}a_1^2a_3^2t_1^2t_2^5t_4 + \frac{3}{2}a_1a_3t_2^4 + \frac{1}{32}a_1^2a_3t_1^4t_2^4 \\
& - \frac{1}{32}a_1a_3^2t_1^4t_2^4 + \frac{1}{2560}a_1^2a_3^2t_1^8t_2^4 - \frac{3}{8}a_1^2a_3t_1^2t_2^4 - \frac{3}{8}a_1a_3^2t_1^2t_2^4 - \frac{1}{8}a_1^2a_3t_2^4 + \frac{1}{8}a_1a_3^2t_2^4 \\
& + \frac{1}{64}a_1^2a_3^2t_1^4t_2^4 - \frac{3}{32}a_1^2a_3^2t_2^4 + \frac{1}{4}a_1^2a_3t_4^3 + \frac{1}{4}a_1a_3^2t_4^3 - \frac{1}{8}a_1^2a_3^2t_1^2t_2^3t_4 + \frac{1}{16}a_1^2a_3^2t_4^4 + a_2t_5 + \frac{1}{2}a_1a_3t_1^3t_5 \\
& + \frac{3}{1120}a_1^2a_3t_1^7t_5 - \frac{3}{1120}a_1a_3^2t_1^7t_5 + \frac{1}{140800}a_1^2a_3^2t_1^{11}t_5 + \frac{1}{80}a_1^2a_3t_1^5t_2t_5 + \frac{1}{80}a_1a_3^2t_1^5t_2t_5 + \frac{1}{8}a_1^2a_3t_1^3t_2^2t_5 \\
& - \frac{1}{8}a_1a_3^2t_1^3t_2^2t_5 + \frac{1}{320}a_1^2a_3^2t_1^7t_2^2t_5 + \frac{1}{4}a_1^2a_3t_1^3t_2^3t_5 + \frac{1}{4}a_1a_3^2t_1^3t_2^3t_5 - \frac{1}{32}a_1^2a_3^2t_1^4t_2^4t_5 + \frac{1}{4}a_1^2a_3t_1^3t_2^4t_5 \\
& + \frac{1}{4}a_1a_3^2t_1^3t_2^4t_5 - \frac{1}{2}a_1^2a_3t_1t_2t_4t_5 + \frac{1}{2}a_1a_3^2t_1t_2t_4t_5 + \frac{1}{80}a_1^2a_3^2t_1^5t_2t_4t_5 + \frac{1}{4}a_1^2a_3t_1^3t_2^3t_4t_5 + \frac{1}{8}a_1^2a_3^2t_1^3t_2^4t_5 \\
& + \frac{1}{4}a_1^2a_3t_1^2t_2^5 - \frac{1}{4}a_1a_3^2t_1^2t_2^5 - \frac{1}{160}a_1^2a_3^2t_1^6t_2^5 - \frac{1}{2}a_1^2a_3t_2^2t_5^2 - \frac{1}{2}a_1a_3^2t_2^2t_5^2 - \frac{1}{8}a_1^2a_3^2t_1^2t_2^2t_5^2 - \frac{1}{2}a_1^2a_3^2t_2^4t_5^2 \\
& - \frac{1}{4}a_1^2a_3^2t_1^3t_5^3 - \frac{1}{40}a_1^2a_3t_1^5t_7 + \frac{1}{40}a_1a_3^2t_1^5t_7 + \frac{1}{2}a_1^2a_3t_1^2t_2^7 - \frac{1}{2}a_1a_3^2t_1^2t_2^7 - \frac{1}{2}a_1^2a_3t_1t_2^7 + \frac{1}{2}a_1a_3^2t_1t_2^7 \\
& - \frac{1}{16}a_1^2a_3t_1^4t_8 - \frac{1}{16}a_1a_3^2t_1^4t_8 - \frac{1}{4}a_1^2a_3t_1^2t_2^8 + \frac{1}{4}a_1a_3^2t_1^2t_2^8 - \frac{1}{160}a_1^2a_3^2t_1^6t_2^8 + \frac{1}{4}a_1^2a_3t_2^8 + \frac{1}{4}a_1a_3^2t_2^8 \\
& + \frac{1}{8}a_1^2a_3^2t_1^2t_2^8 + \frac{1}{2}a_1^2a_3t_4t_8 - \frac{1}{2}a_1a_3^2t_4t_8 - \frac{1}{16}a_1^2a_3^2t_1^4t_4t_8 + \frac{1}{4}a_1^2a_3^2t_2^2t_4t_8 + \frac{1}{2}a_1^2a_3^2t_1t_2t_5t_8 - \frac{1}{4}a_1^2a_3^2t_8^2 \\
& + \frac{1}{80}a_1^2a_3^2t_1^5t_{11} - \frac{1}{4}a_1^2a_3^2t_1^2t_2^2t_{11} + \frac{1}{4}a_1^2a_3^2t_5t_{11}.
\end{aligned}$$

Consider all possible bilinear equations of degree 4:

$$(\alpha_1 D_1^4 + \alpha_2 D_2^2)(\tau, \tau) = 0,$$

Requiring that τ_1 satisfies this ansatz we find that there is only one possible choice:

$$(D_1^4 + 3D_2^2)(\tau, \tau) = 0.$$

Similarly, requiring that τ_1 and τ_2 both satisfy the ansatz of bilinear equation of degree 6, we find that there are only two linearly independent bilinear equations of degree 6:

$$\begin{aligned}
(D_1^6 + 45D_1^2D_2^2 + 90D_2D_4 - 216D_1D_5)(\tau, \tau) &= 0, \\
(D_1^6 + 15D_1^2D_2^2 + 60D_2D_4 - 96D_1D_5)(\tau, \tau) &= 0,
\end{aligned}$$

which are identified to two of the well-known bilinear equations for the hierarchy of type A_2 (the Boussinesq hierarchy) (see Example 4.5.10).

The B_2 case

We choose the matrix realization of the B_2 simple Lie algebra as in [DS84] (cf. page 2032 therein). We consider two explicit examples given respectively by the following matrices¹⁵

¹⁵ X_2 is not the most general upper triangular element of homogeneous degree -1 , as the tau function for the most general case is too big.

$$X_1 = \frac{1}{\lambda} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ a_2 & 0 & 0 & 0 & 0 \\ a_3 & a_5 & 0 & 0 & 0 \\ a_4 & 0 & a_5 & 0 & 0 \\ 0 & a_4 & -a_3 & a_2 & 0 \end{pmatrix}, \quad X_2 = \frac{1}{\lambda} \begin{pmatrix} 0 & 0 & a_3 & a_4 & 0 \\ 0 & 0 & 0 & 0 & a_4 \\ 0 & 0 & 0 & 0 & -a_3 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

The associated tau functions will be denoted by τ_1 and τ_2 . They have the expressions

$$\begin{aligned} \tau_1 = & 1 + \frac{1}{2}a_4t_1 + \frac{1}{4}a_3t_1^2 + \frac{1}{12}a_2t_1^3 - \frac{1}{12}a_5t_1^3 - \frac{1}{192}a_3^2t_1^4 + \frac{1}{96}a_2a_4t_1^4 + \frac{1}{192}a_3a_5t_1^5 + \frac{1}{1920}a_2a_5t_1^6 - \frac{1}{720}a_5^2t_1^6 \\ & - \frac{1}{11520}a_2a_4a_5t_1^7 - \frac{1}{53760}a_2a_3a_5t_1^8 - \frac{1}{483840}a_2^2a_5t_1^9 + \frac{1}{1088640}a_2a_5^2t_1^9 + \frac{1}{3110400}a_2a_4a_5^2t_1^{10} \\ & + \frac{11}{87091200}a_2a_3a_5^2t_1^{11} + \frac{43}{2090188800}a_2^2a_5^2t_1^{12} - \frac{79}{2874009600}a_2a_5^3t_1^{12} - \frac{37}{1931334451200}a_2^2a_5^3t_1^{15} \\ & + \frac{1}{115880067072000}a_2^2a_5^4t_1^{18} + \frac{1}{2}a_2t_3 + a_5t_3 - \frac{1}{8}a_3^2t_1t_3 + \frac{1}{4}a_2a_4t_1t_3 + \frac{1}{8}a_3a_5t_1^2t_3 + \frac{1}{16}a_2a_5t_1^3t_3 \\ & - \frac{1}{24}a_5^2t_1^3t_3 + \frac{1}{192}a_2a_4a_5t_1^4t_3 + \frac{1}{640}a_2a_3a_5t_1^5t_3 + \frac{1}{3840}a_2^2a_5t_1^6t_3 - \frac{1}{1080}a_2a_5^2t_1^6t_3 - \frac{1}{34560}a_2a_4a_5^2t_1^7t_3 \\ & - \frac{1}{193536}a_2a_3a_5^2t_1^8t_3 - \frac{31}{17418240}a_2^2a_5^2t_1^9t_3 + \frac{11}{8709120}a_2a_5^3t_1^9t_3 + \frac{101}{22992076800}a_2^2a_5^3t_1^{12}t_3 \\ & - \frac{29}{6437781504000}a_2^2a_5^4t_1^{15}t_3 + \frac{3}{4}a_2a_5t_3^2 + \frac{1}{4}a_5^2t_3^2 + \frac{1}{16}a_2a_4a_5t_1t_3^2 + \frac{5}{288}a_2a_5^2t_1^2t_3^2 + \frac{1}{1152}a_2a_4a_5^2t_1^4t_3^2 \\ & + \frac{7}{23040}a_2a_3a_5^2t_1^5t_3^2 + \frac{1}{138240}a_2^2a_5^2t_1^6t_3^2 - \frac{17}{69120}a_2a_5^3t_1^6t_3^2 - \frac{13}{34836480}a_2^2a_5^3t_1^9t_3^2 + \frac{1}{2043740160}a_2^2a_5^4t_1^{12}t_3^2 \\ & + \frac{1}{16}a_2^2a_5t_3^3 + \frac{11}{36}a_2a_5^2t_3^3 + \frac{1}{144}a_2a_4a_5^2t_1t_3^3 - \frac{1}{576}a_2a_3a_5^2t_1^2t_3^3 + \frac{1}{3456}a_2^2a_5^2t_1^3t_3^3 + \frac{7}{1728}a_2a_5^3t_1^3t_3^3 \\ & - \frac{11}{414720}a_2^2a_5^3t_1^6t_3^3 - \frac{1}{34836480}a_2^2a_5^4t_1^9t_3^3 + \frac{25}{576}a_2^2a_5^2t_3^4 + \frac{5}{144}a_2a_5^3t_3^4 + \frac{1}{3456}a_2^2a_5^3t_1^3t_3^4 - \frac{1}{552960}a_2^2a_5^4t_1^6t_3^4 \\ & + \frac{5}{576}a_2^2a_5^3t_3^5 + \frac{1}{2304}a_2^2a_5^4t_3^6 - \frac{1}{4}a_2a_5t_1t_5 + \frac{1}{4}a_5^2t_1t_5 - \frac{1}{16}a_2a_4a_5t_1^2t_5 - \frac{1}{48}a_2a_3a_5t_1^3t_5 - \frac{1}{192}a_2^2a_5t_1^4t_5 \\ & + \frac{1}{72}a_2^2a_5^2t_1^4t_5 + \frac{1}{5760}a_2a_4a_5^2t_1^5t_5 - \frac{11}{69120}a_2a_3a_5^2t_1^6t_5 + \frac{1}{483840}a_2^2a_5^2t_1^7t_5 + \frac{41}{483840}a_2a_5^3t_1^7t_5 \\ & - \frac{1}{17418240}a_2^2a_5^3t_1^{10}t_5 + \frac{1}{4379443200}a_2^2a_5^4t_1^{13}t_5 - \frac{1}{8}a_2a_3a_5t_3t_5 - \frac{1}{8}a_2^2a_5t_1t_3t_5 - \frac{1}{24}a_2a_5^2t_1t_3t_5 \\ & - \frac{1}{48}a_2a_4a_5^2t_1^2t_3t_5 - \frac{5}{576}a_2a_3a_5^2t_1^3t_3t_5 - \frac{1}{1152}a_2^2a_5^2t_1^4t_3t_5 + \frac{7}{1152}a_2a_5^3t_1^4t_3t_5 + \frac{29}{967680}a_2^2a_5^3t_1^7t_3t_5 \\ & - \frac{1}{116121600}a_2^2a_5^4t_1^{10}t_3t_5 - \frac{1}{24}a_2a_3a_5^2t_1^2t_3^2t_5 - \frac{7}{96}a_2^2a_5t_1^2t_3^2t_5 - \frac{1}{48}a_2a_5^3t_1^2t_3^2t_5 + \frac{1}{1152}a_2^2a_5^3t_1^4t_3^2t_5 \\ & + \frac{1}{258048}a_2^2a_5^4t_1^7t_3^2t_5 - \frac{11}{576}a_2^2a_5^3t_1^3t_3^2t_5 + \frac{1}{9216}a_2^2a_5^4t_1^4t_3^2t_5 - \frac{1}{768}a_2^2a_5^4t_1^4t_3^2t_5 - \frac{1}{24}a_2a_4a_5^2t_5^2 \\ & - \frac{1}{32}a_2a_3a_5^2t_1t_5^2 + \frac{1}{64}a_2^2a_5^2t_1^2t_5^2 - \frac{1}{96}a_2a_5^3t_1^2t_5^2 - \frac{1}{2304}a_2^2a_5^3t_1^5t_5^2 - \frac{1}{1290240}a_2^2a_5^4t_1^8t_5^2 + \frac{1}{192}a_2^2a_5^3t_1^2t_3^2t_5^2 \\ & - \frac{1}{7680}a_2^2a_5^4t_1^5t_3^2t_5^2 + \frac{a_2^2a_5^4t_1^3t_3^2}{2304} - \frac{1}{192}a_2^2a_5^4t_3^3t_5 + \frac{1}{8}a_2a_3a_5t_1t_7 + \frac{1}{16}a_2^2a_5t_1^2t_7 - \frac{1}{8}a_2a_5^2t_1^2t_7 \\ & + \frac{1}{288}a_2a_3a_5^2t_1^4t_7 + \frac{1}{1440}a_2^2a_5^2t_1^5t_7 - \frac{1}{640}a_2a_5^3t_1^5t_7 - \frac{1}{967680}a_2^2a_5^3t_1^8t_7 - \frac{1}{87091200}a_2^2a_5^4t_1^{11}t_7 \\ & + \frac{1}{12}a_2a_3a_5^2t_1t_3t_7 + \frac{1}{24}a_2^2a_5^2t_1^2t_3t_7 - \frac{1}{16}a_2a_5^3t_1^2t_3t_7 - \frac{1}{11520}a_2^2a_5^3t_1^5t_3t_7 - \frac{1}{1935360}a_2^2a_5^4t_1^8t_3t_7 \\ & + \frac{1}{192}a_2^2a_5^3t_1^2t_3^2t_7 - \frac{a_2^2a_5^4t_1^5t_3^2t_7}{46080} + \frac{a_2^2a_5^4t_1^3t_3^2t_7}{1152} + \frac{1}{12}a_2^2a_5^2t_5t_7 - \frac{1}{12}a_2a_5^3t_5t_7 + \frac{1}{576}a_2^2a_5^3t_1^3t_5t_7 \\ & + \frac{7}{138240}a_2^2a_5^4t_1^6t_5t_7 + \frac{1}{24}a_2^2a_5^3t_3t_5t_7 + \frac{1}{1152}a_2^2a_5^4t_1^3t_3t_5t_7 + \frac{1}{96}a_2^2a_5^4t_3t_5t_7 + \frac{1}{192}a_2^2a_5^4t_1^2t_5^2t_7 \\ & - \frac{1}{48}a_2^2a_5^3t_1t_5^2 - \frac{1}{2304}a_2^2a_5^4t_1^4t_5^2 - \frac{1}{96}a_2^2a_5^4t_1^3t_3t_5^2, \end{aligned}$$

$$\begin{aligned} \tau_2 = & 1 + \frac{1}{288}a_3t_1^6 - \frac{1}{2016}a_4t_1^7 - \frac{1}{11612160}a_3^2t_1^{12} - \frac{1}{12}a_3t_1^3t_3 + \frac{1}{48}a_4t_1^4t_3 - \frac{1}{69120}a_3^2t_1^9t_3 + \frac{1}{2}a_3t_3^2 - \frac{1}{2}a_4t_1t_3^2 \\ & - \frac{1}{1920}a_3^3t_1^6t_3^2 - \frac{1}{48}a_3^2t_1^3t_3^3 + \frac{1}{16}a_3^2t_3^4 + \frac{1}{4032}a_3^2t_1^7t_5 - \frac{1}{96}a_3^2t_1^4t_3t_5 + \frac{1}{4}a_3^2t_1t_3^2t_5 + a_4t_7 + \frac{1}{160}a_3^2t_1^5t_7 \\ & - \frac{1}{2}a_3^2t_5t_7. \end{aligned}$$

Consider all bilinear equations up to degree 4

$$(\alpha_0 + \alpha_1 D_1^2 + \alpha_2 D_1^4 + \alpha_3 D_1 D_3)(\tau, \tau) = 0$$

where $\alpha_0, \dots, \alpha_3$ are constants. Requiring that τ_1 satisfies this ansatz of bilinear equations we find that there is no solution. Similarly, up to degree 8, we find that there are only two possible homogeneous equations (one is of degree 6 and the other is of degree 8). We arrive at

Proposition 4.5.12. *The following dimension estimates hold true*

$$\dim_{\mathbb{C}} \text{HDS}_{B_2}^{[\text{deg} \leq 4]} = 0, \quad \dim_{\mathbb{C}} \text{HDS}_{B_2}^{[\text{deg} \leq 6]} \leq 1, \quad \dim_{\mathbb{C}} \text{HDS}_{B_2}^{[\text{deg} \leq 8]} \leq 2.$$

Moreover, the only possible elements in $\text{HDS}_{B_2}^{[\text{deg} \leq 8]}$ are linear combinations of the following,

$$\begin{aligned} (D_1^6 - 5D_1^3 D_3 - 5D_3^2 + 9D_1 D_5)(\tau, \tau) &= 0, \\ (D_1^8 + 7D_1^5 D_3 - 35D_1^2 D_3^2 - 21D_1^3 D_5 - 42D_3 D_5 + 90D_1 D_7)(\tau, \tau) &= 0. \end{aligned}$$

Remark 4.5.13. As far as we know, explicit bilinear equations for the DS hierarchy of type B_2 are not pointed out in the literature, except that there is a super-variable version given in [KW89]. However, the relationship between the super bilinear equations of Kac–Wakimoto [KW89] and the DS hierarchy of type B_2 is not known. Finding explicit generating series of bilinear equations for the DS hierarchy of type B_2 remains an open question. It is also interesting to remark that the very same equations are contained in [DKM81], as the first two equations of the BKP hierarchy.

The D_4 case

Take the matrix realization of $\mathring{\mathfrak{g}}$ as in [BDY16a] (cf. Example 4.4 therein). Consider the particular point of the Sato Grassmannian of type D_4 given by

$$\gamma = 1 + \lambda E_{\theta}.$$

We put $t_{11} = 0$. It follows from Theorem 4.5.2 that the corresponding tau function is given by

$$\tau = \left(1 - \frac{1}{2} s_{(7|6)} - \frac{1}{2} s_{(6|7)} - \frac{1}{4} s_{(7,6|7,6)} \right)^{\frac{1}{2}}$$

where $s_{(7,6|7,6)} = s_{(7|7)}s_{(6|6)} - s_{(7|6)}s_{(6|7)}$, $s_{(6|6)} = s_{(7|7)} = 0$, and

$$\begin{aligned} s_{(6|7)} = s_{(7|6)} &= \frac{t_1^{11}}{1900800} - \frac{1}{480} t_5 t_1^6 + \frac{1}{160} t_3^2 t_1^5 + \frac{1}{120} t_3' t_1^5 + \frac{1}{80} t_3 t_3' t_1^5 \\ &\quad - \frac{1}{8} t_3^3 t_1^2 - \frac{1}{4} t_3 t_3' t_1^2 - \frac{3}{8} t_3^2 t_3' t_1^2 + \frac{1}{2} t_5^2 t_1 + \frac{3}{4} t_3^2 t_5 + t_3' t_5 + \frac{3}{2} t_3 t_3' t_5. \end{aligned}$$

Hence we have

$$\begin{aligned} \tau &= 1 - \frac{1}{2} s_{(7|6)} = 1 - \frac{t_1^{11}}{3801600} + \frac{1}{960} t_5 t_1^6 - \frac{1}{320} t_3^2 t_1^5 - \frac{1}{240} t_3' t_1^5 - \frac{1}{160} t_3 t_3' t_1^5 \\ &\quad + \frac{1}{16} t_3^3 t_1^2 + \frac{1}{8} t_3 t_3' t_1^2 + \frac{3}{16} t_3^2 t_3' t_1^2 - \frac{1}{4} t_5^2 t_1 - \frac{3}{8} t_3^2 t_5 - \frac{1}{2} t_3' t_5 - \frac{3}{4} t_3 t_3' t_5. \end{aligned}$$

Proposition 4.5.14. *The following dimension estimates hold true*

$$\dim_{\mathbb{C}} \text{HDS}_{D_4}^{[\text{deg} \leq 4]} = 0, \quad \dim_{\mathbb{C}} \text{HDS}_{D_4}^{[\text{deg} \leq 6]} \leq 3.$$

Moreover, the only possible elements in $\text{HDS}_{D_4}^{[\text{deg} \leq 6]}$ are linear combinations of

$$(2D_1^3 D_{3'} + 4D_3 D_{3'} - 3D_{3'}^2)(\tau, \tau) = 0, \quad (4.107)$$

$$(D_1^3 D_3 - D_1^3 D_{3'} + D_3 D_{3'} - D_3^2)(\tau, \tau) = 0, \quad (4.108)$$

$$(D_1^6 + 9D_1 D_5 - 10D_1^3 D_3 + 5D_1^3 D_{3'} - 5D_3 D_{3'})(\tau, \tau) = 0. \quad (4.109)$$

Our last remark is that under the following linear change of time variables

$$\begin{aligned} \partial_{t_1} &\mapsto 2^{-1/6} \partial_{T_1}, & \partial_{t_{3'}} &\mapsto 2^{1/2} \partial_{T_3} + 2^{1/2} 3^{-1/2} \partial_{T_{3'}}, \\ \partial_{t_3} &\mapsto 2^{1/2} \partial_{T_3}, & \partial_{t_5} &\mapsto 2^{7/6} \partial_{T_5} \end{aligned}$$

the bilinear equations (4.107)–(4.109) in the new time variables $T_1, T_3, T_{3'}, T_5$ coincide with those of Kac and Wakimoto [KW89]. Essentially speaking such a change of times is simply a renormalization of flows.

4.5.3 List of generalized Schur polynomials

Take π as in [DS84, BDY16a] (cf. page 2032 of [DS84] and Example 4.4 of [BDY16a]). We list in Table 4.3 the first several Schur polynomials of type (\mathfrak{g}, π) for simple Lie algebras of low ranks.

Type	A_1	A_2	B_2	B_3	C_2	D_4
s_1	t_1	t_1	0	0	t_1	0
s_2	$\frac{1}{2}t_1^2$	$\frac{1}{2}t_1^2 + t_2$	$\frac{1}{2}t_1$	$\frac{1}{2}t_1$	$\frac{1}{2}t_1^2$	$\frac{1}{2}t_1$
s_{1^2}	$\frac{1}{2}t_1^2$	$\frac{1}{2}t_1^2 - t_2$	$-\frac{1}{2}t_1$	$-\frac{1}{2}t_1$	$\frac{1}{2}t_1^2$	$-\frac{1}{2}t_1$
s_3	$\frac{1}{6}t_1^3 + t_3$	$\frac{1}{6}t_1^3 + t_1t_2$	$\frac{1}{4}t_1^2$	$\frac{1}{4}t_1^2$	$\frac{1}{3}t_1^3 + 2t_3$	$\frac{1}{4}t_1^2$
s_{21}	$\frac{1}{3}t_1^3 - t_3$	$\frac{1}{3}t_1^3$	0	0	$\frac{1}{3}t_1^3 - t_3$	0
s_{1^3}	$\frac{1}{6}t_1^3 + t_3$	$\frac{1}{6}t_1^3 - t_1t_2$	$-\frac{1}{4}t_1^2$	$-\frac{1}{4}t_1^2$	$\frac{1}{3}t_1^3 + 2t_3$	$-\frac{1}{4}t_1^2$
s_4	$\frac{1}{24}t_1^4 + t_3t_1$	$\frac{1}{24}t_1^4 + \frac{1}{2}t_2t_1^2 + \frac{1}{2}t_2^2 + t_4$	$\frac{1}{12}t_1^3 + \frac{1}{2}t_3$	$\frac{1}{12}t_1^3 + \frac{1}{2}t_3$	$\frac{1}{12}t_1^4 + 2t_1t_3$	$\frac{1}{12}t_1^3 + \frac{1}{2}t_3 + t_3'$
s_{31}	$\frac{1}{8}t_1^4$	$\frac{1}{8}t_1^4 + \frac{1}{2}t_2^2t_1 - \frac{1}{2}t_2^2 - t_4$	$\frac{1}{12}t_1^3 - t_3$	$\frac{1}{12}t_1^3 - t_3$	$\frac{1}{4}t_1^4$	$\frac{1}{12}t_1^3 - t_3 - t_3'$
s_{2^2}	$\frac{1}{12}t_1^4 - t_1t_3$	$\frac{1}{12}t_1^4 + t_2^2$	$\frac{1}{4}t_1^2$	$\frac{1}{4}t_1^2$	$\frac{1}{12}t_1^4 - t_1t_3$	$\frac{1}{4}t_1^2$
s_{21^2}	$\frac{1}{8}t_1^4$	$\frac{1}{8}t_1^4 - \frac{1}{2}t_2^2t_1 - \frac{1}{2}t_2^2 + t_4$	$-\frac{1}{12}t_1^3 + t_3$	$-\frac{1}{12}t_1^3 + t_3$	$\frac{1}{4}t_1^4$	$-\frac{1}{12}t_1^3 + t_3 + t_3'$
s_{1^4}	$\frac{1}{24}t_1^4 + t_3t_1$	$\frac{1}{24}t_1^4 - \frac{1}{2}t_2t_1^2 + \frac{1}{2}t_2^2 - t_4$	$-\frac{1}{12}t_1^3 - \frac{1}{2}t_3$	$-\frac{1}{12}t_1^3 - \frac{1}{2}t_3$	$\frac{1}{12}t_1^4 + 2t_1t_3$	$-\frac{1}{12}t_1^3 - \frac{1}{2}t_3 - t_3'$

Table 4.3: Simple Lie algebras and Schur polynomials of type (\mathfrak{g}, π)

Chapter 5

Hamiltonian structure of the hierarchy of type D_4 and the strong DR/DZ equivalence

The goal of this chapter is to confirm the strong double ramification / Dubrovin–Zhang (DR/DZ) equivalence in the case of the hierarchy of type D_4 . To do so, we describe the Hamiltonian structure of the DR hierarchy associated to the FJRW theory of type D_4 . Then we describe the Hamiltonian structure of the Drinfeld–Sokolov hierarchy of type D_4 , knowing that it coincides with the corresponding Dubrovin–Zhang hierarchy [GM05, FGM10]. Then using results of [BG16], we conclude on the equivalence (actually coincidence in this case) between the DR and DS (hence DZ) hierarchy of type D_4 . The first section is dedicated to the framework of infinite-dimensional Hamiltonian structures.

5.1 Infinite dimensional Hamiltonian structures

In this section we describe the purely algebraic approach to defining infinite-dimensional Hamiltonian structures. Different authors contributed to constructing this framework, among which Lax [Lax68, Lax75], Gardner, Kruskal, Miura and Zabusky [GKMZ70], Gelfand and Dickey [GD76, Dic03], played major roles. However, important contemporary definitions and notations were introduced by Dubrovin and Zhang [DZ08]. Here we take material from [DZ08, Bur15a, Ros17]. Some notions, such as the algebra of differential polynomials and the space of local functionals, were already introduced in the previous chapters. However, we will generalize these notions.

Differential polynomials, local functionals

Let us fix an integer $N \geq 1$ and consider the following commutative algebra of polynomials in infinitely many variables,

$$\mathcal{A} := \mathbb{C} \left[u_k^\alpha \mid \alpha \in \{1, \dots, N\}, k \geq 0 \right].$$

In general, we will denote the sequence of indeterminates $u_*^\alpha := (u_k^\alpha \mid \alpha \in \{1, \dots, N\}, k \geq 0)$. We define a derivation $\partial_x : \mathcal{A} \rightarrow \mathcal{A}$ by

$$\partial_x := \sum_{\alpha=1}^N \sum_{k \geq 0} u_{k+1}^\alpha \frac{\partial}{\partial u_k^\alpha}.$$

In particular, $u_k^\alpha = \partial_x^k u^\alpha$; we denote $u^\alpha := u_0^\alpha$. Elements of the algebra \mathcal{A} we call *differential polynomials*. It is equipped with a natural gradation given by $\deg u_k^\alpha = k$.

The indeterminates u_*^α are not functions of x but abstract variables. The notation ∂_x we explain as follows: For arbitrary functions $w^1, \dots, w^N \in C^\infty(\mathbb{R}, \mathbb{C})$ define $w_k^\alpha(x) := \partial^k w^\alpha / \partial x^k$, then the evaluation of a differential polynomial $f \in \mathcal{A}$ on the functions $w_*^\alpha(x)$ we define as $f \mapsto f(w_*^\alpha(x))$.

We then define the space of local functionals¹²

$$\mathcal{F} := \mathcal{A} / (\text{Im}(\partial_x) \oplus \mathbb{C}1), \quad (5.1)$$

whose elements are classes of differential polynomials modulo the addition of an exact derivative or a constant term. The class of a polynomial $f \in \mathcal{A}$ we denote alternatively by

$$\bar{f} = \int f = \int f(x) dx.$$

This comes with a integration by part formula $\overline{f \partial_x g} = -\overline{g \partial_x f}$, or more generally

$$\int f \partial_x^k g = (-1)^k \int g \partial_x^k f. \quad (5.2)$$

It is easy to check that for any $\alpha \in \{1, \dots, N\}$ and any $k \geq 1$,

$$\frac{\partial}{\partial u_k^\alpha} \partial_x - \partial_x \frac{\partial}{\partial u_k^\alpha} = \frac{\partial}{\partial u_{k-1}^\alpha}, \quad \frac{\partial}{\partial u^\alpha} \partial_x - \partial_x \frac{\partial}{\partial u^\alpha} = 0. \quad (5.3)$$

In particular, the derivatives $\partial / \partial u^\alpha$ are well defined on \mathcal{F} .

We also introduce the so-called *variational derivatives* which are linear endomorphisms denoted $\delta / \delta u^\alpha : \mathcal{A} \rightarrow \mathcal{A}$, for $\alpha \in \{1, \dots, N\}$, defined by, for $f \in \mathcal{A}$,

$$\frac{\delta f}{\delta u^\alpha} := \sum_{k \geq 0} (-1)^k \partial_x^k \frac{\partial f}{\partial u_k^\alpha}. \quad (5.4)$$

The following proposition is a fundamental and classical result. A purely algebraic proof was in given in [GKMZ70] (Lemma 2), which we follow here.

¹The traditional notation for \mathcal{F} is rather Λ , but we wanted to avoid confusion with the principal cyclic element Λ of an affine algebra, which we will use in the current chapter as well. The notation \mathcal{F} is used in [LWZ10].

²It is also frequent to do not divide by $\mathbb{C}1$ in the quotient (5.1); for our purpose is rather customary and convenient.

Proposition 5.1.1. *Let $f \in \mathcal{A}$ with no constant term. Then f is an exact derivative if and only if all the variational derivatives of f vanish; i.e.*

$$\exists g, f = \partial_x g \Leftrightarrow \forall \alpha, \frac{\delta f}{\delta u^\alpha} = 0.$$

Proof. In what follows, we use Einstein's notation and omit the sum over indices α, β . Let us first prove that if $f \in \text{Im}(\partial_x)$ (and in particular has no constant term), then each variational derivative $\delta f / \delta u^\alpha$ vanish. Let $f = \partial_x g \in \mathcal{A}$, then using Equation (5.3),

$$\begin{aligned} \frac{\delta f}{\delta u^\alpha} &= \sum_{k \geq 0} (-1)^k \partial_x^k \frac{\partial}{\partial u_k^\alpha} \partial_x g \\ &= \sum_{k \geq 1} (-1)^k \partial_x^k \frac{\partial g}{\partial u_{k-1}^\alpha} + \sum_{k \geq 0} (-1)^k \partial_x^{k+1} \frac{\partial g}{\partial u_k^\alpha} = 0. \end{aligned}$$

For the converse, let us introduce the following operator:

$$D_0 := \sum_{k \geq 0} u_k^\alpha \frac{\partial}{\partial u_k^\alpha}.$$

Then it is not hard to see that for any $f \in \mathcal{A}$ with no constant term, there exists a unique $g \in \mathcal{A}$ with no constant term such that $f = D_0 g$; we denote $g = D_0^{-1} f$. Moreover,

$$D_0 \partial_x = \partial_x D_0. \quad (5.5)$$

So let $f \in \mathcal{A}$ with no constant term such that $\delta f / \delta u^\alpha = 0$ for all α . Then using an integration by parts, we get that

$$\begin{aligned} 0 &= u^\alpha \frac{\delta f}{\delta u^\alpha} = u^\alpha \sum_{k \geq 0} (-1)^k \partial_x^k \frac{\partial f}{\partial u_k^\alpha} \\ &= u^\alpha \frac{\partial f}{\partial u^\alpha} - u^\alpha \partial_x \sum_{k \geq 1} (-1)^{k-1} \partial_x^{k-1} \frac{\partial f}{\partial u_k^\alpha} \\ &= u^\alpha \frac{\partial f}{\partial u^\alpha} - \partial_x \left[u^\alpha \sum_{k \geq 1} (-1)^{k-1} \partial_x^{k-1} \frac{\partial f}{\partial u_k^\alpha} \right] + u_1^\alpha \sum_{k \geq 1} (-1)^{k-1} \partial_x^{k-1} \frac{\partial f}{\partial u_k^\alpha}. \end{aligned}$$

We let the reader convince themselves that using the same procedure successively, we obtain that

$$0 = \sum_{k=0}^n u_k^\alpha \frac{\partial f}{\partial u_k^\alpha} - \partial_x \left[\sum_{k=0}^n u_k^\alpha \sum_{j \geq k+1} (-1)^{k-j-1} \frac{\partial f}{\partial u_k^\alpha} \right] + u_{n+1}^\alpha \sum_{k \geq n+1} (-1)^{k-j-1} \frac{\partial f}{\partial u_k^\alpha}.$$

Yet since $f \in \mathcal{A}$ is a polynomial, then the rightmost term in the above equation vanishes for higher values of n . It follows, by Equation (5.5), that for any $f \in \mathcal{A}$ with no constant term and such that $\delta f / \delta u^\alpha = 0$ for all α ,

$$f = \partial_x D_0^{-1} \left[\sum_{k \geq 0} u_k^\alpha \sum_{j \geq k+1} (-1)^{k-j-1} \frac{\partial f}{\partial u_k^\alpha} \right].$$

This concludes the proof of Proposition 5.1.1. \square

Local Poisson structures and Hamiltonian equations

Definition 5.1.2 (Local Poisson structure). A local Poisson structure is a matrix-valued differential operator over differential polynomials $K = (K^{\alpha\beta})$,

$$K^{\alpha\beta} = \sum_{j \geq 0} K_j^{\alpha\beta} \partial_x^j, \quad K_j^{\alpha\beta} \in \mathcal{A},$$

where the sum is finite, along with the action $K : \mathcal{F} \otimes \mathcal{F} \rightarrow \mathcal{F}$ defined by

$$\{\bar{f}, \bar{g}\}_K := \int \sum_{\alpha, \beta=1}^N \frac{\delta f}{\delta u^\alpha} K^{\alpha\beta} \frac{\delta g}{\delta u^\beta},$$

and which satisfies the following properties, for any $\bar{f}, \bar{g}, \bar{h} \in \mathcal{F}$,

- (i) Skew-symmetry: $\{\bar{f}, \bar{g}\}_K + \{\bar{g}, \bar{f}\}_K = 0$;
- (ii) Jacobi identity: $\{\bar{f}, \{\bar{g}, \bar{h}\}_K\}_K + \text{cyclic permutations} = 0$.

An important example is given by the so-called *standard Poisson structure* associated to a constant nondegenerate symmetric bilinear form $\eta^{\alpha\beta}$, namely

$$K^{\alpha\beta} = \eta^{\alpha\beta} \partial_x, \quad \eta^{\alpha\beta} \in \mathbb{C}.$$

The bracket reads

$$\{\bar{f}, \bar{g}\}_K := \int \sum_{\alpha, \beta=1}^N \frac{\delta f}{\delta u^\alpha} \eta^{\alpha\beta} \partial_x \left(\frac{\delta g}{\delta u^\beta} \right). \quad (5.6)$$

It is easy to show that this is indeed a local Poisson structure.

Notice that since $\mathcal{F} = \mathcal{A}/\text{Im}(\partial_x)$ is not a ring (because ∂_x is not a ring homomorphism), we cannot define Leibniz's rule, the latter being included in the definition of a Poisson bracket on a finite-dimensional manifold. Therefore, a local Poisson structure K makes \mathcal{F} into a Lie algebra $(\mathcal{F}, \{-, -\}_K)$, but not into a Poisson algebra.

However, a local Poisson structure K comes with a natural action $K : \mathcal{A} \otimes \mathcal{F} \rightarrow \mathcal{A}$ defined by

$$\{f, \bar{g}\}_K := \sum_{\alpha, \beta=1}^N \sum_{k \geq 0} \frac{\partial f}{\partial u_k^\alpha} \partial_x^k \left(K^{\alpha\beta} \frac{\delta f}{\delta u^\beta} \right), \quad (5.7)$$

which, by the integration by part formula (5.2), satisfies

$$\{\bar{f}, \bar{g}\}_K = \int \{f, \bar{g}\}_K.$$

Now the action (5.7) clearly satisfies Leibniz's rule w.r.t. to the first variable: for any differential polynomials $f, g, h \in \mathcal{A}$,

$$\{fg, \bar{h}\}_K = f\{g, \bar{h}\}_K + g\{f, \bar{h}\}_K.$$

A similar construction can be done w.r.t the second variable. This is as close as we can get to a Poisson algebra structure.

We can now define an infinite dimensional system of Hamiltonian PDEs.

Definition 5.1.3 (Hamiltonian structure). A system of Hamiltonian PDEs on \mathcal{A} is a system of PDEs of the form

$$\frac{\partial u^\alpha}{\partial t_i} = \{u^\alpha, \bar{h}_i\}_K = \sum_{\beta=1}^N K^{\alpha\beta} \frac{\delta \bar{h}_i}{\delta u^\beta},$$

where $\{h_i \mid i \geq 0\}$ is a family of differential polynomials called *Hamiltonian densities* which satisfy, for all $i, j \geq 0$,

$$\{\bar{h}_i, \bar{h}_j\}_K = 0.$$

Poisson algebra of Fourier expansion

Let \mathcal{B} be the subalgebra of $\mathbb{C}[[p_n^\alpha \mid \alpha \in \{1, \dots, N\}, n \in \mathbb{Z}^*]]$, where $\mathbb{Z}^* = \mathbb{Z} \setminus \{0\}$, which consists of formal power series of the form

$$f = \sum_{r \geq 0} \sum_{\substack{\alpha_1, \dots, \alpha_r \\ \in \{1, \dots, N\}}} \sum_{\substack{n_1, \dots, n_r \neq 0 \\ n_1 + \dots + n_r = 0}} f_{\alpha_1, \dots, \alpha_r}^{n_1, \dots, n_r} \prod_{j=1}^r p_{n_j}^{\alpha_j}, \quad f_{\alpha_1, \dots, \alpha_r}^{n_1, \dots, n_r} \in \mathbb{C}.$$

Given a nondegenerate symmetric bilinear form $\eta^{\alpha\beta}$, we endow the algebra \mathcal{B} with a Poisson bracket, defined on the indeterminates by

$$\{p_m^\alpha, p_n^\beta\}_\eta = im\eta^{\alpha\beta} \delta_{m+n, 0}, \quad (5.8)$$

and extended on \mathcal{B} by imposing Leibniz's rule, i.e.

$$\{p_m^\alpha \cdot p_n^\beta, p_r^\gamma\}_\eta := p_m^\alpha \cdot \{p_n^\beta, p_r^\gamma\}_\eta + p_n^\beta \cdot \{p_m^\alpha, p_r^\gamma\}_\eta$$

and so on. Now consider the homomorphism of algebras $T : \mathcal{A} \rightarrow \mathbb{C}[[p_*^*, x]]$ defined on the indeterminates by³

$$T(u_k^\alpha) := \sum_{n \in \mathbb{Z}^*} (in)^k p_k^\alpha e^{inx}, \quad (5.9)$$

and extended on \mathcal{A} by linearity and by imposing $T(fg) = T(f)T(g)$. The above can be thought of as a formal Fourier expansion of the “functions” u^α . For a differential polynomial $f \in \mathcal{A}$, we write

$$T(f) = \sum_{m \in \mathbb{Z}} T_m(f) e^{imx}, \quad T_m(f) \in \mathbb{C}[[p_*^*]],$$

³In Equation (5.9), the expression e^{inx} could have been replaced by z^n for some indeterminate z and therefore does not need to introduce $e^{inx} = \sum (inx)^k / k!$. The reason why we write specifically e^{inx} is because of the factor $(in)^k$ which agrees with the action of ∂_x on \mathcal{A} .

in particular, we consider the map $T_0 : f \mapsto T_0(f)$. Via the homomorphism property and multiple Cauchy product, the Fourier expansion of an arbitrary monomial $u_{k_1}^{\alpha_1} \cdots u_{k_r}^{\alpha_r} \in \mathcal{A}$ reads

$$T(u_{k_1}^{\alpha_1} \cdots u_{k_r}^{\alpha_r}) = \sum_{n \in \mathbb{Z}} e^{inx} \sum_{\substack{n_1 \in \mathbb{Z} \\ n_1 \neq n}} \sum_{\substack{n_2, \dots, n_r \neq 0 \\ n_1 + \dots + n_r = n_1}} (i(n - n_1))^{k_1} p_{n-n_1}^{\alpha_1} \prod_{j=2}^r (in_j)^{k_j} p_{n_j}^{\alpha_j}.$$

Then taking the coefficient in $n = 0$ and changing $n_1 \mapsto -n_1$ in the above sum, we have

$$T_0(u_{k_1}^{\alpha_1} \cdots u_{k_r}^{\alpha_r}) = \sum_{\substack{n_1, \dots, n_r \neq 0 \\ n_1 + \dots + n_r = 0}} \prod_{j=1}^r (in_j)^{k_j} p_{n_j}^{\alpha_j}. \quad (5.10)$$

Therefore, we are naturally brought to consider the subalgebra $\mathcal{B}^{\text{pol}} \subset \mathcal{B}$ which consists of power series of the form

$$f = \sum_{r \geq 0} \sum_{\substack{n_1, \dots, n_r \neq 0 \\ n_1 + \dots + n_r = 0}} f_{\alpha_1, \dots, \alpha_r}(n_1, \dots, n_r) \prod_{j=1}^r p_{n_j}^{\alpha_j}, \quad f_{\alpha_1, \dots, \alpha_r} \in \mathbb{C}_{\leq d}[z_1, \dots, z_r], \quad (5.11)$$

for some $d \geq 0$, meaning that the polynomials $f_{\alpha_1, \dots, \alpha_r}$ have a degree that is bounded from above independently of $\alpha_1, \dots, \alpha_r$. From Equation (5.10), we have

$$T_0(\mathcal{A}) \subset \mathcal{B}^{\text{pol}}.$$

On the other hand, we easily see that the map $T \circ \partial_x : \mathcal{A} \rightarrow \mathbb{C}[[p_*^*, x]]$ is given by

$$T(\partial_x f) = \sum_{m \in \mathbb{Z}} \text{im} f_m e^{imx}. \quad (5.12)$$

Indeed, because T is a homomorphism of algebras, the map $T \circ \partial_x$ is a derivation; therefore it suffices (5.12) to check the relation on the indeterminates u_k^α , for which (5.12) comes from the definition of T (5.9). It follows that $T_0 \circ \partial_x \equiv 0$, which implies that the following map, which we still denote T_0 , is well defined:

$$\begin{aligned} T_0 : \mathcal{F} &\longrightarrow \mathcal{B}^{\text{pol}} \\ \bar{f} &\longmapsto T_0(f). \end{aligned}$$

We give two important lemmas concerning the map T_0 which we take from [Bur15a] (see also [EGH00] § 2.2.3 and 2.9.2); proofs are given in the appendix of [Bur15a].

Lemma 5.1.4. *The map $T_0 : \mathcal{F} \rightarrow \mathcal{B}^{\text{pol}}$ is surjective. Its kernel is given by*

$$\text{Ker}(T_0) = \text{Span} \{\bar{u}^1, \dots, \bar{u}^N\} \subset \mathcal{F}.$$

Consequently, for any power series $f \in \mathcal{B}^{\text{pol}}$, there exists a unique local functional $\bar{h} \in \mathcal{F}$ such that $T_0(\bar{h}) = f$ and which is such that for all $\alpha \in \{1, \dots, N\}$,

$$\left. \frac{\partial \bar{h}}{\partial u^\alpha} \right|_{u_{k>0}^* = 0} = 0.$$

Lemma 5.1.5. *The pullback of the bracket $\{-, -\}_\eta$ on \mathcal{B}^{pol} (5.8) by the homomorphism T_0 is the standard bracket $\{-, -\}_{\eta \partial_x}$ (5.6) on \mathcal{F} .*

Extended spaces and Miura transforms

We now consider generalizations of differential polynomials and local functionals as follows. We define the *algebra of extended differential polynomials* by

$$\widehat{\mathcal{A}}_N := \mathbb{C}[[u^\alpha, \varepsilon]][u_{k>0}^\alpha].$$

We define a grading on it by $\deg \varepsilon = -1$ while $\deg u_k^\alpha = k$ still. We get a decomposition

$$\widehat{\mathcal{A}}_N = \bigoplus_{k \in \mathbb{Z}} \widehat{\mathcal{A}}^{[k]},$$

where $\widehat{\mathcal{A}}^{[k]}$ is the subspace of polynomials of homogeneous degree k . Within the subspace $\widehat{\mathcal{A}}^{[0]}$, the degree in ε counts the number of derivatives which appear in a monomial; in particular, each $u^\alpha \in \widehat{\mathcal{A}}^{[0]}$. For instance, $u_3 \varepsilon^3 + uu_1 \varepsilon \in \widehat{\mathcal{A}}^{[0]}$.

The derivation

$$\partial_x = \sum_{k \geq 0} u_{k+1}^\alpha \frac{\partial}{\partial u_k^\alpha}$$

naturally acts on $\widehat{\mathcal{A}}$ as well so that we generalize local functionals to the *space of extended local functionals* by

$$\widehat{\mathcal{F}} := \widehat{\mathcal{A}} / (\text{Im}(\partial_x) \oplus \mathbb{C}[[\varepsilon]]).$$

This induces the same gradation

$$\mathcal{F} = \bigoplus_{k \in \mathbb{Z}} \mathcal{F}^{[k]}.$$

We can now introduce the group of Miura transforms which plays a key role in the theory we present here. They are named after American mathematician Robert M. Miura, of whom we already cited several major contributions to the domain.

Definition 5.1.6 (Miura transform). A Miura transform is any map $\varphi : \widehat{\mathcal{A}} \rightarrow \widehat{\mathcal{A}}$ which is such that $\tilde{u}^\alpha := \varphi(u^\alpha) \in \widehat{\mathcal{A}}^{[0]}$ and which satisfies

$$\det \left[\frac{\partial \tilde{u}^\alpha}{\partial u^\beta} \Big|_{\varepsilon=0} \right]_{\alpha, \beta \in [1, N]} \neq 0.$$

The group of such maps we call the *Miura group*.

Miura transforms will play the role of change of coordinate on $\widehat{\mathcal{A}}$. We shall emphasize here the importance of the parameter ε . Indeed, for example, the following simple transformation is a Miura transform:

$$\tilde{u} = u + u_1 \varepsilon + u_2 \varepsilon^2 + u_3 \varepsilon^3 + \dots$$

Its inverse is

$$u = \tilde{u} - \tilde{u}_1 \varepsilon - \tilde{u}_2 \varepsilon^2 - \tilde{u}_3 \varepsilon^3 - \dots$$

As we can see, by solving the equation for each power of ε , we find the inverse of the Miura transform. Such transformations would have obviously been impossible to encode in the ring \mathcal{A} . Thus we get a much broader definition of “change of variables.” We can motivate the introduction of such transformation through the applications to the DR hierarchy as we will explain in the next sections.

We now generalize what was done on Fourier expansion to the extended case. For any $d \geq 0$, consider the subalgebra $\mathcal{B}_d^{\text{pol}} \subset \mathcal{B}^{\text{pol}}$ consisting elements of the form of Equation (5.11) such that the polynomials $f_{\alpha_1, \dots, \alpha_n} \in \mathbb{C}[z_1, \dots, z_n]$ are homogeneous of degree d . Then define the algebra $\widehat{\mathcal{B}} := \mathcal{B} \otimes \mathbb{C}[[\varepsilon]]$ and let $\widehat{\mathcal{B}}^{\text{pol}}$ be the subalgebra of $\widehat{\mathcal{B}}$ consisting of power series of the form

$$f = \sum_{i \geq 0} \varepsilon^i f_i, \quad f_i \in \mathcal{B}_i^{\text{pol}}.$$

Then with the exact same procedure we can define a map

$$\widehat{T}_0 : \widehat{\mathcal{F}}^{[0]} \rightarrow \widehat{\mathcal{B}}^{\text{pol}}$$

which is also surjective and which kernel is still spanned by the functionals \bar{u}^α .

Extended Hamiltonian and local Poisson structures

We generalize the notion of local Poisson structure and Hamiltonian densities to the extended context as follows.

Definition 5.1.7 (Poisson and Hamiltonian extended structures). A local Poisson structure on the space of extended local functionals $\widehat{\mathcal{F}}$ is a matrix-valued differential operator $K = (K^{\alpha\beta})$ of the form

$$K = \sum_{j \geq 0} K^{\alpha\beta} \partial_x^j, \quad K_j^{\alpha\beta} \in \mathcal{A}^{[1-j]}.$$

where the sum is finite, along with the action, as before, given by

$$\{\bar{f}, \bar{g}\}_K = \int \sum_{\alpha, \beta=1}^N \frac{\delta f}{\delta u^\alpha} K^{\alpha\beta} \left(\frac{\delta g}{\delta u^\beta} \right),$$

where the variational derivatives are given by the same Equation (5.4). A system of Hamiltonian PDEs on extended differential polynomials $\widehat{\mathcal{A}}$ is a system of PDEs of the form

$$\frac{\partial u^\alpha}{\partial t_i} = \{u^\alpha, \bar{h}_j\}_K = \sum_{\beta=1}^N K^{\alpha\beta} \frac{\delta \bar{h}_j}{\delta u^\beta}, \quad \bar{h}_i \in \widehat{\mathcal{F}}^{[0]},$$

where $\{\bar{h}_i \mid i \geq 0\}$ is a family of extended local functionals $\bar{h}_i \in \widehat{\mathcal{F}}^{[0]}$ which satisfy, for any $i, j \geq 0$,

$$\{\bar{h}_i, \bar{h}_j\} = 0.$$

The skew-symmetry and Jacobi identity conditions on K are non trivial. In particular, the ordinary differential polynomial $K^{\alpha\beta}|_{\varepsilon=0} \in \mathcal{A}_N$ must have the form

$$K^{\alpha\beta}|_{\varepsilon=0} = g^{\alpha\beta}(u)\partial_x + \Gamma_Y^{\alpha\beta}(u)u_1^Y, \quad (5.13)$$

and be such that the following hold [DN83].

- (i) Under Miura transforms $\tilde{u}^\alpha := \varphi(u^\alpha)$ of the type $\partial\tilde{u}^\alpha/\partial u_{k>0}^\beta = 0$, the matrix $g^{\alpha\beta}$ transforms like a (2,0) tensor, i.e.

$$(g \circ \varphi)^{\alpha\beta} = \frac{\partial\tilde{u}^\alpha}{\partial u^\mu} \frac{\partial\tilde{u}^\beta}{\partial u^\nu} g^{\mu\nu},$$

while the matrix $\Gamma_{\alpha\beta}^Y := -g_{\alpha\mu}\Gamma_\beta^{\mu Y}$, where $(g_{\alpha\beta}) := g^{-1}$, transforms like a connection matrix $\nabla = d + \Gamma$, i.e.

$$(\Gamma \circ \varphi)_{\alpha\beta}^Y = \frac{\partial u^\mu}{\partial \tilde{u}^\alpha} \frac{\partial u^\nu}{\partial \tilde{u}^\beta} \Gamma_{\mu\nu}^\sigma \frac{\partial \tilde{u}^Y}{\partial u^\sigma} + \frac{\partial \tilde{u}^Y}{\partial u^\tau} \frac{\partial^2 u^\tau}{\partial \tilde{u}^\alpha \partial \tilde{u}^\beta};$$

- (ii) The skew-symmetry condition $\{\bar{f}, \bar{g}\} + \{\bar{g}, \bar{f}\} = 0$ implies that $g^{\alpha\beta}$ be symmetric and that ∇ be compatible with g in the sense that $\nabla_\gamma g^{\alpha\beta} = 0$ for all γ ;
 (iii) Jacobi identity implies that the torsion and curvature tensors of the connection ∇ vanish. Moreover, ∇ is the Levi-Civita connection associated to the metric g .

We now introduce the formal Dirac delta notation (see e.g. [DZ08]): for any $f, g \in \mathcal{A}$ and $k \geq 0$, we denote

$$\int f(x)\delta^{(k)}(x-y)g(y)dy := f\partial_x^k g, \quad (5.14)$$

$$\int f(x)\delta^{(k)}(x-y)g(y)dx := (-1)^k g\partial_x^k f. \quad (5.15)$$

The difference in the two equation above lies in the variable w.r.t. which we “integrate.” In other words, we introduce a symbol $\delta^{(k)}(x-y)$, where x and y are actually purely notational, which we see as two different bilinear maps $\widehat{\mathcal{A}} \rightarrow \widehat{\mathcal{A}}$.

We shall emphasize here that the integral symbol, in the above, is entirely different from the one used so far. Indeed, in Equations (5.14) and (5.15), the integral symbol is purely notational and does not represent the class modulo derivative. However, if we denote by $\pi : \widehat{\mathcal{A}} \rightarrow \widehat{\mathcal{F}}$ the projection on local functionals, then we get a double integration formula which, thanks to the integration by parts, reads

$$\iint f(x)\delta^{(k)}(x-y)g(y)dydx := \pi \int f(x)\delta^{(k)}(x-y)g(y)dy = \pi \int f(x)\delta^{(k)}(x-y)g(y)dx.$$

Now the above equation is somehow very strange since we have two integral symbols: one for the action of the delta symbol, the other for the projection onto $\widehat{\mathcal{F}}$. But since the two operations commute, it makes sense not to distinguish the two integral symbols.

We can then rewrite the Poisson bracket as

$$\{\bar{f}, \bar{g}\} = \iint \frac{\delta f(x)}{\delta u^\alpha(x)} \sum_{j \geq 0} K_j^{\alpha\beta}(x) \delta^{(j)}(x-y) \frac{\delta g(y)}{\delta u^\beta(y)} dx dy.$$

We denote

$$\{u^\alpha(x), u^\beta(y)\} := \sum_{j \geq 0} K_j^{\alpha\beta}(x) \delta^{(j)}(x-y).$$

It is easy to see, using the rules defined for the symbol δ , that

$$\{u^\alpha(x), u^\beta(y)\} = \iint \frac{\delta u^\alpha(z) \delta(x-z)}{\delta u^\mu(z)} \sum_{j \geq 0} K_j^{\alpha\beta}(u(z)) \delta^{(j)}(z-z') \frac{\delta u^\alpha(z') \delta(z'-y)}{\delta u^\mu(z')} dz dz'. \quad (5.16)$$

Notice the following useful identity, which is only a consequence of Leibniz's rule,

$$\begin{aligned} & \iint f(z) \delta^{(n)}(x-z) g(z) \delta^{(m)}(z-z') h(z') \delta^{(p)}(y-z') dz dz' \\ &= \int f(x) \partial_x^n \left(g(x) \delta^{(m)}(x-z') \right) h(z') \delta^{(p)}(y-z') dz' \\ &= \int (f(x) \partial_z^n \circ g(z') \partial_{z'}^m) h(z') \delta^{(p)}(y-z') dz'. \end{aligned} \quad (5.17)$$

Proposition 5.1.8. Let $\varphi : \widehat{\mathcal{A}} \rightarrow \widehat{\mathcal{A}}$, $\tilde{u}^\alpha := \varphi(u^\alpha)$ be a Miura transform. Given a local Poisson structure $K : \widehat{\mathcal{F}} \otimes \widehat{\mathcal{F}} \rightarrow \widehat{\mathcal{F}}$, the Miura transform φ acts on K via the following formula,

$$(\varphi^* K)^{\alpha\beta} = \sum_{\mu, \nu=1}^N (M^*)_\mu^\alpha \circ K^{\mu\nu} \circ M_\nu^\beta, \quad (5.18)$$

where the matrices M and M^* are given by

$$(M^*)_\mu^\alpha = \sum_{k \geq 0} \frac{\partial \tilde{u}^\alpha}{\partial u_k^\mu} \partial_x^k, \quad M_\nu^\beta = \sum_{\ell \geq 0} (-1)^\ell \partial_x^\ell \circ \frac{\partial \tilde{u}^\beta}{\partial u_\ell^\nu},$$

where the composition of differential operators (in particular, with a function) is still the one defined in Section 3.1.

Proof. Let us start by developing the first variational derivative:

$$\begin{aligned} \{\tilde{u}^\alpha(x), \tilde{u}^\beta(y)\} &:= \iint \frac{\delta \tilde{u}^\alpha(z) \delta(x-z)}{\delta u^\mu(z)} \sum_{j \geq 0} K^{\alpha\beta}(z) \delta^{(j)}(z-z') \frac{\delta \tilde{u}^\beta(z') \delta(y-z')}{\delta u^\nu(z')} dz dz' \\ &= \iint \sum_{k \geq 0} (-1)^k \partial_z^k \left(\frac{\partial \tilde{u}^\alpha(z)}{\partial u_k^\mu(z)} \delta(x-z) \right) \sum_{j \geq 0} K^{\alpha\beta}(z) \delta^{(j)}(z-z') \frac{\delta \tilde{u}^\beta(z') \delta(y-z')}{\delta u^\nu(z')} dz dz' \\ &= \iint \sum_{k \geq 0} \frac{\partial \tilde{u}^\alpha(z)}{\partial u_k^\mu(z)} \delta(x-z) \partial_z^k \left(\sum_{j \geq 0} K^{\alpha\beta}(z) \delta^{(j)}(z-z') \right) \frac{\delta \tilde{u}^\beta(z') \delta(y-z')}{\delta u^\nu(z')} dz dz'. \end{aligned}$$

We can now integrate w.r.t. the variable z and use Equation (5.17) to find

$$\begin{aligned} \{\tilde{u}^\alpha(x), \tilde{u}^\beta(y)\} &= \int \sum_{k \geq 0} \frac{\partial \tilde{u}^\alpha(x)}{\partial u_k^\mu(x)} \partial_x^k \left(\sum_{j \geq 0} K^{\alpha\beta}(x) \delta^{(j)}(x - z') \right) \frac{\delta \tilde{u}^\beta(z') \delta(y - z')}{\delta u^\nu(z')} dz' \\ &= \int \sum_{k \geq 0} \frac{\partial \tilde{u}^\alpha(x)}{\partial u_k^\mu(x)} \partial_{z'}^k \circ \sum_{j \geq 0} K^{\alpha\beta}(z') \partial_{z'}^j \frac{\delta \tilde{u}^\beta(z') \delta(y - z')}{\delta u^\nu(z')} dz' \\ &= \int (M^*)_\mu^\alpha(x, z') \circ K^{\mu\nu}(x, z') \frac{\delta \tilde{u}^\beta(z') \delta(y - z')}{\delta u^\nu(z')} dz', \end{aligned}$$

where

$$K^{\mu\nu}(x, z') := \sum_{j \geq 0} K^{\mu\nu}(x) \partial_{z'}^j,$$

and accordingly for $(M^*)_\mu^\alpha(x, z')$. Now by developing the second variational derivative and using Equation (5.17) again, we find (notice that we can swap y and z' in $\delta(y - z')$)

$$\begin{aligned} \{\tilde{u}^\alpha(x), \tilde{u}^\beta(y)\} &= \int (M^*)_\mu^\alpha(x, z') \circ K^{\mu\nu}(x, z') \sum_{\ell \geq 0} (-1)^\ell \partial_{z'}^\ell \left(\frac{\partial \tilde{u}^\beta(z')}{\partial u_\ell^\nu(z')} \delta(z' - y) \right) dz' \\ &= (M^*)_\mu^\alpha(x, y) \circ K^{\mu\nu}(x, y) \sum_{\ell \geq 0} (-1)^\ell \partial_y^\ell \circ \frac{\partial \tilde{u}^\beta(y)}{\partial u_\ell^\nu(y)}. \end{aligned}$$

Hence the result. \square

In classifying all possible local Poisson structures modulo Miura transforms, the following powerful Darboux-like theorem is due to Erza Getzler.

Theorem 5.1.9 ([Get02]). *Let K be a local Poisson structure with genus 0 part*

$$K^{\alpha\beta}|_{\varepsilon=0} = g^{\alpha\beta}(u) \partial_x + \Gamma_Y^{\alpha\beta}(u) u_1^Y.$$

If $g^{\alpha\beta}(u)$ is nondegenerate, than K is Miura equivalent to a standard Poisson structure. More precisely, there exist a unique constant symmetric invertible matrix $\eta^{\alpha\beta} \in \mathbb{C}$ and a Miura transform $\varphi : \hat{\mathcal{A}} \rightarrow \hat{\mathcal{A}}$ such that

$$(\varphi^* K)^{\alpha\beta}(u) = \eta^{\alpha\beta} \partial_x.$$

Tau symmetry and general notion of tau function

We end this section with the definition of tau structure for Hamiltonian densities as given in [DZ08]. We have already encountered tau symmetry in Chapter 2, §4.3.4, where we followed the approach of Wu [Wu17] to define a tau function for any Drinfeld–Sokolov hierarchy. His achievement was in particular to offer a universal construction (rather than a case by case study). Wu’s construction is then a specific case of a much broader definition of tau function.

Let us make clear the vocabulary thereafter. In the context we are about to describe, a *tau structure* is a specific choice of Hamiltonian densities for a Hamiltonian system which satisfy

the so-called *tau symmetry* (which amounts to the closedness of a 1-form). The *tau function* then naturally arise via Poincaré lemma. As for the most part of what precedes, the following content is taken from [Ros17].

Definition 5.1.10 (Tau structure). Let $K : \widehat{\mathcal{F}} \otimes \widehat{\mathcal{F}} \rightarrow \widehat{\mathcal{F}}$ be a local Poisson structure. Let $\{h_{\alpha,p} \mid \alpha \in \{1, \dots, N\}, p \geq -1\} \subset \widehat{\mathcal{A}}^{[0]}$ be a family of differential polynomials, such that⁴

$$\{\bar{h}_{\alpha,p}, \bar{h}_{\beta,q}\} = 0, \quad p, q \geq -1, \alpha, \beta \in \{1, \dots, N\}.$$

Consider the Hamiltonian system of PDEs it generates, namely

$$\frac{\partial u^\alpha}{\partial t_p^\beta} = \{u^\alpha, \bar{h}_{\beta,p}\} = K^{\alpha\mu} \frac{\delta h_{\beta,p}}{\delta u^\mu}. \quad (5.19)$$

We say that the Hamiltonian densities $h_{\alpha,p}$ form a tau structure of the system (5.19) if the following holds.

(i) The functional $\bar{h}_{1,0}$ is *compatible with spatial translation* in the sense that

$$\frac{\partial u^\alpha}{\partial t_0^1} = u_1^\alpha; \quad (5.20)$$

(ii) The functionals $\bar{h}_{\beta,-1}$ generate trivial flows, i.e. for all $\beta \in \{1, \dots, N\}$,

$$\frac{\partial u^\alpha}{\partial t_{-1}^\beta} = K^{\alpha\mu} \frac{\delta h_{\beta,-1}}{\delta u^\mu} = 0;$$

(iii) The functionals $\bar{h}_{\beta,-1}$ are linearly independent;

(iv) The densities $h_{\alpha,p}$ satisfy the *tau symmetry* property, for all $\alpha, \beta \in \{1, \dots, N\}, p, q \geq 0$,

$$\frac{\partial h_{\alpha,p-1}}{\partial t_q^\beta} = \frac{\partial h_{\beta,q-1}}{\partial t_p^\alpha}. \quad (5.21)$$

Tau symmetry (5.21) is simply asking that the 1-form $\sum h_{\alpha,p-1} dt_p^\alpha$ is closed. It can also be written as

$$\{h_{\alpha,p-1}, \bar{h}_{\beta,q}\} = \{h_{\beta,q-1}, \bar{h}_{\alpha,p}\}.$$

Since the functionals $\bar{h}_{\beta,-1}$ are linearly independent and of degree 0, it follows that the map $\varphi : u^\alpha \mapsto \eta^{\alpha\beta} h_{\beta,-1}$ is a Miura transform, where $\eta^{\alpha\beta}$ is the unique constant symmetric invertible matrix of Getzler's theorem 5.1.9. We call the coordinates

$$\tilde{u}^\alpha := \eta^{\alpha\beta} h_{\beta,-1}(u_*^*, \varepsilon) \quad (5.22)$$

⁴The choice of having two indices α, p for the Hamiltonian densities $h_{\alpha,p}$ seems arbitrary and is not necessary in order to define the tau symmetry property. However, it is actually intrinsically linked to the definition of a tau structure. Moreover, such a way of ordering Hamiltonian densities makes sense when working with the Drinfeld–Sokolov hierarchies since the set of affine exponent has a natural periodic structure modulo the Coxeter number (see Appendix A).

the *normal coordinates* associated to the tau structure. In the normal coordinates, the local Poisson structure takes the following form [DZ08] (see §3.3),

$$(K \circ \varphi)^{\alpha\beta} = \eta^{\alpha\beta} \partial_x + \mathcal{O}(\varepsilon).$$

We now explain how the existence of a tau structure allows to define a tau function in a way entirely similar to what as done in §4.3.4. The commutativity condition $\{\bar{h}_{\alpha,p-1}, \bar{h}_{\beta,q}\} = 0$ amounts to

$$\int \frac{\partial h_{\alpha,p-1}}{\partial t_q^\beta} = 0.$$

Since $\partial h_{\alpha,p-1}/\partial t_q^\beta$ has no constant term, it follows that there exists a unique differential polynomial $\Omega_{\alpha,p;\beta,q} \in \widehat{\mathcal{A}}^{[0]}$ with no constant term either and such that⁵

$$\partial_x \Omega_{\alpha,p;\beta,q} = \frac{\partial h_{\alpha,p-1}}{\partial t_q^\beta}. \quad (5.23)$$

In particular, thanks to the spatial translation compatibility condition (5.20)

$$h_{\alpha,p-1} = \Omega_{\alpha,p;1,0}.$$

The following proposition holds.

Proposition 5.1.11 (Tau function). *Let $u^* \in \mathbb{C}[[t_*^*; \varepsilon]]$ be an arbitrary solution of the Hamiltonian system (5.19). Then there exists a function $\tau \in \mathbb{C}[[t_*^*; \varepsilon]]$ such that for any $\alpha, \beta \in \{1, \dots, N\}$ and $p, q \geq 0$,*

$$\Omega_{\alpha,p;\beta,q} = \frac{\partial^2 \log \tau}{\partial t_p^\alpha \partial t_q^\beta}.$$

In particular, by identifying $x = t_0^1$, if $\tilde{u}^\alpha = \eta^{\alpha\beta} h_{\beta,-1}$ are the normal coordinates of the system (5.19), then

$$\tilde{u}^\alpha = \eta^{\alpha\beta} \frac{\partial^2 \log \tau}{\partial x \partial t_0^\beta}.$$

Notice that, depending on the literature, either τ is called the *tau function* or $F = \exp(\tau)$ is. Oftentimes though, F is called the *potential* or the *free energy* of system (5.19).

Proof. We will proceed by applying Poincaré lemma twice. We start from the defining equation of the Omega function (5.23) which, using the commutativity of the flows $[\partial_x, \partial_{t_r^\gamma}] = 0$, implies

$$\partial_x \frac{\partial}{\partial t_r^\gamma} \Omega_{\alpha,p;\beta,q} = \frac{\partial}{\partial t_r^\gamma} \frac{\partial}{\partial t_q^\beta} h_{\alpha,p-1}. \quad (5.24)$$

⁵This Omega function should of course be compared with the Omega function introduced by Wu; see Equation (4.57). The latter however was matrix-valued.

Now the commutativity of the flows allows to swap t_r^γ and t_q^β , while the tau symmetry condition (5.21) allows to swap $h_{\alpha,p-1}$ and t_q^β . All in all, (5.24) is completely symmetric in permutation of the indices $((\alpha, q), (\beta, q), (\gamma, r))$. Moreover, since we chose $\Omega_{\alpha,p;\beta,q}$ such that it contains no constant term, then $\partial_{t_r^\gamma} \Omega_{\alpha,p;\beta,q}$ is itself completely symmetric.

Let us fix the indices (α, p) . Then we have

$$\frac{\partial}{\partial t_r^\gamma} \Omega_{\alpha,p;\beta,q} = \frac{\partial}{\partial t_q^\beta} \Omega_{\alpha,p;\gamma,r}.$$

This means that the 1-form $\sum \Omega_{\alpha,p;\beta,q} dt_r^\gamma$ is closed. Therefore, by Poincaré lemma, there exists a function $A_{\alpha,p}$ such that

$$\Omega_{\alpha,p;\beta,q} = \frac{\partial A_{\alpha,p}}{\partial t_q^\beta}.$$

Now the tau symmetry property, again, means that $\Omega_{\alpha,p;\beta,q} = \Omega_{\beta,q;\alpha,p}$ so that the 1-form $\sum A_{\alpha,p} dt_p^\alpha$ is also closed. Therefore, there exists another function τ such that

$$\Omega_{\alpha,p;\beta,q} = \frac{\partial^2 \log \tau}{\partial t_p^\alpha \partial t_q^\beta}.$$

□

Finally, we state the following proposition; a proof will be found in [DZ08] (see Theorem 3.1.9 and §3.3). Notice however that checking point (ii) of Proposition 5.1.12 is immediate.

Proposition 5.1.12 (Normal Miura transforms). *Let \tilde{u}^α be the normal coordinates of the tau structure. Let $f(\tilde{u}_*^\alpha) \in \widehat{\mathcal{A}}^{[-2]}$ and let $\tilde{\varphi} : \widehat{\mathcal{A}} \rightarrow \widehat{\mathcal{A}}$, $\tilde{v}^\alpha := \tilde{\varphi}(\tilde{u}^\alpha)$ be the Miura transform defined by*

$$\tilde{v}^\alpha = \tilde{u}^\alpha + \eta^{\alpha\beta} \partial_x \{f, \bar{h}_{\beta,0}\}.$$

Consider the following differential polynomials:

$$g_{\alpha,p}(\tilde{v}_*^\alpha) = h_{\alpha,p} \circ \tilde{\varphi}^{-1}(\tilde{v}_*^\alpha) + \partial_x \{f, \bar{h}_{\beta,q+1}\} \circ \tilde{\varphi}^{-1}(\tilde{v}_*^\alpha).$$

Then the $g_{\alpha,p}$'s form another tau structure of the hierarchy.

Moreover, let $\tilde{u}_*^\alpha(t_*^\alpha) \in \mathbb{C}[[t_*^\alpha; \varepsilon]]$ be any solution of the hierarchy and let $\tilde{v}_*^\alpha(t_*^\alpha) = \tilde{\varphi}(\tilde{u}_*^\alpha(t_*^\alpha))$ be the corresponding solution of the hierarchy modified by the Miura transform. Then the tau function $\tau_{\tilde{u}}(t_*^\alpha)$ associated to $\tilde{u}_*^\alpha(t_*^\alpha)$ and the tau function $\tau_{\tilde{v}}(t_*^\alpha)$ associated to $\tilde{v}_*^\alpha(t_*^\alpha)$ are related by

$$\log \tau_{\tilde{v}}(t_*^\alpha) = \log \tau_{\tilde{u}}(t_*^\alpha) + f(\tilde{u}_*^\alpha(t_*^\alpha; \varepsilon); \varepsilon)|_{x=0}.$$

Such Miura transforms together with the chosen transformation of the Hamiltonian densities form a group whose elements are called normal Miura transforms.

5.2 The double ramification hierarchy of type D_4

The goal of this section is to compute the local Poisson structure $K_{\text{DR}}^{\alpha\beta}$, the Hamiltonian functional $\overline{h}_{1,1}^{\text{DR}}$ and the Hamiltonian densities $h_{0,1}^{\text{DR}}$ to $h_{0,4}^{\text{DR}}$ (in normal coordinates) of the DR hierarchy of type D_4 . Due to results of [BG16], these are the only data that we need to compute in order to confirm the strong DR/DZ equivalence (see Theorem 5.2.24). In first three sections, we give the definitions of cohomological field theories (CohFT) [KM94]; the definition of FJRW theory [FJR11, FJR13] (a special case of CohFT which we use for the hierarchy of type D_4); we describe the construction of the DR hierarchy [Bur15a] and finally, in section 5.2.4 we describe the Hamiltonian structure of the DR hierarchy of type D_4 .

5.2.1 Cohomological field theories, FJRW theory

In this section we introduce the notion of cohomological field theories. We mostly follow the survey [Ros17] and refer to [Zvo12] for more details and proofs. We then introduce Fan–Jarvis–Ruan–(inspired by)Witten (FJRW) theory, taking material from introductions given in [FJ16, LRZ15, Gue17].

Moduli spaces of stable curves

We recall the definition and some properties of the (Deligne–Mumford compactification of the) moduli spaces of stable curves $\overline{\mathcal{M}}_{g,n}$; we will not give much details on the topic and refer to [Zvo12]. First consider the set of all connected compact Riemann surfaces (or equiv. connected projective complex curves) C such that

- ◊ C has genus g ;
- ◊ C is smooth but for a finite number of transverse nodes, say $\{y_1, \dots, y_r\} \subset C$;
- ◊ C has n distinct non nodal marked point $\{x_1, \dots, x_n\} \subset C \setminus \{y_1, \dots, y_r\}$;

where by transverse we mean that each node has local model $xy = 0$. We call a *special point* any marked point or node. We say that two n -marked genus g curves $(C, (x_1, \dots, x_n))$ and $(C', (x'_1, \dots, x'_n))$ are isomorphic if there exists a bi-holomorphism $\phi : C \rightarrow C'$ such that image of a node in C is a node of C' and such that $\phi(x_i) = x'_i$. We say that a curve C is stable if its automorphism group $\text{Aut}(C)$ is finite. The stability condition is equivalent to the following constraint:

$$2g - 2 + n > 0. \tag{5.25}$$

The moduli space $\overline{\mathcal{M}}_{g,n}$, for $g, n \geq 0$ satisfying Equation (5.25), is then the set of all equivalence classes $[C]$ of stable curves isomorphic to C ,

$$\overline{\mathcal{M}}_{g,n} = \{(C; (x_1, \dots, x_n))\} / \sim .$$

It is a compact complex orbifold (or a smooth Deligne–Mumford stack) of complex dimension

$$\dim_{\mathbb{C}} \overline{\mathcal{M}}_{g,n} = 3g - 3 + n. \tag{5.26}$$

On $\overline{\mathcal{M}}_{g,n}$ there is a morphism of orbifolds

$$\pi : \overline{\mathcal{C}}_{g,n} \rightarrow \overline{\mathcal{M}}_{g,n}$$

called the universal curve, whose fibre over a point $x \in \overline{\mathcal{M}}_{g,n}$ is homeomorphic to the curve C of which the point x is the isomorphism class, i.e. $\pi^{-1}(x) \simeq C$. So there are n marked points on $\tilde{x}_1, \dots, \tilde{x}_n$ on $\pi^{-1}(x)$; this gives us n sections

$$s_i : \overline{\mathcal{M}}_{g,n} \rightarrow \overline{\mathcal{C}}_{g,n}$$

which associate to a point $x \in \overline{\mathcal{M}}_{g,n}$ the marked point $\tilde{x}_i \in \pi^{-1}(x)$.

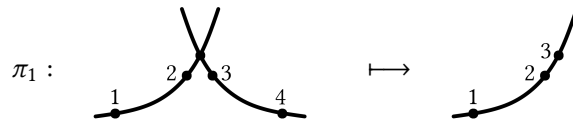
Forgetful, loop and tree morphisms

There are three natural morphisms defined on the moduli spaces of stable curves. The *forgetful morphism*, for any $m \geq 1$,

$$\pi_m : \overline{\mathcal{M}}_{g,n+m} \rightarrow \overline{\mathcal{M}}_{g,n},$$

forgets the last m marked points on a curve and contracts all components of the curve that might have become unstable in the process.

For example, for a curve $C \in \overline{\mathcal{M}}_{0,4}$ consisting of two \mathbb{P}^1 's intersecting at a node with two marked points on each \mathbb{P}^1 component, the forgetful morphism $\pi_1 : \overline{\mathcal{M}}_{0,4} \rightarrow \overline{\mathcal{M}}_{0,3}$ contracts the whole \mathbb{P}^1 component containing the 4th marked point on the node, the latter becomes the third marked point on the resulting \mathbb{P}^1 . This is illustrated below; labelled dots represent marked points, while unlabelled dots represent nodes.

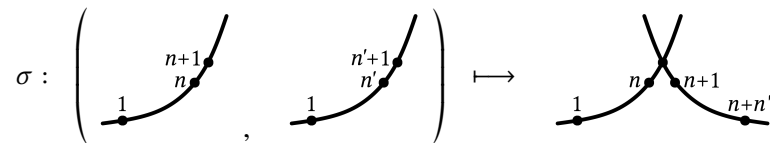


It can be shown that the forgetful map $\pi_1 : \overline{\mathcal{M}}_{g,n+1} \rightarrow \overline{\mathcal{M}}_{g,n}$ coincides with the universal curve $\pi : \overline{\mathcal{C}}_{g,n} \rightarrow \overline{\mathcal{M}}_{g,n}$; hence in what follows we consider indistinctly $\pi = \pi_1$.

The tree morphism

$$\sigma : \overline{\mathcal{M}}_{g,n+1} \times \overline{\mathcal{M}}_{g',n'+1} \rightarrow \overline{\mathcal{M}}_{g+g',n+n'}$$

glues two (classes of) stable curves $y \in \overline{\mathcal{M}}_{g,n+1}$ and $y' \in \overline{\mathcal{M}}_{g',n'+1}$ by identifying their respective last marked points, say x_{n+1} and $x'_{n'+1}$, to a single point in $\sigma(y, y')$ which becomes a stable node. The new node separates⁶ the resulting curve into two components. This is illustrated below.

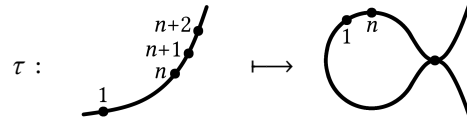


⁶We say that a node separates the curve if its removal disconnects the curve.

The loop morphism

$$\tau : \overline{\mathcal{M}}_{g,n+2} \rightarrow \overline{\mathcal{M}}_{g+1,n}$$

glues one curve onto itself by identifying the two last marked points to a node of the resulting curve. The new node does not separate the resulting curve into two components. Hence, the loop morphism is sometimes called the *non separating gluing map* while the tree morphism is sometimes called the *separating gluing map*. This is illustrated below.



Tautological bundles

On the universal curve $\overline{C}_{g,n}$ we define a line bundle $\varrho : \Omega_{g,n} \rightarrow \overline{C}_{g,n}$ as follows. On the non-nodal points $\tilde{x} \in \overline{C}_{g,n}$, the fibre $\varrho^{-1}(\tilde{x}) \subset \Omega_{g,n}$ is defined as the relative cotangent space w.r.t. the projection $\pi : \overline{C}_{g,n} \rightarrow \overline{\mathcal{M}}_{g,n}$. Then it extends canonically to the singular points and defines an actual line bundle over $\overline{C}_{g,n}$.

We recall the definition of the relative cotangent bundle though. Let \tilde{x} be a nonsingular point in $\overline{C}_{g,n}$. Consider the pullback bundle $\pi^*T^*\overline{\mathcal{M}}_{g,n} \rightarrow \overline{C}_{g,n}$ defined by

$$\pi^*T^*\overline{\mathcal{M}}_{g,n} = \left\{ (\tilde{x}, \alpha) \in \overline{C}_{g,n} \times T^*_{\pi(\tilde{x})}\overline{\mathcal{M}}_{g,n} \right\} \quad (5.27)$$

Then consider the map $\mathring{\pi} : \pi^*T^*\overline{\mathcal{M}}_{g,n} \rightarrow T^*\overline{C}_{g,n}$ defined on fibres by

$$\begin{aligned} \mathring{\pi}_{\tilde{x}} : T^*_{\pi(\tilde{x})}\overline{\mathcal{M}}_{g,n} &\longrightarrow T^*_{\tilde{x}}\overline{C}_{g,n} \\ \alpha_{\pi(\tilde{x})} &\longmapsto \alpha_{\pi(\tilde{x})} \circ \pi_{*,\tilde{x}} \end{aligned}$$

where $\pi_{*,\tilde{x}} : T_{\tilde{x}}\overline{C}_{g,n} \rightarrow T_{\pi(\tilde{x})}\overline{\mathcal{M}}_{g,n}$ is the pushforward associated to π . Then the relative cotangent bundle $\varrho : \Omega_{g,n} \rightarrow \overline{C}_{g,n}$ has fibres defined by

$$\varrho^{-1}(\tilde{x}) := T^*_{\tilde{x}}\overline{C}_{g,n} / \text{Im}(\mathring{\pi}_{\tilde{x}}).$$

Then one can prove that if $x \in \overline{\mathcal{M}}_{g,n}$ and $\tilde{x} \in \pi^{-1}(x)$, then $\varrho^{-1}(\tilde{x})$ is isomorphic to the cotangent to the curve $\pi^{-1}(x) \subset \overline{C}_{g,n}$ at the point \tilde{x} , that is, $\varrho^{-1}(\tilde{x}) \simeq T^*_{\tilde{x}}\pi^{-1}(x)$.

Now recall the sections $s_i : \overline{\mathcal{M}}_{g,n} \rightarrow \overline{C}_{g,n}$ which map a class of curve $x \in \overline{\mathcal{M}}_{g,n}$ to the i -th marked point $\tilde{x}_i = s_i(x)$ on the fibre $\pi^{-1}(x)$. We can pullback the bundle $\Omega_{g,n}$ to a bundle

$$L_i := s_i^*\Omega_{g,n} \rightarrow \overline{\mathcal{M}}_{g,n}$$

whose fibre at $x \in \overline{\mathcal{M}}_{g,n}$ is defined by

$$L_i|_x = \varrho^{-1}(\tilde{x}_i) \simeq T^*_{\tilde{x}_i}\pi^{-1}(x).$$

The fibre bundle L_i we call the i -th *tautological bundle*. Here is the situation in terms of diagrams:

$$\begin{array}{ccc}
\Omega_{g,n} & & L_i = s_i^* \Omega_{g,n} \\
\downarrow \varrho & & \downarrow \\
\overline{C}_{g,n} & \xrightarrow{\pi} & \overline{\mathcal{M}}_{g,n} \\
& \xleftarrow{s_i} &
\end{array}$$

We denote the first Chern class of L_i by

$$\psi_i = c_1(L_i) \in H^2(\overline{\mathcal{M}}_{g,n}, \mathbb{Q}).$$

We call the classes ψ_1, \dots, ψ_n the *psi classes*. (The first Chern class establishes an isomorphism between isomorphism classes of complex line bundles over $\overline{\mathcal{M}}_{g,n}$ and the homology group $H^2(\overline{\mathcal{M}}_{g,n}, \mathbb{Q})$, in particular $c_1(L \otimes L') = c_1(L) + c_1(L')$.)

The Hodge bundle

In a similar way we can define the *Hodge bundle*

$$\mathbb{H} \rightarrow \overline{\mathcal{M}}_{g,n}$$

as the rank- g vector bundle over $\overline{\mathcal{M}}_{g,n}$ whose fibre over $x \in \overline{\mathcal{M}}_{g,n}$ is the space of abelian differentials on the curve $\pi^{-1}(x)$. We recall that an abelian differential on a nodal curve C is a meromorphic differential on C of which the poles satisfy (i) all poles are at a node, (ii) all poles are at most simple (i.e. have order 0 or 1) and (iii) the residue at a node on two branches meeting at this node are opposite to each other.

We denote by

$$\lambda_i = c_i(\mathbb{H}) \in H^{2i}(\overline{\mathcal{M}}_{g,n}, \mathbb{Q}),$$

for $i \in \{1, \dots, g\}$, the i -th Chern class of \mathbb{H} ; we also denote

$$\Lambda(s) = \sum_{i=0}^g s^i \lambda_i \in H(\overline{\mathcal{M}}_{g,n}, \mathbb{Q})[s].$$

More formally, we define \mathbb{H} as the pushforward bundle

$$\mathbb{H} := \pi_* \Omega_{g,n} \rightarrow \overline{\mathcal{M}}_{g,n}.$$

Recall that on nonsingular point $x \in \overline{\mathcal{M}}_{g,n}$, the fibre of $\pi_* \Omega_{g,n}$ is defined as

$$\pi_* \Omega_{g,n} = \left\{ (x, \alpha) \in \overline{\mathcal{M}}_{g,n} \times \Omega_{g,n} \mid \pi \circ \varrho(\alpha) = x \right\}.$$

Then we clearly have, on nonsingular points $x \in \overline{\mathcal{M}}_{g,n}$, that $\mathbb{H}|_x = (\pi_* \Omega_{g,n})|_x = T^* \pi^{-1}(x)$. In the end, the situation is as follows:

$$\begin{array}{ccccc}
\Omega_{g,n} & & \mathbb{H} = \pi_* \Omega_{g,n} & & L_i = s_i^* \Omega_{g,n} \\
\downarrow \rho & & \downarrow & & \swarrow \\
\overline{\mathcal{C}}_{g,n} & \xrightarrow{\pi} & \overline{\mathcal{M}}_{g,n} & & \\
& \searrow & \swarrow & & \\
& & s_i & &
\end{array}$$

We end this section with the definition of the tautological ring. The tautological ring $R^*(\overline{\mathcal{M}}_{g,n})$ can be defined as the smallest \mathbb{Q} -subalgebra of $H^*(\overline{\mathcal{M}}_{g,n}, \mathbb{Q})$ that is closed under the pushforward along any of the forgetful, tree and loop morphisms π, σ, τ defined previously. In particular, it contains the psi classes ψ_1, \dots, ψ_n and the lambda classes $\lambda_1, \dots, \lambda_g$. It is generated by yet other classes; we refer to [Zvo12] for more details.

Cohomological field theories

Definition 5.2.1 (Cohomological field theory). Let $g, n \geq 0$ be such that $2g - 2 + n > 0$. Let V be an N -dimensional complex vector space with basis (e_1, \dots, e_N) and equipped with a symmetric nondegenerate bilinear form $\eta = (\eta^{\alpha\beta})$. A cohomological field theory (CohFT) is a system of linear maps

$$c_{g,n} : V^{\otimes n} \rightarrow H^*(\overline{\mathcal{M}}_{g,n}, \mathbb{C})$$

satisfying the following axioms.

- (i) $c_{g,n}$ is equivariant w.r.t. to simultaneous permutation of copies of V in $V^{\otimes n}$ and of the marked point on the curves;
- (ii) For the specific case $g = 0, n = 3$,

$$c_{0,3}(e_1, e_\alpha, e_\beta) = \eta^{\alpha\beta};$$

- (iii) For the forgetful map $\pi_1 = \pi : \overline{\mathcal{M}}_{g,n+1} \rightarrow \overline{\mathcal{M}}_{g,n}$;

$$\pi^* c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = c_{g,n+1}(e_{\alpha_1}, \dots, e_{\alpha_n}, e_1);$$

- (iv) For the tree morphism $\sigma : \overline{\mathcal{M}}_{g_1, n_1+1} \times \overline{\mathcal{M}}_{g_2, n_2+1} \rightarrow \overline{\mathcal{M}}_{g,n}$, with $g = g_1 + g_2$ and $n = n_1 + n_2$,

$$\sigma^* c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = c_{g_1, n_1+1}(e_1, \dots, e_{n_1}, e_\mu) \eta^{\mu\nu} c_{g_2, n_2+1}(e_\nu, e_1, \dots, e_{n_2});$$

- (v) For the loop morphism $\tau : \overline{\mathcal{M}}_{g,n+2} \rightarrow \overline{\mathcal{M}}_{g+1,n}$,

$$\tau^* c_{g+1,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = \eta^{\mu\nu} c_{g,n+2}(e_\mu, e_\nu, e_{\alpha_1}, \dots, e_{\alpha_n}).$$

Given a CohFT $\{c_{g,n} \mid g, n \geq 0\}$, we define the *intersection numbers* (or *correlators*) associated to the CohFT by

$$\langle \tau_{d_1}(e_{\alpha_1}) \cdots \tau_{d_n}(e_{\alpha_n}) \rangle_g := \int_{\overline{\mathcal{M}}_{g,n}} c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) \cdot \psi_1^{d_1} \cdots \psi_n^{d_n} \in \mathbb{C}.$$

These numbers are 0 if the degree of $c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) \cdot \psi_1^{d_1} \cdots \psi_n^{d_n}$ is not $3g - 3 + n$. We then define the so-called *potential* F of a CohFT as the generating series of the intersection numbers of the CohFT, namely,

$$F(t_*^*; \varepsilon) = \sum_{g \geq 0} \varepsilon^{2g} F_g(t_*^*),$$

where

$$F_g(t_*^*) := \sum_{\substack{n \geq 0 \\ 2g - 2 + n > 0}} \frac{1}{n!} \sum_{d_1, \dots, d_n \geq 0} \langle \tau_{d_1}(e_{\alpha_1}) \cdots \tau_{d_n}(e_{\alpha_n}) \rangle \cdot t_{d_1}^{\alpha_1} \cdots t_{d_n}^{\alpha_n}. \quad (5.28)$$

Example 5.2.2 (Trivial CohFT, Hodge CohFT). The trivial CohFT is the one for which the vector space $V = \mathbb{C}$ is one-dimensional, the metric is $\eta = 1$ and the CohFT is simply $c_{g,n} = 1$. In a similar way, the Hodge CohFT is the one for which, as above, $V = \mathbb{C}$ and $\eta = 1$, yet $c_{g,n} = \Lambda(s) = \sum_{i=0}^g \lambda_i s^i$.

Example 5.2.3 (Gromov–Witten theory). Let us first recall the definition of the moduli spaces of stable maps $\overline{\mathcal{M}}_{g,n}(X, \beta)$. Let X be a smooth projective variety and let $\beta \in H_2(X, \mathbb{Z})$. Consider a stable n -marked genus g curve $(C, (x_1, \dots, x_n))$ together with a morphism $f : C \rightarrow X$ such that $f_*[C] = \beta$. We say that two such triplets $(C, (x_i), f)$ and $(C', (x'_i), f')$ are isomorphic if there exists a biholomorphism $\phi : C \rightarrow C'$ such that $\phi(x_i) = \phi(x'_i)$ (i.e. the two stable curves are isomorphic) and such that $f' \circ \phi = f$. We say that $(C, (x_i), f)$ is a *stable map* if it has a finite number of automorphism for the relation we just defined. The moduli space of stable maps is then the set of isomorphism classes of stable maps,

$$\overline{\mathcal{M}}_{g,n}(X, \beta) := \{(C; (x_1, \dots, x_n); f : C \rightarrow X) \mid f_*[C] = \beta\} / \sim.$$

It is a singular stack and carries a virtual complex dimension

$$\text{vdim}_{\mathbb{C}} \overline{\mathcal{M}}_{g,n}(X, \beta) = 3g - 3 + n + (1 - g) \dim_{\mathbb{C}} X + \int_{\beta} c_1(\text{TX}).$$

We have two natural maps $\overline{\mathcal{M}}_{g,n}(X, \beta)$: the first one,

$$p : \overline{\mathcal{M}}_{g,n}(X, \beta) \rightarrow \overline{\mathcal{M}}_{g,n}$$

takes a stable map $(C, (x_i), f)$ and maps it to the stable curve $(C, (x_i)) \in \overline{\mathcal{M}}_{g,n}$; the second,

$$\text{ev} : \overline{\mathcal{M}}_{g,n}(X, \beta) \rightarrow X^n$$

takes a stable map $(C, (x_i), f)$ and maps it to $(f(x_1), \dots, f(x_n)) \in X^n$.

The associated CohFT is defined as follows. Take the vector space $V = H^*(X, \mathbb{C})$ equipped with the pairing defined by

$$\eta(v, w) = \int_{\mathcal{P}(v)} w,$$

for any $v, w \in H^*(X, \mathbb{C})$ where $\mathcal{P} : H^*(X) \rightarrow H_*(X)$ is the Poincaré duality which gives a ring isomorphism $H^k(X, \mathbb{C}) \rightarrow H_{d-k}(X, \mathbb{C})$, where $d = \dim_{\mathbb{R}} X$. Then define the linear maps

$$c_{g,n}(e_{\alpha_1}, \dots, e_{\alpha_n}) = p_* \text{ev}^*(e_{\alpha_1}, \dots, e_{\alpha_n}).$$

We refer to [BF97] for more details on the virtual fundamental class and how to define the Gromov–Witten classes.

Example 5.2.4 (Witten’s r -spin classes). For this example, we give a very brief description; the next one, Fan–Jarvis–Ruan–Witten theory generalizes Witten’s r -spin classes. Let $r \geq 2$ and consider the vector space $V = \mathbb{C}^{r-1}$. We equip V with the metric $\eta^{\alpha\beta} = \delta_{\alpha+\beta, r}$. In [PV16], Polishchuk and Vaintrob constructed classes

$$c_{g,n}(e_{\alpha_1+1}, \dots, e_{\alpha_n+1}) \in H^*(\overline{\mathcal{M}}_{g,n}, \mathbb{Q})$$

that have degree

$$d_{\alpha_1, \dots, \alpha_r} = \frac{1}{r}(r-2)(g-1) + \frac{1}{r} \sum_{i=1}^n \alpha_i$$

if the $a_i \in \{0, \dots, r-2\}$ are such that $d_{\alpha_1, \dots, \alpha_r}$ is a nonnegative integer, and vanish otherwise. These classes are defined via pushingforward to $\overline{\mathcal{M}}_{g,n}$ Witten’s virtual class on the moduli space of curves with r -spin structures. An r -spin structure on a smooth n -marked curve $(C, (x_1, \dots, x_n))$ is an r -th root L of the twisted canonical bundle $K(\sum \alpha_i x_i)$ of the curve. This notions will be clarified in the next section.

The last example we give is the FJRW theory associated to a pair (W, G) where W is a quasi-homogeneous polynomial and G is a certain group acting on polynomials and leaving W invariant. The computations performed in Section 5.3 relate to the FJRW theory associated to the Dynkin diagram D_4 so we will discuss FJRW theory in a whole section.

Fan–Jarvis–Ruan–Witten theory

In this section we recall the main ideas underlying the Fan–Jarvis–Ruan–(inspired by)Witten (FJRW) theory introduced in [FJR11, FJR13]. We refer, for instance, to the survey [FJ16] by T. J. Jarvis and A. Francis, but also to [Gue17] for a brief review. First we have the following definition.

Definition 5.2.5 (Quasihomogeneity, nondegeneracy). A polynomial $W \in \mathbb{C}[z_1, \dots, z_m]$ is called *quasihomogeneous* if there exist $q_1, \dots, q_m \in \mathbb{Q}$, called *weights*, such that for any $\lambda \in \mathbb{C}^*$,

$$W(\lambda^{q_1} z_1, \dots, \lambda^{q_m} z_m) = \lambda W(z_1, \dots, z_m).$$

A quasihomogeneous polynomial W is called *nondegenerate* if the weights q_i are unique and if W has an isolated singularity at the origin.

By an *isolated singularity at the origin* we mean that the differential of W vanishes at $(0, \dots, 0)$ and that $(0, \dots, 0)$ is a connected component of the zero locus the differential. For any quasihomogeneous polynomial W we define a group called the *maximal diagonal symmetry group* of W defined by

$$G_{\max}^W := \{(\gamma_1, \dots, \gamma_m) \in (\mathbb{C}^*)^m \mid \gamma \cdot W := W(\gamma_1 z_1, \dots, \gamma_m z_m) = W(z_1, \dots, z_m)\}.$$

We will often drop the superscript W when there is no ambiguity. The maximal diagonal symmetry group G_{\max} always contains and special element called the *exponential grading element*,

$$J := (e^{2i\pi q_1}, \dots, e^{2i\pi q_m}) \in G_{\max}.$$

Definition 5.2.6 (Admissible symmetry group). Let $W \in \mathbb{C}[z_1, \dots, z_m]$ be a quasihomogeneous nondegenerate polynomial. A subgroup $G \subset G_{\max}$ of the maximal diagonal symmetry group of W is called *admissible* if $J \in G$.

We call the *central charge* of W the rational number

$$\hat{c}_W := \sum_{i=1}^m (1 - 2q_i).$$

If $\hat{c}_W < 1$ then W is said to have a *simple singularity* at the origin. The latter have been classified into the *ADE* singularities, as described in Table 5.2.1

Table 5.1: Simple singularities and related simply-laced Dynkin diagrams

Type	Dynkin diagram	Singularity	Type	Dynkin diagram	Singularity
$A_n, \quad n \geq 1$		x^{n+1}	E_6		$x^3 + y^4$
$D_n, \quad n \geq 4$		$x^2 y + y^{n-1}$	E_7		$x^3 y + y^3$
			E_8		$x^3 + y^5$

Now let $W \in \mathbb{C}[z_1, \dots, z_m]$ be a quasihomogeneous nondegenerate polynomial and let G be an admissible symmetry group of W . Let $\gamma \in G$ and let us denote

$$\text{Fix}(\gamma) := \{(z_1, \dots, z_n) \in \mathbb{C}^n \mid \gamma \cdot (z_1, \dots, z_n) = (z_1, \dots, z_n)\}.$$

We denote

$$W_\gamma := W|_{\text{Fix}(\gamma)}$$

We can now define the vector space underlying the CohFT of the FJRW theory of a pair (W, G) . In this introduction, we define the vector space of the FJRW CohFT associated to a pair (W, G) as follows.

Definition 5.2.7 (State space). The vector space of the FJRW CohFT associated to a pair (W, G) , which we call the *state space*, we define by

$$\mathcal{H}^{W,G} = \bigoplus_{\gamma \in G} \mathcal{H}_\gamma^{W,G} := \bigoplus_{\gamma \in G} \left(\mathcal{Q}_{W_\gamma} \otimes dz_\gamma \right)^G,$$

where the superscript G denotes the invariant part, $dz_\gamma = \bigwedge_{z_i \in \text{Fix}(\gamma)} dz^i$ and where \mathcal{Q}_{W_γ} is the Jacobian ring of W_γ , i.e.

$$\mathcal{Q}_{W_\gamma} = \mathbb{C}[z_i \in \text{Fix}(\gamma)] / \langle \partial_{z_i} W \mid z_i \in \text{Fix}(\gamma) \rangle.$$

The state space $\mathcal{H}^{W,G}$ is equipped with a nondegenerate symmetric bilinear form $\eta : \mathcal{H}^{W,G} \otimes \mathcal{H}^{W,G} \rightarrow \mathbb{C}$; we refer to [PV16] (Equation (4)) and [CIR14] (Equation (5.12)) for the latter.

Remark 5.2.8. For the special case $\gamma = J$, which always belong to G by definition, it is not hard to see that

$$\dim_{\mathbb{C}} \mathcal{H}_J = 1.$$

Indeed, since $J = (e^{2i\pi q_1}, \dots, e^{2i\pi q_m})$ since each $q_i < 1$, then $B_J = \emptyset$. Therefore, $\dim_{\mathbb{C}} \mathcal{Q}_{W_J} = 1$ and $dz_J = 1$. Consequently, we will denote $\mathbf{1}_J \in \mathcal{H}_J^{W,G}$ the sole generator of $\mathcal{H}_J^{W,G}$.

We now briefly define the moduli space $\mathcal{S}_{g,n}^{W,G}$ of (W, G) -spin curves as it appeared in [FJR11, FJR13]. First, we start with a nodal orbicurve (orbifold curve) C with genus g and n marked points (x_1, \dots, x_n) . We define on C a line bundle $\Omega_{\log, C} \rightarrow C$ by

$$\Omega_{\log, C} := \Omega_C (x_1 + \dots + x_n),$$

where Ω_C is the canonical bundle over C . Roughly speaking, $\Omega_{\log, C}$ consists of meromorphic 1-forms with poles of order 1 at the marked points.

Definition 5.2.9 ((W, G) -spin structure). Let $W \in \mathbb{C}[z_1, \dots, z_m]$ be a quasihomogeneous non-degenerate polynomial and G be an admissible symmetry group. Let us decompose W as a sum of monomials

$$W = \sum_{k=1}^r b_k W_k = \sum_{k=1}^r b_k \prod_{i=1}^m z_i^{a_{i,k}}.$$

Let C be an orbicurve of genus g with n marked points together with m orbifold line bundles $L_1, \dots, L_m \rightarrow C$ such that there exist r isomorphisms of orbifold bundles ϕ_1, \dots, ϕ_r such that

$$\phi_k : W_k(L_1, \dots, L_m) := \bigotimes_k L_i^{\otimes a_{i,k}} \xrightarrow{\sim} \Omega_{\log, C}. \quad (5.29)$$

The data $(C; (x_1, \dots, x_n); (L_1, \dots, L_m); (\phi_1, \dots, \phi_r))$ is called an (W, G) -spin structure.

We say that two (W, G) -spin structures $(C; (x_i); (L_j); (\phi_k))$ and $(C'; (x'_i); (L'_j); (\phi'_k))$ (both of genus g and with n marked points) are isomorphic if there exists an isomorphism of orbicurves $\psi : C \rightarrow C'$ such that $\psi(x_i) = x'_i$ and satisfying the condition

$$\psi^*(L'_i) \simeq L_i \quad (5.30)$$

for all $i \in \{1, \dots, m\}$. Notice that because $\psi : C \rightarrow C'$ is an isomorphism, we automatically have $\Omega_{\log, C} \simeq \Omega_{\log, C'}$ which, together with (5.30), implies that the following diagram commutes,

$$\begin{array}{ccc} W_k(L) & \xrightarrow{\phi_k} & \Omega_{\log, C} \\ \downarrow & & \downarrow \\ W_k(L') & \xrightarrow{\phi'_k} & \Omega_{\log, C'} \end{array}$$

where $W_k(L) = W_k(L_1, \dots, L_m)$ and accordingly for $W_k(L')$. Then we define the moduli space $\mathcal{S}_{g,n}^{W,G}$ of (W, G) -spin structures as the set of equivalence classes modulo isomorphism as described above,

$$\mathcal{S}_{g,n}^{W,G} := \{(C; (x_1, \dots, x_n); (L_1, \dots, L_m); (\phi_1, \dots, \phi_r))\} / \sim.$$

This a compact complex orbifold (or a smooth Deligne–Mumford stack).

There is a natural forgetful morphism which forgets both the orbifold structure of the orbicurve and the (W, G) -spin structure; we denote it

$$\pi : \mathcal{S}_{g,n}^{W,G} \rightarrow \overline{\mathcal{M}}_{g,n}.$$

In [FJR11] it is shown that, as a consequence of Equation (5.29), the orbifold structure of a (W, G) -spin structure on an orbicurve C at its marked points (x_1, \dots, x_n) is characterized by a n -tuple $\vec{\gamma} = (\gamma(1), \dots, \gamma(n)) \in G^{\times n}$. This in turns induces a decomposition

$$\mathcal{S}_{g,n}^{W,G} = \bigoplus_{\vec{\gamma} \in G^{\times n}} \mathcal{S}_{g,n}^{W,G}(\vec{\gamma}),$$

which is such that $\mathcal{S}_{g,n}^{W,G}(\vec{\gamma}) \neq \emptyset$ only if

$$\gamma(1) \cdots \gamma(n) = J^{2g-2+n}. \quad (5.31)$$

Still in [FJR11], Fan, Jarvis and Ruan constructed a *virtual fundamental class* for each $\vec{\gamma} = (\gamma(1), \dots, \gamma(n)) \in G^{\times n}$ such that Equation (5.31) holds (we do not define virtual class here)

$$[\mathcal{S}_{g,n}^{W,G}(\vec{\gamma})]^{\text{vir}} \in H_*(\mathcal{S}_{g,n}^{W,G}, \mathbb{Q}).$$

We can now define the CohFT classes $c_{g,n}$ associated to the FJRW theory of a pair (W, G) we just described.

Theorem 5.2.10 (FJRW CohFT [FJR13]). *Let $W \in \mathbb{C}[z_1, \dots, z_n]$ be a quasihomogeneous non-degenerate polynomial and G be an admissible symmetry group. Let $\vec{\gamma} = (\gamma(1), \dots, \gamma(n)) \in G^{\times n}$ be such that Equation (5.31) holds. For each $i \in \{1, \dots, n\}$ let $\alpha_i \in \mathcal{H}_{\gamma(i)}^{W,G}$. Define the following cohomology classes,*

$$c_{g,n}^{W,G}(\alpha_1, \dots, \alpha_n) = \frac{|G|^g}{\deg(\pi)} \mathcal{P} \circ \pi_* \left([\mathcal{S}_{g,n}^{W,G}(\vec{\gamma})]^{\text{vir}} \cap \alpha_1 \cap \cdots \cap \alpha_n \right),$$

where $|G|$ is the cardinal of G and \mathcal{P} is the Poincaré duality. These classes extend by linearity to maps

$$c_{g,n}^{W,G} : \mathcal{H}^{W,G} \rightarrow H^*(\overline{\mathcal{M}}_{g,n}, \mathbb{Q})$$

which define a CohFT in the sense of Definition 5.2.1 with $e_1 = 1_J \in \mathcal{H}_J^{W,G}$ (see Remark 5.2.8).

5.2.2 The double ramification hierarchies

The double ramification cycles

Here we define the double ramification cycles, introduced in [GV05], following the article [Bur15a]. Let integers $a_1, \dots, a_n \in \mathbb{Z}$ be such that

$$\sum_{i=1}^n a_i = 0. \quad (5.32)$$

To such integers, we will assign a space of *rubber stable maps* with target variety \mathbb{P}^1 relative to $0, \infty \in \mathbb{P}^1$ (stable maps in the sense of Gromov–Witten theory, see Example 5.2.3).

First recall the definition of the moduli space of stable maps associated to a pair (X, β) , where X is a smooth compact complex manifold and $\beta \in H_2(X, \mathbb{Q})$:

$$\overline{\mathcal{M}}_{g,n}(X, \beta) := \{(C; (x_1, \dots, x_n); f : C \rightarrow X) \mid f_*[C] = \beta\} / \sim,$$

where $(C; (x_1, \dots, x_n))$ is a stable curve and $f : C \rightarrow X$ is holomorphic map. Notice that in the particular case $X = \mathbb{P}^1$, we have $\dim H_2(\mathbb{P}^1, \mathbb{Q}) = 1$, hence we can consider $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^1)$ without ambiguity. Let $(C; (x_1, \dots, x_n); f)$ be the representative of a point in $\overline{\mathcal{M}}_{g,n}(\mathbb{P}^1)$; we say that $f : C \rightarrow \mathbb{P}^1$ has *ramification profile* (a_1, \dots, a_n) w.r.t. (x_1, \dots, x_n) if for any $i \in \{1, \dots, n\}$,

$$\begin{cases} a_i > 0 \Rightarrow x_i \text{ is a pole of } f \text{ of order } a_i; \\ a_i < 0 \Rightarrow x_i \text{ is a zero of } f \text{ of order } -a_i. \end{cases} \quad (5.33)$$

In particular, if $a_i > 0$ (resp. $a_i < 0$) then $f(x_i) = 0$ (resp. $f(x_i) = \infty$).⁷ If $a_i = 0$, then x_i is just a regular marked point and $f(x_i) \neq 0, \infty$.

In what follows, we will denote

$$\underline{a} := (a_1, \dots, a_n)$$

with vanishing sum when there is no ambiguity on n . We define the following moduli space of *rubber stable maps*,

$$\overline{\mathcal{M}}_{g;\underline{a}} := \left\{ (C; (x_1, \dots, x_n); f) \in \overline{\mathcal{M}}_{g,n}(\mathbb{P}^1) \mid f : C \rightarrow \mathbb{P}^1 \text{ satisfies (5.33)} \right\} / \mathbb{C}^*,$$

where \mathbb{C}^* acts on the target \mathbb{P}^1 . There is a natural forgetful morphism

$$\phi : \overline{\mathcal{M}}_{g;\underline{a}} \rightarrow \overline{\mathcal{M}}_{g,n},$$

which simply returns the underlying stable curve $(C; (x_1, \dots, x_n))$. The moduli space of rubber stable maps carries a virtual fundamental class

$$[\overline{\mathcal{M}}_{g;\underline{a}}]^{\text{vir}} \in H_{2(2g-3+n)}(\overline{\mathcal{M}}_{g;\underline{a}}, \mathbb{C}).$$

⁷Notice that this explains the vanishing sum condition $\sum a_i = 0$ (5.32). Indeed, for $f : C \rightarrow \mathbb{P}^1$, Cauchy's argument principles imposes that the sum of the order of the zeros of f equals the sum of the order of its poles (see e.g. [Mir95], Corollary 3.18, p. 124).

Definition 5.2.11 (Double ramification cycle). Let $a_1, \dots, a_n \in \mathbb{Z}$ be integers with vanishing sum. We call the *double ramification* (DR) cycle associated to the ramification profile \underline{a} and denote $\text{DR}_g(\underline{a})$ the pushforward of the virtual fundamental class of the moduli space $\overline{\mathcal{M}}_{g,\underline{a}}$ by the forgetful morphism $\pi : \overline{\mathcal{M}}_{g,\underline{a}} \rightarrow \overline{\mathcal{M}}_{g,n}$, i.e.

$$\text{DR}_g(\underline{a}) := \phi_*[\overline{\mathcal{M}}_{g,\underline{a}}]^{\text{vir}} \in H_{2(2g-3+n)}(\overline{\mathcal{M}}_{g,n}, \mathbb{C}).$$

The DR cycles were introduced by Graber and Vakil [GV05]. It follows that the Poincaré dual $\mathcal{P}(\text{DR}_g(\underline{a}))$ of the DR cycle belongs to $H^{2g}(\overline{\mathcal{M}}_{g,n}, \mathbb{C})$. Moreover, it is known that $\mathcal{P}(\text{DR}_g(\underline{a}))$ lies in the tautological ring of $\overline{\mathcal{M}}_{g,n}$ [FP05]. As an immediate consequence of its definition, the pullback of the DR cycle $\text{DR}_g(\underline{a})$ by the forgetful map $\pi_1 : \overline{\mathcal{M}}_{g,n+1} \rightarrow \overline{\mathcal{M}}_{g,n}$ reads

$$\pi_1^* \text{DR}_g(a_1, \dots, a_n) = \text{DR}(a_1, \dots, a_n, 0).$$

Notice that in genus 0, and for $n \geq 3$, the DR cycle coincides with the fundamental class of the moduli space of stable curves [GJV11]:

$$\text{DR}_0(\underline{a}) = [\overline{\mathcal{M}}_{0,n}] \in H_{2(n-3)}(\overline{\mathcal{M}}_{0,n}, \mathbb{C}).$$

Polynomiality and Hain's formula

Let us introduce the moduli space $\mathcal{M}_{g,n}^{\text{ct}} \subset \overline{\mathcal{M}}_{g,n}$ of nodal stable curves of which every node is separating (again, it means that removing any node breaks the curve into two disjoint components); such curves are called of *compact type*. Then Hain [Hai13] showed that, restricted to the moduli space $\mathcal{M}_{g,n}^{\text{ct}}$, the (Poincaré dual of the) DR cycles is tautological and is a polynomial in the a_i 's which is homogeneous of degree $2g$. Namely

$$\mathcal{P}(\text{DR}_g(a_1, \dots, a_n)) \Big|_{\mathcal{M}_{g,n}^{\text{ct}}} = \frac{(-1)^g}{g! 4^g} \left[\sum_{J \subset \{1, \dots, n\}} \sum_{h=0}^g (\underline{a}_J)^2 \delta_h^J \right]^g,$$

where $\underline{a}_J = \sum_{j \in J} a_j$. The divisor $\delta_h^J \in H^2(\overline{\mathcal{M}}_{g,n}, \mathbb{C})$ is the class of curves C with only one node, say $y \in C$, which are such that y separates C into two smooth components, i.e. the normalization of $C \setminus \{y\}$ is a disjoint union $C_1 \sqcup C_2$, and such that C_1 has genus h and contains the marked points $\{x_j\}_{j \in J}$, while C_2 has genus $g-h$ and contains the remaining marked points. Besides we set $\delta_0^{\{i\}} = -\psi_i$.

Because the top lambda class λ_g vanishes outside $\mathcal{M}_{g,n}^{\text{ct}}$, it follows that for any cohomology class $\alpha \in H^*(\overline{\mathcal{M}}_{g,n}, \mathbb{C})$, the pairing

$$\int_{\text{DR}_g(\underline{a})} \lambda_g \alpha \in \mathbb{C}_{2g}[a_1, \dots, a_n], \quad (5.34)$$

by which we mean that it is polynomial in a_1, \dots, a_n of degree $2g$. In general, outside of the moduli space $\mathcal{M}_{g,n}^{\text{ct}}$, the DR are also polynomial but they are not homogeneous. However, they have degree $2g$ and terms of lower degree (e.g. [Bur15a]). Equation (5.34) is crucial in the construction of the DR hierarchy.

The double ramification hierarchy: the Hamiltonian structure

We now define the so-called *double ramification hierarchy* (or DR hierarchy) introduced by Alexander Buryak in [Bur15a]; we follow the approach of [Ros17] though. As already mentioned, the input of the DR hierarchy is a cohomological field theory

$$c_{g,n} : V^{\otimes n} \rightarrow H^*(\overline{\mathcal{M}}_{g,n}, \mathbb{C}),$$

where V is an N -dimensional complex vector space equipped with a nondegenerate symmetric bilinear form (a metric) $\eta : V \otimes V \rightarrow \mathbb{C}$. The classes $c_{g,n}$ must satisfy axioms of Definition 5.2.1. Let integers $\underline{a} = (a_1, \dots, a_n)$ with vanishing sum and let $\underline{\alpha} = (\alpha_1, \dots, \alpha_n) \subset \{1, \dots, N\}$. We introduce the following specific intersection numbers, for any $\alpha \in \{1, \dots, N\}$ and any $d \geq 0$,

$$P_{\alpha, d; \underline{\alpha}}^{g; \underline{a}} := \int_{\text{DR}_g(-\sum a_i, \underline{a})} \lambda_g \psi_1^d c_{g, n+1}(e_\alpha, e_{\alpha_1}, \dots, e_{\alpha_n}). \quad (5.35)$$

Thanks to Equation (5.34), $P_{\alpha, d; \underline{\alpha}}^{g; \underline{a}}$ is a homogeneous polynomial in a_i 's of degree $2g$. Hence, let us write it as such:

$$P_{\alpha, d; \underline{\alpha}}^{g; \underline{a}} = \sum_{\substack{b_1, \dots, b_n \geq 0 \\ \sum b_j = 2g}} \tilde{P}_{\alpha, d; \underline{\alpha}}^{g; \underline{b}} a_1^{b_1} \cdots a_n^{b_n} \in \mathbb{C}_{2g}[a_1, \dots, a_n],$$

where, similarly to previous notations, $\underline{b} = (b_1, \dots, b_n)$. From this we define the following generating series, still for any $\alpha \in \{1, \dots, N\}$ and any $d \geq 0$ (the sum over $1 \leq \alpha_1, \dots, \alpha_n \leq N$ is understood),⁸

$$\hat{g}_{\alpha, d} := \sum_{\substack{g, n \geq 0 \\ 2g-2+n > 0}} \frac{(-1)^g \varepsilon^{2g}}{n!} \sum_{\alpha_1, \dots, \alpha_n \neq 0} P_{\alpha, d; \underline{\alpha}}^{g; \underline{a}} \prod_{j=1}^n p_{\alpha_j}^{\alpha_j} e^{ix \sum a_j}. \quad (5.36)$$

Now consider the following power series (still, the sum over $1 \leq \alpha_1, \dots, \alpha_n \leq N$ is understood):

$$g_{\alpha, d} := \sum_{\substack{g, n \geq 0 \\ 2g-2+n > 0}} \frac{\varepsilon^{2g}}{n!} \sum_{\substack{b_1, \dots, b_n \geq 0 \\ \sum b_j = 2g}} \tilde{P}_{\alpha, d; \underline{\alpha}}^{g; \underline{b}} \prod_{j=1}^n u_{b_j}^{\alpha_j}. \quad (5.37)$$

Because of the constraint $\sum b_j = 2g$ in the above sum, it is clear that $g_{\alpha, d} \in \widehat{\mathcal{A}}^{[0]}$. These differential polynomials will be the Hamiltonian densities of the DR hierarchy. Now it suffices to apply the definition $T(u_k^\alpha) = \sum_{a \neq 0} p_n^\alpha e^{inx}$ to find that

$$T(g_{\alpha, d}) = \hat{g}_{\alpha, d}.$$

⁸In [Bur15a] and other references on the topic (e.g. [BR16b, Ros17]), the series $\hat{g}_{\alpha, d}$ is actually denoted $g_{\alpha, d}$. However, we denote it $\hat{g}_{\alpha, d}$ while $g_{\alpha, d}$ will be kept for the differential polynomial we associate to $\hat{g}_{\alpha, d}$ through the map T_0^{-1} (although both representations are equivalent in virtue of the previous sections' content).

Now, as we will see in the next theorem, besides the somehow simple Hamiltonian densities, we equip the DR hierarchy with the natural standard Poisson bracket

$$\{\bar{f}, \bar{g}\}_{\eta \partial_x} = \int \frac{\delta f}{\delta u^\alpha} \eta^{\alpha\beta} \partial_x \frac{\delta g}{\delta u^\beta},$$

where $\eta^{\alpha\beta}$ is the metric carried by the CohFT. Before giving the main theorem of this section—the integrability of the DR hierarchy—we make an important remark.

Remark 5.2.12. The original construction of Buryak [Bur15a] was actually based not on the series $\hat{g}_{\alpha,d}$ but on $\hat{g}_{\alpha,d}^B := \hat{g}_{\alpha,d;(0)}$ which contains only the terms in $\hat{g}_{\alpha,d}$ for which $\sum a_j = 0$. On the other hand, the homomorphism $T_0 : \widehat{\mathcal{F}} \rightarrow \widehat{\mathcal{B}}^{\text{pol}}$ is surjective and its kernel is spanned by the \bar{u}^α 's. Let us denote $\bar{g}_{\alpha,d}^B$ the functional such that $T_0(\bar{g}_{\alpha,d}^B) = \hat{g}_{\alpha,d}^B$ and which does not depend linearly in the \bar{u}^α 's. Since we obviously have $T_0(g_{\alpha,d}) = \hat{g}_{\alpha,d}^B$, it follows that $\bar{g}_{\alpha,d} = \bar{g}_{\alpha,d}^B$ modulo a linear dependence in the \bar{u}^α 's. Yet the latter are clearly Casimir for the standard Poisson structure and hence, such a dependence does not affect the Hamiltonian structure.

Theorem 5.2.13 (DR hierarchy [Bur15a]). *Let an arbitrary CohFT along with the differential polynomials $\bar{g}_{\alpha,d}$ (5.37). For any $\alpha, \beta \in \{1, \dots, N\}$ and $p, q \geq 0$,*

$$\{\bar{g}_{\alpha,p}, \bar{g}_{\alpha,q}\}_{\eta \partial_x} = 0.$$

The Hamiltonian system generated by the above we call the double ramification integrable hierarchy associated to said CohFT, namely,

$$\frac{\partial u^\alpha}{\partial t_q^\beta} = \eta^{\alpha\beta} \partial_x \frac{\delta \bar{g}_{\beta,q}}{\delta u^\beta}.$$

We refer to [Bur15a] for the proof. As an important special case, we have

$$\bar{g}_{1,0} = \frac{1}{2} \int \eta^{\alpha\beta} u^\alpha u^\beta.$$

Consequently, the compatibility with spatial translation holds:

$$\frac{\partial u^\alpha}{\partial t_0^1} = u_1^\alpha.$$

Recursion relations

Here we give two strong recursion relations of the DR hierarchy.

Theorem 5.2.14 ([BR16b]). *The Hamiltonian densities $g_{\alpha,d}$ of the Double ramification hierarchy satisfy the following recursion:*

$$(D - 1)g_{\alpha,d+1} = \partial_x^{-1} \{g_{\alpha,d}, \bar{g}_{1,1}\}_{\eta \partial_x}, \quad \alpha \in \{1, \dots, n\}, d \geq -1. \quad (5.38)$$

where $D = \varepsilon \partial_\varepsilon + \sum_{k \geq 0} u_k^\alpha \partial_{u_k^\alpha}$.

Remark 5.2.15. Since the Hamiltonian functionals $\bar{g}_{\alpha,d}$ commute, meaning that $\{\bar{g}_{\alpha,d}, \bar{g}_{\beta,e}\} = 0$, we have $\{g_{\alpha,d}, \bar{g}_{\beta,e}\} \in \text{Im}(\partial_x)$. Therefore, Equation (5.39) is meaningful.

Remark 5.2.16. Equation (5.39) allows to compute the density $g_{\alpha,d+1}$. Indeed, by definition, $Du_k^\alpha = (k+1)u_k^\alpha$. So let us write $g_{\alpha,d+1}$ as a plain combination

$$g_{\alpha,d+1} = \sum_{m \geq 0} \sum_{i=1}^m \sum_{1 \leq \alpha_i \leq n} \sum_{k_i \geq 0} C_{k_1, \dots, k_m}^{\alpha, \dots, \alpha_m} \prod_{j=1}^m u_{k_j}^{\alpha_j}.$$

Then

$$(D-1)g_{\alpha,d+1} = \sum_{m \geq 0} \sum_{i=1}^m \sum_{1 \leq \alpha_i \leq n} \sum_{k_i \geq 0} C_{k_1, \dots, k_m}^{\alpha, \dots, \alpha_m} \left(-1 + \sum_{\ell=1}^m (k_\ell + 1) \right) \prod_{j=1}^m u_{k_j}^{\alpha_j}.$$

Now we make the following observations. The local Poisson bracket is determined by the CohFT: $\{u^\alpha(x), u^\beta(y)\} = \eta^{\alpha\beta} \partial_x$. On the other hand, the level -1 Hamiltonian densities $g_{\alpha,-1}$ are given by $g_{\alpha,-1} = \eta_{\alpha\beta} u^\beta$. Therefore, the following corollary holds.

Corollary 5.2.17. *The sole data of $\bar{g}_{1,1}$ entirely determines the DR hierarchy.*

We shall emphasize that, assuming the DR/DZ conjecture holds true, the above reconstruction technique was an unknown feature of the DZ, a fortiori DS, hierarchies. We have a second recursion theorem.

Theorem 5.2.18 ([BR16b]). *The Hamiltonian densities $g_{\alpha,d}$ of the Double ramification satisfy the following recursion:*

$$\frac{\partial g_{\alpha,p+1}}{\partial u^\beta} = \partial_x^{-1} \{g_{\alpha,p}, \bar{g}_{\beta,0}\}_{\eta \partial_x}. \tag{5.39}$$

5.2.3 Strong Double Ramification/Dubrovin–Zhang conjecture

The Dubrovin–Zhang hierarchies

In this section, we briefly describe the construction of the Dubrovin–Zhang (DZ) hierarchies for these are the subject of the DR/DZ conjectures, the weak and the strong. We will not give proofs and refer to [DZ08]. In the specific cases of a CohFT constructed as the FJRW theory of a simple singularity of type ADE, the DZ hierarchy is known to coincide with the DS hierarchy counterpart [LRZ15, GM05, FGM10]. The introduction we give is taken from [Ros17].

The Dubrovin–Zhang hierarchies are hierarchies of PDEs which are defined for any semisimple CohFT. So, to start with, a CohFT with potential F (see Equation (5.28)) is said to be semisimple if the N -dimensional complex associative algebra defined by the structure constants

$$f_{\beta\gamma}^\alpha := \eta^{\alpha\mu} \frac{\partial^3 F_0}{\partial t_0^\mu \partial t_0^\beta \partial t_0^\gamma} \Big|_{t_{p>0}^* = 0}$$

is semisimple. Now let us denote

$$\Omega_{\alpha,p;\beta,q}(t_*^*; \varepsilon) = \frac{\partial^2 F}{\partial t_p^\alpha \partial t_q^\beta} = \sum_{g \geq 0} \Omega_{\alpha,p;\beta,q}^{[2g]} \varepsilon^{2g}.$$

Then we define Hamiltonian densities and local Poisson structure:

$$h_{\alpha,p}^{\text{DZ},(0)}(v^*) := \Omega_{\alpha,p+1;1,0}^{[0]}(t_0^* = v^*, 0, 0, \dots), \quad K_{\text{DZ},(0)}^{\alpha\beta} := \eta^{\alpha\beta} \partial_x.$$

These Hamiltonians indeed commute w.r.t. to the above bracket [DZ08]. They form the so-called genus 0 Dubrovin–Zhang hierarchy, or the *principal hierarchy* of the CohFT. By construction, they automatically form a tau structure.

To construct the full genera DZ hierarchy, we make use of the so-called *topological solution*; it is the one solution v^α of the DZ hierarchy that is determined by the initial data

$$v^\alpha(x, t_*^* = 0) = \delta^{\alpha,1} x,$$

called topological. Then there exist rational functions P_p^α such that [BPS12]

$$F_g(t_0^*, t_1^*, t_2^*, \dots) = F_g(P_0^*(v_0^*, \dots, v_{3g-2}^*), \dots, P_{3g-2}^*(v_0^*, \dots, v_{3g-2}^*), 0, 0, \dots) |_{x=0}.$$

Now consider the change of coordinates

$$\varphi(v^\alpha) = w^\alpha := v^\alpha + \sum_{g \geq 1} \varepsilon^{2g} \frac{\partial^2 F_g(v_0^*, \dots, v_{3g-2}^*)}{\partial t_0^\alpha \partial x}. \quad (5.40)$$

This change of variable is invertible and normal (preserves the tau structure), although it is not a Miura transform because P_*^* are not differential polynomials. The full genera DZ hierarchy is then obtained by applying the inverse transform φ^{-1} to the genus 0 hierarchy. Namely,

$$h_{\alpha,p}^{\text{DZ}}(w_*^*) := h_{\alpha,p}^{\text{DZ},(0)} \circ \varphi^{-1}(w_*^*), \quad K_{\alpha\beta}^{\text{DZ}}(w_*^*) = ((\varphi^{-1})^* K_{\text{DZ},(0)})^{\alpha\beta}(w_*^*),$$

where $(\varphi^{-1})^* K_{\text{DZ},(0)}$ is given in Equation (5.18). By construction, the potential of the CohFT is a tau function of the DZ hierarchy and is associated to the topological solution.

Our presentation of the DZ hierarchies ends here. We retain, as was already stated, that for the ADE CohFTs, the DZ hierarchy is Miura equivalent to the corresponding DS hierarchy [LRZ15].

Tau structure of the DR hierarchy and the strong DR/DZ equivalence conjecture

In [BDGR18], the authors defined a very simple tau structure for the DR hierarchies which is actually only due to the simple standard Poisson structure it is equipped with. First, we add to the Hamiltonian densities $g_{\alpha,d}$ Casimir functionals $g_{\alpha,-1} := \eta_{\alpha\beta} u^\beta$, where $\eta_{\alpha\beta}$ denotes the inverse matrix.

Theorem 5.2.19. *The following Hamiltonian densities are tau symmetric:*

$$h_{\alpha,p}^{\text{DR}} := \frac{\delta g_{\alpha,p+1}}{\delta u^1}, \quad \alpha \in \{1, \dots, N\}, p \geq -1. \quad (5.41)$$

A proof is given in [BDGR18], see Propositions 3.1 and 4.1. Already in [Bur15a], A. Buryak, guided by the examples of the trivial and Hodge hierarchies he computed, made the so-called *weak DR/DZ equivalence conjecture*.

Conjecture 5.2.20 (Weak DR/DZ equivalence). *Given a semisimple CohFT, the associated DR and DZ hierarchies are equivalent through a Miura transform.*

Now take the normal coordinates $\tilde{u}^\alpha = \eta^{\alpha\beta} h_{\beta,-1}^{\text{DR}}$ and consider the special solution given by the initial data $\tilde{u}^\alpha(x, t_*^* = 0; \varepsilon) = \delta^{\alpha,1} x$. It is the same condition as for the topological solution of the DZ hierarchy, however, for the DR hierarchy it does not carry the same geometrical interpretation (since it is not constructed as the potential of the CohFT). This special solution gives a tau function τ which we write as

$$F^{\text{DR}}(t_*^*; \varepsilon) := \log \tau = \sum_{g \geq 0} \varepsilon^{2g} F_g^{\text{DR}}(t_*^*).$$

We also denote its expansion as if it were the generating series of some intersection numbers:

$$F_g^{\text{DR}} =: \sum_{\substack{n \geq 0 \\ 2g-2+n > 0}} \frac{1}{n!} \sum_{d_1, \dots, d_n \geq 0} \langle \tau_{d_1}(e_{\alpha_1}) \cdots \tau_{d_n}(e_{\alpha_n}) \rangle_g^{\text{DR}} \prod_{i=1}^n t_{d_i}^{\alpha_i}.$$

We still call the above the genus g *partition function* of the DR hierarchy and the complex numbers $\langle \tau_{d_1}(e_{\alpha_1}) \cdots \tau_{d_n}(e_{\alpha_n}) \rangle_g^{\text{DR}}$ we also call *correlators*. Even though it is not clear whether the DR partition function carries a geometric interpretation, the following surprising selection rule holds.

Proposition 5.2.21 ([BDGR18]). *The DR correlators $\langle \tau_{d_1}(e_{\alpha_1}) \cdots \tau_{d_n}(e_{\alpha_n}) \rangle_g^{\text{DR}}$ are nonzero only if the following inequalities are satisfied*

$$2g - 3 \leq \sum_{i=1}^n d_i \leq 3g - 3 + n \quad (5.42)$$

The above upper bound shall be compared with the well known selection rule on of the DZ hierarchy which states that the correlators $\langle \tau_{d_1}(e_{\alpha_1}) \cdots \tau_{d_n}(e_{\alpha_n}) \rangle_g$ (the actual intersection numbers of the CohFT) vanish whenever $\sum d_i > 3g - 3 + n$. The latter is only due to the limit on the number of psi classes one integrates.

But the lower bound has no equivalent in the DZ hierarchies! And as it turns out, the lower bound plays a key role in refining the DR/DZ conjecture into its strong form which we state below. First, we have the following theorem.

Theorem 5.2.22 ([BDGR18]). *Let $c_{g,n}$ be a semisimple CohFT and let w^α be the coordinates of the DZ hierarchy in its full form (5.40) and let F be the generating function of the CohFT (also a tau function of the DZ hierarchy). Then there exists a unique differential polynomial $f \in \widehat{\mathcal{A}}^{[-2]}$ such that the normal Miura transform*

$$\tilde{u}^\alpha = w^\alpha + \eta^{\alpha\beta} \partial_x \{f, \bar{h}_{\beta,0}^{\text{DZ}}\}_{\text{DZ}},$$

gives rise to a potential $F^{\text{red}} := F + f|_{x=0}$ such that F^{red} satisfies the selection rules (5.42).

By requiring that the Miura transform is normal, the above entirely determines the Miura transform that is conjectured to hold between the DR and the DZ hierarchy. Hence the strong DR/DZ equivalence conjecture:

Conjecture 5.2.23 (Strong DR/DZ equivalence [BDGR18]). *For any semisimple CohFT, the DR and DZ hierarchies coincide up to the normal Miura transform generated by $f \in \widehat{\mathcal{A}}^{[-2]}$ of Theorem 5.2.22.*

In particular, for the CohFT arising from the FJRW of a simple singularity ADE, the Miura transform is conjectured to be the identity.

We will not prove here that the expected Miura transform reduces to the identity for the ADE cases. The strong DR/DZ equivalence has been confirmed in a number of cases [Bur15a, BG16, BDGR18]:

- ◊ The trivial CohFT;
- ◊ The full Hodge class CohFT;
- ◊ Witten's r -spin classes for $r = 3, 4, 5$ (equivalent to the A_3, A_4, A_5 FJRW theories);
- ◊ The Gromov–Witten theory of \mathbb{P}^1 ;
- ◊ Up to genus 5 for any CohFT of rank 1;
- ◊ Up to genus 2 for any semisimple CohFT.

In the next section, we describe the DR hierarchy for the FJRW of type D_4 . In Section 5.3 we describe the DS hierarchy of type D_4 (which is known to coincide with the DS hierarchy [LRZ15, GM05, FGM10]). We will observe that in normal coordinates, the DS and DR hierarchies coincide, confirming the strong DR/DZ equivalence in the D_4 case.

5.2.4 Hamiltonian structure of the double ramification hierarchy of type D_4

In [BG16] (Theorem 5.3), A. Buryak and J. Guéré proved the following theorem which is of great help in proving the DR/DZ conjecture. First, let us fix some notations. Let a semisimple CohFT and consider the associated Dubrovin–Zhang hierarchy. Let $h_{\alpha,d}^{\text{DZ}} \in \widehat{\mathcal{A}}(\tilde{u}_*^*; \varepsilon)$ be the Hamiltonian densities of the DZ hierarchy and let $K_{\text{DZ}}^{\alpha\beta}$ be the local Poisson structure of the DZ hierarchy.

Theorem 5.2.24 ([BG16]). *Suppose there exists a Miura transform $\varphi : u^\alpha \mapsto \tilde{u}^\alpha(u_*^*; \varepsilon)$, such that*

$$\frac{\partial \tilde{u}^\alpha}{\partial u^1} = \delta^{\alpha,1}, \quad \varphi^*(\eta^{\alpha\beta} \partial_x) = K_{\text{DZ}}^{\alpha\beta}, \quad \overline{g_{1,1} \circ \varphi} = \bar{h}_{1,1}^{\text{DZ}}, \quad (5.43)$$

where $\varphi^*(\eta^{\alpha\beta}\partial_x)$ is given in (5.18). Then the DR hierarchy, under the Miura transform φ , coincides with the DZ hierarchy.

It what follows we give the standard Poisson structure $\eta^{\alpha\beta}\partial_x$ and the first Hamiltonian densities $g_{\alpha,0}$ and $g_{1,1}$ of the DR hierarchy constructed from the CohFT associated to the FJRW theory of the singularity

$$W_{D_4} = x^2y + y^3$$

with symmetry group $G = \langle J \rangle = \langle (e^{2i\pi/3}, e^{2i\pi/3}) \rangle$. Then we give the Poisson structure and the Hamiltonian densities after the Miura transform \tilde{u}^α . The details of the computations are given in the next section.

Hamiltonian densities and Poisson structure in the coordinates u^α of the DR hierarchy of type D_4

The local Poisson structure of the DR hierarchy is the standard one and reads

$$\left(\eta^{\alpha\beta}\right)\partial_x = \begin{pmatrix} 0 & 0 & 6 & 0 \\ 0 & 6 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix} \partial_x.$$

Using Equation (5.37), we computed the Hamiltonian densities $g_{\beta,0}$, for $\beta \in \{1, \dots, 4\}$. The latter, through the tau structure, will allow us to compute the normal coordinates \tilde{u}^α . We obtain the following densities.

$$\begin{aligned} g_{1,0} &= \left(\frac{1}{12}(u^2)^2 + \frac{1}{6}u^1u^3 + \frac{1}{4}(u^4)^2\right) + \left(\frac{1}{216}u^3(u^3)^2 + \frac{1}{6}u_2^1 + \frac{1}{432}(u^3)^2u_2^3\right)\epsilon^2 + \left(\frac{7}{1080}(u_2^3)^2\right. \\ &\quad \left. + \frac{11}{1080}u_1^3u_3^3 + \frac{1}{270}u^3u_4^3\right)\epsilon^4 + \frac{13}{9072}u_6^3\epsilon^6 \\ g_{2,0} &= \left(\frac{1}{6}u^1u^2 - \frac{1}{72}(u^2)^2u^3 + \frac{1}{648}u^2(u^3)^3 + \frac{1}{24}u^3(u^4)^2\right) + \left(-\frac{1}{72}(u_1^2)^2 + \frac{1}{72}u^3u_1^2u_1^3 + \frac{1}{216}u^2(u_1^3)^2\right. \\ &\quad \left. + \frac{1}{24}(u_1^4)^2 - \frac{1}{36}u^2u_2^2 + \frac{1}{144}(u^3)^2u_2^2 + \frac{1}{108}u^2u^3u_2^3 + \frac{1}{12}u^4u_2^4\right)\epsilon^2 + \left(\frac{1}{45}u_2^2u_2^3 + \frac{1}{60}u_1^3u_3^2 + \frac{1}{72}u_1^2u_3^3\right. \\ &\quad \left. + \frac{1}{120}u^3u_4^2 + \frac{1}{216}u^2u_4^3\right)\epsilon^4 + \frac{1}{336}u_6^2\epsilon^6 \end{aligned}$$

$$\begin{aligned}
g_{3,0} = & \left(\frac{1}{12} (u^1)^2 - \frac{1}{216} (u^2)^3 + \frac{1}{432} (u^2)^2 (u^3)^2 + \frac{1}{233280} (u^3)^6 + \frac{1}{24} u^2 (u^4)^2 + \frac{1}{144} (u^3)^2 (u^4)^2 \right) \\
& + \left(\frac{1}{432} u^3 (u_1^2)^2 + \frac{1}{108} u^2 u_1^2 u_1^3 + \frac{1}{3888} (u^3)^3 (u_1^3)^2 + \frac{1}{36} u^4 u_1^3 u_1^4 + \frac{1}{144} u^3 (u_1^4)^2 + \frac{1}{432} (u^3)^2 u_1^4 \right. \\
& + \frac{1}{108} u^2 u^3 u_2^2 + \frac{1}{216} (u^2)^2 u_2^3 + \frac{1}{7776} (u^3)^4 u_2^3 + \frac{1}{72} (u^4)^2 u_2^3 + \frac{1}{36} u^3 u^4 u_2^4 \Big) \epsilon^2 + \left(\frac{5}{15552} (u_1^3)^4 \right. \\
& + \frac{1}{240} (u_2^2)^2 + \frac{11}{3888} u^3 (u_1^3)^2 u_2^3 + \frac{1}{216} u_2^1 u_2^3 + \frac{1}{864} (u^3)^2 (u_2^2)^2 + \frac{1}{80} (u_2^4)^2 + \frac{1}{216} u_1^3 u_3^1 + \frac{1}{216} u_1^2 u_2^2 \\
& + \frac{1}{648} (u^3)^2 u_1^3 u_3^3 + \frac{1}{72} u_1^4 u_3^4 + \frac{1}{270} u^3 u_4^1 + \frac{1}{216} u^2 u_4^2 + \frac{1}{3888} (u^3)^3 u_4^3 + \frac{1}{72} u^4 u_4^4 \Big) \epsilon^4 + \left(\frac{19}{11664} (u_2^3)^3 \right. \\
& + \frac{1}{144} u_1^3 u_2^3 u_3^3 + \frac{5}{2592} u^3 (u_3^3)^2 + \frac{1}{486} (u_1^3)^2 u_4^3 + \frac{25}{7776} u^3 u_2^3 u_4^3 + \frac{1}{648} u^3 u_1^3 u_5^3 + \frac{13}{9072} u_6^1 \\
& + \frac{1}{3888} (u^3)^2 u_6^3 \Big) \epsilon^6 + \left(\frac{23}{15552} (u_4^3)^2 + \frac{19}{7776} u_3^3 u_5^3 + \frac{11}{7776} u_2^3 u_6^3 + \frac{1}{1944} u_1^3 u_7^3 + \frac{1}{7776} u^3 u_8^3 \right) \epsilon^8 \\
& + \frac{1}{38880} u_{10}^3 \epsilon^{10} \\
g_{4,0} = & \left(\frac{1}{2} u^1 u^4 + \frac{1}{12} u^2 u^3 u^4 + \frac{1}{216} (u^3)^3 u^4 \right) + \left(\frac{1}{72} u^4 (u_1^3)^2 + \frac{1}{12} u_1^2 u_1^4 + \frac{1}{24} u^3 u_1^3 u_1^4 + \frac{1}{12} u^4 u_2^2 + \frac{1}{36} u^3 u^4 u_2^3 \right. \\
& \left. + \frac{1}{12} u^2 u_2^4 + \frac{1}{48} (u^3)^2 u_2^4 \right) \epsilon^2 + \left(\frac{1}{15} u_2^3 u_2^4 + \frac{1}{24} u_1^4 u_3^3 + \frac{1}{20} u_1^3 u_3^4 + \frac{1}{72} u^4 u_3^3 + \frac{1}{40} u^3 u_4^4 \right) \epsilon^4 + \frac{1}{112} u_6^4 \epsilon^6.
\end{aligned}$$

Thanks to Theorem 5.2.24, it suffices to compute the Hamiltonian functional $\bar{g}_{1,1}$ to check if the DZ and the DR hierarchy coincide (if the condition of the theorem are satisfied, but we will see that is the case in the next section). We find

$$\begin{aligned}
\bar{g}_{1,1} = & \int \left(\frac{1}{12} u^1 (u^2)^2 + \frac{1}{12} (u^1)^2 u^3 - \frac{1}{108} (u^2)^3 u^3 + \frac{1}{432} (u^2)^2 (u^3)^3 + \frac{1}{326592} (u^3)^7 + \frac{1}{4} u^1 (u^4)^2 \right. \\
& + \frac{1}{12} u^2 u^3 (u^4)^2 + \frac{1}{144} (u^3)^3 (u^4)^2 \Big) + \left(-\frac{1}{6} (u_1^1)^2 + \frac{1}{24} u^2 (u_1^2)^2 - \frac{1}{72} (u^3)^2 (u_1^2)^2 \right. \\
& - \frac{1}{108} (u^3)^2 u_1^1 u_1^3 - \frac{1}{27} u^2 u^3 u_1^2 u_1^3 - \frac{1}{108} (u^2)^2 (u_1^3)^2 - \frac{1}{2592} (u^3)^4 (u_1^3)^2 - \frac{1}{36} (u^4)^2 (u_1^3)^2 \\
& - \frac{1}{4} u^4 u_1^2 u_1^4 - \frac{1}{9} u^3 u^4 u_1^3 u_1^4 - \frac{1}{8} u^2 (u_1^4)^2 - \frac{1}{24} (u^3)^2 (u_1^4)^2 \Big) \epsilon^2 + \left(-\frac{35}{46656} u^3 (u_1^3)^4 \right. \\
& + \frac{1}{144} (u_1^3)^2 u_2^1 + \frac{5}{216} u_1^2 u_1^3 u_2^2 + \frac{1}{48} u^3 (u_2^2)^2 + \frac{1}{54} u^3 u_2^1 u_2^3 + \frac{5}{216} u^2 u_2^2 u_2^3 + \frac{7}{7776} (u^3)^3 (u_2^3)^2 \\
& + \frac{5}{72} u_1^3 u_1^4 u_2^4 + \frac{5}{72} u^4 u_2^3 u_2^4 + \frac{1}{16} u^3 (u_2^4)^2 \Big) \epsilon^4 + \left(\frac{1}{486} (u_1^3)^2 (u_2^3)^2 + \frac{13}{5832} u^3 (u_2^3)^3 \right. \\
& - \frac{1}{112} (u_2^3)^2 - \frac{13 u_3^1 u_3^3}{1512} - \frac{1}{972} (u^3)^2 (u_3^3)^2 - \frac{3}{112} (u_3^4)^2 \Big) \epsilon^6 + \left(-\frac{5}{1728} u_2^3 (u_3^3)^2 + \frac{1}{1728} u^3 (u_4^3)^2 \right) \epsilon^8 \\
& - \frac{1}{7776} (u_5^3)^2 \epsilon^{10}. \tag{5.44}
\end{aligned}$$

Hamiltonian densities and Poisson structure in normal coordinates \tilde{u}^α of the DR hierarchy of type D_4

Now, since the tau structure is given by $h_{\alpha,p}^{\text{DR}} = \frac{\delta g_{\alpha,p+1}}{\delta u^1}$ (5.41), the normal coordinates associated to said tau structure are given by

$$\tilde{u}^\alpha = \eta^{\alpha\beta} \frac{\delta g_{\beta,0}}{\delta u^1}.$$

We find the following Miura transform.

$$\begin{cases} \tilde{u}^1 = u^1 + (-\frac{1}{36}(u_1^3)^2 - \frac{1}{36}u^3u_2^3)\epsilon^2 - \frac{1}{45}u_4^3\epsilon^4 \\ \tilde{u}^2 = u^2 \\ \tilde{u}^3 = u^3 \\ \tilde{u}^4 = u^4 \end{cases} \quad (5.45)$$

In particular, the above satisfies hypothesis $\partial\tilde{u}^\alpha/\partial u^1 = \delta^{\alpha,1}$ of Theorem 5.2.24. The latter is necessary, yet of course not sufficient for the theorem to apply. Using to Equation (5.18), we get the following proposition.

Proposition 5.2.25. *The local Poisson structure of the DR hierarchy of type D_4 in normal coordinates reads*

$$\left(K_{\text{DR}}^{\alpha\beta}\right) = \begin{pmatrix} \left(\frac{1}{6}\tilde{u}_2^3\partial_x + \frac{1}{2}\tilde{u}_1^3\partial_x^2 + \frac{1}{3}\tilde{u}^3\partial_x^3\right)\epsilon^2 + \frac{4}{15}\partial_x^5\epsilon^4 & 0 & 6\partial_x & 0 \\ 0 & 6\partial_x & 0 & 0 \\ 6\partial_x & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\partial \end{pmatrix}. \quad (5.46)$$

Finally, using $h_{\alpha,p}^{\text{DR}} = \frac{\delta g_{\alpha,p+1}}{\delta u^1}$ and the Miura transform (5.45), we can compute the Hamiltonian densities $h_{\alpha,0}$ in normal coordinates. Notice that the Hamiltonian densities $h_{\alpha,0}^{\text{DR}}$ are not necessary to apply Theorem 5.2.24 (in the view of confirming the DR/DZ conjecture); we give them here for the sake of describing the DR hierarchy of type D_4 .

Proposition 5.2.26. *The level 0 tau symmetric Hamiltonian densities $h_{1,0}^{\text{DR}}, \dots, h_{4,0}^{\text{DR}}$ of the DR hierarchy of type D_4 in normal coordinates are given by the following.*

$$\begin{aligned} h_{1,0}^{\text{DR}} &= \left(\frac{1}{12}(\tilde{u}^2)^2 + \frac{1}{6}\tilde{u}^1\tilde{u}^3 + \frac{1}{4}(\tilde{u}^4)^2\right) + \left(\frac{1}{72}\tilde{u}^3(\tilde{u}_1^3)^2 + \frac{1}{3}\tilde{u}_2^1 + \frac{1}{216}(\tilde{u}^3)^2\tilde{u}_2^3\right)\epsilon^2 + \left(\frac{1}{216}(\tilde{u}_2^3)^2 + \frac{1}{72}\tilde{u}_1^3\tilde{u}_3^3\right. \\ &\quad \left. + \frac{1}{180}\tilde{u}^3\tilde{u}_4^3\right)\epsilon^4 + \frac{1}{840}\tilde{u}_6^3\epsilon^6 \\ h_{2,0}^{\text{DR}} &= \left(\frac{1}{6}\tilde{u}^1\tilde{u}^2 - \frac{1}{72}(\tilde{u}^2)^2\tilde{u}^3 + \frac{1}{648}\tilde{u}^2(\tilde{u}^3)^3 + \frac{1}{24}\tilde{u}^3(\tilde{u}^4)^2\right) + \left(-\frac{1}{24}(\tilde{u}_1^2)^2 + \frac{1}{24}\tilde{u}^3\tilde{u}_1^2\tilde{u}_1^3 + \frac{1}{72}\tilde{u}^2(\tilde{u}_1^3)^2\right. \\ &\quad \left. + \frac{1}{8}(\tilde{u}_1^4)^2 - \frac{1}{18}\tilde{u}^2\tilde{u}_2^2 + \frac{1}{72}(\tilde{u}^3)^2\tilde{u}_2^2 + \frac{1}{54}\tilde{u}^2\tilde{u}^3\tilde{u}_2^3 + \frac{1}{6}\tilde{u}^4\tilde{u}_2^4\right)\epsilon^2 + \left(\frac{1}{12}\tilde{u}_2^2\tilde{u}_2^3 + \frac{5}{72}\tilde{u}_1^3\tilde{u}_3^2 + \frac{1}{18}\tilde{u}_1^2\tilde{u}_3^3\right. \\ &\quad \left. + \frac{1}{36}\tilde{u}^3\tilde{u}_4^2 + \frac{2}{135}\tilde{u}^2\tilde{u}_4^3\right)\epsilon^4 + \frac{1}{72}\tilde{u}_6^2\epsilon^6 \end{aligned}$$

$$\begin{aligned}
h_{3,0}^{\text{DR}} &= \left(\frac{1}{12} (\tilde{u}^1)^2 - \frac{1}{216} (\tilde{u}^2)^3 + \frac{1}{432} (\tilde{u}^2)^2 (\tilde{u}^3)^2 + \frac{1}{233280} (\tilde{u}^3)^6 + \frac{1}{24} \tilde{u}^2 (\tilde{u}^4)^2 + \frac{1}{144} (\tilde{u}^3)^2 (\tilde{u}^4)^2 \right) + \\
&\quad \left(\frac{5}{432} \tilde{u}^3 (\tilde{u}_1^2)^2 + \frac{1}{108} \tilde{u}^3 \tilde{u}_1^1 \tilde{u}_1^3 + \frac{1}{36} \tilde{u}^2 \tilde{u}_1^2 \tilde{u}_1^3 + \frac{1}{1296} (\tilde{u}^3)^3 (\tilde{u}_1^3)^2 + \frac{1}{12} \tilde{u}^4 \tilde{u}_1^3 \tilde{u}_1^4 + \frac{5}{144} \tilde{u}^3 (\tilde{u}_1^4)^2 \right. \\
&\quad \left. + \frac{1}{216} (\tilde{u}^3)^2 \tilde{u}_2^1 + \frac{1}{54} \tilde{u}^2 \tilde{u}^3 \tilde{u}_2^2 + \frac{1}{108} (\tilde{u}^2)^2 \tilde{u}_2^3 + \frac{1}{3888} (\tilde{u}^3)^4 \tilde{u}_2^3 + \frac{1}{36} (\tilde{u}^4)^2 \tilde{u}_2^3 + \frac{1}{18} \tilde{u}^3 \tilde{u}^4 \tilde{u}_2^4 \right) \epsilon^2 \\
&\quad + \left(\frac{5}{2592} (\tilde{u}_1^3)^4 + \frac{13}{720} (\tilde{u}_2^2)^2 + \frac{55}{3888} \tilde{u}^3 (\tilde{u}_1^3)^2 \tilde{u}_2^3 + \frac{29}{1080} \tilde{u}_2^1 \tilde{u}_2^3 + \frac{67}{15552} (\tilde{u}^3)^2 (\tilde{u}_2^3)^2 + \frac{13}{240} (\tilde{u}_2^4)^2 \right. \\
&\quad \left. + \frac{1}{30} \tilde{u}_1^3 \tilde{u}_3^1 + \frac{31}{1080} \tilde{u}_1^2 \tilde{u}_3^2 + \frac{11}{1080} \tilde{u}_1^1 \tilde{u}_3^3 + \frac{49}{7776} (\tilde{u}^3)^2 \tilde{u}_1^3 \tilde{u}_3^3 + \frac{31}{360} \tilde{u}_1^4 \tilde{u}_3^4 + \frac{1}{60} \tilde{u}^3 \tilde{u}_1^4 + \frac{2}{135} \tilde{u}^2 \tilde{u}_4^2 \right. \\
&\quad \left. + \frac{1}{1296} (\tilde{u}^3)^3 \tilde{u}_4^3 + \frac{2}{45} \tilde{u}^4 \tilde{u}_4^4 \right) \epsilon^4 + \left(\frac{1129}{116640} (\tilde{u}_2^3)^3 + \frac{1601}{38880} \tilde{u}_1^3 \tilde{u}_2^3 \tilde{u}_3^3 + \frac{1}{120} \tilde{u}^3 (\tilde{u}_3^3)^2 + \frac{19}{1620} (\tilde{u}_1^3)^2 \tilde{u}_4^3 \right. \\
&\quad \left. + \frac{29}{2160} \tilde{u}^3 \tilde{u}_2^3 \tilde{u}_4^3 + \frac{13}{1944} \tilde{u}^3 \tilde{u}_1^3 \tilde{u}_5^3 + \frac{11}{1080} \tilde{u}_6^1 + \frac{17}{19440} (\tilde{u}^3)^2 \tilde{u}_6^3 \right) \epsilon^6 + \left(\frac{191}{43200} (\tilde{u}_4^3)^2 + \frac{1501}{194400} \tilde{u}_3^3 \tilde{u}_5^3 \right. \\
&\quad \left. + \frac{949}{194400} \tilde{u}_2^3 \tilde{u}_6^3 + \frac{127}{64800} \tilde{u}_1^3 \tilde{u}_7^3 + \frac{13}{32400} \tilde{u}^3 \tilde{u}_8^3 \right) \epsilon^8 + \frac{7}{118800} \tilde{u}_{10}^3 \epsilon^{10} \\
h_{4,0}^{\text{DR}} &= \left(\frac{1}{2} \tilde{u}^1 \tilde{u}^4 + \frac{1}{12} \tilde{u}^2 \tilde{u}^3 \tilde{u}^4 + \frac{1}{216} (\tilde{u}^3)^3 \tilde{u}^4 \right) + \left(\frac{1}{24} \tilde{u}^4 (\tilde{u}_1^3)^2 + \frac{1}{4} \tilde{u}_1^2 \tilde{u}_1^4 + \frac{1}{8} \tilde{u}^3 \tilde{u}_1^3 \tilde{u}_1^4 + \frac{1}{6} \tilde{u}^4 \tilde{u}_2^2 \right. \\
&\quad \left. + \frac{1}{18} \tilde{u}^3 \tilde{u}^4 \tilde{u}_2^3 + \frac{1}{6} \tilde{u}^2 \tilde{u}_2^4 + \frac{1}{24} (\tilde{u}^3)^2 \tilde{u}_2^4 \right) \epsilon^2 + \left(\frac{1}{4} \tilde{u}_2^3 \tilde{u}_2^4 + \frac{1}{6} \tilde{u}_1^4 \tilde{u}_3^3 + \frac{5}{24} \tilde{u}_1^3 \tilde{u}_3^4 + \frac{2}{45} \tilde{u}^4 \tilde{u}_4^3 + \frac{1}{12} \tilde{u}^3 \tilde{u}_4^4 \right) \epsilon^4 \\
&\quad + \frac{1}{24} \tilde{u}_6^4 \epsilon^6
\end{aligned}$$

Proposition 5.2.27. *The tau symmetric Hamiltonian functional $\bar{h}_{1,1}^{\text{DR}}$ of the DR hierarchy of type D_4 in normal coordinates is given by the following.*

$$\begin{aligned}
\bar{h}_{1,1}^{\text{DR}} &= \frac{1}{326592} (\tilde{u}^3)^7 + \frac{1}{432} (\tilde{u}^2)^2 (\tilde{u}^3)^3 + \frac{1}{144} (\tilde{u}^4)^2 (\tilde{u}^3)^3 - \frac{1}{108} (\tilde{u}^2)^3 \tilde{u}^3 + \frac{1}{12} (\tilde{u}^1)^2 \tilde{u}^3 + \frac{1}{12} \tilde{u}^2 (\tilde{u}^4)^2 \tilde{u}^3 \\
&\quad + \frac{1}{12} \tilde{u}^1 (\tilde{u}^2)^2 + \frac{1}{4} \tilde{u}^1 (\tilde{u}^4)^2 + \left(-\frac{1}{2592} (\tilde{u}^3)^4 (\tilde{u}_1^3)^2 - \frac{1}{72} (\tilde{u}^3)^2 (\tilde{u}_1^2)^2 - \frac{1}{24} (\tilde{u}^3)^2 (\tilde{u}_1^4)^2 - \frac{1}{108} (\tilde{u}^3)^2 \tilde{u}_1^1 \tilde{u}_1^3 \right. \\
&\quad \left. + -\frac{1}{216} \tilde{u}^1 \tilde{u}^3 (\tilde{u}_1^3)^2 - \frac{1}{27} \tilde{u}^2 \tilde{u}^3 \tilde{u}_1^2 \tilde{u}_1^3 - \frac{1}{9} \tilde{u}^4 \tilde{u}^3 \tilde{u}_1^3 \tilde{u}_1^4 - \frac{1}{6} (\tilde{u}_1^1)^2 + \frac{1}{24} \tilde{u}^2 (\tilde{u}_1^2)^2 - \frac{5}{432} (\tilde{u}^2)^2 (\tilde{u}_1^3)^2 \right. \\
&\quad \left. - \frac{5}{144} (\tilde{u}^4)^2 (\tilde{u}_1^3)^2 - \frac{1}{8} \tilde{u}^2 (\tilde{u}_1^4)^2 - \frac{1}{4} \tilde{u}^4 \tilde{u}_1^2 \tilde{u}_1^4 - \frac{1}{216} \tilde{u}^1 (\tilde{u}^3)^2 \tilde{u}_2^3 - \frac{1}{432} (\tilde{u}^2)^2 \tilde{u}^3 \tilde{u}_2^3 - \frac{1}{144} (\tilde{u}^4)^2 \tilde{u}^3 \tilde{u}_2^3 \right) \epsilon^2 \\
&\quad + \left(\frac{1}{144} (\tilde{u}_1^3)^2 \tilde{u}_2^1 + \frac{7}{7776} (\tilde{u}^3)^2 (\tilde{u}_1^3)^2 \tilde{u}_2^3 + \frac{5}{216} \tilde{u}_1^2 \tilde{u}_1^3 \tilde{u}_2^2 + \frac{1}{36} \tilde{u}_1^1 \tilde{u}_1^3 \tilde{u}_2^3 + \frac{5}{72} \tilde{u}_1^4 \tilde{u}_1^3 \tilde{u}_2^4 - \frac{1}{1458} \tilde{u}^3 (\tilde{u}_1^3)^4 \right. \\
&\quad \left. + \frac{1}{3888} (\tilde{u}^3)^3 \tilde{u}_3^3 \tilde{u}_1^3 + \frac{1}{108} \tilde{u}^3 \tilde{u}_3^3 \tilde{u}_1^1 + \frac{1}{48} \tilde{u}^3 (\tilde{u}_2^2)^2 + \frac{5}{5184} (\tilde{u}^3)^3 (\tilde{u}_2^3)^2 + \frac{1}{16} \tilde{u}^3 (\tilde{u}_2^4)^2 + \frac{1}{54} \tilde{u}^3 \tilde{u}_2^1 \tilde{u}_2^3 \right. \\
&\quad \left. + \frac{5}{216} \tilde{u}^2 \tilde{u}_2^2 \tilde{u}_2^3 + \frac{5}{72} \tilde{u}^4 \tilde{u}_2^3 \tilde{u}_2^4 - \frac{1}{540} \tilde{u}_4^3 (\tilde{u}^2)^2 - \frac{1}{180} \tilde{u}_4^3 (\tilde{u}^4)^2 - \frac{1}{270} \tilde{u}_4^3 \tilde{u}^1 \tilde{u}^3 \right) \epsilon^4 + \left(\frac{1}{144} (\tilde{u}_1^3)^2 \tilde{u}_2^1 \right. \\
&\quad \left. + \frac{7}{7776} (\tilde{u}^3)^2 (\tilde{u}_1^3)^2 \tilde{u}_2^3 + \frac{5}{216} \tilde{u}_1^2 \tilde{u}_1^3 \tilde{u}_2^2 + \frac{1}{36} \tilde{u}_1^1 \tilde{u}_1^3 \tilde{u}_2^3 + \frac{5}{72} \tilde{u}_1^4 \tilde{u}_1^3 \tilde{u}_2^4 - \frac{1}{1458} \tilde{u}^3 (\tilde{u}_1^3)^4 + \frac{1}{3888} (\tilde{u}^3)^3 \tilde{u}_3^3 \tilde{u}_1^3 \right. \\
&\quad \left. + \frac{1}{108} \tilde{u}^3 \tilde{u}_3^3 \tilde{u}_1^1 + \frac{1}{48} \tilde{u}^3 (\tilde{u}_2^2)^2 + \frac{5}{5184} (\tilde{u}^3)^3 (\tilde{u}_2^3)^2 + \frac{1}{16} \tilde{u}^3 (\tilde{u}_2^4)^2 + \frac{1}{54} \tilde{u}^3 \tilde{u}_2^1 \tilde{u}_2^3 + \frac{5}{216} \tilde{u}^2 \tilde{u}_2^2 \tilde{u}_2^3 + \frac{5}{72} \tilde{u}^4 \tilde{u}_2^3 \tilde{u}_2^4 \right. \\
&\quad \left. - \frac{1}{540} \tilde{u}_4^3 (\tilde{u}^2)^2 - \frac{1}{180} \tilde{u}_4^3 (\tilde{u}^4)^2 - \frac{1}{270} \tilde{u}_4^3 \tilde{u}^1 \tilde{u}^3 \right) \epsilon^6 + \left(\frac{65}{54432} \tilde{u}_3^3 \tilde{u}_4^3 \tilde{u}_1^3 - \frac{55}{108864} (\tilde{u}_3^3)^2 \tilde{u}_2^3 + \frac{1}{24300} (\tilde{u}_4^3)^2 \tilde{u}^3 \right. \\
&\quad \left. + \frac{1}{1728} \tilde{u}^3 (\tilde{u}_4^3)^2 + \frac{13}{54432} \tilde{u}^3 \tilde{u}_3^3 \tilde{u}_5^3 \right) \epsilon^8 - \frac{1}{7776} (\tilde{u}_5^3)^2 \epsilon^{10}
\end{aligned}$$

Level 0 flows of the DR hierarchy of type D_4 in normal coordinates

From the data of the level 0 Hamiltonian densities $h_{0,\beta}^{\text{DR}}$, we can compute the level 0 equations of motion in normal coordinates. The computations below agrees with those of [LRZ15] (see

§5.1 on the hierarchy of type D_4). First, as we already know,

$$\left\{ \tilde{u}_{t_0}^\alpha = \tilde{u}_1^\alpha \mid 1 \leq \alpha \leq 4 \right\}.$$

Then we have the following.

$$\left\{ \begin{array}{l} \tilde{u}_{t_0}^1 = \frac{1}{2}\tilde{u}^4\tilde{u}_1^4 + \frac{1}{12}\tilde{u}^3\tilde{u}_3^2\epsilon^2 + \frac{1}{36}(\tilde{u}^3)^2\tilde{u}_1^2 + \frac{1}{18}\tilde{u}^2\tilde{u}_3^3\epsilon^2 - \frac{1}{6}\tilde{u}^2\tilde{u}_1^2 + \frac{1}{18}\tilde{u}^2\tilde{u}_3\tilde{u}_1^3 + \frac{1}{20}\tilde{u}_5^2\epsilon^4 + \frac{1}{6}\tilde{u}_1^3\tilde{u}_2^2\epsilon^2 + \frac{5}{36}\tilde{u}_1^2\tilde{u}_2^3\epsilon^2 \\ \tilde{u}_{t_0}^2 = \frac{1}{36}\tilde{u}^3\tilde{u}_3^3\epsilon^2 - \frac{1}{6}\tilde{u}^3\tilde{u}_1^2 + \frac{1}{36}(\tilde{u}^3)^2\tilde{u}_1^3 - \frac{1}{6}\tilde{u}^2\tilde{u}_1^3 + \frac{1}{180}\tilde{u}_5^3\epsilon^4 + \frac{1}{36}\tilde{u}_1^3\tilde{u}_2^3\epsilon^2 - \frac{1}{6}\tilde{u}_3^2\epsilon^2 + \tilde{u}_1^1 \\ \tilde{u}_{t_0}^3 = \tilde{u}_1^2 \\ \tilde{u}_{t_0}^4 = \frac{1}{6}\tilde{u}^4\tilde{u}_1^3 + \frac{1}{6}\tilde{u}^3\tilde{u}_1^4 + \frac{1}{6}\tilde{u}_3^4\epsilon^2 \end{array} \right.$$

$$\left\{ \begin{array}{l} \tilde{u}_{t_0}^1 = \frac{1}{6}\tilde{u}^4\tilde{u}_3^4\epsilon^2 + \frac{1}{12}(\tilde{u}^4)^2\tilde{u}_1^3 + \frac{1}{540}\tilde{u}^3\tilde{u}_7^3\epsilon^6 + \frac{25}{648}\tilde{u}^3\tilde{u}_2^3\tilde{u}_3^3\epsilon^4 + \frac{5}{216}\tilde{u}^3\tilde{u}_3^3\tilde{u}_4^3\epsilon^4 + \frac{5}{1296}(\tilde{u}^3)^2\tilde{u}_3^3\epsilon^4 \\ + \frac{1}{108}\tilde{u}^3(\tilde{u}_1^3)^3\epsilon^2 + \frac{1}{54}(\tilde{u}^3)^2\tilde{u}_1^3\tilde{u}_2^3\epsilon^2 + \frac{1}{18}\tilde{u}^3\tilde{u}_3^1\epsilon^2 + \frac{1}{324}(\tilde{u}^3)^3\tilde{u}_3^3\epsilon^2 + \frac{1}{1296}(\tilde{u}^3)^4\tilde{u}_1^3 + \frac{1}{18}\tilde{u}^2\tilde{u}_3^2\epsilon^2 \\ + \frac{1}{36}(\tilde{u}^2)^2\tilde{u}_1^3 + \frac{1}{6}\tilde{u}^3\tilde{u}^4\tilde{u}_1^4 + \frac{1}{18}\tilde{u}^2\tilde{u}_3^3\tilde{u}_1^2 + \frac{1}{3600}\tilde{u}_9^3\epsilon^8 + \frac{1}{48}\tilde{u}_3^3\tilde{u}_3^3\epsilon^6 + \frac{101}{6480}\tilde{u}_2^3\tilde{u}_5^3\epsilon^6 + \frac{1}{135}\tilde{u}_1^3\tilde{u}_6^3\epsilon^6 \\ + \frac{59}{1296}\tilde{u}_1^3(\tilde{u}_2^3)^2\epsilon^4 + \frac{7}{216}(\tilde{u}_1^3)^2\tilde{u}_3^3\epsilon^4 + \frac{2}{45}\tilde{u}_5^1\epsilon^4 + \frac{1}{12}\tilde{u}_1^3\tilde{u}_2^1\epsilon^2 + \frac{1}{12}\tilde{u}_1^2\tilde{u}_2^2\epsilon^2 + \frac{1}{36}\tilde{u}_1^1\tilde{u}_2^2\epsilon^2 + \frac{1}{4}\tilde{u}_1^4\tilde{u}_2^4\epsilon^2 \\ \tilde{u}_{t_0}^2 = \frac{1}{2}\tilde{u}^4\tilde{u}_1^4 + \frac{1}{12}\tilde{u}^3\tilde{u}_3^2\epsilon^2 + \frac{1}{36}(\tilde{u}^3)^2\tilde{u}_1^2 + \frac{1}{18}\tilde{u}^2\tilde{u}_3^3\epsilon^2 - \frac{1}{6}\tilde{u}^2\tilde{u}_1^2 + \frac{1}{18}\tilde{u}^2\tilde{u}_3\tilde{u}_1^3 + \frac{1}{20}\tilde{u}_5^2\epsilon^4 + \frac{1}{6}\tilde{u}_1^3\tilde{u}_2^2\epsilon^2 + \frac{5}{36}\tilde{u}_1^2\tilde{u}_2^3\epsilon^2 \\ \tilde{u}_{t_0}^3 = \tilde{u}_1^1 \\ \tilde{u}_{t_0}^4 = \frac{1}{18}\tilde{u}^4\tilde{u}_3^3\epsilon^2 + \frac{1}{6}\tilde{u}^4\tilde{u}_1^2 + \frac{1}{12}\tilde{u}^3\tilde{u}_3^4\epsilon^2 + \frac{1}{36}(\tilde{u}^3)^2\tilde{u}_1^4 + \frac{1}{6}\tilde{u}^2\tilde{u}_1^4 + \frac{1}{18}\tilde{u}^2\tilde{u}_4\tilde{u}_1^3 + \frac{1}{20}\tilde{u}_5^4\epsilon^4 + \frac{5}{36}\tilde{u}_1^4\tilde{u}_2^3\epsilon^2 + \frac{1}{6}\tilde{u}_1^3\tilde{u}_2^4\epsilon^2 \end{array} \right.$$

$$\left\{ \begin{array}{l} \tilde{u}_{t_0}^1 = \frac{1}{6}\tilde{u}^4\tilde{u}_3^3\epsilon^2 + \frac{1}{2}\tilde{u}^4\tilde{u}_1^2 + \frac{1}{4}\tilde{u}^3\tilde{u}_3^4\epsilon^2 + \frac{1}{12}(\tilde{u}^3)^2\tilde{u}_1^4 + \frac{1}{2}\tilde{u}^2\tilde{u}_1^4 + \frac{1}{6}\tilde{u}^3\tilde{u}^4\tilde{u}_1^3 + \frac{3}{20}\tilde{u}_5^4\epsilon^4 + \frac{5}{12}\tilde{u}_1^4\tilde{u}_2^3\epsilon^2 + \frac{1}{2}\tilde{u}_1^3\tilde{u}_2^4\epsilon^2 \\ \tilde{u}_{t_0}^2 = \frac{1}{2}\tilde{u}^4\tilde{u}_1^3 + \frac{1}{2}\tilde{u}^3\tilde{u}_1^4 + \frac{1}{2}\tilde{u}_3^4\epsilon^2 \\ \tilde{u}_{t_0}^3 = 3\tilde{u}_1^4 \\ \tilde{u}_{t_0}^4 = \frac{1}{36}\tilde{u}^3\tilde{u}_3^3\epsilon^2 + \frac{1}{6}\tilde{u}^3\tilde{u}_1^2 + \frac{1}{36}(\tilde{u}^3)^2\tilde{u}_1^3 + \frac{1}{6}\tilde{u}^2\tilde{u}_1^3 + \frac{1}{180}\tilde{u}_5^3\epsilon^4 + \frac{1}{36}\tilde{u}_1^3\tilde{u}_2^3\epsilon^2 + \frac{1}{6}\tilde{u}_3^2\epsilon^2 + \tilde{u}_1^1 \end{array} \right.$$

5.2.5 Details on the computation of the Hamiltonian density $g_{1,1}$

Here we describe how we computed explicitly the fundamental Hamiltonian density $g_{1,1}$, which is the key object in confirming the DR/DZ conjecture here. And the computation of $g_{1,1}$ has to do with the intrinsic grading of the CohFT. We refer to [DR19] for more details.

Consider an arbitrary CohFT with vectors space V and suppose that V is graded with $\deg e_1 = 0$. We say that the CohFT is *homogeneous* if there exists $\delta \in \mathbb{C}$ such that the maps $c_{g,n} : V^{\otimes n} \rightarrow H^*(\overline{\mathcal{M}}_{g,n}, \mathbb{C})$ and $\eta : V^{\otimes 2} \rightarrow \mathbb{C}$ have degree

$$\deg c_{g,n} = \delta(g-1), \quad \deg \eta = -\delta, \quad \delta \in \mathbb{C}, \quad 1 \leq \alpha \leq \dim V.$$

In this case the formal variables u_k^α and ϵ acquire a grading too and the Hamiltonian densities of the resulting DR hierarchy are homogeneous:

$$|u_k^\alpha| = 1 - \deg e_\alpha, \quad |\epsilon| = \frac{1 - \delta}{2}, \quad |g_{\alpha,d}| = d + 3 - \delta - |u^\alpha|.$$

Notice importantly that this grading $|\cdot|$ is not related to the differential grading on $\widehat{\mathcal{A}}$ and $\widehat{\mathcal{F}}$.

Integrable systems of DR type

We now need introduce the concept of integrable hierarchies of double ramification type [BDGR16].

Let $\{\cdot, \cdot\}$ denote the Poisson structure associated to the Hamiltonian operator $\eta^{\mu\nu} \partial_x$. For a local functional $\bar{h} \in \widehat{\mathcal{F}}^{[0]}$ consider the operator $\mathcal{D}_{\bar{h}} : \widehat{\mathcal{A}}[[z]] \rightarrow \widehat{\mathcal{A}}[[z]]$ defined by

$$\mathcal{D}_{\bar{h}} = \partial_x \circ (D - 1) - z\{\cdot, \bar{h}\}, \quad D := \sum_{k \geq 0} u_k^\alpha \frac{\partial}{\partial u_k^\alpha} + \varepsilon \frac{\partial}{\partial \varepsilon}.$$

Suppose there exist $N = \dim V$ solutions $g_\alpha(z) \in \widehat{\mathcal{A}}^{[0]}[[z]]$, with $\alpha = 1, \dots, N$, to $\mathcal{D}_{\bar{h}} g_\alpha(z) = 0$ with initial conditions $g_\alpha(z=0) = \eta_{\alpha\mu} u^\mu$. Then a new vector of solutions in the same class can be found by the following transformation

$$g_\alpha(z) \mapsto a_\alpha^\mu(z) g_\mu(z) + b_\alpha(z), \quad (5.47)$$

where $a_\alpha^\mu(z) = \delta_\alpha^\mu + \sum_{i>0} a_{\alpha,i}^\mu z^i \in \mathbb{C}[[z]]$ and $b_\alpha(z) = \sum_{i>0} b_{\alpha,i} z^i \in \mathbb{C}[[z]]$.

The following result constitutes the main technical tool in the computation the the Hamiltonian density $g_{1,1}$ of DR hierarchy of the D_4 CohFT.

Theorem 5.2.28 ([BDGR16]). *Assume that $\bar{h} \in \widehat{\mathcal{F}}^{[0]}$ has the following properties:*

- (a) *there exist N independent solutions $g_\alpha(z) = \sum_{p \geq 0} g_{\alpha,p-1} z^p \in \widehat{\mathcal{A}}^{[0]}[[z]]$, $\alpha = 1, \dots, N$, to the equation*

$$\mathcal{D}_{\bar{h}} g_\alpha(z) = 0 \quad (5.48)$$

with the initial conditions $g_\alpha(z=0) = \eta_{\alpha\mu} u^\mu$,

- (b) $\frac{\delta \bar{h}}{\delta u^1} = \frac{1}{2} \eta_{\mu\nu} u^\mu u^\nu + \partial_x^2 r$, $r \in \widehat{\mathcal{A}}^{[-2]}$.

Then, up to a transformation of type (5.47), we have

- (i) $g_{1,0} = \frac{1}{2} \eta_{\mu\nu} u^\mu u^\nu + \partial_x^2 (D - 1)^{-1} r$,

- (ii) $\bar{g}_{1,1} = \bar{h}$,

- (iii) $\{\bar{g}_{\alpha,p}, \bar{g}_{\beta,q}\} = 0$, $\alpha, \beta = 1, \dots, N$, $p, q \geq -1$,

- (iv) $\{g_{\alpha,p}, \bar{g}_{\beta,0}\} = \partial_x \frac{\partial g_{\alpha,p+1}}{\partial u^\beta}$, $\beta = 1, \dots, N$, $p \geq -1$,

- (v) $\frac{\partial g_{\alpha,p}}{\partial u^1} = g_{\alpha,p-1}$, $\alpha = 1, \dots, N$, $p \geq -1$,

hence in particular \bar{h} is part of an integrable tau-symmetric hierarchy.

We call a system of densities originating from an Hamiltonian $\bar{h} = \bar{g}_{1,1}$ as in the theorem above an *integrable system of double ramification type*. As proven in [BDGR16], the DR hierarchy of any given cohomological field theory, is always and integrable system of DR type. This fact implies in particular that the entire hierarchy of DR Hamiltonian densities can be reconstructed from $\bar{g}_{1,1} \in \widehat{\mathcal{F}}^{[0]}$ alone by means of the DR recursion equation [BR16a]

$$\partial_x (D - 1) g_{\alpha,d+1} = \{g_{\alpha,d}, \bar{g}_{1,1}\}, \quad g_{\alpha,-1} = \eta_{\alpha\mu} u^\mu, \quad 1 \leq \alpha \leq N, d \geq -1. \quad (5.49)$$

The above result will be used here to effectively compute the Hamiltonian $\bar{g}_{1,1}$ itself, starting from a limited amount of information on the CohFT, as we will show in the next section.

Computation via the Frobenius manifold of type D_4

In [Dub99], based on the work of Kyoji Saito [Sai81, Sai83a, Sai83b], Boris Dubrovin constructed a structure of Frobenius manifold [Dub96] on the space of orbits of any finite irreducible Coxeter group. This Frobenius manifold is generically semisimple and conformal, so Givental-Teleman [Giv01, Tel12] (see also [PPZ15]) theory can be applied to produce a uniquely defined homogeneous cohomological field theory.

As already stated, we are interested in this construction for the case of the D_4 simple singularity, i.e. $W = x^3 + xy^2$. The resulting CohFT $c_{g,n} : V^{\otimes n} \rightarrow \overline{\mathcal{M}}_{g,n}$ has vector space $V = \text{Span}(e_1, \dots, e_4)$ with cohomological degrees given by

$$\deg e_1 = 0, \quad \deg e_2 = \deg e_4 = \frac{1}{3}, \quad \deg e_3 = \frac{2}{3}, \quad \deg c_{g,n} = \delta = \frac{2}{3}.$$

In genus 0 the corresponding Frobenius structure can be described by the Frobenius potential (see for instance [LRZ15])

$$F = \frac{t^1 (t^2)^2}{12} + \frac{(t^1)^2 t^3}{12} - \frac{(t^2)^3 t^3}{216} + \frac{(t^2)^2 (t^3)^3}{1296} + \frac{(t^3)^7}{1632960} + \frac{t^1 (t^4)^2}{4} + \frac{t^2 t^3 (t^4)^2}{24} + \frac{(t^3)^3 (t^4)^2}{432}, \quad (5.50)$$

which, in particular, gives the metric

$$(\eta^{\alpha\beta}) = \left(\frac{\partial^3 F}{\partial t^\alpha \partial t^\beta} \right) = \begin{pmatrix} 0 & 0 & \frac{1}{6} & 0 \\ 0 & \frac{1}{6} & 0 & 0 \\ \frac{1}{6} & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} \end{pmatrix} \quad (5.51)$$

Notice that this CohFT was proved to be isomorphic to the quantum singularity theory of Fan-Jarvis-Ruan-Witten [FJR13] for the simple singularity $W = x^3 + xy^2$, with respect to the non-maximal diagonal symmetry group $\langle J \rangle = \mathbb{Z}/3\mathbb{Z}$ [FFJMR16].

In [GM05, FGM10] it was proved that the Dubrovin-Zhang hierarchy for the Dubrovin-Saito CohFT associated to a Coxeter group coincides with the Drinfeld-Sokolov hierarchy of the corresponding semisimple Lie algebra. As explained above, the $\varepsilon \rightarrow 0$ limit of both the DZ and DR hierarchies of any (semisimple) CohFT coincides with the principal hierarchy of the Frobenius manifold which is, consequently, completely determined, thanks to the recursion equation (5.39), by the Hamiltonian $\bar{g}_{1,1}^{[0]} = \bar{g}_{1,1}|_{\varepsilon=0}$.

To compute the latter in the D_4 case we can use (see [DZ08]) the dilaton equation of the DZ principal hierarchy (which coincides here with the dispersionless DR hierarchy), namely $\bar{g}_{1,1}^{[0]} = \int (D-2)(F(t^*)|_{t^*=u^*}) dx$, to obtain

$$\begin{aligned} \bar{g}_{1,1}^{[0]} = \int & \left[\frac{1}{12} u^1 (u^2)^2 + \frac{1}{12} (u^1)^2 u^3 - \frac{1}{108} (u^2)^3 u^3 + \frac{1}{432} (u^2)^2 (u^3)^3 + \frac{1}{326592} (u^3)^7 + \frac{1}{4} u^1 (u^4)^2 \right. \\ & \left. + \frac{1}{12} u^2 u^3 (u^4)^2 + \frac{1}{144} (u^3)^3 (u^4)^2 \right]. \end{aligned} \quad (5.52)$$

Moreover, thanks to homogeneity of the CohFT, we know that, for any $k \geq 0$,

$$|u_k^1| = 1, \quad |u_k^2| = |u_k^4| = \frac{2}{3}, \quad |u_k^3| = \frac{1}{3}, \quad |\varepsilon| = \frac{1}{6}, \quad |\bar{g}_{1,1}| = \frac{7}{3}. \quad (5.53)$$

In Equation (5.44), we stated that the dispersive expansion of $\text{bar}g_{1,1}$ is given by

$$\begin{aligned} \bar{g}_{1,1} = & \int \left(\frac{1}{12} u^1 (u^2)^2 + \frac{1}{12} (u^1)^2 u^3 - \frac{1}{108} (u^2)^3 u^3 + \frac{1}{432} (u^2)^2 (u^3)^3 + \frac{1}{326592} (u^3)^7 + \frac{1}{4} u^1 (u^4)^2 \right. \\ & + \frac{1}{12} u^2 u^3 (u^4)^2 + \frac{1}{144} (u^3)^3 (u^4)^2 \Big) + \left(-\frac{1}{6} (u_1^1)^2 + \frac{1}{24} u^2 (u_1^2)^2 - \frac{1}{72} (u^3)^2 (u_1^2)^2 \right. \\ & - \frac{1}{108} (u^3)^2 u_1^1 u_1^3 - \frac{1}{27} u^2 u^3 u_1^2 u_1^3 - \frac{1}{108} (u^2)^2 (u_1^3)^2 - \frac{1}{2592} (u^3)^4 (u_1^3)^2 - \frac{1}{36} (u^4)^2 (u_1^3)^2 \\ & - \frac{1}{4} u^4 u_1^2 u_1^4 - \frac{1}{9} u^3 u^4 u_1^3 u_1^4 - \frac{1}{8} u^2 (u_1^4)^2 - \frac{1}{24} (u^3)^2 (u_1^4)^2 \Big) \varepsilon^2 + \left(-\frac{35}{46656} u^3 (u_1^3)^4 \right. \\ & + \frac{1}{144} (u_1^3)^2 u_2^1 + \frac{5}{216} u_1^2 u_1^3 u_2^2 + \frac{1}{48} u^3 (u_2^2)^2 + \frac{1}{54} u^3 u_2^1 u_2^3 + \frac{5}{216} u^2 u_2^2 u_2^3 + \frac{7}{7776} (u^3)^3 (u_2^3)^2 \\ & + \frac{5}{72} u_1^3 u_1^4 u_2^4 + \frac{5}{72} u^4 u_2^3 u_2^4 + \frac{1}{16} u^3 (u_2^4)^2 \Big) \varepsilon^4 + \left(\frac{1}{486} (u_1^3)^2 (u_2^3)^2 + \frac{13}{5832} u^3 (u_2^3)^3 \right. \\ & - \frac{1}{112} (u_3^2)^2 - \frac{13u_3^1 u_3^3}{1512} - \frac{1}{972} (u^3)^2 (u_3^2)^2 - \frac{3}{112} (u_3^4)^2 \Big) \varepsilon^6 + \left(-\frac{5}{1728} u_2^3 (u_3^3)^2 + \frac{1}{1728} u^3 (u_4^3)^2 \right) \varepsilon^8 \\ & - \frac{1}{7776} (u_5^3)^2 \varepsilon^{10}. \end{aligned}$$

We finish this section by explaining how we computed the above started from the data of $\bar{g}_{1,1}^{[0]} = \bar{g}_{1,1}|_{\varepsilon=0}$. By definition, only even powers of ε appear in the Hamiltonian densities of any DR hierarchy. Now remark from (5.53) that all the variables have positive degree, guaranteeing the polynomiality of $\bar{g}_{1,1}$ and, in fact, of all the Hamiltonian densities $g_{\alpha,d}$, for all $1 \leq \alpha \leq N$ and $d \geq -1$. In particular, since $|\varepsilon| = \frac{1}{6}$ and since in each monomials there are as many x -derivatives as powers of ε , we see that any term where the power of ε is bigger than 10 is either trivial or a total x derivative.

One can then verify, by direct computation, that up to rescaling of the variable ε , there exist a unique local functional $\bar{g}_{1,1} \in \widehat{\mathcal{F}}^{[0]}$ such that the following are satisfied:

- ◇ $g_{1,1}^{[0]}$ is given by (5.52);
- ◇ $g_{1,1}$ contains only even powers of ε ;
- ◇ the cohomological degrees satisfy $|\bar{g}_{1,1}| = \frac{7}{3}$;
- ◇ $\bar{g}_{1,1}$ satisfies Equation (5.39).

In practice one writes down the most general polynomial deformation of (5.52) with the given degree and notices that imposing conditions (a) and (b) of theorem 5.2.28 determines, up to rescalings of ε , all the coefficients.

Finally, to determine the correct normalization of ε , it is sufficient to compute the coefficient of a single monomial of $\bar{g}_{1,1}$ containing ε . We can do this by recalling (see for instance [BDGR18]) that, for any CohFT, $\text{Coef}[(u_1^1)^2 \varepsilon^2](\bar{g}_{1,1}) = -\frac{1}{24} \dim V$.

5.3 The Drinfeld–Sokolov hierarchy of type D_4

In Section 4.3, we describe the celebrated Drinfeld–Sokolov procedure [DS84] which associate an integrable hierarchy of PDEs to an arbitrary affine Lie algebra (although in the present

thesis we described it only for untwisted algebras). These hierarchies have the form of matrix Lax pairs (4.43). But as we can see with the example of the KdV hierarchy 4.3.3 (i.e. the DS hierarchy of type A_1), it becomes quickly cumbersome to actually compute the flows of the hierarchy using the matrix Lax pairs.

Already in [DS84], Drinfeld and Sokolov gave a scalar Lax pair for the hierarchies of type $A_n^{(1)}, B_n^{(1)}, C_n^{(1)}$ and $D_n^{(1)}$ (and others), along with a bi-Hamiltonian representation. Yet for the hierarchy of type $D_n^{(1)}$ precisely, only *part* of the hierarchy was represented. The part that has no scalar Lax pair representation we call the *negative flows* and we insist on the fact that it does have a matrix Lax pair though. To make this statement clearer, we need to give a brief description of the untwisted affine algebra $D_n^{(1)}$.

The untwisted affine algebra of type $D_n^{(1)}$

The simple Lie algebras of type D_n are isomorphic to the algebra $\mathfrak{o}(2n)$ of $2n$ -dimensional complex skew-Hermitian matrices, i.e.

$$\mathfrak{o}(2n) = \{M \in \mathfrak{gl}(2n, \mathbb{C}) \mid (M^\dagger)^* + M = 0\}$$

These algebras have Coxeter number

$$h = 2n - 2.$$

The matrices in $\mathfrak{o}(2n)$ are not necessarily traceless and cannot be nilpotent. We take the usual faithful traceless representation of $\mathfrak{o}(2n)$ (see [DS84], but also [LWZ10, CW15b]):

$$\mathfrak{o}(2n) \simeq \{M \in \mathfrak{gl}(2n, \mathbb{C}) \mid MS + S(M^{\text{at}}) = 0\} =: \mathfrak{g},$$

where $(M^{\text{at}})_{i,j} = M_{n+1-j, n+1-i}$ is the antitranspose of M and

$$S = \text{diag}(1, -1, \dots, (-1)^{n-1}, (-1)^{n-1}, \dots, -1, 1).$$

A direct computation shows indeed that $\text{tr}(MS + SM^{\text{at}}) = 2\text{tr} M$ so that $\mathfrak{g} \subset \mathfrak{sl}(2n)$. As before, we denote $\mathfrak{h}, \mathfrak{b}, \mathfrak{n}$ the Cartan, negative Borel and negative nilpotent subalgebras of \mathfrak{g} . We choose the following Weyl generators of the associated affine algebra $\mathfrak{g} = \mathfrak{g} \otimes \mathbb{C}[\lambda, \lambda^{-1}] \oplus \mathbb{C}c$:

$$\begin{aligned} e_i &= E_{i+1,i} + E_{2n+1-i, 2n-i}, & e_n &= \frac{1}{2}(E_{n+1, n-1} + E_{n+2, n}), \\ f_i &= E_{i, i+1} + E_{2n-i, 2n+1-i}, & f_n &= 2(E_{n-1, n+1} + E_{n, n+2}), \\ h_i &= E_{i+1, i+1} - E_{i, i} + E_{2n+1-i, 2n+1-i} - E_{2n-i, 2n-i}, & h_n &= E_{n+1, n+1} - E_{n, n} + E_{n+2, n+2} - E_{n-1, n-1} \end{aligned}$$

and

$$\begin{aligned} e_0 &= \frac{1}{2}(E_{1, 2n-1} + E_{2, 2n}) \otimes \lambda, & f_0 &= 2(E_{2n-1, 1} + E_{2n, 2}) \otimes \lambda^{-1}, \\ h_0 &= E_{1, 1} + E_{2, 2} - E_{2n-1, 2n-1} - E_{2n, 2n}. \end{aligned}$$

The affine algebra \mathfrak{g} has affine exponents $\mathbb{E} = \mathbb{Z}^{\text{odd}} \sqcup (n-1)\mathbb{Z}^{\text{odd}}$. It is often condensed in the following form:

$$\mathbb{E} = \{1, 3, 5, \dots, 2n-3\} \cup \{(n-1)'\} + (2n-2)\mathbb{Z},$$

where the $(n-1)'$ means that if, n is even, every integer that is congruent to $n-1$ modulo $2n-2$ has multiplicity 2; such integers are of the form $a = (n-1)(2b+1)$, with $b \in \mathbb{Z}$. The principal Heisenberg subalgebra admits the following family of generators [Kac90]:

$$\left\{ -\Lambda_k \in \mathfrak{g}^k, \Gamma_k \in \mathfrak{g}^{k(n-1)} \mid k \in \mathbb{Z}^{\text{odd}} \right\},$$

where

$$\Lambda_j := \begin{cases} \Lambda^j & \text{if } j > 0, \\ (\lambda^{-1}\Lambda^{2n-3})^{-j} & \text{if } j < 0, \end{cases} \quad \Gamma_j := \begin{cases} \Gamma^j & \text{if } j > 0, \\ (\lambda^{-1}\Gamma)^{-j} & \text{if } j < 0, \end{cases} \quad (5.54)$$

where $\Lambda = \sum_{i=0}^n e_i$ is the principal cyclic element and

$$\Gamma := \mu \left[E_{n,1} - \frac{1}{2}E_{n+1,1} - \frac{1}{2}E_{n,2n} \otimes \lambda + \frac{1}{4}E_{n+1,2n} \otimes \lambda \right. \quad (5.55)$$

$$\left. + (-1)^n (E_{2n,n+1} - \frac{1}{2}E_{2n,n} - \frac{1}{2}E_{1,n+1} \otimes \lambda + \frac{1}{4}E_{1,n} \otimes \lambda) \right], \quad (5.56)$$

where $\mu = 1$ if n is even and $\mu = \sqrt{-1}$ when n is odd.

The positive and negative flows of the hierarchy of type D_n

Let $\mathcal{L} = \partial + \Lambda + Q \in \text{op}_{\mathfrak{g}}$, that is, $Q \in \mathfrak{b}[[\mathfrak{t}]]$. Let $U \in \widehat{\mathfrak{g}}^-[[\mathfrak{t}]]$ be a function as in Theorem 4.3.7. Following [LRZ15], we introduce the following Drinfeld–Sokolov equations of type D_n :

$$\frac{\partial \mathcal{L}}{\partial t_k} = [\mathcal{L}, (e^{\text{ad}_U} \Lambda_k)_+], \quad (5.57)$$

$$\frac{\partial \mathcal{L}}{\partial \hat{t}_k} = [\mathcal{L}, (e^{\text{ad}_U} \Gamma_k)_+], \quad (5.58)$$

with $k \in \mathbb{Z}_{>0}^{\text{odd}}$. By construction, they generate the whole DS hierarchy. The flows along ∂_{t_k} we call the *positive flows*, while the flows along $\partial_{\hat{t}_k}$ we call the *negative flows*.

In [DS84] (§7.1), Drinfeld and Sokolov introduced the following PDO⁹

$$L = \partial^{2n-2} + \partial^{-1} \sum_{\mu=1}^{n-3} (s^\mu \partial^{2\mu-1} + \partial^{2\mu-1} s^\mu) + \partial^{-1} \varrho \partial^{-1} \varrho, \quad s^n = (\varrho)^2,$$

with the condition $\partial \circ L = L^* \circ \partial$. Drinfeld and Sokolov proved that the positive flows of the hierarchy admit the following scalar Lax pair representation:

$$\frac{\partial L}{\partial t_k} = [L, (P^k)_+], \quad (5.59)$$

where $P = L^{\frac{1}{2n-2}}$, which exists in virtue of Theorem 3.1.8. The precise statement is that there exists a gauge slice $\tilde{\mathcal{L}}$ within the gauge class of the connection \mathcal{L} such that Equation (5.59)

⁹Hereafter, we will use μ, ν to denote the coordinates s^μ , with $\mu \in \{1, \dots, n-1\}$ and α, β to denote the coordinates s^α , with $\alpha \in \{1, \dots, n\}$.

is exactly Equation (5.57) on the connection $\tilde{\mathcal{L}}$. The latter implies that the coordinates s^α are differential polynomials in the coordinates q^α of the function Q .

Now the problem attacked in [LRZ15] is that the negative flows cannot be written using the $(2n-2)$ -th root P of L . To solve this problem, the authors introduced a new class of pseudo-differential operators they called *of the second type*. Doing so, they were able to produce a square root for the operator L . We briefly describe their construction in the next section.

5.3.1 Scalar Lax pair and bi-Hamiltonian structure of the DS hierarchy of type D_n

Pseudo-differential operators of the second type

In this section we describe the construction pseudo-differential operators (PDOs) of the second type, as done in [LWZ10]. We start the algebra of (un-extended) differential polynomials \mathcal{A} with yet another set of indeterminates s_k^1, \dots, s_k^n , with $k \geq 0$, namely

$$\mathcal{A} := \mathcal{A}(s^1, \dots, s^n) = \mathbb{C} \left[s_k^\alpha \mid \alpha \in \{1, \dots, n\}, k \geq 0 \right],$$

equipped with the derivation

$$\partial = \sum_{k \geq 0} \sum_{\alpha=1}^n s_{k+1}^\alpha \frac{\partial}{\partial s_k^\alpha}.$$

In the previous sections, we introduced the extended algebra of differential polynomials $\widehat{\mathcal{A}} = \mathbb{C}[[u_0^*]][[\varepsilon]][[u_{k>0}^*]]$, but for the present construction, they are not needed. They will appear when we will compare the DS and the DR hierarchies. The ring \mathcal{A} is graded by $\deg s_k^\alpha = k$, producing a decomposition

$$\mathcal{A} = \bigoplus_{k \geq 0} \mathcal{A}_k, \quad \mathcal{A}_k \cdot \mathcal{A}_\ell \subset \mathcal{A}_{k+\ell},$$

where \mathcal{A}_k is the subspace of homogeneous elements of degree k . For this gradation, the derivation ∂ has degree 1, i.e. $\partial \mathcal{A}_k \subset \mathcal{A}_{k+1}$. Notice that Liu–Wu–Zhang’s construction [LWZ10] works for an arbitrary commutative ring as long as it is graded, topologically complete and equipped with degree-1 derivation.

Now we consider the following additive group:

$$\widehat{\mathcal{D}} = \left\{ \sum_{k=-\infty}^{+\infty} a^k \partial^k \mid a^k \in \mathcal{A} \right\}$$

Of course, the traditional PDOs (the one we will now call *of the first type*) form a subgroup

$$\mathcal{D} = \left\{ \sum_{k=-\infty}^m a^k \partial^k \mid a^k \in \mathcal{A}, m \in \mathbb{Z} \right\} \subset \widehat{\mathcal{D}}.$$

Clearly, the ring structure (3.1) on \mathcal{D} cannot be extended to $\widehat{\mathcal{D}}$. In what follows, we describe how to produce a new ring of whose elements are allowed to contain infinitely positive and negative powers of ∂ .

Let $A \in \mathcal{D}$ be a PDO of the first type. We say that A is homogeneous of degree k if A has the form

$$A = \sum_{i=-\infty}^k a^i \partial^i, \quad a^i \in \mathcal{A}_{k-i}.$$

We denote by \mathcal{D}_k the subspace of homogeneous operators of degree k . Then clearly

$$\widehat{\mathcal{D}} = \prod_{k \in \mathbb{Z}} \mathcal{D}_k.$$

We introduce the following subgroups of $\widehat{\mathcal{D}}$:

$$\widetilde{\mathcal{D}}_{(d)} = \prod_{k \geq d} \mathcal{D}_k, \quad \widetilde{\mathcal{D}} = \bigcup_{d \in \mathbb{Z}} \widetilde{\mathcal{D}}_{(d)}.$$

Notice that the above induces a decreasing filtration $\widetilde{\mathcal{D}}_{(d)} \subset \widetilde{\mathcal{D}}_{(d-1)}$. Recall that the product of PDOs of the first type is given, for $a, b \in \mathcal{A}$ and $n, m \in \mathbb{Z}$, by

$$a\partial^n \circ b\partial^m = \sum_{k \geq 0} \binom{n}{k} ab_{(k)} \partial^{n+m-k}, \quad \binom{n}{k} = \frac{(n-k+1) \cdots (n-1) \cdot n}{k!}. \quad (5.60)$$

Clearly, if $A \in \mathcal{D}_k$ and $B \in \mathcal{D}_\ell$ then $AB \in \mathcal{D}_{k+\ell}$, so the product can be extended to $\widetilde{\mathcal{D}}$, which makes it into an algebra.

Definition 5.3.1. We call PDOs of the second type the elements of the algebra $\widetilde{\mathcal{D}}$.¹⁰ The elements of the subalgebra $\mathcal{D}^b = \mathcal{D} \cap \widetilde{\mathcal{D}}$ are called bounded operators.

By the gradation on \mathcal{A} , the general form of an element $A \in \widehat{\mathcal{D}}$ reads

$$A = \sum_{i \in \mathbb{Z}} \sum_{j \geq 0} a^{i,j} \partial^i. \quad (5.61)$$

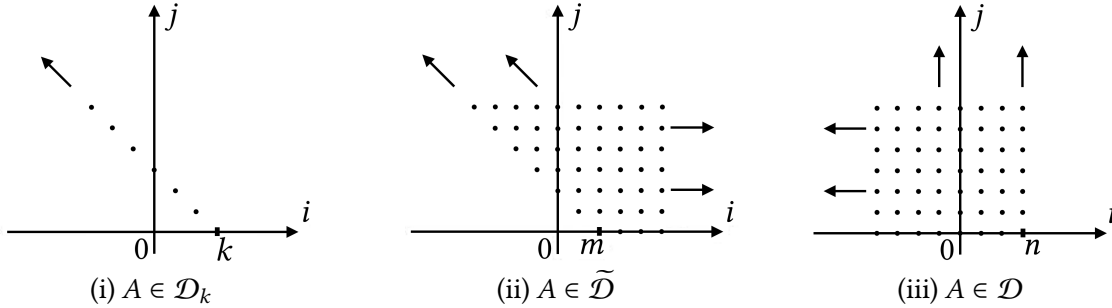
The following lemma gives a nice characterization of \mathcal{D} and $\widetilde{\mathcal{D}}$.

Lemma 5.3.2. Let A be of the general form of Equation (5.61), then the following characterization holds true.

- (i) $A \in \mathcal{D}_k$ iff $a^{i,j} = 0$ outside of the diagonal $\{(i, j) \in \mathbb{Z} \times \mathbb{Z}_{\geq 0} \mid i + j = k\}$;
- (ii) $A \in \widetilde{\mathcal{D}}$ iff there exists m such that $a^{i,j} = 0$ outside of the set $\{(i, j) \in \mathbb{Z} \times \mathbb{Z}_{\geq 0} \mid j \geq m - i\}$;
- (iii) $A \in \mathcal{D}$ iff there exists n such that $a^{i,j} = 0$ outside of the set $\{(i, j) \in \mathbb{Z} \times \mathbb{Z}_{\geq 0} \mid i \leq n\}$.

Proof. Point (i) is the definition of \mathcal{D}_k . Point (ii) follows from the decreasing filtration property $\widetilde{\mathcal{D}}_{(d)} \subset \widetilde{\mathcal{D}}_{(d-1)}$. Point (iii) is the definition of \mathcal{D} . \square

Figure 5.1: Graphical interpretation of Lemma 5.3.2. Each figure shows the lattice $\mathbb{Z} \times \mathbb{Z}_{\geq 0}$; each dot on the lattice indicates the points (i, j) where $a^{i,j}$ needs not vanish.



Lemma 5.3.2 has a nice graphical interpretation, as shown in Figure 5.3.1; the latter is taken from [LRZ15] with minor modifications.

Similarly to the case of PDOs of the first type, we can define the positive and negative parts, the residue as well as the adjoint:

$$\begin{aligned}
 A_+ &= \sum_{i \geq 0} \sum_{j \geq 0} a^{i,j} \partial^i, & A_- &= \sum_{i < 0} \sum_{j \geq 0} a^{i,j} \partial^i, \\
 A^* &= \sum_{i \in \mathbb{Z}} \sum_{j \geq 0} (-1)^i \partial \circ a^{i,j}, & \text{res } A &= \sum_{j \geq 0} a_{-1,j}.
 \end{aligned}$$

We know that if $A \in \mathcal{D}$, then $A_+, A_-, A^* \in \mathcal{D}$; moreover, $A_- \in \tilde{\mathcal{D}}$ as well so that $A_+ \in \mathcal{D}^b$. Similarly, if $A \in \tilde{\mathcal{D}}$, then we easily see that $A_+, A_-, A^* \in \tilde{\mathcal{D}}$ and that $A_- \in \mathcal{D}$ so that $A_- \in \mathcal{D}^b$.

In the next section, we describe how the algebra $\tilde{\mathcal{D}}$ of PDOs of the second type were used by Liu Ruan and Zhang to construct a scalar Lax pair for the negative flows of the DS hierarchy of type D_4 .

Scalar Lax pair of the DS hierarchy of type D_n

Firstly, recall the following operator of the first type,

$$L = \partial^{2n-2} + \partial^{-1} \sum_{\mu=1}^{n-1} (s^\mu \partial^{2\mu-1} + \partial^{2\mu-1} s^\mu) + \partial^{-1} \varrho \partial^{-1} \varrho,$$

$s^n = (\varrho)^2$ with the condition $\partial \circ L = L^* \circ \partial$. The following proposition holds.

Proposition 5.3.3 ([LWZ10]). *There exists a unique PDO of the first type $P \in \mathcal{D}$ of the form*

$$P = \partial + \sum_{k < 0} p^k \partial^k$$

¹⁰In [LWZ10], the algebra \mathcal{D} is denote \mathcal{D}^- while the algebra $\tilde{\mathcal{D}}$ is denoted \mathcal{D}^+ . Here we made a different choice since \mathcal{D}_+ and \mathcal{D}_- were already used in Chapter 3.

such that $P^{2n-2} = L$. Moreover, the operator P satisfies $[P, L] = 0$ and

$$P^* = -\partial P \partial^{-1}.$$

The existence and uniqueness of the $(2n - 2)$ -th root P of L was already established since it involves only PDOs of the first type; see Theorem 3.1.8. However, the additional equations satisfied by P was not established here. We refer to [LRZ15] for the proof. The proposition below is the main tool in building the scalar Lax pair of the negative flows.

Proposition 5.3.4 ([LRZ15]). *There exists a unique PDO of the second type $Q \in \tilde{\mathcal{D}}$ of the form*

$$Q = \partial \circ \varrho + \sum_{m \geq 0} Q^m \circ \partial, \quad (5.62)$$

such that $Q^2 = L$ and where $Q^m \in \mathcal{D}^b$ of order $2m$. Moreover, Q satisfies

$$\begin{aligned} Q^* &= -\partial Q \partial^{-1} \\ Q_+^*(\varrho) &= -\frac{1}{2}(\partial \circ L_+)(1). \end{aligned} \quad (5.63)$$

Now the $(2n - 2)$ -th root $P \in \mathcal{D}$ and the square root $Q \in \tilde{\mathcal{D}}$ together allows to define a hierarchy represented by a scalar Lax pair. The latter is equivalent to the full matrix Lax pairs of the Drinfeld–Sokolov hierarchy of type D_n .

Theorem 5.3.5 ([LWZ10]). *The hierarchy of equations defined by the following flows is consistent and commute pairwise: for $k \in \mathbb{Z}_{>0}^{\text{odd}}$,*

$$\frac{\partial L}{\partial t_k} = \left[L, (P^k)_+ \right], \quad \frac{\partial L}{\partial \hat{t}_k} = \left[L, (Q^k)_+ \right] \quad (5.64)$$

Moreover, it coincides with the flows defined by the matrix Lax pair equations (5.57) and (5.58) respectively.

Bi-Hamiltonian structure of the DS hierarchy of type D_n

We now describe the bi-Hamiltonian structure of the DS hierarchy of type D_n . Notice that the local Poisson structure we will define were already given in [DS84]. What was not, is the Hamiltonian densities for the negative flows. First define the following operator $L' := \partial \circ L \in \mathcal{D}$, which we write in the following form:

$$L' = \partial \circ L =: \partial^{2n-1} + \sum_{\mu=1}^{n-1} (s^\mu \partial^{2\mu-1} + \partial^{2\mu-1} s^\mu) + \varrho \partial^{-1} \circ \varrho.$$

It satisfies the condition $L' = -(L')^*$. We can represent the operator the L' in an alternative form:

$$L' = \partial^{2n-1} + \sum_{\mu=1}^{n-1} (v^\mu \partial^{2\mu-1} + \tilde{v}^\mu \partial^{2\mu-2}) + \varrho \partial^{-1} \circ \varrho. \quad (5.65)$$

Then the functions v^μ are related to the s^μ by an Miura transform (within differential polynomials) while the \tilde{v}^μ are related to the functions v^μ via a non invertible transform thanks to the condition $L' = -(L')^*$.

Given a local functional $\bar{f} \in \mathcal{F}$, we can define its variational derivative w.r.t. L' (or its variational differential) by

$$\frac{\delta \bar{f}}{\delta L'} := \frac{\delta f}{\delta v^n} + \frac{1}{2} \sum_{\mu=1}^{n-1} \partial^{-2\mu} \circ \frac{\delta f}{\delta v^\mu} + \frac{\delta f}{\delta v^\mu} \partial^{-2\mu}. \quad (5.66)$$

Notice that the above expression is well defined since the variational derivatives do not depend on the choice of the representative in the class modulo $\text{Im}(\partial)$. Besides, this choice of variational differential is specific to the algebra of type D_n ; for other algebras, see [DS84] but also [DLZ08] for a particularly clear exposé.

The bi-Hamiltonian structure of the DS hierarchy of type D_n is given by the following local Poisson structures: let two local functionals $\bar{f}, \bar{g} \in \widehat{\mathcal{F}}$ and their variational differentials $X = \frac{\delta \bar{f}}{\delta L'}, Y = \frac{\delta \bar{g}}{\delta L'}$, then

$$\begin{aligned} \{\bar{f}, \bar{g}\}_1 &= \int \text{res} X [(\partial_x Y + L')_- - (L' Y + \partial_x)_- - (\partial_x Y - L')_+ + (L' Y - \partial_x)_+] dx, \\ \{\bar{f}, \bar{g}\}_2 &= \int \text{res} X [(L' Y)_+ L' - L' (Y L')_+] dx. \end{aligned}$$

These brackets are compatible in the sense that for any $\lambda, \mu \in \mathbb{C}$, the map $\lambda\{\cdot, \cdot\}_1 + \mu\{\cdot, \cdot\}_2$ still satisfies Jacobi identity. In [LWZ10], the authors prove the following theorem.

Theorem 5.3.6. *The hierarchy (5.64) admits the following bi-Hamiltonian representation: for any local functional $\bar{f} \in \mathcal{F}$,*

$$\frac{\partial s^\alpha}{\partial t_k} = \{s^\alpha, H_{k+2n-2}\}_1 = \{s^\alpha, H_k\}_2, \quad \frac{\partial s^\alpha}{\partial \hat{t}_k} = \{s^\alpha, \hat{H}_{k+2}\}_1 = \{s^\alpha, \hat{H}_k\}_2, \quad (5.67)$$

where the Hamiltonian functionals are given by

$$H_k = \frac{2n-2}{k} \int \text{res} P^k dx, \quad \hat{H}_k = \frac{2}{k} \int \text{res} Q^k dx$$

Tau structure

Finally, we describe the tau structure given in [LWZ10] by means of these PDOs of the second type. The authors prove that the Hamiltonian functionals $H_1, H_3, \dots, H_{2n-3}$ and \hat{H}_1 are linearly independent and are Casimir of the first Poisson bracket $\{-, -\}_1$.

Next we need to introduce the so-called *topological variables*: for all $\mu \in \{1, \dots, n-1\}$,

$$t_{\mu,p} := \frac{(2n-2)\Gamma(p+1 + \frac{2\alpha-1}{2n-2})}{\Gamma(\frac{2\alpha-1}{2n-2})} t_{(2n-2)p+2\alpha-1}, \quad t_{n,p} := \frac{2\Gamma(p + \frac{3}{2})}{\Gamma(\frac{1}{2})} \hat{t}_{2p+1}.$$

Then define the following differential polynomials: for all $\mu \in \{1, \dots, n-1\}$,

$$h_{\mu, p-1}^{\text{DS}} := \frac{\Gamma(\frac{2\alpha-1}{2n-2})}{(2n-2)\Gamma(p+1+\frac{2\alpha-1}{2n-2})} \text{res } P^{(2n-2)p+2\alpha-1}, \quad h_{n, p-1}^{\text{DS}} := \frac{2\Gamma(\frac{1}{2})}{\Gamma(p+\frac{3}{2})} \text{res } Q^{2p+1} \quad (5.68)$$

Then we can rewrite the Hamiltonian flows (5.67) as

$$\frac{\partial s^\alpha}{\partial t_{\beta, p}} = \{s^\alpha, \bar{h}_{\beta, p}^{\text{DS}}\}_1 = \left(p + \frac{1}{2} + (1 - \delta_{\beta, n}) \frac{2\beta - n}{2n - 2} \right)^{-1} \{s^\alpha, \bar{h}_{\beta, p-1}^{\text{DS}}\}_2$$

These Hamiltonian flows are tau symmetric in the sense that

$$\frac{\partial h_{\alpha, p-1}^{\text{DS}}}{\partial t_q^\beta} = \frac{\partial h_{\beta, q-1}^{\text{DS}}}{\partial t_p^\alpha}.$$

The content described in the present section is all we need to actually compute the first Hamiltonian structure of the DS hierarchy of type D_4 in view of confirming the strong DR/DZ conjecture. This is the goal of the next and last section.

5.3.2 First Hamiltonian structure of the DS hierarchy of type D_4 and strong DR/DZ equivalence

The explicit tau-symmetric Hamiltonian densities.

In this section, we give the explicit formulae for the Hamiltonian densities $h_{\alpha, -1}^{\text{DS}}, h_{\alpha, 0}^{\text{DS}}$, for $1 \leq \alpha \leq 4$, and $h_{1, 1}^{\text{DS}}$. From now, in order to compare the DZ hierarchy with the DR hierarchy, we perform the following substitution:

$$\frac{\partial^k f}{\partial x^k} \mapsto \left(\frac{\varepsilon}{\sqrt{2}} \right)^k \frac{\partial^k f}{\partial x^k}$$

Therefore, we will now work again in the extended spaces. When we reduce to the case of the DS hierarchy of type D_4 , the Lax operator becomes

$$L = \partial_x^6 + \partial_x^{-1} \sum_{\mu=1}^3 \left(s^\mu \partial_x^{2\mu-1} + \partial_x^{2\mu-1} s^\mu \right) + \partial_x^{-1} \varrho \partial_x^{-1} \varrho, \quad (5.69)$$

Recall that the normal coordinates are given by

$$\tilde{u}^\alpha = \eta^{\alpha\beta} h_{\beta, -1}^{\text{DS}}$$

where the matrix $(\eta^{\alpha\beta})$ is given by (see e.g. [LRZ15])

$$(\eta^{\alpha\beta}) = \begin{pmatrix} 0 & 0 & 6 & 0 \\ 0 & 6 & 0 & 0 \\ 6 & 0 & 0 & 0 \\ 0 & 0 & 0 & 2 \end{pmatrix}$$

By Equation (5.68), the normal coordinates read, for $\mu \in \{1, 2, 3\}$,

$$\tilde{u}^\mu = \frac{6}{7-2\mu} \operatorname{res} P^{7-2\mu}, \quad \tilde{u}^4 = 2 \operatorname{res} Q.$$

Because of the form of the operator Q (5.62), its residue is straightforwardly $\operatorname{res} Q = \varrho$, which means that $\tilde{u}^4 = 2\varrho = 2\sqrt{s^4}$. For the remaining normal coordinates, we compute the residues in the coordinates s^α and then inverse the system. We find

$$\begin{cases} s^1 = \frac{1}{2}\tilde{u}^1 + \frac{1}{12}\tilde{u}^2\tilde{u}^3 + \frac{1}{216}(\tilde{u}^3)^3 + \left(-\frac{1}{8}(\tilde{u}^3)^2 - \frac{1}{6}\tilde{u}_2^2 - \frac{1}{9}\tilde{u}^3\tilde{u}_2^3\right)\varepsilon^2 + \frac{23}{90}\tilde{u}_4^3\varepsilon^4, \\ s^2 = \frac{1}{2}\tilde{u}^2 + \frac{1}{8}(\tilde{u}^3)^2 - \frac{1}{2}\tilde{u}_2^3\varepsilon^2, \\ s^3 = \frac{1}{2}\tilde{u}^3, \\ s^4 = \frac{1}{4}(\tilde{u}^4)^2. \end{cases} \quad (5.70)$$

This agrees with the expressions found in [LRZ15], p. 751. Then the Hamiltonian densities $h_{\alpha,0}^{\text{DS}}$ are given by: for $\mu \in \{1, 2, 3\}$,

$$h_{\mu,0}^{\text{DS}} = \frac{6}{(2\mu-1)(2\mu+5)} \operatorname{res} P^{2\mu+5}, \quad h_{4,0}^{\text{DS}} = \frac{2}{3} \operatorname{res} Q^3$$

The computation of $\operatorname{res} P^{2\mu+5}$ can be done straightforwardly by programming. However, the computation $\operatorname{res} Q^3$ requires a little bit more. To do so, we write

$$\begin{aligned} \operatorname{res} Q^3 &= \operatorname{res} QL = \operatorname{res} (\partial_x^{-1} \varrho L) + \operatorname{res} \left(\sum_{m \geq 0} Q_m \partial_x L \right), \\ &= \operatorname{res} (\partial_x^{-1} \varrho L) + \operatorname{res} \left(\sum_{m \geq 0} Q_m \varrho \partial_x^{-1} \varrho \right), \end{aligned}$$

where in the last equation we used the fact that $\partial_x L = L' = L'_+ + \varrho \partial_x^{-1} \varrho$. The term $\operatorname{res} (\partial_x^{-1} \varrho L)$ can also be easily computed by programming. For the remaining part let us first write $Q_m = \sum_{k \leq 2m} q^{m,k} \partial_x^k$, then

$$\begin{aligned} \operatorname{res} \left(\sum_{m \geq 0} Q_m \varrho \partial_x^{-1} \varrho \right) &= \operatorname{res} \left(\sum_{m \geq 0} \sum_{k \leq 2m} q^{m,k} \partial_x^k \varrho \partial_x^{-1} \varrho \right) \\ &= \operatorname{res} \left(\sum_{m \geq 0} \sum_{k \leq 2m} q^{m,k} \sum_{\ell=0}^k \binom{k}{\ell} \left(\frac{\varepsilon}{\sqrt{2}} \right)^\ell \varrho \partial_x^{k-1-\ell} \varrho \right) \\ &= \sum_{m \geq 0} \sum_{k \leq 2m} q^{m,k} \left(\frac{\varepsilon}{\sqrt{2}} \right)^k \varrho_k \varrho = \varrho \sum_{m \geq 0} Q_m(\varrho). \end{aligned}$$

In the above equations, all matters of convergence are resolved by the grading of $\widehat{\mathcal{A}}$ and the fact that $Q \in \mathcal{D}^+$. Now thanks to Equation (5.63), it follows that

$$\sum_{m \geq 0} Q_m(\varrho) = \frac{1}{2} L_+(1) = \sum_{\mu=1}^3 s_{2\mu-2}^\mu \frac{\varepsilon^{2\mu-2}}{2^{\mu-1}}.$$

Using $\tilde{u}^4 = 2\varrho$, we finally find that

$$h_{4,0}^{\text{DS}} = \frac{1}{3} \text{res} (\partial_x^{-1} \tilde{u}^4 L) + \frac{1}{3} \tilde{u}^4 \sum_{\mu=1}^3 s_{2\mu-2}^\mu \frac{\varepsilon^{2\mu-2}}{2^{\mu-1}}.$$

The results below were obtained using the program Wolfram Mathematica. For these programmed computations, the estimates given in Propositions [3.1.7](#) and [3.1.8](#) are of particular interest.

Proposition 5.3.7. *The level 0 tau symmetric Hamiltonian densities $h_{1,0}^{\text{DS}}, \dots, h_{4,0}^{\text{DS}}$ of the DS hierarchy of type D_4 in normal coordinates are given by the following.*

$$\begin{aligned}
h_{1,0}^{\text{DS}} &= \left(\frac{1}{12} (\tilde{u}^2)^2 + \frac{1}{6} \tilde{u}^1 \tilde{u}^3 + \frac{1}{4} (\tilde{u}^4)^2 \right) + \left(\frac{1}{72} \tilde{u}^3 (\tilde{u}_1^3)^2 + \frac{1}{3} \tilde{u}_2^1 + \frac{1}{216} (\tilde{u}^3)^2 \tilde{u}_2^3 \right) \epsilon^2 + \left(\frac{1}{216} (\tilde{u}_2^3)^2 + \frac{1}{72} \tilde{u}_1^3 \tilde{u}_3^3 \right. \\
&\quad \left. + \frac{1}{180} \tilde{u}^3 \tilde{u}_4^3 \right) \epsilon^4 + \frac{1}{840} \tilde{u}_6^3 \epsilon^6 \\
h_{2,0}^{\text{DS}} &= \left(\frac{1}{6} \tilde{u}^1 \tilde{u}^2 - \frac{1}{72} (\tilde{u}^2)^2 \tilde{u}^3 + \frac{1}{648} \tilde{u}^2 (\tilde{u}^3)^3 + \frac{1}{24} \tilde{u}^3 (\tilde{u}^4)^2 \right) + \left(-\frac{1}{24} (\tilde{u}_1^2)^2 + \frac{1}{24} \tilde{u}^3 \tilde{u}_1^2 \tilde{u}_1^3 + \frac{1}{72} \tilde{u}^2 (\tilde{u}_1^3)^2 \right. \\
&\quad \left. + \frac{1}{8} (\tilde{u}_1^4)^2 - \frac{1}{18} \tilde{u}^2 \tilde{u}_2^2 + \frac{1}{72} (\tilde{u}^3)^2 \tilde{u}_2^2 + \frac{1}{54} \tilde{u}^2 \tilde{u}^3 \tilde{u}_2^3 + \frac{1}{6} \tilde{u}^4 \tilde{u}_2^4 \right) \epsilon^2 + \left(\frac{1}{12} \tilde{u}_2^2 \tilde{u}_2^3 + \frac{5}{72} \tilde{u}_1^3 \tilde{u}_3^3 + \frac{1}{18} \tilde{u}_1^2 \tilde{u}_3^3 \right. \\
&\quad \left. + \frac{1}{36} \tilde{u}^3 \tilde{u}_4^2 + \frac{2}{135} \tilde{u}^2 \tilde{u}_4^3 \right) \epsilon^4 + \frac{1}{72} \tilde{u}_6^2 \epsilon^6 \\
h_{3,0}^{\text{DS}} &= \left(\frac{1}{12} (\tilde{u}^1)^2 - \frac{1}{216} (\tilde{u}^2)^3 + \frac{1}{432} (\tilde{u}^2)^2 (\tilde{u}^3)^2 + \frac{1}{233280} (\tilde{u}^3)^6 + \frac{1}{24} \tilde{u}^2 (\tilde{u}^4)^2 + \frac{1}{144} (\tilde{u}^3)^2 (\tilde{u}^4)^2 \right) + \\
&\quad \left(\frac{5}{432} \tilde{u}^3 (\tilde{u}_1^2)^2 + \frac{1}{108} \tilde{u}^3 \tilde{u}_1^1 \tilde{u}_1^3 + \frac{1}{36} \tilde{u}^2 \tilde{u}_1^2 \tilde{u}_1^3 + \frac{1}{1296} (\tilde{u}^3)^3 (\tilde{u}_1^3)^2 + \frac{1}{12} \tilde{u}^4 \tilde{u}_1^3 \tilde{u}_1^4 + \frac{5}{144} \tilde{u}^3 (\tilde{u}_1^4)^2 \right. \\
&\quad \left. + \frac{1}{216} (\tilde{u}^3)^2 \tilde{u}_2^1 + \frac{1}{54} \tilde{u}^2 \tilde{u}^3 \tilde{u}_2^2 + \frac{1}{108} (\tilde{u}^2)^2 \tilde{u}_2^3 + \frac{1}{3888} (\tilde{u}^3)^4 \tilde{u}_2^3 + \frac{1}{36} (\tilde{u}^4)^2 \tilde{u}_2^3 + \frac{1}{18} \tilde{u}^3 \tilde{u}^4 \tilde{u}_2^4 \right) \epsilon^2 \\
&\quad + \left(\frac{5}{2592} (\tilde{u}_1^3)^4 + \frac{13}{720} (\tilde{u}_2^2)^2 + \frac{55}{3888} \tilde{u}^3 (\tilde{u}_1^3)^2 \tilde{u}_2^3 + \frac{29}{1080} \tilde{u}_2^1 \tilde{u}_2^3 + \frac{67}{15552} (\tilde{u}^3)^2 (\tilde{u}_2^2)^2 + \frac{13}{240} (\tilde{u}_2^4)^2 \right. \\
&\quad \left. + \frac{1}{30} \tilde{u}_1^3 \tilde{u}_3^1 + \frac{31}{1080} \tilde{u}_1^2 \tilde{u}_3^2 + \frac{11}{1080} \tilde{u}_1^1 \tilde{u}_3^3 + \frac{49}{7776} (\tilde{u}^3)^2 \tilde{u}_1^3 \tilde{u}_3^3 + \frac{31}{360} \tilde{u}_1^4 \tilde{u}_3^4 + \frac{1}{60} \tilde{u}^3 \tilde{u}_4^1 + \frac{2}{135} \tilde{u}^2 \tilde{u}_4^2 \right. \\
&\quad \left. + \frac{1}{1296} (\tilde{u}^3)^3 \tilde{u}_4^3 + \frac{2}{45} \tilde{u}^4 \tilde{u}_4^4 \right) \epsilon^4 + \left(\frac{1129}{116640} (\tilde{u}_2^3)^3 + \frac{1601}{38880} \tilde{u}_1^3 \tilde{u}_2^3 \tilde{u}_3^3 + \frac{1}{120} \tilde{u}^3 (\tilde{u}_3^3)^2 + \frac{19}{1620} (\tilde{u}_1^3)^2 \tilde{u}_4^3 \right. \\
&\quad \left. + \frac{29}{2160} \tilde{u}^3 \tilde{u}_2^3 \tilde{u}_4^3 + \frac{13}{1944} \tilde{u}^3 \tilde{u}_1^3 \tilde{u}_5^3 + \frac{11}{1080} \tilde{u}_6^1 + \frac{17}{19440} (\tilde{u}^3)^2 \tilde{u}_6^3 \right) \epsilon^6 + \left(\frac{191}{43200} (\tilde{u}_4^3)^2 + \frac{1501}{194400} \tilde{u}_3^3 \tilde{u}_5^3 \right. \\
&\quad \left. + \frac{949}{194400} \tilde{u}_2^3 \tilde{u}_6^3 + \frac{127}{64800} \tilde{u}_1^3 \tilde{u}_7^3 + \frac{13}{32400} \tilde{u}^3 \tilde{u}_8^3 \right) \epsilon^8 + \frac{7}{118800} \tilde{u}_{10}^3 \epsilon^{10} \\
h_{4,0}^{\text{DS}} &= \left(\frac{1}{2} \tilde{u}^1 \tilde{u}^4 + \frac{1}{12} \tilde{u}^2 \tilde{u}^3 \tilde{u}^4 + \frac{1}{216} (\tilde{u}^3)^3 \tilde{u}^4 \right) + \left(\frac{1}{24} \tilde{u}^4 (\tilde{u}_1^3)^2 + \frac{1}{4} \tilde{u}_1^2 \tilde{u}_1^4 + \frac{1}{8} \tilde{u}^3 \tilde{u}_1^3 \tilde{u}_1^4 + \frac{1}{6} \tilde{u}^4 \tilde{u}_2^2 \right. \\
&\quad \left. + \frac{1}{18} \tilde{u}^3 \tilde{u}^4 \tilde{u}_2^3 + \frac{1}{6} \tilde{u}^2 \tilde{u}_2^4 + \frac{1}{24} (\tilde{u}^3)^2 \tilde{u}_2^4 \right) \epsilon^2 + \left(\frac{1}{4} \tilde{u}_2^3 \tilde{u}_2^4 + \frac{1}{6} \tilde{u}_1^4 \tilde{u}_3^3 + \frac{5}{24} \tilde{u}_1^3 \tilde{u}_3^4 + \frac{2}{45} \tilde{u}^4 \tilde{u}_4^3 + \frac{1}{12} \tilde{u}^3 \tilde{u}_4^4 \right) \epsilon^4 \\
&\quad + \frac{1}{24} \tilde{u}_6^4 \epsilon^6
\end{aligned}$$

Finally, the density $h_{1,1}^{\text{DS}}$ is given by

$$h_{1,1}^{\text{DS}} = \frac{36}{91} \text{res } P^{13}.$$

Again, the result was obtained using the program Wolfram Mathematica.

Proposition 5.3.8. *The Hamiltonian density $h_{1,1}^{\text{DS}}$ of the DS hierarchy of type D_4 in normal coordinates is given by the following.*

$$\begin{aligned}
h_{1,1}^{\text{DS}} = & \left(\frac{1}{12} u^1 (u^2)^2 + \frac{1}{12} (u^1)^2 u^3 - \frac{1}{108} (u^2)^3 u^3 + \frac{1}{432} (u^2)^2 (u^3)^3 + \frac{1}{326592} (u^3)^7 + \frac{1}{4} u^1 (u^4)^2 \right. \\
& + \frac{1}{12} u^2 u^3 (u^4)^2 + \frac{1}{144} (u^3)^3 (u^4)^2 \left. \right) + \left(\frac{1}{6} (u^1)^2 + \frac{1}{3} u_2^1 u^1 - \frac{5}{72} (u_2^1)^2 u^2 + \frac{5}{24} (u_1^4)^2 u^2 - \frac{1}{18} u_2^2 (u^2)^2 \right. \\
& + \frac{1}{54} (u_1^3)^2 (u^2)^2 + \frac{1}{72} (u_1^3)^2 u^1 u^3 + \frac{19}{216} u_1^2 u_1^3 u^2 u^3 + \frac{1}{36} u_2^3 (u^2)^2 u^3 + \frac{1}{54} (u_1^2)^2 (u^3)^2 + \frac{1}{72} u_1^1 u_1^3 (u^3)^2 \\
& + \frac{1}{18} (u_1^4)^2 (u^3)^2 + \frac{1}{216} u_2^3 u^1 (u^3)^2 + \frac{7}{216} u_2^2 u^2 (u^3)^2 + \frac{1}{216} u_2^1 (u^3)^3 + \frac{7}{7776} (u_1^3)^2 (u^3)^4 \\
& + \frac{1}{3888} u_2^3 (u^3)^5 + \frac{5}{12} u_1^2 u_1^4 u^4 + \frac{1}{3} u_2^4 u^2 u^4 + \frac{19}{72} u_1^3 u_1^4 u^3 u^4 + \frac{7}{72} u_2^4 (u^3)^2 u^4 + \frac{1}{6} u_2^2 (u^4)^2 \\
& + \frac{1}{18} (u_1^3)^2 (u^4)^2 + \frac{1}{12} u_2^3 u^3 (u^4)^2 \left. \right) \varepsilon^2 + \left(\frac{7}{24} u_1^2 u_2^2 u_1^3 + \frac{5}{72} u_2^1 (u_1^3)^2 + \frac{31}{216} (u_1^2)^2 u_2^3 + \frac{13}{216} u_1^1 u_1^3 u_2^3 \right. \\
& + \frac{31}{72} u_2^3 (u_1^4)^2 + \frac{7}{8} u_1^3 u_1^4 u_2^4 + \frac{1}{216} (u_2^3)^2 u^1 + \frac{1}{72} u_1^3 u_3^3 u^1 + \frac{35}{216} u_3^2 u_1^3 u^2 + \frac{23}{108} u_2^2 u_2^3 u^2 + \frac{4}{27} u_1^2 u_3^3 u^2 \\
& + \frac{7}{270} u_4^3 (u^2)^2 + \frac{13}{144} (u_2^2)^2 u^3 + \frac{29}{216} u_1^2 u_3^2 u^3 + \frac{1}{9} u_1^3 u_3^3 u^3 + \frac{7}{648} (u_1^3)^4 u^3 + \frac{19}{216} u_2^1 u_2^3 u^3 + \frac{7}{216} u_1^1 u_3^3 u^3 \\
& + \frac{13}{48} (u_2^4)^2 u^3 + \frac{29}{72} u_1^4 u_3^4 u^3 + \frac{1}{180} u_4^3 u^1 u^3 + \frac{7}{108} u_4^2 u^2 u^3 + \frac{1}{36} u_4^1 (u^3)^2 + \frac{53}{1296} (u_1^3)^2 u_2^3 (u^3)^2 \\
& + \frac{133}{15552} (u_2^3)^2 (u^3)^3 + \frac{97}{7776} u_1^3 u_3^3 (u^3)^3 + \frac{1}{864} u_4^3 (u^3)^4 + \frac{4}{9} u_3^3 u_1^4 u^4 + \frac{23}{36} u_2^3 u_2^4 u^4 + \frac{35}{72} u_1^3 u_3^4 u^4 \\
& + \frac{7}{36} u_4^4 u^3 u^4 + \frac{7}{90} u_4^3 (u^4)^2 \left. \right) \varepsilon^4 + \left(\frac{349}{3024} (u_2^3)^2 + \frac{4}{21} u_2^2 u_4^2 + \frac{17}{168} u_1^2 u_5^2 + \frac{65}{504} u_5^1 u_1^3 + \frac{151}{756} u_4^1 u_2^3 \right. \\
& + \frac{473}{2592} (u_1^3)^2 (u_2^3)^2 + \frac{31}{168} u_3^1 u_3^3 + \frac{19}{216} (u_1^3)^3 u_3^3 + \frac{19}{210} u_2^2 u_4^3 + \frac{149}{7560} u_1^1 u_5^3 + \frac{349}{1008} (u_3^4)^2 + \frac{4}{7} u_2^4 u_4^4 \\
& + \frac{17}{56} u_1^4 u_5^4 + \frac{1}{840} u_6^3 u^1 + \frac{53}{1512} u_6^2 u^2 + \frac{65}{1512} u_6^1 u^3 + \frac{1741}{23328} (u_2^3)^3 u^3 + \frac{2507}{7776} u_1^3 u_2^3 u_3^3 u^3 + \frac{593}{6480} (u_1^3)^2 u_4^3 u^3 \\
& + \frac{271}{7776} (u_3^2)^2 (u^3)^2 + \frac{2141}{38880} u_2^3 u_4^3 (u^3)^2 + \frac{341}{12960} u_3^3 u_5^3 (u^3)^2 + \frac{43}{19440} u_6^3 (u^3)^3 + \frac{53}{504} u_6^4 u^4 \left. \right) \varepsilon^6 + \left(\frac{u_8^1}{54} \right. \\
& + \frac{653}{1944} u_2^3 (u_3^3)^2 + \frac{17803}{68040} (u_2^3)^2 u_4^3 + \frac{11129}{30240} u_1^3 u_3^3 u_4^3 + \frac{65141}{272160} u_1^3 u_2^3 u_5^3 + \frac{145}{3024} (u_1^3)^2 u_6^3 + \frac{17503}{302400} (u_4^3)^2 u^3 \\
& + \frac{977}{10080} u_3^3 u_5^3 u^3 + \frac{15103}{272160} u_2^3 u_6^3 u^3 + \frac{1831}{90720} u_1^3 u_7^3 u^3 + \frac{19}{9720} u_8^3 (u^3)^2 \left. \right) \varepsilon^8 + \left(\frac{9973}{340200} (u_5^3)^2 + \frac{1301}{25200} u_4^3 u_6^3 \right. \\
& \left. + \frac{347}{10080} u_3^3 u_7^3 + \frac{4427}{272160} u_2^3 u_8^3 + \frac{89}{18144} u_1^3 u_9^3 + \frac{1}{1296} u_{10}^3 u^3 \right) \varepsilon^{10} + \frac{41}{393120} u_{12}^3 \varepsilon^{12}
\end{aligned}$$

The explicit form of the first Poisson structure.

Finally, we compute the components of the first Poisson structure given, for two local functionals $\bar{f}, \bar{g} \in \widehat{\mathcal{F}}$, by

$$\{\bar{f}, \bar{g}\}_1 = \frac{1}{\varepsilon} \int \text{res } X [(\partial_x Y_+ L')_- - (L' Y_+ \partial_x)_- - (\partial_x Y_- L')_+ + (L' Y_- \partial_x)_+] dx, \quad (5.71)$$

where $X = \frac{\delta f}{\delta L'}$, $Y = \frac{\delta g}{\delta L'}$. We aim to express this bracket as bi-vector using Dirac's notation of the components:

$$\{\bar{f}, \bar{g}\}_1 = \iint \frac{\delta \bar{f}(x)}{\delta \tilde{u}^\alpha(x)} \{\tilde{u}^\alpha(x), \tilde{u}^\beta(y)\}_1 \frac{\delta \bar{g}(y)}{\delta \tilde{u}^\beta(y)} dx dy, \quad (5.72)$$

We will use the following identity to compute the Poisson structure (see Equation (5.16)):

$$\{\tilde{u}^\alpha(x), \tilde{u}^\beta(y)\}_1 = \{\tilde{u}^\alpha(z) \delta(x-z), \tilde{u}^\beta \delta(y-z)\}_1.$$

Now the Poisson bracket (5.71) is given in terms of the variational differentials, so that we first need to compute those of the coordinate functionals. To do so, we need to transform the normal coordinates \tilde{u}^α into the coordinates v^α of Equation (5.65). The coordinates \tilde{v}^α are related to the v^α 's by the condition $L^{*\prime} + L' = 0$; they read

$$\begin{cases} \tilde{v}^1 = \frac{1}{2\sqrt{2}}v_1^1\varepsilon - \frac{1}{8\sqrt{2}}v_3^2\varepsilon^3 + \frac{1}{8\sqrt{2}}v_5^3\varepsilon^5, \\ \tilde{v}^2 = \frac{3}{2\sqrt{2}}v_1^2\varepsilon - \frac{5}{4\sqrt{2}}v_3^3\varepsilon^3, \\ \tilde{v}^3 = \frac{5}{2\sqrt{2}}v_1^3\varepsilon. \end{cases} \quad (5.73)$$

Then by identifying the two operators in Equation (5.71) and inverting Equation (5.70), we find the following Miura from the normal coordinates \tilde{u}^α to the coordinates v^α :

$$\begin{cases} \tilde{u}^1 = v^1 - \frac{1}{6}v^2v^3 + \frac{7}{216}(v^3)^3 + \left(\frac{1}{12}(v_1^3)^2 - \frac{5}{12}v_2^2 + \frac{11}{36}v^3v_2^3\right)\varepsilon^2 + \frac{89}{90}v_4^3\varepsilon^4, \\ \tilde{u}^2 = v^2 - \frac{1}{4}(v^3)^2 - \frac{3}{2}v_2^3\varepsilon^2, \\ \tilde{u}^3 = v^3, \\ \tilde{u}^4 = 2\sqrt{v^4}. \end{cases} \quad (5.74)$$

Now recall that the variational differentials are given by

$$\frac{\delta f(z)}{\delta L'(z)} = \frac{\delta f(z)}{\delta v^4(z)} + \frac{1}{2} \sum_{\mu=1}^3 \left(\frac{\delta f(z)}{\delta v^\mu(z)} \partial_z^{-2\mu} + \partial_z^{-2\mu} \frac{\delta f(z)}{\delta v^\mu(z)} \right) \in \mathcal{D}^-.$$

It follows that the variational differentials of the coordinates functionals are given by, for $\mu \in \{1, 2, 3\}$,

$$\frac{\delta \tilde{u}^\mu(z)\delta(x-z)}{\delta L'(z)} = \frac{1}{2} \sum_{\nu=1}^3 \left(\frac{\delta \tilde{u}^\mu(z)\delta(x-z)}{\delta v^\nu(z)} \partial_z^{-2\nu} + \partial_z^{-2\nu} \frac{\delta \tilde{u}^\mu(z)\delta(x-z)}{\delta v^\nu(z)} \right), \quad (5.75)$$

$$\frac{\delta \tilde{u}^4(z)\delta(x-z)}{\delta L'(z)} = \frac{\delta \left(2\sqrt{v^4(z)}\delta(x-z) \right)}{\delta v^4(z)} = \frac{2}{u^4(z)}\delta(x-z). \quad (5.76)$$

Let us denote, for $\alpha \in \{1, 2, 3, 4\}$,

$$X^\alpha(z) := \frac{\delta \tilde{u}^\alpha(z)\delta(\mathbf{x}-z)}{\delta L'(z)}, \quad Y^\alpha(z) := \frac{\delta \tilde{u}^\alpha(z)\delta(\mathbf{y}-z)}{\delta L'(z)}$$

As we can see in Equations (5.75) and (5.76), for any $\mu \in \{1, 2, 3\}$, the pseudo-differential operators (of the first type) Y^μ and Y^4 satisfy

$$(Y^\mu(z))_+ = 0, \quad (Y^4(z))_- = 0.$$

Knowing this, it is therefore not hard to show that the component $\{\tilde{u}^4(x), \tilde{u}^4(y)\}_1$ of the first Poisson structure reads

$$\begin{aligned} \{\tilde{u}^4(x), \tilde{u}^4(y)\}_1 &= \varepsilon^{-1} \int \text{res } X^4(z) \left[(\partial_z Y^4(z) L'(z))_- - (L'(z) Y^4(z) \partial_z)_- \right] dz \\ &= \varepsilon^{-1} \int \text{res } \frac{2}{\tilde{u}^4(z)} \delta(x-z) \left[\left(\partial_z \frac{2}{\tilde{u}^4(z)} \delta(y-z) L'(z) \right)_- - \left(L'(z) \frac{2}{\tilde{u}^4(z)} \delta(y-z) \partial_z \right)_- \right] dz \\ &= 2\delta'(x-y) \end{aligned}$$

(Note that this precise computation does not depend on the integer n of D_n , meaning that similarly, in the D_n case, $\{\tilde{u}^n(x), \tilde{u}^n(y)\}_1 = 2\delta'(x-y)$.) It is easy to see that for any $\mu \in \{1, 2, 3\}$, $\{\tilde{u}^\mu(x), \tilde{u}^4(y)\}_1 = 0$. The remaining components are computed in a similar fashion, this time using

$$\{\tilde{u}^\mu(x), \tilde{u}^\nu(y)\}_1 = \varepsilon^{-1} \int \operatorname{res} X^\mu(z) [(L'(z)Y^\nu(z)\partial_z)_+ - (\partial_z Y^\nu(z)L'(z))_+] dz,$$

for $\mu, \nu \in \{1, 2, 3\}$. Now we equivalently write ∂_x^k instead of $\delta^{(k)}(x-y)$, for their action is identical. We write down the components $\{\tilde{u}^\alpha(x), \tilde{u}^\beta(y)\}_1$ in the matrix

$$K_{\text{DS}}^{\alpha\beta} := \{\tilde{u}^\alpha(x), \tilde{u}^\beta(y)\}_1.$$

Proposition 5.3.9. *The first local Poisson bracket of the DS hierarchy of type D_4 , in normal coordinates reads*

$$(K_{\text{DS}}^{\alpha\beta}) = \begin{pmatrix} (\frac{1}{6}\tilde{u}_2^3\partial_x + \frac{1}{2}\tilde{u}_1^3\partial_x^2 + \frac{1}{3}\tilde{u}^3\partial_x^3)\varepsilon^2 + \frac{4}{15}\partial_x^5\varepsilon^4 & 0 & 6\partial_x & 0 \\ 0 & 6\partial_x & 0 & 0 \\ 6\partial_x & 0 & 0 & 0 \\ 0 & 0 & 0 & 2\partial_x \end{pmatrix}. \quad (5.77)$$

In particular, $K_{\text{DS}}^{\alpha\beta} = \eta^{\alpha\beta}\partial_x + \mathcal{O}(\varepsilon)$.

Remark 5.3.10. Let $\psi : \tilde{u}^\alpha \mapsto v^\alpha$ be the inverse of the Miura transform given by Equation (5.74). Then one can also compute the matrix $K_{\text{DS}}^{\alpha\beta}$ in normal coordinates using Equation (5.18) which we recall here:

$$K_{\text{DS}}^{\alpha\beta} = (M^*)_\mu^\alpha \circ (K_{\text{DS}} \circ \psi)^{\mu\nu} \circ M_\nu^\beta, \quad (5.78)$$

where

$$(M^*)_\mu^\alpha = \sum_{s \geq 0} \frac{\partial \tilde{u}^\alpha}{\partial v_s^\mu} \partial_x^s, \quad M_\nu^\beta = \sum_{s \geq 0} (-1)^s \partial_x^s \circ \frac{\partial \tilde{u}^\beta}{\partial v_s^\nu},$$

and where, by construction, $(K_{\text{DS}} \circ \psi)^{\alpha\beta} = \{v^\alpha(x), v^\beta(y)\}_1$. But then we need to compute the bracket in the coordinates v^α and then use Equation (5.78). It is probably a bit easier in all honesty.

To conclude, comparing (1) the Poisson structures of the DR hierarchy and the DS hierarchy of type D_4 , both in their normal coordinates (Propositions 5.2.25 and 5.3.9 respectively) and (2) the Hamiltonian functional $\bar{h}_{1,1}^{\text{DR}}$ of the DR hierarchy with the Hamiltonian functional $\bar{h}_{1,1}^{\text{DS}}$ of the DS hierarchy (Propositions 5.2.27 and 5.3.8 respectively), we obtain the theorem below. Notice that it holds true in virtue of Theorem 5.2.24.

Theorem 5.3.11 (Strong DR/DZ for D_4 [DR19]). *The strong DR/DZ equivalence holds true for the hierarchy of type D_4 under the normal Miura transform (5.45). More precisely, in the normal coordinates \tilde{u}^α ,*

(i) *The local Poisson structures are the same: $K_{\text{DR}}^{\alpha\beta} = K_{\text{DS}}^{\alpha\beta}$;*

(ii) *The Hamiltonian densities $h_{1,1}^{\text{DR}}$ and $h_{1,1}^{\text{DS}}$ differ by a total x -derivative, i.e. $\bar{h}_{1,1}^{\text{DR}} = \bar{h}_{1,1}^{\text{DS}}$.*

Moreover, the level 0 Hamiltonian densities coincide: $h_{\alpha,0}^{\text{DR}} = h_{\alpha,0}^{\text{DS}}$ for $\alpha \in \{1, \dots, 4\}$.

We recall that, specifically for the hierarchies of type ADE, say of type X_n , the DS hierarchy constructed from the untwisted affine algebra $X_n^{(1)}$ coincides with the DZ hierarchy constructed from the FJRW of the simple singularity of type X_n ; see e.g. [LRZ15]. Besides, we recall that two local functionals coincide if and only if all the variational derivatives $\delta/\delta\tilde{u}^\alpha$ of their difference vanish; see Proposition 5.1.1.

Appendix A

Matrix realizations of classical Kac–Moody algebras

In this appendix, given a $n \times n$ matrix A , we denote by A^{at} the anti-transpose of A , that is, the transpose of A with respect to the antidiagonal (or secondary diagonal). It is given by the formula

$$(A^{\text{at}})_{i,j} = A_{n-j+1, n-i+1}.$$

Besides, in accordance with previous notations, we denote by $E_{i,j}$ the matrix with 1 at the intersection of row i and column j and 0 elsewhere.

When we say that the Lie algebra with (generalized) Cartan matrix A has Weyl generators $\{X_i, Y_i, H_i \mid 1 \leq i \leq r\}$, we mean that they satisfy the relations

$$[X_i, Y_i] = \delta_{i,j} H_i, \quad [H_i, X_i] = A_{i,j} X_i, \quad [H_i, Y_i] = -A_{i,j} Y_i.$$

A.1 Classical simple Lie algebras

Here we give the main features and our chosen matrix realization of the classical simple Lie algebras. By that we mean the algebras whose Dynkin diagram is either of type A, B, C or D . The semisimple Lie algebras correspond exactly to the Kac–Moody algebras of finite type, hence the title of this appendix.

Type $A_n, n \geq 1$

The simple Lie algebras of type A_n are isomorphic to $\mathfrak{sl}(n+1, \mathbb{C})$. They have the following features.

Coxeter number: $h = n + 1$.

Highest root: $\theta = \alpha_1 + \alpha_2 + \cdots + \alpha_n$.

Finite exponents: $\mathbb{F} = \{1, 2, \dots, n\}$.

Dynkin diagram and finite Cartan matrix of size $n \times n$:

$$A_n : \bullet \text{---} \bullet \text{---} \bullet \text{---} \bullet \text{---} \bullet \quad \text{CM}(A_n) = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & & 0 \\ \vdots & \ddots & & 2 & -1 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

Irreducible faithful representation: $A_n \hookrightarrow \{X \in \mathfrak{gl}(n+1, \mathbb{C}) \mid \text{tr}X = 0\} = \mathfrak{sl}(n+1, \mathbb{C})$.
(The trivial one.) Weyl generators of the representation: For $i \in \{1, \dots, n\}$,

$$X_i = E_{i+1,i}, \quad Y_i = E_{i,i+1}, \quad H_i = E_{i+1,i+1} - E_{i,i}.$$

Type B_n , $n \geq 2$

The simple Lie algebras of type B_n are isomorphic to $\mathfrak{o}(2n+1, \mathbb{C})$. They have the following features.

Coxeter number: $h = 2n$.

Highest root: $\theta = \alpha_1 + 2\alpha_2 + \cdots + 2\alpha_n$.

Finite exponents: $\mathbb{F} = \{1, 3, \dots, 2n-1\}$.

Dynkin diagram and finite Cartan matrix of size $n \times n$:

$$B_n : \bullet \text{---} \bullet \text{---} \bullet \text{---} \bullet \text{---} \bullet \rightleftarrows \bullet \quad \text{CM}(B_n) = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & & 0 \\ \vdots & \ddots & -1 & 2 & -1 \\ 0 & \cdots & 0 & -2 & 2 \end{pmatrix}$$

Irreducible faithful representation: $B_n \hookrightarrow \{X \in \mathfrak{sl}(2n+1, \mathbb{C}) \mid XS + SX^{\text{at}} = 0\}$,
where $S = \text{diag}(1, -1, 1, \dots, -1, 1)$.

Weyl generators of the representation: For $i \in \{1, \dots, n-1\}$,

$$\begin{aligned} X_i &= E_{i+1,i} + E_{2n+2-i,2n+1-i}, & X_n &= E_{n+1,n} + E_{n+2,n+1}, \\ Y_i &= E_{i,i+1} + E_{2n+1-i,2n+2-i}, & Y_n &= 2(E_{n,n+1} + E_{n+1,n+2}), \\ H_i &= E_{i+1,i+1} - E_{i,i} + E_{2n+2-i,2n+2-i} - E_{2n+1-i,2n+1-i}, & H_n &= 2(E_{n+2,n+2} - E_{n,n}). \end{aligned}$$

Type C_n , $n \geq 2$

The simple Lie algebras of type C_n are isomorphic to $\mathfrak{sp}(2n)$. They have the following features.

Coxeter number: $h = 2n$.

Highest root: $\theta = \alpha_1 + 2\alpha_2 + \cdots + 2\alpha_{n-1} + \alpha_n$.

Finite exponents: $\mathbb{F} = \{1, 3, \dots, 2n - 1\}$.

Dynkin diagram and finite Cartan matrix of size $n \times n$:

$$C_n : \bullet \cdots \bullet \cdots \bullet \leftarrow \bullet$$

$$\text{CM}(C_n) = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -1 & 2 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & -1 & 0 \\ \vdots & \ddots & & 2 & -2 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

Irreducible faithful representation: $C_n \hookrightarrow \{X \in \mathfrak{sl}(2n) \mid XS + SX^{\text{at}} = 0\}$,

where $S = \text{diag}(1, -1, 1, \dots, 1, -1)$.

Weyl generators of the representation:

$$\begin{aligned} X_i &= E_{i+1,i} + E_{2n+1-i,2n-i}, & X_n &= E_{n+1,n}, \\ Y_i &= E_{i,i+1} + E_{2n-i,2n+1-i}, & Y_n &= E_{n,n+1}, \\ H_i &= E_{i+1,i+1} - E_{i,i} + E_{2n+1-i,2n+1-i} - E_{2n-i,2n-i}, & H_n &= E_{n+1,n+1} - E_{n,n} \end{aligned}$$

Type D_n , $n \geq 2$

The simple Lie algebras of type D_n are isomorphic to $\mathfrak{o}(2n)$. They have the following features.

Coxeter number: $h = 2n - 2$.

Highest root: $\theta = \alpha_1 + 2\alpha_2 + \cdots + 2\alpha_{n-2} + \alpha_{n-1} + \alpha_n$.

Finite exponents:

- ◊ If n is odd: $\mathbb{F} = \{1, 3, \dots, 2n - 3\} \cup \{n - 1\}$;
- ◊ If n is even: $\mathbb{F} = \{1, 3, \dots, 2n - 3\} \cup \{(n - 1)'\}$.

(We mean that when n is even, exponent $n - 1$ has multiplicity 2.)

Dynkin diagram and finite Cartan matrix of size $n \times n$:

$$D_n : \bullet \cdots \bullet \begin{array}{l} \nearrow \bullet \\ \searrow \bullet \end{array}$$

$$\text{CM}(D_n) = \begin{pmatrix} 2 & -1 & 0 & 0 & \cdots & 0 \\ -1 & 2 & \ddots & & \ddots & \vdots \\ 0 & \ddots & \ddots & & & 0 \\ 0 & & & 2 & -1 & -1 \\ \vdots & \ddots & & -1 & 2 & 0 \\ 0 & \cdots & 0 & -1 & 0 & 2 \end{pmatrix}$$

Irreducible faithful representation: $D_n \hookrightarrow \{X \in \mathfrak{sl}(2n) \mid XS + SX^{\text{at}} = 0\}$,

where $S = \text{diag}(1, -1, \dots, (-1)^{n-2}, (-1)^{n-1}, (-1)^{n-1}, (-1)^n, \dots, -1, 1)$.

Weyl generators of the representation:

$$\begin{aligned} X_i &= E_{i+1,i} + E_{2n+1-i,2n-i}, & X_n &= \frac{1}{2}(E_{n+1,n-1} + E_{n+2,n}), \\ Y_i &= E_{i,i+1} + E_{2n-i,2n+1-i}, & Y_n &= 2(E_{n-1,n+1} + E_{n,n+2}), \\ H_i &= E_{i+1,i+1} - E_{i,i} + E_{2n+1-i,2n+1-i} - E_{2n-i,2n-i}, & H_n &= E_{n+1,n+1} - E_{n,n} + E_{n+2,n+2} - E_{n-1,n-1}. \end{aligned}$$

Remark A.1.1. For the D_2 case, the diagram is disconnected as it reads $D_2 : \bullet \bullet$. Therefore, $D_2 \simeq A_1 \oplus A_1$ is semisimple rather than simple. However, the above representation still hold (see [DS84]).

A.1.1 Examples: Rank 2 simple Lie algebras

Type A_2

The simple Lie algebras of type A_2 are isomorphic to $\mathfrak{sl}(3, \mathbb{C})$. They have the following features.

Coxeter number: $h = 3$.

Finite exponents: $\mathbb{F} = \{1, 2\}$.

Dynkin diagram and finite Cartan matrix of size 2×2 :

$$A_2 : \bullet \longleftrightarrow \bullet \qquad \text{CM}(A_2) = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

Weyl generators of our realization:

$$\begin{aligned} X_1 &= \begin{pmatrix} 0 & & \\ 1 & 0 & \\ & & 0 \end{pmatrix} & X_2 &= \begin{pmatrix} 0 & & \\ & 0 & \\ & 1 & 0 \end{pmatrix} \\ Y_1 &= \begin{pmatrix} 0 & 1 & \\ & 0 & \\ & & 0 \end{pmatrix} & Y_2 &= \begin{pmatrix} 0 & & \\ & 0 & 1 \\ & & 0 \end{pmatrix} \\ H_1 &= \begin{pmatrix} -1 & & \\ & 1 & \\ & & 0 \end{pmatrix} & H_2 &= \begin{pmatrix} 0 & & \\ & -1 & \\ & & 1 \end{pmatrix} \end{aligned}$$

Type B_2

The simple Lie algebras of type B_2 are isomorphic to $\mathfrak{so}(5, \mathbb{C})$. They have the following features.

Coxeter number: $h = 4$.

Finite exponents: $\mathbb{F} = \{1, 3\}$.

Dynkin diagram and finite Cartan matrix of size 2×2 :

$$B_2 : \bullet \rightrightarrows \bullet \qquad \text{CM}(B_2) = \begin{pmatrix} 2 & -1 \\ -2 & 2 \end{pmatrix}$$

Weyl generators of our realization:

$$\begin{aligned}
 X_1 &= \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & & 0 & \\ & & & 1 & 0 \\ & & & & & 0 \end{pmatrix} & X_2 &= \begin{pmatrix} 0 & & & \\ & 0 & & \\ 1 & 0 & & \\ & & 1 & 0 \\ & & & & 0 \end{pmatrix} \\
 Y_1 &= \begin{pmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & \\ & & & 0 & 1 \\ & & & & & 0 \end{pmatrix} & Y_2 &= \begin{pmatrix} 0 & & & \\ & 0 & 2 & \\ & & 0 & 2 \\ & & & 0 \\ & & & & 0 \end{pmatrix} \\
 H_1 &= \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & 0 & \\ & & & -1 & \\ & & & & & 1 \end{pmatrix} & H_2 &= \begin{pmatrix} 0 & & & \\ & -2 & & \\ & & 0 & \\ & & & 2 \\ & & & & 0 \end{pmatrix}
 \end{aligned}$$

Type C_2

The simple Lie algebras of type C_2 are isomorphic to $\mathfrak{sp}(4, \mathbb{C})$. They have the following features.

Coxeter number: $h = 4$.

Finite exponents: $\mathbb{F} = \{1, 3\}$.

Dynkin diagram and finite Cartan matrix of size 2×2 :

$$C_2 : \bullet \rightleftarrows \bullet$$

$$CM(C_2) = \begin{pmatrix} 2 & -2 \\ -1 & 2 \end{pmatrix}$$

Weyl generators of our realization:

$$\begin{aligned}
 X_1 &= \begin{pmatrix} 0 & & & \\ 1 & 0 & & \\ & & 0 & \\ & & & 1 & 0 \\ & & & & & 0 \end{pmatrix} & X_2 &= \begin{pmatrix} 0 & & & \\ & 0 & & \\ 1 & 0 & & \\ & & 1 & 0 \\ & & & & 0 \end{pmatrix} \\
 Y_1 &= \begin{pmatrix} 0 & 1 & & \\ & 0 & & \\ & & 0 & 1 \\ & & & & 0 \end{pmatrix} & Y_2 &= \begin{pmatrix} 0 & & & \\ & 0 & 1 & \\ & & 0 & \\ & & & 0 \end{pmatrix} \\
 H_1 &= \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & -1 & \\ & & & & 1 \end{pmatrix} & H_2 &= \begin{pmatrix} 0 & & & \\ & -1 & & \\ & & 1 & \\ & & & & 0 \end{pmatrix}
 \end{aligned}$$

Type D_2

The simple Lie algebras of type D_2 are isomorphic to $\mathfrak{so}(4, \mathbb{C})$. They have the following features.

Coxeter number: $h = 2$.

Faithful representation: Loop algebra of $A_n^{(1)} \hookrightarrow \mathfrak{sl}(n+1, \mathbb{C}) \otimes \mathbb{C}[\lambda, \lambda^{-1}]$.

Weyl generators of the representation: For $i \in \{1, \dots, n\}$ (the $\otimes 1$ is implicit),

$$e_i = E_{i+1,i}, \quad f_i = E_{i,i+1}, \quad h_i = E_{i+1,i+1} - e_{i,i}.$$

To which we add the highest root triplet:

$$e_0 = E_{1,n} \otimes \lambda, \quad f_0 = E_{n,1} \otimes \lambda^{-1}, \quad h_0 = E_{1,1} - E_{n,n}.$$

Normalization constant of the standard bilinear form: $\kappa = 1$, i.e., $(X, Y) = \text{tr}(\pi(X)\pi(Y))$.

Generators of the principal Heisenberg subalgebra: $\Lambda_j = \Lambda^j$, for $j \in \mathbb{Z} \setminus (n+1)\mathbb{Z}$.

Principal cyclic element: $\Lambda = e_0 + \dots + e_n$,

$$\Lambda = \begin{pmatrix} 0 & \cdots & 0 & \lambda \\ 1 & \ddots & & 0 \\ 0 & 1 & & \vdots \\ \vdots & \ddots & \ddots & \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Type $B_n^{(1)}$, $n \geq 2$

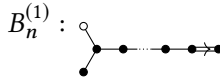
The affine Kac–Moody algebras of type $B_n^{(1)}$ have the following features.

Coxeter number: $h = 2n$.

Highest root: $\theta = \alpha_1 + 2\alpha_2 + \dots + 2\alpha_n$.

Affine exponents: $\mathbb{E} = \mathbb{Z}^{\text{odd}} = \mathbb{Z} \setminus 2\mathbb{Z}$.

Extended Dynkin diagram and affine generalized Cartan matrix of size $(n+1) \times (n+1)$:

$B_n^{(1)}$: 

$\text{GCM}(B_n^{(1)}) = \begin{pmatrix} 2 & 0 & -1 & 0 & \cdots & 0 \\ 0 & 2 & -1 & & \ddots & \vdots \\ -1 & -1 & 2 & \ddots & & \\ 0 & & \ddots & \ddots & & 0 \\ \vdots & \ddots & & & & 0 \\ & & & & -1 & 2 & -1 \\ 0 & \cdots & 0 & 0 & -2 & 2 \end{pmatrix}$

Faithful representation: Loop algebra of $B_n^{(1)} \hookrightarrow \mathfrak{o}(2n+1) \otimes \mathbb{C}[\lambda, \lambda^{-1}]$.

Weyl generators of the representation: For $i \in \{1, \dots, n-1\}$ (the $\otimes 1$ is implicit),

$$\begin{aligned} e_i &= E_{i+1,i} + E_{2n+2-i,2n+1-i}, & e_n &= E_{n+1,n} + E_{n+2,n+1}, \\ f_i &= E_{i,i+1} + E_{2n+1-i,2n+2-i}, & f_n &= 2(E_{n,n+1} + E_{n+1,n+2}), \\ h_i &= E_{i+1,i+1} - E_{i,i} + E_{2n+2-i,2n+2-i} - E_{2n+1-i,2n+1-i}, & h_n &= 2(E_{n+2,n+2} - E_{n,n}). \end{aligned}$$

To which we add the highest root triplet:

$$\begin{aligned} e_0 &= \frac{1}{2}(E_{1,2n} + E_{2,2n+1}) \otimes \lambda, & f_0 &= 2(E_{2n,1} + E_{2n+1,2}) \otimes \lambda^{-1}, \\ h_0 &= E_{1,1} + E_{2,2} - E_{2n,2n} - E_{2n+1,2n+1}. \end{aligned}$$

Normalization constant of the standard bilinear form: $\kappa = \frac{1}{2}$, i.e., $(X, Y) = \frac{1}{2}\text{tr}(\pi(X)\pi(Y))$.

Generators of the principal Heisenberg subalgebra:

- ◊ If $j \in \mathbb{Z}_{>0}^{\text{odd}}$, then $\Lambda_j = \sqrt{2}\Lambda^j$;
- ◊ If $j \in \mathbb{Z}_{<0}^{\text{odd}}$, then $\Lambda_j = \sqrt{2}z^j(\Lambda^{2n-1})^{-j}$.

Principal cyclic element: $\Lambda = e_0 + \cdots + e_n$,

$$\Lambda = \begin{pmatrix} 0 & \cdots & 0 & \lambda & 0 \\ 1 & \ddots & & 0 & \lambda \\ 0 & 1 & & & 0 \\ \vdots & \ddots & \ddots & & \vdots \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Type $C_n^{(1)}$, $n \geq 2$

The affine Kac–Moody algebras of type $C_n^{(1)}$ have the following features.

Coxeter number: $h = 2n$.

Highest root: $\theta = \alpha_1 + 2\alpha_2 + \cdots + 2\alpha_{n-1} + \alpha_n$.

Affine exponents: $\mathbb{E} = \mathbb{Z}^{\text{odd}} = \mathbb{Z} \setminus 2\mathbb{Z}$.

Extended Dynkin diagram and affine generalized Cartan matrix of size $(n+1) \times (n+1)$:

$$C_n^{(1)} : \begin{array}{c} \circ \rightleftarrows \bullet \cdots \bullet \leftleftarrows \circ \end{array} \quad \text{GCM}(C_n^{(1)}) = \begin{pmatrix} 2 & -1 & 0 & \cdots & 0 \\ -2 & 2 & \ddots & \ddots & \vdots \\ 0 & -1 & \ddots & & 0 \\ \vdots & \ddots & \ddots & & -1 & 0 \\ & & & & 2 & -2 \\ 0 & \cdots & 0 & -1 & 2 \end{pmatrix}$$

Faithful representation: Loop algebra of $C_n^{(1)} \hookrightarrow \mathfrak{sp}(2n) \otimes \mathbb{C}[\lambda, \lambda^{-1}]$.

Weyl generators of the representation: For $i \in \{1, \dots, n-1\}$ (the $\otimes 1$ is implicit),

$$\begin{aligned} e_i &= E_{i+1,i} + E_{2n+1-i,2n-i}, & e_n &= E_{n+1,n}, \\ f_i &= E_{i,i+1} + E_{2n-i,2n+1-i}, & f_n &= E_{n,n+1}, \\ h_i &= E_{i+1,i+1} - E_{i,i} + E_{2n+1-i,2n+1-i} - E_{2n-i,2n-i}, & h_n &= E_{n+1,n+1} - E_{n,n} \end{aligned}$$

To which we add the highest root triplet:

$$e_0 = E_{1,n} \otimes \lambda, \quad f_0 = E_{n,1} \otimes \lambda^{-1}, \quad h_0 = E_{1,1} - E_{2n,2n}.$$

Normalization constant of the standard bilinear form: $\kappa = 1$, i.e., $(X, Y) = \text{tr}(\pi(X)\pi(Y))$.

Generators of the principal Heisenberg subalgebra: $\Lambda_j = \Lambda^j$, for $j \in \mathbb{Z}^{\text{odd}}$.

Principal cyclic element: $\Lambda = e_0 + \cdots + e_n$,

$$\Lambda = \begin{pmatrix} 0 & \cdots & \lambda & \cdots & 0 \\ 1 & \ddots & & & \vdots \\ 0 & 1 & & & \\ \vdots & \ddots & \ddots & & \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

Type $D_n^{(1)}$, $n \geq 2$

The affine Kac–Moody algebras of type $D_n^{(1)}$ have the following features.

Coxeter number: $h = 2n - 2$.

Highest root: $\theta = \alpha_1 + 2\alpha_2 + \cdots + 2\alpha_{n-2} + \alpha_{n-1} + \alpha_n$.

Affine exponents:

- ◊ If n is odd: $\mathbb{E} = \mathbb{Z}^{\text{odd}} \cup (n-1)\mathbb{Z}^{\text{odd}}$;
- ◊ If n is even: $\mathbb{E} = \mathbb{Z}^{\text{odd}} \cup ((n-1)\mathbb{Z}^{\text{odd}})'$.

(When n is even, exponents of the form $(n-1)k$, with $k \in \mathbb{Z}^{\text{odd}}$, have multiplicity 2.)

Extended Dynkin diagram and affine generalized Cartan matrix of size $(n+1) \times (n+1)$:

$D_n^{(1)}$:

GCM($D_n^{(1)}$) =

$$\begin{pmatrix} 2 & 0 & -1 & 0 & \cdots & 0 \\ 0 & 2 & -1 & & \ddots & \vdots \\ -1 & -1 & 2 & \ddots & & \\ 0 & & \ddots & \ddots & & 0 \\ \vdots & \ddots & & & 2 & -1 & -1 \\ & & & & -1 & 2 & 0 \\ 0 & \cdots & 0 & -1 & 0 & 2 \end{pmatrix}$$

Faithful representation: Loop algebra of $D_n^{(1)} \hookrightarrow \mathfrak{o}(2n) \otimes \mathbb{C}[\lambda, \lambda^{-1}]$.

Weyl generators of the representation: For $i \in \{1, \dots, n-1\}$ (the $\otimes 1$ is implicit),

$$\begin{aligned} e_i &= E_{i+1, i} + E_{2n+1-i, 2n-i}, & e_n &= \frac{1}{2}(E_{n+1, n-1} + E_{n+2, n}), \\ f_i &= E_{i, i+1} + E_{2n-i, 2n+1-i}, & f_n &= 2(E_{n-1, n+1} + E_{n, n+2}), \\ h_i &= E_{i+1, i+1} - E_{i, i} + E_{2n+1-i, 2n+1-i} - E_{2n-i, 2n-i}, & h_n &= E_{n+1, n+1} - E_{n, n} + E_{n+2, n+2} - E_{n-1, n-1}. \end{aligned}$$

To which we add the highest root triplet:

$$\begin{aligned} e_0 &= \frac{1}{2}(E_{1, 2n-1} + E_{2, 2n}) \otimes \lambda, & f_0 &= 2(E_{2n-1, 1} + E_{2n, 2}) \otimes \lambda^{-1}, \\ h_0 &= E_{1, 1} + E_{2, 2} - E_{2n-1, 2n-1} - E_{2n, 2n}. \end{aligned}$$

Normalization constant of the standard bilinear form: $\kappa = \frac{1}{2}$, i.e., $(X, Y) = \frac{1}{2} \text{tr}(\pi(X)\pi(Y))$.

Generators of the principal Heisenberg subalgebra: $\Lambda_j = \sqrt{2}\Lambda^j$ and $\Lambda_{((n-1)j)'} = \sqrt{2n-2}\Gamma^j$ for $j \in \mathbb{Z}^{\text{odd}}$, where

$$\Gamma = \mu \left[E_{n,1} - \frac{1}{2}E_{n+1,1} - \frac{1}{2}E_{n,2n} \otimes \lambda + \frac{1}{4}E_{n+1,2n} \otimes \lambda \right. \\ \left. + (-1)^n \left(E_{2n,n+1} - \frac{1}{2}E_{2n,n} - \frac{1}{2}E_{1,n+1} \otimes \lambda + \frac{1}{4}E_{1,n} \otimes \lambda \right) \right],$$

where $\mu = -1$ if n is even and $\mu = -\sqrt{-1}$ if n is odd.

Principal cyclic element: $\Lambda = e_0 + \cdots + e_n$,

$$\Lambda = \begin{pmatrix} 0 & \cdots & & & & 0 & \lambda & 0 \\ 1 & \ddots & & & & & 0 & \lambda \\ 0 & \ddots & & & & & & 0 \\ \vdots & & 1 & & & & & \vdots \\ & & \frac{1}{2} & 0 & & & & \\ & & & \frac{1}{2} & 0 & & & \\ & & & & 1 & & & \\ & & & & & \ddots & & \\ 0 & \cdots & & & & & 0 & 1 & 0 \end{pmatrix}.$$

Bibliography

- [AC91] M. J. Ablowitz and P. A. Clarkson. *Solitons, Nonlinear Evolutions Equations and Inverse Scattering*. LMS Lecture Notes Series 149. Cambridge University Press, 1991.
- [ACvM12] M. Adler, M. Cafasso, and P. van Moerbeke. Non-linear PDEs for gap probabilities in random matrices and KP theory. *Physica D*, 241:2265–2284, 2012.
- [AKNS74] M. J. Ablowitz, D. J. Kaup, A. C. Newell, , and H. Segur. The inverse scattering transform – Fourier analysis for nonlinear problems. *Stud. Appl. Math*, 53, 1974.
- [AM78] M. Adler and J. Moser. On a Class of Polynomials Connected with the Korteweg–De Vries Equation. *Commun. math. Phys.*, 61:1–30, 1978.
- [AMcKM77] H. Airault, H.P. McKean, and J. Moser. Rational and elliptic solutions of the Korteweg–de Vries equation and a related many body problem. *Comm. Pure Appl. Math*, 30:95–148, 1977.
- [Adl79] M. Adler. On a Trace Functional for Formal Pseudo-Differential Operators and the Symplectic Structure of the Korteweg–de Vries Type Equations. *Inventiones mathematicae*, 50:219–248, 1979.
- [BBT03] O. Babelon, D. Bernard, and M. Talon. *Introduction to Classical Integrable Systems*. Cambridge Monographs on Mathematical Physics. Cambridge University Press, 2003.
- [BDGR16] A. Buryak, B. Dubrovin, J. Guere, and P. Rossi. Integrable systems of double ramification type . 2016. arxiv:1609.04059.
- [BDGR18] A. Buryak, B. Dubrovin, J. Guere, and P. Rossi. Tau-structure for the Double Ramification Hierarchies. *Commun. Math. Phys.*, 363:191–260, 2018.
- [BDY16a] M. Bertola, B. Dubrovin, and D. Yang. Correlation functions of the KdV hierarchy and applications to intersection numbers over $\mathcal{M}_{g,n}$. *Physica D: Nonlinear Phenomena*, 327:30–57, 2016.
- [BF97] K. Behrend and B. Fantechi. The intrinsic normal cone. *Invent. Math.*, 128 (1):45–88, 1997.

- [BG16] A. Buryak and Guéré. Towards a description of the double ramification hierarchy for Witten's r -spin class. *Journ. de Math. Pures et Appliquées*, 106:837–865, 2016.
- [BGR17] A. Buryak, J. Guéré, and P. Rossi. DR/DZ equivalence conjecture and tautological relations. arXiv preprint. 2017.
- [BO00] A. Borodin and A. Okounkov. A Fredholm determinant formula for Toeplitz determinants. *Integral Equations and Operator Theory*, 37:386–396, 2000.
- [BPS12] A. Buryak, H. Posthuma, and S. Shadrin. On deformations of quasi-Miura transformations and the Dubrovin–Zhang bracket. *Journ. of Geom. and Phys.*, 62(7):1639–1651, 2012.
- [BR16a] A. Buryak and P. Rossi. Double Ramification Cycles and Quantum Integrable Systems. *Letters Math. Phys.*, 106:289–317, 2016.
- [BR16b] A. Buryak and P. Rossi. Recursion Relations for Double Ramification Hierarchies. *Commun. Math. Phys.*, 342:533–568, 2016.
- [BS99] A Böttcher and Silberman B. *Introduction to Large Truncated Toeplitz Matrices*. Berlin, Heidelberg, New Yor: Springer, 1999.
- [BW00] E. L. Baso and H. Widom. On a Toeplitz determinant identity of Borodin and Okounkov. *Integral Equations and Operator Theory*, 37(104):397–401, 2000.
- [BY17] J. Balog and D. Yang. Geometric interpretation of Zhou's explicit formula for the Witten–Kontsevich tau function. *Lett. Math. Phys.*, 107(10):1837–1857, 2017.
- [Bak02] H. Baker. *Proc Lond Math Soc (Ser 1)*. 34:347–360, 1902; *Proc Lond Math Soc (Ser 1)*. 35:333–374, 1903; *Proc Lond Math Soc (Ser 2)*. 3:24–47, 1905.
- [Bou77] J. Boussinesq. Essai sur la theorie des eaux courantes. *Mémoires présentés par divers savants à l'Acadadémie des Sciences de l'Institut Natational de France*, XXIII, XXIV, 1877.
- [Bou88] N. Bourbaki. *Algèbre, tome IV*. Springer Verlag, 1988.
- [Bur15a] A. Buryak. Double ramification cycles and integrable hierarchies. *Commun. Math. Phys.*, 336(3):1085–1107, 2015.
- [CDY18] M. Cafasso, A. du Crest de Villeneuve, and D. Yang. Drinfeld-Sokolov hierarchies, tau functions, and generalized Schur polynomials. *SIGMA*, 14, 104, 2018.
- [CIR14] A. Chiodo, H Iritani, and Y. Ruan. Landau-Ginzburg/Calabi-Yau correspondence, global mirror symmetry and Orlov equivalence. *Pub. Math. IHES*, 119(1):127–216, 2014.
- [CW15a] M. Cafasso and C.-Z. Wu. Tau functions and the limit of block Toeplitz determinants. *IMRN*, 2015(20):10339–10366, 2015.

-
- [CW15b] M. Cafasso and C.-Z. Wu. Borodin-Okounkov formula, string equation and topological solutions of Drinfeld-Sokolov hierarchies. 2015. arXiv preprint.
- [Caf08] M. Cafasso. Block Toeplitz Determinants, Constrained KP and Gelfand–Dickey Hierarchies. *Math Phys Anal Geom*, 11:11–51, 2008.
- [Cam97] J. Campbell. *Proc Lond Math Soc.* 28:381–390, 1897; *Proc Lond Math Soc.* 29:14–32. 1898.
- [Car05] R. Carter. *Lie Algebras of Finite and Affine Type*. Cambridge studies in advanced mathematics 96. Cambridge University Press, 2005.
- [DJKM81a] E. Date, M. Jimbo, M. Kashiwara, and T. Miwa. Transformation groups for soliton equations III: Operator approach to the Kadomtsev–Pevniashvili equation. *Journ. Phys. Soc. Japan*, 50 (11):3806–3812, 1982.
- [DJKM81b] E. Date, M. Jimbo, M. Kashiwara, and T. Miwa. Transformation groups for soliton equations VI: KP hierarchies of orthogonal and symplectic type. *Journ. Phys. Soc. Japan*, 50 (11):3813–3818, 1982.
- [DJKM83] E. Date, M. Jimbo, M. Kashiwara, and T. Miwa. Transformation groups for soliton equations, in “Nonlinear Integrable Systems—Classical Theory and Quantum Theory,” Kyoto, 1981. pages 39–119, 1983.
- [DJM00] E. Date, M. Jimbo, and T. Miwa. *Solitons, Differential equations, symmetries and infinite dimensional algebras*. Cambridge tracts in mathematics 135. Cambridge University Press, 2000.
- [DKM81] E. Date, M. Kashiwara, and T. Miwa. Transformation groups for soliton equations II: Vertex operators and tau functions. *Proc. of the Japan Academy, Ser. A, Math. Sciences*, 57:387–392, 1981.
- [DLZ08] B. Dubrovin, S.-Q. Liu, and Y. Zhang. Frobenius manifolds and central invariants for the Drinfeld-Sokolov bihamiltonian structures. *Adv. Math.*, 219(3):780–837, 2008.
- [DN83] B. Dubrovin and S. Novikov. The Hamiltonian formalism of one-dimensional systems of hydrodynamic type and the Bogolyubov–Whitham averaging method. *Proc. of the USSR Acad. Sciences (Doklady Akademii Nauk SSSR)*, 270 (4):781–785, 1983.
- [DR19] A. du Crest de Villeneuve and P. Rossi. Quantization of the D_4 Drinfeld–Sokolov hierarchy and Fan-Jarvis-Ruan-Witten theory. Accepted for publication in the *Journal of Geometry and Physics*, 2019.
- [DS84] V. G. Drinfeld and V. V. Sokolov. Lie algebras and equations of Korteweg–de Vries type. *J. Soviet Math.*, 24:81–180, 1984.

- [DZ03] B. Dubrovin and Y. Zhang. Virasoro Symmetries of the Extended Toda Hierarchy. *Proc USSR Academy of Sciences*, 250 (1):161–193, 2003.
- [DZ08] B. Dubrovin and Y. Zhang. Normal forms of hierarchies of integrable PDEs, Frobenius manifolds and Gromov–Witten invariants. Unpublished article; arXiv preprint (v. 2008): math/0108160. 2001.
- [Dic03] L.A. Dickey. *Soliton Equations and Hamiltonian Systems*. World Scientific, 2003.
- [Dub96] B. Dubrovin. Geometry of 2D topological field theories. In “Integrable Systems and Quantum Groups” (Montecatini Terme, 1993), Editors: Francaviglia, M., Greco, S. *Springer Lecture Notes in Math.*, 1620:120–348, 1996.
- [Dub99] B. Dubrovin. Differential geometry of the space of orbits of a Coxeter group. *Surveys in Differential Geometry*, 4:181–212, 1999.
- [Duc17] A. du Crest de Villeneuve. From the Adler–Moser polynomials to the polynomial tau functions of KdV. *Journal of Integrable Systems*, 2(1):xyx012, 2017.
- [Dyn47] E. B. Dynkin. Calculation of the coefficients in the Campbell–Hausdorff formula. *Proc USSR Academy of Sciences*, 57:323–326, 1947.
- [EGH00] Y. Eliashberg, A. Givental, and H. Hofer. Introduction to symplectic field theory. GAFA 2000 Visions in Mathematics special volume, part II. pages 560–673, 2000.
- [EH11] V. Z. Enolskii and J. Harnad. Schur function expansions of KP tau-functions associated to algebraic curves. *Russian Mathematical Surveys*, 66 (4):767, 2011.
- [EW06] K. Erdmann and M. J. Wildon. *Lie Algebras of Finite and Affine Type*. Springer Undergraduate Mathematics Series. Springer, 2006.
- [FFJMR16] H. Fan, A. Francis, T. Jarvis, E. Merrel, and Y. Ruan. Witten’s D 4 Integrable Hierarchies Conjecture. *Chinese Annals of Mathematics, Series B*, 37 (2):175–192, 2016.
- [FGM10] E. Frenkel, A. Givental, and T. Milanov. Soliton equations, vertex operators, and simple singularities. *Funct. Anal. Other Math.*, 3:47–63, 2010.
- [FJ16] A. Francis and T. J. Jarvis. A Brief Survey of FJRW Theory. 2016. arXiv preprint.
- [FJR11] H. Fan, T. J. Jarvis, and Y. Ruan. The Witten equation and its virtual fundamental cycle. 2011. arXiv preprint.
- [FJR13] H. Fan, T. J. Jarvis, and Y. Ruan. The Witten equation, mirror symmetry, and quantum singularity theory. *Annals of Math.*, 178:1–106, 2013.
- [FP05] C. Faber and R. Pandharipande. Relative maps and tautological classes. *Journ. European Math. Soc.*, 7:13–49, 2005.

-
- [FZ71] L. D. Faddeev and V. E. Zakharov. Korteweg–de Vries Equation: A Completely Integrable Hamiltonian System. *Funct. Anal. Appl.*, 5, 1971.
- [GD76] I. M. Gelfand and L. A. Dickey. Fractional powers of operators and Hamiltonian systems. *Funk. Anal. Priloz.*, 10:13–29, 1976. (Russian).
- [GHM92] M. F. de Groot, T. J. Hollowood, and J. L. Miramontes. Generalized Drinfel’d–Sokolov hierarchies. *Comm. Math. Phys.*, 145 (1):57–84, 1992.
- [GJV11] I. P. Goulden, D. M. Jackson, and R. Vakil. The Moduli Space of Curves, Double Hurwitz Numbers, and Faber’s Intersection Number Conjecture. *Annals of Combinatorics*, 14:381–436, 2011.
- [GKM68] C. S. Gardner, M. D. Kruskal, and R. M. Miura. Korteweg–de Vries Equation and Generalizations. II. Existence of Conservation Laws and Constants of Motion. *Journal of Mathematical Physics*, 9 (8):1204–1209, 1968.
- [GKMZ70] C. S. Gardner, M. D. Kruskal, R. M. Miura, and N. J. Zabusky. Korteweg–de Vries Equation and Generalizations. V. Uniqueness and Nonexistence of Polynomial Conservation Laws. *Journal of Mathematical Physics*, 11 (3):952–960, 1970.
- [GM05] A. Givental and T. Milanov. Simple singularities and integrable hierarchies. In “The breadth of symplectic and Poisson geometry”. *Progress in Mathematics*, 232:173–201, Birkhauser Boston, Boston, MA, 1970.
- [GV05] T. Graber and R. Vakil. Relative virtual localization and vanishing of tautological classes on moduli spaces of curves. *Duke Math. Journ.*, 130 (1):1–37, 2005.
- [Gan00] F. R. Gantmacher. The Theory of Matrices, vols I and II. *Russian Mathematical Surveys*, 2000.
- [Gar71] C. S. Gardner. Korteweg–de Vries Equation and Generalizations. IV. The Korteweg–de Vries Equation as a Hamiltonian System. *J. Math. Phys.*, 12, 1971.
- [Get02] E. Getzler. A Darboux theorem for Hamiltonian operators in the formal calculus of variations. *Duke Math. Journ.*, 111 (3):535–561, 2002.
- [Giv01] A. Givental. Semisimple Frobenius structures at higher genus. *Internat. Math. Res. Notices*, 23:1265–1286, 2001.
- [Gue17] J. Guéré. Hodge Integrals in FJRW Theory. *Michigan Math. J.*, 66 (4):831–854, 2017.
- [HM93] T. J. Hollowood and J. L. Miramontes. Tau functions and generalized integrable hierarchies. *Communication in Mathematical Physics*, 157 (1):99–117, 1993.
- [HMSG94] T. J. Hollowood, J. L. Miramontes, and J. Sanchez Guillen. Additional symmetries of generalized integrable hierarchies. *J. Phys. A*, 27 (13):4629–4644, 1994.

- [Hai13] R. Hain. Normal functions and the geometry of moduli spaces of curves. In: Handbook of Moduli, vol. I. *Adv. Lect. Math.*, 24:527–578, 2013.
- [Hau06] F. Hausdorff. *Ber Verh Saechs Akad Wiss Leipzig*, 58:19–48, 1906.
- [Hir04] R. Hirota. *The Direct Method in Soliton Theory*. Cambridge tracts in mathematics 155. Cambridge University Press, 2004.
- [Hir76] R. Hirota. Direct method of finding exact solutions of nonlinear evolution equations, volume 515 of *Lecture Notes in Mathematics*. Springer, Berlin, Heidelberg, 1976.
- [Hum72] J. Humphreys. *Introduction to Lie Algebras and Representation Theory*. Graduate Texts in Mathematics no 9. Springer-Verlag New York, 1972.
- [IZ92] C. Itzykson and J.-B. Zuber. Combinatorics of the modular group. II. The Kontsevich integrals. *Internat. J. Modern Phys. A*, 7:5661–5705, 1992.
- [Igu11] K. Igusa. Lie Algebras. (Unpublished notes). 2011. Notes for Brandeis University’s course Math 223A.
- [JMS78] M. Jimbo, T. Miwa and M. Sato. *RIMS*. 14:223, 1978; *RIMS*. 15:201, 577, 871, 1979; *RIMS*. 16:531, 1980. 1978.
- [JMW81] Jimbo, M. and Miwa, T. and Ueno, K. Monodromy preserving deformations of linear ordinary differential equations with rational coefficients i. *Physica D*, 2:306–352, 1981.
- [KM81] M. Kashiwara and T. Miwa. Transformation groups for soliton equations I: The tau functions of the Kadomtsev–Petviashvili equation. *Proc. Japan Acad.*, 57:342–347, 1981.
- [KM94] Manin Y. Kontsevich, M. Gromov-Witten classes, quantum cohomology, and enumerative geometry. *Proc. Japan Acad.*, 57:342–347, 1981.
- [KP70] B. B. Kadomtsev and V. I. Petviashvili. On the stability of solitary waves in weakly dispersive media. *Sov. Phys. Dokl.*, 15:539–541, 1970.
- [KW89] V. G. Kac and M. Wakimoto. Exceptional hierarchies of soliton equations. In Proceedings of symposia in pure mathematics. 49:191, 1989.
- [Kac68] V. Kac. Simple irreducible graded lie algebras of finite growth. *Iz vestija AN USSR (ser. Math.)*, 32:1923–1967, 1968.
- [Kac90] V. Kac. *Infinite dimensional Lie algebras*. Cambridge University Press, 3 edition, 1990.
- [KdV95] D. J. Korteweg and G. de Vries. On the Change of Form of Long Waves Advancing in a Rectangular Canal, and on a New Type of Long Stationary Waves. *Philosophical Magazine*, 39 (240):422–443, 1895.

-
- [Kon92] M. Kontsevich. Intersection theory on the moduli space of curves and the matrix Airy function. *Comm. Math. Phys.*, 147 (1):1–23, 1992.
- [Kos59] B. Kostant. The principal three-dimensional subgroup and the Betti numbers of a complex simple Lie group. *Am. J. Math.*, 81, No 2:973–1032, 1959.
- [KvdL18] V. Kac and J. van de Leur. Equivalence of formulations of the MKP hierarchy and its polynomial tau-functions. *Japanese Journal of Mathematics*, 13 (2):235–271, 2018.
- [LRZ15] S.-Q. Liu, Y. Ruan, and Y. Zhang. BCFG Drinfeld–Sokolov hierarchies and FJRW-theory. *Invent. math.*, 201:711–772, 2015.
- [LVY18] S. Lacroix, B. Vicedo, and C. Young. Affine Gaudin models and hypergeometric functions on affine opers. 2018. arXiv preprint.
- [LWZ10] S.-Q. Liu, C.-Z. Wu, and Y. Zhang. On the Drinfeld–Sokolov hierarchies of D type. *International Mathematics Research Notices*, 2011, No 8:1952–1996, 2010.
- [Lax68] P. D. Lax. Integrals of nonlinear equations of evolution and solitary waves. *Com. Pure and App. Math.*, 21:467–490, 1968.
- [Lax75] P. D. Lax. Periodic Solutions of the KdV Equation. *Com. Pure and App. Math.*, 28:141–188, 1975.
- [Mac95] I. G. Macdonald. *Symmetric functions and Hall polynomials*. Oxford Mathematical Monographs. Oxford University Press Inc., New York, 1995.
- [Mag78] F. Magri. A simple model of the integrable Hamiltonian equations. *J. Math. Phys.*, 19 (5):1156, 1978.
- [Mir95] R. Miranda. *Algebraic Curves and Riemann Surfaces*. Graduate Studies in Mathematics. American Mathematical Society, 1995.
- [Mir99] J. L. Miramontes. Tau-functions generating the conservation laws for generalized integrable hierarchies of KdV and affine toda type. *Nuclear Phys. B*, 547:623–663, 1999.
- [Miu68] R. M. Miura. Korteweg–deVries Equation and Generalizations. I. Explicit Nonlinear Transformation. *Journal of Mathematical Physics*, 9 (8):1202–1204, 1967.
- [Miw81] Miwa, T. Painlevé property of monodromy preserving equations and the analyticity of τ -function. *Publ. RIMS*, 17 (2):703–721, 1981.
- [Moo68] R. V. Moody. A New Class of Lie Algebras. *Journal of Algebra*, 10 (2):211–230, 1968.
- [PPZ15] R. Pandharipande, A. Pixton, and D. Zvonkine. Relations on $\overline{\mathcal{M}}_{g,n}$ via 3-spin structures. *Journal of the American Mathematical Society*, 28, 1:279–309, 2015.

- [PS05] J. Praught and R. G. Smirnov. Andrew Lenard: A Mystery Unraveled. *SIGMA*, 1, 2005.
- [PV16] A. Polishchuk and A. Vaintrob. Matrix factorizations and cohomological field theories. *Journ. reine angew. Math.*, 714:1–122, 2016.
- [Ros02] W. Rossmann. *Lie Groups – An Introduction Through Linear Groups*. Oxford Graduate Texts in Mathematics 5. Oxford Science Publications, 2002.
- [Ros17] P. Rossi. Integrability, Quantization and Moduli Spaces of Curves. *SIGMA*, 13, 2017.
- [SW85] G. Segal and G. Wilson. Loop groups and equations of KdV type. *Publications mathématiques de l’I.H.E.S.*, 61:5–65, 1985.
- [Sai81] K. Saito. Primitive forms for a universal unfolding of a function with an isolated critical point. *Journal of the Faculty of Science: University of Tokyo, Section IA Mathematics*, 28 (3):775–792, 1981.
- [Sai83a] K. Saito. The higher residue pairings $K_F^{(k)}$ for a family of hypersurface singular points. In “Singularities, Part 2 (Arcata, CA, 1981)”. *Proc. Sympos. Pure Math.*, 4 (3):441–463. Amer. Math. Soc., Providence, RI., 1983.
- [Sai83b] K. Saito. Period mapping associated to a primitive form. *Publ. Res. Inst. Math. Sci.*, 19 (3):1231–1264, 1983.
- [Sat81] M. Sato. Soliton Equations as Dynamical Systems on an Infinite Dimensional Grassmann Manifolds. *RIMS Kokyuroku*, 439:30–46, 1981.
- [Sch01] Issai Schur. Über eine Klasse von Matrizen, die sich einer gegebenen Matrix zuordnen lassen. Doctoral dissertation, Universität Berlin, 1901. Reprinted in Brauer k Rohrbach, 71:1–72, 1973.
- [Ser92] J.-P. Serre. *Lie Algebras and Lie Groups: 1964 Lectures given at Harvard University*. Lecture Notes in Mathematics 1500. Springer-Verlag Berlin Heidelberg, 2 edition, 1992.
- [Sim05] B. Simon. *Trace Ideals and Their Applications: Second Edition*. Mathematical Surveys and Monographs, vol. 120. Providence, RI: AMS, 2 edition, 2005.
- [Tel12] C. Teleman. The structure of 2D semi-simple field theories. *Inventiones Mathematicae*, 188, 3:525–588, 2012.
- [Wid76] H. Widom. Asymptotic behavior of block Toeplitz matrices and determinants II. *Advances in Mathematics*, 21 (1):1–29, 1976.
- [Wil98] G. Wilson. Collisions of Calogero-Moser particles and an adelic Grassmannian. *Invent. Math.*, 133:1–41, 1998.

-
- [Wit91] E. Witten. Two-dimensional gravity and intersection theory on moduli space. *Surveys in differential geometry*, 1:243–310, 1991.
- [Wit93] E. Witten. Algebraic geometry associated with matrix models of two-dimensional gravity. *Topological methods in modern mathematics (Stony Brook, NY, 1991)*, 340 (2-3):235–269, 1993.
- [Wu17] C.-Z. Wu. Tau functions and Virasoro symmetries for Drinfeld–Sokolov hierarchies. *Advances in Mathematics*, 306:603–652, 2017.
- [Zvo12] D. Zvonkine. An introduction to moduli spaces of curves and its intersection theory. In “Handbook of Teichmüller Theory, Volume III”. pages 667–716, 2012.