Advances in Mathematical Physics

Nonlinear and Noncommutative Mathematics: New Developments and Applications in Quantum Physics

Guest Editors: M. Norbert Hounkonnou, S. Twareque Ali, Gerald A. Goldin, Richard Kerner, Kalyan B. Sinha, and Akira Yoshioka

Nonlinear and Noncommutative Mathematics: New Developments and Applications in Quantum Physics

Nonlinear and Noncommutative Mathematics: New Developments and Applications in Quantum Physics

Guest Editors: M. Norbert Hounkonnou, S. Twareque Ali, Gerald A. Goldin, Richard Kerner, Kalyan B. Sinha, and Akira Yoshioka

Editorial Board

Contents

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 497040, 4 pages
doi:10.1155/2010/497040

Editorial

Nonlinear and Noncommutative Mathematics: New Developments and Applications in Quantum Physics

M. Norbert Hounkonnou,[1] S. Twareque Ali,[2] Gerald A. Goldin,[3] Richard Kerner,[4] Kalyan B. Sinha,[5] and Akira Yoshioka[6]

[1] *International Chair in Mathematical Physics and Applications (ICMPA-UNESCO Chair), University of Abomey-Calavi (UAC), 072 BP 50, Cotonou, Benin*
[2] *Department of Mathematics and Statistics, Concordia University, 1455 de Maisonneuve Blvd. West, Montreal, Quebec, Canada H3G 1M8*
[3] *Departments of Mathematics and Physics, Rutgers University, Piscataway, NJ 08854, USA*
[4] *Laboratoire de Physique Théorique de la Matière Condensée (LPTMC), Université Pierre et Marie Curie, 4 Place Jussieu, 75005 Paris, France*
[5] *Jawaharlal Nehru Centre for Advanced Scientific Research, Indian Institute of Science, Bangalore, India*
[6] *Department of Mathemtaics, Faculty of Science, Tokyo University of Science, Kagurazaka 1-3, Shinjyuku-ku, Tokyo 162-8601, Japan*

Correspondence should be addressed to M. Norbert Hounkonnou, norbert.hounkonnou@cipma.uac.bj

Received 8 September 2010; Accepted 8 September 2010

1. Introduction to the Special Issue

This Special Issue of Advances in Mathematical Physics addresses a number of contemporary topics, centered in the domains (broadly construed) of nonlinear and noncommutative mathematics. These encompass some rapidly developing, versatile areas without, of course, well-defined boundaries. They not only suggest mathematical generalizations in many different directions, but their applications in theoretical and mathematical physics are diverse—in the fields of quantization, quantum theory, and quantum information, in high-energy physics and quantum gravity, in plasma physics and condensed matter theory, and in other areas.

The ten articles in the Special Issue can address but a small subset of these fields, making use of a cross-section of mathematical methods. Some of the papers are intended to present reviews of important areas, while others are more specifically focused. In the latter case, we nevertheless sought to encourage the authors to provide some general background.

2. Quantization and Quantum Field Theory

The issue begins with a set of four articles directly pertinent to problems of quantization and quantum field theory, while the methods described also have many other important applications. The review by Martin Schlichenmaier provides a succinct yet detailed exposition of the key concepts in Berezin-Toeplitz and the related deformation quantization for compact Kähler manifolds. This method of quantization has been theoretically one of the most useful quantization techniques, the strongest results in the field coming from compact manifolds. The author also points out possible uses of this technique in the study of noncommutative geometry, for example, fuzzy spheres. Next Jean-Pierre Antoine and Camillo Trapani present an overview of partial inner product spaces and operators on them, with the results illustrated by families of function spaces that are important in signal analysis as well as mathematical physics. While the Hilbert space as the state space is germane to quantum mechanics, the use of operators with continuous spectra is traditionally handled within the Dirac bra-ket scheme which, mathematically speaking, involves the use of Gelfand triples. A mathematically rigorous and adequate framework for such a structure is the more general concept of a partial inner product space. Similarly, in signal processing, where continuous scales of smoothness are often required, the use of partial inner product spaces becomes particularly useful.

In the third article, John Klauder defines and analyzes coherent states associated with covariant, non-renormalizable, self-interacting scalar fields quantized so as to satisfy affine rather than canonical commutation relations. Coherent states are well known objects in physics and in analysis. The present article analyzes affine coherent states, built using the position and dilatation operators, starting from the ground state of a lattice version of a covariant, non-renormalizable, self-interacting, scalar quantum field. Then Ivan Todorov reviews his joint work with B. Bakalov, N. M. Nikolov and K.-H. Rehren on gauge symmetry and Howe duality in a local scalar field theory of conformal dimension two in four-dimensional space-time. The review extends the results obtained in the two-dimensional conformal field theories to four dimensions. Possible models are surveyed and the infinite-dimensional Lie algebras that naturally arise in this context are investigated. One of the most important results is the example of realization of the Doplicher-Haag-Roberts' theory of superselection sectors and compact gauge groups. This survey article has great pedagogical value and will become one of the best sources for all those who are interested in conformal field theories.

3. Two-Dimensional Projective Geomeotry

Following these extensive reviews, we include a short mathematical article by Kyousuke Uchino using lattice theory to explain V. Arnold's two-dimensional projective geometry. Arnold considered the set of all homogeneous quadratic forms of two variables, which is a Poisson algebra by the canonical Poisson bracket. The point of Arnold's idea is that the points of the projective plane are replaced with quadratic forms. Using the canonical Poisson bracket, Arnold showed that the Jacobi identity of the Poisson algebra gives the altitude theorem. Extending Arnold's idea Tomihisa gave proofs of fundamental theorems of projective planes, such as Pappus' theorem, Pascal's theorem and Brianchon's theorem by means of the Poisson algebra. Uchino gives a clear view of these arguments by means of the lattice structure of a Lie algebra. Using the Plucker embedding, Uchino shows that the two-dimensional projective geometry is encoded in $sl(2)$. As an application, the classical

Yang-Baxter equation is translated into the geometry of a quadratic curve on the projective plane.

4. Infinite-Dimensional Groups and Algebras, Gauge Theories, Topological String Theory, and Instantons

The next paper in this special issue is a wide-ranging survey by Rudolf Schmid of infinite-dimensional Lie groups and algebras, including diffeomorphism groups, gauge groups, and loop groups, together with some of their areas of application in mathematical physics. Mathematically, infinite-dimensional Lie groups, being non-locally compact, do not satisfy many of the elegant properties of finite dimensional groups. Physically, however, they can describe the symmetries of systems (including nonlinear systems) having infinitely-many degrees of freedom; and they can encode mathematically the important idea of local symmetry. The areas of application discussed in this article range from electromagnetism (Maxwell's equations), hydrodynamics (Euler's equations), and plasma physics (Maxwell-Vlasov equations) to the Korteweg-de Vries equation, BRST symmetry, the standard model, gravity, and supersymmetry.

We then present a review by Richard Szabo of some results in gauge theories and topological string theory. Connections are developed between instanton counting in maximally supersymmetric gauge theories, and invariants of smooth varieties. The counting problems discussed by Szabo connect six-dimensional topological string theory/gauge theory, four-dimensional supersymmetric gauge theories, three-dimensional Chern Simons theory, and two-dimensional q-deformed Yang-Mills theory, as well as two-dimensional conformal field theory—with relationships to the entropy of supersymmetric black holes and, for example, Donaldson-Thomas and Gromov-Witten invariants. These topological models, obtained as topological twists of a given physical theory, can serve as exactly solvable systems which capture the physical content of certain sectors of more elaborate systems with local propagating degrees of freedom. They describe the BPS sectors of physical models, and compute nonperturbative effects therein. Thus, for certain supersymmetric charged black holes, the microscopic Bekenstein-Hawking-Wald entropy is computed by the Witten index of the relevant supersymmetric gauge theory. This is equivalent to the counting of stable BPS bound states of D-branes in the pertinent geometry, and is related to invariants of three-folds via the OSV conjecture.

There follows Akifumi Sako's review of four-dimensional noncommutative instantons, focusing on some recent perspectives and showing, for example, how the noncommutative instanton charges coincide with the charges of the instanton solutions prior to noncommutative deformation.

5. Additional Topics and Innovative Methods

The final two articles in this Special Issue again present more specific results. Rémi Léandre discusses white noise analysis and quantum probability (described by means of a bosonic Fock space), and Malliavin Calculus (for the Wiener measure in L^p space), constructing a Poisson structure and a Lie algebroid on the space of smooth 1-forms on the Wiener space using the Nualart-Pardoux definition of Anticipative Stratanovich integrals. The development, enriched by a good list of relevant biographical references, contains interesting comments on the implications of relevant results and clarifies various concepts in the field.

Finally, A. Bostan et al. draw our attention to a variety of essential concepts as they exemplify exact representations of the renormalization group as isogenies of elliptic curves using linear differential equations covariant under commuting rational transformations of infinite order. These novel and original results are at the crossroads of many domains of mathematical physics, like the theory of elliptic curves and modular forms, hypergeometric function theory and differential algebra, to name just a few. The concepts and ideas introduced in this seminal paper provide new and powerful tools for deeper investigation and analysis of lattice models of statistical mechanics, creating an unexpected bridge to Calabi-Yau manifolds and string theory. Last but not least, this paper provides an example of a fruitful collaboration, in which the skills and deep knowledge of applied mathematics, fundamental mathematics and theoretical physics brought in by co-authors coming from different countries and continents merge in a very harmonious way to give a beautiful and useful result.

Many of the papers also include substantial bibliographic resources. We expect these expositions to be a rich resource, of interest both to mathematicians and to theoretical physicists.

M. Norbert Hounkonnou
S. Twareque Ali
Gerald A. Goldin
Richard Kerner
Kalyan B. Sinha
Akira Yoshioka

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 927280, 38 pages
doi:10.1155/2010/927280

Review Article

Berezin-Toeplitz Quantization for Compact Kähler Manifolds. A Review of Results

Martin Schlichenmaier

Mathematics Research Unit, University of Luxembourg, FSTC, 6, rue Coudenhove-Kalergi 1359 Luxembourg-Kirchberg, Luxembourg

Correspondence should be addressed to Martin Schlichenmaier, martin.schlichenmaier@uni.lu

Received 30 December 2009; Accepted 5 March 2010

Academic Editor: S. T. Ali

This article is a review on Berezin-Toeplitz operator and Berezin-Toeplitz deformation quantization for compact quantizable Kähler manifolds. The basic objects, concepts, and results are given. This concerns the correct semiclassical limit behaviour of the operator quantization, the unique Berezin-Toeplitz deformation quantization (star product), covariant and contravariant Berezin symbols, and Berezin transform. Other related objects and constructions are also discussed.

1. Introduction

For quantizable Kähler manifolds the Berezin-Toeplitz (BT) quantization scheme, both the operator quantization and the deformation quantization, supplies canonically defined quantizations. Some time ago, in joint work with Martin Bordemann and Eckhard Meinrenken, the author of this review showed that for compact Kähler manifolds it is a well-defined quantization scheme with correct semiclassical limit [1].

What makes the Berezin-Toeplitz quantization scheme so attractive is that it does not depend on further choices and that it does not only produce a formal deformation quantization, but one which is deeply related to some operator calculus.

From the point of view of classical mechanics, compact Kähler manifolds appear as phase space manifolds of restricted systems or of reduced systems. A typical example of its appearance is given by the spherical pendulum which after reduction has as phase-space the complex projective space.

Very recently, inspired by fruitful applications of the basic techniques of the Berezin-Toeplitz scheme beyond the quantization of classical systems, the interest in it revived considerably.

For example, these techniques show up in a noncommutative geometry. More precisely, they appear in the approach to noncommutative geometry using fuzzy manifolds. The quantum spaces of the BT quantization of level m, defined in Section 3 further down, are finite-dimensional, and the quantum operator of level m constitutes finite-dimensional noncommutative matrix algebras. This is the arena of noncommutative fuzzy manifolds and gauge theories over them. The classical limit, the commutative manifold, is obtained as limit $m \to \infty$. The name *fuzzy sphere* was coined by Madore [2] for a certain quantized version of the Riemann sphere. It turned out to be a quite productive direction in the noncommutative geometry approach to quantum field theory. It is impossible to give a rather complete list of people working in this approach. The following is a rather erratic and random choice of [3–10].

Another appearance of Berezin-Toeplitz quantization techniques as basic ingredients is in the pioneering work of Jørgen Andersen on the mapping class group (MCG) of surfaces in the context of Topological Quantum Field Theory (TQFT). Beside other results, he was able to prove the asymptotic faithfulness of the mapping group action on the space of covariantly constant sections of the Verlinde bundle with respect to the Axelrod-Witten-de la Pietra and Witten connection [11, 12]; see also [13]. Furthermore, he showed that the MCG does not have Kazhdan's property T. Roughly speaking, a group which has *property T* says that the identity representation is isolated in the space of all unitary representations of the group [14]. In these applications, the manifolds to be quantized are the moduli spaces of certain flat connections on Riemann surfaces or, equivalently, the moduli space of stable algebraic vector bundles over smooth projective curves. Here further exciting research is going on, in particular, in the realm of TQFT and the construction of modular functors [15–17].

In general, quite often moduli spaces come with a Kähler structure which is quantizable. Hence, it is not surprising that the Berezin-Toeplitz quantization scheme is of importance in moduli space problems. Noncommutative deformations and a quantization being a noncommutative deformation, yield also information about the commutative situation. These aspects clearly need further investigations.

There are a lot of other applications on which work has already been done, recently started, or can be expected. As the Berezin-Toeplitz scheme has become a basic tool, this seems the right time to collect the techniques and results in such a review. We deliberately concentrate on the case of compact Kähler manifolds. In particular, we stress the methods and results valid for all of them. Due to "space-time" limitations, we will not deal with the noncompact situation. In this situation, case by case studies of the examples or class of examples are needed. See Section 3.7 for references to some of them. Also we have to skip presenting recent attempts to deal with special singular situations, like orbifolds, but see at least [18–20].

Of course, there are other reviews presenting similar quantization schemes. A very incomplete list is the following [21–25].

This review is self-contained in the following sense. I try to explain all notions and concepts needed to understand the results and theorems only assuming some background in modern geometry and analysis. And as such it should be accessible for a newcomer to the field (both for mathematicians as for physicists) and help him to enter these interesting research directions. It is not self-contained in the strict sense as it does supply only those proofs or sketches of proofs which are either not available elsewhere or are essential for the understanding of the statements and concepts. The review does not require a background in quantum physics as only mathematical aspects of quantizations are touched on.

2. The Set-Up of Geometric Quantization

In the following, I will recall the principal set-up of geometric quantization which is usually done for symplectic manifolds in the case when the manifold is a Kähler manifold.

2.1. Kähler Manifolds

We will only consider phase-space manifolds which carry the structure of a Kähler manifold (M, ω). Recall that in this case M is a complex manifold and ω, the Kähler form, is a nondegenerate closed positive $(1,1)$-form.

If the complex dimension of M is n, then the Kähler form ω can be written with respect to local holomorphic coordinates $\{z_i\}_{i=1,\dots,n}$ as

$$\omega = \mathrm{i} \sum_{i,j=1}^{n} g_{ij}(z) dz_i \wedge d\overline{z}_j, \tag{2.1}$$

with local functions $g_{ij}(z)$ such that the matrix $(g_{ij}(z))_{i,j=1,\dots,n}$ is hermitian and positive definite.

Later on we will assume that M is a compact Kähler manifold.

2.2. Poisson Algebra

Denote by $C^\infty(M)$ the algebra of complex-valued (arbitrary often) differentiable functions with the point-wise multiplication as an associative product. A symplectic form on a differentiable manifold is a closed nondegenerate 2-form. In particular, we can consider our Kähler form ω as a symplectic form.

For symplectic manifolds, we can introduce on $C^\infty(M)$ a Lie algebra structure, the Poisson bracket *Poisson bracket* $\{\cdot,\cdot\}$, in the following way. First we a assign to every $f \in C^\infty(M)$ its *Hamiltonian vector field* X_f, and then to every pair of functions f and g the *Poisson bracket* $\{\cdot,\cdot\}$ via

$$\omega(X_f, \cdot) = df(\cdot), \qquad \{f, g\} := \omega(X_f, X_g). \tag{2.2}$$

One verifies that this defines a Lie algebra structures and that furthermore, we have the Leibniz rule

$$\{fg, h\} = f\{g, h\} + \{f, h\}g, \quad \forall f, g, h \in C^\infty(M). \tag{2.3}$$

Such a compatible structure is called a *Poisson algebra*.

2.3. Quantum Line Bundles

A *quantum line bundle* for a given symplectic manifold (M, ω) is a triple (L, h, ∇), where L is a complex line bundle, h a Hermitian metric on L, and ∇ a connection compatible with the metric h such that the (pre)quantum condition

$$\begin{gathered}\operatorname{curv}_{L,\nabla}(X, Y) := \nabla_X \nabla_Y - \nabla_Y \nabla_X - \nabla_{[X,Y]} = -\mathrm{i}\,\omega(X, Y), \\ \text{resp.,} \quad \operatorname{curv}_{L,\nabla} = -\mathrm{i}\,\omega\end{gathered} \tag{2.4}$$

is fulfilled. A symplectic manifold is called *quantizable* if there exists a quantum line bundle.

In the situation of Kähler manifolds, we require for a quantum line bundle to be holomorphic and that the connection is compatible both with the metric h and the complex structure of the bundle. In fact, by this requirement ∇ will be uniquely fixed. If we choose local holomorphic coordinates on the manifold and a local holomorphic frame of the bundle, the metric h will be represented by a function $\widehat{h}$. In this case, the curvature in the bundle can be given by $\overline{\partial}\partial \log \widehat{h}$ and the quantum condition reads as

$$\mathrm{i}\overline{\partial}\partial \log \widehat{h} = \omega. \tag{2.5}$$

2.4. Example: The Riemann Sphere

The Riemann sphere is the complex projective line $\mathbb{P}^1(\mathbb{C}) = \mathbb{C} \cup \{\infty\} \cong S^2$. With respect to the quasiglobal coordinate z, the form can be given as

$$\omega = \frac{\mathrm{i}}{(1 + z\overline{z})^2} dz \wedge d\overline{z}. \tag{2.6}$$

For the Poisson bracket, one obtains

$$\{f, g\} = \mathrm{i}(1 + z\overline{z})^2 \left(\frac{\partial f}{\partial \overline{z}} \cdot \frac{\partial g}{\partial z} - \frac{\partial f}{\partial z} \frac{\partial g}{\partial \overline{z}} \right). \tag{2.7}$$

Recall that the points in $\mathbb{P}^1(\mathbb{C})$ correspond to lines in $\mathbb{C}^2$ passing through the origin. If we assign to every point in $\mathbb{P}^1(\mathbb{C})$ the line it represents, we obtain a holomorphic line bundle, called the tautological line bundle. The hyper plane section bundle is dual to the tautological bundle. It turns out that it is a quantum line bundle. Hence $\mathbb{P}^1(\mathbb{C})$ is quantizable.

2.5. Example: The Complex Projective Space

Next we consider the n-dimensional complex projective space $\mathbb{P}^n(\mathbb{C})$. The example above can be extended to the projective space of any dimension. The Kähler form is given by the Fubini-Study form

$$\omega_{\mathrm{FS}} := \mathrm{i}\frac{\left(1+|w|^2\right)\sum_{i=1}^{n} dw_i \wedge d\overline{w}_i - \sum_{i,j=1}^{n} \overline{w}_i w_j dw_i \wedge d\overline{w}_j}{\left(1+|w|^2\right)^2}. \tag{2.8}$$

The coordinates w_j, $j = 1,\ldots,n$, are affine coordinates $w_j = z_j/z_0$ on the affine chart $U_0 := \{(z_0 : z_1 : \cdots : z_n) \mid z_0 \neq 0\}$. Again, $\mathbb{P}^n(\mathbb{C})$ is quantizable with the hyper plane section bundle as a quantum line bundle.

2.6. Example: The Torus

The (complex-) one-dimensional torus can be given as $M = \mathbb{C}/\Gamma_\tau$, where $\Gamma_\tau := \{n + m\tau \mid n, m \in \mathbb{Z}\}$ is a lattice with $\operatorname{Im}\tau > 0$. As Kähler form, we take

$$\omega = \frac{\mathrm{i}\pi}{\operatorname{Im}\tau} dz \wedge d\overline{z}, \tag{2.9}$$

with respect to the coordinate z on the covering space $\mathbb{C}$. Clearly this form is invariant under the lattice Γ_τ and hence well defined on M. For the Poisson bracket, one obtains

$$\{f, g\} = \mathrm{i}\frac{\operatorname{Im}\tau}{\pi}\left(\frac{\partial f}{\partial \overline{z}} \cdot \frac{\partial g}{\partial z} - \frac{\partial f}{\partial z}\frac{\partial g}{\partial \overline{z}}\right). \tag{2.10}$$

The corresponding quantum line bundle is the theta line bundle of degree 1, that is, the bundle whose global sections are scalar multiples of the Riemann theta function.

2.7. Example: The Unit Disc and Riemann Surfaces

The unit disc

$$\mathfrak{D} := \{z \in \mathbb{C} \mid |z| < 1\} \tag{2.11}$$

is a noncompact Kähler manifold. The Kähler form is given by

$$\omega = \frac{2\mathrm{i}}{(1 - z\overline{z})^2} dz \wedge d\overline{z}. \tag{2.12}$$

Every compact Riemann surface M of genus $g \geq 2$ can be given as the quotient of the unit disc under the fractional linear transformations of a Fuchsian subgroup of $SU(1,1)$. If $R = ((a/b)\ (b/a))$ with $|a|^2 - |b|^2 = 1$ (as an element of $SU(1,1)$), then the action is

$$z \longmapsto R(z) := \frac{az+b}{\overline{b}z+\overline{a}}. \tag{2.13}$$

The Kähler form (2.12) is invariant under the fractional linear transformations. Hence it defines a Kähler form on M. The quantum bundle is the canonical bundle, that is, the bundle whose local sections are the holomorphic differentials. Its global sections can be identified with the automorphic forms of weight 2 with respect to the Fuchsian group.

2.8. Consequences of Quantizability

The above examples might create the wrong impression that every Kähler manifold is quantizable. This is not the case. For example, only those higher-dimensional tori complex tori are quantizable which are abelian varieties, that is, which admit enough theta functions. It is well known that for $n \geq 2$ a generic torus will not be an abelian variety. Why this implies that they will not be quantizable, we will see in a moment.

In the language of differential geometry, a line bundle is called a positive line bundle if its curvature form (up to a factor of $1/\mathrm{i}$) is a positive form. As the Kähler form is positive, the quantum condition (2.4) yields that a quantum line bundle L is a positive line bundle.

2.9. Embedding into Projective Space

In the following, we assume that M is a quantizable compact Kähler manifold with quantum line bundle L. Kodaira's embedding theorem says that L is ample, that is, that there exists a certain tensor power L^{m_0} of L such that the global holomorphic sections of L^{m_0} can be used to embed the phase space manifold M into the projective space of suitable dimension. The embedding is defined as follows. Let $\Gamma_{\mathrm{hol}}(M, L^{m_0})$ be the vector space of global holomorphic sections of the bundle L^{m_0}. Fix a basis $s_0, s_1, \ldots, s_N$. We choose local holomorphic coordinates z for M and a local holomorphic frame $e(z)$ for the bundle L. After these choices, the basis elements can be uniquely described by local holomorphic functions $\widehat{s}_0, \widehat{s}_1, \ldots, \widehat{s}_N$ defined via $s_j(z) = \widehat{s}_j(z)e(z)$. The embedding is given by the map

$$M \hookrightarrow \mathbb{P}^N(\mathbb{C}), \qquad z \longmapsto \phi(z) = (\widehat{s}_0(z) : \widehat{s}_1(z) : \cdots : \widehat{s}_N(z)). \tag{2.14}$$

Note that the point $\phi(z)$ in projective space neither depends on the choice of local coordinates nor on the choice of the local frame for the bundle L. Furthermore, a different choice of basis corresponds to a $\mathrm{PGL}(N,\mathbb{C})$ action on the embedding space and hence the embeddings are projectively equivalent.

By this embedding, quantizable compact Kähler manifolds are complex submanifolds of projective spaces. By Chow's theorem [26], they can be given as zero sets of homogenous polynomials, that is, they are smooth projective varieties. The converse is also true. Given a smooth subvariety M of $\mathbb{P}^n(\mathbb{C})$, it will become a Kähler manifold by restricting the Fubini-Study form. The restriction of the hyper plane section bundle will be an associated quantum line bundle.

At this place a warning is necessary. The embedding is only an embedding as complex manifolds are not an isometric embedding as Kähler manifolds. This means that in general $\phi^{-1}(\omega_{\mathrm{FS}}) \neq \omega$. See Section 7.6 for results on an "asymptotic expansion" of the pullback.

A line bundle, whose global holomorphic sections will define an embedding into projective space, is called a *very ample line bundle*. In the following, we will assume that L is already very ample. If L is not very ample, we choose $m_0 \in \mathbb{N}$ such that the bundle L^{m_0} is very ample and take this bundle as quantum line bundle with respect to the rescaled Kähler form $m_0\,\omega$ on M. The underlying complex manifold structure will not change.

3. Berezin-Toeplitz Operators

In this section, we will consider an operator quantization. This says that we will assign to each differentiable (differentiable will always mean differentiable to any order) function f on our Kähler manifold M (i.e., on our "phase space") the Berezin-Toeplitz (BT) quantum operator T_f. More precisely, we will consider a whole family of operators $T_f^{(m)}$. These operators are defined in a canonical way. As we know from the Groenewold-van Howe theorem, we cannot expect that the Poisson bracket on M can be represented by the Lie algebra of operators if we require certain desirable conditions see [27] for further details. The best we can expect is that we obtain it at least "asymptotically". In fact, this is true.

In our context also the operator of geometric quantization exists. At the end of this section, we will discuss its relation to the BT quantum operator. It will turn out that if we take for the geometric quantization the Kähler polarization then they have the same asymptotic behaviour.

3.1. *Tensor Powers of the Quantum Line Bundle*

Let (M, ω) be a compact quantizable Kähler manifold and (L, h, ∇) a quantum line bundle. We assume that L is already very ample. We consider all its tensor powers

$$\left(L^m, h^{(m)}, \nabla^{(m)}\right). \tag{3.1}$$

Here $L^m := L^{\otimes m}$. If $\widehat{h}$ corresponds to the metric h with respect to a local holomorphic frame e of the bundle L, then $\widehat{h}^m$ corresponds to the metric $h^{(m)}$ with respect to the frame $e^{\otimes m}$ for the bundle L^m. The connection $\nabla^{(m)}$ will be the induced connection.

We introduce a scalar product on the space of sections. In this review, we adopt the convention that a hermitian metric (and a scalar product) is antilinear in the first argument and linear in the second argument. First we take the Liouville form $\Omega = (1/n!)\omega^{\wedge n}$ as a volume form on M and then set for the scalar product and the norm

$$\langle \varphi, \psi \rangle := \int_M h^{(m)}(\varphi, \psi)\Omega, \qquad \|\varphi\| := \sqrt{\langle \varphi, \varphi \rangle}, \tag{3.2}$$

on the space $\Gamma_\infty(M, L^m)$ of global C^∞-sections. Let $L^2(M, L^m)$ be the L^2-completion of $\Gamma_\infty(M, L^m)$, and $\Gamma_{\mathrm{hol}}(M, L^m)$ its (due to the compactness of M) finite-dimensional closed subspace of global holomorphic sections. Let

$$\Pi^{(m)} : \mathrm{L}^2(M, L^m) \longrightarrow \Gamma_{\mathrm{hol}}(M, L^m) \tag{3.3}$$

be the projection onto this subspace.

Definition 3.1. For $f \in C^\infty(M)$, the *Toeplitz operator* $T_f^{(m)}$ *(of level m)* is defined by

$$T_f^{(m)} := \Pi^{(m)}(f\cdot) : \Gamma_{\mathrm{hol}}(M, L^m) \longrightarrow \Gamma_{\mathrm{hol}}(M, L^m). \tag{3.4}$$

In words, one takes a holomorphic section s and multiplies it with the differentiable function f. The resulting section $f \cdot s$ will only be differentiable. To obtain a holomorphic section, one has to project it back on the subspace of holomorphic sections.

The linear map

$$T^{(m)} : C^\infty(M) \longrightarrow \mathrm{End}(\Gamma_{\mathrm{hol}}(M, L^m)), \quad f \longrightarrow T_f^{(m)} = \Pi^{(m)}(f\cdot), \quad m \in \mathbb{N}_0 \tag{3.5}$$

is the *Toeplitz* or *Berezin-Toeplitz quantization map (of level m)*. It will neither be a Lie algebra homomorphism nor an associative algebra homomorphism as in general

$$T_f^{(m)} T_g^{(m)} = \Pi^{(m)}(f\cdot)\Pi^{(m)}(g\cdot)\Pi^{(m)} \neq \Pi^{(m)}(fg\cdot)\Pi = T_{fg}^{(m)}. \tag{3.6}$$

Furthermore, on a fixed level m, it is a map from the infinite-dimensional commutative algebra of functions to a noncommutative finite-dimensional (matrix) algebra. The finite-dimensionality is due to the compactness of M. A lot of classical information will get lost. To recover this information, one has to consider not just a single level m but all levels together.

Definition 3.2. The Berezin-Toeplitz quantization is the map

$$C^\infty(M) \longrightarrow \prod_{m \in \mathbb{N}_0} \mathrm{End}\Big(\Gamma_{\mathrm{hol}}\Big(M, L^{(m)}\Big)\Big), \qquad f \longrightarrow \Big(T_f^{(m)}\Big)_{m \in \mathbb{N}_0}. \tag{3.7}$$

We obtain a family of finite-dimensional (matrix) algebras and a family of maps. This infinite family should in some sense "approximate" the algebra $C^\infty(M)$.

3.2. *Approximation Results*

Denote for $f \in C^\infty(M)$ by $|f|_\infty$, the supnorm of f on M and by

$$\left\| T_f^{(m)} \right\| := \sup_{\substack{s \in \Gamma_{\mathrm{hol}}(M, L^m) \\ s \neq 0}} \frac{\left\| T_f^{(m)} s \right\|}{\|s\|}, \tag{3.8}$$

the operator norm with respect to the norm (3.2) on $\Gamma_{\mathrm{hol}}(M, L^m)$. The following theorem was shown in 1994.

Theorem 3.3 (Bordemann et al. [1]). (a) *For every $f \in C^\infty(M)$, there exists a $C > 0$ such that*

$$|f|_\infty - \frac{C}{m} \le \left\| T_f^{(m)} \right\| \le |f|_\infty. \tag{3.9}$$

In particular, $\lim_{m\to\infty} \| T_f^{(m)} \| = |f|_\infty$.

(b) *For every $f, g \in C^\infty(M)$,*

$$\left\| m\,\mathrm{i}\left[T_f^{(m)}, T_g^{(m)} \right] - T_{\{f,g\}}^{(m)} \right\| = O\left(\frac{1}{m} \right). \tag{3.10}$$

(c) *For every $f, g \in C^\infty(M)$,*

$$\left\| T_f^{(m)} T_g^{(m)} - T_{f\cdot g}^{(m)} \right\| = O\left(\frac{1}{m} \right). \tag{3.11}$$

These results are contained in Theorems 4.1, 4.2, and in Section 5 in [1]. We will indicate the proof for (b) and (c) in Section 5. It will make reference to the symbol calculus of generalised Toeplitz operators as developed by Boutet de Monvel and Guillemin [28]. The original proof of (a) was quite involved and required Hermite distributions and related objects. On the basis of the asymptotic expansion of the Berezin transform [29], a more direct proof can be given. I will discuss this in Section 7.3.

Only on the basis of this theorem, we are allowed to call our scheme a quantizing scheme. The properties in the theorem might be rephrased as *the BT operator quantization has the correct semiclassical limit*.

3.3. Further Properties

From Theorem 3.3 (c), we have the following proposition.

Proposition 3.4. *Let $f_1, f_2, \ldots, f_r \in C^\infty(M)$; then*

$$\left\| T_{f_1\cdots f_r}^{(m)} - T_{f_1}^{(m)} \cdots T_{f_r}^{(m)} \right\| = O\left(m^{-1} \right) \tag{3.12}$$

follows directly.

Proposition 3.5.

$$\lim_{m\to\infty} \left\| \left[T_f^{(m)}, T_g^{(m)} \right] \right\| = 0. \tag{3.13}$$

Proof. Using the left side of the triangle inequality, from Theorem 3.3 (b), it follows that

$$\left| m \left\| \left[T_f^{(m)}, T_g^{(m)} \right] \right\| - \left\| T_{\{f,g\}}^{(m)} \right\| \right| \leq \left\| m\,\mathrm{i}\left[T_f^{(m)}, T_g^{(m)} \right] - T_{\{f,g\}}^{(m)} \right\| = O\left(\frac{1}{m}\right). \tag{3.14}$$

By part (a) of the theorem $\|T_{\{f,g\}}^{(m)}\| \to |\{f,g\}|_\infty$, and it stays finite. Hence $\|[T_f^{(m)}, T_g^{(m)}]\|$ has to be a zero sequence. □

Proposition 3.6. *The Toeplitz map*

$$C^\infty(M) \longrightarrow \mathrm{End}\left(\Gamma_{\mathrm{hol}}\left(M, L^{(m)}\right)\right), \qquad f \longrightarrow T_f^{(m)}, \tag{3.15}$$

is surjective.

For a proof, see [1, Proposition 4.2].

This proposition says that for a fixed m every operator $A \in \mathrm{End}(\Gamma_{\mathrm{hol}}(M, L^{(m)}))$ is the Toeplitz operator of a function f_m. In the language of Berezin's co- and contravariant symbols, f_m will be the contravariant symbol of A. We will discuss this in Section 6.2.

Proposition 3.7. *For all $f \in C^\infty(M)$,*

$$T_f^{(m)*} = T_{\overline{f}}^{(m)}. \tag{3.16}$$

In particular, for real valued functions f the associated Toeplitz operator is self-adjoint.

Proof. Take $s, t \in \Gamma_{\mathrm{hol}}(M, L^m)$; then

$$\left\langle s, T_f^{(m)} t \right\rangle = \left\langle s, \Pi^{(m)}(f \cdot t) \right\rangle = \langle s, f \cdot t \rangle = \left\langle \overline{f} \cdot s, t \right\rangle = \left\langle T_{\overline{f}}^{(m)} s, t \right\rangle. \tag{3.17}$$

□

The opposite of the last statement of the above proposition is also true in the following sense.

Proposition 3.8. *Let $A \in \mathrm{End}(\Gamma_{\mathrm{hol}}(M, L^{(m)}))$ be a self-adjoint operator; then there exists a real valued function f, such that $A = T_f^{(m)}$.*

Proof. By the surjectivity of the Toeplitz map $A = T_f^{(m)}$ with a complex-valued function $f = f_0 + \mathrm{i}\, f_1$ with real functions f_0 and f_1. As $T_f^{(m)} = A = A^* = T_{\overline{f}}^{(m)}$, it follows that $T_{f-\overline{f}} = 0$ and hence $T_{f_1}^{(m)} = 0$. From this we conclude that $A = T_f^{(m)} = T_{f_1}^{(m)}$. □

We like to stress the fact that the Toeplitz map is never injective on a fixed level m. Only if $\|T_{f-g}^{(m)}\| \to 0$ for $m \to 0$, we can conclude that $f = g$.

Proposition 3.9. *Let $f \in C^\infty(M)$ and $n = \dim_{\mathbb{C}} M$. Denote the trace on $\mathrm{End}(\Gamma_{\mathrm{hol}}(M, L^m))$ by $\mathrm{Tr}^{(m)}$, then*

$$\mathrm{Tr}^{(m)}\left(T_f^{(m)}\right) = m^n \left(\frac{1}{\mathrm{vol}(\mathbb{P}^n(\mathbb{C}))} \int_M f\,\Omega + O\left(m^{-1}\right) \right). \tag{3.18}$$

See [1], respectively [30] for a detailed proof.

3.4. Strict Quantization

The asymptotic results of Theorem 3.3 say that the BT operator quantization is a strict quantization in the sense of Rieffel [31] as formulated in the book by Landsman [32]. We take as base space $X = \{0\} \cup \{1/m \mid m \in \mathbb{N}\}$, with its induced topology coming from $\mathbb{R}$. Note that $\{0\}$ is an accumulation point of the set $\{1/m \mid m \in \mathbb{N}\}$. As C^* algebras above the points $\{1/m\}$, we take the algebras $\mathrm{End}(\Gamma_{\mathrm{hol}}(M, L^{(m)}))$ and above $\{0\}$ the algebra $C^\infty(M)$. For $f \in C^\infty(M)$, we assign $0 \mapsto f$ and $1/m \mapsto T_f^{(m)}$. Now the property (a) in Theorem 3.3 is called in [32] Rieffel's condition, (b) Dirac's condition, and (c) von Neumann's condition. Completeness is true by Propositions 3.6 and 3.8.

This definition is closely related to the notion of continuous fields of C^*-algebras; see [32].

3.5. Relation to Geometric Quantization

There exists another quantum operator in the geometric setting, the operator of geometric quantization introduced by Kostant and Souriau. In a first step, the prequantum operator associated to the bundle L^m for the function $f \in C^\infty(M)$ is defined as

$$P_f^{(m)} := \nabla^{(m)}_{X_f^{(m)}} + \mathrm{i}\, f \cdot id. \tag{3.19}$$

Here $\nabla^{(m)}$ is the connection in L^m, and $X_f^{(m)}$ the Hamiltonian vector field of f with respect to the Kähler form $\omega^{(m)} = m \cdot \omega$, that is, $m\omega(X_f^{(m)}, \cdot) = df(\cdot)$. This operator $P_f^{(m)}$ acts on the space of differentiable global sections of the line bundle L^m. The sections depend at every point on $2n$ local coordinates and one has to restrict the space to sections covariantly constant along the excessive dimensions. In technical terms, one chooses a *polarization*. In general such a polarization is not unique. But in our complex situation, there is a canonical one by only taking the holomorphic sections. This polarization is called *Kähler polarization*. The operator of geometric quantization is then defined by the following proposition.

$$Q_f^{(m)} := \Pi^{(m)} P_f^{(m)}. \tag{3.20}$$

The Toeplitz operator and the operator of geometric quantization (with respect to the Kähler polarization) are related by the following.

Proposition 3.10 (Tuynman Lemma). *Let M be a compact quantizable Kähler manifold; then*

$$Q_f^{(m)} = \mathrm{i} \cdot T^{(m)}_{f-(1/2m)\Delta f}, \tag{3.21}$$

where Δ is the Laplacian with respect to the Kähler metric given by ω.

For the proof, see [33, 34] for a coordinate independent proof.

In particular, the $Q_f^{(m)}$ and the $T_f^{(m)}$ have the same asymptotic behaviour. We obtain for $Q_f^{(m)}$ similar results as in Theorem 3.3. For details see [35]. It should be noted that for (3.21) the compactness of M is essential.

3.6. L_α Approximation

In [34] the notion of L_α, respectively, $\mathrm{gl}(N)$, respectively, $\mathrm{su}(N)$ quasilimit were introduced. It was conjectured in [34] that for every compact quantizable Kähler manifold, the Poisson algebra of functions is a $\mathrm{gl}(N)$ quasilimit. In fact, the conjecture follows from Theorem 3.3; see [1, 35] for details.

3.7. The Noncompact Situation

Berezin-Toeplitz operators can be introduced for noncompact Kähler manifolds. In this case the L^2 spaces are the space of bounded sections and for the subspaces of holomorphic sections one can only consider the bounded holomorphic sections. Unfortunately, in this context the proofs of Theorem 3.3 do not work. One has to study examples or classes of examples case by case in order to see whether the corresponding properties are correct.

In the following, we give a very incomplete list of references. Berezin himself studied bounded complex-symmetric domains [36]. In this case the manifold is an open domain in $\mathbb{C}^n$. Instead of sections one studies functions which are integrable with respect to a suitable measure depending on $\hbar$. Then $1/\hbar$ corresponds to the tensor power of our bundle. Such Toeplitz operators were studied extensively by Upmeier in a series of works [37–40]. See also the book of Upmeier [41]. For $\mathbb{C}^n$ see Berger and Coburn [42, 43]. Klimek and Lesniewski [44, 45] studied the Berezin-Toeplitz quantization on the unit disc. Using automorphic forms and the universal covering, they obtain results for Riemann surfaces of genus $g \geq 2$. The names of Borthwick, Klimek, Lesniewski, Rinaldi, and Upmeier should be mentioned in the context of BT quantization for Cartan domains and super Hermitian spaces.

A quite different approach to Berezin-Toeplitz quantization is based on the asymptotic expansion of the Bergman kernel outside the diagonal. This was also used by the author together with Karabegov [29] for the compact Kähler case. See Section 7 for some details. Engliš [46] showed similar results for bounded pseudoconvex domains in $\mathbb{C}^N$. Ma and Marinescu [18, 19] developed a theory of Bergman kernels for the symplectic case, which yields also results on the Berezin-Toeplitz operators for certain noncompact Kähler manifolds and even orbifolds.

4. Berezin-Toeplitz Deformation Quantization

There is another approach to quantization. Instead of assigning noncommutative operators to commuting functions, one might think about “deforming” the pointwise commutative multiplication of functions into a noncommutative product. It is required to remain associative, the commutator of two elements should relate to the Poisson bracket of the elements, and it should reduce in the “classical limit” to the commutative situation.

It turns out that such a deformation which is valid for all differentiable functions cannot exist. A way out is to enlarge the algebra of functions by considering formal power series over them and to deform the product inside this bigger algebra. A first systematic treatment and applications in physics of this idea were given in 1978 by Bayen et al. [47, 48]. There the notion of *deformation quantization* and *star products* were introduced. Earlier versions of these concepts were around due to Berezin [49], Moyal [50], and Weyl [51]. For a presentation of the history, see [24].

We will show that for compact Kähler manifolds M, there is a natural star product.

4.1. Definition of Star Products

We start with a Poisson manifold $(M, \{\cdot,\cdot\})$, that is, a differentiable manifold with a Poisson bracket for the function such that $(C^\infty(M), \cdot, \{\cdot,\cdot\})$ is a Poisson algebra. Let $\mathcal{A} = C^\infty(M)[[\nu]]$ be the algebra of formal power series in the variable ν over the algebra $C^\infty(M)$.

Definition 4.1. A product $\star$ on $\mathcal{A}$ is called a (formal) star product for M (or for $C^\infty(M)$) if it is an associative $\mathbb{C}[[\nu]]$-linear product which is ν-adically continuous such that

(1) $\mathcal{A}/\nu\mathcal{A} \cong C^\infty(M)$, that is, $f \star g \bmod \nu = f \cdot g$,

(2) $(1/\nu)(f \star g - g \star f) \bmod \nu = -\mathrm{i}\{f, g\}$,

where $f, g \in C^\infty(M)$.

Alternatively, we can write

$$f \star g = \sum_{j=0}^{\infty} C_j(f, g)\nu^j, \tag{4.1}$$

with $C_j(f, g) \in C^\infty(M)$ such that the C_j are bilinear in the entries f and g. The conditions (1) and (2) can be reformulated as

$$C_0(f, g) = f \cdot g, \qquad C_1(f, g) - C_1(g, f) = -\mathrm{i}\{f, g\}. \tag{4.2}$$

By the ν-adic continuity, (4.1) fixes $\star$ on $\mathcal{A}$. A *(formal) deformation quantization* is given by a *(formal) star product*. I will use both terms interchangeable.

There are certain additional conditions for a star product which are sometimes useful.

(1) We call it “null on constants” if $1 \star f = f \star 1 = f$, which is equivalent to the fact that the constant function 1 will remain the unit in $\mathcal{A}$. In terms of the coefficients, it can be formulated as $C_k(f, 1) = C_k(1, f) = 0$ for $k \geq 1$. In this review, we always assume this to be the case for star products.

(2) We call it self-adjoint if $\overline{f \star g} = \overline{g} \star \overline{f}$, where we assume $\overline{\nu} = \nu$.

(3) We call it local if

$$\operatorname{supp} C_j(f, g) \subseteq \operatorname{supp} f \cap \operatorname{supp} g, \quad \forall f, g \in C^\infty(M). \tag{4.3}$$

From the locality property, it follows that the C_j are bidifferential operators and that the global star product defines for every open subset U of M a star product for the Poisson algebra $C^\infty(U)$. Such local star products are also called *differential star products*.

4.2. Existence of Star Products

In the usual setting of deformation theory, there always exists a trivial deformation. This is not the case here, as the trivial deformation of $C^\infty(M)$ to $\mathcal{A}$, which is nothing else as extending the point-wise product to the power series, is not allowed as it does not fulfil Condition (2) in Definition 4.1 (at least not if the Poisson bracket is nontrivial). In fact, the existence problem is highly nontrivial. In the symplectic case, different existence proofs, from different perspectives, were given by Marc De Wilde and Lecomte [52], Omori et al. [53, 54], and Fedosov [55, 56]. The general Poisson case was settled by Kontsevich [57].

4.3. Equivalence and Classification of Star Products

Definition 4.2. Given a Poisson manifold $(M, \{\cdot,\cdot\})$. Two star products $\star$ and $\star'$ associated to the Poisson structure $\{\cdot,\cdot\}$ are called equivalent if and only if there exists a formal series of linear operators

$$B = \sum_{i=0}^{\infty} B_i \nu^i, \quad B_i : C^\infty(M) \longrightarrow C^\infty(M), \tag{4.4}$$

with $B_0 = id$ such that

$$B(f) \star' B(g) = B(f \star g). \tag{4.5}$$

For local star products in the general Poisson setting, there are complete classification results. Here I will only consider the symplectic case.

To each local star product $\star$, its *Fedosov-Deligne class*

$$\mathrm{cl}(\star) \in \frac{1}{\mathrm{i}\,\nu}[\omega] + H^2_{\mathrm{dR}}(M)[[\nu]] \tag{4.6}$$

can be assigned. Here $H^2_{\mathrm{dR}}(M)$ denotes the 2nd deRham cohomology class of closed 2-forms modulo exact forms and $H^2_{\mathrm{dR}}(M)[[\nu]]$ the formal power series with such classes as coefficients. Such formal power series are called *formal deRham classes*. In general we will use $[\alpha]$ for the cohomology class of a form α.

This assignment gives a 1 : 1 correspondence between the formal deRham classes and the equivalence classes of star products.

For contractible manifolds, we have $H^2_{\mathrm{dR}}(M) = 0$ and hence there is up to equivalence exactly one local star product. This yields that locally all local star products of a manifold are equivalent to a certain fixed one, which is called the Moyal product. For these and related classification results, see [58–62].

4.4. Star Products with Separation of Variables

For our compact Kähler manifolds, we will have many different and even nonequivalent star products. The question is the following: is there a star product which is given in a natural way? The answer will be yes: the Berezin-Toeplitz star product to be introduced below. First we consider star products respecting the complex structure in a certain sense.

Definition 4.3 (Karabegov [63]). A star product is called *star product with separation of variables* if and only if

$$f \star h = f \cdot h, \qquad h \star g = h \cdot g, \tag{4.7}$$

for every locally defined holomorphic function g, antiholomorphic function f, and arbitrary function h.

Recall that a local star product $\star$ for M defines a star product for every open subset U of M. We have just to take the bidifferential operators defining $\star$. Hence it makes sense to talk about $\star$-multiplying with local functions.

Proposition 4.4. *A local $\star$ product has the separation of variables property if and only if in the bidifferential operators $C_k(\cdot,\cdot)$ for $k \geq 1$ in the first argument only derivatives in holomorphic and in the second argument only derivatives in antiholomorphic directions appear.*

In Karabegov's original notation the rôles of the holomorphic and antiholomorphic functions are switched. Bordemann and Waldmann [64] called such star products *star products of Wick type*. Both Karabegov and Bordemann-Waldmann proved that there exist for every Kähler manifold star products of separation of variables type. In Section 4.8, we will give more details on Karabegov's construction. Bordemann and Waldmann modified Fedosov's method [55, 56] to obtain such a star product. See also Reshetikhin and Takhtajan [65] for yet another construction. But I like to point out that in all these constructions the result is only a formal star product without any relation to an operator calculus, which will be given by the Berezin-Toeplitz star product introduced in the next section.

Another warning is in order. The property of being a star product of separation of variables type will not be kept by equivalence transformations.

4.5. Berezin-Toeplitz Star Product

Theorem 4.5. *There exists a unique (formal) star product $\star_{\mathrm{BT}}$ for M*

$$f \star_{\mathrm{BT}} g := \sum_{j=0}^{\infty} \nu^j C_j(f,g), \quad C_j(f,g) \in C^{\infty}(M), \tag{4.8}$$

in such a way that for $f, g \in C^{\infty}(M)$ and for every $N \in \mathbb{N}$ we have with suitable constants $K_N(f,g)$ for all m

$$\left\| T_f^{(m)} T_g^{(m)} - \sum_{0 \le j < N} \left(\frac{1}{m}\right)^j T_{C_j(f,g)}^{(m)} \right\| \le K_N(f,g) \left(\frac{1}{m}\right)^N. \tag{4.9}$$

The star product is null on constants and self-adjoint.

This theorem has been proven immediately after [1] was finished. It has been announced in [66, 67] and the proof was written up in German in [35]. A complete proof published in English can be found in [30].

For simplicity we might write

$$T_f^{(m)} \cdot T_g^{(m)} \sim \sum_{j=0}^{\infty} \left(\frac{1}{m}\right)^j T_{C_j(f,g)}^{(m)} \quad (m \longrightarrow \infty), \tag{4.10}$$

but we will always assume the strong and precise statement of (4.9). The same is assumed for other asymptotic formulas appearing further down in this review.

Next we want to identify this star product. Let K_M be the canonical line bundle of M, that is, the nth exterior power of the holomorphic 1-differentials. The canonical class δ is the first Chern class of this line bundle, that is, $\delta := c_1(K_M)$. If we take in K_M the fibre metric coming from the Liouville form Ω, then this defines a unique connection and further a unique curvature $(1,1)$-form ω_{can}. In our sign conventions, we have $\delta = [\omega_{\mathrm{can}}]$.

Together with Karabegov the author showed the following theorem.

Theorem 4.6 (see [29]). (a) *The Berezin-Toeplitz star product is a local star product which is of separation of variable type.*

(b) *Its classifying Deligne-Fedosov class is*

$$\mathrm{cl}(\star_{\mathrm{BT}}) = \frac{1}{\mathrm{i}} \left(\frac{1}{\nu}[\omega] - \frac{\delta}{2} \right) \tag{4.11}$$

for the characteristic class of the star product $\star_{BT}$.

(c) *The classifying Karabegov form associated to the Berezin-Toeplitz star product is*

$$-\frac{1}{\nu}\omega + \omega_{\mathrm{can}}. \tag{4.12}$$

The Karabegov form has not yet defined here. We will introduce it below in Section 4.8. Using K-theoretic methods, the formula for $\mathrm{cl}(\star_{\mathrm{BT}})$ was also given by Hawkins [68].

4.6. Star Product of Geometric Quantization

Tuynman's result (3.21) relates the operators of geometric quantization with Kähler polarization and the BT operators. As the latter define a star product, it can be used to give also a star product $\star_{\mathrm{GQ}}$ associated to geometric quantization. Details can be found in [30]. This star product will be equivalent to the BT star product, but it is not of the separation of variables type. The equivalence is given by the $\mathbb{C}[[\nu]]$-linear map induced by

$$B(f) := f - \nu\frac{\Delta}{2}f = \left(id - \nu\frac{\Delta}{2}\right)f. \tag{4.13}$$

We obtain $B(f)\star_{\mathrm{BT}}B(g) = B(f\star_{\mathrm{GQ}}g)$.

4.7. Trace for the BT Star Product

From (3.18) the following complete asymptotic expansion for $m \to \infty$ can be deduced [30, 69]:

$$\mathrm{Tr}^{(m)}\left(T_f^{(m)}\right) \sim m^n\left(\sum_{j=0}^{\infty}\left(\frac{1}{m}\right)^j \tau_j(f)\right), \quad \text{with } \tau_j(f) \in \mathbb{C}. \tag{4.14}$$

We define the $\mathbb{C}[[\nu]]$-linear map

$$\mathrm{Tr} : C^\infty(M)[[\nu]] \longrightarrow \nu^{-n}\mathbb{C}[[\nu]], \qquad \mathrm{Tr}\, f := \nu^{-n}\sum_{j=0}^{\infty}\nu^j\tau_j(f), \tag{4.15}$$

where the $\tau_j(f)$ are given by the asymptotic expansion (4.14) for $f \in C^\infty(M)$ and for arbitrary elements by $\mathbb{C}[[\nu]]$-linear extension.

Proposition 4.7 (see [30]). *The map* Tr *is a trace, that is, we have*

$$\mathrm{Tr}(f \star g) = \mathrm{Tr}(g \star f). \tag{4.16}$$

4.8. Karabegov Quantization

In [63, 70] Karabegov not only gave the notion of *separation of variables type*, but also a proof of existence of such formal star products for any Kähler manifold, whether compact, noncompact, quantizable, or nonquantizable. Moreover, he classified them completely as individual star product not only up to equivalence.

He starts with (M, ω_{-1}) a pseudo-Kähler manifold, that is, a complex manifold with a nondegenerate closed $(1,1)$-form not necessarily positive.

A formal form $\widehat{\omega} = (1/\nu)\omega_{-1} + \omega_0 + \nu\omega_1 + \cdots$ is called a formal deformation of the form $(1/\nu)\omega_{-1}$ if the forms ω_r, $r \geq 0$, are closed but not necessarily nondegenerate $(1,1)$-forms on M. It was shown in [63] that all deformation quantizations with separation of variables on the pseudo-Kähler manifold (M, ω_{-1}) are bijectively parameterized by the formal deformations of the form $(1/\nu)\omega_{-1}$.

Assume that we have such a star product $(\mathcal{A} := C^\infty(M)[[\nu]], \star)$. Then for $f, g \in \mathcal{A}$ the operators of left and right multiplications L_f, R_g are given by $L_f g = f \star g = R_g f$. The associativity of the star-product $\star$ is equivalent to the fact that L_f commutes with R_g for all $f, g \in \mathcal{A}$. If a star product is differential, then L_f, R_g are formal differential operators.

Karabegov constructs his star product associated to the deformation $\widehat{\omega}$ in the following way. First he chooses on every contractible coordinate chart $U \subset M$ (with holomorphic coordinates $\{z_k\}$) its formal potential

$$\widehat{\Phi} = \left(\frac{1}{\nu}\right)\Phi_{-1} + \Phi_0 + \nu\Phi_1 + \cdots, \qquad \widehat{\omega} = i\partial\overline{\partial}\widehat{\Phi}. \tag{4.17}$$

Then construction is done in such a way that we have for the left (right) multiplication operators on U

$$L_{\partial\Phi/\partial z_k} = \frac{\partial\Phi}{\partial z_k} + \frac{\partial}{\partial z_k}, \qquad R_{\partial\Phi/\partial\overline{z}_l} = \frac{\partial\Phi}{\partial\overline{z}_l} + \frac{\partial}{\partial\overline{z}_l}. \tag{4.18}$$

The set $\mathcal{L}(U)$ of all left multiplication operators on U is completely described as the set of all formal differential operators commuting with the point-wise multiplication operators by antiholomorphic coordinates $R_{\overline{z}_l} = \overline{z}_l$ and the operators $R_{\partial\Phi/\partial\overline{z}_l}$. From the knowledge of $\mathcal{L}(U)$, the star product on U can be reconstructed. The local star-products agree on the intersections of the charts and define the global star-product $\star$ on M.

We have to mention that this original construction of Karabegov will yield a star product of separation of variable type but with the role of holomorphic and antiholomorphic variables switched. This says for any open subset $U \subset M$ and any holomorphic function a and antiholomorphic function b on U that the operators L_a and R_b are the operators of point-wise multiplication by a and b, respectively, that is, $L_a = a$ and $R_b = b$.

4.9. Karabegov's Formal Berezin Transform

Given such a star products $\star$, Karabegov introduced the formal *Berezin transform I* as the unique formal differential operator on M such that for any open subset $U \subset M$, holomorphic functions a, and antiholomorphic functions b on U, the relation $I(a \cdot b) = b \star a$ holds (see [71]). He shows that $I = 1 + \nu\Delta + \cdots$, where Δ is the Laplace operator corresponding to the pseudo-Kähler metric on M.

Karabegov considered the following associated star products. First the *dual* star-product $\widetilde{\star}$ on M is defined for $f, g \in \mathcal{A}$ by the formula

$$f \,\widetilde{\star}\, g = I^{-1}(Ig \star If). \tag{4.19}$$

It is a star-product with separation of variables on the pseudo-Kähler manifold $(M, -\omega_{-1})$. Its formal Berezin transform equals I^{-1}, and thus the dual to $\widetilde{\star}$ is $\star$. Note that it is not a star product of the same pseudo-Kähler manifold. Denote by $\widetilde{\omega} = -(1/\nu)\omega_{-1} + \widetilde{\omega}_0 + \nu\widetilde{\omega}_1 + \cdots$ the formal form parameterizing the star-product $\widetilde{\star}$.

Next, the opposite of the dual star-product, $\star' = \widetilde{\star}^{\mathrm{op}}$, is given by the formula

$$f \star' g = I^{-1}(If \star Ig). \tag{4.20}$$

It defines a deformation quantization with separation of variables on M, but with the roles of holomorphic and antiholomorphic variables swapped—with respect to $\star$. It could be described also as a deformation quantization with separation of variables on the pseudo-Kähler manifold $(\overline{M}, \omega_{-1})$, where $\overline{M}$ is the manifold M with the opposite complex structure. But now the pseudo-Kähler form will be the same. Indeed the formal Berezin transform I establishes an equivalence of deformation quantizations $(\mathcal{A}, \star)$ and $(\mathcal{A}, \star')$.

How is the relation to the Berezin-Toeplitz star product $\star_{\mathrm{BT}}$ of Theorem 4.5? There exists a certain formal deformation $\widehat{\omega}$ of the form $(1/\nu)\omega$ which yields a star product $\star$ in the Karabegov sense. The opposite of its dual will be equal to the Berezin-Toeplitz star product, that is,

$$\star_{\mathrm{BT}} = \widetilde{\star}^{\mathrm{op}} = \star'. \tag{4.21}$$

The classifying Karabegov form $\widetilde{\omega}$ of $\widetilde{\star}$ will be the form (4.12). Note as $\star$ and $\star_{\mathrm{BT}}$ are equivalent via I, we have $\mathrm{cl}(\star) = \mathrm{cl}(\star_{\mathrm{BT}})$; see the formula (4.11). We will identify $\widehat{\omega}$ in Section 7.5.

5. The Disc Bundle and Global Operators

In this section, we identify the bundles L^m over the Kähler manifold M as associated line bundles of one unique S^1-bundle over M. The Toeplitz operator will appear as "modes" of a global Toeplitz operator. A detailed analysis of this global operator will yield a proof of Theorem 3.3 part (b) and part (c).

Moreover, we will need this set-up to discuss coherent states, Berezin symbols, and the Berezin transform in the next sections. For a more detailed presentation, see [35].

5.1. The Disc Bundle

We will assume that the quantum line bundle L is already very ample, that is, it has enough global holomorphic sections to embed M into projective space. From the bundle (as the connection ∇ will not be needed anymore, I will drop it in the notation) (L, h), we pass to its dual $(U, k) := (L^*, h^{-1})$ with dual metric k. Inside of the total space U, we consider the circle bundle

$$Q := \{\lambda \in U \mid k(\lambda, \lambda) = 1\}, \tag{5.1}$$

the (open) disc bundle, and (closed) disc bundle, respectively

$$D := \{\lambda \in U \mid k(\lambda,\lambda) < 1\}, \qquad \overline{D} := \{\lambda \in U \mid k(\lambda,\lambda) \le 1\}. \tag{5.2}$$

Let $\tau : U \to M$ be the projection (maybe restricted to the subbundles).

For the projective space $\mathbb{P}^N(\mathbb{C})$ with the hyperplane section bundle H as quantum line bundle, the bundle U is just the tautological bundle. Its fibre over the point $z \in \mathbb{P}^N(\mathbb{C})$ consists of the line in $\mathbb{C}^{N+1}$ which is represented by z. In particular, for the projective space the total space of U with the zero section removed can be identified with $\mathbb{C}^{N+1} \setminus \{0\}$. The same picture remains true for the via the very ample quantum line bundle in projective space embedded manifold M. The quantum line bundle will be the pull-back of H (i.e., its restriction to the embedded manifold) and its dual is the pull-back of the tautological bundle.

In the following we use $E \setminus 0$ to denote the total space of a vector bundle E with the image of the zero section removed. Starting from the real-valued function $\widehat{k}(\lambda) := k(\lambda,\lambda)$ on U, we define $\widetilde{a} := (1/2\mathrm{i})(\partial - \overline{\partial}) \log \widehat{k}$ on $U \setminus 0$ (the derivation is taken with respect to the complex structure on U) and denote by α its restriction to Q. With the help of the quantization condition (2.4), we obtain $d\alpha = \tau^* \omega$ (with the deRham differential $d = d_Q$) and that in fact $\mu = (1/2\pi)\tau^*\Omega \wedge \alpha$ is a volume form on Q. Indeed α is a contact form for the contact manifold Q. As far as the integration is concerned we get

$$\int_Q (\tau^* f)\mu = \int_M f\Omega, \quad \forall f \in C^\infty(M). \tag{5.3}$$

Recall that Ω is the Liouville volume form on M.

5.2. The Generalized Hardy Space

With respect to μ, we take the L^2-completion $\mathrm{L}^2(Q,\mu)$ of the space of functions on Q. The generalized *Hardy space* $\mathscr{H}$ is the closure of the space of those functions in $\mathrm{L}^2(Q,\mu)$ which can be extended to holomorphic functions on the whole disc bundle $\overline{D}$. The generalized *Szegö projector* is the projection

$$\Pi : \mathrm{L}^2(Q,\mu) \longrightarrow \mathscr{H}. \tag{5.4}$$

By the natural circle action, the bundle Q is an S^1-bundle and the tensor powers of U can be viewed as associated line bundles. The space $\mathscr{H}$ is preserved by the S^1-action. It can be decomposed into eigenspaces $\mathscr{H} = \prod_{m=0}^{\infty} \mathscr{H}^{(m)}$, where $c \in S^1$ acts on $\mathscr{H}^{(m)}$ as multiplication by c^m. The Szegö projector is S^1 invariant and can be decomposed into its components, the Bergman projectors

$$\widehat{\Pi}^{(m)} : \mathrm{L}^2(Q,\mu) \longrightarrow \mathscr{H}^{(m)}. \tag{5.5}$$

Sections of $L^m = U^{-m}$ can be identified with functions ψ on Q which satisfy the equivariance condition $\psi(c\lambda) = c^m \psi(\lambda)$, that is, which are homogeneous of degree m. This identification is given via the map

$$\gamma_m : \mathrm{L}^2(M, L^m) \longrightarrow \mathrm{L}^2(Q, \mu), \qquad s \longmapsto \psi_s, \quad \text{where } \psi_s(\alpha) = \alpha^{\otimes m}(s(\tau(\alpha))), \tag{5.6}$$

which turns out to be an isometry onto its image. On $\mathrm{L}^2(M, L^m)$, we have the scalar product (3.2). Restricted to the holomorphic sections, we obtain the isometry

$$\gamma_m : \Gamma_{\mathrm{hol}}(M, L^m) \cong \mathcal{H}^{(m)}. \tag{5.7}$$

In the case of $\mathbb{P}^N(\mathbb{C})$, this correspondence is nothing else as the identification of the global sections of the mth tensor powers of the hyper plane section bundle with the homogenous polynomial functions of degree m on $\mathbb{C}^{N+1}$.

5.3. The Toeplitz Structure

There is the notion of Toeplitz structure (Π, Σ) as developed by Boutet de Monvel and Guillemin in [28, 72]. I do not want to present the general theory but only the specialization to our situation. Here Π is the Szegö projector (5.4) and Σ is the submanifold

$$\Sigma := \{t\alpha(\lambda) \mid \lambda \in Q,\ t > 0\} \subset T^*Q \setminus 0 \tag{5.8}$$

of the tangent bundle of Q defined with the help of the 1-form α. It turns out that Σ is a symplectic submanifold, called a symplectic cone.

A (generalized) *Toeplitz operator* of order k is an operator $A : \mathcal{H} \to \mathcal{H}$ of the form $A = \Pi \cdot R \cdot \Pi$, where R is a pseudodifferential operator (ΨDO) of order k on Q. The Toeplitz operators constitute a ring. The symbol of A is the restriction of the principal symbol of R (which lives on T^*Q) to Σ. Note that R is not fixed by A, but Guillemin and Boutet de Monvel showed that the symbols are well defined and that they obey the same rules as the symbols of ΨDOs. In particular, the following relations are valid:

$$\sigma(A_1 A_2) = \sigma(A_1)\sigma(A_2), \qquad \sigma([A_1, A_2]) = \mathrm{i}\{\sigma(A_1), \sigma(A_2)\}_\Sigma. \tag{5.9}$$

Here $\{\cdot,\cdot\}_\Sigma$ is the restriction of the canonical Poisson structure of T^*Q to Σ coming from the canonical symplectic form on T^*Q.

5.4. A Sketch of the Proof of Theorem 3.3

For this we need only to consider the following two generalized Toeplitz operators.

(1) The generator of the circle action gives the operator $D_\varphi = (1/\mathrm{i})(\partial/\partial\varphi)$, where φ is the angular variable. It is an operator of order 1 with symbol t. It operates on $\mathcal{H}^{(m)}$ as multiplication by m.

(2) For $f \in C^\infty(M)$, let M_f be the operator on $\mathrm{L}^2(Q,\mu)$ corresponding to multiplication with $\tau^* f$. We set

$$T_f = \Pi \cdot M_f \cdot \Pi : \mathcal{H} \longrightarrow \mathcal{H}. \tag{5.10}$$

As M_f is constant along the fibres of τ, the operator T_f commutes with the circle action. Hence we can decompose

$$T_f = \prod_{m=0}^{\infty} T_f^{(m)}, \tag{5.11}$$

where $T_f^{(m)}$ denotes the restriction of T_f to $\mathcal{H}^{(m)}$. After the identification of $\mathcal{H}^{(m)}$ with $\Gamma_{\mathrm{hol}}(M, L^m)$, we see that these $T_f^{(m)}$ are exactly the Toeplitz operators $T_f^{(m)}$ introduced in Section 3. We call T_f the global Toeplitz operator and the $T_f^{(m)}$ the local Toeplitz operators. The operator T_f is of order 0. Let us denote by $\tau_\Sigma : \Sigma \subseteq T^*Q \to Q \to M$ the composition, then we obtain for its symbol $\sigma(T_f) = \tau_\Sigma^*(f)$.

Now we are able to prove (3.10). First we introduce for a fixed $t > 0$

$$\Sigma_t := \{ t \cdot \alpha(\lambda) \mid \lambda \in Q \} \subseteq \Sigma. \tag{5.12}$$

It turns out that $\omega_{\Sigma|\Sigma_t} = -t\tau_\Sigma^* \omega$. The commutator $[T_f, T_g]$ is a Toeplitz operator of order -1. From the above, we obtain with (5.9) that the symbol of the commutator equals

$$\sigma\big([T_f, T_g]\big)(t\alpha(\lambda)) = \mathrm{i}\{\tau_\Sigma^* f, \tau_\Sigma^* g\}_\Sigma(t\alpha(\lambda)) = -\mathrm{i}\, t^{-1}\{f, g\}_M(\tau(\lambda)). \tag{5.13}$$

We consider the Toeplitz operator

$$A := D_\varphi^2 [T_f, T_g] + \mathrm{i}\, D_\varphi T_{\{f,g\}}. \tag{5.14}$$

Formally this is an operator of order 1. Using $\sigma(T_{\{f,g\}}) = \tau_\Sigma^* \{f, g\}$ and $\sigma(D_\varphi) = t$, we see that its principal symbol vanishes. Hence it is an operator of order 0. Now M and hence also Q are compact manifolds. This implies that A is a bounded operator (ΨDOs of order 0 on compact manifolds are bounded). It is obviously S^1-invariant and we can write $A = \prod_{m=0}^\infty A^{(m)}$, where $A^{(m)}$ is the restriction of A on the space $\mathcal{H}^{(m)}$. For the norms we get $\|A^{(m)}\| \le \|A\|$. But

$$A^{(m)} = A_{|\mathcal{H}^{(m)}} = m^2 \Big[T_f^{(m)}, T_g^{(m)}\Big] + \mathrm{i}\, m T_{\{f,g\}}^{(m)}. \tag{5.15}$$

Taking the norm bound and dividing it by m, we get part (b) of Theorem 3.3. Using (5.7), the norms involved indeed coincide.

Quite similar, one can prove part (c) of Theorem 3.3 and more general the existence of the coefficients $C_j(f, g)$ for the Berezin-Toeplitz star product of Theorem 4.5. See [30, 35] for the details.

6. Coherent States and Berezin Symbols

6.1. Coherent States

Let the situation be as in the previous section. In particular, L is assumed to be already very ample, $U = L^*$ is the dual of the quantum line bundle, $Q \subset U$ the unit circle bundle, and $\tau : Q \to M$ the projection. In particular, recall the correspondence (5.6) $\psi_s(\alpha) = \alpha^{\otimes m}(s\tau(\alpha))$ of m-homogeneous functions ψ_s on U with sections of L^m. To obtain this correspondence, we fixed the section s and varied a.

Now we do the opposite. We fix $\alpha \in U \setminus 0$ and vary the section s. Obviously, this yields a linear form on $\Gamma_{\mathrm{hol}}(M, L^m)$ and hence with the help of the scalar product (3.2), we make the following.

Definition 6.1. (a) The *coherent vector (of level m)* associated to the point $\alpha \in U \setminus 0$ is the unique element $e_\alpha^{(m)}$ of $\Gamma_{\mathrm{hol}}(M, L^m)$ such that

$$\left\langle e_\alpha^{(m)}, s \right\rangle = \psi_s(\alpha) = \alpha^{\otimes m}(s(\tau(\alpha))) \tag{6.1}$$

for all $s \in \Gamma_{\mathrm{hol}}(M, L^m)$.

(b) The *coherent state (of level m)* associated to $x \in M$ is the projective class

$$\mathrm{e}_x^{(m)} := \left[e_\alpha^{(m)}\right] \in \mathbb{P}(\Gamma_{\mathrm{hol}}(M, L^m)), \quad \alpha \in \tau^{-1}(x),\ \alpha \neq 0. \tag{6.2}$$

Of course, we have to show that the object in (b) is well defined. Recall that $\langle \cdot, \cdot \rangle$ denotes the scalar product on the space of global sections $\Gamma_\infty(M, L^m)$. In the convention of this review, it will be antilinear in the first argument and linear in the second argument. The coherent vectors are antiholomorphic in α and fulfil

$$e_{c\alpha}^{(m)} = \overline{c}^m \cdot e_\alpha^{(m)}, \quad c \in \mathbb{C}^* := \mathbb{C} \setminus \{0\}. \tag{6.3}$$

Note that $e_\alpha^{(m)} \equiv 0$ would imply that all sections will vanish at the point $x = \tau(\alpha)$. Hence, the sections of L cannot be used to embed M into a projective space, which is a contradiction to the very ampleness of L. Hence, $e_\alpha^{(m)} \not\equiv 0$ and due to (6.3) the class

$$\left[e_\alpha^{(m)}\right] := \left\{ s \in \Gamma_{\mathrm{hol}}(M, L^m) \mid \exists c \in \mathbb{C}^* : s = c \cdot e_\alpha^{(m)} \right\} \tag{6.4}$$

is a well-defined element of the projective space $\mathbb{P}(\Gamma_{\mathrm{hol}}(M, L^m))$, only depending on $x = \tau(\alpha) \in M$.

This kind of coherent states goes back to Berezin. A coordinate independent version and extensions to line bundles were given by Rawnsley [73]. It plays an important role in the work of Cahen et al. on the quantization of Kähler manifolds [74–77], via Berezin's covariant symbols. I will return to this in Section 6.5. In these works, the coherent vectors are parameterized by the elements of $L \setminus 0$. The definition here uses the points of the total space of the dual bundle U. It has the advantage that one can consider all tensor powers of L together on an equal footing.

Definition 6.2. The *coherent state embedding* is the antiholomorphic embedding

$$M \longrightarrow \mathbb{P}(\Gamma_{\text{hol}}(M, L^m)) \cong \mathbb{P}^N(\mathbb{C}), \qquad x \longmapsto \left[e^{(m)}_{\tau^{-1}(x)}\right]. \tag{6.5}$$

Here $N = \dim \Gamma_{\text{hol}}(M, L^m) - 1$. In this review, in abuse of notation, $\tau^{-1}(x)$ will always denote a non-zero element of the fiber over x. The coherent state embedding is up to conjugation the embedding of Section 2.9 with respect to an orthonormal basis of the sections. In [78] further results on the geometry of the coherent state embedding are given.

6.2. Covariant Berezin Symbols

We start with the following definition.

Definition 6.3. The *covariant Berezin symbol* $\sigma^{(m)}(A)$ *(of level* m*)* of an operator $A \in \text{End}(\Gamma_{\text{hol}}(M, L^{(m)}))$ is defined as

$$\sigma^{(m)}(A) : M \longrightarrow \mathbb{C}, \qquad x \longmapsto \sigma^{(m)}(A)(x) := \frac{\left\langle e^{(m)}_\alpha, Ae^{(m)}_\alpha \right\rangle}{\left\langle e^{(m)}_\alpha, e^{(m)}_\alpha \right\rangle}, \quad \alpha \in \tau^{-1}(x). \tag{6.6}$$

As the factors appearing in (6.3) will cancel, it is a well-defined function on M. If the level m is clear from the context, I will sometimes drop it in the notation.

We consider also the *coherent projectors* used by Rawnsley

$$P^{(m)}_x = \frac{\left|e^{(m)}_\alpha\right\rangle\left\langle e^{(m)}_\alpha\right|}{\left\langle e^{(m)}_\alpha, e^{(m)}_\alpha \right\rangle}, \quad \alpha \in \tau^{-1}(x). \tag{6.7}$$

Here we used the convenient bra-ket notation of the physicists. Recall, if s is a section, then

$$P^{(m)}_x s = \frac{\left|e^{(m)}_\alpha\right\rangle\left\langle e^{(m)}_\alpha, s\right\rangle}{\left\langle e^{(m)}_\alpha, e^{(m)}_\alpha \right\rangle} = \frac{\left\langle e^{(m)}_\alpha, s\right\rangle}{\left\langle e^{(m)}_\alpha, e^{(m)}_\alpha \right\rangle} e^{(m)}_\alpha. \tag{6.8}$$

Again the projector is well defined on M. With its help, the covariant symbol can be expressed as

$$\sigma^{(m)}(A) = \text{Tr}\left(AP^{(m)}_x\right). \tag{6.9}$$

From the definition of the symbol, it follows that $\sigma^{(m)}(A)$ is real analytic and

$$\sigma^{(m)}(A^*) = \overline{\sigma^{(m)}(A)}. \tag{6.10}$$

6.3. Rawnsley's ϵ Function

Rawnsley [73] introduced a very helpful function on the manifold M relating the local metric in the bundle with the scalar product on coherent states. In our dual description, we define it in the following way.

Definition 6.4. Rawnsley's epsilon function is the function

$$M \longrightarrow C^\infty(M), \qquad x \longmapsto \epsilon^{(m)}(x) := \frac{h^{(m)}\left(e_\alpha^{(m)}, e_\alpha^{(m)}\right)(x)}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle}, \quad \alpha \in \tau^{-1}(x). \tag{6.11}$$

With (6.3), it is clear that it is a well-defined function on M. Furthermore, using (6.1)

$$0 \neq \left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle = \alpha^{\otimes m}\left(e_\alpha^{(m)}(\tau(\alpha))\right), \tag{6.12}$$

it follows that

$$e_\alpha^{(m)}(x) \neq 0, \quad \text{for } x = \tau(\alpha), \text{ and } \epsilon^{(m)} > 0. \tag{6.13}$$

Hence, we can define the modified measure

$$\Omega_\epsilon^{(m)}(x) := \epsilon^{(m)}(x)\Omega(x) \tag{6.14}$$

for the space of functions on M and obtain a modified scalar product $\langle \cdot, \cdot \rangle_\epsilon^{(m)}$ for $C^\infty(M)$.

Proposition 6.5. *For $s_1, s_2 \in \Gamma_{\mathrm{hol}}(M, L^m)$, we have*

$$\begin{aligned} h^{(m)}(s_1, s_2)(x) &= \frac{\overline{\left\langle e_\alpha^{(m)}, s_1 \right\rangle} \left\langle e_\alpha^{(m)}, s_2 \right\rangle}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle} \cdot \epsilon^{(m)}(x) \\ &= \left\langle s_1, P_x^{(m)} s_2 \right\rangle \cdot \epsilon^{(m)}(x). \end{aligned} \tag{6.15}$$

Proof. Due to (6.13), we can represent every section s locally at x as $s(x) = \widehat{s}(x) e_\alpha^{(m)}$ with a local function $\widehat{s}$. Now

$$\left\langle e_\alpha^{(m)}, s \right\rangle = \alpha^{(m)}\left(\widehat{s}(x) e_\alpha^{(m)}(x)\right) = \widehat{s}(x)\alpha^{(m)}\left(e_\alpha^{(m)}(x)\right) = \widehat{s}(x)\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle. \tag{6.16}$$

We rewrite $h^{(m)}(s_1, s_2)(x) = \overline{\widehat{s}_1} \widehat{s}_2 h^{(m)}(e_\alpha^{(m)}, e_\alpha^{(m)})(x)$, and obtain

$$h^{(m)}(s_1, s_2)(x) = \frac{\overline{\left\langle e_\alpha^{(m)}, s_1 \right\rangle}}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle} \frac{\left\langle e_\alpha^{(m)}, s_2 \right\rangle}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle} \cdot h^{(m)}\left(e_\alpha^{(m)}, e_\alpha^{(m)}\right)(x). \tag{6.17}$$

From the definition (6.11), the first relation follows. Obviously, it can be rewritten with the coherent projector to obtain the second relation. □

There exists another useful description of the epsilon function.

Proposition 6.6. *Let $s_1, s_2, \ldots, s_k$ be an arbitrary orthonormal basis of $\Gamma_{\mathrm{hol}}(M, L^m)$. Then*

$$\epsilon^{(m)}(x) = \sum_{j=1}^{k} h^{(m)}(s_j, s_j)(x). \tag{6.18}$$

Proof. For every vector ψ in a finite-dimensional hermitian vector space with orthonormal basis s_j, $j = 1, \ldots, k$, the coefficient with respect to the basis element s_j is given by $\psi_j = \langle s_j, \psi \rangle$. Furthermore, $\langle \psi, \psi \rangle = \|\psi\|^2 = \sum_j \overline{\psi_j} \psi_j$. Using the relation (6.15) we can rewrite

$$\sum_{j=1}^{k} h(s_j, s_j)(x) = \frac{\epsilon^{(m)}(x)}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle} \sum_{j=1}^{k} \overline{\left\langle e_\alpha^{(m)}, s_j \right\rangle} \left\langle e_\alpha^{(m)}, s_j \right\rangle. \tag{6.19}$$

Hence the claim follows. □

In certain special cases, the functions $\epsilon^{(m)}$ will be constant as a function of the points of the manifold. In this case, we can apply Proposition 6.11 below for $A = id$, the identity operator, and obtain

$$\epsilon^{(m)} = \frac{\dim \Gamma_{\mathrm{hol}}(M, L^m)}{\mathrm{vol}\, M}. \tag{6.20}$$

Here $\mathrm{vol}\, M$ denotes the volume of the manifold with respect to the Liouville measure. Now the question arises when $\epsilon^{(m)}$ will be constant, respectively, when the measure $\Omega_\epsilon^{(m)}$ will be the standard measure (up to a scalar). From Proposition 6.6, it follows that if there is a transitive group action on the manifold and everything, for example, Kähler form, bundle, metric, is homogenous with respect to the action this will be the case. An example is given by $M = \mathbb{P}^N(\mathbb{C})$. By a result of Rawnsley [73], respectively, Cahen et al. [74], $\epsilon^{(m)} \equiv \mathrm{const}$ if and only if the quantization is projectively induced. This means that under the conjugate of the coherent state embedding, the Kähler form ω of M coincides with the pull-back of the Fubini-Study form. Note that in general this is not the case; see Section 7.6.

6.4. Contravariant Berezin Symbols

Recall the modified Liouville measure (6.14) and modified scalar product for the functions on M introduced in the last subsection.

Definition 6.7. Given an operator $A \in \mathrm{End}(\Gamma_{\mathrm{hol}}(M, L^{(m)}))$, then a *contravariant Berezin symbol* $\check{\sigma}^{(m)}(A) \in C^\infty(M)$ of A is defined by the representation of the operator A as integral

$$A = \int_M \check{\sigma}^{(m)}(A)(x) P_x^{(m)} \Omega_\epsilon^{(m)}(x) \tag{6.21}$$

if such a representation exists.

Proposition 6.8. *The Toeplitz operator $T_f^{(m)}$ admits a representation* (6.21) *with*

$$\check{\sigma}^{(m)}\left(T_f^{(m)}\right) = f, \tag{6.22}$$

that is, the function f is a contravariant symbol of the Toeplitz operator $T_f^{(m)}$. Moreover, every operator $A \in \mathrm{End}(\Gamma_{\mathrm{hol}}(M, L^{(m)}))$, has a contravariant symbol.

Proof. Let $f \in C^\infty(M)$ and set

$$A := \int_M f(x) P_x^{(m)} \Omega_\epsilon^{(m)}(x), \tag{6.23}$$

then $\check{\sigma}^{(m)}(A) = f$. For arbitrary $s_1, s_2 \in \Gamma_{\mathrm{hol}}(M, L^m)$, we calculate (using (6.15))

$$\begin{aligned}
\langle s_1, A s_2 \rangle &= \int_M f(x) \left\langle s_1, P_x^{(m)} s_2 \right\rangle \Omega_\epsilon^{(m)}(x) \\
&= \int_M f(x) h^{(m)}(s_1, s_2)(x) \Omega(x) \\
&= \int_M h^{(m)}(s_1, f s_2)(x) \Omega(x) \\
&= \langle s_1, f s_2 \rangle = \left\langle s_1, T_f^{(m)} s_2 \right\rangle.
\end{aligned} \tag{6.24}$$

Hence $T_f^{(m)} = A$. As the Toeplitz map is surjective (Proposition 3.6), every operator is a Toeplitz operator, hence has a contravariant symbol. □

Note that given an operator its contravariant symbol on a fixed level m is not uniquely defined.

We introduce on $\mathrm{End}(\Gamma_{\mathrm{hol}}(M, L^{(m)}))$ the Hilbert-Schmidt norm

$$\langle A, C \rangle_{\mathrm{HS}} = \mathrm{Tr}(A^* \cdot C). \tag{6.25}$$

Theorem 6.9. *The Toeplitz map $f \to T_f^{(m)}$ and the covariant symbol map $A \to \sigma^{(m)}(A)$ are adjoint:*

$$\left\langle A, T_f^{(m)} \right\rangle_{\mathrm{HS}} = \left\langle \sigma^{(m)}(A), f \right\rangle_\epsilon^{(m)}. \tag{6.26}$$

Proof.

$$\left\langle A, T_f^{(m)} \right\rangle = \operatorname{Tr}\left(A^* \cdot T_f^{(m)}\right) = \operatorname{Tr}\left(A^* \int_M f(x) P_x^{(m)} \Omega_\epsilon^{(m)}(x)\right) = \int_M f(x) \operatorname{Tr}\left(A^* \cdot P_x^{(m)}\right) \Omega_\epsilon^{(m)}(x). \tag{6.27}$$

Now applying Definition 6.7 and (6.10)

$$\left\langle A, T_f^{(m)} \right\rangle = \int_M f(x) \sigma^{(m)}(A^*) \Omega_\epsilon^{(m)}(x) = \int_M \overline{\sigma^{(m)}(A)}(x) f(x) \Omega_\epsilon^{(m)}(x) = \left\langle \sigma^{(m)}(A), f(x) \right\rangle_\epsilon^{(m)}. \tag{6.28}$$

□

As every operator has a contravariant symbol, we can also conclude

$$\langle A, B \rangle_{\mathrm{HS}} = \left\langle \sigma^{(m)}(A), \check{\sigma}^{(m)}(B) \right\rangle_\epsilon^{(m)}. \tag{6.29}$$

From Theorem 6.9 by using the surjectivity of the Toeplitz map, we get the following proposition.

Proposition 6.10. *The covariant symbol map $\sigma^{(m)}$ is injective.*

Another application is the following.

Proposition 6.11.

$$\operatorname{Tr} A = \int_M \sigma^{(m)}(A) \Omega_\epsilon^{(m)}. \tag{6.30}$$

Proof. We use $\mathrm{Id} = T_1$ and by (6.26) $\operatorname{Tr} A = \langle A, \mathrm{Id} \rangle_{\mathrm{HS}} = \langle \sigma^{(m)}(A), 1 \rangle_\epsilon^{(m)}$. □

6.5. *Berezin Star Product*

Under certain very restrictive conditions, Berezin covariant symbols can be used to construct a star product, called the *Berezin star product*. Recall that Proposition 6.10 says that the linear symbol map

$$\sigma^{(m)} : \operatorname{End}\left(\Gamma_{\mathrm{hol}}\left(M, L^{(m)}\right)\right) \longrightarrow C^\infty(M) \tag{6.31}$$

is injective. Its image is a subspace $\mathcal{A}^{(m)}$ of $C^\infty(M)$, called the subspace of covariant symbols of level m. If $\sigma^{(m)}(A)$ and $\sigma^{(m)}(B)$ are elements of this subspace the operators, A and B will be uniquely fixed. Hence also $\sigma^{(m)}(A \cdot B)$. Now one takes

$$\sigma^{(m)}(A) \star_{(m)} \sigma^{(m)}(B) := \sigma^{(m)}(A \cdot B) \tag{6.32}$$

as a definition for an associative and noncommutative product $\star_{(m)}$ on $\mathcal{A}^{(m)}$.

It is even possible to give an analytic expression for the resulting symbol. For this we introduce the *two-point function*

$$\psi^{(m)}(x, y) = \frac{\left\langle e_\alpha^{(m)}, e_\beta^{(m)} \right\rangle \left\langle e_\beta^{(m)}, e_\alpha^{(m)} \right\rangle}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle \left\langle e_\beta^{(m)}, e_\beta^{(m)} \right\rangle} \tag{6.33}$$

with $\alpha = \tau^{-1}(x) = x$ and $\beta = \tau^{-1}(y)$. This function is well defined on $M \times M$. Furthermore, we have the *two-point symbol*

$$\sigma^{(m)}(A)(x, y) = \frac{\left\langle e_\alpha^{(m)}, A e_\beta^{(m)} \right\rangle}{\left\langle e_\alpha^{(m)}, e_\beta^{(m)} \right\rangle}. \tag{6.34}$$

It is the analytic extension of the real-analytic covariant symbol. It is well defined on an open dense subset of $M \times M$ containing the diagonal. Using (6.15), we express

$$\begin{aligned}
\sigma^{(m)}(A \cdot B)(x) &= \frac{\left\langle e_\alpha^{(m)}, A \cdot B e_\alpha^{(m)} \right\rangle}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle} \\
&= \frac{\left\langle A^* e_\alpha^{(m)}, B e_\alpha^{(m)} \right\rangle}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle} \\
&= \int_M h^{(m)}\left(A^* e_\alpha^{(m)}, B e_\alpha^{(m)}\right)(y) \frac{\Omega(y)}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle} \\
&= \int_M \frac{\left\langle e_\alpha^{(m)}, A e_\beta^{(m)} \right\rangle \left\langle e_\beta^{(m)}, B e_\alpha^{(m)} \right\rangle}{\left\langle e_\beta^{(m)}, e_\beta^{(m)} \right\rangle} \frac{\epsilon^{(m)}(y)\Omega(y)}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle} \\
&= \int_M \sigma^{(m)}(A)(x, y) \sigma^{(m)}(B)(y, x) \cdot \psi^{(m)}(x, y) \cdot \epsilon^{(m)}(y)\Omega(y).
\end{aligned} \tag{6.35}$$

The crucial problem is how to relate different levels m to define for all possible symbols a unique product not depending on m. In certain special situations like these studied by Berezin himself [36] and Cahen et al. [74], the subspaces are nested into each other and the union $\mathcal{A} = \bigcup_{m \in \mathbb{N}} \mathcal{A}^{(m)}$ is a dense subalgebra of $C^\infty(M)$. Indeed, in the cases considered, the manifold is a homogenous manifold and the epsilon function $\epsilon^{(m)}$ is a constant. A detailed analysis shows that in this case a star product is given.

For further examples, for which this method works (not necessarily compact), see other articles by Cahen et al. [75–77]. For related results, see also work of Moreno and Ortega-Navarro [79, 80]. In particular, also the work of Engliš [46, 81–83]. Reshetikhin and Takhtajan [65] gave a construction of a (formal) star product using formal integrals in the spirit of the Berezin's covariant symbol construction.

7. Berezin Transform

7.1. The Definition

Starting from $f \in C^\infty(M)$, we can assign to it its Toeplitz operator $T_f^{(m)} \in \mathrm{End}(\Gamma_{\mathrm{hol}}(M, L^{(m)}))$ and then assign to $T_f^{(m)}$ the covariant symbol $\sigma^{(m)}(T_f^{(m)})$. It is again an element of $C^\infty(M)$.

Definition 7.1. The map

$$C^\infty(M) \longrightarrow C^\infty(M), \qquad f \longmapsto I^{(m)}(f) := \sigma^{(m)}\left(T_f^{(m)}\right) \tag{7.1}$$

is called the *Berezin transform (of level m)*.

From the point of view of Berezin's approach, the operator $T_f^{(m)}$ has as a contravariant symbol f. Hence $I^{(m)}$ gives a correspondence between contravariant symbols and covariant symbols of operators. The Berezin transform was introduced and studied by Berezin [36] for certain classical symmetric domains in $\mathbb{C}^n$. These results were extended by Unterberger and Upmeier [84]; see also Engliš [46, 81, 82] and Engliš and Peetre [85]. Obviously, the Berezin transform makes perfect sense in the compact Kähler case which we consider here.

7.2. The Asymptotic Expansion

The results presented here are joint work with Karabegov [29]. See also [86] for an overview.

Theorem 7.2. *Given $x \in M$, then the Berezin transform $I^{(m)}(f)$ evaluated at the point x has a complete asymptotic expansion in powers of $1/m$ as $m \to \infty$*

$$I^{(m)}(f)(x) \sim \sum_{i=0}^{\infty} I_i(f)(x) \frac{1}{m^i}, \tag{7.2}$$

where $I_i : C^\infty(M) \to C^\infty(M)$ are maps with

$$I_0(f) = f, \qquad I_1(f) = \Delta f. \tag{7.3}$$

Here the Δ is the usual Laplacian with respect to the metric given by the Kähler form ω.

Complete asymptotic expansion means the following. Given $f \in C^\infty(M)$, $x \in M$, and an $r \in \mathbb{N}$, then there exists a positive constant A such that

$$\left| I^{(m)}(f)(x) - \sum_{i=0}^{r-1} I_i(f)(x) \frac{1}{m^i} \right|_\infty \le \frac{A}{m^r}. \tag{7.4}$$

In Section 7.4, I will give some remarks on the proof but before I present you a nice application.

7.3. Norm Preservation of the BT Operators

In [87] I conjectured (7.2) (which is now a mathematical result) and showed how such an asymptotic expansion supplies a different proof of Theorem 3.3, part (a). For completeness, I reproduce the proof here.

Proposition 7.3.

$$\left| I^{(m)}(f) \right|_\infty = \left| \sigma^{(m)}\left(T_f^{(m)}\right) \right|_\infty \le \left\| T_f^{(m)} \right\| \le |f|_\infty. \tag{7.5}$$

Proof. Using Cauchy-Schwarz inequality, we calculate ($x = \tau(\alpha)$)

$$\left| \sigma^{(m)}\left(T_f^{(m)}\right)(x) \right|^2 = \frac{\left| \left\langle e_\alpha^{(m)}, T_f^{(m)} e_\alpha^{(m)} \right\rangle \right|^2}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle^2} \le \frac{\left\langle T_f^{(m)} e_\alpha^{(m)}, T_f^{(m)} e_\alpha^{(m)} \right\rangle}{\left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle} \le \left\| T_f^{(m)} \right\|^2. \tag{7.6}$$

Here the last inequality follows from the definition of the operator norm. This shows the first inequality in (7.5). For the second inequality, introduce the multiplication operator $M_f^{(m)}$ on $\Gamma_\infty(M, L^m)$. Then $\|T_f^{(m)}\| = \|\Pi^{(m)} M_f^{(m)} \Pi^{(m)}\| \le \|M_f^{(m)}\|$ and for $\varphi \in \Gamma_\infty(M, L^m)$, $\varphi \ne 0$

$$\frac{\left\| M_f^{(m)} \varphi \right\|^2}{\|\varphi\|^2} = \frac{\int_M h^{(m)}(f\varphi, f\varphi)\Omega}{\int_M h^{(m)}(\varphi, \varphi)\Omega} = \frac{\int_M f(z)\overline{f(z)} h^{(m)}(\varphi, \varphi)\Omega}{\int_M h^{(m)}(\varphi, \varphi)\Omega} \le |f|_\infty^2. \tag{7.7}$$

Hence,

$$\left\| T_f^{(m)} \right\| \le \left\| M_f^{(m)} \right\| = \sup_{\varphi \ne 0} \frac{\left\| M_f^{(m)} \varphi \right\|}{\|\varphi\|} \le |f|_\infty. \tag{7.8}$$

□

Proof (Theorem 3.3 part (a)). Choose as $x_e \in M$ a point with $|f(x_e)| = |f|_\infty$. From the fact that the Berezin transform has as a leading term the identity, it follows that $|(I^{(m)} f)(x_e) - f(x_e)| \le C/m$ with a suitable constant C. Hence, $||f(x_e)| - |(I^{(m)} f)(x_e)|| \le C/m$ and

$$|f|_\infty - \frac{C}{m} = |f(x_e)| - \frac{C}{m} \le \left| \left(I^{(m)} f \right)(x_e) \right| \le \left| I^{(m)} f \right|_\infty. \tag{7.9}$$

Putting (7.5) and (7.9) together, we obtain

$$|f|_\infty - \frac{C}{m} \le \left\| T_f^{(m)} \right\| \le |f|_\infty. \tag{7.10}$$

□

7.4. Bergman Kernel

To understand the Berezin transform better, we have to study the Bergman kernel. Recall from Section 5, the Szegö projectors $\Pi : \mathrm{L}^2(Q,\mu) \to \mathcal{H}$ and its components $\widehat{\Pi}^{(m)} : \mathrm{L}^2(Q,\mu) \to \mathcal{H}^{(m)}$, the Bergman projectors. The Bergman projectors have smooth integral kernels, the *Bergman kernels* $\mathcal{B}_m(\alpha,\beta)$ defined on $Q \times Q$, that is,

$$\widehat{\Pi}^{(m)}(\psi)(\alpha) = \int_Q \mathcal{B}_m(\alpha,\beta)\psi(\beta)\mu(\beta). \tag{7.11}$$

The Bergman kernels can be expressed with the help of the coherent vectors.

Proposition 7.4.

$$\mathcal{B}_m(\alpha,\beta) = \psi_{e_\beta^{(m)}}(\alpha) = \overline{\psi_{e_\alpha^{(m)}}(\beta)} = \left\langle e_\alpha^{(m)}, e_\beta^{(m)} \right\rangle. \tag{7.12}$$

For the proofs of this and the following propositions, see [29] or [86].

Let $x, y \in M$ and choose $\alpha, \beta \in Q$ with $\tau(\alpha) = x$ and $\tau(\beta) = y$, then the functions

$$u_m(x) := \mathcal{B}_m(\alpha,\alpha) = \left\langle e_\alpha^{(m)}, e_\alpha^{(m)} \right\rangle, \tag{7.13}$$

$$v_m(x,y) := \mathcal{B}_m(\alpha,\beta) \cdot \mathcal{B}_m(\beta,\alpha) = \left\langle e_\alpha^{(m)}, e_\beta^{(m)} \right\rangle \cdot \left\langle e_\beta^{(m)}, e_\alpha^{(m)} \right\rangle \tag{7.14}$$

are well defined on M and on $M \times M$, respectively. The following proposition gives an integral representation of the Berezin transform.

Proposition 7.5.

$$\begin{aligned} \left(I^{(m)}(f)\right)(x) &= \frac{1}{\mathcal{B}_m(\alpha,\alpha)} \int_Q \mathcal{B}_m(\alpha,\beta)\mathcal{B}_m(\beta,\alpha)\tau^* f(\beta)\mu(\beta) \\ &= \frac{1}{u_m(x)} \int_M v_m(x,y) f(y)\Omega(y). \end{aligned} \tag{7.15}$$

Typically, asymptotic expansions can be obtained using stationary phase integrals. But for such an asymptotic expansion of the integral representation of the Berezin transform, we will not only need an asymptotic expansion of the Bergman kernel along the diagonal (which is well known) but in a neighbourhood of it. This is one of the key results obtained in [29]. It is based on works of Boutet de Monvel and Sjöstrand [88] on the Szegö kernel and in generalization of a result of Zelditch [89] on the Bergman kernel on the diagonal. The integral representation is used then to prove the existence of the asymptotic expansion of the Berezin transform.

Having such an asymptotic expansion, it still remains to identify its terms. As it was explaining in Section 4.8, Karabegov assigns to every formal deformation quantizations with the "separation of variables" property a *formal Berezin transform* I. In [29] it is shown that there is an explicitly specified star product $\star$ (see [29, Theorem 5.9]) with associated formal Berezin transform such that if we replace $1/m$ by the formal variable ν in the asymptotic expansion of the Berezin transform $I^{(m)}f(x)$ we obtain $I(f)(x)$. This finally proves Theorem 7.2. We will exhibit the star product $\star$ in the next section.

7.5. Identification of the BT Star Product

Moreover in [29] there is another object introduced, the *twisted product*

$$R^{(m)}(f,g) := \sigma^{(m)}\left(T_f^{(m)} \cdot T_g^{(m)}\right). \tag{7.16}$$

Also for it the existence of a complete asymptotic expansion was shown. It was identified with a twisted formal product. This allows the identification of the BT star product with a special star product within the classification of Karabegov. From this identification, the properties of Theorem 4.6 of locality, separation of variables type, and the calculation to the classifying forms and classes for the BT star product follow.

As already announced in Section 4.8, the BT star product $\star_{BT}$ is the opposite of the dual star product of a certain star product $\star$. To identify $\star$ we will give its classifying Karabegov form $\widehat{\omega}$. As already mentioned above, Zelditch [89] proved that the the function u_m (7.13) has a complete asymptotic expansion in powers of $1/m$. In detail he showed

$$u_m(x) \sim m^n \sum_{k=0}^{\infty} \frac{1}{m^k} b_k(x), \quad b_0 = 1. \tag{7.17}$$

If we replace in the expansion $1/m$ by the formal variable ν, we obtain a formal function s defined by

$$e^s(x) = \sum_{k=0}^{\infty} \nu^k b_k(x). \tag{7.18}$$

Now take as formal potential (4.17)

$$\widehat{\Phi} = \frac{1}{\nu}\Phi_{-1} + s, \tag{7.19}$$

where Φ_{-1} is the local Kähler potential of the Kähler form $\omega = \omega_{-1}$. Then $\widehat{\omega} = \mathrm{i}\,\partial\bar{\partial}\widehat{\Phi}$. It might be also written in the form

$$\widehat{\omega} = \frac{1}{\nu}\omega + \mathbb{F}\left(\mathrm{i}\,\partial\bar{\partial}\log \mathcal{B}_m(\alpha,\alpha)\right). \tag{7.20}$$

Here we denote the replacement of $1/m$ by the formal variable ν by the symbol $\mathbb{F}$.

7.6. Pullback of the Fubini-Study Form

Starting from the Kähler manifold (M,ω) and after choosing an orthonormal basis of the space $\Gamma_{\mathrm{hol}}(M,L^m)$, we obtain an embedding

$$\phi^{(m)} : M \longrightarrow \mathbb{P}^{N(m)} \tag{7.21}$$

of M into projective space of dimension $N(m)$. On $\mathbb{P}^{N(m)}$ we have the standard Kähler form, the Fubini-Study form ω_{FS}. The pull-back $(\phi^{(m)})^*\omega_{\mathrm{FS}}$ will not depend on the orthogonal basis chosen for the embedding. But in general it will not coincide with a scalar multiple of the Kähler form ω we started with (see [78] for a thorough discussion of the situation).

It was shown by Zelditch [89], by generalizing a result of Tian [90], that $(\Phi^{(m)})^*\omega_{\mathrm{FS}}$ admits a complete asymptotic expansion in powers of $1/m$ as $m \to \infty$. In fact it is related to the asymptotic expansion of the Bergman kernel (7.13) along the diagonal. The pull-back can be given as [89, Proposition 9]

$$\left(\phi^{(m)}\right)^*\omega_{\mathrm{FS}} = m\omega + \mathrm{i}\,\partial\bar{\partial}\log u_m(x). \tag{7.22}$$

If we again replace $1/m$ by ν, we obtain via (7.20) the Karabegov form introduced in Section 4.8

$$\widehat{\omega} = \mathbb{F}\left(\left(\phi^{(m)}\right)^*\omega_{\mathrm{FS}}\right). \tag{7.23}$$

References

[1] M. Bordemann, E. Meinrenken, and M. Schlichenmaier, "Toeplitz quantization of Kähler manifolds and gl(n), $n \to \infty$ limits," *Communications in Mathematical Physics*, vol. 165, no. 2, pp. 281–296, 1994.
[2] J. Madore, "The fuzzy sphere," *Classical and Quantum Gravity*, vol. 9, no. 1, pp. 69–87, 1992.
[3] A. P. Balachandran, B. P. Dolan, J. Lee, X. M., and D. O'Connor, "Fuzzy complex projective spaces and their star-products," *Journal of Geometry and Physics*, vol. 43, no. 2-3, pp. 184–204, 2002.
[4] A. P. Balachandran, S. Kurkcuoglu, and S. Vaidya, *Lectures on Fuzzy and Fuzzy SUSY Physics*, World Scientific, Singapore, 2007.
[5] C. I. Lazaroiu, D. McNamee, and Sämann, "Generalized Berezin quantization, Bergman metric and fuzzy laplacians," *Journal of High Energy Physics*. In press.
[6] B. P. Dolan, I. Huet, S. Murray, and D. O'Connor, "Noncommutative vector bundles over fuzzy $\mathbb{CP}^N$ and their covariant derivatives," *Journal of High Energy Physics*, vol. 7, article 007, 2007.
[7] P. Aschieri, J. Madore, P. Manousselis, and G. Zoupanos, "Dimensional reduction over fuzzy coset spaces," *Journal of High Energy Physics*, vol. 2004, no. 4, article 034, 2004.
[8] H. Grosse, M. Maceda, J. Madore, and H. Steinacker, "Fuzzy instantons," *International Journal of Modern Physics A*, vol. 17, no. 15, pp. 2095–2111, 2002.

[9] H. Grosse and H. Steinacker, "Finite gauge theory on fuzzy $\mathbb{CP}^2$," *Nuclear Physics B*, vol. 704, p. 145, 2005.

[10] U. Carow-Watamura, H. Steinacker, and S. Watamura, "Monopole bundles over fuzzy complex projective spaces," *Journal of Geometry and Physics*, vol. 54, no. 4, pp. 373–399, 2005.

[11] J. E. Andersen, "Asymptotic faithfulness of the quantum $SU(n)$ representations of the mapping class groups," *Annals of Mathematics*, vol. 163, no. 1, pp. 347–368, 2006.

[12] J. E. Andersen, "Deformation quantization and geometric quantization of abelian moduli spaces," *Communications in Mathematical Physics*, vol. 255, no. 3, pp. 727–745, 2005.

[13] M. Schlichenmaier, "Berezin-Toeplitz quantization of the moduli space of flat $SU(N)$ connections," *Journal of Geometry and Symmetry in Physics*, vol. 9, pp. 33–44, 2007.

[14] J. E. Andersen, "Mapping class groups do not have property (T)," http://arxiv.org/abs/0706.2184.

[15] J. E. Andersen and N. L. Gammelgaard, "Hitchin's projectively flat connection, Toeplitz operators and the asymptotic expansion of TQFT curve operators," http://arxiv.org/abs/0903.4091.

[16] J. E. Andersen and K. Ueno, "Abelian conformal field theory and determinant bundles," *International Journal of Mathematics*, vol. 18, no. 8, pp. 919–993, 2007.

[17] J. E. Andersen and K. Ueno, "Geometric construction of modular functors from conformal field theory," *Journal of Knot Theory and Its Ramifications*, vol. 16, no. 2, pp. 127–202, 2007.

[18] X. Ma and G. Marinescu, *Holomorphic Morse Inequalities and Bergman Kernels*, vol. 254 of *Progress in Mathematics*, Birkhäuser, Basel, Switzerland, 2007.

[19] X. Ma and G. Marinescu, "Toeplitz operators on symplectic manifolds," *Journal of Geometric Analysis*, vol. 18, no. 2, pp. 565–611, 2008.

[20] L. Charles, "Toeplitz operators and Hamiltonian torus actions," *Journal of Functional Analysis*, vol. 236, no. 1, pp. 299–350, 2006.

[21] S. T. Ali and H.-D. Doebner, "Ordering problem in quantum mechanics: prime quantization and a physical interpretation," *Physical Review A*, vol. 41, no. 3, pp. 1199–1210, 1990.

[22] S. T. Ali and M. Engliš, "Quantization methods: a guide for physicists and analysts," *Reviews in Mathematical Physics*, vol. 17, no. 4, pp. 391–490, 2005.

[23] S. T. Ali, "Quantization techniques: a quick overview," in *Contemporary Problems in Mathematical Physics*, pp. 3–78, World Scientific, River Edge, NJ, USA, 2002.

[24] D. Sternheimer, "Deformation quantization: twenty years after," in *Particles, Fields, and Gravitation*, vol. 453 of *AIP Conference Proceedings*, pp. 107–145, The American Institute of Physics, Woodbury, NY, USA, 1998.

[25] G. Dito and D. Sternheimer, "Deformation quantization: genesis, developments and metamorphoses," in *Deformation Quantization*, vol. 1 of *IRMA Lectures in Mathematics and Theoretical Physics*, pp. 9–54, Walter de Gruyter, Berlin, Germany, 2002.

[26] M. Schlichenmaier, *An Introduction to Riemann Surfaces, Algebraic Curves and Moduli Spaces*, Theoretical and Mathematical Physics, Springer, Berlin, Germany, 2nd edition, 2007.

[27] R. Abraham and J. E. Marsden, *Foundations of Mechanics*, Addison Wesley, Redwood City, Calif, USA, 1985.

[28] L. Boutet de Monvel and V. Guillemin, *The Spectral Theory of Toeplitz Operators*, Princeton University Press, Princeton, NJ, USA, 1981.

[29] A. V. Karabegov and M. Schlichenmaier, "Identification of Berezin-Toeplitz deformation quantization," *Journal für die Reine und Angewandte Mathematik*, vol. 540, pp. 49–76, 2001.

[30] M. Schlichenmaier, "Deformation quantization of compact Kähler manifolds by Berezin-Toeplitz quantization," in *Proceedings of the Conférence Moshé Flato*, G. Dito and D. Sternheimer, Eds., vol. 22 of *Mathematical Physics Studies*, pp. 289–306, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2000.

[31] M. Rieffel, "Questions on quantization," in *Operator Algebras and Operator Theory*, vol. 228 of *Contemporary Mathematics*, pp. 315–326, American Mathematical Society, Providence, RI, USA, 1998.

[32] N. P. Landsman, *Mathematical Topics between Classical and Quantum Mechanics*, Springer Monographs in Mathematics, Springer, New York, NY, USA, 1998.

[33] G. M. Tuynman, "Generalized Bergman kernels and geometric quantization," *Journal of Mathematical Physics*, vol. 28, no. 3, pp. 573–583, 1987.

[34] M. Bordemann, J. Hoppe, P. Schaller, and M. Schlichenmaier, "$\mathrm{gl}(\infty)$ and geometric quantization," *Communications in Mathematical Physics*, vol. 138, no. 2, pp. 209–244, 1991.

[35] M. Schlichenmaier, *Zwei Anwendungen algebraisch-geometrischer Methoden in der theoretischen Physik: Berezin-Toeplitz-Quantisierung und globale Algebren der zweidimensionalen konformen Feldtheorie*, Habiliationsschrift Universität Mannheim, 1996.

[36] F. A. Berezin, "Quantization in complex symmetric spaces," *Izvestiya Akademii Nauk SSSR*, vol. 9, no. 2, pp. 341–379, 1975.

[37] H. Upmeier, "Toeplitz C^*-algebras on bounded symmetric domains," *Annals of Mathematics*, vol. 119, no. 3, pp. 549–576, 1984.

[38] H. Upmeier, "Toeplitz operators on symmetric Siegel domains," *Mathematische Annalen*, vol. 271, no. 3, pp. 401–414, 1985.

[39] H. Upmeier, "Fredholm indices for Toeplitz operators on bounded symmetric domains," *American Journal of Mathematics*, vol. 110, no. 5, pp. 811–832, 1988.

[40] H. Upmeier, "Toeplitz C^*-algebras and noncommutative duality," *Journal of Operator Theory*, vol. 26, no. 2, pp. 407–432, 1991.

[41] H. Upmeier, *Toeplitz Operators and Index Theory in Several Complex Variables*, vol. 81 of *Operator Theory: Advances and Applications*, Birkhäuser, Basel, Switzerland, 1996.

[42] C. A. Berger and L. A. Coburn, "Toeplitz operators and quantum mechanics," *Journal of Functional Analysis*, vol. 68, no. 3, pp. 273–299, 1986.

[43] L. A. Coburn, "Deformation estimates for the Berezin-Toeplitz quantization," *Communications in Mathematical Physics*, vol. 149, no. 2, pp. 415–424, 1992.

[44] S. Klimek and A. Lesniewski, "Quantum Riemann surfaces. I. The unit disc," *Communications in Mathematical Physics*, vol. 146, no. 1, pp. 103–122, 1992.

[45] S. Klimek and A. Lesniewski, "Quantum Riemann surfaces. II. The discrete series," *Communications in Mathematical Physics*, vol. 24, pp. 125–139, 1992.

[46] M. Engliš, "Weighted Bergman kernels and quantization," *Communications in Mathematical Physics*, vol. 227, no. 2, pp. 211–241, 2002.

[47] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, "Quantum mechanics as a deformation of classical mechanics—part I," *Letters in Mathematical Physics*, vol. 1, no. 6, pp. 521–530, 1977.

[48] F. Bayen, M. Flato, C. Fronsdal, A. Lichnerowicz, and D. Sternheimer, "Deformation theory and quantization—parts II and III," *Annals of Physics*, vol. 111, pp. 61–110, 111–115, 1978.

[49] F. A. Berezin, "Quantization," *Mathematics of the USSR-Izvestiya*, vol. 8, pp. 1109–1165, 1974.

[50] J. E. Moyal, "Quantum mechanics as a statistical theory," vol. 45, pp. 99–124, 1949.

[51] H. Weyl, *Gruppentheorie und Quantenmechanik*, Wissenschaftliche Buchgesellschaft, Darmstadt, Germany, 2nd edition, 1977.

[52] Marc De Wilde and P. B. A. Lecomte, "Existence of star-products and of formal deformations of the Poisson Lie algebra of arbitrary symplectic manifolds," *Letters in Mathematical Physics*, vol. 7, no. 6, pp. 487–496, 1983.

[53] H. Omori, Y. Maeda, and A. Yoshioka, "Weyl manifolds and deformation quantization," *Advances in Mathematics*, vol. 85, no. 2, pp. 224–255, 1991.

[54] H. Omori, Y. Maeda, and A. Yoshioka, "Existence of closed star-products," *Letters in Mathematical Physics*, vol. 26, pp. 284–294, 1992.

[55] B. V. Fedosov, "Deformation quantization and asymptotic operator representation," *Funktional Anal i Prilozhen*, vol. 25, no. 2, pp. 184–194, 1990.

[56] B. V. Fedosov, "A simple geometrical construction of deformation quantization," *Journal of Differential Geometry*, vol. 40, no. 2, pp. 213–238, 1994.

[57] M. Kontsevich, "Deformation quantization of Poisson manifolds," *Letters in Mathematical Physics*, vol. 66, no. 3, pp. 157–216, 2003.

[58] P. Deligne, "Déformations de l'algèbre des fonctions d'une varièté symplectique: comparaison entre Fedosov et De Wilde, Lecomte," *Selecta Mathematica. New Series*, vol. 1, no. 4, pp. 667–697, 1995.

[59] S. Gutt and J. Rawnsley, "Equivalence of star products on a symplectic manifold; an introduction to Deligne's Čech cohomology classes," *Journal of Geometry and Physics*, vol. 29, no. 4, pp. 347–392, 1999.

[60] M. Bertelson, M. Cahen, and S. Gutt, "Equivalence of star products," *Classical and Quantum Gravity*, vol. 14, no. 1, pp. A93–A107, 1997.

[61] R. Nest and B. Tsygan, "Algebraic index theorem," *Communications in Mathematical Physics*, vol. 172, no. 2, pp. 223–262, 1995.

[62] R. Nest and B. Tsygan, “Algebraic index theory for families,” *Advances in Mathematics*, vol. 113, pp. 151–205, 1995.

[63] A. V. Karabegov, “Deformation quantizations with separation of variables on a Kähler manifold,” *Communications in Mathematical Physics*, vol. 180, no. 3, pp. 745–755, 1996.

[64] M. Bordemann and St. Waldmann, “A Fedosov star product of the Wick type for Kähler manifolds,” *Letters in Mathematical Physics*, vol. 41, no. 3, pp. 243–253, 1997.

[65] N. Reshetikhin and L. A. Takhtajan, “Deformation quantization of Kähler manifolds,” in *L. D. Faddeev’s Seminar on Mathematical Physics*, vol. 201 of *American Mathematical Society Translations: Series 2*, pp. 257–276, American Mathematical Society, Providence, RI, USA, 2000.

[66] M. Schlichenmaier, “Berezin-Toeplitz quantization of compact Kähler manifolds,” in *Quantization, Coherent States and Poisson Structures*, A. Strasburger, S. T. Ali, J.-P. Antoine, J.-P. Gazeau, and A. Odzijewicz, Eds., pp. 101–115, Polish Scientific Publisher PWN, 1998, Proceedings of the 14th Workshop on Geometric Methods in Physics (Białowieża, Poland, July 1995).

[67] M. Schlichenmaier, “Deformation quantization of compact Kähler manifolds via Berezin-Toeplitz operators,” in *Proceedings of the 21st International Colloquium on Group Theoretical Methods in Physics*, H.-D. Doebner, P. Nattermann, and W. Scherer, Eds., pp. 369–400, World Scientific, Goslar, Germany, 1996.

[68] E. Hawkins, “The correspondence between geometric quantization and formal deformation quantization,” http://arxiv.org/abs/physics/9811049.

[69] D. Borthwick, T. Paul, and A. Uribe, “Semiclassical spectral estimates for Toeplitz operators,” *Annales de l’Institut Fourier*, vol. 48, no. 4, pp. 1189–1229, 1998.

[70] A. V. Karabegov, “Cohomological classification of deformation quantizations with separation of variables,” *Letters in Mathematical Physics*, vol. 43, no. 4, pp. 347–357, 1998.

[71] A. V. Karabegov, “On the canonical normalization of a trace density of deformation quantization,” *Letters in Mathematical Physics*, vol. 45, no. 3, pp. 217–228, 1998.

[72] V. Guillemin, “Some classical theorems in spectral theory revisited,” in *Seminar on Singularities of Solutions of Linear Partial Differential Equations*, L. Hörmander, Ed., vol. 91 of *Annals of Mathematics Studies*, pp. 219–259, Princeton University Press, Princeton, NJ, USA, 1979.

[73] J. H. Rawnsley, “Coherent states and Kähler manifolds,” *The Quarterly Journal of Mathematics. Oxford. Second Series*, vol. 28, no. 112, pp. 403–415, 1977.

[74] J. Rawnsley, M. Cahen, and S. Gutt, “Quantization of Kähler manifolds. I. Geometric interpretation of Berezin’s quantization,” *Journal of Geometry and Physics*, vol. 7, no. 1, pp. 45–62, 1990.

[75] M. Cahen, S. Gutt, and J. Rawnsley, “Quantization of Kähler manifolds. II,” *Transactions of the American Mathematical Society*, vol. 337, no. 1, pp. 73–98, 1993.

[76] M. Cahen, S. Gutt, and J. Rawnsley, “Quantization of Kähler manifolds. III,” *Letters in Mathematical Physics*, vol. 30, no. 4, pp. 291–305, 1994.

[77] M. Cahen, S. Gutt, and J. Rawnsley, “Quantization of Kähler manifolds. IV,” *Letters in Mathematical Physics*, vol. 34, no. 2, pp. 159–168, 1995.

[78] S. Berceanu and M. Schlichenmaier, “Coherent state embeddings, polar divisors and Cauchy formulas,” *Journal of Geometry and Physics*, vol. 34, no. 3-4, pp. 336–358, 2000.

[79] C. Moreno and P. Ortega-Navarro, “$*$-products on $D1(\mathbb{C})$, S^2 and related spectral analysis,” *Letters in Mathematical Physics*, vol. 7, no. 3, pp. 181–193, 1983.

[80] C. Moreno, “$*$-products on some Kähler manifolds,” *Letters in Mathematical Physics*, vol. 11, no. 4, pp. 361–372, 1986.

[81] M. Engliš, “Berezin quantization and reproducing kernels on complex domains,” *Transactions of the American Mathematical Society*, vol. 348, no. 2, pp. 411–479, 1996.

[82] M. Engliš, “Asymptotics of the Berezin transform and quantization on planar domains,” *Duke Mathematical Journal*, vol. 79, no. 1, pp. 57–76, 1995.

[83] M. Engliš, “The asymptotics of a Laplace integral on a Kähler manifold,” *Journal für die Reine und Angewandte Mathematik*, vol. 528, pp. 1–39, 2000.

[84] A. Unterberger and H. Upmeier, “The Berezin transform and invariant differential operators,” *Communications in Mathematical Physics*, vol. 164, no. 3, pp. 563–597, 1994.

[85] M. Engliš and J. Peetre, “On the correspondence principle for the quantized annulus,” *Mathematica Scandinavica*, vol. 78, no. 2, pp. 183–206, 1996.

[86] M. Schlichenmaier, “Berezin-Toeplitz quantization and Berezin transform,” in *Long Time Behaviour of Classical and Quantum Systems*, vol. 1, pp. 271–287, World Scientific, River Edge, NJ, USA, 2001.

[87] M. Schlichenmaier, "Berezin-Toeplitz quantization and Berezin symbols for arbitrary compact Kähler manifolds," in *Proceedings of the 17th Workshop on Geometric Methods in Physics*, M. Schlichenmaier, et al., Ed., pp. 45–56, Warsaw University Press, Białowieża, Poland, 1998.

[88] L. Boutet de Monvel and J. Sjöstrand, "Sur la singularité des noyaux de Bergman et de Szegö," in *Journées: Équations aux Dérivées Partielles de Rennes*, Asterisque, pp. 123–164, Société Mathématique de France, Paris, France, 1976.

[89] S. Zelditch, "Szegö kernels and a theorem of Tian," *International Mathematics Research Notices*, no. 6, pp. 317–331, 1998.

[90] G. Tian, "On a set of polarized Kähler metrics on algebraic manifolds," *Journal of Differential Geometry*, vol. 32, no. 1, pp. 99–130, 1990.

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 457635, 37 pages
doi:10.1155/2010/457635

Review Article

The Partial Inner Product Space Method: A Quick Overview

Jean-Pierre Antoine[1] and Camillo Trapani[2]

[1] *Institut de Recherche en Mathématique et Physique, Université Catholique de Louvain, 1348 Louvain-la-Neuve, Belgium*
[2] *Dipartimento di Matematica ed Applicazioni, Università di Palermo, 90123 Palermo, Italy*

Correspondence should be addressed to Jean-Pierre Antoine, jean-pierre.antoine@uclouvain.be

Received 16 December 2009; Accepted 15 April 2010

Academic Editor: S. T. Ali

Many families of function spaces play a central role in analysis, in particular, in signal processing (e.g., wavelet or Gabor analysis). Typical are L^p spaces, Besov spaces, amalgam spaces, or modulation spaces. In all these cases, the parameter indexing the family measures the behavior (regularity, decay properties) of particular functions or operators. It turns out that all these space families are, or contain, scales or lattices of Banach spaces, which are special cases of *partial inner product spaces* (PIP-*spaces*). In this context, it is often said that such families should be taken as a whole and operators, bases, and frames on them should be defined globally, for the whole family, instead of individual spaces. In this paper, we will give an overview of PIP-spaces and operators on them, illustrating the results by space families of interest in mathematical physics and signal analysis. The interesting fact is that they allow a global definition of operators, and various operator classes on them have been defined.

1. Motivation

In the course of their curriculum, physics and mathematics students are usually taught the basics of Hilbert space, including operators of various types. The justification of this choice is twofold. On the mathematical side, Hilbert space is the example of an infinite-dimensional topological vector space that more closely resembles the familiar Euclidean space and thus it offers the student a smooth introduction into functional analysis. On the physics side, the fact is simply that Hilbert space is the daily language of quantum theory; therefore, mastering it is an essential tool for the quantum physicist.

However, the tool in question is actually insufficient. A pure Hilbert space formulation of quantum mechanics is both inconvenient and foreign to the daily behavior of most physicists, who stick to the more suggestive version of Dirac, although it lacks a rigorous formulation. On the other hand, the interesting solutions of most partial differential equations are seldom smooth or square integrable. Physically meaningful events correspond to changes of regime, which mean discontinuities and/or distributions. Shock waves are a typical example. Actually this state of affairs was recognized long ago by authors like Leray or Sobolev, whence they introduced the notion of *weak solution*. Thus it is no coincidence that many textbooks on PDEs begin with a thorough study of distribution theory [1–4].

All this naturally leads to the introduction of Rigged Hilbert Spaces (RHS) [5]. In a nutshell, a RHS is a triplet:

$$\Phi \hookrightarrow \mathscr{H} \hookrightarrow \Phi^{\times}, \tag{1.1}$$

where $\mathscr{H}$ is a Hilbert space, Φ is a dense subspace of the $\mathscr{H}$, equipped with a locally convex topology, finer than the norm topology inherited from $\mathscr{H}$, and $\Phi^{\times}$ is the space of continuous conjugate linear functionals on Φ, endowed with the strong dual topology. By duality, each space in (1.1) is dense in the next one and all embeddings are linear and continuous. In addition, the space Φ is in general required to be reflexive and nuclear. Standard examples of rigged Hilbert spaces are the Schwartz distribution spaces over $\mathbb{R}$ or $\mathbb{R}^n$, namely $\mathcal{S} \subset L^2 \subset \mathcal{S}^{\times}$ or $\mathfrak{D} \subset L^2 \subset \mathfrak{D}^{\times}$ [5–8].

The problem with the RHS (1.1) is that, besides the Hilbert space vectors, it contains only two types of elements: "very good" functions in Φ and "very bad" ones in $\Phi^{\times}$. If one wants a fine control on the behavior of individual elements, one has to interpolate somehow between the two extreme spaces. In the case of the Schwartz triplet, $\mathcal{S} \subset L^2 \subset \mathcal{S}^{\times}$, a well-known solution is given by a chain of Hilbert spaces, the so-called Hermite representation of tempered distributions [9].

In fact, this is not at all an isolated case. Indeed many function spaces that play a central role in analysis come in the form of families, indexed by one or several parameters that characterize the behavior of functions (smoothness, behavior at infinity, ...). The typical structure is a chain or a scale of Hilbert spaces, or a chain of (reflexive) Banach spaces (a discrete chain of Hilbert spaces $\{\mathscr{H}_n\}_{n\in\mathbb{Z}}$ is called a *scale* if there exists a self-adjoint operator $B \geqslant 1$ such that $\mathscr{H}_n = D(B^n)$, for all $n \in \mathbb{Z}$, with the graph norm $\|f\|_n = \|B^n f\|$. A similar definition holds for a continuous chain $\{\mathscr{H}_\alpha\}_{\alpha\in\mathbb{R}}$.). Let us give two familiar examples.

(i) First, consider the Lebesgue the Lebesgue L^p spaces on a finite interval, for example, $\mathcal{I} = \{L^p([0,1], dx), 1 \leqslant p \leqslant \infty\}$:

$$L^{\infty} \subset \cdots \subset L^{\overline{q}} \subset L^{\overline{r}} \subset \cdots \subset L^2 \subset \cdots \subset L^r \subset L^q \subset \cdots \subset L^1, \tag{1.2}$$

where $1 < q < r < 2$. Here L^q and $L^{\overline{q}}$ are dual to each other $(1/q + 1/\overline{q} = 1)$, and similarly are $L^r, L^{\overline{r}}$ $(1/r + 1/\overline{r} = 1)$. By the Hölder inequality, the (L^2) inner product

$$\langle f \mid g \rangle = \int_0^1 \overline{f(x)} g(x) dx \tag{1.3}$$

is well defined if $f \in L^q, g \in L^{\overline{q}}$. However, it is *not* well defined for two arbitrary functions $f, g \in L^1$. Take, for instance, $f(x) = g(x) = x^{-1/2} : f \in L^1$, but $fg = f^2 \notin L^1$.

Thus, on L^1, (1.3) defines only a *partial* inner product. The same result holds for any compact subset of $\mathbb{R}$ instead of [0,1].

(ii) As a second example, take the scale of Hilbert spaces built on the powers of a positive self-adjoint operator $A \geqslant 1$ in a Hilbert space $\mathcal{H}_0$. Let $\mathcal{H}_n$ be $D(A^n)$, the domain of A^n, equipped with the graph norm $\|f\|_n = \|A^n f\|$, $f \in D(A^n)$, for $n \in \mathbb{N}$ or $n \in \mathbb{R}^+$, and $\mathcal{H}_{\overline{n}} := \mathcal{H}_{-n} = \mathcal{H}_n^\times$ (conjugate dual)

$$\mathfrak{D}^\infty(A) := \bigcap_n \mathcal{H}_n \subset \ldots \subset \mathcal{H}_2 \subset \mathcal{H}_1 \subset \mathcal{H}_0 \subset \mathcal{H}_{\overline{1}} \subset \mathcal{H}_{\overline{2}} \cdots \subset \mathfrak{D}_{\overline{\infty}}(A) := \bigcup_n \mathcal{H}_n. \tag{1.4}$$

Note that, in the second example (ii), the index n could also be taken as real, the link between the two cases being established by the spectral theorem for self-adjoint operators. Here again the inner product of $\mathcal{H}_0$ extends to each pair $\mathcal{H}_n, \mathcal{H}_{-n}$, but on $\mathfrak{D}_{\overline{\infty}}(A)$ it yields only a partial inner product. The following examples are standard:

(i) $(A_{\mathrm{p}} f)(x) = (1 + x^2) f(x)$ in $L^2(\mathbb{R}, dx)$,

(ii) $(A_{\mathrm{m}} f)(x) = (1 - d^2/dx^2) f(x)$ in $L^2(\mathbb{R}, dx)$,

(iii) $(A_{\mathrm{osc}} f)(x) = (1 + x^2 - d^2/dx^2) f(x)$ in $L^2(\mathbb{R}, dx)$.

(The notation is suggested by the operators of position, momentum and harmonic oscillator energy in quantum mechanics, resp.). Note that both $\mathfrak{D}^\infty(A_{\mathrm{p}}) \cap \mathfrak{D}^\infty(A_{\mathrm{m}})$ and $\mathfrak{D}^\infty(A_{\mathrm{osc}})$ coincide with the Schwartz space $\mathcal{S}(\mathbb{R})$ of smooth functions of fast decay, and $\mathfrak{D}_{\overline{\infty}}(A_{\mathrm{osc}})$ with the space $\mathcal{S}^\times(\mathbb{R})$ of tempered distributions (considered here as continuous *conjugate linear* functionals on $\mathcal{S}$). As for the operator A_{m}, it generates the scale of Sobolev spaces $H^s(\mathbb{R})$, $s \in \mathbb{Z}$ or $\mathbb{R}$.

However, a moment's reflection shows that the total-order relation inherent in a chain is in fact an unnecessary restriction; partially ordered structures are sufficient, and indeed necessary in practice. For instance, in order to get a better control on the behavior of individual functions, one may consider the lattice built on the powers of A_{p} and A_{m} simultaneously. Then the extreme spaces are still $\mathcal{S}(\mathbb{R})$ and $\mathcal{S}^\times(\mathbb{R})$. Similarly, in the case of several variables, controlling the behavior of a function in each variable separately requires a nonordered set of spaces. This is in fact a statement about tensor products (remember that $L^2(X \times Y) \simeq L^2(X) \otimes L^2(Y)$). Indeed the tensor product of two chains of Hilbert spaces, $\{\mathcal{H}_n\} \otimes \{\mathcal{K}_m\}$, is naturally a lattice $\{\mathcal{H}_n \otimes \mathcal{K}_m\}$ of Hilbert spaces. For instance, in the example above, for two variables x, y, that would mean considering intermediate Hilbert spaces corresponding to the product of two operators, $(A_{\mathrm{m}}(x))^n (A_{\mathrm{m}}(y))^m$.

Thus the structure to analyze is that of *lattices of Hilbert or Banach spaces*, interpolating between the extreme spaces of an RHS, as in (1.1). Many examples can be given, for instance, the lattice generated by the spaces $L^p(\mathbb{R}, dx)$, the amalgam spaces $W(L^p, \ell^q)$, the mixed-norm spaces $L_m^{p,q}(\mathbb{R}, dx)$, and many more. In all these cases, which contain most families of function spaces of interest in analysis and in signal processing, a common structure emerges for the "large" space V, defined as the union of all individual spaces. There is a lattice of Hilbert or reflexive Banach spaces V_r, with an (order-reversing) involution $V_r \leftrightarrow V_{\overline{r}}$, where $V_{\overline{r}} = V_r^\times$ (the space of continuous conjugate linear functionals on V_r), a central Hilbert space $V_o \simeq V_{\overline{o}}$, and a partial inner product on V that extends the inner product of V_o to pairs of dual spaces V_r, $V_{\overline{r}}$.

Moreover, many operators should be considered globally, for the whole scale or lattice, instead of on individual spaces. In the case of the spaces $L^p(\mathbb{R})$, such are, for instance, operators implementing translations ($x \mapsto x - y$) or dilations ($x \mapsto x/a$), convolution operators, Fourier transform, and so forth. In the same spirit, it is often useful to have a *common* basis for the whole family of spaces, such as the Haar basis for the spaces $L^p(\mathbb{R})$, $1 < p < \infty$. Thus we need a notion of operator and basis defined globally for the scale or lattice itself.

This state of affairs prompted A. Grossmann and one of us (the first author) to systematize this approach, and this led to the concept of *partial inner product space* or PIP-space [10–13]. After many years and various developments, we devoted a full monograph [14] to a detailed survey of the theory. The aim of this paper is to present the formalism of PIP-spaces, which indeed answers these questions. In a first part, the structure of PIP-space is derived systematically from the abstract notion of compatibility and then particularized to the examples listed above. In a second part, operators on PIP-spaces are introduced and illustrated by several operators commonly used in Gabor or wavelet analysis. Finally we describe a number of applications of PIP-spaces in mathematical physics and in signal processing. Of course, the treatment is sketchy, for lack of space. For a complete information, we refer the reader to our monograph [14].

2. Partial Inner Product Spaces

2.1. Basic Definitions

The basic question is how to generate PIP-spaces in a systematic fashion. In order to answer, we may reformulate it as follows: given a vector space V and two vectors $f, g \in V$, when does their inner product make sense? A way of formalizing the answer is given by the idea of *compatibility*.

Definition 2.1. A *linear compatibility relation* on a vector space V is a symmetric binary relation $f\#g$ which preserves linearity:

$$\begin{gathered} f\#g \Longleftrightarrow g\#f, \quad \forall f, g \in V, \\ f\#g,\, f\#h \Longrightarrow f\#(\alpha g + \beta h), \quad \forall f, g, h \in V,\ \forall \alpha, \beta \in \mathbb{C}. \end{gathered} \tag{2.1}$$

As a consequence, for every subset $S \subset V$, the set $S^{\#} = \{g \in V : g\#f, \text{for all } f \in S\}$ is a vector subspace of V and one has

$$S^{\#\#} = \left(S^{\#}\right)^{\#} \supseteq S, \qquad S^{\#\#\#} = S^{\#}. \tag{2.2}$$

Thus one gets the following equivalences:

$$\begin{aligned} f\#g &\Longleftrightarrow f \in \{g\}^{\#} \Longleftrightarrow \{f\}^{\#\#} \subseteq \{g\}^{\#} \\ &\Longleftrightarrow g \in \{f\}^{\#} \Longleftrightarrow \{g\}^{\#\#} \subseteq \{f\}^{\#}. \end{aligned} \tag{2.3}$$

From now on, we will call *assaying subspace* of V a subspace S such that $S^{\#\#} = S$ and denote by $\mathcal{F}(V, \#)$ the family of all assaying subsets of V, ordered by inclusion. Let F be the isomorphy class of $\mathcal{F}$, that is, $\mathcal{F}$ is considered as an abstract partially ordered set. Elements of F will be denoted by $r, q, \ldots$, and the corresponding assaying subsets by $V_r, V_q, \ldots$. By definition, $q \leqslant r$ if and only if $V_q \subseteqq V_r$. We also write $V_{\overline{r}} = V_r^{\#},\ r \in F$. Thus the relations (2.3) mean that $f\#g$ if and only if there is an index $r \in F$ such that $f \in V_r,\ g \in V_{\overline{r}}$. In other words, vectors should not be considered individually, but only in terms of assaying subspaces, which are the building blocks of the whole structure.

It is easy to see that the map $S \mapsto S^{\#\#}$ is a closure, in the sense of universal algebra, so that the assaying subspaces are precisely the "closed" subsets. Therefore one has the following standard result.

Theorem 2.2. *The family $\mathcal{F}(V,\#) \equiv \{V_r, r \in F\}$, ordered by inclusion, is a complete involutive lattice, that is, it is stable under the following operations, arbitrarily iterated:*

(i) *involution:* $V_r \leftrightarrow V_{\overline{r}} = (V_r)^{\#}$,

(ii) *infimum:* $V_{p\wedge q} \equiv V_p \wedge V_q = V_p \cap V_q,\ (p, q, r \in F)$,

(iii) *supremum:* $V_{p\vee q} \equiv V_p \vee V_q = (V_p + V_q)^{\#\#}$.

The smallest element of $\mathcal{F}(V,\#)$ is $V^{\#} = \bigcap_r V_r$ and the greatest element is $V = \bigcup_r V_r$. By definition, the index set F is also a complete involutive lattice; for instance,

$$\left(V_{p\wedge q}\right)^{\#} = V_{\overline{p\wedge q}} = V_{\overline{p}\vee\overline{q}} = V_{\overline{p}} \vee V_{\overline{q}}. \tag{2.4}$$

Definition 2.3. A *partial inner product* on $(V,\#)$ is a Hermitian form $\langle\cdot \mid \cdot\rangle$ defined exactly on compatible pairs of vectors. A *partial inner product space* (PIP-space) is a vector space V equipped with a linear compatibility and a partial inner product.

Note that the partial inner product is not required to be positive definite.

The partial inner product clearly defines a notion of *orthogonality:* $f \perp g$ if and only if $f\#g$ and $\langle f \mid g\rangle = 0$.

Definition 2.4. The PIP-space $(V, \#, \langle\cdot \mid \cdot\rangle)$ is *nondegenerate* if $(V^{\#})^{\perp} = \{0\}$, that is, if $\langle f \mid g\rangle = 0$ for all $f \in V^{\#}$ implies that $g = 0$.

We will assume henceforth that our PIP-space $(V, \#, \langle\cdot \mid \cdot\rangle)$ is nondegenerate. As a consequence, $(V^{\#}, V)$ and every couple $(V_r, V_{\overline{r}}),\ r \in F$, are dual pairs in the sense of topological vector spaces [15]. We also assume that the partial inner product is positive definite.

Now one wants the topological structure to match the algebraic structure, in particular, the topology τ_r on V_r should be such that its conjugate dual be $V_{\overline{r}}$: $(V_r[\tau_r])^{\times} = V_{\overline{r}},$ for all $r \in F$. This implies that the topology τ_r must be finer than the weak topology $\sigma(V_r, V_{\overline{r}})$ and coarser than the Mackey topology $\tau(V_r, V_{\overline{r}})$:

$$\sigma(V_r, V_{\overline{r}}) \preceq \tau_r \preceq \tau(V_r, V_{\overline{r}}). \tag{2.5}$$

From here on, we will assume that every V_r carries its Mackey topology $\tau(V_r, V_{\overline{r}})$. This choice has two interesting consequences. First, if $V_r[\tau_r]$ is a Hilbert space or a reflexive Banach space, then $\tau(V_r, V_{\overline{r}})$ coincides with the norm topology. Next, $r < s$ implies that $V_r \subset V_s$, and the embedding operator $E_{sr} : V_r \to V_s$ is continuous and has dense range. In particular, $V^\#$ is dense in every V_r.

2.2. Examples

2.2.1. Sequence Spaces

Let V be the space ω of *all* complex sequences $x = (x_n)$ and define on it (i) a compatibility relation by $x\#y \Leftrightarrow \sum_{n=1}^{\infty} |x_n y_n| < \infty$ and (ii) a partial inner product $\langle x \mid y \rangle = \sum_{n=1}^{\infty} \overline{x_n} y_n$.

Then $\omega^\# = \varphi$, the space of finite sequences, and the complete lattice $\mathcal{F}(\omega, \#)$ consists of Köthe's perfect sequence spaces [15, § 30]. Among these, typical assaying subspaces are the weighted Hilbert spaces

$$\ell^2(r) = \left\{ (x_n) : \sum_{n=1}^{\infty} |x_n|^2 r_n^{-2} < \infty \right\}, \tag{2.6}$$

where $r = (r_n), r_n > 0$, is a sequence of positive numbers. The involution is $\ell^2(r) \leftrightarrow \ell^2(\overline{r}) = \ell^2(r)^\times$, where $\overline{r}_n = 1/r_n$. In addition, there is a central, self-dual Hilbert space, namely, $\ell^2(1) = \ell^2(\overline{1}) = \ell^2$, where $1 = (1)$.

2.2.2. Spaces of Locally Integrable Functions

Let now V be $L^1_{\text{loc}}(\mathbb{R}, dx)$, the space of Lebesgue measurable functions, integrable over compact subsets, and define a compatibility relation on it by $f\#g \Leftrightarrow \int_{\mathbb{R}} |f(x)g(x)| dx < \infty$ and a partial inner product $\langle f \mid g \rangle = \int_{\mathbb{R}} \overline{f(x)} g(x) dx$.

Then $V^\# = L^\infty_c(\mathbb{R})$, the space of bounded measurable functions of compact support. The complete lattice $\mathcal{F}(L^1_{\text{loc}}, \#)$ consists of Köthe function spaces [16, 17]. Here again, typical assaying subspaces are weighted Hilbert spaces

$$L^2(r) = \left\{ f \in L^1_{\text{loc}}(\mathbb{R}, dx) : \int_{\mathbb{R}} |f(x)|^2 r(x)^{-2} dx < \infty \right\}, \tag{2.7}$$

with $r, r^{-1} \in L^2_{\text{loc}}(\mathbb{R}, dx),\ r(x) > 0$ a.e. The involution is $L^2(r) \leftrightarrow L^2(\overline{r})$, with $\overline{r} = r^{-1}$, and the central, self-dual Hilbert space is $L^2(\mathbb{R}, dx)$.

2.2.3. Nested Hilbert Spaces

This is the original construction of Grossmann [18] for finding an "easy" substitute to distributions, and actually one of the motivations for introducing PIP-spaces. And indeed the two are closely related; see [14, Section 2.4.1]

2.2.4. Rigged Hilbert Spaces

This is the simplest example of PIP-space, but it is a rather poor one. Indeed, in the RHS (1.1), two elements are compatible if both belong to $\mathscr{H}$, or one of them belongs to Φ. Thus the three defining spaces are the only assaying subspaces. The partial inner product is, of course, simply that of $\mathscr{H}$, provided the sesquilinear form that puts Φ and $\Phi^\times$ in duality has been correctly normalized.

3. Lattices of Hilbert or Banach Spaces

From the previous examples, we learn that $\mathcal{F}(V,\#)$ is a huge lattice (it is complete!) and that assaying subspaces may be complicated, such as Fréchet spaces, nonmetrizable spaces, and so forth. This situation suggests to choose an involutive sublattice $\mathcal{I} \subset \mathcal{F}$, indexed by I, such that

(i) $\mathcal{I}$ is *generating*:

$$f\#g \Longleftrightarrow \exists\ r \in I \text{ such that } f \in V_r,\ g \in V_{\overline{r}}, \tag{3.1}$$

(ii) every $V_r,\ r \in I$, is a Hilbert space or a reflexive Banach space,

(iii) there is a unique self-dual assaying subspace $V_o = V_{\overline{o}}$, which is a Hilbert space.

In that case, the structure $V_I := (V, \mathcal{I}, \langle\cdot \mid \cdot\rangle)$ is called, respectively, a *lattice of Hilbert spaces* (LHS) or a *lattice of Banach spaces* (LBS). Both types are particular cases of the so-called indexed PIP-spaces [14]. Note that $V^\#$, V themselves usually do *not* belong to the family $\{V_r, r \in I\}$, but they can be recovered as

$$V^\# = \bigcap_{r\in I} V_r, \quad V = \sum_{r\in I} V_r. \tag{3.2}$$

In the LBS case, the lattice structure takes the following forms:

(i) $V_{p\wedge q} = V_p \cap V_q$, with the *projective* norm

$$\|f\|_{p\wedge q} = \|f\|_p + \|f\|_q, \tag{3.3}$$

(ii) $V_{p\vee q} = V_p + V_q$, with the *inductive* norm

$$\|f\|_{p\vee q} = \inf_{f=g+h}\left(\|g\|_p + \|h\|_q\right), \quad g \in V_p,\ h \in V_q. \tag{3.4}$$

These norms are usual in interpolation theory [19]. In the LHS case, one takes similar definitions with squared norms, in order to get Hilbert norms throughout.

In the rest of this section, we will list a series of concrete examples of LHS/LBSs. Some more examples, which are of particular interest in signal processing, will be given in Section 6.2. For simplicity, we will restrict ourselves to one dimension, although most spaces may be defined on $\mathbb{R}^n$, $n > 1$, as well.

3.1. Chains of Hilbert or Banach Spaces

Typical are the two examples described in Section 1.

(1) The chain of Lebesgue spaces on a finite interval $\mathcal{I} = \{L^p([0,1], dx),\ 1 < p < \infty\}$. The chain (1.2) is a (totally ordered) lattice. The corresponding lattice completion is obtained by adding "nonstandard" spaces such as

$$L^{p-} = \bigcap_{1<q<p} L^q \quad \text{(non-normable Fréchet)}, \qquad L^{p+} = \bigcup_{p<q<\infty} L^q \quad \text{(nonmetrizable)}. \tag{3.5}$$

(2) The scale (1.4) of Hilbert spaces $\{\mathcal{H}_n, n \in \mathbb{Z}\}$ built on powers of $A = A^* \geqslant 1$. The lattice completion is similar to the previous one, introducing analogous "nonstandard" spaces [14, Section 5.1].

3.2. Sequence Spaces

3.2.1. A LHS of Weighted ℓ^2 Spaces

In ω, with the compatibility # and the partial inner product defined in Section 2.2.1, we may take the lattice $\mathcal{I} = \{\ell^2(r)\}$ of the weighted Hilbert spaces defined in (2.6), with lattice operations:

(i) infimum: $\ell^2(p \wedge q) = \ell^2(p) \wedge \ell^2(q) = \ell^2(r),\ r_n = \min(p_n, q_n)$,

(ii) supremum: $\ell^2(p \vee q) = \ell^2(p) \vee \ell^2(q) = \ell^2(s),\ s_n = \max(p_n, q_n)$,

(iii) duality: $\ell^2(p \wedge q) \leftrightarrow \ell^2(\overline{p} \vee \overline{q}), \ell^2(p \vee q) \leftrightarrow \ell^2(\overline{p} \wedge \overline{q})$.

As a matter of fact, the norms above are equivalent to the projective and inductive norms, respectively. Then, it is easy to show that the lattice $\mathcal{I} = \{\ell^2(r)\}$ is generating in $\mathcal{F}(\omega, \#)$.

3.2.2. Köthe Perfect Sequence Spaces

We have already noticed that the complete lattice $\mathcal{F}(\omega, \#)$ consists precisely of all Köthe perfect sequence spaces. Indeed, these are defined as the assaying subspaces corresponding to the compatibility #, which is called *α-duality* [15]. Among these, there is an interesting class, the so-called ℓ_ϕ spaces associated to symmetric norming functions.

Definition 3.1. A real-valued function ϕ defined on the space φ of finite sequences is said to be a *norming function* if

(n_1) $\phi(x) > 0$ for every sequence $x \in \varphi$, $x \neq 0$,

(n_2) $\phi(\alpha x) = |\alpha|\phi(x),$ for all $x \in \varphi$, for all $\alpha \in \mathbb{C}$,

(n_3) $\phi(x + y) \leqslant \phi(x) + \phi(y)$, for all $x, y \in \varphi$,

(n_4) $\phi(1, 0, 0, 0, \ldots) = 1$.

A norming function ϕ is *symmetric* if

(n_5) $\phi(x_1, x_2, \ldots, x_n, 0, 0, \ldots) = \phi(|x_{j_1}|, |x_{j_2}|, \ldots, |x_{j_n}|, 0, 0, \ldots)$,

where $j_1, j_2, \ldots, j_n$ is an arbitrary permutation of $1, 2, \ldots, n$.

From property (n_5), it is clear that a symmetric norming function ϕ is entirely determined by its values on the set $[\varphi]$ of finite, positive, nonincreasing sequences. Hence, from conditions (n_2) and (n_4), we deduce that

$$\phi_\infty(x) \leqslant \phi(x) \leqslant \phi_1(x), \quad \forall\, x \in \varphi, \tag{3.6}$$

where $\phi_\infty(x) = \max_{i=1,\ldots,n}|x_i|$ and $\phi_1(x) = \sum_{i=1}^n |x_i|$.

To every symmetric norming function ϕ, one can associate a Banach space ℓ_ϕ as follows. Given a sequence $x \in \omega$, define its nth section as $x^{(n)} = (x_1, x_2, \ldots, x_n, 0, 0, \ldots)$. Then the sequence $(\phi(x^{(n)}))$ is nondecreasing, so that one can define

$$\ell_\phi = \left\{ x \in \omega : \sup_n \phi\left(x^{(n)}\right) < \infty \right\} \tag{3.7}$$

and then extend the norming function ϕ to the whole of ℓ_ϕ by putting $\phi(x) = \lim_n \phi(x^{(n)})$. This relation defines a norm ϕ on ℓ_ϕ, for which it is complete, hence, a Banach space. In other words, we can also say that $\ell_\phi = \{x \in \omega : \phi(x) < \infty\}$ is the natural domain of definition of the extended norming function ϕ. Clearly, one has $\ell_{\phi_\infty} = \ell^\infty$ and $\ell_{\phi_1} = \ell^1$. Similarly, $\ell^p = \ell_{\phi_p}$, where $\phi_p(x) = \left(\sum_n |x_n|^p\right)^{1/p}$. Thus every space ℓ_ϕ contains ℓ^1 and is contained in ℓ^∞.

In addition, the set of Banach spaces ℓ_ϕ constitutes a lattice. Given two symmetric norming functions ϕ and ψ, one defines their infimum and supremum, exactly as for the general case:

(i) $\phi \wedge \psi := \max\{\phi, \psi\}$, which defines on the space $\ell_{\phi\wedge\psi} := \ell_\phi \cap \ell_\psi$ a norm equivalent to $\phi(x) + \psi(x)$,

(ii) $\phi \vee \psi := \min\{\phi, \psi\}$, which defines on the space $\ell_{\phi\vee\psi} := \ell_\phi + \ell_\psi$ a norm equivalent to $\inf_{x=y+z}\{\phi(y) + \psi(z)\},\ x \in \ell_\phi + \ell_\psi,\ y \in \ell_\phi,\ z \in \ell_\psi$.

It remains to analyze the relationship of the spaces ℓ_ϕ with the PIP-space structure of ω. Define, for any finite, positive, nonincreasing sequence $y \in [\varphi]$,

$$\overline{\phi}(y) := \max_{x\in[\varphi]} \frac{\langle x \mid y \rangle}{\phi(x)}. \tag{3.8}$$

The function $\overline{\phi}$ thus defined is a symmetric norming function; hence, it can be extended to the corresponding Banach space $\ell_{\overline{\phi}}$. The function $\overline{\phi}$ is said to be *conjugate* to ϕ and the space $\ell_{\overline{\phi}}$ is the conjugate dual of ℓ_ϕ with respect to the partial inner product, that is, $\ell_{\overline{\phi}} = (\ell_\phi)^\#$. Clearly one has $\overline{\overline{\phi}} = \phi$; hence, $\ell_{\overline{\overline{\phi}}} = (\ell_\phi)^{\#\#} = \ell_\phi$.

In addition, it is easy to show that $\ell_{\overline{\phi\wedge\psi}} = \ell_{\overline{\phi}\vee\overline{\psi}}$ and $\ell_{\overline{\phi\vee\psi}} = \ell_{\overline{\phi}\wedge\overline{\psi}}$. In other words, one gets the following result.

Proposition 3.2. *The family of Banach spaces ℓ_ϕ, where ϕ is a symmetric norming function, is an involutive sublattice of the lattice $\mathcal{F}(\omega, \#)$ and a LBS.*

Actually, since every ϕ satisfies the inclusions $\ell^1 \subset \ell_\phi \subset \ell^\infty$, the family $\{\ell_\phi\}$ is also an involutive sublattice of the lattice $\mathcal{F}(\ell^\infty, \#)$ obtained by restricting to ℓ^∞ the PIP-space structure of ω.

These spaces $\{\ell_\phi\}$ may be generalized further to what is called the theory of Banach ideals of sequences. See [14, Section 4.3] for more details.

3.3. *Spaces of Locally Integrable Functions*

3.3.1. A LHS of Weighted L^2 Spaces

In $L^1_{\mathrm{loc}}(\mathbb{R}, dx)$, we may take the lattice $\mathcal{I} = \{L^2(r)\}$ of the weighted Hilbert spaces defined in (2.7), with

(i) infimum: $L^2(p \wedge q) = L^2(p) \wedge L^2(q) = L^2(r),\ r(x) = \min(p(x), q(x))$,

(ii) supremum: $L^2(p \vee q) = L^2(p) \vee L^2(q) = L^2(s),\ s(x) = \max(p(x), q(x))$,

(iii) duality: $L^2(p \wedge q) \leftrightarrow L^2(\overline{p} \vee \overline{q}),\ L^2(p \vee q) \leftrightarrow L^2(\overline{p} \wedge \overline{q})$.

Here too, these norms are equivalent to the projective and inductive norms, respectively.

3.3.2. The Spaces $L^p(\mathbb{R}, dx),\ 1 < p < \infty$

The spaces $L^p(\mathbb{R}, dx),\ 1 < p < \infty$ do not constitute a scale, since one has only the inclusions $L^p \cap L^q \subset L^s, p < s < q$. Thus one has to consider the lattice they generate, with the following lattice operations:

(i) $L^p \wedge L^q = L^p \cap L^q$, with projective norm,

(ii) $L^p \vee L^q = L^p + L^q$, with inductive norm.

For $1 < p, q < \infty$, both spaces $L^p \wedge L^q$ and $L^p \vee L^q$ are reflexive Banach spaces and their conjugate duals are, respectively, $(L^p \wedge L^q)^\times = L^{\overline{p}} \vee L^{\overline{q}}$ and $(L^p \vee L^q)^\times = L^{\overline{p}} \wedge L^{\overline{q}}$.

It is convenient to introduce the following unified notation:

$$L^{(p,q)} = \begin{cases} L^p \wedge L^q, & \text{if } p \geqslant q, \\ L^p \vee L^q, & \text{if } p \leqslant q. \end{cases} \tag{3.9}$$

Then, for $1 < p, q < \infty$, $L^{(p,q)}$ is a reflexive Banach space, with conjugate dual $L^{(\overline{p},\overline{q})}$.

Next, if we represent (p, q) by the point of coordinates $(1/p, 1/q)$, we may associate all the spaces $L^{(p,q)}$ $(1 \leqslant p, q \leqslant \infty)$ in a one-to-one fashion with the points of a unit square $\mathrm{J} = [0,1] \times [0,1]$ (see Figure 1). Thus, in this picture, the spaces L^p are on the main diagonal, intersections $L^p \cap L^q$ above it and sums $L^p + L^q$ below.

The space $L^{(p,q)}$ is contained in $L^{(p',q')}$ if (p, q) is on the left and/or above (p', q'). Thus the smallest space is

$$V_{\mathrm{J}}^{\#} = L^{(\infty,1)} = L^\infty \cap L^1 \tag{3.10}$$

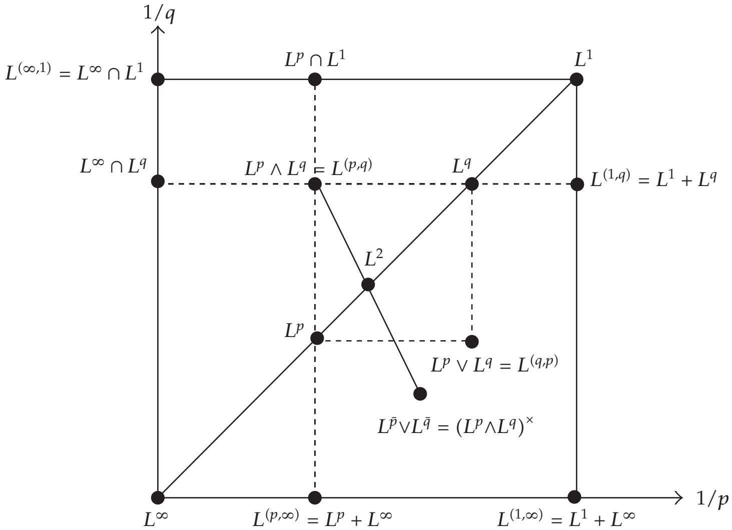


Figure 1: The unit square describing the lattice $\mathcal{J}$.

and it corresponds to the upper-left corner, while the largest one is

$$V_{\mathcal{J}} = L^{(1,\infty)} = L^1 + L^\infty, \tag{3.11}$$

corresponding to the lower-right corner. Inside the square, duality corresponds to (geometrical) symmetry with respect to the center (1/2,1/2) of the square, which represents the space L^2. The ordering of the spaces corresponds to the following rule:

$$L^{(p,q)} \subset L^{(p',q')} \Longleftrightarrow (p,q) \leqslant (p',q') \Longleftrightarrow p \geqslant p', \quad q \leqslant q'. \tag{3.12}$$

With respect to this ordering, $\mathcal{J}$ is an involutive lattice with the operations

$$\begin{gathered}(p,q) \wedge (p',q') = (p \vee p', q \wedge q'),\\ (p,q) \vee (p',q') = (p \wedge p', q \vee q'),\\ \overline{(p,q)} = (\overline{p},\overline{q}),\end{gathered} \tag{3.13}$$

where $p \wedge p' = \min\{p,p'\}$, $p \vee p' = \max\{p,p'\}$. It is remarkable that the lattice $\mathcal{J}$ generated by $\mathcal{I} = \{L^p\}$ is obtained at the first "generation". One has, for instance, $L^{(r,s)} \wedge L^{(a,b)} = L^{(r \vee a, s \wedge b)}$, both as sets and as topological vector spaces.

3.3.3. Mixed-Norm Lebesgue Spaces $L_m^{p,q}$

An interesting class of function spaces, close relatives to the Lebesgue L^p spaces, consists of the so-called L^P spaces with mixed norm. Let (X,μ) and (Y,ν) be two σ-finite measure spaces and $1 \leqslant p,q \leqslant \infty$ (in the general case, one considers n such spaces and n-tuples

$P := (p_1, p_2, \dots, p_n)$). Then, a function $f(x, y)$ measurable on the product space $X \times Y$ is said to belong to $L^{(p,q)}(X \times Y)$ if the number obtained by taking successively the p-norm in x and the q-norm in y, in that order, is finite (exchanging the order of the two norms leads in general to a different space). If $p, q < \infty$, the norm reads

$$\|f\|_{(p,q)} = \left(\int_Y \left(\int_X |f(x,y)|^p d\mu(x)\right)^{q/p} d\nu(y)\right)^{1/q}. \tag{3.14}$$

The analogous norm for p or $q = \infty$ is obvious. For $p = q$, one gets the usual space $L^p(X \times Y)$.

These spaces enjoy a number of properties similar to those of the L^p spaces: (i) each space $L^{(p,q)}$ is a Banach space and it is reflexive if and only if $1 < p, q < \infty$; (ii) the conjugate dual of $L^{(p,q)}$ is $L^{(\overline{p},\overline{q})}$, where, as usual, $p^{-1} + \overline{p}^{-1} = 1$, $q^{-1} + \overline{q}^{-1} = 1$; thus the topological conjugate dual coincides with the Köthe dual; (iii) the mixed-norm spaces satisfy a generalized Hölder inequality and have nice interpolation properties.

The case $X = Y = \mathbb{R}^d$ with Lebesgue measure is the important one for signal processing [20, Section 11.1]. More generally, one can add a weight function m and obtain the spaces $L^{p,q}_m(\mathbb{R}^d)$ (we switch to a notation more suitable for the applications):

$$\|f\|^{p,q}_m = \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |f(x,\omega)|^p \, m(x,\omega)^p dx\right)^{q/p} d\omega\right)^{1/q}. \tag{3.15}$$

Here the weight function m is a nonnegative locally integrable function on $\mathbb{R}^{2d}$, assumed to be v-moderate, that is, $m(z_1 + z_2) \leqslant v(z_1)m(z_2)$, for all $z_1, z_2 \in \mathbb{R}^{2d}$, with v a submultiplicative weight function, that is, $v(z_1 + z_2) \leqslant v(z_1)v(z_2)$, for all $z_1, z_2 \in \mathbb{R}^{2d}$. The typical weights are of polynomial growth: $v_s(z) = (1 + |z|)^s$, $s \geqslant 0$.

The space $L^{p,q}_m(\mathbb{R}^{2d})$ is a Banach space for the norm $\|\cdot\|^{p,q}_m$. The duality property is, as expected, $(L^{p,q}_m)^\times = L^{\overline{p},\overline{q}}_{1/m}$. Of course, things simplify when $p = q$: $L^{p,p}_m(\mathbb{R}^{2d}) = L^p_m(\mathbb{R}^{2d})$, a weighted L^p space.

Concerning lattice properties of the family of $L^{p,q}_m$ spaces, we cannot expect more than for the L^p spaces. Two $L^{p,q}_m$ spaces are never comparable, even for the same weight m, so one has to take the lattice generated by intersection and duality.

A different type of mixed-norm spaces is obtained if one takes $X = Y = \mathbb{Z}^d$, with the counting measure. Thus one gets the space $\ell^{p,q}_m(\mathbb{Z}^{2d})$, which consists of all sequences $a = (a_{kn})$, $k, n \in \mathbb{Z}^d$, for which the following norm is finite:

$$\|a\|_{\ell^{p,q}_m} := \left(\sum_{n\in\mathbb{Z}^d} \left(\sum_{k\in\mathbb{Z}^d} |a_{kn}|^p \, m(k,n)^p\right)^{q/p}\right)^{1/q}. \tag{3.16}$$

Contrary to the continuous case, here we do have inclusion relations: if $p_1 \leqslant p_2, q_1 \leqslant q_2$ and $m_2 \leqslant Cm_1$, then $\ell^{p_1,q_1}_{m_1} \subseteq \ell^{p_2,q_2}_{m_2}$.

Discrete mixed-norm spaces have been used extensively in functional analysis and signal processing. For instance, they are key to the proof that certain operators are bounded between two given function spaces, such as modulation spaces (see below) or ℓ^p spaces. In general, a mixed-norm space will prove useful whenever one has a signal consisting of sequences labeled by two indices that play different roles. An obvious example is time-frequency or time-scale analysis: a Gabor or wavelet basis (or frame) is written as $\{\psi_{j,k}, j, k \in \mathbb{Z}\}$, where j indexes the scale or frequency and k the time. More generally, this applies whenever signals are expanded with respect to a dictionary with two indices.

3.3.4. Köthe Function Spaces

The mixed-norm Lebesgue spaces $L_m^{p,q}$ are special cases of a very general class, the so-called *Köthe function spaces*. These have been introduced (and given that name) by Dieudonné [16] and further studied by Luxemburg-Zaanen [21]. The procedure here is entirely parallel to that used in Section 3.2.2 above for introducing the sequence spaces ℓ_{ϕ}.

Let (X, μ) be a σ-finite measure space and M^+ the set of all measurable, non-negative functions on X, where two functions are identified if they differ at most on a μ-null set. A *function norm* is a mapping $\rho : M^+ \to \overline{\mathbb{R}}$ such that

(i) $0 \leqslant \rho(f) \leqslant \infty$, for all $f \in M^+$ and $\rho(f) = 0$ if and only if $f = 0$,

(ii) $\rho(f_1 + f_2) \leqslant \rho(f_1) + \rho(f_2)$, for all $f_1, f_2 \in M^+$,

(iii) $\rho(af) = a\rho(f)$, for all $f \in M^+$, for all $a \geqslant 0$,

(iv) $f_1 \leqslant f_2 \Rightarrow \rho(f_1) \leqslant \rho(f_2)$, for all $f_1, f_2 \in M^+$.

A function norm ρ is said to have the *Fatou property* if and only if $0 \leqslant f_1 \leqslant f_2 \leqslant \dots,\ f_n \in M^+$ and $f_n \to f$ pointwise implies that $\rho(f_n) \to \rho(f)$.

Given a function norm ρ, it can be extended to all complex measurable functions on X by defining $\rho(f) = \rho(|f|)$. Denote by L_ρ the set of all measurable f such that $\rho(f) < \infty$. With the norm $\|f\| = \rho(f), L_\rho$ is a normed space and a subspace of the vector space V of all measurable μ-a.e. finite, functions on X. Furthermore, if ρ has the Fatou property, L_ρ is complete, that is, a Banach space.

A function norm ρ is said to be *saturated* if, for any measurable set $E \subset X$ of positive measure, there exists a measurable subset $F \subset E$ such that $\mu(F) > 0$ and $\rho(\chi_F) < \infty$ (χ_F is the characteristic function of F).

Let ρ be a saturated function norm with the Fatou property. Define

$$\rho'(f) = \sup\left\{\int_X |fg| d\mu : \rho(g) \leqslant 1\right\}. \tag{3.17}$$

Then ρ' is a saturated function norm with the Fatou property and $\rho'' \equiv (\rho')' = \rho$. Hence, $L_{\rho'}$ is a Banach space. Moreover, one has also

$$\rho'(f) = \sup\left\{\left|\int_X fg\, d\mu\right| : \rho(g) \leqslant 1\right\}. \tag{3.18}$$

For each ρ as above, L_ρ is a Banach space and $L_{\rho'} = (L_\rho)^{\#}$, that is, each L_ρ is assaying. The pair $\langle L_\rho, L_{\rho'}\rangle$ is actually a dual pair, although $\langle V^{\#}, V\rangle$ is not. The space $L_{\rho'}$ is called the *Köthe dual* or *α-dual* of L_ρ and denoted by $(L_\rho)^{\alpha}$.

However, $L_{\rho'}$ is in general only a closed subspace of the Banach conjugate dual $(L_\rho)^{\times}$; thus, the Mackey topology $\tau(L_\rho, L_{\rho'})$ is coarser than the ρ-norm topology, which is

$\tau(L_\rho, (L_\rho)^\times)$. This defect can be remedied by further restricting ρ. A function norm ρ is called *absolutely continuous* if $\rho(f_n) \searrow 0$ for every sequence $f_n \in L_\rho$ such that $f_1 \geqslant f_2 \geqslant \ldots \searrow 0$ pointwise a.e. on X. For instance, the Lebesgue L^p-norm is absolutely continuous for $1 \leqslant p < \infty$, but the L^∞-norm is *not*! Also, even if ρ is absolutely continuous, ρ' need not be. Yet, this is the appropriate concept, in view of the following results:

(i) $L_{\rho'} = (L_\rho)^\alpha = (L_\rho)^\times$ if and only if ρ is absolutely continuous;

(ii) L_ρ is reflexive if and only if ρ *and* ρ' are absolutely continuous and ρ has the Fatou property.

Let ρ be a saturated, absolutely continuous function norm on X, with the Fatou property and such that ρ' is also absolutely continuous. Then $\langle L_\rho, L_{\rho'}\rangle$ is a reflexive dual pair of Banach spaces. In addition, the set J of all function norms with these properties is an involutive lattice with respect to the following partial order: $\rho_1 \leqslant \rho_2$ if and only if $\rho_1(f) \leqslant \rho_2(f)$, for every measurable f. The lattice operations are the following:

(i) $(\rho_1 \vee \rho_2)(f) = \max\{\rho_1(f), \rho_2(f)\}$,

(ii) $(\rho_1 \wedge \rho_2)(f) = \inf\{\rho_1(f_1) + \rho_2(f_2); f_1, f_2 \in M^+, f_1 + f_2 = |f|\}$,

(iii) involution : $\rho \leftrightarrow \rho'$.

For the corresponding Banach spaces, we have the relations

$$L_{(\rho_1\vee\rho_2)} = \left(L_{\rho_1} \cap L_{\rho_2}\right)_{\text{proj}}, \quad L_{(\rho_1\wedge\rho_2)} = \left(L_{\rho_1} + L_{\rho_2}\right)_{\text{ind}}. \tag{3.19}$$

Consider now the usual space $V = L^1_{\text{loc}}(X, d\mu)$, with the compatibility and partial inner product defined in Section 2.2.2, so that $V^\# = L^\infty_c(X, d\mu)$. Then the construction outlined above provides $L^1_{\text{loc}}(X, d\mu)$ with the structure of a LBS. Indeed, one has the following result.

Proposition 3.3. *Let J be the set of saturated, absolutely continuous function norms ρ on X, with the Fatou property and such that ρ' is also absolutely continuous. Let $\mathcal{I}$ denote the set $\mathcal{I} := \{L_\rho : \rho \in J$ and $L_\rho \subset L^1_{\text{loc}}\}$. Then $\mathcal{I}$ is a LBS, with the lattice operations defined above.*

More general situations may be considered, for which we refer to [14, Section 4.4].

4. Comparing PIP-Spaces

The definition of LBS/LHS given in Section 3 leads to the proper notion of comparison between two linear compatibilities on the same vector space. Namely, we shall say that a compatibility $\#_1$ is *finer* than $\#_2$, or that $\#_2$ is *coarser* than $\#_1$, if $\mathcal{F}(V, \#_2)$ is an involutive cofinal sublattice of $\mathcal{F}(V, \#_1)$ (given a partially ordered set F, a subset $K \subset F$ is *cofinal* to F if, for any element $x \in F$, there is an element $k \in K$ such that $x \leqslant k$).

Now, suppose that a linear compatiblity # is given on V. Then, every involutive cofinal sublattice of $\mathcal{F}(V, \#)$ defines a coarser PIP-space, and *vice versa*. Thus coarsening is always possible, and will ultimately lead to a minimal PIP-space, consisting of $V^\#$ and V only, that is, the situation of distribution spaces. However, the operation of refining is not always possible; in particular there is no canonical solution, *a fortiori* no unique maximal solution. There are exceptions, however, for instance, when one is given explicitly a larger set of assaying subspaces that also form, or generate, a larger involutive sublattice. To give an example, the weighted L^2 spaces of Section 3.3.1 form an involutive sublattice of the involutive lattice $\mathcal{I}$ of Köthe function spaces of Section 3.3.4; thus, $\mathcal{I}$ is a genuine refinement of the original LHS.

In the case of a LHS, refining is possible, with infinitely many solutions, by use of interpolation methods or the spectral theorem for self-adjoint operators, which are essentially equivalent in this case. In particular, one may always refine a *discrete* scale of Hilbert spaces into a (nonunique) *continuous* one. Indeed, for the scale described in Section 1, Example (ii), one has, by definition, $\mathcal{H}_n = D(A^n)$, the domain of A^n, equipped with the graph norm $\|f\|_n = \|A^n f\|,\ f \in D(A^n)$, for $n \in \mathbb{N}$. Then, for each $0 \leqslant \alpha \leqslant 1$, one may define

$$\mathcal{H}_{n+\alpha} := \left\{ f \in \mathcal{H}_0 : \int_1^\infty s^{2n+2\alpha}\, d\langle f \mid E(s) f\rangle < \infty \right\}, \tag{4.1}$$

where $\{E(s), 1 \leq s < \infty\}$ is the spectral family of A. With the inner product

$$\langle f \mid g\rangle_{n+\alpha} = \langle A^{n+\alpha} f \mid A^{n+\alpha} g\rangle, \quad f, g \in \mathcal{H}_{n+\alpha}, \tag{4.2}$$

$\mathcal{H}_{n+\alpha}$ is a Hilbert space and one has the continuous embeddings

$$\mathcal{H}_{n+1} \hookrightarrow \mathcal{H}_{n+\beta} \hookrightarrow \mathcal{H}_{n+\alpha} \hookrightarrow \mathcal{H}_n, \quad 0 \leqslant \alpha \leqslant \beta \leqslant 1. \tag{4.3}$$

One may go further, as follows. Let φ be any continuous, positive function on $[1, \infty)$ such that $\varphi(t)$ is unbounded for $t \to \infty$, but increases slower than any power t^α $(0 < \alpha 1)$. An example is $\varphi(t) = \log\, t$ $(t \geqslant 1)$. Then $\varphi(A)$ is a well-defined self-adjoint operator, with domain

$$D(\varphi(A)) = \left\{ f \in \mathcal{H}_0 : \int_1^\infty (1 + \varphi(s))^2\, d\langle f \mid E(s) f\rangle < \infty \right\}. \tag{4.4}$$

With the corresponding inner product

$$\langle f \mid g\rangle_\varphi = \langle f \mid g\rangle + \langle \varphi(A) f \mid \varphi(A) g\rangle, \tag{4.5}$$

$D(\varphi(A))$ becomes a Hilbert space $\mathcal{H}_\varphi$. For every α, $0 < \alpha \leqslant 1$, one has, with proper inclusions and continuous embeddings,

$$\mathcal{H}_\alpha \hookrightarrow \mathcal{H}_\varphi \hookrightarrow \mathcal{H}_0. \tag{4.6}$$

This can be continued as far as one wants, with the result that every scale of Hilbert spaces possesses infinitely many proper refinements which are themselves chains of Hilbert spaces [14, Chapter 5].

Another type of refinement consists in refining a RHS $\Phi \subset \mathcal{H} \subset \Phi^\times$, by inserting a number of intermediate spaces, called *interspaces*, namely, spaces $\mathcal{E}$ such that $\Phi \hookrightarrow \mathcal{E} \hookrightarrow \Phi^\times$ (which implies that the conjugate dual $\mathcal{E}^\times$ is also an interspace). Upon some additional conditions, the most important of which being that Φ be dense in $\mathcal{E} \cap \mathcal{F}$ with its projective topology, for any pair $\mathcal{E}, \mathcal{F}$ of interspaces, one obtains in that way a proper refining of the original RHS. With this construction, which goes under the name of *multiplication framework*, one succeeds, for instance, in defining a valid (partial) multiplication between distributions. A thorough analysis may be found in [14, Section 6.3].

5. Operators on PIP-Spaces

5.1. General Definitions

As already mentioned, the basic idea of (indexed) PIP-spaces is that vectors should not be considered individually, but only in terms of the subspaces V_r ($r \in F$ or $r \in I$), the building blocks of the structure; see (3.1). Correspondingly, an operator on a PIP-space should be defined in terms of assaying subspaces only, with the proviso that only bounded operators between Hilbert or Banach spaces are allowed. Thus an operator is a *coherent collection* of bounded operators. More precisely, one has the following.

Definition 5.1. Given a LHS or LBS $V_I = \{V_r, r \in I\}$, an *operator* on V_I is a map $A : \mathfrak{D}(A) \to V$, such that

(i) $\mathfrak{D}(A) = \bigcup_{q \in \mathrm{d}(A)} V_q$, where $\mathrm{d}(A)$ is a nonempty subset of I,

(ii) for every $q \in \mathrm{d}(A)$, there exists a $p \in I$ such that the restriction of A to V_q is linear and continuous into V_p (we denote this restriction by A_{pq}),

(iii) A has no proper extension satisfying (i) and (ii).

The linear bounded operator $A_{pq} : V_q \to V_p$ is called a *representative* of A. In terms of the latter, the operator A may be characterized by the set $\mathrm{j}(A) = \{(q,p) \in I \times I : A_{pq} \text{ exists}\}$. Thus the operator A may be identified with the collection of its representatives:

$$A \simeq \{A_{pq} : V_q \longrightarrow V_p : (q,p) \in \mathrm{j}(A)\}. \tag{5.1}$$

By condition (ii), the set $\mathrm{d}(A)$ is obtained by projecting $\mathrm{j}(A)$ on the "first coordinate" axis. The projection $\mathrm{i}(A)$ on the "second coordinate" axis plays, in a sense, the role of the range of A. More precisely,

$$\begin{aligned} \mathrm{d}(A) &= \{q \in I : \text{there is a } p \text{ such that } A_{pq} \text{ exists}\}, \\ \mathrm{i}(A) &= \{p \in I : \text{there is a } q \text{ such that } A_{pq} \text{ exists}\}. \end{aligned} \tag{5.2}$$

The following properties are immediate see the (see Figure 2):

(i) $\mathtt{d}(A)$ is an initial subset of I: if $q \in \mathtt{d}(A)$ and $q' < q$, then $q' \in \mathtt{d}(A)$, and $A_{pq'} = A_{pq}E_{qq'}$, where $E_{qq'}$ is a representative of the unit operator (this is what we mean by a 'coherent' collection),

(ii) $\mathtt{i}(A)$ is a final subset of I: if $p \in \mathtt{i}(A)$ and $p' > p$, then $p' \in \mathtt{i}(A)$ and $A_{p'q} = E_{p'p}A_{pq}$.

(iii) $\mathtt{j}(A) \subset \mathtt{d}(A) \times \mathtt{i}(A)$, with strict inclusion in general.

We denote by $\mathrm{Op}(V_I)$ the set of all operators on V_I. Of course, a similar definition may be given for operators $A : V_I \to Y_K$ between two LHSs or LBSs.

Since $V^{\#}$ is dense in V_r, for every $r \in I$, an operator may be identified with a separately continuous sesquilinear form on $V^{\#} \times V^{\#}$. Indeed, the restriction of any representative A_{pq}

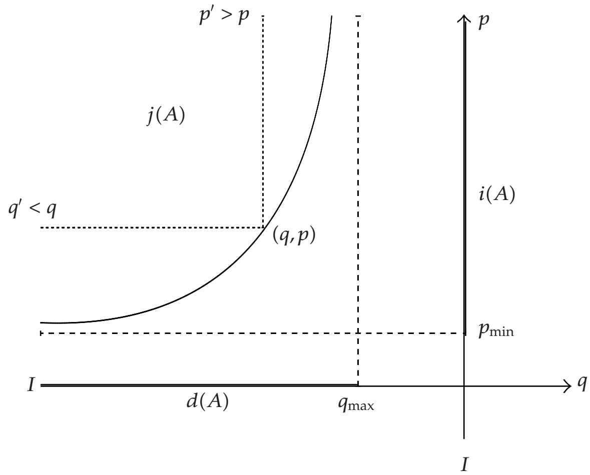


Figure 2: Characterization of the operator A, in the case of a scale.

to $V^{\#} \times V^{\#}$ is such a form, and all these restrictions coincide. Equivalently, an operator may be identified with a continuous linear map from $V^{\#}$ into V (continuity with respect to the respective Mackey topologies).

But the idea behind the notion of operator is to keep also the *algebraic operations* on operators; namely, we define the following operations:

(i) *Adjoint*: Every $A \in \mathrm{Op}(V_I)$ has a unique adjoint $A^{\times} \in \mathrm{Op}(V_I)$, defined by the relation

$$\langle A^{\times} x \mid y \rangle = \langle x \mid Ay \rangle, \quad \text{for } y \in V_r,\ r \in \mathrm{d}(A),\ x \in V_{\overline{s}},\ s \in \mathrm{i}(A), \tag{5.3}$$

that is, $(A^{\times})_{\overline{rs}} = (A_{sr})^*$ (usual Hilbert/Banach space adjoint). ¡list-item¿¡label/¿

It follows that $A^{\times\times} = A$, for every $A \in \mathrm{Op}(V_I)$: no extension is allowed, by the maximality condition (iii) of Definition 5.1.

(ii) *Partial Multiplication*: The product AB is defined if and only if there is a $q \in \mathrm{i}(B) \cap \mathrm{d}(A)$, that is, if and only if there is a continuous factorization through some V_q:

$$V_r \xrightarrow{B} V_q \xrightarrow{A} V_s, \quad \text{that is,}\ (AB)_{sr} = A_{sq} B_{qr}. \tag{5.4}$$

It is worth noting that, for a LHS/LBS, the domain $\mathfrak{D}(A)$ is always a vector subspace of V (this is not true for a general PIP-space). Therefore, $\mathrm{Op}(V_I)$ is a vector space and a *partial *-algebra* [22].

The concept of PIP-space operator is very simple, yet it is a far-reaching generalization of bounded operators. It allows indeed to treat on the same footing all kinds of operators, from bounded ones to very singular ones. By this, we mean the following, loosely speaking. Take

$$V_r \subset V_o \simeq V_{\overline{o}} \subset V_s \quad (V_o = \text{Hilbert space}). \tag{5.5}$$

Three cases may arise:

(i) if A_{oo} exists, then A corresponds to a bounded operator $V_o \to V_o$,

(ii) if A_{oo} does not exist, but only $A_{or} : V_r \to V_o$, with $r < o$, then A corresponds to an unbounded operator, with domain $D(A) \supset V_r$,

(iii) if no A_{or} exists, but only $A_{sr} : V_r \to V_s$, with $r < o < s$, then A corresponds to a singular operator, with Hilbert space domain possibly reduced to $\{0\}$.

5.2. Special Classes of Operators on PIP-Spaces

Exactly as for Hilbert or Banach spaces, one may define various types of operators between PIP-spaces, in particular LBS/LHSs. We discuss briefly the most important classes, namely, regular operators, homomorphisms and isomorphisms, unitary operators, symmetric operators, and orthogonal projections. Further details may be found in the monograph [14].

5.2.1. Regular and Totally Regular Operators

An operator A on a nondegenerate PIP-space V_I, with positive-definite partial inner product, in particular, a LBS/LHS, is called *regular* if $\mathtt{d}(A) = \mathtt{i}(A) = I$ or, equivalently, if $A : V^{\#} \to V^{\#}$ and $A : V \to V$ continuously for the respective Mackey topologies. This notion depends only on the pair $(V^{\#}, V)$, not on the particular compatibility #. The set of all regular operators $V_I \to V_I$ is denoted by $\mathrm{Reg}(V_I)$. Thus a regular operator may be multiplied both on the left and on the right by an arbitrary operator. Clearly, the set $\mathrm{Reg}(V_I)$ is a $*$-algebra and can often be identified with an O^*-algebra [22, 23].

We give two examples.

(1) If $V = \omega$, $V^{\#} = \varphi$, then $\mathrm{Op}(\omega)$ consists of arbitrary infinite matrices and $\mathrm{Reg}(\omega)$ of infinite matrices with finite rows and finite columns.

(2) If $V = \mathcal{S}^{\times}$, $V^{\#} = \mathcal{S}$, then $\mathrm{Op}(\mathcal{S}^{\times})$ consists of arbitrary tempered kernels, while $\mathrm{Reg}(\mathcal{S}^{\times})$ contains those kernels that can be extended to $\mathcal{S}^{\times}$ and map $\mathcal{S}$ into itself. A nice example is the Fourier transform.

An operator A is called *totally regular* if $\mathtt{j}(A)$ contains the diagonal of $I \times I$, that is, A_{rr} exists for every $r \in I$ or A maps every V_r into itself continuously.

5.2.2. Homomorphisms

Let V_I, Y_K be two LHSs or LBSs. An operator $A \in \mathrm{Op}(V_I, Y_K)$ is called a *homomorphism* if

(i) $\mathtt{j}(A) = I \times K$ and $\mathtt{j}(A^{\times}) = K \times I$,

(ii) $f \#_I g$ implies that $Af \#_K Ag$.

We denote by $\mathrm{Hom}(V_I, Y_K)$ the set of all homomorphisms from V_I into Y_K. The following properties are easy to prove:

(i) $A \in \mathrm{Hom}(V_I, Y_K)$ if and only if $A^\times \in \mathrm{Hom}(Y_K, V_I)$,

(ii) if $A \in \mathrm{Hom}(V_I, Y_K)$, then $\mathsf{j}(A^\times A)$ contains the diagonal of $I \times I$ and $\mathsf{j}(AA^\times)$ contains the diagonal of $K \times K$.

The homomorphism $M \in \mathrm{Hom}(W_I, Y_K)$ is a *monomorphism* if $MA = MB$ implies that $A = B$, for any two elements of $A, B \in \mathrm{Hom}(V_I, W_L)$, where V_I is any PIP-space. Typical examples of monomorphisms are the inclusion maps resulting from the restriction of a support. Take for instance, $L^1_{\mathrm{loc}}(X, d\mu)$, the space of locally integrable functions on a measure space (X, μ). Let Ω be a measurable subset of X and Ω' its complement, both of nonzero measure, and construct the space $L^1_{loc}(\Omega, d\mu)$, which is a PIP-subspace of $L^1_{\mathrm{loc}}(X, d\mu)$ (see Section 5.2.5). Given $f \in L^1_{\mathrm{loc}}(X, d\mu)$, define $f^{(\Omega)} = f\chi_\Omega$, where χ_Ω is the characteristic function of χ_Ω. Then we obtain an injection monomorphism $M^{(\Omega)} : L^1_{\mathrm{loc}}(\Omega, d\mu) \to L^1_{\mathrm{loc}}(X, d\mu)$ as follows:

$$\left(M^{(\Omega)} f^{(\Omega)}\right)(x) = \begin{cases} f^{(\Omega)}(x), & \text{if } x \in \Omega, \\ 0, & \text{if } x \notin \Omega, \end{cases} \qquad f^{(\Omega)} \in L^1_{\mathrm{loc}}(\Omega, d\mu). \tag{5.6}$$

If we consider the lattice of weighted Hilbert spaces $\{L^2(r)\}$ in this PIP-space, then the correspondence $r \leftrightarrow r^{(\Omega)} = r\chi_\Omega$ is a bijection between the corresponding involutive lattices.

The homomorphism $A \in \mathrm{Hom}(V_I, Y_K)$ is an *isomorphism* if there exists a homomorphism $B \in \mathrm{Hom}(Y_K, V_I)$ such that $BA = 1_V$, $AB = 1_Y$, where 1_V, 1_Y denote the identity operators on V_I, Y_K, respectively.

For instance, the Fourier transform is an isomorphism of the Schwartz RHS $\mathcal{S} \subset L^2 \subset \mathcal{S}^\times$ onto itself and, similarly, of the Feichtinger triplet (6.16) onto itself. In both cases, the property extends to several scales of lattices interpolating between the two extreme spaces, for instance, the Hilbert chain of the Hermite representation of tempered distributions.

5.2.3. Unitary Operators and Group Representations

The operator $U \in \mathrm{Op}(V_I, Y_K)$ is *unitary* if $U^\times U$ and $UU^\times$ are defined and $U^\times U = 1_V, UU^\times = 1_Y$. We emphasize that unitary operators need *not* be homomorphisms ! In fact, unitarity is a rather weak property and it is insufficient for group representations.

Thus a unitary representation of a group G into a PIP-space V_I is defined as a homomorphism of G into the *unitary isomorphisms* of V_I. Given such a unitary representation U of G into V_I, where the latter has the central Hilbert space $\mathscr{H}_0$, consider the representative $U_{00}(g)$ of $U(g)$ in $\mathscr{H}_0$. Then $g \mapsto U_{00}(g)$ is a unitary representation of G into $\mathscr{H}_0$, in the usual sense.

To give an example, let V_I be the scale built on the powers of the operator (Hamiltonian) $H = -\Delta + v(\mathbf{r})$ on $L^2(\mathbb{R}^3, d\mathbf{x})$, where Δ is the Laplacian on $\mathbb{R}^3$ and v is a (nice) rotation invariant potential. The system admits as symmetry group $G = \mathrm{SO}(3)$ (the full symmetry group might be larger; for instance, the Coulomb potential admits SO(4) as symmetry group for its bound states.) and the representation U_{00} is the natural representation of SO(3) in $L^2(\mathbb{R}^3)$:

$$[U_{00}(\rho)\psi](\mathbf{x}) = \psi\left(\rho^{-1}\mathbf{x}\right), \quad \rho \in \mathrm{SO}(3). \tag{5.7}$$

Then U_{00} extends to a unitary representation U by totally regular isomorphisms of V_I. Angular momentum decompositions, corresponding to irreducible representations of SO(3), extend to V_I as well. In addition, this is a good setting also for representations of the Lie algebra $\mathfrak{so}(3)$. Notice that the representation U is totally regular, but this need not be the case. For instance, if the potential v is not rotation invariant, U will no longer be totally regular, although it is still an isomorphism.

5.2.4. Symmetric Operators

An operator $A \in \mathrm{Op}(V_I)$ is *symmetric* if $A^\times = A$. Since one has $A^{\times\times} = A$, no extension is allowed, by the maximality condition. Thus symmetric operators behave essentially like bounded self-adjoint operators in a Hilbert space. Yet, they can be very singular, as indicated above, for a chain

$$V^{\#} \subset \cdots \subset V_r \subset V_o = V_{\overline{o}} \subset V_s \subset \cdots \subset V. \tag{5.8}$$

In this case, the question is whether a symmetric operator $A \in \mathrm{Op}(V_I)$ has a self-adjoint restriction to the central Hilbert space V_o. In a Hilbert space context, an answer is given by the celebrated KLMN theorem (KLMN stands for Kato, Lax, Lions, Milgram, Nelson). Actually, this classical result already has a distinct PIP-space flavor. Thus is not surprising that the KLMN theorem has a natural generalization to a LHS or a PIP-space with positive-definite partial inner product and central Hilbert space $V_o = V_{\overline{o}}$, and a quadratic form version as well [14, Section 3.3.5].

An interesting application is a correct description of very singular Hamiltonians in quantum mechanics, typically with δ or δ' interactions. For instance, one can treat in this way the Kronig-Penney crystal model, which consists of a δ potential at each node of a lattice, in one, two, or three dimensions [24, 25].

5.2.5. Orthogonal Projections, Bases, Frames

An operator P on a nondegenerate PIP-space V, respectively, a LBS/LHS V_I, is an *orthogonal projection* if $P \in \mathrm{Hom}(V_I)$ and $P^2 = P = P^\times$. It follows immediately from the definition that an orthogonal projection is totally regular, that is, $\mathsf{j}(P)$ contains the diagonal $I \times I$, or still that P leaves every assaying subspace invariant. Equivalently, P is an orthogonal projection if P is an idempotent operator (that is, $P^2 = P$) such that $\{Pf\}^{\#} \supseteq \{f\}^{\#}$ for every $f \in V$ and $\langle g \mid Pf\rangle = \langle Pg \mid f\rangle$ whenever $f\#g$. We denote by $\mathrm{Proj}(V)$ the set of all orthogonal projections in V and similarly for $\mathrm{Proj}(V_I)$.

These projection operators enjoy several properties similar to those of Hilbert space projectors. Two of them are of special interest in the present context.

(i) Given a nondegenerate PIP-space V, there is a natural notion of subspace, called *orthocomplemented*, which guarantees that such a subspace W of V is again a nondegenerate PIP-space with the induced compatibility relation and the restriction of the partial inner product. There are equivalent topological conditions, so that orthocomplemented subspaces are the proper PIP-subspaces [26]. Then the basic theorem about projections states that a subspace W of V is orthocomplemented if and only if W is the range of an orthogonal projection $P \in \mathrm{Proj}(V)$, that is, $W = PV$. Then $V = W \oplus Z$, where Z is another orthocomplemented subspace.

(ii) An orthogonal projection P is of finite rank if and only if $W = \mathrm{Ran}\, P \subset V^{\#}$ and $W \cap W^{\perp} = \{0\}$ (this condition is, of course, superfluous when the partial inner product is positive definite).

There is a natural partial order on the set of projections:

$$P_W \leq P_Y \quad \text{if and only if} \quad W \subseteq Y, \tag{5.9}$$

but the lattice properties of $\mathrm{Proj}(V)$ are unknown. Thus we expect that quantum logic may be reformulated in a PIP-space language only under additional restrictions on V.

Property (ii) has important consequences for the structure of bases. First we recall that a sequence $\{e_n, n = 1, 2, \ldots\}$ of vectors in a Banach space E is a *Schauder basis* if, for every $f \in E$, there exists a unique sequence of scalar coefficients $\{c_k, k = 1, 2, \ldots\}$ such that $\lim_{m\to\infty} \| f - \sum_{k=1}^{m} c_k e_k \| = 0$. Then one may write

$$f = \sum_{k=1}^{\infty} c_k e_k. \tag{5.10}$$

The basis is *unconditional* if the series (5.10) converges unconditionally in E (i.e., it keeps converging after an arbitrary permutation of its terms).

A standard problem is to find, for instance, a sequence of functions that is an unconditional basis for *all* the spaces $L^p(\mathbb{R}), 1 < p < \infty$. In the PIP-space language, this statement means that the basis vectors must belong to $V^{\#} = \bigcap_{1<p<\infty} L^p(\mathbb{R})$. Also, since (5.10) means that f may be approximated by finite sums, the property (ii) of orthogonal projections implies that all the basis vectors must belong to $V^{\#}$. Some examples are given in Section 6.2.5.

Actually, in the context of signal processing, orthogonal (in the Hilbert sense) bases are not enough; one needs also biorthogonal bases and, more generally, *frames*. We recall that a countable family of vectors $\{\psi_n\}$ in a Hilbert space $\mathscr{H}$ is called a *frame* if there are two positive constants $\mathtt{m}, \mathtt{M}$, with $0 < \mathtt{m} \leqslant \mathtt{M} < \infty$, such that

$$\mathrm{m}\|f\|^2 \leqslant \sum_{n=1}^{\infty} |\langle \psi_n \mid f \rangle|^2 \leqslant \mathrm{M}\|f\|^2, \quad \forall f \in \mathscr{H}. \tag{5.11}$$

The two constants $\mathtt{m}, \mathtt{M}$ are called *frame bounds*. If $\mathtt{m} = \mathtt{M}$, the frame is said to be *tight*. Consider the analysis operator $C : \mathscr{H} \to \ell^2$ defined by $C : f \mapsto \{|\langle \psi_n \mid f \rangle|\}$ and the frame operator

$S = C^*C$. Then the vectors $\widetilde{\psi}_n = S^{-1}\psi_n$ also constitute a frame, called the *canonical dual frame*, and one has the (strongly converging) expansions

$$f = \sum_{n=1}^{\infty} \left| \langle \psi_n \mid f \rangle \right| \widetilde{\psi}_n = \sum_{n=1}^{\infty} \left| \langle \widetilde{\psi}_n \mid f \rangle \right| \psi_n. \tag{5.12}$$

Then the considerations made above for bases should apply to frame vectors as well, that is, the vectors $\psi_n,\ \widetilde{\psi}_n$ should also belong to $V^{\#}$.

6. Applications of PIP-Spaces

6.1. Applications in Mathematical Physics

PIP-spaces have found many applications in mathematical physics, mostly in quantum physics. We will sketch a few of them in this section. Most of what follows is described in detail in [14].

6.1.1. Dirac Formalism in Quantum Mechanics

The mathematical description of a quantum system rests on three basic principles: (i) The *superposition principle*, which implies that the set of states of the system has a linear structure; (ii) The notion of *transition amplitude*, given by an inner product: $A(\psi_1 \to \psi_2) = \langle \psi_2 \mid \psi_1 \rangle$, which yields transition probabilities by $P(\psi_1 \to \psi_2) = |\langle \psi_2 \mid \psi_1 \rangle|^2$; and (iii) The *probabilistic interpretation*, which requires that $\langle \psi \mid \psi \rangle = \|\psi\|^2 > 0$, whenever $\psi \neq 0$.

Combining these basic principles implies that the set of states of the system is a positive-definite inner product space Φ, that is, a pre-Hilbert space. On this basis, Dirac developed a formalism for quantum physics with great computational capacity and broad predictive power. The essential features of Dirac's formalism are the following.

(i) Physical observables are represented by linear operators in the space Φ and these operators form an algebra. Therefore, it makes sense to arbitrarily add and multiply operators to form new operators.

(ii) For a given quantum physical system, there exist complete systems of commuting observables (CSCO) in the algebra of observables. The system of eigenvectors for a chosen CSCO provides a basis for the space Φ, that is, every vector $\phi \in \Phi$ can be expanded into the eigenvectors of the CSCO.

In the simplest case, a CSCO consists of only one observable A, with a mixed spectrum consisting of discrete eigenvalues $\{\lambda_n\} = \sigma_p(A)$ and a continuous part $\{\lambda\} = \sigma_c(A)$. The corresponding eigenvectors, written as $|\lambda_n\rangle,\ |\lambda\rangle$, respectively, obey "orthogonality" relations

$$\langle \lambda_m \mid \lambda_n \rangle = \delta_{mn}, \quad \langle \lambda \mid \lambda_n \rangle = 0, \quad \langle \lambda \mid \lambda' \rangle = \delta(\lambda - \lambda'). \tag{6.1}$$

Then every $\phi \in \Phi$ can be expanded as

$$\phi = \sum_n |\lambda_n\rangle\langle \lambda_n \mid \phi \rangle + \int_{\sigma_c(A)} d\lambda |\lambda\rangle\langle \lambda \mid \phi \rangle. \tag{6.2}$$

Clearly the eigenvectors $|\lambda\rangle$ cannot belong to the pre-Hilbert space Φ, nor to its completion $\mathscr{H}$. Thus Dirac's formalism, while extremely practical and used by physicists on a daily basis, is not mathematically well defined.

For that reason, von Neumann formulated a rigorous version of quantum mechanics, in a pure Hilbert space language. His formulation consists in the following two axioms: (i) *Pure states* are represented by rays (i.e., one-dimensional subspaces) in a Hilbert space $\mathscr{H}$; and (ii) *Observables* are represented by self-adjoint operators in $\mathscr{H}$. This formulation is well defined mathematically, but too restrictive. Nonnormalizable eigenvectors, corresponding to points of a continuous spectrum, cannot belong to $\mathscr{H}$, yet they are extremely useful and often have a clear physical meaning (plane waves, for instance). Observables may be unbounded, so that domain considerations must be taken into account. In particular, unbounded operators may not always be multiplied. Thus it is understandable that the large majority of physicists stay with Dirac's formalism.

This had prompted several authors [27–31] to propose a rigorous version in terms of a RHS $\Phi \subset \mathscr{H} \subset \Phi^\times$. In this scheme, the space Φ is constructed from the basic observables (*labeled* observables) of the system at hand and is interpreted as the space of all physically preparable states. The conjugate dual $\Phi^\times$ contains idealized states (probes), identified with measurement devices. In that context, let A be an observable, represented by a self-adjoint operator in $\mathscr{H}$ such that $A : \Phi \to \Phi$ continuously. Then A may be transposed by duality to a linear operator $A^\times : \Phi^\times \to \Phi^\times$, which is an extension of $A^\dagger := A^* \upharpoonright \Phi$, where A^* is the usual Hilbert space adjoint operator, namely,

$$\langle \phi \mid A^\times F \rangle = \langle A\phi \mid F \rangle, \quad \forall \phi \in \Phi,\ F \in \Phi^\times. \tag{6.3}$$

For such an operator, the vector $\xi_\lambda \in \Phi^\times$ is called a *generalized eigenvector* of A, with eigenvalue $\lambda \in \mathbb{C}$, if it satisfies

$$\langle \phi \mid A^\times \xi_\lambda \rangle := A^\times \xi_\lambda(\phi) = \overline{\lambda}\ \xi_\lambda(\phi) \equiv \overline{\lambda} \langle \phi \mid \xi_\lambda \rangle, \quad \forall\, \phi \in \Phi. \tag{6.4}$$

Then the spectral theorem of Gel'fand-Maurin [5, 6] states that A possesses in $\Phi^\times$ a complete orthonormal set of generalized eigenvectors $\xi_\lambda \in \Phi^\times,\ \lambda \in \mathbb{R}$. This means that, for any two $\phi, \psi \in \Phi$, one has (we split again into the discrete and the continuous part of the spectrum of A)

$$\langle \phi \mid \psi \rangle = \sum_n \langle \phi \mid \lambda_n \rangle \langle \lambda_n \mid \psi \rangle + \int \langle \phi \mid \lambda \rangle \langle \lambda \mid \psi \rangle \rho(\lambda) d\lambda, \tag{6.5}$$

where ρ is a non-negative integrable function. In that way one recovers essentially Dirac's bra-and-ket formalism. This approach is based on a RHS, but the construction is such that a PIP-space version is easily obtained—and is in fact closer to Dirac's spirit. For more details, see [14, Section 7.1.1].

6.1.2. Symmetries, Singular Interactions in Quantum Mechanics

Several other topics in quantum mechanics can be advantageously formulated in a RHS or PIP-space language, for instance, the implementation of symmetries, with the two dual points of view, the active one and the passive one [32]. A symmetry group is represented by a unitary representation in $\mathscr{H}$ that extends to a unitary representation U in the enveloping PIP-space, in the sense defined in Section 5.2.3. Then, in accordance with the physical interpretation given above, the active point of view corresponds to the action of U in $V^{\#} \equiv \Phi$, the passive one to the action on $V \equiv \Phi^{\times}$.

Another problem is a correct definition of a Hamiltonian with a singular interaction, already mentioned in Section 5.2.4. In the simplest case, the standard definition is $H = -(\Delta/2m) + V$, where the interaction V is given by some reasonable function (potential). However, there are cases where a singular interaction is needed, for instance when V is replaced formally by a δ function or a δ' function, with support in a point (or several) or a manifold of lower dimension. Then the usual formulation is based on von Neumann's theory of self-adjoint extensions of symmetric operators, sometimes coupled with Krein's formula [33]. But here the PIP-space approach is a convenient substitute to that approach, as shown in [24, 25] and [14, Section 7.1.3].

6.1.3. *Quantum Scattering Theory*

In scattering theory, it is common to use scales of Hilbert spaces built on the powers of $A_1 := (1 + |\mathbf{x}|^2)$ or $A_2 := (1 + |\mathbf{p}|^2)$, and the LHS obtained by combinations of both. This example contains the Sobolev spaces (the scale built on A_2), the weighted spaces L^2_s (the scale built on A_1), and spaces of mixed type. In particular, operators of the form $f(\mathbf{x})g(\mathbf{p})$, for suitable functions f, g, play an essential role in the so-called phase-space approach to scattering theory and they may be controlled by this LHS. For instance, their trace ideal properties may be derived in this way and they are used for proving the absence of singular continuous spectrum by the limiting absorption principle.

On the other hand, the Weinberg-van Winter (WVW) formulation of scattering theory [34–36] has a very natural interpretation in terms of a LHS, whose components, including the extreme ones, are Hilbert spaces consisting of functions analytic in a sector; thus the indexing parameter is the opening angle of that sector. This technique has allowed to show that the WVW formalism is a particular case of the Complex Scaling Method [14, Section 7.2.3], a result hitherto unknown.

6.1.4. *Quantum Field Theory*

Mathematically rigorous formulations of QFT rely heavily on a RHS or a PIP-space approach, primarily Wightman's axiomatic formulation. There, indeed, a quantum field is defined as an operator-valued distribution, which is customarily written in terms of an unsmeared field (field at a point) $A(x)$, as

$$A(f) = \int_{\mathbb{R}^4} A(x)f(x)dx, \quad f \in \mathcal{S}(\mathbb{R}^4). \tag{6.6}$$

Under quite reasonable assumptions, the unsmeared field can be defined as a map from $\mathcal{S}(\mathbb{R}^4)$ into $\mathrm{Op}(V)$, where V is a conveniently chosen PIP-space. This allows to give the previous formula a rigorous mathematical meaning [14, Section 7.3.1].

Another PIP-space version of QFT is the Fock construction (tensor algebra) over the RHS

$$\mathcal{S}(\mathcal{V}_m^+) \hookrightarrow L^2(\mathcal{V}_m^+, d\mu) \hookrightarrow \mathcal{S}^\times(\mathcal{V}_m^+), \tag{6.7}$$

where $\mathcal{V}_m^+$ denotes the forward mass shell $\mathcal{V}_m^+ = \{p \in \mathbb{R}^4 : p^2 = m^2,\ p_0 > 0\}$ and $d\mu$ the Lorentz invariant measure d^3p/p_0 on it. Write $\Phi_1 = \mathcal{S}(\mathcal{V}_m^+)$, the space of "good" one-particle states. Then define

$$\Phi_n = \Phi_1^{\bigotimes_s n}, \tag{6.8}$$

where the right-hand side denotes the symmetrized tensor product of n copies of Φ_1, corresponding to n-boson states. Again Φ_n is reflexive, complete, and nuclear with respect to its natural topology, and it can be described as the end space of a scale of Hilbert spaces. Finally, define

$$\Phi = \bigoplus_{n=0}^{\infty} \Phi_n, \tag{6.9}$$

that is, the topological direct sum of the component spaces. Elements of Φ are finite sequences $f = \{f_0, f_1, \ldots, f_n, \ldots\}$, $f_0 \in \mathbb{C}$, $f_n \in \Phi_n$, that is, totally symmetric functions of Schwartz type. The space Φ is reflexive, complete, and nuclear with respect to the direct sum topology. Its dual is the topological product

$$\Phi^\times = \prod_{n=0}^{\infty} \Phi_n^\times. \tag{6.10}$$

Thus we get a suitable RHS, in which the central Hilbert space $\mathcal{H}$ is Fock space, that is, the tensor algebra $\Gamma(\Phi_1)$ over Φ_1.

Other examples are the construction of QFT via the Borchers algebra, Nelson's Euclidean field theory, or the precise treatment of unsmeared fields (fields at a point). See [14, Section 7.3] for a detailed presentation.

6.1.5. Representations of Lie Groups and Lie Algebras

Let us return to the situation described in Section 5.2.3. We start with a strongly continuous unitary representation U_{00} of a Lie group G in a Hilbert space $\mathcal{H}_0$ and seek to build a PIP-space V_I, with $\mathcal{H}_0$ being its central Hilbert space, such that U_{00} extends to a unitary representation U into V_I.

The solution of this problem is well known from Nelson's theory of analytic vectors. Let $\overline{\Delta}$ be the closure of the Nelson operator $\Delta := \sum_{j=1}^n X_j^2$, where $\{X_j, j = 1, \ldots, n\}$ are the representatives under U_{00} of the elements of a basis of the Lie algebra $\mathfrak{g}$ of G. Δ is essentially self-adjoint on the so-called Gårding domain $\mathscr{H}_0^G$, $\overline{\Delta}$ is self-adjoint, and $\overline{\Delta} \geqslant 0$. Define $V_I := \{\mathscr{H}_n, n \in \mathbb{Z}\}$ as the canonical scale of Hilbert spaces generated by the operator $(\overline{\Delta} + 1)$:

$$V^{\#} := \mathfrak{D}^{\infty}\left(\overline{\Delta}\right) \hookrightarrow \mathscr{H}_0 \hookrightarrow V := \mathfrak{D}_{\overline{\infty}}\left(\overline{\Delta}\right). \tag{6.11}$$

First, one has $\mathfrak{D}^{\infty}(\overline{\Delta}) = \bigcap_{n=1}^{\infty} D(\overline{\Delta}^n) = \mathscr{H}_0^{\infty}$, the space of C^{∞}-vectors of U_{00}. Next, for every $g \in G$, $U_{00}(g)$ leaves each $\mathscr{H}_n$, $n \in \mathbb{N}$, invariant and its restriction $U_{nn}(g) : \mathscr{H}_n \to \mathscr{H}_n$ is continuous; thus it can be transposed to a continuous map $U_{nn}(g^{-1}) : \mathscr{H}_{\overline{n}} \to \mathscr{H}_{\overline{n}}$. It follows that U_{00} extends to a unitary representation U in the LHS V_I. Corresponding to the triplet (6.11), we have three representations of G, namely, U_{00}, its restriction $U_{\infty\infty}$, and the dual $U_{\overline{\infty\infty}}$ of the latter, which is an extension of the first two. All three are continuous. Moreover, if one of the three is topologically irreducible (i.e., there is no proper invariant closed subspace), so are the other two.

In addition to the representations of the group G, the scale V_I is the natural tool for studying the properties of the operators representing elements of the Lie algebra $\mathfrak{g}$ or the universal enveloping algebra $\mathfrak{U}(\mathfrak{g})$ of G. For every element $x \in \mathfrak{g}$ or $L \in \mathfrak{U}(\mathfrak{g})$, the representative $U(x)$, respectively $U(L)$, originally defined on $\mathscr{H}_0^G$, extends to a regular operator on V_I. These regular operators have in general no $\{0,0\}$-representative, since x and L are represented in $\mathscr{H}_0$ by unbounded operators. As in the case of the group G, one gets three *-representations of the enveloping algebra $\mathfrak{U}(\mathfrak{g})$, and in particular of the Lie algebra $\mathfrak{g}$, in the three spaces of the triplet (6.11). Namely, one has, for every $L, L_1, L_2 \in \mathfrak{U}(\mathfrak{g})$,

$$\begin{gathered} U(L_1)U(L_2) = U(L_1 L_2), \\ U(L^+) = U(L)^{\times}, \end{gathered} \tag{6.12}$$

where $L \leftrightarrow L^+$ is the involution on $\mathfrak{U}(\mathfrak{g})$. These representations have the same irreducibility properties as the corresponding ones of the group. See [14, Section 7] for further details.

6.2. Applications in Analysis and Signal Processing

Many families of function spaces of interest in analysis or signal processing are, or contain, lattices of Banach spaces. To quote a few: amalgam spaces, modulation spaces, Besov spaces, α-modulation spaces, coorbit spaces, which contain many of the previous cases. We shall describe them briefly in succession. For further information about those spaces, we refer to our monograph [14, Chapters 4 and 8].

6.2.1. Amalgam Spaces

A situation intermediate between the mixed-norm spaces $L^{p,q}(\mathbb{R}^{2d})$ (for $m \equiv 1$) and the spaces $\ell^{\,p,q}(\mathbb{Z}^{2d})$ is that of the so-called *amalgam spaces*. They were introduced specifically to overcome the inability of the L^p norms to distinguish between the local and the global behavior of functions. The simplest ones are the spaces $W(L^p, \ell^q)$ of Wiener [37], which consist of functions on $\mathbb{R}$ which are locally in L^p with ℓ^q behavior at infinity, in the sense that the L^p norms over the intervals $(n, n+1)$ form an ℓ^q sequence. It is a Banach space for the norm

$$\|f\|_{p,q} = \left\{ \sum_{n=-\infty}^{\infty} \left[\int_n^{n+1} |f(x)|^p dx \right]^{q/p} \right\}^{1/q}, \quad 1 \leqslant p, q < \infty. \tag{6.13}$$

The following inclusion relations, with all embeddings continuous, derive immediately from those of the L^p and the ℓ^q spaces.

(i) If $q_1 \leqslant q_2$, then $W(L^p, \ell^{q_1}) \subset W(L^p, \ell^{q_2})$.

(ii) If $p_1 \leqslant p_2$, then $W(L^{p_2}, \ell^q) \subset W(L^{p_1}, \ell^q)$.

Thus the smallest space is $W(L^\infty, \ell^1)$ and the largest space is $W(L^1, \ell^\infty)$. As for duality, one has $W(L^p, \ell^q)^\times = W(L^{\overline{p}}, \ell^{\overline{q}})$, for $1 \leqslant p, q < \infty$.

The interesting fact is that, for $1 \leqslant p, q \leqslant \infty$, the set $\mathcal{Q}$ of all amalgam spaces $\{W(L^p, \ell^q)\}$ may be represented by the points (p, q) of the *same* unit square J as in the example of the L^p spaces, with the *same* order structure. However, $\mathcal{Q}$ is *not* a lattice with respect to the order (3.12). One has indeed

$$\begin{aligned} W(L^p, \ell^q) \wedge W\left(L^{p'}, \ell^{q'}\right) &\supset W\left(L^{p \vee p'}, \ell^{q \wedge q'}\right), \\ W(L^p, \ell^q) \vee W\left(L^{p'}, \ell^{q'}\right) &\subset W\left(L^{p \wedge p'}, \ell^{q \vee q'}\right), \end{aligned} \tag{6.14}$$

where again $\wedge$ means intersection with projective norm and $\vee$ means vector sum with inductive norm, but equality is not obtained. Thus, as in the previous case, one gets chains by varying either p or q, but not both.

A very useful class of amalgam spaces is the family $W(\mathcal{F}L^p, \ell^q), 1 \leqslant p, q \leqslant \infty$, where $\mathcal{F}L^p$ denotes the set of Fourier transforms of L^p functions (one may even add weights on both spaces). These spaces have, for instance, nice inclusion and convolution properties.

Among these, the most interesting one is $\mathcal{S}_0 = W(\mathcal{F}L^1, \ell^1)$, called the *Feichtinger algebra*. The space $\mathcal{S}_0$ has many interesting properties; for instance, one has the following.

(i) $\mathcal{S}_0$ is a Banach space for the norm $\|f\|_{\mathcal{S}_0} = \|V_{g_0} f\|_1$, and $\mathcal{S} \hookrightarrow \mathcal{S}_0 \hookrightarrow L^2$, with all embeddings continuous with dense range. Here g_0 is the Gaussian and $V_g f$ denotes the Short-Time Fourier (or Gabor) Transform of $f \in L^2(\mathbb{R}^d)$, given in (6.17) below.

(ii) $\mathcal{S}_0$ is a Banach algebra with respect to pointwise multiplication and convolution.

(iii) Time-frequency shifts $T_x M_\omega$ are isometric on $\mathcal{S}_0$: $\|T_x M_\omega f\|_{\mathcal{S}_0} = \|f\|_{\mathcal{S}_0}$, where $(T_x g)(y) = g(y - x)$ (translation) and $(M_\omega h)(y) = e^{2\pi i y \omega} h(y)$ (modulation). $\mathcal{S}_0$ is continuously embedded in any Banach space with the same property and containing g_0; thus it is the smallest Banach space with this property.

(iv) The Fourier transform is an isometry on $\mathcal{S}_0$: $\|\mathcal{F} f\|_{\mathcal{S}_0} = \|f\|_{\mathcal{S}_0}$.

As for the (conjugate) dual $\mathcal{S}_0^\times$ of $\mathcal{S}_0$, it is a Banach space with norm $\|f\|_{\mathcal{S}_0^\times} = \|V_g f\|_\infty$. The space $\mathcal{S}_0^\times$ contains both the δ function and the pure frequency $\chi_\omega(x) = e^{-2\pi i x \omega}$.

In virtue of (i) above, we have

$$\mathcal{S} \hookrightarrow \mathcal{S}_0 \hookrightarrow L^2 \hookrightarrow \mathcal{S}_0^\times \hookrightarrow \mathcal{S}^\times, \tag{6.15}$$

where all embeddings are continuous and have dense range. It turns out that the triplet

$$\mathcal{S}_0(\mathbb{R}^d) \hookrightarrow L^2(\mathbb{R}^d) \hookrightarrow \mathcal{S}_0^\times(\mathbb{R}^d) \tag{6.16}$$

is the prototype of a *Banach Gel'fand triple*, that is, a RHS (or LBS) in which the extreme spaces are (nonreflexive) Banach spaces. This is often a very convenient substitute for Schwartz' RHS and it is widely used in signal processing.

6.2.2. Modulation Spaces and Gabor Analysis (Time-Frequency Analysis)

Modulation spaces are closely linked to, and in fact defined in terms of, the Short-Time Fourier (or Gabor) Transform. Given a C^∞ window function $g \neq 0$, the *Short-Time Fourier Transform* (STFT) of $f \in L^2(\mathbb{R}^d)$ is defined by

$$(V_g f)(x,\omega) = \langle M_\omega T_x g \mid f \rangle := \int_{\mathbb{R}^d} \overline{g(y-x)} f(y) e^{-2\pi i y\omega} dy, \quad x,\omega \in \mathbb{R}^d, \tag{6.17}$$

where, as usual, $(T_x g)(y) = g(y-x)$ (translation) and $(M_\omega h)(y) = e^{2\pi i y\omega} h(y)$ (modulation).

Then, given a v-moderate weight function $m(x,\omega)$, (see Section 3.3.3) the modulation space $M_m^{p,q}$ is defined in terms of a mixed norm of an STFT:

$$M_m^{p,q}(\mathbb{R}^d) = \{f \in \mathcal{S}^\times(\mathbb{R}^d) : V_g f \in L_m^{p,q}(\mathbb{R}^{2d})\}, \quad 1 \leqslant p,q \leqslant \infty. \tag{6.18}$$

For $p = q$, one writes $M_m^p \equiv M_m^{p,p}$. The space $M_m^{p,q}$ is a Banach space for the norm

$$\|f\|_{M_m^{p,q}} := \|V_g f\|_{L_m^{p,q}}. \tag{6.19}$$

Actually, there is a slightly more restrictive definition, which uses the weight function $m_s(x,\omega) \equiv w_s(\omega) = (1+|\omega|)^s, s \geqslant 0$, (or, equivalently, $\widetilde{m}_s(x,\omega) = (1+|\omega|^2)^{s/2}$), so that the norm reads

$$\|f\|_{M_{w_s}^{p,q}} = \left(\int_{\mathbb{R}^d} \left(\int_{\mathbb{R}^d} |\langle M_\omega T_x g \mid f \rangle|^p dx\right)^{q/p} (1+|\omega|)^{sq}\, d\omega\right)^{1/q}. \tag{6.20}$$

Equivalently, one may define a modulation space as the inverse Fourier transform of a Wiener amalgam space:

$$M_{w_s}^{p,q} = \mathcal{F}^{-1}\left(W\left(L^p, \ell_{w_s}^q\right)\right). \tag{6.21}$$

This space is independent of the choice of window g, in the sense that different window functions define equivalent norms.

The lattice properties of the family $\{M_m^{p,q}, 1 \leqslant p,q \leqslant \infty\}$ are, of course, the same as those of the mixed-norm spaces $L_m^{p,q}$. As for duality, one has $(M_m^{p,q})^\times = M_{1/m}^{\overline{p},\overline{q}}$. Inclusion relations hold, leading again to a lattice structure: if $p_1 \leqslant p_2$, $q_1 \leqslant q_2$, and $m_2 \leqslant Cm_1$, for some constant $C > 0$, then $M^{p_1,q_1}_{m_1} \subseteq M^{p_2,q_2}_{m_2}$. In particular, one has

$$M^1_v \subseteq M^{p,q}_m \subseteq M^{\infty}_{1/v}. \tag{6.22}$$

The class of modulation spaces $M^{p,q}_{w_s}$ contains several well-known spaces, such as the following:

(i) The Bessel potential spaces or fractional Sobolev spaces $H^s = M^2_{\widetilde{m}_s}$:

$$H^s(\mathbb{R}^d) = M^2_{\widetilde{m}_s}(\mathbb{R}^d) = \left\{ f \in \mathcal{S}^{\times} : \int_{\mathbb{R}^d} \left|\widehat{f}(t)\right|^2 \left(1 + |t|^2\right)^s dt < \infty \right\}, \quad s \in \mathbb{R}. \tag{6.23}$$

(ii) $L^2(\mathbb{R}^d) = M^2(\mathbb{R}^d)$.

(iii) The Feichtinger algebra $\mathcal{S}_0 = M^1$ and its dual $\mathcal{S}^{\times}_0 = M^{\infty}$.

By construction, modulation spaces are function spaces well adapted to *Gabor analysis*, although they can often be replaced by amalgam spaces. A wealth of information about the spaces and their application in Gabor analysis may be found in the monograph of Gröchenig [20]. Here we just indicate a few relevant points, especially those that are of a PIP-space nature. We consider in particular the action of several types of operators on such spaces.

(i) *Translation and Modulation Operators*

(a) Every amalgam space $W(L^p, \ell^q)$ and every mixed-norm space $L^{p,q}_m$ are invariant under translation, that is, T_y is a totally regular operator in the corresponding PIP-space.

(b) Every modulation space $M^{p,q}_m$ is invariant under time-frequency shifts (translation and modulation), that is, T_y and M_{ξ} are totally regular operators.

(ii) *Fourier Transform*

(a) For $1 \leqslant p, q \leqslant 2, \mathcal{F}$ maps $W(L^p, \ell^q)$ into $(WL^{\overline{q}}, \ell^{\overline{p}})$ continuously, that is, $J(\mathcal{F})$ contains every pair (p, q), $(\overline{q}, \overline{p})$.

(b) If $m(\xi, -x) \leqslant Cm(x, \xi)$, then every space M^p_m is invariant under Fourier transform.

(iii) *Gabor Frame Operators*

Given a nonzero window function $g \in L^2(\mathbb{R}^d)$ and lattice parameters $\alpha, \beta > 0$, the set of vectors

$$\mathcal{G}(g, \alpha, \beta) = \left\{ M_{n\beta} T_{k\alpha} g,\ k, n \in \mathbb{Z}^d \right\} \tag{6.24}$$

is called a *Gabor system*. The system $\mathcal{G}(g, \alpha, \beta)$ is a *Gabor frame* if there exist two constants $\mathtt{m} > 0$ and $\mathtt{M} < \infty$ such that

$$\mathrm{m}\|f\|^2 \leqslant \sum_{k,n\in\mathbb{Z}^d} \left|\langle \mathrm{M}_{n\beta} T_{k\alpha} g \mid f \rangle\right|^2 \leqslant \mathrm{M}\|f\|^2, \quad \forall f \in L^2(\mathbb{R}^d). \tag{6.25}$$

The associated Gabor frame operator $S_{g,g}$ is given by

$$S_{g,g} f := \sum_{k,n \in \mathbb{Z}^d} \langle M_{n\beta} T_{k\alpha} g \mid f \rangle M_{n\beta} T_{k\alpha} g. \tag{6.26}$$

The main results of the Gabor time-frequency analysis stem from the following proposition.

Proposition 6.1. *If $\mathcal{G}(g, \alpha, \beta)$ is a Gabor frame, there exists a dual window $\breve{g} = S_{gg}^{-1} g$ such that $\mathcal{G}(\breve{g}, \alpha, \beta)$ is a frame, called the dual frame. Then one has, for every $f \in L^2(\mathbb{R}^d)$,*

$$f = \sum_{k,n \in \mathbb{Z}^d} \langle M_{n\beta} T_{k\alpha} g \mid f \rangle M_{n\beta} T_{k\alpha} \breve{g} = \sum_{k,n \in \mathbb{Z}^d} \langle M_{n\beta} T_{k\alpha} \breve{g} \mid f \rangle M_{n\beta} T_{k\alpha} g, \tag{6.27}$$

with unconditional convergence in $L^2(\mathbb{R}^d)$.

The outcome of the theory is that the modulation spaces $M_m^{p,q}$ turn out to be the natural class of function spaces for Gabor analysis [20]. Define indeed the following operator, generalizing (6.26) slightly:

$$S_{g,g'} f := \sum_{k,n \in \mathbb{Z}^d} \langle M_{n\beta} T_{k\alpha} g \mid f \rangle M_{n\beta} T_{k\alpha} g', \quad g, g' \in L^2(\mathbb{R}^{2d}). \tag{6.28}$$

Then one has the following results (they are highly nontrivial and their proof requires deep analysis)

(i) If $g, g' \in W(L^\infty, \ell^1)$, then the Gabor frame operator $S_{g,g'}$ is bounded on every $L^p(\mathbb{R}^{2d})$, $1 \leqslant p \leqslant \infty$.

(ii) If $g, g' \in M_v^1$, then $S_{g,g'}$ is bounded on $M_m^{p,q}$ for all $1 \leqslant p, q \leqslant \infty$, all v-moderate weights m, and all α, β.

(iii) If $\breve{g}$ is a dual window of g, that is, $S_{g,\breve{g}} = 1$ on L^2, then the two expansions in (6.27) converge unconditionally in $M_m^{p,q}$ if $p, q < \infty$.

Clearly statements (i) and (ii) can be translated into PIP-space language, by saying that $S_{g,g'}$ is a totally regular operator in the chain $\{L^p, 1 \leqslant p \leqslant \infty\}$, respectively, any PIP-space built from modulation spaces.

These results should suffice to convince the reader that the modulation spaces $M_m^{p,q}$ are the "natural" spaces for Gabor analysis. Actually, most of this remains true if one replaces modulation spaces by amalgam spaces $W(L^p, \ell_m^q)$. Second, it is obvious that most of the statements have a distinctly PIP-space flavor: it is not some individual space $M_m^{p,q}$ or $W(L^p, \ell_m^q)$ that counts, but the whole family, with many operators being regular in the sense of PIP-spaces.

6.2.3. Besov Spaces and Wavelet Analysis (Time-Scale Analysis)

Besov spaces were introduced around 1960 for providing a precise control on the smoothness of solutions of certain partial differential equations. Later on, it was discovered that they are closely linked to wavelet analysis, exactly as the (much more recent) modulation spaces are structurally adapted to Gabor analysis. In fact, there are many equivalent definitions of Besov spaces. We restrict ourselves to a “discrete” formulation, based on a dyadic partition of unity. Other definitions may be found in the literature quoted in [14], in particular [19, 38–40].

Let us consider a weight function $\varphi \in \mathcal{S}(\mathbb{R})$ with the following properties:

(i) $\operatorname{supp} \varphi = \{\xi : 2^{-1} \leqslant |\xi| \leqslant 2\}$,

(ii) $\varphi(\xi) > 0$ for $2^{-1} < |\xi| < 2$,

(iii) $\sum_{j=-\infty}^{\infty} \varphi(2^{-j}\xi) = 1 \ (\xi \neq 0)$.

Then one defines the following functions by their Fourier transform:

(i) $\widehat{\varphi_j}(\xi) = \varphi(2^{-j}\xi), \ j \in \mathbb{Z}$: high “frequency” for $j > 0$, low “frequency” for $j < 0$,

(ii) $\widehat{\psi}(\xi) = 1 - \sum_{j=1}^{\infty} \varphi(2^{-j}\xi)$: low “frequency” part.

Given the weight function φ, the inhomogeneous Besov space B_{pq}^s is defined as

$$B_{pq}^s = \left\{ f \in \mathcal{S}^\times : \|f\|_{pq}^s < \infty \right\}, \tag{6.29}$$

where $\|\cdot\|_{pq}^s$ denotes the norm

$$\|f\|_{pq}^s := \|\psi * f\|_p + \left(\sum_{j=1}^{\infty} \left(2^{sj} \|\varphi_j * f\|_p \right)^q \right)^{1/q}, \quad s \in \mathbb{R}, \ 1 \leqslant p, q \leqslant \infty. \tag{6.30}$$

The space B_{pq}^s is a Banach space and it does not depend on the choice of the weight function φ, since a different choice defines an equivalent norm. Note that $B_{22}^s = H^s$, the (fractional) Sobolev space or Bessel potential space.

For $f \in B_{pq}^s$, one may write the following (weakly converging) expansion, known as a *dyadic Littlewood-Paley decomposition*:

$$f = \psi * f + \sum_{j=1}^{\infty} \varphi_j * f. \tag{6.31}$$

Clearly the first term represents the (relatively uninteresting) low-“frequency” part of the function, whereas the second term analyzes in detail the high-“frequency” component.

Besov spaces enjoy many familiar properties (for more details, we refer to the literature, e.g., [19, Section 6.2] or [39, Chapter 2]).

(i) Inclusion Relations

The following relations hold, where all embeddings are continuous:

(a) $\mathcal{S} \hookrightarrow B_{pq}^s \hookrightarrow \mathcal{S}^\times$,

(b) $B_{pq}^s \hookrightarrow L^p$, if $1 \leqslant p, q \leqslant \infty$ and $s > 0$,

(c) for $s_1 < s_2$, $B^{s_2}_{pq} \hookrightarrow B^{s_1}_{pq}$ $(1 \leqslant q, p \leqslant \infty)$,

(d) for $1 \leqslant q_1 < q_2 \leqslant \infty$, $B^{s}_{pq_1} \hookrightarrow B^{s}_{pq_2}$ $(s \in \mathbb{R}, 1 \leqslant p \leqslant \infty)$,

(e) for $s - 1/p = s_1 - 1/p_1, B^{s}_{pq} \hookrightarrow B^{s_1}_{p_1 q_1}$ $(s, s_1 \in \mathbb{R},\ 1 \leqslant p \leqslant p_1 \leqslant \infty,\ 1 \leqslant q \leqslant q_1 \leqslant \infty)$.

In the terminology of Section 4, the first statement means that the spaces B^{s}_{pq} are interspaces for the RHS $\mathcal{S} \hookrightarrow L^2 \hookrightarrow \mathcal{S}^\times$. The inclusion relations above mean that the family of spaces B^{s}_{pq} contains again many chains of Banach spaces.

(ii) Interpolation

Besov spaces enjoy nice interpolation properties, in all three parameters $s,\ p,\ q$.

(iii) Duality

One has $(B^{s}_{pq})^\times = B^{-s}_{\overline{p}\,\overline{q}}$ $(s \in \mathbb{R})$.

(iv) Translation and Dilation Invariance

Every space B^{s}_{pq} is invariant under translation and dilation (the unitary dilation operator in L^2 reads $(D_a f)(x) = a^{-1/2} f(x/a)$.)

(v) Regularity Shift

Let $J^\sigma : \mathcal{S}^\times \to \mathcal{S}^\times$ denote the operator $J^s f = \mathcal{F}^{-1}\{(1 + |\cdot|^2)^{s/2} \mathcal{F} f\}$, $s \in \mathbb{R}$. Then J^σ is an isomorphism from B^{s}_{pq} onto $B^{s-\sigma}_{pq}$. Thus J^σ is totally regular for $\sigma \leqslant 0$, but not for $\sigma > 0$.

It is also useful to consider the homogeneous Besov space $\dot{B}^{s}_{pq}$, defined as the set of all $f \in \mathcal{S}^\times$ for which $\|f\dot{\|}^{s}_{pq} < \infty$, where the quasinorm $\|\cdot\dot{\|}^{s}_{pq}$ is defined by

$$\|f\dot{\|}^{s}_{pq} := \left(\sum_{j=-\infty}^{\infty} \left(2^{sj} \|\varphi_j * f\|_p \right)^q \right)^{1/q}. \tag{6.32}$$

(This is only a quasi-norm since $\|f\dot{\|}^{s}_{pq} = 0$ if and only if supp $\widehat{f} = \{0\}$, i.e., f is a polynomial.) Note that, if $0 \notin$ supp $\widehat{f}$, then $f \in \dot{B}^{s}_{pq}$ if and only if $f \in B^{s}_{pq}$.

The spaces $\dot{B}^{s}_{pq}$ have properties similar to the previous ones and, in addition, one has $B^{s}_{pq} = L^p \cap \dot{B}^{s}_{pq}$ for $s > 0$, $1 \leqslant p, q \leqslant \infty$. In particular, every space $\dot{B}^{s}_{pq}$ is invariant under translation and dilation, which is not surprising, since these spaces are in fact based on the $ax + b$ group, consisting precisely of dilations and translations of the real line, via the coorbit space construction (see Section 6.2.4 below).

Besov spaces are well adapted to *wavelet analysis*, because the definition (6.29) essentially relies on a dyadic partition (powers of 2). Historically, the connection was made with the *discrete* wavelet analysis, for that reason. Indeed, there exists an equivalent definition given in terms of decay of wavelet coefficients. More precisely, if a function f is expanded in a wavelet basis, the decay properties of the wavelet coefficients allow to characterize precisely to which Besov space the function f belongs, as we shall see below. In addition, the Besov spaces may also be characterized in terms of the *continuous* wavelet transform (see [14, Section 8.4]).

In order to go into details, we have to recall some basic facts about the wavelet transform (for simplicity, we restrict ourselves to one dimension). Whereas the STFT is defined in terms of translation and modulation, the continuous wavelet transform is based on translations and dilations

$$(W_\psi s)(b,a) = a^{-1}\int_{-\infty}^{\infty} \overline{\psi\left(a^{-1}(x-b)\right)}s(x)dx, \quad a>0,\ b\in\mathbb{R},\ s\in L^2(\mathbb{R}). \tag{6.33}$$

Note that we use here the so-called L^1 normalization. It is more frequent to use the L^2 normalization, in which the prefactor is $a^{-1/2}$ instead of a^{-1}. In this relation, the wavelet ψ is assumed to satisfy the admissibility condition

$$c_\psi := \int_{-\infty}^{\infty} d\omega|\omega|^{-1}\left|\widehat{\psi}(\omega)\right|^2 < \infty, \tag{6.34}$$

which implies that $\int_{-\infty}^{\infty}\psi(x)dx = 0$. This condition is only necessary, but becomes sufficient under some mild restrictions, so that it is commonly used as admissibility condition in practice.

However, discretizing the two parameters a and b in (6.33) leads in general only to frames. In order to get orthogonal wavelet bases, one relies on the so-called *multiresolution analysis* of $L^2(\mathbb{R})$. This is defined as an increasing sequence of closed subspaces of $L^2(\mathbb{R})$:

$$\cdots \subset \mathcal{V}_{-2} \subset \mathcal{V}_{-1} \subset \mathcal{V}_0 \subset \mathcal{V}_1 \subset \mathcal{V}_2 \subset \cdots \tag{6.35}$$

with $\bigcap_{j\in\mathbb{Z}}\mathcal{V}_j = \{0\}$ and $\bigcup_{j\in\mathbb{Z}}\mathcal{V}_j$ dense in $L^2(\mathbb{R})$ and such that

(1) $f(x)\in\mathcal{V}_j \Leftrightarrow f(2x)\in\mathcal{V}_{j+1}$,

(2) there exists a function $\phi\in\mathcal{V}_0$, called a *scaling function*, such that the family $\{\phi(\cdot-k),\ k\in\mathbb{Z}\}$ is an orthonormal basis of $\mathcal{V}_0$.

Combining conditions (1) and (2), one sees that $\{\phi_{jk}\equiv 2^{j/2}\phi(2^j\cdot-k),\ k\in\mathbb{Z}\}$ is an orthonormal basis of $\mathcal{V}_j$. The space $\mathcal{V}_j$ can be interpreted as an *approximation* space at resolution 2^j. Defining $\mathcal{W}_j$ as the orthogonal complement of $\mathcal{V}_j$ in $\mathcal{V}_{j+1}$, that is, $\mathcal{V}_j\oplus\mathcal{W}_j = \mathcal{V}_{j+1}$, we see that $\mathcal{W}_j$ contains the additional *details* needed to improve the resolution from 2^j to 2^{j+1}. Thus one gets the decomposition $L^2(\mathbb{R}) = \oplus_{j\in\mathbb{Z}}\mathcal{W}_j$. The crucial theorem then asserts the existence of a function ψ, called the *mother wavelet*, explicitly computable from ϕ, such that $\{\psi_{jk}\equiv 2^{j/2}\psi(2^j\cdot-k),\ k\in\mathbb{Z}\}$ constitutes an orthonormal basis of $\mathcal{W}_j$ and thus $\{\psi_{jk}\equiv 2^{j/2}\psi(2^j\cdot -k),\ j,k\in\mathbb{Z}\}$ is an orthonormal basis of $L^2(\mathbb{R})$: these are the *orthonormal wavelets*. Thus the expansion of an arbitrary function $f\in L^2$ into an orthogonal wavelet basis $\{\psi_{jk},\ j,k\in\mathbb{Z}\}$ reads

$$f = \sum_{j,k\in\mathbb{Z}} c_{jk}\psi_{jk}, \quad \text{with}\, c_{jk} = \langle\psi_{jk}\mid f\rangle. \tag{6.36}$$

Additional regularity conditions can be imposed to the scaling function ϕ. Given $r \in \mathbb{N}$, the multiresolution analysis corresponding to ϕ is called *r-regular* if

$$\left|\frac{d^n \phi}{dx^n}\right| \leqslant c_m \left(1+|x|^m\right), \quad \text{for all } n \leqslant r \text{ and all integers } m \in \mathbb{N}. \tag{6.37}$$

Well-known examples include the Haar wavelets, the B-splines, and the various Daubechies wavelets.

As a result of the "dyadic" definition (6.29)-(6.30), it is natural that Besov spaces can be characterized in terms of an r-regular multiresolution analysis $\{\mathcal{V}_j\}$. Let $E_j : L^2 \to \mathcal{V}_j$ be the orthogonal projection on $\mathcal{V}_j$ and $D_j = E_{j+1} - E_j$ that on $\mathcal{W}_j$. Let $0 < s < r$ and $f \in L^p(\mathbb{R})$. Then, $f \in B^s_{pq}(\mathbb{R})$ if and only if $\|D_j f\|_p = 2^{-js}\delta_j$, where $(\delta_j) \in \ell^q(\mathbb{N})$, and one has ($\asymp$ means equivalence of norms)

$$\|f\|^s_{pq} \asymp \|E_0 f\|_p + \left(\sum_{j\in\mathbb{Z}} 2^{jsq} \|D_j f\|^q_p\right)^{1/q}. \tag{6.38}$$

Specializing to $p = q = 2$, one gets a similar result for Sobolev spaces: given $f \in H^{-r}(\mathbb{R})$ and $|s| < r$, $f \in H^s(\mathbb{R})$ if, and only if, $E_0 f \in L^2(\mathbb{R})$ and $\|D_j f\|_2 = 2^{-js}\epsilon_j$, $j \in \mathbb{N}$, where $(\epsilon_j) \in \ell^2(\mathbb{N})$.

But there is more. Indeed, modulation spaces and Besov spaces admit decomposition of elements into wavelet bases and each space can be uniquely characterized by the decay properties of the wavelet coefficients. To be precise, let $\{\psi_{jk},\ j,k \in \mathbb{Z}\}$ be an orthogonal wavelet basis coming from an r-regular multiresolution analysis based on the scaling function ϕ. Then the following results are typical [41, Chapters II.9 and VI.10].

(i) *Inhomogeneous Besov Spaces:* $f \in B^s_{pq}(\mathbb{R})$ if it can be written as

$$f(x) = \sum_{k\in\mathbb{Z}} \beta_k \phi(x-k) + \sum_{j\geqslant 0, k\in\mathbb{Z}} c_{jk}\psi_{jk}, \tag{6.39}$$

where $(\beta_k) \in \ell^p$ and $\left(\sum_{k\in\mathbb{Z}} |c_{jk}|^p\right)^{1/p} = 2^{-j(s+1/2-1/p)}\gamma_j$, with $(\gamma_j) \in \ell^q(\mathbb{Z})$.

(ii) *Homogeneous Besov Spaces:* Let $|s| < r$. Then, if $f \in \dot{B}^s_{pq}(\mathbb{R})$, its wavelet coefficients c_{jk} verify $\left(\sum_{k\in\mathbb{Z}} |c_{jk}|^p\right)^{1/p} = 2^{-j(s+1/2-1/p)}\ \gamma_j$, where $(\gamma_j) \in \ell^q(\mathbb{Z})$. Conversely, if this condition is satisfied, then $f = g + P$, where $g \in \dot{B}^s_{pq}$ and P is a polynomial.

6.2.4. α-Modulation Spaces, Coorbit Spaces

The α-modulation spaces ($\alpha \in [0,1]$) are spaces intermediate between modulation and Besov spaces, to which they reduce for $\alpha = 0$ and $\alpha \to 1$, respectively. As for coorbit spaces, they are a far-reaching generalization, based on integrable group representations [42]. They contain most of the previous spaces, but we will refrain from describing them in detail, for lack of space.

6.2.5. Unconditional Bases

We conclude this section with some examples of unconditional wavelet bases, as announced in Section 5.2.5. Actually, the concept of wavelet basis can be further generalized to *biorthogonal bases*, obtained by considering two scales of the type (6.35) and imposing cross-orthogonality relations [43, Section 8.3]. For precise definitions, we refer to the literature.

(i) The Haar wavelet basis is defined by the scaling function $\phi_H = \chi_{[0,1]}$ and the mother wavelet $\psi_H = \chi_{[0,1/2]} - \chi_{[1/2,1]}$. It is a standard result that the Haar system is an unconditional basis for every $L^p(\mathbb{R})$, $1 < p < \infty$.

(ii) The Lemarié-Meyer wavelet basis is an unconditional basis for all L^p spaces, Sobolev spaces, and homogeneous Besov spaces $\dot{B}^s_{pq}$ $(1 \leqslant p, q < \infty)$ [44].

(iii) There is a class of wavelet bases (Wilson bases of exponential decay) that are unconditional bases for every modulation space $M^{p,q}_m$, $1 \leqslant p, q < \infty$, but *not* for L^p, $1 < p < \infty$, $p \neq 2$ [45].

7. Conclusion

Most families of function spaces used in analysis and in signal processing come in scales or lattices and in fact are, or contain, PIP-spaces. The (lattice) indices defining the (partial) order characterize the properties of the corresponding functions or distributions: smoothness, local integrability, decay at infinity. Thus it seems natural to formulate the properties of various operators globally, using the theory of PIP-space operators; in particular the set $\mathsf{j}(A)$ of an operator encodes its properties in a very convenient and visual fashion. In addition, it is often possible to determine uniquely whether a function belongs to one of those spaces simply by estimating the (asymptotic) behavior of its Gabor or wavelet coefficients, a real breakthrough in functional analysis [41].

A legitimate question is whether there are instances where a PIP-space is really *needed*, or a RHS could suffice. The answer is that there are plenty of examples, among the applications we have enumerated in Section 6 (without details, for lack of space). We may therefore expect that the PIP-space formalism will play a significant role in Gabor/wavelet analysis, as well as in mathematical physics.

Concerning the applications in mathematical physics, in almost all cases, the relevant structure is a scale or a chain of Hilbert spaces, which allows a finer control on the behavior of operators. For instance (all details may be found in our monograph [14]) the following are considered.

(i) For the singular interactions in quantum mechanics (δ or δ' potentials), the approach of Grossmann et al. [18, 24, 25] is definitely the most appropriate; a RHS would be irrelevant.

(ii) The very formulation of the WVW approach to quantum scattering theory [34–36] requires a LHS, whose end spaces are in fact Hilbert spaces.

(iii) For quantum field theory, the energy bounds of Fredenhagen and Hertel [46] rely in an essential way on the scale generated by the Hamiltonian, and so does Nelson's approach to Euclidean field theory [47].

As for the applications in signal processing, all families of spaces routinely used are, or contain, chains of Banach spaces, which are needed for a fine tuning of elements (usually, distributions) and operators on them. Such are, for instance, L^p spaces, amalgam spaces, modulation spaces, Besov spaces, or coorbit spaces, mentioned above. Here again, a RHS, even with Banach end spaces like the Feichtinger algebra and its conjugate dual, is clearly not sufficient. The whole Section 6.2 illustrates the statement.

References

[1] L. Hörmander, *Linear Partial Differential Operators*, Die Grundlehren der Mathematischen Wissenschaften, Bd. 116, Springer, Berlin, Germany, 1963.
[2] J.-L. Lions and E. Magenes, *Problèmes aux Limites non Homogènes et Applications*, vol. 1, Dunod, Paris, France, 1968.
[3] J.-L. Lions and E. Magenes, *Problèmes aux Limites non Homogènes et Applications*, vol. 2, Dunod, Paris, France, 1968.
[4] J.-L. Lions and E. Magenes, *Problèmes aux Limites non Homogènes et Applications*, vol. 3, Dunod, Paris, France, 1970.
[5] I. M. Gel'fand and N. Ya. Vilenkin, *Generalized Functions IV: Applications of Harmonic Analysis*, Academic Press, New York, NY, USA, 1964.
[6] K. Maurin, *General Eigenfunction Expansions and Unitary Representations of Topological Groups*, Monografie Matematyczne, Tom 48, PWN-Polish Scientific, Warsaw, Poland, 1968.
[7] L. Schwartz, *Théorie des Distributions. I*, Hermann, Paris, France, 1957.
[8] L. Schwartz, *Théorie des Distributions. II*, Hermann, Paris, France, 1959.
[9] B. Simon, "Distributions and their Hermite expansions," *Journal of Mathematical Physics*, vol. 12, pp. 140–148, 1971.
[10] J.-P. Antoine and A. Grossmann, "Partial inner product spaces. I. General properties," *Journal of Functional Analalysis*, vol. 23, no. 4, pp. 369–378, 1976.
[11] J.-P. Antoine and A. Grossmann, "Partial inner product spaces. II. Operators," *Journal of Functional Analysis*, vol. 23, no. 4, pp. 379–391, 1976.
[12] J.-P. Antoine, "Partial inner product spaces. III. Compatibility relations revisited," *Journal of Mathematical Physics*, vol. 21, no. 2, pp. 268–279, 1980.
[13] J.-P. Antoine, "Partial inner product spaces. IV. Topological considerations," *Journal of Mathematical Physics*, vol. 21, no. 8, pp. 2067–2079, 1980.
[14] J.-P. Antoine and C. Trapani, *Partial Inner Product Spaces: Theory and Applications*, vol. 1986 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 2010.
[15] G. Köthe, *Topological Vector Spaces. I*, Die Grundlehren der Mathematischen Wissenschaften, Bd. 159, Berlin, Germany, 1966.
[16] J. Dieudonné, "Sur les espaces de Köthe," *Journal d'Analyse Mathématique*, vol. 1, pp. 81–115, 1951.
[17] S. Goes and R. Welland, "Some remarks on Köthe spaces," *Mathematische Annalen*, vol. 175, pp. 127–131, 1968.
[18] A. Grossmann, "Elementary properties of nested Hilbert spaces," *Communications in Mathematical Physics*, vol. 2, pp. 1–30, 1966.
[19] J. Bergh and J. Löfström, *Interpolation Spaces: An Introduction*, Grundlehren der Mathematischen Wissenschaften, no. 22, Springer, Berlin, Germany, 1976.
[20] K. Gröchenig, *Foundations of Time-Frequency Analysis*, Applied and Numerical Harmonic Analysis, Birkhäuser, Boston, Mass, USA, 2001.
[21] A. C. Zaanen, *Integration*, chapter 15, North-Holland, Amsterdam, The Netherlands, 2nd edition, 1961.
[22] J.-P. Antoine, A. Inoue, and C. Trapani, *Partial *-Algebras and Their Operator Realizations*, vol. 553 of *Mathematics and Its Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2002.
[23] J.-P. Antoine and F. Mathot, "Regular operators on partial inner product spaces," *Annales de l'Institut Henri Poincaré. Section A*, vol. 37, no. 1, pp. 29–50, 1982.
[24] A. Grossmann, R. Høegh-Krohn, and M. Mebkhout, "A class of explicitly soluble, local, many-center Hamiltonians for one-particle quantum mechanics in two and three dimensions. I," *Journal of Mathematical Physics*, vol. 21, no. 9, pp. 2376–2386, 1980.
[25] A. Grossmann, R. Høegh-Krohn, and M. Mebkhout, "The one particle theory of periodic point Interactions-Polymers, monomolecular layers, and crystals," *Communications in Mathematical Physics*, vol. 77, no. 1, pp. 87–110, 1980.

[26] J.-P. Antoine and A. Grossmann, "Orthocomplemented subspaces of nondenenerate partial inner product spaces," *Journal of Mathematical Physics*, vol. 19, no. 1, pp. 329–335, 1978.
[27] J.-P. Antoine, "Dirac formalism and symmetry problems in quantum mechanics. I. General Dirac formalism," *Journal of Mathematical Physics*, vol. 10, pp. 53–69, 1969.
[28] J.-P. Antoine, "Dirac formalism and symmetry problems in quantum mechanics. II. Symmetry problems," *Journal of Mathematical Physics*, vol. 10, no. 12, pp. 2276–2290, 1969.
[29] A. Böhm, "The Rigged Hilbert space and mathematical description of physical systems," in *Mathematical Methods of Theoretical Physics*, W. A. Brittin, et al., Ed., Lectures in Theoretical Physics, vol. IX A, pp. 255–317, Gordon and Breach, New York, NY, USA, 1967.
[30] J. E. Roberts, "The Dirac bra and ket formalism," *Journal of Mathematical Physics*, vol. 7, pp. 1097–1104, 1966.
[31] J. E. Roberts, "Rigged Hilbert spaces in quantum mechanics," *Communications in Mathematical Physics*, vol. 3, no. 2, pp. 98–119, 1966.
[32] J.-P. Antoine, "Symmetries of quantum systems: A partial inner product space approach," *Journal of Physics A: Mathematical and Theoretical*, vol. 41, no. 30, Article ID 304037, 13 pages, 2008.
[33] S. Albeverio, F. Gesztesy, R. Høegh-Krohn, and H. Holden, *Solvable Models in Quantum Mechanics*, Texts and Monographs in Physics, Springer, New York, NY, USA, 1988.
[34] C. van Winter, "Fredholm equations on a Hilbert space of analytic functions," *Transactions of the American Mathematical Society*, vol. 162, pp. 103–139, 1971.
[35] C. van Winter, "Complex dynamical variables for multiparticle systems with analytic interactions. I," *Journal of Mathematical Analysis and Applications*, vol. 47, no. 3, pp. 633–670, 1974.
[36] C. van Winter, "Complex dynamical variables for multiparticle systems with analytic interactions. II," *Journal of Mathematical Analysis and Applications*, vol. 48, no. 2, pp. 368–399, 1974.
[37] J. J. F. Fournier and J. Stewart, "Amalgams of L^p and ℓ^q," *Bulletin of the American Mathematical Society*, vol. 13, no. 1, pp. 1–21, 1985.
[38] H. Triebel, *Function Spaces I*, Birkhäuser, Basel, Switzerland, 1978.
[39] H. Triebel, *Function Spaces II*, Birkhäuser, Basel, Switzerland, 1992.
[40] H. Triebel, *Function Spaces III*, Birkhäuser, Basel, Switzerland, 2006.
[41] Y. Meyer, *Ondelettes et Opérateurs. I*, Actualités Mathématiques, Hermann, Paris, France, 1990.
[42] H. G. Feichtinger and K. H. Gröchenig, "Banach spaces related to integrable group representations and their atomic decompositions. I," *Journal of Functional Analysis*, vol. 86, no. 2, pp. 307–340, 1989.
[43] I. Daubechies, *Ten Lectures on Wavelets*, vol. 61 of *CBMS-NSF Regional Conference Series in Applied Mathematics*, SIAM, Philadelphia, PA, USA, 1992.
[44] K. Gröchenig, "Unconditional bases in translation and dilation invariant function spaces on $\mathbb{R}^n$," in *Constructive Theory of Functions (Varna, 1987)*, pp. 174–183, Bulgarian Academy of Sciences, Sofia, Bulgaria, 1988.
[45] H. G. Feichtinger, K. Gröchenig, and D. Walnut, "Wilson bases and modulation spaces," *Mathematische Nachrichten*, vol. 155, pp. 7–17, 1992.
[46] K. Fredenhagen and J. Hertel, "Local algebras of observables and pointlike localized fields," *Communications in Mathematical Physics*, vol. 80, no. 4, pp. 555–561, 1981.
[47] E. Nelson, "Time-ordered operator products of sharp-time quadratic forms," *Journal of Mathematical Physics*, vol. 5, pp. 1190–1197, 1964.

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 191529, 16 pages
doi:10.1155/2010/191529

Review Article

New Affine Coherent States Based on Elements of Nonrenormalizable Scalar Field Models

John R. Klauder[1, 2]

[1] *Department of Physics, University of Florida, Gainesville, FL 32611, USA*
[2] *Department of Mathematics, University of Florida, Gainesville, FL 32611, USA*

Correspondence should be addressed to John R. Klauder, john.klauder@gmail.com

Received 23 December 2009; Accepted 15 April 2010

Academic Editor: S. T. Ali

Recent proposals for a nontrivial quantization of covariant, nonrenormalizable, self-interacting, scalar quantum fields have emphasized the importance of quantum fields that obey affine commutation relations rather than canonical commutation relations. When formulated on a spacetime lattice, such models have a lattice version of the associated ground state, and this vector is used as the fiducial vector for the definition of the associated affine coherent states, thus ensuring that in the continuum limit, the affine field operators are compatible with the system Hamiltonian. In this article, we define and analyze the associated affine coherent states as well as briefly review the author's approach to nontrivial formulations of such nonrenormalizable models.

1. Introduction and Overview

The author has been interested for long in finding a meaningful and nontrivial quantization of nonrenormalizable quantum field theories [1]. Quantum gravity is the most important example of this kind, but it is kinematically simpler to try to understand nonrenormalizable, self-interacting scalar fields as test cases initially. The subject of this article continues that quest by proposing coherent states based on affine field algebras associated with a recent scheme to quantize scalar nonrenormalizable models. Suitable coherent states have the virtue of bridging the classical-quantum divide, and thus offer a useful tool in the study of such problems.

In Section 2 we introduce the concept of affine coherent states first starting with simple few degree of freedom systems. Next, we extend that discussion to affine fields in a 1-dimensional and later to an s-dimensional Euclidean space. To get a better handle on such representations, we study cutoff field examples by replacing space by a finite, periodic lattice of points with $L < \infty$ points on a side each separated by a lattice spacing $a > 0$. Each point on the lattice carries a set of affine fields that obey a discrete form of the affine field algebra and thus represent a finite but large number of independent, one-dimensional affine systems. In the continuum limit, of course, such systems become affine field algebras, but studying them before taking the continuum limit offers certain advantages. In addition, we choose very special vectors to serve as fiducial vectors for the associated affine coherent states. These fiducial vectors are chosen as ground states for suitable model field theories that in Section 3 are also formulated on spacetime lattices with similar properties. These quantum field models have been chosen to achieve a nontrivial quantization of quartic, self-interacting, covariant, nonrenormalizable models, specifically the so-called φ_n^4 models in spacetime dimensions $n \geq 5$. These models are sketched in Section 3 to an extent necessary so as to understand the choice of vectors selected as fiducial vectors for the discussion in Section 2. A brief commentary appears in Section 4 setting the present work in the context of efforts to study quantum gravity.

2. New Affine Coherent States

2.1. Single Degree of Freedom

For a single degree of freedom, the canonical, self adjoint, quantum variables are P and Q, which obey the canonical commutation relation $[Q, P] = i\hbar\mathbb{1}$. The usual irreducible representation of these operators involves spectra for both operators that runs from $-\infty$ to $+\infty$. The affine variables are formally obtained from the canonical ones by multiplying the canonical commutation relation by the operator Q, which leads to the relation

$$Q[Q,P] = i\hbar Q = [Q,(QP+PQ)/2] \equiv [Q,D], \tag{2.1}$$

an expression that introduces the affine quantum variables Q and D which satisfy the affine commutation relation $[Q, D] = i\hbar Q$. It is noteworthy that the affine commutation relation has three inequivalent, irreducible representations with self-adjoint affine variables: one for which $Q > 0$, one for which $Q < 0$, and one for which $Q = 0$. For the first two representations, D has a spectrum that covers the whole real line. For the irreducible representation for which $Q > 0$, a suitable set of coherent states may be given by

$$|p,q+\rangle \equiv e^{ipQ/\hbar}\, e^{-i\ln(q)D/\hbar}\, |\eta+\rangle, \tag{2.2}$$

where $p \in \mathbb{R} = (-\infty,\infty)$, $q \in \mathbb{R}^+ = (0,\infty)$, and the unit vector $|\eta+\rangle$ is called the fiducial vector. In this formulation q is dimensionless while p has the dimensions of $\hbar$. In the representation in which Q is diagonalized, that is, $Q|x\rangle = x|x\rangle$, $x > 0$, it follows that the function $\eta_+(x) = \langle x|\eta+\rangle$ is supported only on the positive real line. One possible choice of fiducial vector is given by

$$\eta_+(x) = N_+ \exp\left[-\left(B/x^2\right) - C(x-1)^2/\hbar\right], \quad x > 0, \tag{2.3}$$

where $B > 0$, $C > 0$, and N_+ normalizes the expression so that $\||\eta+\rangle\| = 1$. If instead we were interested in the second irreducible representation for which $Q < 0$, then a suitable set of coherent states would be given (for $q < 0$) by

$$|p,q-\rangle \equiv e^{ipQ/\hbar}\, e^{-i\ln(-q)D/\hbar}\, |\eta-\rangle. \tag{2.4}$$

In this case $Q|x\rangle = x|x\rangle$, $x < 0$, and we may choose

$$\eta_-(x) = N_- \exp\left[-\left(B/x^2\right) - C(x+1)^2/\hbar\right], \tag{2.5}$$

from which we conclude that $\||\eta-\rangle\| = 1$.

2.2. Reducible Representation

Finally, when we wish to include both irreducible representations, where $Q > 0$ and $Q < 0$, we may choose a set of coherent states—now for $p \in \mathbb{R}, q \in \mathbb{R} \setminus 0$—given by

$$|p, q\pm\rangle \equiv \theta(q)|p, q+\rangle + \theta(-q)|p, q-\rangle, \tag{2.6}$$

where $\theta(q) \equiv 1$ if $q > 0$ and $\theta(q) \equiv 0$ if $q < 0$. It follows that

$$\||p, q\pm\rangle\|^2 = \theta(q)\||p, q\pm\rangle\|^2 + \theta(-q)\||p, q-\rangle\|^2 = 1, \tag{2.7}$$

and thus the coherent states $|p, q\pm\rangle$ are all normalized vectors.

It is useful to review some algebraic properties of affine variables and some properties of affine coherent state matrix elements. In particular, if we introduce ($q \neq 0$)

$$A[p, q] \equiv e^{ipQ/\hbar} e^{-i\ln(|q|)D/\hbar}, \tag{2.8}$$

we have the relations

$$\begin{aligned}
A[p, q]^\dagger D A[p, q] &= D + p|q|Q, \\
A[p, q]^\dagger P A[p, q] &= |q|^{-1}P + p, \\
A[p, q]^\dagger Q A[p, q] &= |q|Q, \\
\langle p, q\pm|B(D, P, Q)|p, q\pm\rangle &= \theta(q)\langle\eta+|B\left(D + pqQ, q^{-1}P + p, qQ\right)|\eta+\rangle \\
&\quad + \theta(-q)\langle\eta-|B\left(D + p(-q)Q, (-q)^{-1}P + p, (-q)Q\right)|\eta-\rangle.
\end{aligned} \tag{2.9}$$

Introducing $\langle(\cdot)\rangle_\pm \equiv \langle\eta\pm|(\cdot)|\eta\pm\rangle$, it follows that

$$\langle p, q\pm|Q|p, q\pm\rangle = \theta(q)q\langle Q\rangle_+ + \theta(-q)(-q)\langle Q\rangle_-, \tag{2.10}$$

and we note that if we choose the parameters B and C such that $\langle Q\rangle_+ = 1$, it follows that $\langle Q\rangle_- = -1$, and we find that

$$\langle p, q\pm|Q|p, q\pm\rangle = \theta(q)q - \theta(-q)(-q) \equiv q, \tag{2.11}$$

a desirable value indeed. Additionally, for whatever values we assign to B and C it follows that

$$\langle p, q\pm | Q | p, q\pm \rangle = q + O(\hbar), \tag{2.12}$$

and, furthermore, it is not difficult to show that

$$\langle p, q\pm | Q^r | p, q\pm \rangle = q^r + O(\hbar) \tag{2.13}$$

for all r.

In addition,

$$\begin{aligned} \langle p, q\pm | D | p, q\pm \rangle &= \theta(q)\,[pq\langle Q\rangle_+ + \langle D\rangle_+] + \theta(-q)\,[p(-q)\langle Q\rangle_- + \langle D\rangle_-] \\ &= pq, \end{aligned} \tag{2.14}$$

where we have chosen the physically natural values that $\langle D\rangle_\pm = 0$, which follow directly when the parameters B and C are both real. These expectation values enable us to identify q as the mean value of Q and pq as the mean value of D in the coherent states. This identification introduces a natural connection between the quantum and classical variables, but in no sense is this particular connection required.

2.3. Alternative Affine Coherent States

The construction of $|p, q\pm\rangle$ entails two distinct parts proportional to $\theta(q)$ and $\theta(-q)$, respectively. If J degrees of freedom are defined this way, then 2^J disjoint parts are involved, which becomes very large when $J \gg 1$. To avoid this aspect, we next focus on an alternative construction.

In this new version, we choose to give up some features of the set of coherent states defined above and instead define

$$|p, q\rangle \equiv e^{ipQ/\hbar} e^{-i\ln(|q|)D/\hbar} |\eta\rangle, \tag{2.15}$$

where the spectrum of Q is taken to be the whole real line. Observe, in this case, that the state $|p, -q\rangle = |p, q\rangle$, namely, the coherent states for $q > 0$ and $q < 0$ are identical to each other. Additionally, in this case we consider $Q|x\rangle = x|x\rangle$. where $-\infty < x < 0$ and $0 < x < \infty$, and $\eta(x) = \langle x|\eta\rangle$ is conveniently chosen as an even function, that is, $\eta(-x) = \eta(x)$. Consequently, using the abbreviation $\langle(\cdot)\rangle \equiv \langle\eta|(\cdot)|\eta\rangle$, we find that

$$\langle p, q | Q | p, q \rangle = |q| \langle Q\rangle = 0 \tag{2.16}$$

for all (p, q), and likewise $\langle p, q|Q^{2r+1}|p, q\rangle = 0$ for any odd power.

However, let us instead consider

$$\left\langle p, q \middle| Q^2 \middle| p, q \right\rangle = q^2 \left\langle Q^2 \right\rangle, \tag{2.17}$$

which, along with the modest requirement that $\langle Q^2 \rangle = 1$, leads to

$$\left\langle p,q \middle| Q^2 \middle| p,q \right\rangle = q^2, \tag{2.18}$$

from which we can conclude that q^2 is the mean value of Q^2. By a suitable choice of the fiducial vector $|\eta\rangle$, we can also arrange that

$$\left\langle p,q \middle| Q^{2r} \middle| p,q \right\rangle = q^{2r} + O(\hbar). \tag{2.19}$$

In the same states $|p,q\rangle$, it follows that

$$\langle p,q|B(D,P,Q)|p,q\rangle = \left\langle B\left(D + p|q|Q, |q|^{-1}P + p, |q|\mathrm{Q}\right)\right\rangle. \tag{2.20}$$

If the function $\eta(x) = \langle x|\eta\rangle$ is real, then besides $\langle Q\rangle = 0$ it also follows that $\langle D\rangle = 0$ and $\langle P\rangle = 0$, which leads to

$$\langle p,q|D|p,q\rangle = 0, \qquad \langle p,q|P|p,q\rangle = p, \tag{2.21}$$

and thus implies that p is the mean value of P in the coherent state $|p,q\rangle$. Additionally, we see that

$$\langle p,q|P^2|p,q\rangle = \left\langle \left(|q|^{-1}P + p\right)^2 \right\rangle = p^2 + q^{-2}\left\langle P^2 \right\rangle, \tag{2.22}$$

where the factor $\langle P^2\rangle = \hbar^2 c$ for a dimensionless constant $c > 0$.

With the foregoing expectation values, we can argue that the expression given by

$$\begin{aligned}\langle p,q|\mathscr{H}|p,q\rangle &\equiv \langle p,q|\frac{1}{2}\left(P^2 + \omega^2 Q^2\right) + \lambda Q^4|p,q\rangle \\ &= \frac{1}{2}\left(p^2 + \hbar^2 c/q^2 + \omega^2 q^2\right) + \lambda_0 q^4 \\ &= \frac{1}{2}\left(p^2 + \omega^2 q^2\right) + \lambda q^4 + O(\hbar)\end{aligned} \tag{2.23}$$

seems to provide a reasonable connection between suitable quantum and classical Hamiltonians.

Traditionally, when dealing with coherent states, one also speaks about a resolution of unity in the form

$$\mathbb{1} = \int |p,q\rangle\langle p,q|w(p,q)\,dp\,dq, \tag{2.24}$$

for some weight function $w(p,q)$ that is positive almost everywhere. However, in this article we shall not focus on this form of the resolution of unity. Instead, we shall rely on the fact that the coherent state overlap function $\langle p'', q'' \mid p', q' \rangle$ serves as a reproducing kernel for a reproducing kernel Hilbert space representation of the underlying abstract Hilbert space $\mathfrak{H}$; see [2–4].

2.4. Many Dimensional Affine Coherent States

We now extend the preceding analysis to a discussion of many dimensional affine coherent states. Consider a set of J (J being *odd*) independent affine fields such as Q_j and D_j, $j = 0, \pm 1, \pm 2, \ldots, \pm J^*$, where $J^* \equiv (J-1)/2$, and the only nonvanishing commutator is given by

$$[Q_l, D_j] = i\hbar \delta_{l,j} Q_l, \tag{2.25}$$

and for each j the spectrum of Q_j is the entire real line save for zero. The coherent states for this system are taken as

$$|p, q\rangle \equiv e^{i\Sigma_j p_j Q_j/\hbar} e^{-i\Sigma_j \ln(|q_j|) D_j/\hbar} |\eta\rangle. \tag{2.26}$$

In terms of the states $|x\rangle$, where $Q_j|x\rangle = x_j|x\rangle$, for all j, it follows that

$$\langle x \mid p, q\rangle = \left[\Pi_j |q_j|^{-1/2}\right] e^{i\Sigma p_j x_j/\hbar}\, \eta(x/|q|), \tag{2.27}$$

where $\eta(x/|q|) \equiv \eta(x_{-J^*}/|q_{-J^*}|, \ldots, x_0/|q_0|, \ldots, x_{J^*}/|q_{J^*}|)$.

As a first example, we assume (with $\hbar = 1$) that

$$\eta(x) = \langle x \mid \eta\rangle = N e^{-(1/2)\omega \Sigma_j x_j^2}. \tag{2.28}$$

In this case,

$$\langle x \mid p, q\rangle = N \left[\Pi_j |q_j|^{-1}\right]^{1/2} e^{i\Sigma_j p_j x_j - (1/2)\omega \Sigma_j (x_j^2/q_j^2)}, \tag{2.29}$$

and the overlap of two such coherent states is given by

$$\langle p', q' \mid p, q\rangle = \prod_{j=-J^*}^{J^*} \frac{\left[2|q'_j|^{-1}|q_j|^{-1}\right]^{1/2}}{\left[q'^{-2}_j + q_j^{-2}\right]^{1/2}} e^{-(1/2\omega)(p'_j - p_j)^2/[q'^{-2}_j + q_j^{-2}]}. \tag{2.30}$$

Let us extend our present example to an infinite number of degrees of freedom, for which $J \to \infty$. This leads to the expression

$$\langle p', q' \mid p, q\rangle = \prod_{j=-\infty}^{\infty} \frac{\left[2|q'_j|^{-1}|q_j|^{-1}\right]^{1/2}}{\left[q'^{-2}_j + q_j^{-2}\right]^{1/2}} e^{-(1/2\omega)(p'_j - p_j)^2/[q'^{-2}_j + q_j^{-2}]}, \tag{2.31}$$

which provides a well-defined product representation for affine coherent states provided that the variables $\{q'_j, q'_j\}_{j=-\infty}^{\infty}$ and $\{p_j, q_j\}_{j=-\infty}^{\infty}$ are well chosen.

2.5. Representation for a Field

A one-dimensional affine field theory involves field operators that satisfy the affine commutation relation

$$[\varphi(x), \rho(y)] = i\hbar\delta(x-y)\varphi(x), \quad x, y \in \mathbb{R}, \tag{2.32}$$

where $\varphi(x)$ is the generalization of Q and $\rho(y)$ generalizes D. Let us regularize this field formulation by introducing a one-dimensional lattice space with $J < \infty$ lattice points (again with J conveniently chosen as an *odd* number) each separated by a lattice spacing $a > 0$, which leads to a regularized affine field representation given by fields φ_k and ρ_k, where x has been replaced by the integer k and $x = ka$. Here, $-J^* \leq k \leq J^*$, and $J^* \equiv (J-1)/2$ as before. These operators obey the affine commutation relation given in the form

$$[\varphi_j, \rho_k] = i\hbar\, a^{-1}\delta_{j,k}\,\varphi_j. \tag{2.33}$$

The coherent states for such a regularized field are given by

$$|p, q\rangle \equiv e^{i\Sigma_j p_j \varphi_j a/\hbar}\, e^{-\Sigma_k \ln(|q_k|)\rho_k a/\hbar} |\eta\rangle. \tag{2.34}$$

For our *first example*, we choose $|\eta\rangle$ so that (with $\hbar = 1$)

$$\langle \phi \mid \eta \rangle = M e^{-(1/2)\omega \Sigma_k \phi_k^2 a}, \tag{2.35}$$

here, the vector $|\phi\rangle$ replaces $|x\rangle$ as used before, where $\varphi(x)|\phi\rangle = \phi(x)|\phi\rangle$. In the present case the overlap function of two coherent states becomes

$$\langle p', q' \mid p, q \rangle = \prod_{j=-J^*}^{J^*} \frac{\left[2\left|q'_j\right|^{-1}\left|q_j\right|^{-1}\right]^{1/2}}{\left[q'^{-2}_j + q_j^{-2}\right]^{1/2}} e^{-(1/2\omega)(p'_j - p_j)^2 a/[q'^{-2}_j + q_j^{-2}]}. \tag{2.36}$$

Next we investigate a possible continuum limit in which $J^* \to \infty$, $a \to 0$, and initially we require that $(2J^* + 1)a = Ja \equiv \mathrm{X}$ may be large but finite; a subsequent limit in which $\mathrm{X} \to \infty$ is taken later. Moreover, in this limit we also insist that $p_j \to p(x)$ and $q_j \to q(x)$, and that both functions are *continuous*. To ensure that we focus on the representations induced by the given fiducial vector, we restrict attention to functions $p(x)$ that have compact support and functions $q(x)$ such that $\ln|q(x)|$ also has compact support, or stated otherwise, $|q(x)| = 1$ outside a compact region. It is clear that the exponential factor exhibits a satisfactory continuum limit given by

$$\Sigma_j \left(q'_j - p_j\right)^2 a / \left[q'^{-2}_j + q_j^{-2}\right] \longrightarrow \int \left[p'(x) - p(x)\right]^2 / \left[q'(x)^{-2} + q(x)^{-2}\right] dx. \tag{2.37}$$

However, the continuum limit of the prefactor turns out to be *identically zero* unless $|q'(x)| = |q(x)|$ for all x! This implies that $\langle p', q' \mid p, q\rangle = 0$ whenever $|q'(x)| \not\equiv |q(x)|$ (as befits a nonseparable Hilbert space!), and thus the operator representation with this fiducial vector has turned out to be highly singular. Stated otherwise, the affine field algebra fails to have an acceptable continuum limit for a strictly Gaussian fiducial vector with the indicated form.

As a *second example*, suppose that

$$\eta(\phi) = \langle \phi \mid \eta\rangle = N \frac{e^{-(1/2)\omega\Sigma_j \phi_j^2 a}}{\left[\Pi_j |\phi_j|\right]^{(J-1)/2J}}. \tag{2.38}$$

This unusual expression leads to a normalizable fiducial vector, that is,

$$N^2 \int \frac{e^{-\omega\Sigma_j \phi_j^2 a}}{\left[\Pi_j |\phi_j|\right]^{(J-1)/J}} \Pi_j d\phi_j = 1, \tag{2.39}$$

the finiteness of which is clear.

For the second example, we find that

$$\langle \phi \mid p, q\rangle = N\left[\Pi_j |q_j|^{-1}\right]^{1/2J} \frac{e^{i\Sigma_k p_k \phi_k a - (1/2)\omega\Sigma_j (\phi_j^2/q_j^2) a}}{\left[\Pi_j |\phi_j|\right]^{(J-1)/2J}}, \tag{2.40}$$

and the coherent state overlap function $\langle p', q' \mid p, q\rangle$ for the second example is given by the integral

$$\begin{aligned}
\langle p', q' \mid p, q\rangle &= N^2\left[\Pi_k |q'_k|^{-1} |q_k|^{-1}\right]^{1/2J} \int e^{-i\Sigma_k (p'_k - p_k)\phi_k a} \\
&\quad \times e^{-(1/2)\omega\Sigma_k \phi_k^2 [q'^{-2}_k + q_k^{-2}] a} \frac{1}{\left[\Pi_j |\phi_j|\right]^{(J-1)/J}} \Pi_k d\phi_k \\
&= N^2 \prod_k \frac{\left[2|q'_k|^{-1} |q_k|^{-1}\right]^{1/2J}}{\left[q'^{-2}_k + q_k^{-2}\right]^{1/2J}} \\
&\quad \times \int \frac{e^{-i\Sigma_k (p'_k - p_k)\phi_k a / [(q'^{-2}_k + q_k^{-2})/2]^{1/2}} e^{-\omega\Sigma_k \phi_k^2 a}}{\left[\Pi_j |\phi_j|\right]^{(J-1)/J}} \Pi_k d\phi_k.
\end{aligned} \tag{2.41}$$

The normalization factor in this relation follows from the fact that

$$1 = \langle p, q \mid p, q\rangle = N^2 \int \frac{e^{-\omega\Sigma_k \phi_k^2 a}}{\left[\Pi_j |\phi_j|\right]^{(J-1)/J}} \Pi_k d\phi_k. \tag{2.42}$$

In considering the continuum limit, we again focus on continuous functions $p(x)$ and $\ln|q(x)|$ that have compact support, and we initially restrict attention to a finite overall interval $X = Ja$. In this case, the new prefactor involves the $1/J$th root of the previous prefactor, and this new version leads to an acceptable continuum limit. In particular, it is possible to evaluate the prefactor itself exactly—which also equals the coherent state matrix element $\langle p, q' \mid p, q\rangle$—as

$$\begin{aligned}\langle p, q' \mid p, q\rangle &= \lim_{a\to 0}\prod_j \frac{\left[2\left|q'_j\right|^{-1}\left|q_j\right|^{-1}\right]^{1/2J}}{\left[q_j'^{-2} + q_j^{-2}\right]^{1/2J}} \\ &= \lim_{a\to 0}\exp\left[-1/(2Ja)\Sigma_j\left\{\ln\left|q'_j\right| + \ln\left|q_j\right| + \ln\left[\left(q_j'^{-2} + q_j^{-2}\right)/2\right]\right\}a\right] \\ &= \exp\left[-(1/2X)\int\left\{\ln\left|q'(x)\right| + \ln\left|q(x)\right| + \ln\left[q'(x)^{-2} + q(x)^{-2}/2\right]\right\}dx\right].\end{aligned} \tag{2.43}$$

For arguments of compact support, it follows in the limit that $X \to \infty$, that the prefactor becomes *unity* and thus $\langle p, q' \mid p, q\rangle = 1$ for all arguments! While unusual, this result is perfectly acceptable from a reproducing kernel point of view. (Indeed, this result is surely more acceptable than the conclusion for the first example where $\langle p, q' \mid p, q\rangle$ was identically zero except when $|q'(x)| = |q(x)|$ for all x.) Moreover, for the second example, the expression $\langle p', q' \mid p, q\rangle$ is well defined and is continuous in its labels in the continuum limit when X is finite as well as in the further limit that $X \to \infty$. Thus, although we have not explicitly evaluated the coherent state overlap function in the general case for the second example, it is clearly a continuous function of positive type suitable to be a reproducing kernel for the (separable) Hilbert space of interest.

2.6. Construction of Field Theoretic Affine Coherent States

In light of the foregoing discussion, it is but a small step to introduce the affine coherent states of interest in the study of nonrenormalizable scalar field models. In the rest of this section, we introduce the ground state of our proposed models and define the associated affine coherent states using that ground state as the fiducial vector; in the following section we give a brief discussion that outlines the motivation for the particular choice of the ground state for these models.

Our analysis is aimed at a field theory model regularized by a spacetime lattice that has one time dimension and s space dimensions. Focusing on the spatial aspects, we put $L < \infty$ lattice points in each of the s directions with a lattice spacing of $a > 0$. Each lattice point is labeled by a multi-integer $k = (k_1, k_2, \ldots, k_s)$, where each $k_j \in \mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\}$. The total number of lattice points in a spatial slice of the lattice is given by $N' \equiv L^s$, a factor which plays the role that J played in examples one and two above. In the representation in which the field operator is diagonal, there is a field associated with each lattice point, ϕ_k, much as in examples one and two above, save for the fact that now the label k is s-dimensional. In the present case, the ground state of the system Hamiltonian is chosen (with $\hbar = 1$) as

$$\Psi_0(\phi) = M\frac{e^{-(1/2)\Sigma'_{k,l}\phi_k A_{k-l}\phi_l a^{2s} - (1/2)W(\phi a^{(s-1)/2})}}{\Pi'_k\left[\Sigma'_l J_{k,l}\phi_l^2\right]^{(N'-1)/4N'}}, \tag{2.44}$$

where primes on products and sums signify they apply only to a given spatial slice. In this expression, A_{k-l} is a numerical matrix proportional to $a^{-(1+s)}$ and $J_{k,l} = 1/(2s+1)$ for $l = k$ and when l is any of the $2s$ nearest neighbors of k in the same spatial slice; otherwise $J_{k,l} = 0$. The role of $J_{k,l}$ is to provide a local average of field values in the sense that $\overline{\phi_k^2} \equiv \Sigma_l J_{k,l}\phi_l^2$, which means that the denominator exhibits an integrable singularity for the ground state distribution even in the limit $N' \to \infty$. The A factor in the exponent and the J factor in the denominator well represent the functional form of the ground state for large and small field values, respectively. The unspecified function W (discussed in the following section) is needed to modify intermediate field values.

The affine coherent states for this fiducial vector are given by

$$\begin{aligned}\langle \phi \mid p, q\rangle &= M\frac{\Pi'_k\left[|q_k|^{-1/2}\right]}{\Pi'_k\left[\Sigma'_l J_{k,l}(\phi_l^2/q_l^2)\right]^{(N'-1)/4N'}} \\ &\quad\times e^{i\Sigma'_k p_k\phi_k a^s-(1/2)\Sigma'_{k,l}(\phi_k/|q_k|)A_{k-l}(\phi_l/|q_l|)a^{2s}-(1/2)W((\phi/|q|)a^{(s-1)/2})},\end{aligned} \tag{2.45}$$

which leads to the overlap expression

$$\begin{aligned}\langle p', q' \mid p, q\rangle &= M^2\,\Pi'_k\left\{\left[|q'_k|^{-1/2}\right]\left[|q_k|^{-1/2}\right]\right\}\int \Pi'_k d\phi_k e^{-i\Sigma'_k(p'_k-p_k)\phi_k a} \\ &\quad\times e^{-(1/2)\Sigma'_{k,l}(\phi_k/|q'_k|)A_{k-l}(\phi_l/|q'_l|)a^{2s}-(1/2)W((\phi/|q'|)a^{(s-1)/2})} \\ &\quad\times e^{-(1/2)\Sigma'_{k,l}(\phi_k/|q_k|)A_{k-l}(\phi_l/|q_l|)a^{2s}-(1/2)W((\phi/|q|)a^{(s-1)/2})} \\ &\quad\times\frac{1}{\Pi'_k\left[\Sigma'_l J_{k,l}\left(\phi_l^2/q_l'^{-2}\right)\right]^{(N'-1)/4N'}\;\Pi'_k\left[\Sigma'_l J_{k,l}(\phi_l^2/q_l^2)\right]^{(N'-1)/4N'}}.\end{aligned} \tag{2.46}$$

In the present case, and when $p'_k = p_k$ for all k, observe that a simple change of variables does not eliminate the q' and q variables from the integrand as was the case for example two. However, we do expect that in the continuum limit, the prefactor will again cancel with the coefficients q' and q from the denominator factor for the simple reason that in the continuum limit, all neighboring q' and q values will become equal to ensure a continuous label function $q'(x)$ and $q(x)$, and as such they will emerge from the denominator and join the prefactor leading, in the case of an infinite spatial volume, to a factor of unity, much as was the case for example two. Thus we expect the continuum limit for an infinite volume also to be given by the expression

$$\begin{aligned}\langle p', q' \mid p, q\rangle &= \lim_{a\to 0} M^2\int \Pi'_k d\phi_k\, e^{-i\Sigma'_k(p'_k-p_k)\phi_k a} \\ &\quad\times e^{-(1/2)\Sigma'_{k,l}(\phi_k/|q'_k|)A_{k-l}(\phi_l/|q'_l|)a^{2s}-(1/2)W((\phi/|q'|)a^{(s-1)/2})} \\ &\quad\times e^{-(1/2)\Sigma'_{k,l}(\phi_k/|q_k|)A_{k-l}(\phi_l/|q_l|)a^{2s}-(1/2)W((\phi/|q|)a^{(s-1)/2})} \\ &\quad\times\frac{1}{\left[\Pi'_k\left[\Sigma'_l J_{k,l}\phi_l^2\right)\right]^{(N'-1)/2N'}}.\end{aligned} \tag{2.47}$$

On the other hand, in a finite spatial volume, the preceding expression for the coherent state overlap function would be multiplied by the factor

$$e^{-(1/2V')\int[\ln|q'(x)|+\ln|q(x)|]d^s x}, \tag{2.48}$$

where $V' \equiv N'\,a^s = (La)^s < \infty$ is the volume of the spatial slice.

This concludes our discussion of the affine coherent states relevant for the field theory models of interest.

3. Brief Review of the Author's Approach to Nonrenormalizable Models

We now outline the author's program to study the Euclidean-space formal functional integral given (for $\hbar = 1$ and $n \geq 5$) by

$$S(h) = \mathcal{N}\int \exp\left(\int\left\{h\phi - \frac{1}{2}\left[(\nabla\phi)^2 + m_0^2\phi^2\right] - g_0\phi^4\right\}d^n x\right)\mathfrak{D}\phi. \tag{3.1}$$

A study of such an expression via conventional perturbation theory is not acceptable because of the unlimited number of new counterterms needed such as $\phi^6, \phi^8, \phi^{10}, \ldots$, and so forth, as well as higher derivative terms as well. We need a radically new approach.

3.1. A Lattice Framework to Build on

We now sketch our alternative approach. First, we adopt a finite, periodic, hypercubic, spacetime lattice with $L < \infty$ sites on a side, a lattice spacing of $a > 0$, and lattice sites labeled by a multi-integer $k = \{k_0, k_1, k_2, \ldots, k_s\}$, $k_j \in \mathbb{Z}$, where $s = n - 1$ is the number of spatial dimensions and k_0 denotes the Euclidean time direction, which will become the true time direction after a Wick rotation. Second, we approximate the former equation by a lattice functional integral given by

$$N\int \exp\left(\Sigma_k\left\{h_k\phi_k a^n - \frac{1}{2}(\phi_{k^*} - \phi_k)^2 a^{n-2} - \frac{1}{2}m_0^2\phi_k^2 a^n - g_0\phi_0^4 a^n\right\}\right)\Pi_k d\phi_k, \tag{3.2}$$

where k^* denotes each of the n next nearest neighbors in a positive sense and a summation over such points is implicit. As it stands, this lattice expression represents a lattice cutoff of the formal continuum functional integral. We need to introduce a counterterm into this expression to account for the needed renormalizations that will appear in the continuum limit. Choosing counterterms on the basis of perturbation theory is inappropriate, and we need an alternative principle to choose the counterterms.

3.2. Hard-Core Interactions

There are strong reasons to believe that a perturbation theory *about the free theory* does not hold true for nonrenormalizable models. In fact, the author has long argued [1] that the nonlinear interaction term for nonrenormalizable models acts partially like a hard core, which, in a functional integral formulation, acts to project out certain field histories that would otherwise be allowed by the free theory alone. An argument favoring this explanation is provided by the Sobolev-type inequality

$$\left\{\int \phi(x)^4 d^n x\right\}^{1/2} \leq C\left\{\int\left[(\nabla\phi(x))^2 + m^2\phi(x)^2\right] d^n x\right\}, \tag{3.3}$$

which for $n \leq 4$ holds for $C = 4/3$, while for $n \geq 5$ holds only for $C = \infty$ [5]. In the latter case, this means that there are functions $\phi(x)$ such that the left hand side diverges while the right hand side is finite. One example of such a function is given by

$$\phi_{\text{singular}}(x) = |x|^{-p} e^{-x^2}, \qquad n/4 \leq p < n/2 - 1. \tag{3.4}$$

The issues discussed above are even more self evident for a one dimensional system with the classical action

$$I = \int\left[\frac{1}{2}\left(\dot{x}(t)^2 - x(t)^2\right) - g x(t)^{-4}\right] dt, \tag{3.5}$$

for which, when $g > 0$, the paths are unable to penetrate the barrier at $x = 0$, and thus the family of theories for $g > 0$ also exhibit a hard-core behavior and they do *not* converge to the free theory as $g \to 0$. This hard-core behavior holds in both the classical and quantum theories for this example.

3.3. Pseudofree Models

If the interacting theories do not pass to the free theory as the nonlinear coupling constant reduces to zero, to what limit do they converge? We have introduced the term *pseudofree model* to label the limit of the interacting theories when the coupling constant goes to zero. It is the pseudofree theory about which a perturbation exists, if one exists at all, and not about the usual free theory. Our initial goal in understanding nonrenormalizable theories is to get a handle on the pseudofree theory, a theory that is fundamentally different from the usual free theory.

As a result of the lack of any connection of the interacting theories with the usual free theory, our procedure to choose the proper counterterm for nonrenormalizable models will turn out to be somewhat circuitous.

3.4. Role of Sharp Time Averages

Let us consider the average of powers of the expression

$$\Sigma_{k_0} F(\phi, a) a, \tag{3.6}$$

where $F(\phi, a)$ is a function of lattice points all at a fixed value of k_0, in any lattice spacetime, based on the distribution generated by the exponential of the action, which we denote by $\langle(\cdot)\rangle$. We write the average of the pth power of this expression as

$$\langle[\Sigma_{k_0}F(\phi,a)a]^p\rangle = \Sigma_{k_{0_1},k_{0_2},\ldots,k_{0_p}}a^p\langle F(\phi_1,a)F(\phi_2,a)\cdots F(\phi_p,a)\rangle, \tag{3.7}$$

where ϕ_j here refers to the fact that "$k_0 = j$" in this term. A straightforward inequality leads to

$$\begin{aligned}&|\langle F(\phi_1,a)F(\phi_2,a)\cdots F(\phi_p,a)\rangle|\\ &\quad\le |\langle[F(\phi_1,a)]^p\rangle\langle[F(\phi_2,a)]^p\rangle\cdots\langle[F(\phi_p,a)]^p\rangle|^{1/p},\end{aligned} \tag{3.8}$$

which casts the problem into one at a sharp time. For sufficiently large L, it follows that this sharp time expression may be given by

$$\langle[F(\phi,a)]^p\rangle = \int [F(\phi,a)]^p\Psi_0(\phi)^2\Pi'_k d\phi_k, \tag{3.9}$$

where the integral is taken over fields at a fixed value of k_0, $\Psi_0(\phi)^2$ denotes the ground state distribution, and Π'_k denotes a product over the spatial lattice at a fixed value of k_0. Thus we have arrived at the important conclusion that if the sharp time average is finite, then the full spacetime average is also finite.

Both the inequality noted above and the argument involving the ground state distribution may be found in [6–8].

3.5. Choosing the Ground State

Attention now turns to finding the ground state, or more to the point, *choosing* the ground state so that expressions we desire to be finite actually become finite. In particular, this remark means that the ground state is *tailored* or *designed* so that those quantities that are divergent in the usual perturbation theory are in fact rendered finite. Once the ground state is chosen, one defines the associated lattice Hamiltonian for the system by means of the expression

$$\mathscr{H} \equiv -\frac{1}{2}\hbar^2\Sigma'_k\partial^2/\partial\phi_k^2 + \frac{1}{2}\hbar^2\ [1/\Psi_0(\phi)]\Sigma'_k\partial^2\Psi_0(\phi)/\partial\phi_k^2, \tag{3.10}$$

where Σ'_k denotes a sum over a spatial lattice at a fixed value of k_0. Note the appearance of $\hbar^2$ in both factors. In turn, the lattice Hamiltonian readily leads to the lattice action. In summary, we focus first on the desired modification of the ground state for the system, which leads to the lattice Hamiltonian, and finally to the desired lattice action.

It has been argued that the ground state for a free system ($g_0 = 0$) and with no counterterm is clearly a Gaussian. Moreover, such a function leads to divergences for several quantities of interest, such as those expressions basic to a mass perturbation, namely, when $F(\phi, a) = \Sigma'_k\phi_k^2 a^s$. It has also been argued that the *source* of those divergences can be traced to a specific, single factor when the integrals involved are reexpressed in hyperspherical coordinates, which are defined by

$$\phi_k \equiv \kappa\eta_k, \quad \Sigma'_k\phi_k^2 = \kappa^2, \quad \Sigma'_k\eta_k^2 = 1, \\ \kappa \geq 0, \; -1 \leq \eta_k \leq 1. \tag{3.11}$$

As an example, consider a typical (Gaussian) integral of interest given by

$$\begin{aligned} K\int \left[\Sigma'_k\phi_k^2 a^s\right]^p e^{-\Sigma'_{k,l}\phi_k A_{k-l}\phi_l a^{2s}}\Pi'_k d\phi_k \\ = 2K\int \kappa^{2p}a^{sp}e^{-\kappa^2\Sigma'_{k,l}\eta_k A_{k-l}\eta_l s^{2s}}\kappa^{(N'-1)}d\kappa\, \delta\left(1-\Sigma'_k\eta_k^2\right)\Pi'_k d\eta_k. \end{aligned} \tag{3.12}$$

Here, $A_{k-l} \propto a^{-(1+s)}$ is a matrix responsible for the spatial gradient and other suitable quadratic terms in the lattice Hamiltonian, and N' is the total number of lattice sites in a spatial slice of the lattice. In the continuum limit, in which $a \to 0$, it follows that $L \to \infty$ in such a way, initially, that $V' \equiv (La)^s = N'a^s < \infty$. Later, one may extend the continuum limit procedure so that $V' \to \infty$ as well. In short, in the continuum limit, it follows that $N' \to \infty$. It is not difficult to show that in the foregoing integral, as displayed in hyperspherical coordinates, it is the term N' in the measure factor $\kappa^{(N'-1)}$ that is *the very source of the divergences!* If one could change that factor to one that remains *finite* in the continuum limit, *the divergences would disappear!*

How is one to change a factor that has arisen from a bona-fide coordinate transformation? The answer is as follows: it cannot be done *directly*, but it can be done *indirectly*. The way to do so is to choose a different, non-Gaussian ground state distribution, one that has roughly the form

$$K'\frac{1}{\kappa^{(N'-1)}}e^{-\Sigma'_{k,l}\phi_k A_{k-l}\phi_l a^{2s}}, \tag{3.13}$$

a form that has an additional factor in the denominator to cancel the measure factor $\kappa^{(N'-1)}$ altogether; in fact, it does not need to cancel it all, but it is a reasonable place to begin. Such a ground state distribution arises from a Hamiltonian that is not quadratic but contains another component that we identify as the desired counterterm. Finally, that counterterm is taken over to the lattice action, and thereby we have determined the counterterm in this convoluted manner!

There are many ways to choose the Hamiltonian so that the counterterm leads to a modification of the ground state of the desired form. A large class of ground state modifications may be given by

$$M\frac{1}{\Pi'_k\left[\Sigma'_l J_{k,l}\phi_l^2\right]^{(N'-1)/4N'}}e^{-(1/2)\Sigma'_{k,l}\phi_k A_{k-l}\phi_l-(1/2)W(\phi a^{(s-1)/2})} \tag{3.14}$$

for various choices of the constant coefficients $J_{k,l}$. For example, one may choose $J_{k,l} = \delta_{k,l}$, and this is appropriate to discuss ultralocal models, which may be described as relativistic models with their spatial gradient terms removed. As we have shown elsewhere, such a choice leads to a Poisson ground state distribution, which, although appropriate and correct for ultralocal models, is not desirable for truly relativistic models. For relativistic models, on the other hand, it has been proposed [9] to choose the expression

$$J_{k,l} = \frac{1}{2s+1}\delta_{k,l\in\{k\cup\mathrm{nn}\}}, \tag{3.15}$$

where the expression $l \in \{k \cup \mathrm{nn}\}$ means that $l = k$ and all the spatial nearest neighbors to k, $J_{k,l} = 0$ elsewhere in the spatial slice, and $\Sigma'_l J_{k,l} = 1$.

With $J_{k,l}$ so chosen, it follows that the Hamiltonian does not represent a local interaction in the continuum limit due to cross terms coming from one derivative each of the A and J terms. To fix that, the unwanted cross terms are removed from the lattice Hamiltonian by means of a suitable, auxiliary term $W(\phi a^{(s-1)/2})$ in the ground state that effects mid-level field values. Readers who may be interested in what form of counterterm for the lattice action such a modified ground state leads to should consult [6–8].

This is not the place to debate the merits of the suggested proposal for the relativistic models and their proposed ground state wave functions. Rather, in this paper, we accept the suggested form of the ground state, and, consequently, we are then led to the set of affine coherent states with the indicated ground state chosen as the fiducial vector that was discussed in the previous section.

4. Commentary

The discussion of affine coherent states for a single degree of freedom emphasized the difficulty in ensuring that $\langle p,q|Q|p,q\rangle = q$ for all $q \in \mathbb{R} \setminus 0$. The solution involved a superposition of disjoint states. When generalized to infinitely many degrees of freedom, this disjoint feature would have led to an infinite number of unitarily inequivalent representations of the affine variables. To avoid this situation, we instead accepted the requirement that $\langle p,q|Q^2|p,q\rangle = q^2$, a compromise which eventually led to a suitable representation for infinitely many degrees of freedom. (It is interesting to observe that this modification has some similarities with the Wilson construction for wavelets which also involves symmetric fiducial vectors; e.g., see [10].)

For quantum gravity, which is the principal nonrenormalizable model of interest, it is noteworthy that the classical field variable $g_{ab}(x)$, the 3×3 spatial metric, forms a positive-definite matrix and thus it is possible to define affine coherent states for quantum gravity such that

$$\langle \pi, g|\hat{g}_{ab}(x)|\pi, g\rangle = g_{ab}(x), \tag{4.1}$$

with an irreducible representation of the appropriate affine variables; see [11, 12]. This fact means that for the gravitational field we have the best situation we could hope for!

References

[1] J. R. Klauder, "Remarks on nonrenormalizable interactions," *Physics Letters B*, vol. 47, no. 6, pp. 523–525, 1973.

[2] N. Aronszajn, "La théorie des noyaux reproduisants et ses applications, Premiere Partie," *Proceedings of the Cambridge Philosophical Society*, vol. 39, pp. 133–153, 1943.

[3] N. Aronszajn, "Note additionnelle l'article La théorie des noyaux reproduisants et ses applications," *Proceedings of the Cambridge Philosophical Society*, vol. 39, p. 205, 1943.
[4] N. Aronszajn, "Theory of reproducing Kernels," *Transactions of the American Mathematical Society*, vol. 68, pp. 337–404, 1950.
[5] J. R. Klauder, *Beyond Conventional Quantization*, Cambridge University Press, Cambridge, Mass, USA, 2000.
[6] J. R. Klauder, "Divergence-free nonrenormalizable models," *Journal of Physics A*, vol. 41, no. 33, Article ID 335206, 15 pages, 2008.
[7] J. R. Klauder, "Taming nonrenormalizability," *Journal of Physics A*, vol. 42, no. 33, Article ID 335208, 7 pages, 2009.
[8] J. R. Klauder, "The legacy of Alladi Ramakrishnan in the mathematical sciences," (K. Alladi, J. R. Klauder, C. R. Rao, Eds.), Springer (2010, to appear); http://arxiv.org/abs/0904.2869.
[9] J. R. Klauder, "A new approach to nonrenormalizable models," *Annals of Physics*, vol. 322, no. 11, pp. 2569–2602, 2007.
[10] I. Daubechies, S. Jaffard, and J.-L. Journé, "A simple Wilson orthonormal basis with exponential decay," *SIAM Journal on Mathematical Analysis*, vol. 22, no. 2, pp. 554–573, 1991.
[11] J. R. Klauder, "Noncanonical quantization of gravity. I. Foundations of affine quantum gravity," *Journal of Mathematical Physics*, vol. 40, no. 11, pp. 5860–5882, 1999.
[12] J. R. Klauder, "Noncanonical quantization of gravity. II. Constraints and the physical Hilbert space," *Journal of Mathematical Physics*, vol. 42, no. 9, pp. 4440–4465, 2001.

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 509538, 12 pages
doi:10.1155/2010/509538

Review Article

Gauge Symmetry and Howe Duality in 4D Conformal Field Theory Models

Ivan Todorov

Institute for Nuclear Research and Nuclear Energy, Tsarigradsko Chaussee 72, 1784 Sofia, Bulgaria

Correspondence should be addressed to Ivan Todorov, todorov@inrne.bas.bg

Received 17 September 2009; Accepted 28 October 2009

Academic Editor: Richard Kerner

It is known that there are no local scalar Lie fields in more than two dimensions. Bilocal fields, however, which naturally arise in conformal operator product expansions, do generate infinite Lie algebras. It is demonstrated that these Lie algebras of local observables admit (highly reducible) unitary positive energy representations in a Fock space. The multiplicity of their irreducible components is governed by a compact gauge group. The mutually commuting observable algebra and gauge group form a dual pair in the sense of Howe. In a theory of local scalar fields of conformal dimension two in four space-time dimensions the associated dual pairs are constructed and classified.

1. Introduction

We review results of [1–7] on 4D conformal field theory (CFT) models, which can be summed up as follows. The requirement of *global conformal invariance* (GCI) in compactified Minkowski space together with the Wightman axioms [8] implies the Huygens principle (3.6) and rationality of correlation functions [1]. A class of 4D GCI quantum field theory models gives rise to a (reducible) Fock space representation of a pair consisting of an infinite-dimensional Lie algebra $\mathcal{L}$ and a commuting with its compact Lie group U. The state space $\mathcal{F}$ splits into a direct sum of irreducible $\mathcal{L} \times U$ modules, so that each irreducible representation (IR) of $\mathcal{L}$ appears with a multiplicity equal to the dimension of an associated IR of U. The pair $(\mathcal{L}, U)$ illustrates a interconnects two independent developments: (i) it appears as a *reductive dual pair* [9, 10], within (a central extension of) an infinite-dimensional symplectic Lie algebra; (ii) it provides a representation theoretic realization of the Doplicher-Haag-Roberts' (DHR) theory of superselection sectors and compact gauge groups, [11, 12]. I will first briefly recall Howe's and DHR's theories, then (in Section 2) I will explain how some 2D CFT technics can be extended to four space-time dimensions (in spite of persistent doubts that this is at all possible). After these preliminaries we will proceed with our survey of 4D CFT models and associated infinite-dimensional Lie algebras which relate the two independent developments.

1.1. Reductive Dual Pairs

The notion of a (reductive) dual pair was introduced by Roger Howe in an influential preprint of the 1970s that was eventually published in [10]. It was previewed in two earlier papers of Howe [9, 13], highlightening the role of the Heisenberg group and the applications of dual pairs to physics. For Howe a dual pair, the counterpart for groups and for Lie algebras of the mutual commutants of von Neumann algebras [12] is a (highly structured) concept that plays a unifying role in such widely different topics as Weil's metaplectic group approach [14, 15] to θ functions and automorphic forms (an important chapter in number theory) and the quantum mechanical Heisenberg group along with the description of massless particles in terms of the ladder representations of $U(2,2)$ [16], among others (in physics).

Howe begins in [9] with a $2n$-dimensional real symplectic manifold $\mathcal{W} = \mathcal{V} + \mathcal{V}'$ where $\mathcal{V}$ is spanned by n symbols a_i, $i = 1, \dots, n$, called *annihilation operators* and $\mathcal{V}'$ is spanned by their conjugate, the *creation operators* a_i^* satisfying the canonical commutation relations (CCR)

$$\left[a_i, a_j\right] = 0 = \left[a_i^*, a_j^*\right], \quad \left[a_i, a_j^*\right] = \delta_{ij}. \tag{1.1}$$

The commutator of two elements of the real vector space $\mathcal{W}$ being a real number it defines a (nondegenerate, skew-symmetric) bilinear form on it which vanishes on $\mathcal{V}$ and on $\mathcal{V}'$ separately and for which $\mathcal{V}'$ appears as the dual space to $\mathcal{V}$ (the space of linear functionals on $\mathcal{V}$). The real symplectic Lie algebra $sp(2n, \mathbb{R})$ spanned by antihermitean quadratic combinations of a_i and a_j^* acts by commutators on $\mathcal{W}$ preserving its reality and the above bilinear form. This action extends to the *Fock space* $\mathcal{F}$ (unitary, irreducible) representation of the CCR. It is, however, only exponentiated to the double cover of $Sp(2n, \mathbb{R})$, the *metaplectic group* $Mp(2n)$ (that is not a matrix group, i.e., has no faithful finite-dimensional representation; we can view its Fock space, called by Howe [9] *oscillator representation* as the defining one). Two subgroups G and G' of $Mp(2n)$ are said to form a *(reductive) dual pair* if they act reductively on $\mathcal{F}$ (that is automatic for a unitary representation like the one considered here) and each of them is the full centralizer of the other in $Mp(2n)$. The oscillator representation of $Mp(2n)$ displays a *minimality* property, [17, 18] that keeps attracting the attention of both physicists and mathematicians, see, for example, [19–21].

1.2. Local Observables Determine a Compact Gauge Group

Observables (unlike charge carrying fields) are left invariant by (global) gauge transformations. This is, in fact, part of the definition of a gauge symmetry or a *superselection rule* as explained by Wick et al. [22]. It required the nontrivial vision of Rudolf Haag to predict in the 1960s that a local net of obsevable algebras should determine the compact gauge group that governs the structure of its superselection sectors (for a review and references to the original work see [12]). It took over 20 years and the courage and dedication of Haag's (then) young collaborators, Doplicher and Roberts [11], to carry out this program to completion. They proved that all superselection sectors of a local QFT $\mathcal{A}$ with a mass gap are contained in the vacuum representation of a canonically associated (graded local) field extension $\mathcal{E}$, and they are in a one-to-one correspondence with the unitary irreducible representations (IRs) of a compact gauge group G of internal symmetries of $\mathcal{E}$, so that $\mathcal{A}$ consists of the fixed points of $\mathcal{E}$ under G. The pair $(\mathcal{A}, \mathcal{G})$ in $\mathcal{E}$ provides a general realization of a dual pair in a local quantu theory.

2. How Do 2D CFT Methods Work in Higher Dimensions?

A number of reasons are given why 2-dimensional conformal field theory is, in a way, exceptional so that extending its methods to higher dimensions appears to be hopeless.

(1) The $2D$ conformal group is infinite-dimensional. It is the direct product of the diffeomorphism groups of the left and right (compactified) light rays. (In the euclidean picture it is the group of analytic and antianalytic conformal mappings.) By contrast, for $D > 2$, according to the Liouville theorem, the quantum mechanical conformal group in D space-time dimensions is finite (in fact, $(D+1)(D+2)/2$)-dimensional. It is (a covering of) the spin group Spin $(D, 2)$.

(2) The representation theory of affine Kac-Moody algebras [23] and of the Virasoro algebra [24] is playing a crucial role in constructing soluble $2D$ models of (rational) CFT. There are, on the other hand, no local Lie fields in higher dimensions. After an inconclusive attempt by Robinson [25] (criticized in [26]) this was proven for scalar fields by Baumann [27].

(3) The light cone in two dimensions is the direct product of two light rays. This geometric fact is the basis of splitting $2D$ variables into right- and left-movers' *chiral variables*. No such splitting seems to be available in higher dimensions.

(4) There are chiral algebras in $2D$ CFT whose *local currents* satisfy the axioms of *vertex algebras* (As a mathematical subject vertex algebras were anticipated by Frenkel and Kac [28] and introduced by Borcherds [29]; for reviews and further references see, e.g., [30, 31].) and have rational correlation functions. It was believed for a long time that they have no physically interesting higher-dimensional CFT analogue.

(5) Furthermore, the chiral currents in a $2D$ CFT on a torus have elliptic correlation functions [32], the 1-point function of the stress energy tensor appearing as a modular form (these can be also interpreted as finite temperature correlation functions and a thermal energy mean value on the Riemann sphere). Again, there seemed to be no good reason to expect higher-dimensional analogues of these attractive properties.

We will argue that each of the listed features of $2D$ CFT does have, when properly understood, a higher-dimensional counterpart.

(1) The presence of a conformal anomaly (a nonzero Virasoro central charge c) tells us that the infinite conformal symmetry in $1 + 1$ dimension is, in fact, broken. What is actually used in $2D$ CFT are the (conformal) *operator product expansions* (OPEs) which can be derived for any D and allow to extend the notion of a primary field (e.g., with respect to the stress-energy tensor).

(2) For $D = 4$, infinite-dimensional Lie algebras are generated by *bifields* $V_{ij}(x_1, x_2)$ which naturally arise in the OPE of a (finite) set of (say, hermitean, scalar) local fields ϕ_i of dimension $d(> 1)$:

$$\left(x_{12}^2\right)^d \phi_i(x_1)\phi_j(x_2) = N_{ij} + x_{12}^2 V_{ij}(x_1, x_2) + O\left(\left(x_{12}^2\right)^2\right), \tag{2.1}$$

$$x_{12} = x_1 - x_2, \quad x^2 = \mathbf{x}^2 - x^{0^2}, \quad N_{ij} = N_{ji} \in \mathbb{R}, \tag{2.2}$$

where V_{ij} are defined as (infinite) sums of OPE contributions of (twist two) conserved local tensor currents (and the real symmetric matrix (N_{ij}) is positive definite). We say more on this in what follows (reviewing results of [2–7]).

(3) We will exhibit a factorization of higher-dimensional intervals by using the following parametrization of the conformally compactified space-time ([33–36]):

$$\overline{M} = \left\{ z_\alpha = e^{it} u_\alpha, \alpha = 1, \ldots, D;\ t, u_\alpha \in \mathbb{R};\ u^2 = \sum_{\alpha=1}^{D} u_\alpha^2 = 1 \right\} = \frac{\mathbb{S}^{D-1} \times \mathbb{S}^1}{\{1, -1\}}. \tag{2.3}$$

The real interval between two points $z_1 = e^{it_1} u_1$, $z_2 = e^{it_2} u_2$ is given by

$$z_{12}^2 \left(z_1^2 z_2^2\right)^{-1/2} = 2(\cos t_{12} - \cos \alpha) = -4 \sin t_+ \sin t_-, \quad z_{12} = z_1 - z_2, \tag{2.4}$$

$$t_\pm = \frac{1}{2}(t_{12} \pm \alpha), \quad u_1 \cdot u_2 = \cos \alpha, \quad t_{12} = t_1 - t_2. \tag{2.5}$$

Thus t_+ and t_- are the compact picture counterparts of "left" and "right" chiral variables (see [36]). The factorization of $2D$ cross-ratios into chiral parts again has a higher-dimensional analogue [37]:

$$s := \frac{x_{12}^2 x_{34}^2}{x_{13}^2 x_{24}^2} = u_+ u_-, \qquad t := \frac{x_{14}^2 x_{23}^2}{x_{13}^2 x_{24}^2} = (1 - u_+)(1 - u_-), \qquad x_{ij} = x_i - x_j, \tag{2.6}$$

which yields a separation of variables in the d'Alembert equation (cf. equation (2.1)). One should, in fact, be able to derive the factorization (2.6) from (2.4).

(4) It turns out that the requirement of *global conformal invariance* (GCI) in Minkowski space together with the standard Wightman axioms of local commutativity and energy positivity entails the rationality of correlation functions in any even number of space-time dimensions [1]. Indeed, GCI and local commutativity of Bose fields (for space-like separations of the arguments) imply the *Huygens principle* and, in fact, the strong (algebraic) locality condition

$$\left(x_{12}^2\right)^n \left[\phi_i(x_1), \phi_j(x_2)\right] = 0 \quad \text{for } n \text{ sufficiently large}, \tag{2.7}$$

a condition only consistent with the theory of free fields for an even number of space time dimensions. It is this Huygens locality condition which allows the introduction of higher-dimensional vertex algebras [35, 36, 38].

(5) Local GCI fields have elliptic thermal correlation functions with respect to the (differences of) conformal time variables in any even number of space-time dimensions; the corresponding energy mean values in a Gibbs (KMS) state (see, e.g., [12]) are expressed as linear combinations of modular forms [36].

The rest of the paper is organized as follows. In Section 3 we reproduce the general form of the 4-point function of the bifield V and the leading term in its conformal partial wave expansion. The case of a theory of scalar fields of dimension $d = 2$ is singled out, in which the bifields (and the unit operator) close a commutator algebra. In Section 4 we classify the arising infinite-dimensional Lie algebras $\mathcal{L}$ in terms of the three real division rings $\mathbb{F} = \mathbb{R}, \mathbb{C}, \mathbb{H}$. In Section 5 we formulate the main result of [6, 7] on the Fock space representations of the Lie algebra $\mathcal{L}(\mathbb{F})$ coupled to the (dual, in the sense of Howe [9]) compact gauge group $U(N, \mathbb{F})$, where N is the central charge of $\mathcal{L}$.

3. Four-Point Functions and Conformal Partial Wave Expansions

The conformal bifields $V(x_1, x_2)$ of dimension $(1, 1)$ which arise in the OPE (2.2) (as sums of integrals of conserved tensor currents) satisfy the d'Alembert equation in each argument [3]; we will call them *harmonic bifields*. Their correlation functions depend on the dimension d of the local scalar fields ϕ. For $d = 1$ one is actually dealing with the theory of a free massless field. We will, therefore, assume $d > 1$. A basis $\{f_{\nu i}, \nu = 0, 1, \ldots, d-2, i = 1, 2\}$ of invariant amplitudes $F(s, t)$ such that

$$\begin{gathered} \langle 0 \mid V_1(x_1, x_2) V_2(x_3, x_4) \mid 0 \rangle = \frac{1}{\rho_{13}\rho_{24}} F(s, t), \\ \rho_{ij} = x_{ij}^2 + i0x_{ij}^0, \quad x^2 = \mathbf{x}^2 - \left(x^0\right)^2 \end{gathered} \tag{3.1}$$

is given by

$$\begin{gathered} (u_+ - u_-) f_{\nu 1}(s, t) = \frac{u_+^{\nu+1}}{(1-u_+)^{\nu+1}} - \frac{u_-^{\nu+1}}{(1-u_-)^{\nu+1}}, \\ (u_+ - u_-) f_{\nu 2}(s, t) = (-1)^{\nu} \left(u_+^{\nu+1} - u_-^{\nu+1}\right), \quad \nu = 0, 1, \ldots, d-2, \end{gathered} \tag{3.2}$$

where $u_\pm$ are the "chiral variables" (2.6)

$$\begin{gathered} f_{01} = \frac{1}{t}, \qquad f_{02} = 1; \qquad f_{11} = \frac{1-s-t}{t^2}, \qquad f_{12} = t - s - 1 \\ f_{21} = \frac{(1-t)^2 - s(2-t) + s^2}{t^3}, \qquad f_{\nu 2}(s, t) = \frac{1}{t} f_{\nu 1}\left(\frac{s}{t}, \frac{1}{t}\right), \end{gathered} \tag{3.3}$$

$f_{\nu,i}$, $i = 1,2$, corresponding to *single pole terms* [5] in the 4-point correlation functions $w_{\nu i}(x_1,\dots,x_4) = f_{\nu i}(s,t)/\rho_{13}\rho_{24}$:

$$
\begin{aligned}
w_{01} &= \frac{1}{\rho_{14}\rho_{23}}, \qquad w_{02} = \frac{1}{\rho_{13}\rho_{24}},\\
w_{11} &= \frac{\rho_{13}\rho_{24}-\rho_{14}\rho_{23}-\rho_{12}\rho_{34}}{\rho_{14}^2\rho_{23}^2}, \qquad w_{12} = \frac{\rho_{14}\rho_{23}-\rho_{13}\rho_{24}-\rho_{12}\rho_{34}}{\rho_{13}^2\rho_{24}^2},\\
w_{21} &= \frac{(\rho_{13}\rho_{24}-\rho_{14}\rho_{23})^2-\rho_{12}\rho_{34}(2\rho_{13}\rho_{24}-\rho_{14}\rho_{23})+\rho_{12}^2\rho_{34}^2}{\rho_{14}^3\rho_{23}^3},\\
w_{22} &= \frac{(\rho_{14}\rho_{23}-\rho_{13}\rho_{24})^2-\rho_{12}\rho_{34}(2\rho_{14}\rho_{23}-\rho_{13}\rho_{24})+\rho_{12}^2\rho_{34}^2}{\rho_{13}^3\rho_{24}^3}.
\end{aligned}
\tag{3.4}
$$

We have $w_{\nu 2} = P_{34}w_{\nu 1}(= P_{12}w_{\nu 1})$ where P_{ij} stands for the substitution of the arguments x_i and x_j. Clearly, for $x_1 = x_2$ (or $s = 0$, $t = 1$) only the amplitudes f_{0i} contribute to the 4-point function (3.1). It has been demonstrated in [4] that the lowest angular momentum (ℓ) contribution to $f_{\nu i}$ corresponds to $\ell = \nu$. The corresponding OPE of the bifield V starts with a local scalar field ϕ of dimension $d = 2$ for $\nu = 0$, with a conserved current j_μ (of $d = 3$) for $\nu = 1$, with the stress energy tensor $T_{\lambda\mu}$ for $\nu = 2$. Indeed, the amplitude $f_{\nu 1}$ admits an expansion in twist two (the twist of a symmetric traceless tensor is defined as the difference between its dimension and its rank. All conserved symmetric tensors in $4D$ have twist two.) *conformal partial waves* $\beta_\ell(s,t)$ [39] starting with (for a derivation see [4, Appendix B])

$$
\beta_\nu(s,t) = \frac{G_{\nu+1}(u_+)-G_{\nu+1}(u_-)}{u_+-u_-}, \qquad G_\mu(u) = u^\mu F(\mu,\mu;2\mu;u). \tag{3.5}
$$

Remark 3.1. Equations (3.2) and (3.5) provide examples of solutions of the d'Alambert equation in any of the arguments x_i, $i = 1,2,3,4$. In fact, the general conformal covariant (of dimension 1 in each argument) such solution has the form of the right-hand side of (3.1) with

$$
F(s,t) = \frac{f(u_+)-f(u_-)}{u_+-u_-}. \tag{3.6}
$$

Remark 3.2. We note that albeit each individual conformal partial wave is a transcendental function (like (3.5)) the sum of all such twist two contributions is the rational function $f_{\nu 1}(s,t)$.

It can be deduced from the analysis of 4-point functions that the commutator algebra of a set of harmonic bifields generated by OPE of scalar fields of dimension d can only close on the V's and the unit operator for $d = 2$. In this case the bifields V are proven, in addition, to be *Huygens bilocal* [5].

Remark 3.3. In general, irreducible positive energy representations of the (connected) conformal group are labeled by triples $(d; j_1, j_2)$ including the dimension d and the Lorentz weight $(j_1, j_2)(2j_i \in \mathbb{N})$, [40]. It turns out that for $d = 3$ there is a spin-tensor bifield of weight $((3/2;1/2,0),(3/2;0,1/2))$ whose commutator algebra does close; for $d=4$ there is a conformal tensor bifield of weight $((2;1,0),(2;0,1))$ with this property. These bifields may be termed *lefthanded*. They are analogues of chiral $2D$ currents; a set of bifields invariant under space reflections would also involve their righthanded counterparts (of weights $((3/2;0,1/2),(3/2;1/2,0))$, and $((2;0,1),(2;1,0))$, resp.).

4. Infinite-Dimensional Lie Algebras and Real Division Rings

Our starting point is the following result of [5].

Proposition 4.1. *The harmonic bilocal fields V arising in the OPEs of a (finite) set of local hermitean scalar fields of dimension $d=2$ can be labeled by the elements M of an unital algebra $\mathcal{M}\subset \mathrm{Mat}(L,\mathbb{R})$ of real matrices closed under transposition, $M\to {}^tM$, in such a way that the following commutation relations (CR) hold:*

$$\begin{aligned}
[V_{M_1}(x_1,x_2),V_{M_2}(x_3,x_4)] = {} & \Delta_{13}V_{{}^tM_1M_2}(x_2,x_4)+\Delta_{24}V_{M_1{}^tM_2}(x_1,x_3)\\
& +\Delta_{23}V_{M_1M_2}(x_1,x_4)+\Delta_{14}V_{M_2M_1}(x_3,x_2)\\
& +\mathrm{tr}(M_1M_2)\Delta_{12,34}+\mathrm{tr}\left({}^tM_1M_2\right)\Delta_{12,43},
\end{aligned} \tag{4.1}$$

here Δ_{ij} is the free field commutator, $\Delta_{ij}:=\Delta^+_{ij}-\Delta^+_{ji}$, and $\Delta_{12,ij}=\Delta^+_{1i}\Delta^+_{2j}-\Delta^+_{i1}\Delta^+_{j2}$, where $\Delta^+_{ij}=\Delta^+(x_i-x_j)$ is the 2-point Wightman function of a free massless scalar field.

We call the set of bilocal fields closed under the CR (4.1) a *Lie system*. The types of Lie systems are determined by the corresponding *t-algebras*, that is, real associative matrix algebras $\mathcal{M}$ closed under transposition. We first observe that each such $\mathcal{M}$ can be equipped with a *Frobenius inner product*

$$\langle M_1,M_2\rangle=\mathrm{tr}\left({}^tM_1M_2\right)=\sum_{ij}(M_1)_{ij}(M_2)_{ij}, \tag{4.2}$$

which is symmetric, positive definite, and has the property $\langle M_1M_2,M_3\rangle=\langle M_1,M_3{}^tM_2\rangle$. This implies that for every right ideal $\mathcal{I}\subset\mathcal{M}$ its orthogonal complement is again a right ideal while its transposed ${}^t\mathcal{I}$ is a left ideal. Therefore, $\mathcal{M}$ is a *semisimple* algebra so that every module over $\mathcal{M}$ is a direct sum of irreducible modules.

Let now $\mathcal{M}$ be irreducible. It then follows from the Schur's lemma (whose real version [41] is richer but less popular than the complex one) that its commutant $\mathcal{M}'$ in $\mathrm{Mat}(L,\mathbb{R})$ coincides with one of the three *real division rings* (or not necessarily commutative *fields*): the fields of real and complex numbers $\mathbb{R}$ and $\mathbb{C}$, and the noncommutative division ring $\mathbb{H}$ of quaternions. In each case the Lie algebra of bilocal fields is a central extension of an infinite-dimensional Lie algebra that admits a discrete series of highest weight representations. Finite dimensional simple Lie groups G with this property have been extensively studied by mathematicians (for a review and references, see [42]); for an extension to the infinite-dimensional case, see [43]. If Z is the centre of G and K is a closed maximal subgroup of G such that K/Z is compact then G is characterized by the property that (G,K) is a *hermitean symmetric pair*. Such groups give rise to *simple space-time symmetries* in the sense of [44] (see also earlier work—in particular by Günaydin—cited there).

It was proven, first in the theory of a single scalar field ϕ (of dimension two) [2], and eventually for an arbitrary set of such fields [5], that the bilocal fields V_M can be written as linear combinations of normal products of free massless scalar fields $\varphi_i(x)$:

$$V_M(x_1,x_2) = \sum_{i,j=1}^{L} M^{ij} : \varphi_i(x_1)\varphi_j(x_2). \tag{4.3}$$

For each of the above types of Lie systems V_M has a canonical form, namely,

$$\begin{aligned}
\mathbb{R} : V(x_1,x_2) &= \sum_{i=1}^{N} : \varphi_i(x_1)\varphi_i(x_2),\\
\mathbb{C} : W(x_1,x_2) &= \sum_{j=1}^{N} : \varphi_j^*(x_1)\varphi_j(x_2),\\
\mathbb{H} : Y(x_1,x_2) &= \sum_{m=1}^{N} : \varphi_m^+(x_1)\varphi_m(x_2),
\end{aligned} \tag{4.4}$$

where φ_i are real, φ_j are complex, and φ_m are quaternionic valued fields (corresponding to (3.2) with $L = N, 2N$, and $4N$, resp.). We will denote the associated infinite-dimensional Lie algebra by $\mathcal{L}(\mathbb{F})$, $\mathbb{F} = \mathbb{R}, \mathbb{C}$, or $\mathbb{H}$.

Remark 4.2. We note that the quaternions (represented by 4×4 real matrices) appear both in the definition of Y—that is, of the matrix algebra $\mathcal{M}$, and of its commutant $\mathcal{M}'$, the two mutually commuting sets of imaginary quaternionic units ℓ_i and r_j corresponding to the splitting of the Lie algebra $so(4)$ of real skew-symmetric 4×4 matrices into a direct sum of "a left and a right" $so(3)$ Lie subalgebras:

$$\begin{gathered}
\ell_1 = \sigma_3 \otimes \epsilon, \qquad \ell_2 = \epsilon \otimes \mathbf{1}, \qquad \ell_3 = \ell_1\ell_2 = \sigma_1 \otimes \epsilon,\\
\left(\ell_j\right)_{\alpha\beta} = \delta_{\alpha 0}\delta_{j\beta} - \delta_{\alpha j}\delta_{0\beta} - \varepsilon_{0j\alpha\beta}, \quad \alpha,\beta = 0,1,2,3, \quad j = 1,2,3,\\
r_1 = \epsilon \otimes \sigma_3, \qquad r_2 = \mathbf{1} \otimes \epsilon, \qquad r_3 = r_1 r_2 = \epsilon \otimes \sigma_1,
\end{gathered} \tag{4.5}$$

where σ_k are the Pauli matrices, $\epsilon = i\sigma_2$, $\varepsilon_{\mu\nu\alpha\beta}$ is the totally antisymmetric Levi-Civita tensor normalized by $\varepsilon_{0123} = 1$. We have

$$\begin{aligned}
Y(x_1,x_2) &= V_0(x_1,x_2)\mathbf{1} + V_1(x_1,x_2)\ell_1 + V_2(x_1,x_2)\ell_2 + V_3(x_1,x_2)\ell_3\\
&= Y(x_2,x_1)^+ \quad \left(\ell_i^+ = -\ell_i, \left[\ell_i, r_j\right] = 0\right),\\
V_\kappa(x_1,x_2) &= \sum_{m=1}^{N} : \varphi_m^\alpha(x_1)(\ell_\kappa)_{\alpha\beta}\varphi_m^\beta(x_2), \quad \ell_0 = \mathbf{1}.
\end{aligned} \tag{4.6}$$

In order to determine the Lie algebra corresponding to the CR (4.1) in each of the three cases (4.5) we choose a discrete basis and specify the topology of the resulting infinite matrix algebra in such a way that the generators of the conformal Lie algebra (most importantly, the conformal Hamiltonian H) belong to it. The basis, say (X_{mn}) where m, n are multi-indices, corresponds to the expansion [34] of a free massless scalar field φ in creation and annihilation operators of fixed energy states

$$\varphi(z) = \sum_{\ell=0}^{\infty} \sum_{\mu=1}^{(\ell+1)^2} \left(\left(z^2\right)^{-\ell-1} \varphi_{\ell+1,\mu} + \varphi_{-\ell-1,\mu} \right) h_{\ell\mu}(z), \tag{4.7}$$

where $(h_{\ell\mu}(z), \mu = 1, \ldots, (\ell+1)^2)$ form a basis of homogeneous harmonic polynomials of degree ℓ in the complex 4-vector z (of the parametrization (2.3) of $\overline{M}$). The generators of the conformal Lie algebra $su(2,2)$ are expressed as infinite sums in X_{mn} with a finite number of diagonals (cf. Appendix B in [6]). The requirement $su(2,2) \subset \mathcal{L}$ thus restricts the topology of $\mathcal{L}$ implying that the last (c-number) term in (4.1) gives rise to a nontrivial central extension of $\mathcal{L}$.

The analysis of [6, 7] yields the following.

Proposition 4.3. *The Lie algebras $\mathcal{L}(\mathbb{F})$, $\mathbb{F} = \mathbb{R}, \mathbb{C}, \mathbb{H}$ are 1-parameter central extensions of appropriate completions of the following inductive limits of matrix algebras:*

$$\begin{aligned} \mathbb{R} &: sp(\infty, \mathbb{R}) = \lim_{n\to\infty} sp(2n, \mathbb{R}), \\ \mathbb{C} &: u(\infty, \infty) = \lim_{n\to\infty} u(n, n), \\ \mathbb{H} &: so^*(4\infty) = \lim_{n\to\infty} so^*(4n). \end{aligned} \tag{4.8}$$

In the free field realization (4.4) *the suitably normalized central charge coincides with the positive integer N.*

5. Fock Space Representation of the Dual Pair $\mathcal{L}(\mathbb{F}) \times U(N, \mathbb{F})$

To summarize the discussion of the last section, there are three infinite-dimensional irreducible Lie algebras, $\mathcal{L}(\mathbb{F})$, that are generated in a theory of GCI scalar fields of dimension $d = 2$ and correspond to the three real division rings $\mathbb{F}$ (Proposition 4.3). For an integer central charge N they admit a free field realization of type (4.3) and a Fock space representation with (compact) gauge group $U(N, \mathbb{F})$:

$$U(N, \mathbb{R}) = O(N), \qquad U(N, \mathbb{C}) = U(N), \qquad U(N, \mathbb{H}) = Sp(2N)\big(= USp(2N)\big). \tag{5.1}$$

It is remarkable that this result holds in general.

Theorem 5.1. (i) *In any unitary irreducible positive energy representation (UIPER) of $\mathcal{L}(\mathbb{F})$ the central charge N is a positive integer.*

(ii) *All UIPERs of $\mathcal{L}(\mathbb{F})$ are realized (with multiplicities) in the Fock space $\mathcal{F}$ of $N \dim_{\mathbb{R}} \mathbb{F}$ free hermitean massless scalar fields.*

(iii) *The ground states of equivalent UIPERs in* $\mathcal{F}$ *form irreducible representations of the gauge group* $U(N, \mathbb{F})$ (5.1). *This establishes a one-to-one correspondence between UIPERs of* $\mathcal{L}(\mathbb{F})$ *occurring in the Fock space and the irreducible representations of* $U(N, \mathbb{F})$.

The *proof* of this theorem for $\mathbb{F} = \mathbb{R}, \mathbb{C}$ is given in [6] (the proof of (i) is already contained in [2]); the proof for $\mathbb{F} = \mathbb{H}$ is given in [7].

Remark 5.2. Theorem 5.1 is also valid—and its proof becomes technically simpler—for a 2-dimensional chiral theory (in which the local fields are functions of a single complex variable). For $\mathbb{F} = \mathbb{C}$ the representation theory of the resulting infinite-dimensional Lie algebra $u(\infty, \infty)$ is then essentially equivalent to that of the vertex algebra $W_{1+\infty}$ studied in [45] (see the introduction in [6] for a more precise comparison).

Theorem 5.1 provides a link between two parallel developments, one in the study of the highest weight modules of reductive Lie groups (and of related dual pairs—see Section 1.1) [42, 43, 46, 47] (and [9, 10]), the other in the work of Doplicher and Roberts [11] and Haag [12] on the theory of (global) gauge groups and superselection sectors—see Section 1.2. (They both originate—in the paper of Irving Segal and Rudolf Haag, resp.—at the same Lille 1957 conference on mathematical problems in quantum field theory.) Albeit the settings are not equivalent the results match. The observable algebra (in our case, the commutator algebra generated by the set of bilocal fields V_M) determines the (compact) gauge group and the structure of the superselection sectors of the theory. (For a more careful comparison between the two approaches, see [6, Sections 1 and 4].)

The infinite-dimensional Lie algebra $\mathcal{L}(\mathbb{F})$ and the compact gauge group $U(N, \mathbb{F})$ appear as a rather special (limit-) case of a *dual pair* in the sense of Howe [9, 10]. It would be interesting to explore whether other (inequivalent) pairs would appear in the study of commutator algebras of (spin)tensor bifields (discussed in Remark 3.3) and of their supersymmetric extension (e.g., a limit as $m, n \to \infty$ of the series of Lie superalgebras $osp(4m^* \mid 2n)$ studied in [48]).

Acknowledgments

The author would like to thank his coauthors Bojko Bakalov, Nikolay M. Nikolov, and Karl-Henning Rehren. All results (reported in Sections 3–5) of this paper have been obtained in collaboration with them. He also acknowledges a partial support from the Bulgarian National Council for Scientific Research under Contracts Ph-1406 and DO-02-257.

References

[1] N. M. Nikolov and I. T. Todorov, "Rationality of conformally invariant local correlation functions on compactified Minkowski space," *Communications in Mathematical Physics*, vol. 218, no. 2, pp. 417–436, 2001.

[2] N. M. Nikolov, Ya. S. Stanev, and I. T. Todorov, "Four-dimensional conformal field theory models with rational correlation functions," *Journal of Physics A*, vol. 35, no. 12, pp. 2985–3007, 2002.

[3] N. M. Nikolov, Ya. S. Stanev, and I. T. Todorov, "Globally conformal invariant gauge field theory with rational correlation functions," *Nuclear Physics B*, vol. 670, no. 3, pp. 373–400, 2003.

[4] N. M. Nikolov, K.-H. Rehren, and I. T. Todorov, "Partial wave expansion and Wightman positivity in conformal field theory," *Nuclear Physics B*, vol. 722, no. 3, pp. 266–296, 2005.

[5] N. M. Nikolov, K.-H. Rehren, and I. Todorov, “Harmonic bilocal fields generated by globally conformal invariant scalar fields,” *Communications in Mathematical Physics*, vol. 279, no. 1, pp. 225–250, 2008.

[6] B. Bakalov, N. M. Nikolov, K.-H. Rehren, and I. Todorov, “Unitary positive-energy representations of scalar bilocal quantum fields,” *Communications in Mathematical Physics*, vol. 271, no. 1, pp. 223–246, 2007.

[7] B. Bakalov, N. M. Nikolov, K.-H. Rehren, and I. Todorov, “Infinite-dimensional Lie algebras in 4D conformal quantum field theory,” *Journal of Physics A*, vol. 41, no. 19, Article ID 194002, 12 pages, 2008.

[8] R. F. Streater and A. S. Wightman, *PCT, Spin and Statistics, and All That*, Princeton Landmarks in Physics, Princeton University Press, Princeton, NJ, USA, 2000.

[9] R. Howe, “Dual pairs in physics: harmonic oscillators, photons, electrons, and singletons,” in *Applications of Group Theory in Physics and Mathematical Physics*, M. Flato, P. Sally, and G. Zuckerman, Eds., vol. 21 of *Lectures in Applied Mathematics*, pp. 179–207, American Mathematical Society, Providence, RI, USA, 1985.

[10] R. Howe, “Transcending classical invariant theory,” *Journal of the American Mathematical Society*, vol. 2, no. 3, pp. 535–552, 1989.

[11] S. Doplicher and J. E. Roberts, “Why there is a field algebra with a compact gauge group describing the superselection structure in particle physics,” *Communications in Mathematical Physics*, vol. 131, no. 1, pp. 51–107, 1990.

[12] R. Haag, *Local Quantum Physics: Fields, Particles, Algebras*, Texts and Monographs in Physics, Springer, Berlin, Germany, 1992.

[13] R. Howe, “On the role of the Heisenberg group in harmonic analysis,” *Bulletin of the American Mathematical Society*, vol. 3, no. 2, pp. 821–843, 1980.

[14] A. Weil, “Sur certains groupes d’opérateurs unitaires,” *Acta Mathematica*, vol. 111, no. 1, pp. 143–211, 1964.

[15] A. Weil, “Sur la formule de Siegel dans la théorie des groupes classiques,” *Acta Mathematica*, vol. 113, no. 1, pp. 1–87, 1965.

[16] G. Mack and I. T. Todorov, “Irreducibility of the ladder representations of $U(2,2)$ when restricted to the Poincaré subgroup,” *Journal of Mathematical Physics*, vol. 10, pp. 2078–2085, 1969.

[17] A. Joseph, “Minimal realizations and spectrum generating algebras,” *Communications in Mathematical Physics*, vol. 36, pp. 325–338, 1974.

[18] A. Joseph, “The minimal orbit in a simple Lie algebra and its associated maximal ideal,” *Annales Scientifiques de l’École Normale Supérieure Série 4*, vol. 9, no. 1, pp. 1–29, 1976.

[19] D. Kazhdan, B. Pioline, and A. Waldron, “Minimal representations, spherical vectors and exceptional theta series,” *Communications in Mathematical Physics*, vol. 226, no. 1, pp. 1–40, 2002.

[20] M. Günaydin and O. Pavlyk, “A unified approach to the minimal unitary realizations of noncompact groups and supergroups,” *Journal of High Energy Physics*, vol. 2006, no. 9, article 050, 2006.

[21] T. Kobayashi and G. Mano, “The Schrödinger model for the minimal representation of the indefinite orthogonal group $O(p,q)$,” to appear in *Memoirs of the American Mathematical Society*.

[22] G. C. Wick, A. S. Wightman, and E. P. Wigner, “The intrinsic parity of elementary particles,” *Physical Review*, vol. 88, pp. 101–105, 1952.

[23] V. G. Kac, *Infinite-Dimensional Lie Algebras*, Cambridge University Press, Cambridge, UK, 3rd edition, 1990.

[24] V. G. Kac and A. K. Raina, *Bombay Lectures on Highest Weight Representations of Infinite-Dimensional Lie Algebras*, vol. 2 of *Advanced Series in Mathematical Physics*, World Scientific, Teaneck, NJ, USA, 1987.

[25] D. W. Robinson, “On a soluble model of relativistic field theory,” *Physics Letters B*, vol. 9, pp. 189–190, 1964.

[26] J. H. Lowenstein, “The existence of scalar Lie fields,” *Communications in Mathematical Physics*, vol. 6, pp. 49–60, 1967.

[27] K. Baumann, “There are no scalar Lie fields in three or more dimensional space-time,” *Communications in Mathematical Physics*, vol. 47, no. 1, pp. 69–74, 1976.

[28] I. B. Frenkel and V. G. Kac, “Basic representations of affine Lie algebras and dual resonance models,” *Inventiones Mathematicae*, vol. 62, no. 1, pp. 23–66, 1980.

[29] R. E. Borcherds, "Vertex algebras, Kac-Moody algebras, and the Monster," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 83, no. 10, pp. 3068–3071, 1986.

[30] V. Kac, *Vertex Algebras for Beginners*, vol. 10 of *University Lecture Series*, American Mathematical Society, Providence, RI, USA, 2nd edition, 1998.

[31] E. Frenkel and D. Ben-Zvi, *Vertex Algebras and Algebraic Curves*, vol. 88 of *Mathematical Surveys and Monographs*, American Mathematical Society, Providence, RI, USA, 2nd edition, 2004.

[32] Y. Zhu, "Modular invariance of characters of vertex operator algebras," *Journal of the American Mathematical Society*, vol. 9, no. 1, pp. 237–302, 1996.

[33] A. Uhlmann, "The closure of Minkowski space," *Acta Physica Polonica*, vol. 24, pp. 295–296, 1963.

[34] I. T. Todorov, "Infinite-dimensional Lie algebras in conformal QFT models," in *Conformal Groups and Related Symmetries: Physical Results and Mathematical Background*, A. O. Barut and H.-D. Doebner, Eds., vol. 261 of *Lecture Notes in Physics*, pp. 387–443, Springer, Berlin, Germany, 1986.

[35] N. M. Nikolov, "Vertex algebras in higher dimensions and globally conformal invariant quantum field theory," *Communications in Mathematical Physics*, vol. 253, no. 2, pp. 283–322, 2005.

[36] N. M. Nikolov and I. T. Todorov, "Elliptic thermal correlation functions and modular forms in a globally conformal invariant QFT," *Reviews in Mathematical Physics*, vol. 17, no. 6, pp. 613–667, 2005.

[37] F. A. Dolan and H. Osborn, "Conformal four point functions and the operator product expansion," *Nuclear Physics B*, vol. 599, no. 1-2, pp. 459–496, 2001.

[38] B. Bakalov and N. M. Nikolov, "Jacobi identity for vertex algebras in higher dimensions," *Journal of Mathematical Physics*, vol. 47, no. 5, Article ID 053505, 30 pages, 2006.

[39] V. K. Dobrev, G. Mack, V. B. Petkova, S. G. Petrova, and I. T. Todorov, *Harmonic Analysis on the n-Dimensional Lorentz Group and Its Application to Conformal Quantum Field Theory*, vol. 63 of *Lecture Notes in Physics*, Springer, Berlin, Germany, 1977.

[40] G. Mack, "All unitary ray representations of the conformal group $SU(2,2)$ with positive energy," *Communications in Mathematical Physics*, vol. 55, no. 1, pp. 1–28, 1977.

[41] S. Lang, *Algebra*, vol. 211 of *Graduate Texts in Mathematics*, Springer, New York, NY, USA, 3rd edition, 2002.

[42] T. Enright, R. Howe, and N. Wallach, "A classification of unitary highest weight modules," in *Representation Theory of Reductive Groups*, vol. 40 of *Progress in Mathematics*, pp. 97–143, Birkhäuser, Boston, Mass, USA, 1983.

[43] M. U. Schmidt, "Lowest weight representations of some infinite-dimensional groups on Fock spaces," *Acta Applicandae Mathematicae*, vol. 18, no. 1, pp. 59–84, 1990.

[44] G. Mack and M. de Riese, "Simple space-time symmetries: generalizing conformal field theory," *Journal of Mathematical Physics*, vol. 48, no. 5, Article ID 052304, 21 pages, 2007.

[45] V. Kac and A. Radul, "Representation theory of the vertex algebra $W_{1+\infty}$," *Transformation Groups*, vol. 1, no. 1-2, pp. 41–70, 1996.

[46] M. Kashiwara and M. Vergne, "On the Segal-Shale-Weil representations and harmonic polynomials," *Inventiones Mathematicae*, vol. 44, no. 1, pp. 1–47, 1978.

[47] H. P. Jakobsen, "The last possible place of unitarity for certain highest weight modules," *Mathematische Annalen*, vol. 256, no. 4, pp. 439–447, 1981.

[48] M. Günaydin and R. J. Scalise, "Unitary lowest weight representations of the noncompact supergroup $OSp(2m^*|2n)$," *Journal of Mathematical Physics*, vol. 32, no. 3, pp. 599–606, 1991.

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 956128, 9 pages
doi:10.1155/2010/956128

Research Article

Arnold's Projective Plane and r-Matrices

K. Uchino

Tokyo University of Science, Wakamiya, 26 Shinjyuku, Tokyo, Japan

Correspondence should be addressed to K. Uchino, k_uchino@oct.rikadai.jp

Received 19 December 2009; Accepted 27 February 2010

Academic Editor: A. Yoshioka

We will explain Arnold's 2-dimensional (shortly, 2D) projective geometry (Arnold, 2005) by means of lattice theory. It will be shown that the projection of the set of nontrivial triangular r-matrices is the pencil of tangent lines of a quadratic curve on Arnold's projective plane.

1. Introduction

We briefly describe Arnold's projective geometry [1]. We recall the ordinary projective plane $\mathbb{P}(\mathbb{R}^3)$ over the real field. We replace $\mathbb{R}^3$ to the set of the quadratic functions $\widehat{Q} := ap^2 + 2bpq + cq^2$ on the canonical symplectic plane $(\mathbb{R}^2 : p, q)$. The quadratic functions form into a Lie subalgebra of the canonical Poisson algebra on the symplectic plane. This Lie algebra is isomorphic to $sl(2,\mathbb{R})$; that is, Arnold introduced a projective plane $\mathbb{P}(sl(2,\mathbb{R}))$. By the projection, a (nontrivial) quadratic function $\widehat{Q}$ corresponds to a point Q on the projective plane, and the Killing form on $sl(2,\mathbb{R})$ corresponds to a quadratic curve on the projective plane, because it is a symmetric bilinear form. Since the Killing form is nondegenerate, the associated curve defines a duality (so-called polar system) between the projective lines and the projective points. Arnold showed that the Poisson bracket $\{\widehat{Q}_1, \widehat{Q}_2\}$ corresponds to the pole point of the projective line through Q_1 and Q_2. As an application, it was shown that given a good triangle composed of three points (Q_1, Q_2, Q_3), the three altitudes intersect the same point (altitude theorem). Interestingly, the altitude theorem is shown by the Jacobi identity (see Figure 1).

The monomials of double brackets $\{\{\widehat{Q}_i, \widehat{Q}_j\}, \widehat{Q}_k\}$ correspond to the three bold lines, via the line-point duality. By the Jacobi identity, the three monomials are linearly dependent. This implies that the three lines intersect the same point.

We suppose that 2D projective geometry is encoded in the Lie algebra. However some information on the Lie algebra is lost in the process of constructing projective geometry;

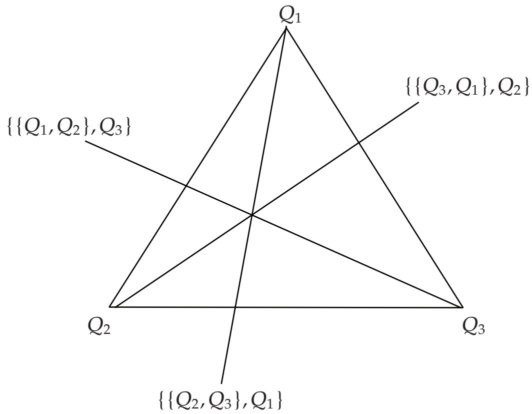


Figure 1

besides, it is not clear why it has to happen, conceptually. So we will reformulate the Arnold construction by means of lattice theory. Since the lattice is an algebra, the problem becomes more clear. We will prove, when the characteristic of the ground field is not 2, that each 3D simple Lie algebra admits a modular lattice structure. This proposition explains why 2D projective geometry can be encoded in $sl(2)$.

It is crucial to consider the *Plücker* embedding for the algebraic Arnold construction. When $\mathfrak{g}$ is 3D and simple, the Lie algebra multiplication $\mu : \mathfrak{g} \wedge \mathfrak{g} \to \mathfrak{g}$, $x \wedge y \mapsto [x, y]$ is an isomorphism, which induces an isomorphism $\mathbb{P}(\mathfrak{g} \wedge \mathfrak{g}) \cong \mathbb{P}(\mathfrak{g})$. We will show that the line-point duality on $\mathbb{P}(\mathfrak{g})$ is equivalent with the isomorphism $\mathbb{P}(\mathfrak{g} \wedge \mathfrak{g}) \cong \mathbb{P}(\mathfrak{g})$, up to the *Plücker* embedding.

As an application we study a projective geometry of triangular r-matrices, when $\mathfrak{g} = sl(2)$. We will prove that the projection of the set of nontrivial triangular r-matrices is equivalent to the pencil of tangent lines on the quadratic curve made from the Killing form. This proposition is equivalently translated as follows: the classical Yang-Baxter equation $\{r, r\} = 0$ on $sl(2)$ is equivalent to the quadratic curve on the projective plane.

2. Algebraic Arnold Construction

2.1. Subspace Lattices

A lattice is, by definition, a set equipped with two commutative associative multiplications, $\smile$ and $\frown$, satisfying the following two identities:

$$\begin{aligned} x \smile (x \frown y) &= x, \\ x \frown (x \smile y) &= x. \end{aligned} \tag{2.1}$$

A lattice has a canonical order $\leq$ defined by

$$x \leq y \Longleftrightarrow x = x \frown y \quad (\Longleftrightarrow y = x \smile y). \tag{2.2}$$

A lattice is called a modular lattice when it satisfies the inequality (modular rule)

$$x \leq z \Longrightarrow x \smile (y \frown z) = (x \smile y) \frown z \tag{2.3}$$

The notion of lattice morphism is defined by the usual manner.

Example 2.1 (subspace lattices). Let V be a vector space. Consider the set of subspaces of V: $\mathrm{Latt}(V) := \{S \mid S \subset V\}$. Define two natural multiplications (cup- and cap-products) on $\mathrm{Latt}(V)$ by

$$\begin{aligned} S_1 \smile S_2 &:= S_1 + S_2, \\ S_1 \frown S_2 &:= S_1 \cap S_2, \end{aligned} \tag{2.4}$$

where $S_1, S_2 \in \mathrm{Latt}(V)$. Then $\mathrm{Latt}(V)$ becomes a modular lattice. The induced order is the natural inclusion relation $S_1 \subset S_2$. Given a linear injection $f : V_1 \to V_2$, an associated lattice morphism $\mathrm{Latt}(f)$ is naturally defined by $\mathrm{Latt}(f)(S) := f(S)$.

The subspace lattices are *complementary*; that is, the zero space $\mathbf{0}$ is the unit element with respect to $\smile$ and the total space V is $\frown$. If V is split for each S, then there exists a cosubspace $\overline{S}$ satisfying $\overline{S} \smile S = V$ and $\overline{S} \frown S = \mathbf{0}$. The subspace $\overline{S}$ (resp., S) is called a complement[1] of S (resp., $\overline{S}$). Such a lattice is called a *complemented lattice*. If V is finite dimensional, then the subspace lattice is a *complemented-modular-lattice*. A projective geometry is axiomatically defined as a complemented-modular-lattice satisfying some additional properties.

Definition 2.2. When V is $(n+1)$-dimensional, the subspace lattice $\mathrm{Latt}(V)$ is called an n-dimensional projective geometry over V.

The 1D subspaces are regarded as projective points, 2D subspaces are projective lines and so on. The zero space $\mathbf{0}$ is regarded as the empty set. For instance, given two 2D subspaces S_1 and S_2, the intersection $S_1 \frown S_2 (= S_1 \cap S_2)$ is the common point of two projective lines (if it exists).

2.2. Lie Algebra Construction of Projective Plane

Let $(\mathfrak{g}, [-,-])$ be a 3D $\mathbb{K}$-Lie algebra with a nondegenerate symmetric invariant pairing $(-,-)$, where $\mathbb{K}$ is the ground field $\mathrm{char}(\mathbb{K}) \neq 2$. We assume that $[\mathfrak{g}, \mathfrak{g}] = \mathfrak{g}$, or equivalently, the Lie bracket is an isomorphism from $\mathfrak{g} \wedge \mathfrak{g}$ to $\mathfrak{g}$. This assumption is needed in order to construct projective geometry. Such a Lie algebra is simple. In particular, when $\mathbb{K}$ is a closed field, $\mathfrak{g}$ is isomorphic to $sl(2, \mathbb{K})$.

We denote by $p(x)$ the 1D subspace generated by $x \in \mathfrak{g}$ and denote by $l(x, y)$ the 2D subspace generated by $x, y \in \mathfrak{g}$. We define new cup-and cap-products $\smile'$ and $\frown'$ on $\mathrm{Latt}(\mathfrak{g})$.

Definition 2.3 (Arnold products). (i) $p(x) \smile' p(y) := p^{\perp}([x, y]), p(x) \neq p(y)$, where $\perp$ means the orthogonal space with respect to the invariant pairing $(-,-)$.

(ii) $l(x, y) \frown' l(z, w) := p([[x, y], [z, w]]), l(x, y) \neq l(z, w)$.

(iii) $S_1 \smile' S_2 := S_1 \smile S_2$ and $S_1 \frown' S_2 := S_1 \frown S_2$, in all other cases.

Proposition 2.4. *The set* $\mathrm{Latt}(\mathfrak{g})$ *with Arnold products becomes a lattice and it is the same as the classical subspace lattice.*

Proof. By the invariancy of the pairing, we have $([x,y],x) = ([x,y],y) = 0$. This gives

$$p^{\perp}([x,y]) = p(x) + p(y) = l(x,y). \tag{2.5}$$

Hence we have $p(x) \smile' p(y) = p(x) \smile p(y)$. We prove that $l(x,y) \cap l(z,w) = p([[x,y],[z,w]])$. Since $\perp\perp = id$, we have $l^{\perp}(x,y) = p([x,y])$, $l^{\perp}(z,w) = p([z,w])$, and $p^{\perp}([[x,y],[z,w]]) = l([x,y],[z,w])$. Thus we obtain

$$\begin{aligned} p^{\perp}([[x,y],[z,w]]) &= l([x,y],[z,w]) \\ &= p([x,y]) + p([z,w]) \\ &= l^{\perp}(x,y) + l^{\perp}(z,w) \\ &= (l(x,y) \cap l(z,w))^{\perp}, \end{aligned} \tag{2.6}$$

which gives $l(x,y) \frown' l(z,w) = l(x,y) \frown l(z,w)$. □

In the following, we omit the "prime" on the Arnold products.

Remark 2.5 (Lie algebra identities versus lattice identities). We put $l([x,y]) := l(x,y)$. Then the Arnold products are coherent with the Lie bracket, namely,

$$\begin{aligned} p(x) \smile p(y) &= l([x,y]), \\ l(x) \frown l(y) &= p([x,y]). \end{aligned} \tag{2.7}$$

Tomihisa [2] discovered an interesting identity on $sl(2,\mathbb{R})$:

$$[[[x_1,y],[x_2,z]],x_3] + \text{cyclic permutation w.r.t. } (1,2,3) = 0, \tag{2.8}$$

where y, z are fixed. The Tomihisa identity induces a lattice identity

$$\begin{aligned} &((l_1 \frown l_y) \smile (l_2 \frown l_z)) \frown l_3 \\ &\quad \le (((l_2 \frown l_y) \smile (l_3 \frown l_z)) \frown l_1) \smile (((l_3 \frown l_y) \smile (l_1 \frown l_z)) \frown l_2), \end{aligned} \tag{2.9}$$

where $l_i := l(x_i)$, $l_y := l(y)$, and $l_z := l(z)$. In a study by Aicardi in [3], it was shown that the projection of Tomihisa identity is equivalent to the Pappus theorem. (The author also proved this proposition around winter 2007.) We leave it to the reader to write down the lattice identity associated with the Jacobi identity.

Given a coordinate (ξ_1, ξ_2, ξ_3) on the projective plane, the symmetric pairing is regarded as a defining equation of a nondegenerate quadratic curve (so-called polar system):

$$(-,-) = \sum a_{ij}\xi_i\xi_j. \tag{2.10}$$

We give an example of the quadratic curve made from the pairing.

Example 2.6 (see, [1, 2]). We assume that $\mathfrak{g} := sl(2,\mathbb{R})$ generated by the standard basis (X, Y, H) satisfying the following relations:

$$[H, X] = 2X, \qquad [X, Y] = H, \qquad [H, Y] = -2Y. \tag{2.11}$$

The symmetric pairing is equivalent with the Killing form. We define the scale of the form by

$$(X, Y) := \frac{1}{2}, \qquad (H, H) := 1, \tag{2.12}$$

and all others zero. We set a point $Q := \xi_1 Y + \xi_2 H + \xi_3 X$, $\xi_1, \xi_2, \xi_3 \in \mathbb{R}$. It is on the quadratic curve, that is, $(Q, Q) = 0$ if and only if $\xi_1\xi_3 - (\xi_2)^2 = 0$, and this condition is equivalent with $(x_0)^2 = (x_1)^2 + (x_2)^2$, via the coordinate transformation

$$x_0 + x_1 = \xi_1, \qquad x_2 = \xi_2, \qquad x_0 - x_1 = \xi_3. \tag{2.13}$$

Hence the quadratic curve is regarded as a circle on the projective plane with coordinate $[1 : x_1 : x_2]$.

Remark 2.7. The pairing induces a metric $(- + +)$ on $(\mathbb{R}^3 : x_0, x_1, x_2)$. It is well known that the inside of the circle (so-called timelike subspace) is a hyperbolic plane. The altitude theorem is strictly a theorem on the hyperbolic plane because a metric is needed to define the notion of altitude.

Given a point $p = (p_1, p_2, p_3)$, a line is defined by

$$(p, -) = \sum a_{ij} p_i \xi_j = 0. \tag{2.14}$$

This line is called the *polar line* of the point; conversely the point is called the *pole* of the line. Namely, the orthogonal space $p := l^{\perp}$ is the pole of the line l.

Corollary 2.8 (see [1]). *Given a line $l(x, y)$, $p([x, y])$ is the pole of the line.*

Figure 2 is depicting the duality defined by an ellipse.

The line l is the polar line of the point p which is inside an ellipse. The line and point are connected by the tangent lines and chords of the ellipse. This figure has been drawn in a book by Kawada in [4].

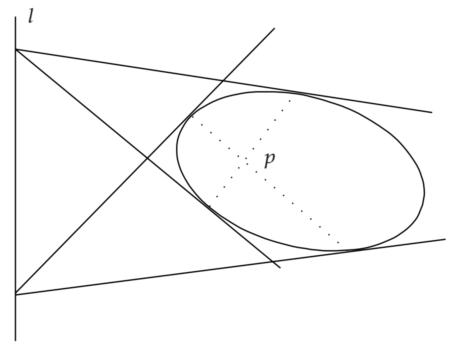


Figure 2

We did not use Jacobi identity in this section. So we will discuss the Jacobi identity on the Lie algebra in the next section.

2.3. Jacobi Identity

Proposition 2.9. *Let V be a 3D $\mathbb{K}$-vector space equipped with a skew-symmetric bracket product $[-,-]$ and a nondegenerate symmetric invariant pairing $(-,-)$. Then the bracket satisfies the Jacobi identity.*

Proof. When $x, y, z \in V$ are linearly dependent, Jacobi identity holds. So we assume that x, y, z is linear independent.

Case 1. By the invariancy of the pairing, we have

$$\left([[x,y],z] + [y,[x,z]],x\right) = \left([x,y],[z,x]\right) - \left([x,z],[y,x]\right) = 0. \tag{2.15}$$

We assume that $[x,[y,z]] = 0$. Then we obtain

$$\begin{aligned}\left([[x,y],z] + [y,[x,z]],y\right) &= \left([x,y],[z,y]\right) = \left(y,[x,[y,z]]\right) = 0,\\ \left([[x,y],z] + [y,[x,z]],z\right) &= -\left([x,z],[y,z]\right) = \left(z,[x,[y,z]]\right) = 0.\end{aligned} \tag{2.16}$$

Since the pairing is nondegenerate, we have $[[x,y],z] + [y,[x,z]] = 0$, which gives the Jacobi identity $[x,[y,z]] = [[x,y],z] + [y,[x,z]](= 0)$.

Case 2. Assume that $[x,[y,z]] \neq 0$. We already saw that $([[x,y],z] + [y,[x,z]],x) = 0$. We show that $([[x,y],z]+[y,[x,z]],[y,z]) = 0$. Since $\dim(V) = 3$, one can put $[y,z] = ax+by+cz$, for some $a, b, c \in \mathbb{K}$. Then we have

$$\begin{aligned}
([[x,y],z] + [y,[x,z]], [y,z]) &= ([[x,y],z] + [y,[x,z]], ax + by + cz) \\
&= ([[x,y],z] + [y,[x,z]], by + cz) \\
&= ([[x,y],z], by) + ([y,[x,z]], cz) \\
&= b([x,y],[z,y]) - c([x,z],[y,z]) \\
&= b([x,y], -cz) - c([x,z], by) \\
&= -bc(x,[y,z]) - cb(x,[z,y]) = 0.
\end{aligned} \tag{2.17}$$

Therefore $[[x,y],z]+[y,[x,z]]$ is orthogonal with the independent two elements x and $[y,z]$. (If x and $[y,z]$ are linearly dependent, then $[x,[y,z]] = 0$.) The monomial $[x,[y,z]]$ is also orthogonal with x and $[y,z]$. Thus we obtain

$$[x,[y,z]] = \lambda([[x,y],z] + [y,[x,z]]), \quad \lambda(\neq 0) \in \mathbb{K}. \tag{2.18}$$

We can assume that $([x,[y,z]],y) \neq 0$ or $([x,[y,z]],z) \neq 0$, because the pairing is nondegenerate. We assume that $([x,[y,z]],y) \neq 0$ without loss of generality. Then we obtain $\lambda = 1$, because

$$([x,[y,z]],y) = -([y,z],[x,y]) = \lambda([[x,y],z],y) = \lambda([x,y],[z,y]). \tag{2.19}$$

□

The proposition above indicates that the Jacobi identity is a priori invested in the projective plane.

2.4. Duality Principle

Let μ be the Lie algebra structure on $\mathfrak{g}$; that is, $\mu : \mathfrak{g} \wedge \mathfrak{g} \to \mathfrak{g}$, $x \wedge y \mapsto [x,y]$. Since μ is an isomorphism, it induces a lattice isomorphism

$$\mathrm{Latt}(\mu) : \mathrm{Latt}(\mathfrak{g} \wedge \mathfrak{g}) \longrightarrow \mathrm{Latt}(\mathfrak{g}). \tag{2.20}$$

Let Line $:= \{p(x) \smile p(y)\}$ be the set of all lines in $\mathrm{Latt}(\mathfrak{g})$. One can define an injection pl : Line $\to \mathrm{Latt}(\mathfrak{g} \wedge \mathfrak{g})$ as

$$pl : p(x) \smile p(y) \longmapsto p(x \wedge y). \tag{2.21}$$

This mapping is called a *Plücker* embedding.

Proposition 2.10. *The diagram below is commutative:*

$$\begin{tikzcd} \text{Line} \arrow[r, "\perp"] \arrow[d, "pl"'] & \text{Point} \arrow[d, "i"] \\ \text{Latt}(\mathfrak{g} \wedge \mathfrak{g}) \arrow[r, "\text{Latt}(\mu)"] & \text{Latt}(\mathfrak{g}) \end{tikzcd} \tag{2.22}$$

where Point is the set of points on the projective plane and $\perp$ is the duality correspondence.

Namely, the Lie algebra multiplication is equivalent to the duality principle.

3. r-Matrices

In this section, we assume that $\mathfrak{g} = sl(2, \mathbb{K})$, $\mathbb{Q} \subset \mathbb{K}$. We consider a graded commutative algebra

$$\bigwedge \mathfrak{g} := \bigwedge^3 \mathfrak{g} \oplus \bigwedge^2 \mathfrak{g} \oplus \mathfrak{g}. \tag{3.1}$$

A graded Poisson bracket $\{-,-\}$ with degree -1 is uniquely defined on $\bigwedge\mathfrak{g}$ by the axioms of graded Poisson algebra and the condition

$$\{A, B\} := [A, B], \quad \text{if } A, B \in \mathfrak{g}. \tag{3.2}$$

This bracket is called a Schouten-Nijenhuis bracket. Let r be a 2 tensor in $\bigwedge^2\mathfrak{g}$. The Maurer-Cartan (MC) equation $\{r, r\} = 0$ is called a classical Yang-Baxter equation and the solution is called a classical *triangular* r-matrix. For instance, $r := X \wedge H$ is a solution of the MC-equation.

Remark 3.1. When $\mathfrak{g} = o(3, \mathbb{R})$, there is no nontrivial triangular r-matrix.

Proposition 3.2. *The set of points $p(r)$ associated with nontrivial triangular r-matrices*

$$\mathbb{P}(r) := \{p(r) \in \text{Latt}(\mathfrak{g} \wedge \mathfrak{g}) \mid \{r, r\} = 0,\ r \neq 0\} \tag{3.3}$$

bijectively corresponds to the pencil of tangent lines of the quadratic curve made from the symmetric pairing, or, equivalently, $\mathbb{P}(r)$ bijectively corresponds to the quadratic curve.

Proof. Since $\dim \mathfrak{g} = 3$, one can write $r = r_1 \wedge r_2$ for some $r_1, r_2 \in \mathfrak{g}$. By the biderivation property of the Schouten-Nijenhuis bracket, we have

$$\{r, r\} = (\pm) r_1 \wedge [r_1, r_2] \wedge r_2. \tag{3.4}$$

Hence $\{r, r\} = 0$ if and only if $[r_1, r_2]$ is linearly dependent on r_1 and r_2.

Lemma 3.3. *The set of lines $l(x, y)$ including the pole $p([x, y])$ corresponds to $\mathbb{P}(r)$, via the Plücker embedding.*

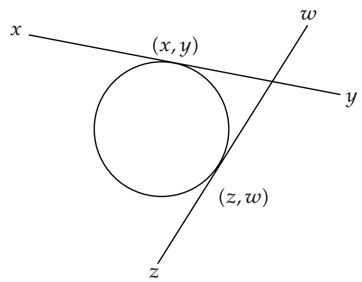


Figure 3

If $p([x, y]) \leq l(x, y)$, then the pairing $([x, y], [x, y])$ vanishes, because $([x, y], [x, y]) = ([x, y], \lambda x + \mu y) = 0$. Hence $p([x, y])$ is on the quadratic curve. It is easy to check that $l(x, y)$ is not cross over the curve. Hence the line is tangent to the curve at $p([x, y])$.

Lemma 3.4. *A line is tangent to the quadratic curve if and only if the pole is on the line, and the tangent point is the pole of the line (see Figure 3).*

Therefore $\mathbb{P}(r)$ is identified with the pencil of tangent lines of the quadratic curve. The proof of the proposition is completed. □

Acknowledgment

The author would like to express his thanks to Professor Akira Yoshioka and Dr. Toshio Tomihisa for their helpful comments.

Endnotes

1. The complement is not unique in general.

References

[1] V. G. Arnold, "Lobachevsky triangle altitudes theorem as the Jacobi identity in the Lie algebra of quadratic forms on symplectic plane," *Journal of Geometry and Physics*, vol. 53, no. 4, pp. 421–427, 2005.

[2] T. Tomihisa, "Geometry of projective plane and Poisson structure," *Journal of Geometry and Physics*, vol. 59, no. 5, pp. 673–684, 2009.

[3] F. Aicardi, "Projective geometry from Poisson algebras," preprint.

[4] Y. Kawada, *Affine Geometry, Projective Geometry*, Iwanami kouza kiso suugaku, Iwanami Syoten, Tokyo, Japan, 2nd edition, 1976.

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 280362, 35 pages
doi:10.1155/2010/280362

Review Article

Infinite-Dimensional Lie Groups and Algebras in Mathematical Physics

Rudolf Schmid

Department of Mathematics, Emory University, Atlanta, GA 30032, USA

Correspondence should be addressed to Rudolf Schmid, rudolf.schmid@emory.edu

Received 8 March 2010; Accepted 16 May 2010

Academic Editor: G. A. Goldin

We give a review of infinite-dimensional Lie groups and algebras and show some applications and examples in mathematical physics. This includes diffeomorphism groups and their natural subgroups like volume-preserving and symplectic transformations, as well as gauge groups and loop groups. Applications include fluid dynamics, Maxwell's equations, and plasma physics. We discuss applications in quantum field theory and relativity (gravity) including BRST and supersymmetries.

1. Introduction

Lie groups play an important role in physical systems both as phase spaces and as symmetry groups. Infinite-dimensional Lie groups occur in the study of dynamical systems with an infinite number of degrees of freedom such as PDEs and in field theories. For such infinite-dimensional dynamical systems, diffeomorphism groups and various extensions and variations thereof, such as gauge groups, loop groups, and groups of Fourier integral operators, occur as symmetry groups and phase spaces. Symmetries are fundamental for Hamiltonian systems. They provide conservation laws (Noether currents) and reduce the number of degrees of freedom, that is, the dimension of the phase space.

The topics selected for review aim to illustrate some of the ways infinite-dimensional geometry and global analysis can be used in mathematical problems of physical interest. The topics selected are the following.

(1) Infinite-Dimensional Lie Groups.

(2) Lie Groups as Symmetry Groups of Hamiltonian Systems.

(3) Applications.

(4) Gauge Theories, the Standard Model, and Gravity.

(5) SUSY (supersymmetry).

2. Infinite-Dimensional Lie Groups

2.1. Basic Definitions

A general theory of infinite-dimensional Lie groups is hardly developed. Even Bourbaki [1] only develops a theory of infinite-dimensional manifolds, but all of the important theorems about Lie groups are stated for finite-dimensional ones.

An infinite-dimensional *Lie group* $\mathcal{G}$ is a group and an infinite-dimensional manifold with smooth group operations

$$m : \mathcal{G} \times \mathcal{G} \longrightarrow \mathcal{G}, \quad m(g,h) = g \cdot h, \quad C^{\infty}, \tag{2.1}$$

$$i : \mathcal{G} \longrightarrow \mathcal{G}, \quad i(g) = g^{-1}, \quad C^{\infty}. \tag{2.2}$$

Such a Lie group $\mathcal{G}$ is locally diffeomorphic to an infinite-dimensional vector space. This can be a Banach space whose topology is given by a norm $\|\cdot\|$, a Hilbert space whose topology is given by an inner product $\langle\cdot,\cdot\rangle$, or a Frechet space whose topology is given by a metric but not by a norm. Depending on the choice of the topology on $\mathcal{G}$, we talk about Banach, Hilbert, or Frechet Lie groups, respectively.

The *Lie algebra* $\mathfrak{g}$ of a Lie group $\mathcal{G}$ is defined as $\mathfrak{g} = \{$left invariant vector fields on $\mathcal{G}\} \simeq T_e\mathcal{G}$ (tangent space at the identity e). The isomorphism is given (as in finite dimensions) by

$$\xi \in T_e\mathcal{G} \longmapsto X_{\xi} \in \mathfrak{g}, \quad X_{\xi}(g) := T_e L_g(\xi), \tag{2.3}$$

and the Lie bracket on $\mathfrak{g}$ is induced by the Lie bracket of left invariant vector fields $[\xi,\eta] = [X_{\xi}, X_{\eta}](e), \xi, \eta \in T_e\mathcal{G}$.

These definitions in infinite dimensions are identical with the definitions in finite dimensions. The big difference although is that infinite-dimensional manifolds, hence Lie groups, are *not* locally compact. For Frechet Lie groups, we have the additional nontrivial difficulty of the question how to define differentiability of functions defined on a Frechet space; see the study by Keller in [2]. Hence the very definition of a Frechet manifold is not canonical. This problem does not arise for Banach- and Hilbert-Lie groups; the differential calculus extends in a straightforward manner from $\mathbb{R}^n$ to Banach and Hilbert spaces, but not to Frechet spaces.

2.2. Finite- versus Infinite-Dimensional Lie Groups

Infinite-dimensional Lie groups are *NOT* locally compact. This causes some deficiencies of the Lie theory in infinite dimensions. We summarize some classical results in *finite* dimensions which are *NOT* true in general in infinite dimensions as follows.

(1) There is NO Implicit Function Theorem or Inverse Function Theorem in infinite dimensions! (except Nash-Moser-type theorems).

(2) If G is a finite-dimensional Lie group, the *exponential map* $\exp : \mathfrak{g} \to G$ is defined as follows. To each $\xi \in \mathfrak{g}$, we assign the corresponding left invariant vector field X_ξ defined by (2.3). We take the flow $\varphi_\xi(t)$ of X_ξ and define $\exp(\xi) = \varphi_\xi(1)$. The exponential map is a local diffeomorphism from a neighborhood of zero in $\mathfrak{g}$ onto a neighborhood of the identity in G; hence exp defines canonical coordinates on the Lie group G. This is not true in infinite dimensions.

(3) If $f_1, f_2 : G_1 \to G_2$ are smooth Lie group homomorphisms (i.e., $f_i(g \cdot h) = f_i(g) \cdot f_i(h)$, $i = 1, 2$) with $T_e f_1 = T_e f_2$, then locally $f_1 = f_2$. This is not true in infinite dimensions.

(4) If $f : G \to H$ is a continuous group homomorphism between finite-dimensional Lie groups, then f is smooth. This is not true in infinite dimensions.

(5) If $\mathfrak{g}$ is any finite-dimensional Lie algebra, then there exists a connected finite-dimensional Lie group G with $\mathfrak{g}$ as its Lie algebra; that is, $\mathfrak{g} \simeq T_e G$. This is not true in infinite dimensions.

(6) If G is a finite-dimensional Lie group and $H \subset G$ is a *closed* subgroup, then H is a Lie subgroup (i.e., Lie group and submanifold). This is not true in infinite dimensions.

(7) If G is a finite-dimensional Lie group with Lie algebra $\mathfrak{g}$ and $\mathfrak{h} \subset \mathfrak{g}$ is a subalgebra, then there exists a unique connected Lie subgroup $H \subset G$ with $\mathfrak{h}$ as its Lie algebra; that is, $\mathfrak{h} \simeq T_e H$. This is not true in infinite dimensions.

Some classical examples of *finite*-dimensional Lie groups are the matrix groups $GL(n)$, $SL(n)$, $O(n)$, $SO(n)$, $U(n)$, $SU(n)$, and $Sp(n)$ with smooth group operations given by matrix multiplication and matrix inversion. The Lie algebra bracket is the commutator $[A, B] = AB - BA$ with exponential map given by $\exp(A) = \sum_{i=0}^{\infty}(1/i!)A^i = e^A$.

2.3. Examples of Infinite-Dimensional Lie Groups

2.3.1. The Vector Groups $\mathcal{G} = (V, +)$

Let V be a Banach space and take $\mathcal{G} = V$ with $m(x, y) = x + y$, $i(x) = -x$, and $e = 0$, which makes $\mathcal{G}$ into an Abelian Lie group; that is, $m(x, y) = m(y, x)$. For the Lie algebra we have $\mathfrak{g} \simeq T_e V \simeq V$. For $u \in T_e V$ the corresponding left invariant vector field X_u is given by $X_u(v) = u$, $\forall v \in V$; that is, $X_u = \text{const}$. Hence the Lie algebra $\mathfrak{g} = V$ with the trivial Lie bracket $[u, v] = 0$ is Abelian. For the exponential map we get $\exp : \mathfrak{g} = V \to \mathcal{G} = V$, $\exp = id_V$.

2.3.2. The General Linear Group $\mathcal{G} = (GL(V), \circ)$

Let V be a Banach space and $L(V, V)$ the space of bounded linear operators $A : V \to V$. Then $L(V, V)$ is a Banach space with the operator norm $\|A\| = \sup_{\|x\| \leq 1} \|A(x)\|$, and the group $\mathcal{G} = GL(V)$ of all invertible elements is open in $L(V, V)$. So $GL(V)$ is a smooth Lie group with $m(f, g) = f \circ g$, $i(f) = f^{-1}$, and $e = id_V$. Its Lie algebra is $\mathfrak{g} = L(V, V)$ with the commutator bracket $[A, B] = AB - BA$ and exponential map $\exp A = e^A$.

2.3.3. *The Abelian Gauge Group* $\mathcal{G} = (C^\infty(M), +)$

Let M be a finite-dimensional manifold and let $\mathcal{G} = C^\infty(M)$ (smooth functions on M). With group operation being addition, that is, $m(f, g) = f + g$, $i(f) = -f$, and $e = 0$. $\mathcal{G}$ is an Abelian C^∞ (addition is smooth) Frechet Lie group with Lie algebra $\mathfrak{g} = T_e C^\infty(M) \simeq C^\infty(M)$, with trivial bracket $[\xi, \eta] = 0$, and $\exp = id$. If we complete these spaces in the C^k-norm, $k < \infty$ (denoted by $\mathcal{G}^k$), then $\mathcal{G}^k$ is a Banach-Lie group, and if we complete in the H^s-Sobolev norm with $s > (1/2)\dim M$ then $\mathcal{G}^s$ is a Hilbert-Lie group.

2.3.4. *The Abelian Gauge Group* $\mathcal{G} = (C^\infty(M, \mathbb{R} - \{0\}), \cdot)$

Let M be a finite-dimensional manifold and let $\mathcal{G} = C^\infty(M, \mathbb{R} - \{0\})$, with group operation being multiplication; that is, $m(f, g) = f \cdot g$, $i(f) = f^{-1}$, and $e = 1$. For $k < \infty$, $C^k(M, \mathbb{R} - \{0\})$ is open in $C^\infty(M, \mathbb{R})$, and if M is compact, then $C^k(M, \mathbb{R} - \{0\})$ is a Banach-Lie group. If $s > (1/2)\dim M$, then $H^s(M, \mathbb{R} - \{0\})$ is closed under multiplication, and if M is compact, then $H^s(M, \mathbb{R} - \{0\})$ is a Hilbert-Lie group.

2.3.5. *Loop Group* $\mathcal{G} = (C^k(M, G), \cdot)$

We generalize the Abelian example (see Section 2.3.4) by replacing $\mathbb{R} - \{0\}$ with any finite-dimensional (non-Abelian) Lie group G. Let $\mathcal{G} = C^k(M, G)$ with pointwise group operations $m(f, g)(x) = f(x) \cdot g(x)$, $x \in M$, and $i(f)(x) = (f(x))^{-1}$, where "$\cdot$" and "$(\cdot)^{-1}$" are the operations in G. If $k < \infty$ then $C^k(M, G)$ is a Banach-Lie group. Let g denote the Lie algebra of G, then the Lie algebra of $\mathcal{G} = C^k(M, G)$ is $\mathfrak{g} = C^k(M, g)$, with pointwise Lie bracket $[\xi, \eta](x) = [\xi(x), \eta(x)]$, $x \in M$, the latter bracket being the Lie bracket in g. The exponential map $\exp : g \to G$ defines the exponential map $\mathrm{EXP} : \mathfrak{g} = C^k(M, g) \to \mathcal{G} = C^k(M, G)$, $\mathrm{EXP}(\xi) = \exp \circ \xi$, which is a local diffeomorphism. The same holds for $H^s(M, G)$ if $s > (1/2)\dim M$.

Applications of these infinite-dimensional Lie groups are in gauge theories and quantum field theory, where they appear as groups of gauge transformations. We will discuss these in Section 5.

Special Case: $\mathcal{G} = (C^k(S^1, G), \cdot)$

As a special case of example mentioned in Section 2.3.5 we take $M = S^1$, the circle. Then $\mathcal{G} = C^k(S^1, G) = \mathcal{L}^k(G)$ is called a *loop group* and $\mathfrak{g} = C^k(S^1, g) = \mathfrak{l}^k(g)$ is its loop algebra. They find applications in the theory of affine Lie algebras, Kac-Moody Lie algebras (central extensions), completely integrable systems, soliton equations (Toda, KdV, KP), and quantum field theory; see, for example, [3] and Section 5. Central extensions of loop algebras are examples of infinite-dimensional Lie algebras which need not have a corresponding Lie group.

Certain subgroups of loop groups play an important role in quantum field theory as groups of gauge transformations. We will discuss these in Section 2.4.4.

2.4. *Diffeomorphism Groups*

Among the most important "classical" infinite-dimensional Lie groups are the diffeomorphism groups of manifolds. Their differential structure is not the one of a Banach Lie group as defined above. Nevertheless they have important applications.

Let M be a compact manifold (the noncompact case is technically much more complicated but similar results are true; see the study by Eichhorn and Schmid in [4]) and let $\mathcal{G} = \mathrm{Diff}^{\infty}(M)$ be the group of all smooth diffeomorphisms on M, with group operation being composition; that is, $m(f, g) = f \circ g$, $i(f) = f^{-1}$, and $e = id_M$. For C^{∞} diffeomorphisms, $\mathrm{Diff}^{\infty}(M)$ is a Frechet manifold and there are nontrivial problems with the notion of smooth maps between Frechet spaces. There is no canonical extension of the differential calculus from Banach spaces (which is the same as for $\mathbb{R}^n$) to Frechet spaces; see the study by Keller in [2]. One possibility is to generalize the notion of differentiability. For example, if we use the so-called C^{∞}_{Γ} differentiability, then $\mathcal{G} = \mathrm{Diff}^{\infty}(M)$ becomes a C^{∞}_{Γ} Lie group with C^{∞}_{Γ} differentiable group operations. These notions of differentiability are difficult to apply to concrete examples. Another possibility is to complete $\mathrm{Diff}^{\infty}(M)$ in the Banach C^k-norm, $0 \leq k < \infty$, or in the Sobolev H^s-norm, $s > (1/2)\dim M$. Then $\mathrm{Diff}^k(M)$ and $\mathrm{Diff}^s(M)$ become Banach and Hilbert manifolds, respectively. Then we consider the inverse limits of these Banach- and Hilbert-Lie groups, respectively:

$$\mathrm{Diff}^{\infty}(M) = \lim_{\leftarrow} \mathrm{Diff}^k(M) \tag{2.4}$$

becomes a so-called ILB- (Inverse Limit of Banach) Lie group, or with the Sobolev topologies

$$Dif f^{\infty}(M) = \lim_{\leftarrow} \mathrm{Diff}^s(M) \tag{2.5}$$

becomes a so-called ILH- (Inverse Limit of Hilbert) Lie group. See the study by Omori in [5] for details. Nevertheless, the group operations are not smooth, but have the following differentiability properties. If we equip the diffeomorphism group with the Sobolev H^s-topology, then $\mathrm{Diff}^s(M)$ becomes a C^{∞} Hilbert manifold if $s > (1/2)\dim M$ and the group multiplication

$$m : \mathrm{Diff}^{s+k}(M) \times \mathrm{Diff}^s(M) \longrightarrow \mathrm{Diff}^s(M) \tag{2.6}$$

is C^k differentiable; hence for $k = 0$, m is only continuous on $Dif f^s(M)$. The inversion

$$i : \mathrm{Diff}^{s+k}(M) \longrightarrow \mathrm{Diff}^s(M) \tag{2.7}$$

is C^k differentiable; hence for $k = 0$, i is only continuous on $\mathrm{Diff}^s(M)$. The same differentiability properties of m and i hold in the C^k topology.

The *Lie algebra* of $\mathrm{Diff}^{\infty}(M)$ is given by $\mathfrak{g} = T_e\mathrm{Diff}^{\infty}(M) \simeq \mathrm{Vec}^{\infty}(M)$ being the space of smooth vector fields on M. Note that the space $\mathrm{Vec}(M)$ of all vector fields is a Lie algebra only for C^{∞} vector fields, but not for C^k or H^s vector fields if $k < \infty$, $s < \infty$, because one loses derivatives by taking brackets.

The exponential map on the diffeomorphism group is given as follows. For any vector field $X \in \mathrm{Vec}^{\infty}(M)$, take its flow $\varphi_t \in \mathrm{Diff}^{\infty}(M)$, then define $\mathrm{EXP} : \mathrm{Vec}^{\infty}(M) \to \mathrm{Diff}^{\infty}(M) : X \mapsto \varphi_1$, the flow at time $t = 1$. The exponential map EXP is *NOT* a local diffeomorphism; it is not even locally surjective.

We see that the diffeomorphism groups are not Lie groups in the classical sense, but what we call *nested Lie groups*. Nevertheless they have important applications as we will see.

2.4.1. *Subgroups of Diff$^\infty(M)$*

Several subgroups of $\mathrm{Diff}^\infty(M)$ have important applications.

2.4.2. *Group of Volume-Preserving Diffeomorphisms*

Let μ be a volume on M and

$$\mathcal{G} = \mathrm{Diff}^\infty_\mu(M) = \left\{ f \in \mathrm{Diff}^\infty(M) \mid f^*\mu = \mu \right\} \tag{2.8}$$

the group of volume-preserving diffeomorphisms. $\mathrm{Diff}^\infty_\mu(M)$ is a closed subgroup of $\mathrm{Diff}^\infty(M)$ with Lie algebra

$$\mathfrak{g} = \mathrm{Vec}^\infty_\mu(M) = \left\{ X \in \mathrm{Vec}^\infty(M) \mid \mathrm{div}_\mu X = 0 \right\} \tag{2.9}$$

being the space of divergence-free vector fields on M. $\mathrm{Vec}^\infty_\mu(M)$ is a Lie subalgebra of $\mathrm{Vec}^\infty(M)$.

Remark 2.1. We cannot apply the finite-dimensional theorem that if $\mathrm{Vec}^\infty_\mu(M)$ is Lie algebra then there exists a Lie group whose Lie algebra it is; nor the one that if $\mathrm{Diff}^\infty_\mu(M) \subset \mathrm{Diff}(M)$ is a closed subgroup then it is an Lie subgroup.

Nevertheless $\mathrm{Diff}^\infty_\mu(M)$ is an ILH-Lie group.

2.4.3. *Symplectomorphism Group*

Let ω be a symplectic 2-form on M and

$$\mathcal{G} = \mathrm{Diff}^\infty_\omega(M) = \left\{ f \in \mathrm{xDiff}^\infty(M) \mid f^*\omega = \omega \right\} \tag{2.10}$$

the group of canonical transformations (or symplectomorphisms). $\mathrm{Diff}^\infty_\omega(M)$ is a closed subgroup of $\mathrm{Diff}^\infty(M)$ with Lie algebra

$$\mathfrak{g} = \mathrm{Vec}^\infty_\omega(M) = \{ X \in \mathrm{Vec}^\infty(M) \mid L_X\omega = 0 \} \tag{2.11}$$

being the space of locally Hamiltonian vector fields on M. $\mathrm{Vec}^\infty_\omega(M)$ is a Lie subalgebra of $\mathrm{Vec}^\infty(M)$. Again $\mathrm{Diff}^\infty_\omega(M)$ is an ILH-Lie group.

2.4.4. *Group of Gauge Transformations*

The diffeomorphism subgroups that arise in gauge theories as gauge groups behave nicely because they are isomorphic to subgroups of loop groups which are not only ILH-Lie groups but actually Hilbert-Lie groups.

Let $\pi : P \to M$ be a principal G bundle with G being a finite-dimensional Lie group (structure group) acting on P from the right $p \in P$, $g \in G$, and $p \cdot g \in P$.

The *Gauge group* $\mathcal{G}$ is the group of gauge transformations defined by

$$\mathcal{G} = \{\phi \in \mathrm{Diff}^{\infty}(P);\ \phi(p\cdot g) = \phi(p)\cdot g,\ \pi(\phi(p)) = \pi(p)\}. \tag{2.12}$$

$\mathcal{G}$ is a group under composition, hence a subgroup of the diffeomorphism group $\mathrm{Diff}^{\infty}(P)$. Since a gauge transformation $\phi \in \mathcal{G}$ preserves fibers, we can realize each such $\phi \in \mathcal{G}$ via $\phi(p) = p\cdot\tau(p)$, where $\tau : P \to G$ satisfies $\tau(p\cdot g) = g^{-1}\tau(p)g$, for $p \in P$, $g \in G$. Let

$$\mathrm{Gau}(P) = \left\{\tau \in C^{\infty}(P,G);\ \tau(p\cdot g) = g^{-1}\tau(p)g\right\}. \tag{2.13}$$

$\mathrm{Gau}(P)$ is a group under pointwise multiplication, hence a subgroup of the loop group $C^{\infty}(P,G)$ (see Section 2.4.3), which extends to a Hilbert-Lie group if equipped with the H^s-Sobolev topology. We give $\mathrm{Gau}(P)$ the induced topology and extend it to a Hilbert-Lie group denoted by $\mathrm{Gau}^s(P)$. Another interpretation is that $\mathrm{Gau}(P)$ is isomorphic to $C^{\infty}(Ad\,P)$ the space of sections of the associated vector bundle $Ad(P) = P \times_G G$. Completed in the H^s Sobolev topology, we get $\mathrm{Gau}^s(P) \simeq H^s(Ad\,P)$.

Let $\mathbf{g}$ denote the Lie algebra of G. Then the *Lie algebra* $\mathfrak{gau}(P)$ of $\mathrm{Gau}(P)$ is a subalgebra of the loop algebra $H^s(P,\mathbf{g})$ under pointwise bracket in $\mathbf{g}$, the finite-dimensional Lie algebra of G; that is, for any $\xi,\eta \in H^s(P,\mathbf{g})$ the bracket is defined by $[\xi,\eta]_{\mathfrak{gau}(P)}(p) = [\xi(p),\eta(p)]_{\mathbf{g}}$, $p \in P$. Then $\mathfrak{gau}^s(P)$ is the subalgebra of Ad-invariant $\mathbf{g}$-valued functions on P; that is,

$$\mathfrak{gau}(P) = \{\xi \in C^{\infty}(P,\mathbf{g});\ \xi(p\cdot g) = Ad_{g^{-1}}\xi(p)\}. \tag{2.14}$$

The Lie algebra $\mathfrak{lie}\ \mathcal{G}$ (running out of symbols) of the gauge group $\mathcal{G}$ is the Lie subalgebra of $\mathfrak{X}^{\infty}(P)$ consisting of all G-invariant vertical vector fields X on P; that is,

$$\mathfrak{lie}\,\mathcal{G} = \left\{X \in \mathfrak{X}^{\infty}(P);\ R_g^{*}X = X,\ X(p) \in \mathbf{g},\ g \in G,\ p \in P\right\} \tag{2.15}$$

with commutator bracket $[X_1,X_2] = X_1X_2 - X_2X_1 \in \mathfrak{lie}\ \mathcal{G}$.

On the other hand, the Lie algebra of $C^{\infty}(Ad\,P)$ is $C^{\infty}(ad(P))$ being the space of sections of the associated vector bundle $ad(P) \equiv (P\times_G g) \to M$ with pointwise bracket.

We have three versions of gauge groups: $\mathcal{G}$, $\mathrm{Gau}(P)$, and $C^{\infty}(Ad\,P)$. They are all group isomorphic. There is a natural group isomorphism $\mathrm{Gau}(P) \to \mathcal{G} : \tau \mapsto \phi$ defined by $\phi(p) = p\cdot\tau(p)$, $p \in P$, which preserves the product $\tau_1\cdot\tau_2 \mapsto \phi_1\circ\phi_2$. Identifying $\mathcal{G}$ with $\mathrm{Gau}(P)$, we can avoid the troubles with diffeomorphism groups and we can extend $\mathcal{G}$ to a Hilbert-Lie group $\mathcal{G}^s$. So $\mathcal{G}^s$ is actually a Hilbert-Lie group in the classical sense; that is, the group operations are C^{∞}. Also the three Lie algebras $\mathfrak{lie}\,\mathcal{G}$, $\mathfrak{gau}(P)$, and $C^{\infty}(ad\,P)$ are canonically isomorphic. Indeed, for $s \in C^{\infty}(ad\,P)$ define $\xi \in \mathfrak{gau}(P)\xi : P \to g$ by $\xi(p\cdot a) := Ad_{a^{-1}}\xi(p)$; and for $\xi \in \mathfrak{gau}(P)$ define $s \in C^{\infty}(ad\,P)$ by $s(\pi(p)) := [p,\xi(p)]$.

On the other hand, for $\xi \in \mathfrak{gau}(P)$ define $Z_{\xi} \in \mathfrak{lie}\,\mathcal{G}$ by

$$Z_{\xi}(p) = \frac{d}{dt}|_{t=0}R(p,\exp t\xi(p)) \quad (= \xi(p)^{*}(p)), \tag{2.16}$$

that is, Z_{ξ} is the fundamental vector field on P, generated by $\xi \in \mathbf{g}$. Z_{ξ} is invariant if and only if $\xi(p\cdot g) = Ad_{g^{-1}}\xi(p)$.

To topologize $\mathfrak{lie}\ \mathcal{G}$, we complete $C^\infty(ad\,P)$ in the H^s-Sobolev norm. If $s > (1/2)\dim M$, then $\mathfrak{lie}\,\mathcal{G}^s \simeq H^s(ad\,P) \simeq \mathfrak{gau}^s(P)$ are isomorphic Hilbert-Lie algebras.

There is a natural exponential map $\mathrm{Exp} : \mathfrak{gau}(P) \to \mathrm{Gau}(P)$, which is a local diffeomorphism. Let $\exp : \mathbf{g} \to G$ be the finite-dimensional exponential map. Then define

$$\mathrm{Exp} : \mathfrak{gau}^s(P) \to \mathrm{Gau}^s(P) : (\mathrm{Exp}\,\xi)(p) = \exp(\xi(p)), \quad \xi \in \mathfrak{gau}^s(P). \tag{2.17}$$

Or in terms of $\mathcal{G}$, $\mathrm{Exp} : \mathfrak{lie}\ \mathcal{G}^s \to \mathcal{G}^s : (\mathrm{Exp}\,\xi)(p) = p \cdot \exp(\xi_p)$.

We have the following theorem (Schmid [6]).

Theorem 2.2. *For $s > (1/2)\dim M$,*

$$\mathcal{G}^s \simeq \mathrm{Gau}^s(P) \simeq H^s(Ad\,P) \tag{2.18}$$

is a smooth Hilbert-Lie group with Lie algebra

$$\mathfrak{lie}\,\mathcal{G}^s \simeq \mathfrak{gau}^s(P) \simeq H^s(ad\,P) \tag{2.19}$$

and smooth exponential map, which is a local diffeomorphism,

$$\mathrm{EXP} : \mathfrak{lie}\,\mathcal{G}^s \longrightarrow \mathcal{G}^s : (\mathrm{EXP}\,\xi)(p) = p \cdot \exp(\xi(p)). \tag{2.20}$$

See [1–5, 7–19].

3. Lie Groups as Symmetry Groups of Hamiltonian Systems

A short introduction and "crash course" to geometric mechanics can be found in the studies by Abraham and Marsden [20], Marsden [21], as well as Marsden and Ratiu [22]. For the general theory of infinite-dimensional manifolds and global analysis, see, for example, the studies by Bourbaki [9], Lang [14], as well as Palais [18].

3.1. Hamilton's Equations on Poisson Manifolds

A *Poisson manifold* is a manifold P (in general infinite-dimensional) equipped with a bilinear operation $\{\cdot,\cdot\}$, called *Poisson bracket*, on the space $C^\infty(P)$ of smooth functions on P satisfying the following.

(i) $(C^\infty(P), \{\cdot,\cdot\})$ is a Lie algebra; that is, $\{\cdot,\cdot\} : C^\infty(P) \times C^\infty(P) \to C^\infty(P)$ is bilinear, skew symmetric and satisfies the Jacobi identity $\{\{F,G\},H\} + \{\{H,F\},G\} + \{\{G,H\},F\} = 0$ for all $F, G, H \in C^\infty(P)$.

(ii) $\{\cdot,\cdot\}$ satisfies the Leibniz rule; that is, $\{\cdot,\cdot\}$ is a derivation in each factor: $\{F \cdot G, H\} = F \cdot \{G,H\} + G \cdot \{F,H\}$, for all $F, G, H \in C^\infty(P)$.

The notion of Poisson manifolds was rediscovered many times under different names, starting with Lie, Dirac, Pauli, and others. The name *Poisson manifold* was coined by Lichnerowicz.

For any $H \in C^\infty(P)$ we define the *Hamiltonian vector field* X_H by

$$X_H(F) = \{F, H\}, \quad F \in C^\infty(P). \tag{3.1}$$

It follows from (ii) that indeed X_H defines a derivation on $C^\infty(P)$, hence a vector field on P. *Hamilton's equations* of motion for a function $F \in C^\infty(P)$ with Hamiltonian $H \in C^\infty(P)$ (energy function) are then defined by the flow (integral curves) of the vector field X_H; that is,

$$\dot{F} = X_H(F) = \{F, H\}, \quad \text{where } \dot{} = \frac{d}{dt}. \tag{3.2}$$

We then call F a *Hamiltonian system* on P with energy (Hamiltonian function) H.

3.2. Examples of Poisson Manifolds and Hamilton's Equations

Poisson manifolds are a generalization of symplectic manifolds on which Hamilton's equations have a canonical formulated.

3.2.1. Finite-Dimensional Classical Mechanics

For finite-dimensional classical mechanics we take $P = \mathbb{R}^{2n}$ with coordinates $(q^1, \ldots, q^n, p_1, \ldots, p_n)$ with the standard Poisson bracket for any two functions $F(q^i, p_i)$, $H(q^i, p_i)$ given by

$$\{F, H\} = \sum_{i=1}^{n} \frac{\partial F}{\partial p_i}\frac{\partial H}{\partial q^i} - \frac{\partial H}{\partial p_i}\frac{\partial F}{\partial q^i}. \tag{3.3}$$

Then the *classical Hamilton's equations* are

$$\dot{q}^i = \left\{q^i, H\right\} = \frac{\partial H}{\partial p_i}, \quad \dot{p}_i = \{p_i, H\} = -\frac{\partial H}{\partial q^i}, \tag{3.4}$$

where $i = 1, \ldots, n$. This finite-dimensional Hamiltonian system is a system of ordinary differential equations for which there are well-known existence and uniqueness theorems; that is, it has locally unique smooth solutions, depending smoothly on the initial conditions.

Example 3.1 (Harmonic Oscillator). As a concrete example we consider the harmonic oscillator. Here $P = \mathbb{R}^2$ and the Hamiltonian (energy) is $H(q, p) = (1/2)(q^2 + p^2)$. Then Hamilton's equations are

$$\dot{q} = p, \qquad \dot{p} = -q. \tag{3.5}$$

3.2.2. Infinite-Dimensional Classical Field Theory

Let V be a Banach space and V^* its dual space with respect to a pairing $\langle\cdot,\cdot\rangle : V\times V^* \to \mathbb{R}$ (i.e., $\langle\cdot,\cdot\rangle$ is a symmetric, bilinear, nondegenerate function). On $P = V \times V^*$ we have the canonical Poisson bracket for $F, H \in C^\infty(P)$, $\varphi \in V$, and $\pi \in V^*$, given by

$$\{F, H\} = \left\langle \frac{\delta F}{\delta \pi}, \frac{\delta H}{\delta \varphi} \right\rangle - \left\langle \frac{\delta H}{\delta \pi}, \frac{\delta F}{\delta \varphi} \right\rangle, \tag{3.6}$$

where the functional derivatives $\delta F/\delta\pi \in V$, $\delta F/\delta\varphi \in V^*$ are the "duals" under the pairing $\langle\cdot,\cdot\rangle$ of the partial gradients $D_1F(\pi) \in V^*, D_2F(\varphi) \in V^{**} \simeq V$. The corresponding Hamilton's equations are

$$\dot{\varphi} = \{\varphi, H\} = \frac{\delta H}{\delta \pi}, \quad \dot{\pi} = \{\pi, H\} = -\frac{\delta H}{\delta \varphi}. \tag{3.7}$$

As a special case in finite dimensions, if $V \simeq \mathbb{R}^n$, so that $V^* \simeq \mathbb{R}^n$ and $P = V \times V^* \simeq \mathbb{R}^{2n}$, and the pairing is the standard inner product in $\mathbb{R}^n$, then the Poisson bracket (3.6) and Hamilton's equations (3.7) are identical with (3.3) and (3.4), respectively.

Example 3.2 (Wave Equations). As a concrete example we consider the wave equations. Let $V = C^\infty(\mathbb{R}^3)$ and $V^* = \mathrm{Den}(\mathbb{R}^3)$ (densities) and the L^2 pairing $\langle\varphi, \pi\rangle = \int \varphi(x)\pi(x)dx$. We take the Hamiltonian to be $H(\varphi, \pi) = \int((1/2)\pi^2 + (1/2)|\nabla\varphi|^2 + F(\varphi))dx$, where F is some function on V. Then Hamilton's (3.7) become

$$\dot{\varphi} = \pi, \qquad \dot{\pi} = \nabla^2\varphi - F'(\varphi), \quad \text{where } ' = \frac{d}{d\varphi}, \tag{3.8}$$

which imply the wave equation $\partial^2\varphi/\partial t^2 = \nabla^2\varphi - F'(\varphi)$. Different choices of F give different wave equations; for example, for $F = 0$ we get the linear wave equation $\partial^2\varphi/\partial t^2 = \nabla^2\varphi$. For $F = (1/2)m\varphi$ we get the Klein-Gordon equation $\nabla^2\varphi - \partial^2\varphi/\partial t^2 = m\varphi$. So these wave equations and the Klein-Gordon equation are infinite-dimensional Hamiltonian systems on $P = C^\infty(\mathbb{R}^3) \times \mathrm{Den}(\mathbb{R}^3)$.

3.2.3. Cotangent Bundles

The finite-dimensional examples of Poisson brackets (3.3) and Hamilton's (3.4) and the infinite-dimensional examples (3.6) and (3.7) are the local versions of the general case where $P = T^*Q$ is the cotangent bundle (phase space) of a manifold Q (configuration space). If Q is an n-dimensional manifold, then T^*Q is a $2n$-Poisson manifold locally isomorphic to $\mathbb{R}^{2n}$ whose Poisson bracket is locally given by (3.3) and Hamilton's equations are locally given by (3.4). If Q is an infinite-dimensional Banach manifold, then T^*Q is a Poisson manifold locally isomorphic to $V \times V^*$ whose Poisson bracket is given by (3.6) and Hamilton's equations are locally given by (3.7).

3.2.4. Symplectic Manifolds

All the examples above are special cases of symplectic manifolds (P, ω). That means that P is equipped with a symplectic structure ω which is a closed ($d\omega = 0$), (weakly) nondegenerate 2-form on the manifold P. Then for any $H \in C^\infty(P)$ the corresponding Hamiltonian vector field X_H is defined by $dH = \omega(X_H, \cdot)$ and the canonical Poisson bracket is given by

$$\{F, H\} = \omega(X_F, X_H), \quad F, H \in C^\infty(P). \tag{3.9}$$

For example, on $\mathbb{R}^{2n}$ the canonical symplectic structure ω is given by $\omega = \sum_{i=1}^n dp_i \wedge dq^i = d\theta$, where $\theta = \sum_{i=1}^n p_i \wedge dq^i$. The same formula for ω holds locally in T^*Q for any finite-dimensional Q (Darboux's Lemma). For the infinite-dimensional example $P = V \times V^*$, the symplectic form ω is given by $\omega((\varphi_1, \pi_1), (\varphi_2, \pi_2)) = \langle \varphi_1, \pi_2 \rangle - \langle \varphi_2, \pi_1 \rangle$. Again these two formulas for ω are identical if $V = \mathbb{R}^n$.

Remark 3.3. (A) If P is a finite-dimensional symplectic manifold, then P is even dimensional.

(B) If the Poisson bracket $\{\cdot, \cdot\}$ is nondegenerate, then $\{\cdot, \cdot\}$ comes from a symplectic form ω; that is, $\{\cdot, \cdot\}$ is given by (3.9).

3.2.5. The Lie-Poisson Bracket

Not all Poisson brackets are of the form given in the above examples (3.3), (3.6), and (3.9); that is, not all Poisson manifolds are symplectic manifolds. An important class of Poisson bracket is the so-called *Lie-Poisson bracket*. It is defined on the dual of any Lie algebra. Let G be a Lie group with Lie algebra $\mathfrak{g} = T_eG \simeq$ {left invariant vector fields on G}, and let $[\cdot, \cdot]$ denote the Lie bracket (commutator) on $\mathfrak{g}$. Let $\mathfrak{g}^*$ be the dual of a $\mathfrak{g}$ with respect to a pairing $\langle \cdot, \cdot \rangle : \mathfrak{g}^* \times \mathfrak{g} \to \mathbb{R}$. Then for any $F, H \in C^\infty(\mathfrak{g}^*)$ and $\mu \in \mathfrak{g}^*$, the *Lie-Poisson bracket* is defined by

$$\{F, H\}(\mu) = \pm \left\langle \mu, \left[\frac{\delta F}{\delta \mu}, \frac{\delta H}{\delta \mu} \right] \right\rangle, \tag{3.10}$$

where $\delta F/\delta\mu, \delta H/\delta\mu \in \mathfrak{g}$ are the "duals" of the gradients $DF(\mu), DH(\mu) \in \mathfrak{g}^{**} \simeq \mathfrak{g}$ under the pairing $\langle \cdot, \cdot \rangle$. Note that the Lie-Poisson bracket is degenerate in general; for example, for $G = SO(3)$ the vector space $\mathfrak{g}^*$ is 3 dimensional, so the Poisson bracket (3.10) cannot come from a symplectic structure. This Lie-Poisson bracket can also be obtained in a different way by taking the canonical Poisson bracket on T^*G (locally given by (3.3) and (3.6)) and then restricting it to the fiber at the identity $T_e^*G = \mathfrak{g}^*$. In this sense the Lie-Poisson bracket (3.10) is induced from the canonical Poisson bracket on T^*G. It is induced by the symmetry of left multiplication as we will discuss in Section 3.3.

Example 3.4 (Rigid Body). A concrete example of the Lie-Poisson bracket is given by the rigid body. Here $G = SO(3)$ is the configuration space of a free rigid body. Identifying the Lie algebra $(\mathfrak{so}(3), [\cdot, \cdot])$ with $(\mathbb{R}^3, \times)$, where $\times$ is the vector product on $\mathbb{R}^3$, and $\mathfrak{g}^* = \mathfrak{so}(3)^* \simeq \mathbb{R}^3$, the Lie-Poisson bracket translates into

$$\{F, H\}(m) = -m \cdot (\nabla F \times \nabla H). \tag{3.11}$$

For any $F \in C^\infty(\mathfrak{so}(3)^*)$, we have $(dF/dt)(m) = \nabla F \cdot \dot{m} = \{F, H\}(m) = -m \cdot (\nabla F \times \nabla H) = \nabla F \cdot (m \times \nabla H)$; hence $\dot{m} = m \times \nabla H$. With the Hamiltonian $H = (1/2)(m_1^2/I_1^2 + m_2^2/I_2^2 + m_3^2/I_3^2)$ we get Hamilton's equation as

$$\dot{m}_1 = \frac{I_2 - I_3}{I_2 I_3} m_2 m_3, \qquad \dot{m}_2 = \frac{I_3 - I_1}{I_3 I_1} m_3 m_1, \qquad \dot{m}_3 = \frac{I_1 - I_2}{I_1 I_2} m_1 m_2. \tag{3.12}$$

These are *Euler's* equations for the free rigid body.

3.3. *Reduction by Symmetries*

The examples we have discussed so far are all canonical examples of Poisson brackets, defined either on a symplectic manifold (P, ω) or T^*Q, or on the dual of a Lie algebra $\mathfrak{g}^*$. Different, noncanonical Poisson brackets can arise from *symmetries.* Assume that a Lie group G is acting in a Hamiltonian way on the Poisson manifold $(P, \{\cdot, \cdot\})$. That means that we have a smooth map $\varphi : G \times P \to P : \varphi(g, p) = g \cdot p$ such that the induced maps $\varphi_g = \varphi(g, \cdot) : P \to P$ are *canonical transformations*, for each $g \in G$. In terms of Poisson manifolds, a canonical transformation is a smooth map that preserves the Poisson bracket. So the action of G on P is a *Hamiltonian action* if $\varphi_g^*\{F, H\} = \{\varphi_g^* F, \varphi_g^* H\}$, for all $F, H \in C^\infty(P), g \in G$. For any $\xi \in \mathfrak{g}$ the canonical transformations $\varphi_{\exp(t\xi)}$ generate a Hamiltonian vector field ξ_F on P and a momentum map $J : P \to \mathfrak{g}^*$ given by $J(x)(\xi) = F(x)$, which is Ad^* equivariant.

If a Hamiltonian system X_H is invariant under a Lie group action, that is, $H(\varphi_g(x)) = H(x)$, then we obtain a reduced Hamiltonian system on a reduced phase space (reduced Poisson manifold). We recall the following Marsden-Weinstein reduction theorem [23].

Theorem 3.5 (Reduction Theorem). *For a Hamiltonian action of a Lie group G on a Poisson manifold $(P, \{\cdot, \cdot\})$, there is an equivariant momentum map $J : P \to \mathfrak{g}^*$ and for every regular $\mu \in \mathfrak{g}^*$ the reduced phase space $P_\mu \equiv J^{-1}(\mu)/G_\mu$ carries an induced Poisson structure $\{\cdot, \cdot\}_\mu$ (G_μ being the isotropy group). Any G-invariant Hamiltonian H on P defines a Hamiltonian H_μ on the reduced phase space P_μ, and the integral curves of the vector field X_H project onto integral curves of the induced vector field $\widehat{X}_{H_\mu}$ on the reduced space P_μ.*

Example 3.6 (Rigid Body). The rigid body discussed above can be viewed as an example of this reduction theorem. If $P = T^*G$ and G is acting on T^*G by the cotangent lift of the left translation $l_g : G \to G$, $l_g(h) = gh$, then the momentum map $J : T^*G \to \mathfrak{g}^*$ is given by $J(\alpha_g) = T_e^* R_g(\alpha_g)$ and the reduced phase space $(T^*G)_\mu = J^{-1}(\mu)/G_\mu$ is isomorphic to the coadjoint orbit $\mathcal{O}_\mu$ through $\mu \in \mathfrak{g}^*$. Each coadjoint orbit $\mathcal{O}_\mu$ carries a natural symplectic structure ω_μ, and in this case, the reduced Lie-Poisson bracket $\{\cdot, \cdot\}_\mu$ on the coadjoint orbit $\mathcal{O}_\mu$ is induced by the symplectic form ω_μ on $\mathcal{O}_\mu$ as in (3.9). Furthermore $T^*G/G \simeq \mathfrak{g}^*$ and the induced Poisson bracket $\{\cdot, \cdot\}_\mu$ on $\mathcal{O}_\mu$ are identical with the Lie-Poisson bracket restricted to the coadjoint orbit $\mathcal{O}_\mu \subset \mathfrak{g}^*$. For the rigid body we apply this construction to $G = SO(3)$.

See [1, 8, 10, 17, 19–31].

4. Applications

We now discuss some infinite-dimensional examples of reduced Hamiltonian systems.

4.1. Maxwell's Equations

Maxwell's equations of electromagnetism are a reduced Hamiltonian system with the Lie group $\mathcal{G} = (C^\infty(M), +)$ discussed in Section 2.3.3 as symmetry group.

Let E, B be the electric and magnetic fields on $\mathbb{R}^3$, then Maxwell's equations for a charge density ρ are

$$\dot{E} = \operatorname{curl} B, \qquad \dot{B} = -\operatorname{curl} E, \tag{4.1}$$

$$\operatorname{div} B = 0, \qquad \operatorname{div} E = \rho. \tag{4.2}$$

Let A be the magnetic potential such that $B = -\operatorname{curl} A$. As configuration space we take $V = \operatorname{Vec}(\mathbb{R}^3)$, vector fields (potentials) on $\mathbb{R}^3$, so $A \in V$, and as phase space we have $P = T^*V \simeq V \times V^* \ni (A, E)$, with the standard L^2 pairing $\langle A, E\rangle = \int A(x)E(x)dx$, and canonical Poisson bracket given by (3.6), which becomes

$$\{F, H\}(A, E) = \int \left(\frac{\delta F}{\delta A} \frac{\delta H}{\delta E} - \frac{\delta H}{\delta A} \frac{\delta F}{\delta E} \right) dx. \tag{4.3}$$

As Hamiltonian we take the total electromagnetic energy

$$H(A, E) = \frac{1}{2} \int \left(|\operatorname{curl} A|^2 + |E|^2 \right) dx. \tag{4.4}$$

Then Hamilton's equations in the canonical variables A and E are $\dot{A} = \delta H/\delta E = E \Rightarrow \dot{B} = -\operatorname{curl} E$ and $\dot{E} = -\delta H/\delta A = -\operatorname{curl}\operatorname{curl} A = \operatorname{curl} B$. So the first two equations of Maxwell's equations (4.1) are Hamilton's equations; we get the third one automatically from the potential $\operatorname{div} B = -\operatorname{div}\operatorname{curl} A = 0$ and we obtain the 4th equation $\operatorname{div} E = \rho$ through the following symmetry (gauge invariance). The Lie group $\mathcal{G} = (C^\infty(\mathbb{R}^3), +)$ acts on V by $\varphi \cdot A = A + \nabla\varphi$, $\varphi \in \mathcal{G}$, $A \in V$. The lifted action to $V \times V^*$ becomes $\varphi \cdot (A, E) = (A + \nabla\varphi, E)$, and has the momentum map $J : V \times V^* \longrightarrow \mathfrak{g}^* \simeq \{\text{charge densities}\}$:

$$J(A, E) = \operatorname{div} E. \tag{4.5}$$

With $\mathfrak{g} = C^\infty(\mathbb{R}^3)$ and $\mathfrak{g}^* = \operatorname{Den}(\mathbb{R}^3)$, we identify elements of $\mathfrak{g}^*$ with charge densities. The Hamiltonian H is $\mathcal{G}$ invariant; that is, $H(\varphi \cdot (A, E)) = H(A + \nabla\varphi, E) = H(A, E)$. Then the reduced phase space for $\rho \in \mathfrak{g}^*$ is $(V \times V^*)_\rho = J^{-1}(\rho)/G = \{(E, B) \mid \operatorname{div} E = \rho, \operatorname{div} B = 0\}$ and the reduced Hamiltonian is

$$H_\rho(E, B) = \frac{1}{2} \int \left(|E|^2 + |B|^2 \right) dx. \tag{4.6}$$

The reduced Poisson bracket becomes for any functions F, H on $(V \times V^*)_\rho$

$$\{F, H\}_\rho(E, B) = \int \left(\frac{\delta F}{\delta E} \cdot \operatorname{curl} \frac{\delta H}{\delta B} - \frac{\delta H}{\delta E} \cdot \operatorname{curl} \frac{\delta F}{\delta B} \right) dx, \tag{4.7}$$

and a straightforward computation shows that

$$\dot{F} = \{F, H_\rho\}_\rho \Longleftrightarrow \begin{cases} \dot{E} = \operatorname{curl} B, & \dot{B} = -\operatorname{curl} E, \\ \operatorname{div} B = 0, & \operatorname{div} E = \rho. \end{cases} \tag{4.8}$$

So Maxwell's equations (4.1), (4.2) are an infinite-dimensional Hamiltonian system on this reduced phase space with respect to the reduced Poisson bracket.

4.2. *Fluid Dynamics*

Euler's equations for an incompressible fluid

$$\frac{\partial u}{\partial t} + u \cdot \nabla u = -\nabla p, \quad \operatorname{div} u = 0 \tag{4.9}$$

are equivalent to the equations of geodesics on $\mathrm{Diff}_\mu^\infty(M)$. See the study by Marsden et al. in [15] for details.

4.3. *Plasma Physics*

The Maxwell-Vlasov's equations are a reduced Hamiltonian system on a more complicated reduced space. See the study by Marsden et al. in [32] for details.

Maxwell-Vlasov's equations for a plasma density $f(x, v, t)$ generating the electric and magnetic fields E and B are the following set of equations:

$$\begin{aligned}
&\frac{\partial f}{\partial t} + v \cdot \frac{\partial f}{\partial x} + (E + v \times B)\frac{\partial f}{\partial v} = 0, \\
&\frac{\partial B}{\partial t} = -\operatorname{curl} E, \\
&\frac{\partial E}{\partial t} = \operatorname{curl} B - J_{\mathrm{f}}, \quad J_f = \text{current density}, \\
&\operatorname{div} E = \rho_f, \quad \rho_f = \text{charge density}, \\
&\operatorname{div} B = 0.
\end{aligned} \tag{4.10}$$

This coupled nonlinear system of evolution equations is an infinite-dimensional Hamiltonian system of the form $\dot{F} = \{F, H\}_{\rho_f}$ on the reduced phase space

$$\mathcal{MV} = \left(T^*\mathrm{Diff}_\omega^\infty\left(\mathbb{R}^6\right) \times T^*V\right) / C^\infty\left(\mathbb{R}^6\right) \tag{4.11}$$

(V being the same space as in the example of Maxwell's equations) with respect to the following reduced Poisson bracket, which is induced via gauge symmetry from the canonical Poisson bracket on $T^*\mathrm{Diff}^\infty_\omega(\mathbb{R}^6) \times T^*V$:

$$
\begin{aligned}
\{F,G\}_{\rho_f}(f,E,B) = &\int f\left\{\frac{\delta F}{\delta f},\frac{\delta G}{\delta f}\right\}dx\,dv \\
&+\int\left(\frac{\delta F}{\delta E}\cdot \operatorname{curl}\frac{\delta G}{\delta B}-\frac{\delta G}{\delta E}\cdot \operatorname{curl}\frac{\delta F}{\delta B}\right)dx\,dv \\
&+\int\left(\frac{\delta F}{\delta E}\cdot\frac{\partial f}{\partial v}\frac{\delta G}{\delta f}-\frac{\delta G}{\delta E}\cdot\frac{\partial f}{\partial v}\frac{\delta F}{\delta f}\right)dx\,dv \\
&+\int fB\cdot\left(\frac{\partial}{\partial v}\frac{\delta F}{\delta f}\times\frac{\partial}{\partial v}\frac{\delta G}{\delta f}\right)dx\,dv,
\end{aligned}
\tag{4.12}
$$

and with Hamiltonian

$$
H(f,E,B) = \frac{1}{2}\int v^2 f(x,v,t)dv + \frac{1}{2}\int\left(|E|^2+|B|^2\right)dx. \tag{4.13}
$$

More complicated plasma models are formulated as Hamiltonian systems. For example, for the two-fluid model the phase space is a coadjoint orbit of the semidirect product ($\ltimes$) of the group $\mathcal{G} = \mathrm{Diff}^\infty(\mathbb{R}^6) \ltimes (C^\infty(\mathbb{R}^6) \times C^\infty(\mathbb{R}^6))$. For the MHD model, $\mathcal{G} = \mathrm{Diff}^\infty(\mathbb{R}^6) \ltimes (C^\infty(\mathbb{R}^6) \times \Omega^2(\mathbb{R}^3))$.

4.4. The KdV Equation and Fourier Integral Operators

There are many known examples of PDEs which are infinite-dimensional Hamiltonian systems, such as the Benjamin-Ono, Boussinesq, Harry Dym, KdV, KP equations, and others. In many cases the Poisson structures and Hamiltonians are given *ad* hoc on a formal level. We illustrate this with the KdV equation, where at least one of the three known Hamiltonian structures is well understood [33].

The Korteweg-deVries (KdV) equation

$$
u_t + 6uu_x + u_{xxx} = 0 \tag{4.14}
$$

is an infinite-dimensional Hamiltonian system with the Lie group of invertible Fourier integral operators as symmetry group. Gardner found that with the bracket

$$
\{F,G\} = \int_0^{2\pi}\frac{\delta F}{\delta u}\frac{\partial}{\partial x}\frac{\delta G}{\delta u}\,dx \tag{4.15}
$$

and Hamiltonian

$$
H(u) = \int_0^{2\pi}\left(u^3+\frac{1}{2}u_x^3\right)dx \tag{4.16}
$$

u satisfies the KdV equation (4.14) if and only if

$$\dot{u} = \{u, H\}. \tag{4.17}$$

The question is where this Poisson bracket (4.15) and Hamiltonian (4.16) come from? We showed [33–35] that this bracket is the Lie-Poisson bracket on a coadjoint orbit of Lie group $\mathcal{G}$ = FIO of invertible Fourier integral operators on the circle S^1. We briefly summarize the following.

A Fourier integral operators on a compact manifold M is an operator

$$A : C^\infty(M) \longrightarrow C^\infty(M) \tag{4.18}$$

locally given by

$$A(u)(x) = (2\pi)^{-n} \iint e^{i\varphi(x,y,\xi)} a(x,\xi) u(y) dy\, d\xi, \tag{4.19}$$

where $\varphi(x, y, \xi)$ is a phase function with certain properties and the symbol $a(x, \xi)$ belongs to a certain symbol class. A pseudodifferential operator is a special kind of Fourier integral operators, locally of the form

$$P(u)(x) = (2\pi)^{-n} \iint e^{i(x-y)\cdot\xi} p(x,\xi) u(y) dy\, d\xi. \tag{4.20}$$

Denote by FIO and ΨDO the groups under composition (operator product) of invertible Fourier integral operators and invertible pseudodifferential operators on M, respectively. We have the following results.

Both groups ΨDO and FIO are smooth infinite-dimensional ILH-Lie groups. The smoothness properties of the group operations (operator multiplication and inversion) are similar to the case of diffeomorphism groups (2.6), (2.7). The Lie algebras of both ILH-Lie groups ΨDO and FIO are the Lie algebras of all pseudodifferential operators under the commutator bracket. Moreover, FIO is a smooth infinite-dimensional principal fiber bundle over the diffeomorphism group of canonical transformations $\mathrm{Diff}^\infty_\omega(T^*M - \{0\})$ with structure group (gauge group) ΨDO.

For the KdV equation we take the special case where $M = S^1$. Then the Gardner bracket (4.15) is the Lie-Poisson bracket on the coadjoint orbit of FIO through the Schrodinger operator $P \in \Psi DO$. Complete integrability of the KdV equation follows from the infinite system of conserved integral in involution given by $H_k = \mathrm{Trace}(P^k)$; in particular the Hamiltonian (4.16) equals $H = H_2$.

See the study by Adams et al. in [34, 35] for details.

See [10, 15, 31–40].

5. Gauge Theories, the Standard Model, and Gravity

Here we will encounter various infinite-dimensional Lie groups and algebras such as diffeomorphism groups, loop groups, groups of gauge transformations, and their cohomologies.

5.1. Gauge Theories: Yang-Mills, QED, and QCD

Consider a principal G-bundle $\pi : P \to M$, with M being a compact, orientable Riemannian manifold (e.g., $M = S^4, T^4$) and G a compact non-Abelian gauge group with Lie algebra $\mathbf{g}$. Let $\mathcal{A}$ be the infinite-dimensional affine space of connection 1-forms on P. So each $A \in \mathcal{A}$ is a $\mathbf{g}$-valued, equivariant 1-form on P (also called vector potential) and defines the covariant derivative of any field φ by $D_A\varphi = d\varphi + (1/2)[A, \varphi]$. The curvature 2-form F_A (or field strength) is a $\mathbf{g}$-valued 2-form and is defined as $F_A = D_A A = dA + (1/2)[A, A]$. They are locally given by $A = A_\mu dx^\mu$ and $F = (1/2)F_{\mu\nu}dx^\mu \wedge dx^\nu$, where $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$.

In pure Yang-Mills theory the action functional is given by

$$S(A) = \frac{1}{2}\|F_A\|^2 = \frac{1}{2}\int_M \mathrm{Tr}\left(F_{\mu\nu}F^{\mu\nu}\right), \tag{5.1}$$

and the Yang-Mills equations become globally

$$d * F_A = 0. \tag{5.2}$$

With added fermionic field ψ interaction, the action becomes

$$S(A, \psi) = \frac{1}{2}\|F_A\|^2 + \langle \partial\!\!\!/_A \psi, \psi\rangle, \tag{5.3}$$

where ψ is a section of the spin bundle $\mathcal{S}pin^{\pm}(M)$ and, $\partial\!\!\!/_A : \mathcal{S}pin^{\pm}(M) \to \mathcal{S}pin^{\mp}(M)$ is the induced Dirac operator.

5.1.1. Gauge Invariance

In gauge theories the symmetry group is the group of gauge transformations. The diffeomorphism subgroups that arise in gauge theories as gauge groups behave nicely because they are isomorphic to subgroups of loop groups, as discussed in Section 2.4.4.

The group $\mathcal{G}$ of gauge transformations of the principal G-bundle $\pi : P \to M$ is given by

$$\begin{aligned} \mathcal{G} &= \left\{\phi \in \mathrm{Diff}^\infty(P);\ \phi(p \cdot g) = \phi(p) \cdot g,\ \pi\phi(p) = \pi(p)\right\} \\ &\cong \left\{\tau \in C^\infty(P, G);\ \tau(p \cdot g) = g^{-1}\tau(p)g\right\} = \mathrm{Gau}(P) \end{aligned} \tag{5.4}$$

which is a smooth Hilbert-Lie group with smooth group operations [6].

We only sketch here what role this infinite-dimensional gauge group $\mathcal{G}$ plays in these quantum field theories. A good reference for this topic is the study by Deligne et al. in [41, 42].

The gauge group $\mathcal{G}$ acts on $\mathcal{A}$ via pullback $\phi \in \mathcal{G}$, $A \in \mathcal{A}$, $\phi \cdot A = (\phi^{-1})^* A \in \mathcal{A}$, or under the isomorphism (see Section 2.4.4) $\mathcal{G} \cong \mathrm{Gau}(P)$, $\phi \Leftrightarrow \tau$ we have $\mathrm{Gau}(P)$ acting on $\mathcal{A}$ by $\tau \cdot A = \tau A \tau^{-1} + \tau d\tau^{-1}$. Hence the covariant derivative transforms as $D_{\tau\cdot A} = \tau D_A \tau^{-1}$, and the action on the field is $\tau \cdot F_A := F_{\tau\cdot A} = \tau F_A \tau^{-1}$.

The action functional (the Yang-Mills functional) is $S(A) = \|F_A\|^2$, locally given by $\|F_A\|^2 = (1/2)\int_M \mathrm{Tr}(F_{\mu\nu}F^{\mu\nu})$. This action is gauge invariant $S(\phi \cdot A) = S(A)$, $\phi \in \mathcal{G}$, so the

Yang-Mills functional is defined on the orbit space $\mathcal{M} = \mathcal{A}/\mathcal{G}$. The space $\mathcal{M}$ is in general not a manifold since the action of $\mathcal{G}$ on $\mathcal{A}$ is not free. If we restrict to irreducible connections, then $\mathcal{M}$ is a smooth infinite-dimensional manifold and $\mathcal{A} \to \mathcal{M}$ is an infinite-dimensional principal fiber bundle with structure group $\mathcal{G}$.

For self-dual connections $F_A = *F_A$ (instantons) on a compact 4-manifold, the moduli space $\mathcal{M} = \{A \in \mathcal{A}; A \text{ self-dual}\}/\mathcal{G}$ is a smooth finite-dimensional manifold. Self-dual connections absolutely minimize the Yang-Mills action integral

$$\mathbf{YM}(A) = \int_{\Omega} \|F_A\|^2, \quad \Omega \subset M \text{ compact}. \tag{5.5}$$

The Feynman path integral quantizes the action and we get the probability amplitude

$$W(f) = \int_{\mathcal{A}/\mathcal{G}} e^{-S(A)} f(A) \mathfrak{D}(A) \tag{5.6}$$

for any gauge-invariant functional $f(A)$.

Let $\mathcal{G}$ be the group of gauge transformations. So $\phi \in \mathcal{G} \Leftrightarrow \phi : P \to P$ is a diffeomorphism over id_M; that is, $\phi(p \cdot g) = \phi(p) \cdot g$, $p \in P$, $g \in G$. Then $\mathcal{G}$ acts on $\mathcal{A}$ and $\mathcal{S}pin^{\pm}(M)$ by $\phi \cdot A = (\phi^{-1})^* A$ and $\phi \cdot \psi = (\phi^{-1})^* \psi$. The action functionals S are gauge invariant:

$$\text{Yang-Mills: } S(\phi \cdot A) = S(A), \quad A \in \mathcal{A}, \phi \in \mathcal{G}, \tag{5.7}$$

$$\text{QED: } S(\phi \cdot A, \phi \cdot \psi) = S(A, \psi), \quad A \in \mathcal{A}, \psi \in \mathcal{S}pin^{\pm}(M), \phi \in \mathcal{G}. \tag{5.8}$$

5.1.2. Quantum Electrodynamics (QED) and Quantum Chromodynamics (QCD)

In classical field theory, one considers a Lagrangian $\mathcal{L}(\phi_i, \partial_\mu \phi_i)$ of the fields $\varphi_i : \mathbb{R}^n \to \mathbb{R}$, $i = 1, \dots, k$, and $\partial_\mu = \partial/\partial x_\mu$ and the corresponding action functional $S = \int \mathcal{L}(\phi_i, \partial_\mu \phi_i) d^n x$. The variational principle $\delta S = 0$ then leads to the Euler-Lagrange equations of motion

$$\frac{\partial \mathcal{L}}{\partial \phi_i} - \partial_\mu \frac{\partial \mathcal{L}}{\partial(\partial_\mu \phi_i)} = 0. \tag{5.9}$$

In QED and QCD the Lagrangian is more complicated of the form

$$\mathcal{L}(A, \psi, \varphi) = -\frac{1}{4g^2} \operatorname{Tr} F_{\mu\nu} F^{\mu\nu} - i\overline{\psi}[\gamma^\mu(\partial_\mu + ieA_\mu) + m]\psi + \left(D_A^\mu \varphi\right)^\dagger \left(D_A^\mu \varphi\right) - m^2 \varphi^\dagger \varphi, \tag{5.10}$$

where $A_\mu(x)$ is a potential 1-form (boson), and the field strength F is given by $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$. In QED the gauge group of the principal bundle is $G = U(1)$, and in QCD we have $G = SU(2)$. The Dirac γ-matrices are $\gamma^i = \left(\begin{smallmatrix} 0 & -\sigma_i \\ \sigma_i & 0 \end{smallmatrix}\right)$, where σ_i are the Pauli matrices (canonical basis of $\mathfrak{su}(2)$) and $\overline{\psi} = \psi^\dagger \gamma^o$ is the Pauli adjoint with $\gamma^o = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, m is the electron mass, e is the electron charge, and g is a coupling constant.

5.1.3. The Equations of Motion

The variational principle of the Lagrangian (5.10) with respect to the fields A, ψ, and φ gives the corresponding Euler-Lagrange equations of motion. They describe, for instance, the motion of an electron $\psi(x)$ (fermion, spinor) in an electromagnetic field F, interacting with a bosonic field φ. We get, from the variational principle, $\delta S/\delta A_\mu = 0 \Rightarrow \partial_\mu F^{\mu\nu} = e\overline{\psi}\gamma^\nu\psi$, which are Maxwell's equations for $G = U(1)$.

In the free case, that is, when $\psi = 0$, we get $\partial_\mu F^{\mu\nu} = 0$, the vacuum Maxwell equations.

For $G = SU(2)$ these equations become $D^\mu F_{\mu\nu} = 0$, the Yang-Mills equations. Moreover, $\delta S/\delta\psi = 0 \Rightarrow i(\not{\partial}_A - m)\psi = 0$, which are Dirac's equations, where $\not{\partial}_A = \gamma^\mu(\partial_\mu + ieA_\mu) = \gamma^\mu D_A^\mu$. In the free case, that is, when $A = 0$, we get $i(\not{\partial}_A - m)\psi = 0$, the classical Dirac equation.

5.1.4. Chiral Symmetry

The chiral symmetry is the symmetry that leads to anomalies and the BRST invariance. In QCD the chiral symmetry of the Fermi field ψ is given by $\psi \mapsto e^{i\beta\gamma_5}\psi$, where β is a constant and $\gamma_5 = i\gamma^o\gamma^1\gamma^2\gamma^3$. The classical Noether current of this symmetry is given by $J_\mu = \overline{\psi}\gamma_\mu\gamma_5\psi$ which is conserved; that is, $\partial^\mu J_\mu = 0$.

This conservation law breaks down after quantization; one gets

$$\partial^\mu J_\mu = 2\mathrm{im}\,\overline{\psi}\gamma_5\psi - \frac{g^2}{8\pi^2}\operatorname{Tr} F_{\mu\nu}F^{\mu\nu} \equiv \omega \neq 0. \tag{5.11}$$

This value ω is called the *chiral anomaly*.

5.2. Quantization

The quantization is given by the Feynman path integral:

$$\int\int_{\mathcal{A}/\mathcal{G}\times\mathcal{S}pin} e^{iS(A,\psi)}\mathcal{F}(A,\psi)\mathfrak{D}A\mathfrak{D}\psi = \langle\mathcal{F}(A,\psi)\rangle \tag{5.12}$$

which computes the expectation value $\langle\mathcal{F}(A,\psi)\rangle$ of the function $\mathcal{F}(A,\psi)$. This is an integral over two infinite-dimensional spaces: the gauge orbit space $\mathcal{A}/\mathcal{G}$ and the fermionic Berezin integral over the spin space $\mathcal{S}pin^\pm(M)$. These integrals are mathematically not defined but physicists compute them by gauge fixing; that is, fixing a section $\sigma : \mathcal{A}/\mathcal{G} \to \mathcal{A}$, (e.g., $\sigma(A) = \partial_\mu A^\mu = 0$, the Lorentz gauge) and then integrating over the section σ. Such a section does not exist globally, but only locally (Gribov ambiguity!). The effect of such a gauge fixing is that one gets extra terms in the Lagrangian (gauge-fixing terms) and one has to introduce new fields, so-called ghost fields η via the Faddeev-Popov procedure. The such obtained effective Lagrangian is no longer gauge invariant. This *effective Lagrangian* has the form in QCD:

$$\begin{aligned}
\mathcal{L}_{\text{eff}}(A,\psi,\eta) = {} & \frac{1}{2}\operatorname{Tr}(F_{\mu\nu}F^{\mu\nu}) && \text{kinetic energy} \\
& + \frac{1}{2\alpha}\operatorname{Tr}(\partial_\mu A^\mu)^2 && \text{gauge-fixing term} \\
& - g\partial_\mu \overline{\eta} D_A^\mu \eta && \text{ghost term} \\
& + \cdots && \text{interaction terms.}
\end{aligned} \tag{5.13}$$

We can write this globally as

$$\mathcal{L}_{\text{eff}} = \frac{1}{2}\|F_A\|^2 + \frac{1}{2}\|\sigma(A)\|^2 + \overline{\eta}\mathcal{M}\eta + \cdots, \tag{5.14}$$

where $\mathcal{M} = (\delta/\delta\phi)(\sigma(\phi\cdot A))$ is the Faddeev-Popov determinant, acting like the Jacobian of the global gauge variation $\delta/\delta\phi$ over the section σ. Writing this term in the exponent of the action functional like a "fermionic Gaussian integral" leads to the Faddeev-Popov ghost fields $\eta,\overline{\eta}$ in the form $\det\mathcal{M} = \int e^{-\overline{\eta}\mathcal{M}\eta} d\overline{\eta}\, d\eta$.

The effective Lagrangian $\mathcal{L}_{\text{eff}}$ is NOT gauge invariant but has a new symmetry, called BRST symmetry.

5.3. BRST Symmetry

Named after Becchi et al. [43] and Tyutin who discovered this invariance in 1975-76, the BRST operator $\mathfrak{s}$ is given as follows:

$$\left.\begin{aligned}
\mathfrak{s}A &= d\eta + [A,\eta] \\
\mathfrak{s}\eta &= -\frac{1}{2}[\eta,\eta]
\end{aligned}\right\} \quad \mathcal{L}_{\text{eff}} \text{ is BRST invariant.} \tag{5.15}$$

Note that the BRST operator $\mathfrak{s}$ mixes bosons (A) and fermions (η). This is an example of supersymmetry which we will discuss in Section 6. Also, the BRST operator $\mathfrak{s}$ is nilpotent; that is, $\mathfrak{s}^2 = 0$. The question arises whether this operator $\mathfrak{s}$ is the coboundary operator of some kind of cohomology. The affirmative answer is given by the following theorem (Schmid [6, 44]).

Theorem 5.1. *Let $C^{q,p}(\mathfrak{lie}\,\mathcal{G},\Omega_{\text{loc}})$ be the Chevalley-Eilenberg complex of the Lie algebra $\mathfrak{lie}\,\mathcal{G}$ of infinitesimal gauge transformations, with respect to the induced adjoint representation on local forms Ω_{loc}, with boundary operator*

$$\delta_{\text{loc}} : C^{q,p}(\mathfrak{lie}\,\mathcal{G},\Omega_{\text{loc}}) \longrightarrow C^{q+1,p}(\mathfrak{lie}\,\mathcal{G},\Omega_{\text{loc}}), \qquad \delta_{\text{loc}}^2 = 0. \tag{5.16}$$

Then with $\mathfrak{s} := (-1)^{p+1}/(q+1)\delta_{\text{loc}}$*, one has* $\mathfrak{s}^2 = 0$ *and the following.*

(1) *For* $q = 0,\ p = 1,\ A \in \mathcal{A} \subset \mathcal{C}^{0,1}$*, then* $\mathfrak{s}A = d\eta + [A, \eta]$*.*

(2) *For* $q = 1,\ p = 0,\ \eta \in \mathcal{C}^{1,0}$*, then* $\mathfrak{s}\eta = -(1/2)[\eta, \eta]$*, the Maurer-Cartan form.*

(3) *The chiral anomaly* ω *(given by* (5.11)*) is represented as cohomology class of this complex* $[\omega] \in \mathcal{H}^{1,0}_{\text{BRST}}(\mathfrak{lie}\,\mathcal{G}, \Omega_{\text{loc}})$*.*

5.3.1. The Chevalley-Eilenberg Cohomology

We are now going to explain the previous theorem, in particular the general definition of the Chevalley-Eilenberg [45] complex and the corresponding cohomology.

Let G be a Lie group with Lie algebra $\mathfrak{g}$ and let σ be a representation of $\mathfrak{g}$ on the vector space W. Denote by $\mathcal{C}^q(\mathfrak{g}, W)$ the space of W-valued q-cochains on $\mathfrak{g}$ and define the coboundary operator $\delta : \mathcal{C}^q(\mathfrak{g}, W) \to \mathcal{C}^{q+1}(\mathfrak{g}, W)$ by

$$\begin{aligned}\delta\Phi(\xi_o, \dots, \xi_q) = &\sum_{i=0}^{q}(-1)^i\sigma(\xi_i)\Phi\Big(\xi_o, \dots, \widehat{\xi}_i, \dots, \xi_q\Big)\\ &+\sum_{i<j}(-1)^{i+j}\Phi\Big(\sigma(\xi_i)\xi_j, \dots, \widehat{\xi}_i, \dots, \widehat{\xi}_j, \dots, \xi_q\Big).\end{aligned} \tag{5.17}$$

We have $\delta^2 = 0$, and define the Lie algebra cohomology of $\mathfrak{g}$ with respect to (σ, W) as $\mathcal{H}^*(\mathfrak{g}, W) = \ker\delta/\operatorname{im}\delta$. This is called the Chevalley-Eilenberg cohomology [45] of the Lie algebra $\mathfrak{g}$ with respect to the representation σ.

5.3.2. Anomalies

The Noether current induced by the chiral symmetry (after quantization) for the free case $(\psi = 0)$, that is, for pure Yang-Mills becomes

$$\begin{aligned}\partial^\mu J_\mu &= -\frac{g^2}{8\pi^2}\varepsilon^{\mu\nu\rho\tau}\operatorname{Tr} F_{\mu\nu}F_{\rho\tau}\\ &= -\frac{1}{4}\pi^2\varepsilon^{\mu\nu\rho\tau}\operatorname{Tr}\partial_\mu\left(A_\nu\partial_\rho A_\tau + \frac{2}{3}A_\mu A_\rho A_\tau\right)\\ &= \omega \neq 0 \quad \text{anomaly}.\end{aligned} \tag{5.18}$$

See (5.11).

Note the similarity with the Chern-Simon Lagrangian

$$\mathcal{L}(A) = \int_M \operatorname{Tr}\left(Ad\,A + \frac{2}{3}A^3\right). \tag{5.19}$$

We are going to derive a representation of the chiral anomaly ω in the BRST cohomology that is $[\omega] \in \mathcal{H}^{1,0}_{\text{BRST}}(\mathfrak{lie}\,\mathcal{G}, \Omega_{\text{loc}})$.

The question is "if $\mathfrak{s}\omega = 0$, does there exist a *local* functional $F(A)$, such that $\omega = \mathfrak{s}(F(A))$? That is, is ω BRST $\mathfrak{s}$-exact? The answer in general is NO; that is, ω represents a nontrivial cohomology class. This class is given by the *Chern-Weil homotopy*.

Let $\widetilde{A} = A + \eta \in \mathcal{C}^{0,1} \times \mathcal{C}^{1,0}$ and $\widetilde{F} \equiv s\widetilde{A} + \widetilde{A}^2 = F_{\widetilde{A}}$. For $t \in [0,1]$, let $\widetilde{F}_t = t\widetilde{F} + (t^2 - t)\widetilde{A}^2$ and define the Chern-Simons form

$$\omega_{2q-1} \equiv q \int_0^1 \mathrm{Tr}\left(\widetilde{A}\widetilde{F}_t^{q-1}\right) dt, \tag{5.20}$$

we get

$$\mathfrak{s}\omega_{2q-1} = \mathrm{Tr}\, \widetilde{F}^q. \tag{5.21}$$

We write ω_{2q-1} as sum of homogeneous terms in ghost number (upper index) and degree (lower index) $\omega_{2q-1} = \omega^0_{2q-1} + \omega^1_{2q-2} + \omega^2_{2q-3} + \cdots + \omega^{2q-1}_0$. Let $\omega(X, A) = \int_M \omega^1_{2q-2}(X)$.

Theorem 5.2 (see Schmid [46]). *The form* $\omega(X, A) = \int_M \int_0^1 \widetilde{A}\widetilde{F}_t^{q-1}(X)dt$ *satisfies the Wess-Zumino consistency condition* $(\mathfrak{s}\omega)(X_0, X_1, A) = 0$ *and represents the chiral anomaly* $[\omega] \in \mathscr{H}^{1,0}_{\mathrm{BRST}}(\mathfrak{lie}\,\mathcal{G}, \Omega_{\mathrm{loc}})$.

We have an explicit form of the anomaly in $(2q - 2)$ dimensions:

$$\omega^1_{2q-2} = q(q-1) \int_0^1 (1-t)\, \mathrm{Tr}\left(\eta\delta_{\mathrm{loc}}\left(\widetilde{A}\widetilde{F}_t^{q-2}\right)\right) dt. \tag{5.22}$$

So for $q = 2$ the non-Abelian anomaly in 2 dimensions becomes $\omega^1_2 = \mathrm{Tr}(\eta\delta_{\mathrm{loc}}\widetilde{A})$, and for $q = 3$ the non-Abelian anomaly in 4 dimensions becomes

$$\omega^1_4 = \mathrm{Tr}\left(\eta\delta_{\mathrm{loc}}\left(\widetilde{A}\delta_{\mathrm{loc}}\widetilde{A} + \frac{1}{2}\widetilde{A}^3\right)\right). \tag{5.23}$$

5.4. The Standard Model

The standard model is a Yang-Mills gauge theory. Recall that the free *Yang-Mills* equations are $D^*_A F = 0$, where A is a connection 1-form (vector potential), and F is the associated curvature 2-form (field) on the principal bundle P. The connection A defines the covariant derivative D_A and the curvature F given by $F = D_A A = dA + (1/2)[A, A]$, or locally $F_{\mu\nu} = \partial_\mu A_\nu - \partial_\nu A_\mu + [A_\mu, A_\nu]$, and $D^\mu F_{\mu\nu} = 0$. Again the connection A is the fundamental object.

For different choices of the gauge Lie group G, we obtain the 3 theories that make up the standard model. For $G = U(1)$ on a trivial bundle (i.e., global symmetry, which gives charge conservation) the curvature 2-form F is simply the electromagnetic field, and the Yang-Mills equations $D^*_A F = 0$ are *Maxwell's equations* $dF = 0$, locally $\partial_\mu F^{\mu\nu} = 0$. For $G = U(1)$ as local gauge group we get the quantum mechanical symmetry and the equations of motion are *Dirac's equations*. Combing the two, we get QED as a $U(1)$ gauge theory. For $G = SU(N)$ we get the full non-Abelian Yang-Mills equations $D^*_A F = 0$. For weak interactions with $G = SU(2)$ and combining the two (spontaneous symmetry breaking, Higgs), we get the

Glashow-Weinberg-Salam model as $SU(2)\times U(1)$ Yang-Mills theory of *electroweak interactions*. For $G = SU(3)$ we obtain the Yang-Mills equations $D_A^* F = 0$ for *strong interactions* and the equations of motion for QCD. Finally that standard model is a $SU(3) \times SU(2) \times U(1)$ gauge theory governed by the corresponding Yang-Mills equations $D_A^* F = 0$. Recall that F is the curvature in the corresponding principal bundle determined by the connection A.

For interactions, all the relevant fields involved can be considered as sections of corresponding associated vector bundles induced by representations of the gauge groups, for example, the Dirac operator on the associated spin bundle (induced by the spin representation of $SU(2)$) acting on spinors (sections of this bundle). The vector potentials are the corresponding connection 1-forms and the Yang-Mills fields are the corresponding curvature 2-forms on these bundles over spacetime.

Again we do not need the metric and the curvature is determined by the potential, so the potential is the fundamental object.

5.5. Gravity

5.5.1. Stop Looking for Gravitons

Stop looking for the graviton, not because it had been found but because it does not exist. The graviton is supposed to be the particle that communicates the gravitational force. But the gravitational force is not a fundamental force. *Gravity is geometry*. One might as well search for the Corioliston for the coriolis force or the Centrifugiton for the centrifugal force.

Since Einstein in the 1920s, physicists have tried to unify what are considered the four fundamental forces, namely, electromagnetism, weak and strong nuclear forces, and the gravitational force. In the 1970s, the three nongravitational forces were unified in the standard model. At high enough energy (about 10^{15} GeV) they become the same force.

Since then, with all the string theory, SUSY, branes, and extra dimensions, the gravitational force could not be incorporated into GUT that includes all 4 forces and no graviton has been found experimentally. The reason is simple: not many people, including Einstein himself, take/took the general theory of relativity seriously enough, according to which we know that the gravitational force does not exist as fundamental force but as geometry! We do not feel it. What we feel is the resistance of the solid ground on which we stand. In general relativity, free-falling objects follow geodesics of spacetime, and what we perceive as the force of gravity is instead a result of our being unable to follow those geodesics because of the mechanical resistance of matter. Newton's apple falls downward because the spacetime in which we exist is *curved*. The "gravitational force" is *not* a force but it is the *geometry of spacetime* as Einstein observed in [47, page 137]:

> *"Die Koeffizienten ($g_{\mu\nu}$) dieser Mertik beschreiben in Bezug auf das gewählte Koordinatensystem zugleich das Gravitationsfeld."*

("The coefficients ($g_{\mu\nu}$) of this metric with respect to the chosen coordinate system describe at the same time the gravitational field") [47, page 146]:

> *"Aus pysikalischen Gründen bestand die Überzeugung, dass das metrische Feld zugleich das Gravitationsfeld sei."*

("For physical reasons there was the conviction that the metric field was at the same time the gravitational field").

Therefore GUT, the grand unified theory had been completed since the 1970s with the standard model. Since the gravitational force does not exist as a fundamental force, there is nothing more to unify as forces. If we want to unify all four theories, then it has to be done in a geometric way. The equations governing gravity as well as the standard model are all *curvature* equations, Einstein's equation, and the Yang-Mills equations.

5.5.2. Einstein's Vacuum Field Equations

Let (M, g) be spacetime with Lorentzian metric g. Then *Einstein's vacuum field equations* are

$$\mathrm{Ric} = 0, \tag{5.24}$$

where Ric is the Ricci curvature of the Lorentz metric g. These are the Euler-Lagrange equations for the Lagrangian $\mathcal{L}(g) = \int R(g)\mu(g)$, where $\mu(g) = \sqrt{-\det g}d^4x$ and $R(g)$ is the scalar curvature of g.

Or in general, locally, in terms of the stress-energy tensor $T_{\mu\nu}$, Einstein's equations are $G_{\mu\nu} = \kappa T_{\mu\nu}$ with the Einstein tensor $G_{\mu\nu} = R_{\mu\nu} - (1/2)g_{\mu\nu}R$. The stress-energy tensor $T_{\mu\nu}$ is the conserved Noether current corresponding to spacetime translation invariance.

The Levi-Civita connection Γ of the Riemannian metric g is given by

$$\Gamma^{\lambda}_{\mu\nu} = \frac{1}{2}g^{\lambda\sigma}\left(\frac{\partial g_{\sigma\mu}}{\partial x^{\nu}} + \frac{\partial g_{\sigma\nu}}{\partial x^{\mu}} - \frac{\partial g_{\mu\nu}}{\partial x^{\sigma}}\right). \tag{5.25}$$

The curvature tensor R and the Ricci curvature Ric in Einstein's field equations are completely determined by the connection Γ.

First the curvature tensor R is locally given by

$$R^{\lambda}_{\mu\nu\kappa} = \left(\frac{\partial\Gamma^{\lambda}_{\mu\nu}}{\partial x^{\kappa}} - \frac{\partial\Gamma^{\lambda}_{\mu\kappa}}{\partial x^{\nu}}\right) + \left(\Gamma^{\eta}_{\mu\nu}\Gamma^{\lambda}_{\kappa\eta} - \Gamma^{\eta}_{\mu\kappa}\Gamma^{\lambda}_{\nu\eta}\right). \tag{5.26}$$

Taking its trace, we get the Rici tensor Ric as $(\mathrm{Ric})^{\mu\nu} = R^{\lambda}_{\mu\lambda\nu}$.

So we can express Einstein's equations completely in terms of the connection (potential) Γ; we do not need the metric g; also the curvature R is determined by the potential Γ. So the potential Γ is the fundamental object.

The free motion in spacetime is along geodesic curves $\gamma(t)$ which again are expressed in terms of the connection by

$$\ddot{\gamma}^{\alpha} + \Gamma^{\alpha}_{\beta\nu}\dot{\gamma}^{\beta}\dot{\gamma}^{\nu} = 0. \tag{5.27}$$

5.5.3. Symmetry

In general relativity the diffeomorphism group plays the role of a symmetry group of coordinate transformations. Then the vacuum Einstein's field equations $\mathrm{Ric}(g) = 0$ are invariant under coordinate transformations, that is, under the action of $\mathrm{Diff}^{\infty}(M)$. Denote by $\mathcal{M}$ the space of all metrics g on M. Then, Einstein's field equations $\mathrm{Ric}(g) = 0$ are a Hamiltonian system on the reduced space $P = M/\mathrm{Diff}^{\infty}(M)$; see the study by Marsden et al. in [15] for details.

5.6. Conclusions

The relation between the connection Γ occurring in Einstein's equations and the connection A in Yang-Mills equations is as follows. A is a Lie algebra-valued 1-form on a principal G-bundle (P, π, M) over spacetime (M, g) (or any associated vector bundle given by representations of G). The Levi-Civita connection Γ is a connection 1-form in this sense on the tangent bundle TM (frame bundle) with $\mathrm{G} = GL(n)$. So in this sense general relativity and the standard model are Yang-Mills gauge theories.

Therefore all four theories, electromagnetism, weak interaction, strong interaction, and gravity, are unified as curvature equations in vector bundles over spacetime. Different interactions require different bundles.

There is no hierarchy problem because there is no fundamental gravitational force. The question why gravitational interaction is so much weaker than electroweak and strong interactions is meaningless, comparing apples with oranges. Why are so many physicists still talking about gravitational force? It is like as if we are still talking about "sun rise" and "sun set", 500 years after Copernicus! only worse; these are serious scientists trying to unify all four "forces" to a TOE.

I am not saying that there are no open problems in physics. Of course there is still the problem of unifying quantum mechanics and general relativity on a geometric level (not as forces). The question is "how does spacetime look at the Planck scale? Do we have to modify spacetime to incorporate quantum mechanics or quantum mechanics to accommodate spacetime, or both? We need a theory of quantum gravity. There are several theories in the developing stage that promise to accomplish this.

(i) Superstring theory by E. Witten et al.

(ii) Discrete spacetime at Planck length by R. Loll in "Causal dynamical triangulation" and by J. Ambjorn, J. Jurkiewicz, and R. Loll in "The Universe from Scratch" [arXiv: hep-th/0509010].

(iii) Spacetime quantization: loop quantum gravity by L. Smolin in "Three Roads to Quantum Gravity" (London: Weidenfeld and Nicholson, 2000) and by S. O. Bilson-Thompson, F. Markopoulou, and L. Smolin in "Quantum Gravity and the Standard Model", preprint 2006.

(iv) Geometric formulation of quantum mechanics by A. Ashtekar and T. A. Schilling [arXiv: gr-qc/9706069].

(v) Deterministic quantum mechanics at Planck scale by G. t'Hooft in "Quantum Gravity as a Dissipative Deterministic System."

(vi) Branes and new dimensions: parallel universes by L. Randell et al., D. Deutsch, PS. In the brane world, gravity is again singled out as the only force not confined to one brane.

(vii) Noncommutative Geometry. A. Connes describes the standard model form general relativity.

(viii) The most recent new development is by Verlinde [48]. He agrees that gravity is not a fundamental force, but explains it as an emergent force (entropic force) caused by a change in the amount of information (entropy) associated with the positions of bodies of matter.

See [6, 11, 27, 36, 41–59].

6. SUSY (Supersymmetry)

Supersymmetry (SUSY) is an important idea in quantum filed theory and string theories. The BRST symmetry we described in Section 5.3 is an example of SUSY. Now we give a summary of a mathematical description of super Hamiltonian systems on supersymplectic supermanifolds, state a generalization of the Marsden-Weinstein reduction theorem in this context, and illustrate the method with examples. This is a very technical topic, so we only give a brief sketch; for details see the studies Glimm in [60] and Tuynman in [61].

The classical Marsden-Weinstein reduction theorem is a geometrical result stating that if a Lie group G acts on a symplectic manifold P by symplectomorphisms and admits an equivariant momentum map $J : P \to \mathfrak{g}^*$, then, for any regular value $\mu \in \mathfrak{g}$ of J, the quotient $P_\mu = J^{-1}(\mu)/G_\mu$ of the preimage $J^{-1}(\mu)$ by the isotropy group G_μ of μ has a natural symplectic structure. A dynamical interpretation of the Marsden-Weinstein reduction theorem gives the following. If a given Hamiltonian $H \in C^\infty(P)$ is invariant under the action of the group G, then it projects to a reduced Hamiltonian H_μ on the reduced space P_μ. The integral curves of the Hamiltonian vector field X_H project to the integral curves of X_{H_μ}. In this sense, one has reduced the system by symmetries. This reduction procedure unifies many methods and results concerning the use of symmetries in classical mechanics, some dating back to the time of Euler and Lagrange.

In [60], Glimm generalizes this result to the setting of supermanifolds, using the analytic construction of supercalculus and supergeometry due to DeWitt [62] and Tuynman [61]. The general idea of supermanifolds and superanalysis is to do geometry and analysis over a graded algebra of even supernumbers rather than $\mathbb{R}$. In "supermathematics," we have even and odd variables. Two variables a, b are called *even*, or commuting, or bosonic if $a \cdot b = b \cdot a$, whereas ξ, χ are called *odd*, or anticommuting, or fermionic if $\xi \cdot \chi = -\chi \cdot \xi$. The problem is doing analysis with even and odd variables.

Many classes of differential equations have extensions that involve odd variables. These are called *superized* versions, or *supersymmetric* extensions. An active area of research is in particular the construction of supersymmetric integrable systems.

As an example, consider the Korteweg-de Vries equation $u_t = -u_{xxx} + 6uu_x$. A possible supersymmetric extension is the following system of an even variable $u(t, x)$ and an odd variable $\xi(t, x)$:

$$\begin{aligned} u_t &= -u_{xxx} + 6uu_x - 3\xi\xi_{xx}, \\ \xi_t &= -\xi_{xxx} + 3u\xi_x + 3\xi u_x. \end{aligned} \tag{6.1}$$

There is a different way of writing system (6.1). For this, one considers the so-called 2|1-dimensional *superspace*. This is a space with coordinates (x, t, ϑ), where x and t are even numbers as before and ϑ is an odd number. We can now gather the components $u(x,t)$ and $\xi(x,t)$ into a *superfield* $\Phi(t,x,\vartheta)$ defined on superspace via

$$\Phi(t,x,\vartheta) = \xi(t,x) + \vartheta \cdot u(t,x). \tag{6.2}$$

The function $\Phi(t,x,\vartheta)$ takes as values odd numbers; it is thus called an odd function. Note how the right-hand side can be seen as a Taylor series in ϑ; indeed, all higher powers of ϑ are zero because ϑ is anticommuting. One defines the odd differential operator $D = \partial/\partial\vartheta + \vartheta \cdot (\partial/\partial x)$, acting on superfields. A computation yields that system (6.1) is equivalent to

$$\Phi_t = -D^6\Phi + 3D^2\Phi D\Phi. \tag{6.3}$$

(Note that $D^2 = \partial_x$.) System (6.1) is called the *component formulation*; (6.3) is called the *superspace formulation*. In our example at hand, a justification for calling the system an "extension" of KdV would be that if one takes (6.3) and writes it in component form (6.1), one recovers the "usual" KdV by setting ξ to zero. Also, it can be shown that (6.3) is invariant under transformations $\Phi(t,x,\vartheta) \mapsto \Phi(t,x-\eta\vartheta,\vartheta+\eta)$, where η is an odd parameter. The infinitesimal version of this transformation is $\delta\Phi = \eta(\partial_\vartheta - \vartheta\partial_x)\Phi$. This transformation is called a "supersymmetry" in the present context. In components, it reads

$$\delta u = \eta\xi_x, \qquad \delta\xi = \eta u. \tag{6.4}$$

These equations illustrate a characterization of supersymmetries: supersymmetries (as opposed to regular symmetries) "mix" even and odd variables.

One needs some concept of supermanifold even if one only works with the component formulation. For example, one has implicitly in system (6.1) the space of all $(u(t,x),\xi(t,x))$ on which the equations are defined; this is some kind of superspace itself. Also, there is some kind of supersubmanifold of those (u,ξ) which solve the equations.

In [60], Glimm proves a comprehensive result on supersymplectic reduction. He uses an analytic-geometric approach to the theory of supermanifolds, and not the Kostant theory of graded manifolds [49]. The Poisson bracket induced by odd supersymplectic forms is *not* a super Lie bracket on the space of supersmooth functions. This stands of course in contrast to both the usual ungraded case and the super case with even supersymplectic forms. While this makes the algebraic approach conceptually more difficult, no such problems arise in the analytic approach. Also, we do not require that the action be free and proper, but have the weaker requirement that the quotient space only has a manifold structure.

There are different approaches to supermanifolds. The *Algebraic approach* (Kostant [49], Berezin-Leĭtes, late 1970s) takes "superfunctions" as fundamental object. A *graded manifold* is a pair (M,A), where M is a conventional manifold and $U \mapsto A(U)$ is the following sheaf over M:

$$A(U) \simeq C^\infty(U) \otimes \bigwedge \mathbb{R}^n, \quad U \subseteq M. \tag{6.5}$$

So a superfunction $f \in A(U)$ can be written as

$$f = f_\emptyset + \sum_i f_{\{i\}} \vartheta^i + \sum_{i<j} f_{\{i,j\}} \vartheta^i \vartheta^j + \cdots . \tag{6.6}$$

The *Geometric approach* (DeWitt [62]) takes points as fundamental objects. Let $\mathcal{A}$ be the ring generated by L generators $\theta^1, \theta^2, \ldots, \theta^L$ (Grassmann numbers) with relations $\theta^i \theta^j = -\theta^j \theta^i$.

An element is written as

$$\begin{gathered} a = \underbrace{a_\emptyset}_{=\mathrm{B}a\ (\text{“Body”})} \underbrace{+ \sum_i a_{\{i\}} \theta^i + \sum_{i<j} a_{\{i,j\}} \theta^i \theta^j + \cdots}_{\text{nilpotent part}} \quad (a_\omega \in \mathbb{R}), \\ a \in \mathcal{A}_0 \Longleftrightarrow a = a_\emptyset + \sum a_{\{i,j\}} \theta^i \theta^j + \text{terms with even \# of } \theta\text{s}, \\ a \in \mathcal{A}_1 \Longleftrightarrow a = \sum a_{\{i\}} \theta^i + \sum a_{\{i,j,k\}} \theta^i \theta^j \theta^k + \text{terms with odd \# of } \theta\text{s}. \end{gathered} \tag{6.7}$$

For calculus we "replace" reals $\mathbb{R}$ by Grassmann numbers $\mathcal{A}$. A *DeWitt supermanifold* is a topological space M which is locally superdiffeomorphic to $\mathcal{A}^{m|n} \overset{\text{def}}{=} (\mathcal{A}^{m \oplus n})_0 = \mathcal{A}_0^m \times \mathcal{A}_1^n$.

The two approaches are equivalent. There is a one-to-one correspondence between isomorphism classes of $m|n$-dimensional DeWitt supermanifolds whose body is a fixed smooth m-dimensional manifold X and isomorphism classes of $m|n$-dimensional graded manifolds over X.

The DeWitt topology of $\mathcal{A}^{m|n}$ is defined as follows.

$$U \subseteq \mathcal{A}^{m|n} \text{ is open if and only if } B(U) \text{ is open in } B\mathcal{A}^{m \oplus n} = \mathbb{R}^m,\ U = B^{-1}(B(U)).$$

Smooth functions $\mathcal{A}^{m|n} \to \mathcal{A}$ are defined as follows. Let $U \subseteq \mathcal{A}^{m|n}$ be an open set. A function $f : U \to \mathcal{A}$ is called *smooth* if there is a collection of smooth real functions defined on $BU \subseteq \mathbb{R}^m$:

$$f_{i_1 \ldots i_n} : BU \to \mathbb{R}, \quad \text{for } i_1, \ldots, i_n = 0, 1 \tag{6.8}$$

such that

$$f(x_1, \ldots, x_m, \xi_1, \ldots, \xi_n) = \sum_{i_1, \ldots, i_n = 0}^{1} \xi_1^{i_1} \ldots \xi_n^{i_n} \cdot Z f_{i_1 \ldots i_n}(x_1, \ldots, x_m), \tag{6.9}$$

where Zg is defined as follows. If $\mathbf{x} \in \mathcal{A}^{m|0}$ has the decomposition $\mathbf{x} = B\mathbf{x} + \mathbf{n}$, then

$$Zg(\mathbf{x}) = \sum_{k=1}^{\infty} \frac{1}{k!} D^k g(B\mathbf{x})[\mathbf{n}, \ldots, \mathbf{n}]. \tag{6.10}$$

Note. Smooth functions map the body to the body!

Superdifferential Geometry

Versions of the inverse function theorem and implicit function theorem are still valid. Concepts of tangent space, vector fields, flows, and Lie groups can be developed as in the ungraded case.

Lie Supergroups

A Lie supergroup is a supermanifold G that is a group and for which the group operations of multiplication and inversion are smooth. If a Lie supergroup G acts freely and properly on a supermanifold M, then the quotient M/G can be given the structure of a supermanifold such that the projection $\pi : M \to M/G$ is a surjective submersion. The structure of $T(M/G)$ is given by the following.

Theorem 6.1. *Let $\varphi : M \times G \to M$ be an action. Suppose that M/G has a supermanifold structure such that $\pi : M \to M/G$ is a surjective submersion; that is, $T\pi(p)$ is onto for every $p \in M$. For any $p \in M$, one has*

$$\ker T\pi(p) = \{\mathfrak{X}_M(p) \mid \mathfrak{X} \in \mathfrak{g}\} \tag{6.11}$$

and this is a proper subspace of T_pM. In particular,

$$T_{[p]}(M/G) \simeq T_pM/\ \{\mathfrak{X}_M(p) \mid \mathfrak{X} \in \mathfrak{g}\}. \tag{6.12}$$

Remark 6.2. There are examples where M/G does *not* have a supermanifold structure and $\{\mathfrak{X}_M(p) \mid \mathfrak{X} \in \mathfrak{g}\}$ fails to be a free submodule.

Supersymplectic Structures

A supersymplectic supermanifold (M, ω) is a supermanifold M together with a closed (i.e., $d\omega = 0$) nondegenerate homogeneous left 2-form $\omega \in \Omega^2_L(M)$.

Examples 6.3. (1) $\mathscr{A}^{2m|n}$ with coordinates (q^i, p^i, ξ^j),

$$\omega = \sum dq^i \wedge dp^i + \frac{1}{2}\sum d\xi^j \wedge d\xi^j \tag{6.13}$$

defines an *even* supersymplectic form.

(2) On $\mathscr{A}^{m|m}$ with coordinates (x^i, ξ^j),

$$\omega = \sum dx^i \wedge d\xi^i \tag{6.14}$$

defines an *odd* supersymplectic form.

Let $\omega \in \Omega^2_L(M)$ be a 2-form on the supermanifold M and $p \in M$. Then $\omega(p) \in \mathrm{Alt}^2_L(T_pM)$ is nondegenerate if and only if the real 2-form $B\omega(Bp) = \omega(Bp)|_{B(T_{Bp}M)} \in \mathrm{Alt}^2(B(T_{Bp}M))$ is nondegenerate.

Hamiltonian Supermechanics

A smooth vector field $X \in \mathfrak{X}(M)$ is called (globally) *Hamiltonian* if there is some function $H \in C^\infty(M, \mathcal{A})$ such that $i_X \omega = dH$. For $f, g \in C^\infty(M, \mathcal{A})$ define the Super-Poisson bracket by

$$\{f, g\} = \langle X_f \ , X_g \mid \omega \rangle \in C^\infty(M, \mathcal{A}). \tag{6.15}$$

Fact. If ω is even, then $C^\infty(M, \mathcal{A})$ is a Lie superalgebra with respect to $\{\cdot, \cdot\}$. This is *false* if ω is odd.

Momentum Maps

Let $\varphi : M \times G \to M$ be an action of the Lie supergroup G on the supersymplectic manifold (M, ω) which preserves ω. Recall that, in the ungraded case, a momentum map is an $\mathbb{R}$-linear map $\widehat{J} : \mathfrak{g} \to C^\infty(M)$ such that $X_{\widehat{J}(\mathfrak{X})} = \mathfrak{X}_M$.

Superversion. $\widehat{J} \in C^\infty(\mathfrak{g} \times M, \mathcal{A})$ is a *momentum map* for the action of G on M if $\widehat{J}$ is leftlinear in the first argument and

$$\left\langle (0, v) \mid d\widehat{J}(\mathfrak{X}, x) \right\rangle = \langle v, \mathfrak{X}_M(x) \mid \omega(x) \rangle \tag{6.16}$$

for all $x \in M$, for all $v \in T_x M$. Instead of $\widehat{J} \in C^\infty(\mathfrak{g} \times M, \mathcal{A})$, one can consider $\mathbf{J} : M \to \mathfrak{g}^*$ defined through $\langle \mathfrak{X} \mid \mathbf{J}(x) \rangle = \widehat{J}(\mathfrak{X}, x)$.

Theorem 6.4. *Let $H \in C^\infty(M, \mathcal{A})$ be a Hamiltonian with vector field X_H such that $[X_H, X_H] = 0$. If H is G-invariant, then $\mathbf{J}$ is preserved by the flow ϕ of X_H; that is,*

$$\mathbf{J} \circ \phi_{t,\tau} = \mathbf{J} \quad \forall (t, \tau) \in \mathcal{A}^{1|1} \ (\textit{where defined}). \tag{6.17}$$

Suppose that the momentum map is Ad^*-equivariant. Let $\mu \in B\mathfrak{g}^*_{p(\omega)}$ be a regular value of $\mathbf{J}$. Let G_μ be the isotropy group of μ. Suppose that the quotient space $P_\mu = \mathbf{J}^{-1}(\mu)/G_\mu$ can be given a supermanifold structure such that the projection $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ is a surjective submersion.

Superreduction Theorem

The supermanifold P_μ has a unique supersymplectic form $\omega_\mu \in \Omega^2_L(P_\mu)$ with the property

$$\pi^*_\mu \omega_\mu = \imath^*_\mu \omega, \tag{6.18}$$

where $\pi_\mu : \mathbf{J}^{-1}(\mu) \to P_\mu$ is the projection and $\imath_\mu : \mathbf{J}^{-1}(\mu) \to M$ is the inclusion. The form ω_μ has the same parity as ω.

Example 6.5 (The Bose-Fermi Oscillator). We consider the phase space $P = \mathcal{A}^{2|2}$ with supersymplectic form $\omega = dp \wedge dq + (1/2)(d\xi_1 \wedge d\xi_1 + d\xi_2 \wedge \delta\xi_2)$. The Hamiltonian $H(q,p,\xi_1,\xi_2) = (1/2)(p^2+q^2) - \xi_1\xi_2$ defines the Hamiltonian vector field $X_H = p\partial_q - q\partial_p - \xi_2\partial_{\xi_1} + \xi_1\partial_{\xi_2}$.

The SUSY algebra of the Bose-Fermi oscillator is given as follows. The Lie supergroup of invertible 2|2-supermatrices acts on $\mathcal{A}^{2|2}$ via

$$\mathcal{A}^{2|2} \times GL(2|2) \longrightarrow \mathcal{A}^{2|2}, \qquad (\mathbf{q}, \mathrm{G}) \longmapsto \mathbf{q}^{ST} G. \tag{6.19}$$

The algebra of Bose-Fermi supersymmetry is the intersection of the Lie superalgebras of the stabilizer of ω and the stabilizer of H:

$$\mathfrak{bf}(2|2) = \mathfrak{stab}(H) \cap \mathfrak{stab}(\omega). \tag{6.20}$$

The SUSY algebra of the Bose-Fermi oscillator $\mathfrak{bf}(2|2)$ is generated by A_1, A_2, C_1, C_2, where

$$\begin{gathered} A_i = \begin{pmatrix} 0 & 0 & \mathbf{e}_i^T\boldsymbol{\gamma}^0 \\ 0 & 0 & -\mathbf{e}_i^T \\ \mathbf{e}_i & -\boldsymbol{\gamma}^0\mathbf{e}_i & \mathbf{0} \end{pmatrix} \quad \text{for } i = 1,2, \\ C_1 = \frac{1}{2}\begin{pmatrix} -\boldsymbol{\gamma}^0 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\gamma}^0 \end{pmatrix}, \quad C_2 = \frac{1}{2}\begin{pmatrix} \boldsymbol{\gamma}^0 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\gamma}^0 \end{pmatrix}, \quad \boldsymbol{\gamma}^0 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}. \end{gathered} \tag{6.21}$$

Momentum Map and Quotient

The action of $BF(2|2)$ on $\mathcal{A}^{2|2}$ admits a momentum map

$$\widehat{J} : \mathfrak{bf}(2|2) \times \mathcal{A}^{2|2} \longrightarrow \mathcal{A}, \qquad (\mathrm{X}, \mathbf{q}) \longmapsto \frac{1}{2}\mathbf{q}^{ST}\mathrm{X}\Omega\mathbf{q}. \tag{6.22}$$

In components, we write $X = \sum_{i=1,2} a_iA_i + c_iC_i$. Then

$$\widehat{J}(X,q,p,\boldsymbol{\xi}) = (c_2 - c_1)\frac{1}{4}\left(p^2+q^2\right) - (c_2+c_1)\frac{1}{4}\boldsymbol{\xi}^T\boldsymbol{\gamma}^0\boldsymbol{\xi} - \mathbf{a}^T\left(p\mathbb{I}_2 - q\boldsymbol{\gamma}^0\right)\boldsymbol{\xi}, \tag{6.23}$$

where $\mathbf{a} = (a_1, a_2)^T$.

Now let $\boldsymbol{\mu} \in B(\mathfrak{bf}(2|2)_0^*) \simeq \mathbb{R}^2$. The isotropy group of $\boldsymbol{\mu}$ is the whole group $BF(2|2)$. The action of $BF(2|2)$ on $\mathbf{J}^{-1}(\boldsymbol{\mu})$ is transitive; that is, the quotient space $\mathbf{J}^{-1}(\boldsymbol{\mu})/BF(2|2)$ is a single point.

Example 6.6 (Wess-Zumino Model in 2+1-Dimensional Spacetime). Let $\mathcal{S}(\mathcal{A})$ be the Schwartz space of functions $\mathbb{R}^2 \to \mathcal{A}$. Consider the phase space $\mathcal{S}(\mathcal{A})^{2|2} = [\mathcal{S}(\mathcal{A})_0] \times [\mathcal{S}(\mathcal{A})_0] \times [\mathcal{S}(\mathcal{A})_1]^2 \ni (\phi, \pi, \psi)$ with supersymplectic form $\omega = d\pi \wedge d\phi + (1/2) d\psi \wedge d\psi$ and Hamiltonian

$$\mathcal{H}(\phi, \pi, \psi) = \frac{1}{2} \int_{\mathbb{R}^2} \left(\|\nabla \phi\|^2 + \pi^2 + \overline{\psi} \mathfrak{g}^i d_i \psi \right) d^2 x, \tag{6.24}$$

where $\mathfrak{g}^1 = \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right)$, $\mathfrak{g}^2 = \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right)$, and $\overline{\psi} = \psi^T \gamma^0$.

Then Hamilton's equations are the following:

$$\dot{\phi} = \frac{\delta \mathcal{H}}{\delta \pi} = \pi, \qquad \dot{\pi} = -\frac{\delta \mathcal{H}}{\delta \phi} = \Delta \phi, \qquad \dot{\psi} = \frac{\delta \mathcal{H}}{\delta \psi} = \mathfrak{g}^0 \mathfrak{g}^i \partial_i \psi. \tag{6.25}$$

This is equivalent to $(\partial_0^2 - \partial_1^2 - \partial_2^2)\phi = 0$, $\mathfrak{g}^\mu \partial_\mu \psi = 0$ which are the massless Klein-Gordon and Dirac equations in $2+1$-dimensional spacetime.

Remark 6.7. The well-known SUSY algebra from the Lagrangian description can be "exported" to the Hamiltonian setup.

We reduce by an Abelian subgroup of the SUSY group:

$$\begin{gathered} \mathcal{S}(\mathcal{A})^{2|2} \times \mathcal{A}^{2|0} \longrightarrow \mathcal{S}(\mathcal{A})^{2|2}, \\ (\Phi, r) \longmapsto [S_r(\Phi)](x) = \sum_k \frac{1}{k!} D^k \Phi(x + Br)[n, \dots, n], \end{gathered} \tag{6.26}$$

where $r = Br + n$, is a "superspatial shift". The momentum map becomes

$$\mathbf{J}(\Phi) = \left(\int \left(\pi \partial_i \Phi + \frac{1}{2} \psi^T \partial_i \psi \right) \right)_{i=1,2}. \tag{6.27}$$

We determine the *reduced phase space*. The center of mass of $\Phi = (\phi, \pi, \psi) \in \mathcal{S}(\mathcal{A})^{2|2}$ is $\mathcal{C}(\Phi) = (1/M(\Phi)) \int_{\mathbb{R}^2} x \cdot |\Phi|^2 d^2 x \in \mathcal{A}^{2|0}$, where $M(\Phi) = \int |\Phi|^2 d^2 x \in \mathcal{A}_0$. We may identify the quotient space $\mathbf{J}^{-1}(\boldsymbol{\mu}) / \mathcal{A}^{2|0}$ with the subset

$$P_\mu = \left\{ \widetilde{\Phi} \in \mathbf{J}^{-1}(\mu) \mid \mathcal{C}\left(\widetilde{\Phi}\right) = 0 \right\} \subseteq \mathcal{S}(\mathcal{A})^{2|2}. \tag{6.28}$$

Finally the *reduced equations become*

$$\begin{aligned}
\dot{\phi} &= \pi + F^i(\phi, \pi, \psi)\partial_i \phi, \\
\dot{\pi} &= \Delta\phi + F^i(\phi, \pi, \psi)\partial_i \pi, \\
\dot{\psi} &= \mathfrak{g}^0 \mathfrak{g}^i \partial_i \psi + F^i(\phi, \pi, \psi)\partial_i \psi,
\end{aligned} \tag{6.29}$$

where $F^i(\Phi) = (1/M(\Phi)) \int x^i \cdot (\phi\pi + \pi\Delta\phi - \psi^T \mathfrak{g}^j \partial_j \psi) d^2x$.

See [41, 42, 49, 60–63].

References

[1] N. Bourbaki, *Lie Groups and Lie Algebras*, Herman, Paris, France, 1975.

[2] H. H. Keller, *Differential Calculus in Locally Convex Spaces*, Lecture Notes in Mathematics, Vol. 417, Springer, Berlin, Germany, 1974.

[3] A. Pressley and G. Segal, *Loop Groups*, Oxford Mathematical Monographs, Oxford University Press, New York, NY, USA, 1986.

[4] J. Eichhorn and R. Schmid, "Form preserving diffeomorphisms on open manifolds," *Annals of Global Analysis and Geometry*, vol. 14, no. 2, pp. 147–176, 1996.

[5] H. Omori, *Infinite-Dimensional Lie Groups*, vol. 158 of *Translations of Mathematical Monographs*, American Mathematical Society, Providence, RI, USA, 1997.

[6] R. Schmid, "The geometry of BRS transformations," *Illinois Journal of Mathematics*, vol. 34, no. 1, pp. 87–97, 1990.

[7] A. Banyaga, *The Structure of Classical Diffeomorphism Groups*, vol. 400 of *Mathematics and Its Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1997.

[8] D. Bleecker, *Gauge Theory and Variational Principles*, vol. 1 of *Global Analysis Pure and Applied Series A*, Addison-Wesley, Reading, Mass, USA, 1981.

[9] N. Bourbaki, *Variétés Différentielles et Analytiques. Fascicule de Résultats*, Actualités Scientifiques et Industrielles, No. 1347, Hermann, Paris, France, 1971.

[10] D. G. Ebin and J. Marsden, "Groups of diffeomorphisms and the motion of an incompressible fluid," *Annals of Mathematics*, vol. 92, pp. 102–163, 1970.

[11] G. A. Goldin, "Lectures on diffeomorphism groups in quantum physics," in *Contemporary Problems in Mathematical Physics*, pp. 3–93, World Scientific, Hackensack, NJ, USA, 2004.

[12] R. S. Hamilton, "The inverse function theorem of Nash and Moser," *Bulletin of the American Mathematical Society*, vol. 7, no. 1, pp. 65–222, 1982.

[13] V. G. Kac, *Infinite-Dimensional Lie Algebras*, Cambridge University Press, Cambridge, Mass, USA, 2nd edition, 1985.

[14] S. Lang, *Differential Manifolds*, Addison-Wesley, Reading, Mass, USA, 1972.

[15] J. E. Marsden, D. G. Ebin, and A. E. Fischer, "Diffeomorphism groups, hydrodynamics and relativity," in *Proceedings of the 13th Biennial Seminar of the Canadian Mathematical Congress Differential Geometry and Applications*, J. R. Vanstone, Ed., vol. 1, pp. 135–279, 1972.

[16] P. W. Michor, *Manifolds of Differentiable Mappings*, vol. 3 of *Shiva Mathematics Series*, Shiva, Nantwich, UK, 1980.

[17] P. J. Olver, *Applications of Lie Groups to Differential Equations*, vol. 107 of *Graduate Texts in Mathematics*, Springer, New York, NY, USA, 2nd edition, 1993.

[18] R. Palais, *Foundations of Global Nonlinear Analysis*, Addison-Wesley, Reading, Mass, USA, 1968.

[19] R. Schmid, *Infinite-Dimensional Hamiltonian Systems*, vol. 3 of *Monographs and Textbooks in Physical Science. Lecture Notes*, Bibliopolis, Naples, Italy, 1987.

[20] R. Abraham and J. E. Marsden, *Foundations of Mechanics*, Benjamin/Cummings Publishing, Reading, Mass, USA, 2nd edition, 1978.

[21] J. E. Marsden, *Lectures on Mechanics*, vol. 174 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge, UK, 1992.

[22] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry: A Basic Exposition of Classical Mechanical System*, vol. 17 of *Texts in Applied Mathematics*, Springer, New York, NY, USA, 1994.

[23] J. Marsden and A. Weinstein, "Reduction of symplectic manifolds with symmetry," *Reports on Mathematical Physics*, vol. 5, no. 1, pp. 121–130, 1974.

[24] R. Abraham, J. E. Marsden, and T. Ratiu, *Manifolds, Tensor Analysis, and Applications*, vol. 2 of *Global Analysis Pure and Applied: Series B*, Addison-Wesley, Reading, Mass, USA, 1983.

[25] P. R. Chernoff and J. E. Marsden, *Properties of Infinite Dimensional Hamiltonian Systems*, Lecture Notes in Mathematics, Vol. 425, Springer, Berlin, Germany, 1974.

[26] Y. Choquet-Bruhat, C. DeWitt-Morette, and M. Dillard-Bleick, *Analysis, Manifolds and Physics*, North-Holland, Amsterdam, The Netherlands, 2nd edition, 1982.

[27] F. Gay-Balmaz and T. S. Ratiu, "Reduced Lagrangian and Hamiltonian formulations of Euler-Yang-Mills fluids," *The Journal of Symplectic Geometry*, vol. 6, no. 2, pp. 189–237, 2008.

[28] F. Gay-Balmaz and T. S. Ratiu, "A new Lagrangian dynamic reduction in field theory," to appear in *Annales de l'Institut Fourier*.

[29] V. Guillemin and S. Sternberg, *Symplectic Techniques in Physics*, Cambridge University Press, Cambridge, Mass, USA, 1984.

[30] R. Temam, *Infinite-Dimensional Dynamical Systems in Mechanics and Physics*, vol. 68 of *Applied Mathematical Sciences*, Springer, New York, NY, USA, 1988.

[31] S. Vasylkevych and J. E. Marsden, "The Lie-Poisson structure of the Euler equations of an ideal fluid," *Dynamics of Partial Differential Equations*, vol. 2, no. 4, pp. 281–300, 2005.

[32] J. E. Marsden, A. Weinstein, T. Ratiu, R. Schmid, and R. G. Spencer, "Hamiltonian systems with symmetry, coadjoint orbits and plasma physics," in *Proceedings of the IUTAM-ISIMM Symposium on Modern Developments in Analytical Mechanics*, vol. 1, pp. 289–340, Atti della Accademia delle Scienze di Torino. Classe di Scienze Fisiche, Matematiche e Naturali, Torino, Italy, 1982.

[33] M. Adams, T. Ratiu, and R. Schmid, "The Lie group structure of diffeomorphism groups and invertible Fourier integral operators, with applications," in *Infinite-Dimensional Groups with Applications*, vol. 4 of *Math. Sci. Res. Inst. Publ.*, pp. 1–69, Springer, New York, NY, USA, 1985.

[34] M. Adams, T. Ratiu, and R. Schmid, "A Lie group structure for pseudodifferential operators," *Mathematische Annalen*, vol. 273, no. 4, pp. 529–551, 1986.

[35] M. Adams, T. Ratiu, and R. Schmid, "A Lie group structure for Fourier integral operators," *Mathematische Annalen*, vol. 276, no. 1, pp. 19–41, 1986.

[36] A. E. Fischer and J. E. Marsden, "The Einstein equations of evolution—a geometric approach," *Journal of Mathematical Physics*, vol. 13, pp. 546–568, 1972.

[37] F. Gay-Balmaz and T. S. Ratiu, "The Lie-Poisson structure of the LAE-α equation," *Dynamics of Partial Differential Equations*, vol. 2, no. 1, pp. 25–57, 2005.

[38] J. E. Marsden and A. Weinstein, "The Hamiltonian structure of the Maxwell-Vlasov equations," *Physica D*, vol. 4, no. 3, pp. 394–406, 1982.

[39] F. Trèves, *Introduction to Pseudodifferential and Fourier Integral Operators: Pseudodifferential Operators*, vol. 1 of *The University Series in Mathematic*, Plenum Press, New York, NY, USA, 1980.

[40] F. Trèves, *Introduction to Pseudodifferential and Fourier Integral Operators: Fourier Integral Operators*, vol. 2 of *The University Series in Mathematic*, Plenum Press, New York, NY, USA, 1980.

[41] P. Deligne et al., Ed., *Quantum Fields and Strings: A Course for Mathematicians*, vol. 1, American Mathematical Society, Providence, RI, USA, 1999.

[42] P. Deligne et al., Ed., *Quantum Fields and Strings: A Course for Mathematicians*, vol. 2, American Mathematical Society, Providence, RI, USA, 1999.

[43] C. Becchi, A. Rouet, and R. Stora, "Renormalization of gauge theories," *Annals of Physics*, vol. 98, no. 2, pp. 287–321, 1976.

[44] R. Schmid, "Local cohomology in gauge theories, BRST transformations and anomalies," *Differential Geometry and Its Applications*, vol. 4, no. 2, pp. 107–116, 1994.

[45] C. Chevalley and S. Eilenberg, "Cohomology theory of Lie groups and Lie algebras," *Transactions of the American Mathematical Society*, vol. 63, pp. 85–124, 1948.

[46] R. Schmid, "BRST cohomology and anomalies," in *Geometrical and Algebraic Aspects of Nonlinear Field Theory*, North-Holland Delta Ser., pp. 159–172, North-Holland, Amsterdam, The Netherlands, 1989.

[47] A. Einstein, *Mein Weltbild*, Ullstein, Frankfurt, Germany, 1934.

[48] E. Verlinde, "On the origin of gravity and the laws of Newton," http://arxiv.org/abs/1001.0785.

[49] B. Kostant, "Graded manifolds, graded Lie theory, and prequantization," in *Differential Geometrical Methods in Mathematical Physics*, vol. 570 of *Lecture Notes in Mathematics*, pp. 177–306, Springer, Berlin, Germany, 1977.

[50] E. Binz, J. Śniatycki, and H. Fischer, *Geometry of Classical Fields*, vol. 154 of *North-Holland Mathematics Studies*, North-Holland, Amsterdam, The Netherlands, 1988.

[51] L. Bonora and P. Cotta-Ramusino, "Some remarks on BRS transformations, anomalies and the cohomology of the Lie algebra of the group of gauge transformations," *Communications in Mathematical Physics*, vol. 87, no. 4, pp. 589–603, 1982/83.

[52] M. Dubois-Violette, "The Weil-BRS algebra of a Lie algebra and the anomalous terms in gauge theory," *Journal of Geometry and Physics*, vol. 3, no. 4, pp. 525–565, 1986.

[53] T. Eguchi, P. B. Gilkey, and A. J. Hanson, "Gravitation, gauge theories and differential geometry," *Physics Reports*, vol. 66, no. 6, pp. 213–393, 1980.

[54] L. D. Faddeev and A. A. Slavnov, *Gauge Fields: Introduction to Quantum Theory*, vol. 50 of *Frontiers in Physics*, Benjamin/Cummings Publishing, Reading, Mass, USA, 1980.

[55] G. A. Goldin, R. Menikoff, and D. H. Sharp, "Diffeomorphism groups, gauge groups, and quantum theory," *Physical Review Letters*, vol. 51, no. 25, pp. 2246–2249, 1983.

[56] G. A. Goldin, "The diffeomorphism group approach to nonlinear quantum systems," *International Journal of Modern Physics B*, vol. 6, no. 11-12, pp. 1905–1916, 1992.

[57] C. Itzykson and J. B. Zuber, *Quantum Field Theory*, International Series in Pure and Applied Physic, McGraw-Hill, New York, NY, USA, 1980.

[58] D. Kastler and R. Stora, "A differential geometric setting for BRS transformations and anomalies. I, II," *Journal of Geometry and Physics*, vol. 3, no. 4, pp. 437–505, 1986.

[59] B. Kostant and S. Sternberg, "Symplectic reduction, BRS cohomology, and infinite-dimensional Clifford algebras," *Annals of Physics*, vol. 176, no. 1, pp. 49–113, 1987.

[60] T. Glimm, *Supersymplectic reduction*, Ph.D. thesis, Emory University, 2003, http://www.mathcs.emory.edu/~rudolf/Glimm.pdf.

[61] G. M. Tuynman, *Supermanifolds and Supergroups: Basic Theory*, vol. 570 of *Mathematics and Its Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 2004.

[62] B. DeWitt, *Supermanifolds*, Cambridge Monographs on Mathematical Physics, Cambridge University Press, Cambridge, Mass, USA, 2nd edition, 1992.

[63] T. Glimm and R. Schmid, "On the supersymmetry group of the classical Bose-Fermi oscillator," *Journal of Geometry and Symmetry in Physics*, vol. 4, pp. 45–58, 2005.

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 107857, 70 pages
doi:10.1155/2010/107857

Review Article

Instantons, Topological Strings, and Enumerative Geometry

Richard J. Szabo[1, 2]

[1] *Department of Mathematics, Heriot-Watt University, Colin Maclaurin Building, Riccarton, Edinburgh EH14 4AS, UK*
[2] *Maxwell Institute for Mathematical Sciences, Edinburgh, UK*

Correspondence should be addressed to Richard J. Szabo, r.j.szabo@ma.hw.ac.uk

Received 21 December 2009; Accepted 12 March 2010

Academic Editor: M. N. Hounkonnou

We review and elaborate on certain aspects of the connections between instanton counting in maximally supersymmetric gauge theories and the computation of enumerative invariants of smooth varieties. We study in detail three instances of gauge theories in six, four, and two dimensions which naturally arise in the context of topological string theory on certain noncompact threefolds. We describe how the instanton counting in these gauge theories is related to the computation of the entropy of supersymmetric black holes and how these results are related to wall-crossing properties of enumerative invariants such as Donaldson-Thomas and Gromov-Witten invariants. Some features of moduli spaces of torsion-free sheaves and the computation of their Euler characteristics are also elucidated.

1. Introduction

Topological theories in physics usually relate BPS quantities to geometrical invariants of the underlying manifolds on which the physical theory is defined. For the purposes of the present article, we will focus on two particular and well-known instances of this. The first is instanton counting in supersymmetric gauge theories in four dimensions, which gives the Seiberg-Witten and Donaldson-Witten invariants of four-manifolds. The second is topological string theory, which is related to the enumerative geometry of Calabi-Yau threefolds and computes, for example, Gromov-Witten invariants, Donaldson-Thomas invariants, Gopakumar-Vafa BPS invariants, and key aspects of Kontsevich's homological mirror symmetry conjecture.

From a physical perspective, these topological models are not simply of academic interest, but they also serve as exactly solvable systems which capture the physical content of certain sectors of more elaborate systems with local propagating degrees of freedom. Such is the case for the models we will consider in this paper, which are obtained as topological twists of a given physical theory. The topologically twisted theories describe the BPS sectors of physical models, and compute nonperturbative effects therein. For example, for certain supersymmetric charged black holes, the microscopic Bekenstein-Hawking-Wald entropy is computed by the Witten index of the relevant supersymmetric gauge theory. This is equivalent to the counting of stable BPS bound states of D-branes in the pertinent geometry, and is related to invariants of threefolds via the OSV conjecture [1].

From a mathematical perspective, we are interested in counting invariants associated to moduli spaces of coherent sheaves on a smooth complex projective variety X. To define such invariants, we need moduli spaces that are varieties rather than algebraic stacks. The standard method is to choose a polarization on X and restrict attention to semistable sheaves. If X is a Kähler manifold, then a natural choice of polarization is provided by a fixed Kähler two-form on X. Geometric invariant theory then constructs a projective variety which is a coarse moduli space for semistable sheaves of fixed Chern character. In this paper we will be interested in the computation of suitably defined Euler characteristics of certain moduli spaces, which are the basic enumerative invariants. We will also compute more sophisticated holomorphic curve counting invariants of a Calabi-Yau threefold X, which can be defined using virtual cycles of the pertinent moduli spaces and are invariant under deformations of X. In some instances the two types of invariants coincide.

An alternative approach to constructing moduli varieties is through framed sheaves. Then there is a projective *Quot* scheme which is a fine moduli space for sheaves with a given framing. A framed sheaf can be regarded as a geometric realization of an instanton in a noncommutative gauge theory on X [2–4] which asymptotes to a fixed connection at infinity. The noncommutative gauge theory in question arises as the worldvolume field theory on a suitable arrangement of D-branes in the geometry. In Nekrasov's approach [5], the set of observables that enter in the instanton counting are captured by the infrared dynamics of the topologically twisted gauge theory, and they compute the intersection theory of the (compactified) moduli spaces. The purpose of this paper is to overview the enumeration of such noncommutative instantons and its relation to the standard counting invariants of X.

In the following we will describe the computation of BPS indices of stable D-brane states via instanton counting in certain noncommutative and q-deformations of gauge theories on branes in various dimensions. We will pay particular attention to three noncompact examples which each arise in the context of Type IIA string theory.

(1) D6-D2-D0 bound states in D6-brane gauge theory—These compute Donaldson-Thomas invariants and describe atomic configurations in a melting crystal model [6]. This also provides a solid example of a (topological) gauge theory/string theory duality. The counting of noncommutative instantons in the pertinent topological gauge theory is described in detail in [7, 8].

(2) D4-D2-D0 bound states in D4-brane gauge theory—These count black hole microstates and allow us to probe the OSV conjecture. Their generating functions also appear to be intimately related to the two-dimensional rational conformal field theory.

(3) D2-D0 bound states in D2-brane gauge theory—These compute Gromov-Witten invariants of local curves. Instanton counting in the two-dimensional gauge theory on the base of the fibration is intimately related to instanton counting in the four-dimensional gauge theory obtained by wrapping supersymmetric D4-branes around certain noncompact four-cycles C, and also to the enumeration of flat connections in Chern-Simons theory on the boundary of C. These interrelationships are explored in detail in [9–13].

These counting problems provide a beautiful hierarchy of relationships between topological string theory/gauge theory in six dimensions, four-dimensional supersymmetric gauge theories, Chern-Simons theory in three dimensions, and a certain q-deformation of two-dimensional Yang-Mills theory. They are also intimately related to two-dimensional conformal field theory.

2. Topological String Theory

The basic setting in which to describe all gauge theories that we will analyse in this paper within a unified framework is through topological string theory, although many aspects of these models are independent of their connection to topological strings. In this section, we briefly discuss some physical and mathematical aspects of topological string theory, and how they naturally relate to the gauge theories that we are ultimately interested in. Further details about topological string theory can be found in, for example, [14, 15], or in [16] which includes a more general introduction. Introductory and advanced aspects of toric geometry are treated in the classic text [17] and in the reviews [18, 19]. The standard reference for the sheaf theory that we use is the book [20], while a more physicist-geared introduction with applications to string theory can be found in the review [21].

2.1. Topological Strings and Gromov-Witten Theory

Topological string theory may be regarded as a theory whose state space is a “subspace” of that of the full physical Type II string theory. It is designed so that it can resolve the mathematical problem of counting maps

$$f : \Sigma_g \longrightarrow X \tag{2.1}$$

from a closed oriented Riemann surface Σ_g of genus g into some target space X. In the physical Type II theory, any harmonic map f, with respect to chosen metrics on Σ_g and X, is allowed. They are described by solutions to second-order partial differential equations, the Euler-Lagrange equations obtained from a variational principle based on a sigma-model. The simplification supplied by topological strings is that one replaces this sigma-model on the worldsheet Σ_g by a two-dimensional topological field theory, which can be realized as a topological twist of the original $\mathcal{N} = 2$ superconformal field theory. In this reduction, the state space descends to its BRST cohomology defined with respect to the $\mathcal{N} = 2$ supercharges, which naturally carries a Frobenius algebra structure. This defines a consistent quantum theory if and only if the target space X is a Calabi-Yau threefold, that is, a complex Kähler manifold of dimension $\dim_{\mathbb{C}}(X) = 3$ with trivial canonical holomorphic line bundle K_X, or equivalently trivial first Chern class $c_1(X) := c_1(K_X) = 0$. We fix a closed nondegenerate Kähler $(1,1)$-form ω on X.

The corresponding topological string amplitudes F_g have interpretations in compactifications of Type II string theory on the product of the target space X with four-dimensional Minkowski space $\mathbb{R}^{3,1}$. For instance, at genus zero the amplitude F_0 is the prepotential for vector multiplets of $\mathcal{N} = 2$ supergravity in four dimensions. The higher genus contributions F_g, $g \geq 1$ correspond to higher derivative corrections of the schematic form $R^2 T^{2g-2}$, where R

is the curvature and T is the graviphoton field strength. We will now explain how to compute the amplitudes F_g. There are two types of topological string theories that we consider in turn.

2.1.1. A-Model

The A-model topological string theory isolates symplectic structure aspects of the Calabi-Yau threefold X. It is built on holomorphically embedded curves (2.1). The holomorphic string maps f in this case are called *worldsheet instantons*. They are classified topologically by their homology classes

$$\beta = f_*[\Sigma_g] \in H_2(X, \mathbb{Z}). \tag{2.2}$$

With respect to a basis S_i of two cycles on X, one can write

$$\beta = \sum_{i=1}^{b_2(X)} n_i [S_i], \tag{2.3}$$

where the Betti number $b_2(X)$ is the rank of the second homology group $H_2(X, \mathbb{Z})$, and $n_i \in \mathbb{Z}$. Due to the topological nature of the sigma-model, the string theory functional integral localizes equivariantly (with respect to the BRST cohomology) onto contributions from worldsheet instantons.

The sum over all maps can be encoded in a generating function $F^X_{\text{top}}(g_s, \mathbf{Q})$ which depends on the string coupling g_s and a vector of variables $\mathbf{Q} = (Q_1, \ldots, Q_{b_2(X)})$ defined as follows. Let

$$t_i = \int_{S_i} \omega \tag{2.4}$$

be the complex Kähler parameters of X with respect to the basis S_i. They appear in the values of the sigma-model action evaluated on a worldsheet instanton. For an instanton in curve class (2.3), the corresponding Boltzmann weight is

$$\mathbf{Q}^\beta := \prod_{i=1}^{b_2(X)} (Q_i)^{n_i} \quad \text{with } Q_i := \mathrm{e}^{-t_i}. \tag{2.5}$$

Then the quantum string theory is described by a genus expansion of the free energy

$$F^X_{\text{top}}(g_s, \mathbf{Q}) = \sum_{g=0}^{\infty} g_s^{2g-2} F_g(\mathbf{Q}) \tag{2.6}$$

weighted by the Euler characteristic $\chi(\Sigma_g) = 2 - 2g$ of Σ_g, where the genus g contribution to the statistical sum is given by

$$F_g(\mathbf{Q}) = \sum_{\beta \in H_2(X, \mathbb{Z})} N_{g,\beta}(X) \mathbf{Q}^\beta, \tag{2.7}$$

and in this formula the classes $\beta \neq 0$ correspond to worldsheets of genus g. The numbers $N_{g,\beta}(X)$ are called the *Gromov-Witten invariants* of X and they "count", in an appropriate sense, the number of worldsheet instantons (holomorphic curves) of genus g in the two-homology class β. They can be defined as follows.

A worldsheet instanton (2.1) is said to be *stable* if the automorphism group $\mathrm{Aut}(\Sigma_g, f)$ is finite. Let $\mathfrak{M}_g(X, \beta)$ be the (compactified) moduli space of isomorphism classes of stable holomorphic maps (2.1) from connected genus g curves to X representing β. This is the *instanton moduli space* onto which the path integral of topological string theory localizes. It is a proper Deligne-Mumford stack over $\mathbb{C}$ which generalizes the moduli space $\mathfrak{M}_g$ of "stable" curves of genus g. While the dimension of $\mathfrak{M}_g$ is $3g-3$, the moduli space $\mathfrak{M}_g(X, \beta)$ is in general reducible and of impure dimension, as all possible stable maps occur. However, there is a perfect obstruction theory [22] which generically has virtual dimension

$$(1-g)(\dim_{\mathbb{C}}(X)-3)+\int_{\beta} c_1(X). \tag{2.8}$$

When X is a Calabi-Yau threefold, this integer vanishes and there is a virtual fundamental class [22]

$$[\mathfrak{M}_g(X, \beta)]^{\mathrm{vir}} \in CH_0(\mathfrak{M}_g(X, \beta)) \tag{2.9}$$

in the degree zero Chow group. In this case, we define

$$N_{g,\beta}(X) := \int_{[\mathfrak{M}_g(X,\beta)]^{\mathrm{vir}}} 1, \tag{2.10}$$

and so the Gromov-Witten invariants give the "virtual numbers" of worldsheet instantons. One generically has $N_{g,\beta}(X) \in \mathbb{Q}$ due to the orbifold nature of the moduli stack $\mathfrak{M}_g(X, \beta)$. One can also define invariants by integrating the Euler class of an *obstruction bundle* over $\mathfrak{M}_g(X, \beta)$. There are precise recipes for computing the Gromov-Witten invariants $N_{g,\beta}(X)$ for *toric* varieties X.

2.1.2. B-Model

The B-model topological string theory isolates complex structure aspects of the Calabi-Yau threefold X. It enumerates the *constant* string maps which send the entire surface Σ_g to a fixed point in X, and hence have trivial curve class $\beta = 0$. The Gromov-Witten invariants in this case are completely understood. There is a natural isomorphism

$$\mathfrak{M}_g(X, 0) \cong \mathfrak{M}_g \times X, \tag{2.11}$$

and the degree zero Gromov-Witten invariants $N_{g,0}(X)$ involve only the classical cohomology ring $H^{\bullet}(X, \mathbb{Q})$ and "Hodge integrals" over the moduli spaces of Riemann surfaces $\mathfrak{M}_g$ defined as follows.

There is a canonical stack line bundle $\mathfrak{L} \to \mathfrak{M}_g$ with fibre $T^*_{\Sigma_g}$ over the moduli point $[\Sigma_g]$, the cotangent space of Σ_g at some fixed point. We define the tautological class ψ to be the first Chern class of $\mathfrak{L}$, $\psi := c_1(\mathfrak{L}) \in H^2(\mathfrak{M}_g, \mathbb{Q})$. The *Hodge bundle* $\mathfrak{E} \to \mathfrak{M}_g$ is the complex vector bundle of rank g whose fibre over a point Σ_g is the space $H^0(\Sigma_g, K_{\Sigma_g})$ of holomorphic sections of the canonical line bundle $K_{\Sigma_g} \to \Sigma_g$. Let $\lambda_j := c_j(\mathfrak{E}) \in H^{2j}(\mathfrak{M}_g, \mathbb{Q})$. A *Hodge integral* over $\mathfrak{M}_g$ is an integral of products of the classes ψ and λ_j.

Explicit expressions for $N_{g,0}(X)$ for *generic* threefolds X are then obtained as follows. Let $\{\gamma_a\}_{a\in A}$ be a basis for $H^\bullet(X, \mathbb{Z})$ (modulo torsion), and let $D_2 \subset A$ index the generators of degree two. Then one has

$$\begin{aligned} N_{0,0}(X) &= \sum_{a_i \in A} \frac{1}{3!} \int_X \left(\gamma_{a_1} \smile \gamma_{a_2} \smile \gamma_{a_3}\right), \\ N_{1,0}(X) &= -\sum_{a \in D_2} \frac{1}{24} \int_X \gamma_a \smile c_2(X), \\ N_{g\geq 2,0}(X) &= \frac{(-1)^g}{2} \int_X \left(c_3(X) - c_1(X) \smile c_2(X)\right) \int_{\mathfrak{M}_g} \lambda_{g-1}^3, \end{aligned} \tag{2.12}$$

where the Hodge integral can be expressed in terms of Bernoulli numbers as

$$\int_{\mathfrak{M}_g} \lambda_{g-1}^3 = \frac{|B_{2g}|}{2g} \frac{|B_{2g-2}|}{2g-2} \frac{1}{(2g-2)!}. \tag{2.13}$$

Note that $c_1(X) = 0$ above when X is Calabi-Yau.

Thus we know how to compute everything in the B-model, and it is completely under control. Our main interest is thus in extending these computations to the A-model. In analogy with the above considerations, one can note that there is a natural forgetful map

$$\pi : \mathfrak{M}_g(X, \beta) \longrightarrow \mathfrak{M}_g, \qquad (f, \Sigma_g) \longmapsto \Sigma_g, \tag{2.14}$$

and then reduce any integral over $\mathfrak{M}_g(X, \beta)$ to $\mathfrak{M}_g$ using the corresponding Gysin push-forward map $\pi_!$. However, this is difficult to do explicitly in most cases. The Gromov-Witten theory of X is the study of tautological intersections in the moduli spaces $\mathfrak{M}_g(X, \beta)$. There is a string duality between the A-model and the B-model which is related to homological mirror symmetry.

2.2. Open Topological Strings

An open topological string in X is described by a holomorphic embedding $f : \Sigma_{g,h} \to X$ of a curve $\Sigma_{g,h}$ of genus g with h holes. A D-brane in X is a choice of Dirichlet boundary condition on these string maps, which ensures that the Cauchy problem for the Euler-Lagrange equations on $\Sigma_{g,h}$ locally has a unique solution. They correspond to *Lagrangian* submanifolds L of the Calabi-Yau threefold X, that is, $\omega|_L = 0$. If $\partial\Sigma_{g,h} = \sigma_1 \cup \cdots \cup \sigma_h$, then we consider holomorphic maps such that

$$f(\sigma_i) \subset L. \tag{2.15}$$

This defines open string instantons, which are labelled by their *relative* homology classes

$$f_*\left[\Sigma_{g,h}\right] = \beta \in H_2(X, L). \tag{2.16}$$

If we assume that $b_1(L) = 1$, so that $H_1(L, \mathbb{Z})$ is generated by a single nontrivial one-cycle γ, then

$$f_*[\sigma_i] = w_i \gamma, \tag{2.17}$$

where $w_i \in \mathbb{Z}$, $i = 1, \dots, h$, are the winding numbers of the boundary maps $f|_{\sigma_i}$.

The free energy of the A-model open topological string theory at genus g is given by

$$F_{\mathbf{w},g}(\mathbf{Q}) = \sum_{\beta} N_{\mathbf{w},g,\beta}(X)\mathbf{Q}^{\beta}, \tag{2.18}$$

where $\mathbf{w} = (w_1, \dots, w_h)$ and the numbers $N_{\mathbf{w},g,\beta}(X)$ are called *relative Gromov-Witten invariants*. To incorporate all topological sectors, in addition to the string coupling g_s weighting the Euler characteristics $\chi(\Sigma_{g,h}) = 2 - 2g - h$, we introduce an $N \times N$ Hermitean matrix V to weight the different winding numbers. This matrix is associated to the holonomy of a gauge connection (Wilson line) on the D-brane. Then, taking into account that the holes are indistinguishable, the complete genus expansion of the generating function is

$$F^X_{\mathrm{top}}(g_s, \mathbf{Q}; V) = \sum_{g=0}^{\infty} \sum_{h=1}^{\infty} \sum_{\mathbf{w} \in \mathbb{Z}^h} \frac{1}{h!}\, g_s^{2g-2+h}\; F_{\mathbf{w},g}(\mathbf{Q}) \prod_{i=1}^{h} \mathrm{Tr}\left(V^{w_i}\right). \tag{2.19}$$

The traces are computed by formally taking the limit $N \to \infty$ and expanding in irreducible representations R of the D-brane gauge group $U(\infty)$.

2.3. Black Hole Microstates and D-Brane Gauge Theory

When X is a Calabi-Yau threefold, certain BPS black holes on $X \times \mathbb{R}^{3,1}$ can be constructed by D-brane engineering. D-branes in X correspond to submanifolds of X equipped with vector bundles with connection, the Chan-Paton gauge bundles, and they carry *charges* associated with the Chern characters of these bundles. This data defines a class in the differential K-theory of X, which provides a topological classification of D-branes in X.

The microscopic black hole entropy can be computed by counting stable bound states of D0–D2–D4–D6 branes wrapping holomorphic cycles of X with the following configurations:

(i) D6-brane charge Q_6,

(ii) D4-branes wrapping an ample divisor

$$[C] = \sum_{i=1}^{b_2(X)} Q_4^i\; [C_i] \in H_4(X, \mathbb{Z}) \tag{2.20}$$

with respect to a basis of four-cycles C_i, $i = 1, \dots, b_4(X) = b_2(X)$, of X,

(iii) D2-branes wrapping a two-cycle

$$[S] = \sum_{i=1}^{b_2(X)} Q_2^i \; [S_i] \in H_2(X, \mathbb{Z}), \tag{2.21}$$

(iv) D0-brane charge Q_0.

These D-brane charges give the black hole its charge quantum numbers. If we consider large enough numbers of D-branes in this system, then they form bound states which become large black holes with smooth event horizons, that can be counted and therefore account for the microscopic black hole entropy. In this scenario, $p^I = (Q_6, Q_4^i)$ are interpreted as magnetic charges and $q_I = (Q_0, Q_2^i)$ as electric charges. The thermal partition function defined via a canonical ensemble for the D0 and D2 branes with chemical potentials $\mu^I = (\phi^0, \phi_i^2)$, and a microcanonical ensemble for the D4 and D6 branes, is given by

$$Z_{\mathrm{BH}}\left(Q_6, \boldsymbol{Q}_4, \boldsymbol{\phi}^2, \phi^0\right) = \sum_{Q_0, Q_2^i} \Omega(Q_0, \boldsymbol{Q}_2, \boldsymbol{Q}_4, Q_6) \; \mathrm{e}^{-Q_0\phi^0 - Q_2^i \phi_i^2}, \tag{2.22}$$

where Ω is the degeneracy of BPS states with the given D-brane charges.

As we mentioned in Section 2.1., the closed topological string amplitudes F_g are related to supergravity quantities on Minkowski spacetime $\mathbb{R}^{3,1}$. The fact that the genus zero free energy F_0 for topological strings on X is a prepotential for BPS black hole charges in $\mathcal{N} = 2$ supergravity determines the entropy $S_{\mathrm{BH}}(p, q)$ of an extremal black hole as a Legendre transformation of F_0, provided that one fixes the charge moduli by the attractor mechanism. The genus zero topological string amplitude F_0 is a homogeneous function of degree two in the $\mathcal{N} = 2$ vector multiplet fields X^I. The black hole entropy in the supergravity approximation is then

$$S_{\mathrm{BH}}(p, q) = \mu^I q_I - \operatorname{Im} F_0\left(X^I = p^I + \mathrm{i}\mu^I\right), \tag{2.23}$$

where the chemical potentials μ^I are determined by the charges p^I and q_I by solving the equation

$$q_I = \frac{\partial \operatorname{Im} F_0}{\partial \mu^I}. \tag{2.24}$$

Further analyses of the entropy of $\mathcal{N} = 2$ BPS black holes on $\mathbb{R}^{3,1}$ have been extended to higher genus and suggest the relationship

$$Z_{\mathrm{BH}}\left(Q_6, \boldsymbol{Q}_4, \boldsymbol{\phi}^2, \phi^0\right) = \left| Z_{\mathrm{top}}^X(g_s, \mathbf{Q}) \right|^2 \tag{2.25}$$

between the black hole partition function (2.22) and the topological string partition function

$$Z_{\mathrm{top}}^X(g_s, \mathbf{Q}) = \exp F_{\mathrm{top}}^X(g_s, \mathbf{Q}), \tag{2.26}$$

where the moduli on both sides of this equation are related by their fixing at the attractor point

$$g_s = \frac{4\pi \mathrm{i}}{(\mathrm{i}/\pi)\ \phi^0 + Q_6}, \qquad t_i = \frac{2\phi_i^2 + \mathrm{i}Q_4^i}{(\mathrm{i}/\pi)\ \phi^0 + Q_6}. \tag{2.27}$$

The remarkable relationship (2.25) is called the OSV conjecture [1]. It provides a means of using the perturbation expansion of topological strings and Gromov-Witten theory to compute black hole entropy to all orders. Alternatively, although the evidence for this proposal is derived for large black hole charge, the left-hand side of the expression (2.25) makes sense for finite charges and in some cases is explicitly computable in closed form. It can thus be used to define *nonperturbative* topological string amplitudes, and hence a nonperturbative completion of a string theory.

In the following, we will focus on the computation of the black hole partition function (2.22). The fact that this partition function is computable in a D-brane gauge theory will then give a physical interpretation of the enumerative invariants of X in terms of black hole entropy. Suppose that we have a collection of Dp-branes wrapping a submanifold $M_{p+1} \subset X$, with $\dim_{\mathbb{R}}(M_{p+1}) = p+1$ and Chan-Paton gauge field strength F. D-branes are charged with respect to supergravity differential form fields, the Ramond-Ramond fields, which are also classified topologically by differential K-theory. Recall that such an array couples to all n-form Ramond-Ramond fields $C_{(n)}$ through anomalous Chern-Simons couplings

$$\int_{M_{p+1}} \sum_{n \geq 0} C_{(n)} \wedge \mathrm{Tr}\ \exp(2\pi\alpha' F), \tag{2.28}$$

where $\sqrt{\alpha'}$ is the string length. In particular, these couplings contain all terms

$$\int_{M_{p+1}} C_{(p+1-2m)} \wedge \mathrm{Tr}(F^m), \tag{2.29}$$

and so the topological charge $\mathrm{ch}_m(\mathcal{E})$ of a Chan-Paton gauge bundle $\mathcal{E} \to M_{p+1}$ on a Dp-brane is equivalent to D$(p-2m)$-brane charge. A prominent example of this, which will be considered in detail later on, is the coupling $\int_{M_{p+1}} C_{(p-3)} \wedge \mathrm{Tr}(F \wedge F)$. For $p=3$, this shows that the counting of D4-D0 brane bound states is equivalent to the enumeration of instantons on the four-dimensional part of the D4-brane in X. The remaining sections of this paper look at these relationships from the point of view of various BPS configurations of these D-branes. We will study the enumeration problems from the point of view of gauge theories on the D-branes in order of decreasing dimensionality, stressing the analogies between each description.

3. D6-Brane Gauge Theory and Donaldson-Thomas Invariants

In this section we will look at a single D6-brane ($Q_6 = 1$) and turn off all D4-brane charges ($Q_4^i = 0$). We will discuss various physical theories which are modelled by the D6-brane gauge theory in this case, but otherwise have no *a priori* relation to string theory. These will include a tractable model for quantum gravity and the statistical mechanics of certain atomic crystal configurations. From the perspective of enumerative geometry, these partition functions will compute the Donaldson-Thomas theory of X.

3.1. Kähler Quantum Gravity

We will construct a model of quantum gravity on any Kähler threefold X, which will motivate the sorts of counting problems that we consider in this section. The partition function is defined by

$$Z = \sum_{\substack{\text{quantized}\\ \omega}} \mathrm{e}^{-S}, \tag{3.1}$$

where

$$S = \frac{1}{g_s^2} \int_X \frac{1}{3!}\, \omega \wedge \omega \wedge \omega. \tag{3.2}$$

The sum is over "quantized" Kähler two-forms on X, in the following sense. We decompose the "macroscopic" form ω into a fixed "background" Kähler two-form ω_0 on X and the curvature F of a holomorphic line bundle $\mathcal{L}$ over X as

$$\omega = \omega_0 + g_s F. \tag{3.3}$$

To satisfy the requirement that there are no D4-branes in X, we impose the fluctuation condition

$$\int_\beta F = 0 \tag{3.4}$$

for all two-cycles $\beta \in H_2(X, \mathbb{Z})$.

Substituting (3.3) together with (3.4) into (3.2) gives the action

$$S = \frac{1}{g_s^2} \frac{1}{3!} \int_X \omega_0^3 + \frac{1}{2} \int_X F \wedge F \wedge \omega_0 + g_s \int_X \frac{1}{3!} F \wedge F \wedge F. \tag{3.5}$$

The statistical sum (3.1) thus becomes (dropping an irrelevant constant term)

$$Z = \sum_{\substack{\text{line}\\ \text{bundles } \mathcal{L}}} q^{\mathrm{ch}_3(\mathcal{L})} \prod_{i=1}^{b_2(X)} (Q_i)^{\int_{C_i} \mathrm{ch}_2(\mathcal{L})}, \tag{3.6}$$

where $q = -\mathrm{e}^{-g_s}$, $Q_i = \mathrm{e}^{-t_i}$, and $\mathrm{ch}_m(\mathcal{L})$ denotes the mth Chern character of the given line bundle $\mathcal{L} \to X$. Note the formal similarity with the A-model topological string partition function constructed in Section 2.1. However, there is a problem with the way in which we have thus far set up this model. The fluctuation condition (3.4) on F implies $\mathrm{ch}_2(\mathcal{L}) = \mathrm{ch}_3(\mathcal{L}) = 0$. Hence only trivial line bundles can contribute to the sum (3.6), and the partition function is trivial.

The resolution to this problem is to enlarge the range of summation in (3.6) to include singular gauge fields and ideal sheaves. Namely, we take F to correspond to a *singular* $U(1)$ gauge field A on X. This can be realized in two (related) possible ways:

(1) we can make a singular gauge field A nonsingular on the blow-up

$$\widehat{X} \longrightarrow X \tag{3.7}$$

of the target space, obtained by blowing up the singular points of A on X into copies of the complex projective plane $\mathbb{P}^2$. This means that the quantum gravitational path integral induces a topology change of the target space X. This is referred to as "quantum foam" in [23, 24], or

(2) we can relax the notion of line bundle to ideal sheaf. Ideal sheaves lift to line bundles on $\widehat{X}$. However, there are "more" sheaves on X than blow-ups $\widehat{X}$ of X.

In this paper we will take the second point of view. Recall that torsion-free sheaves $\mathcal{E}$ on X can be defined by the property that they sit inside short exact sequences of sheaves of the form

$$0 \longrightarrow \mathcal{E} \longrightarrow \mathcal{F} \longrightarrow \mathcal{S}_Z \longrightarrow 0, \tag{3.8}$$

where $\mathcal{F}$ is a holomorphic vector bundle on X, and $\mathcal{S}_Z$ is a coherent sheaf supported at the singular points $Z \subset X$ of a gauge connection A of $\mathcal{F}$. Applying the Chern character to (3.8) and using its additivity on exact sequences give

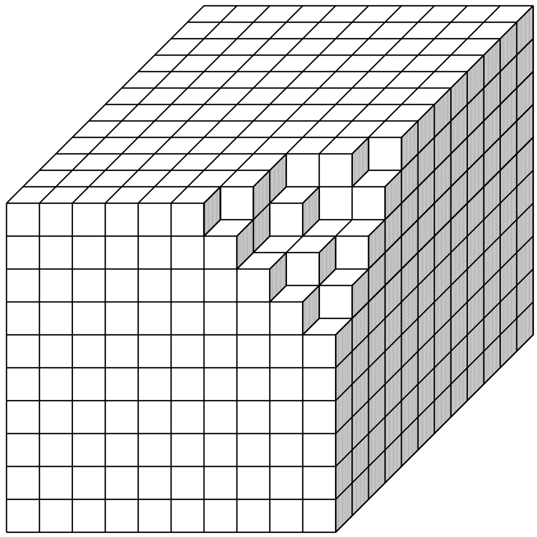
$$\mathrm{ch}_m(\mathcal{E}) = \mathrm{ch}_m(\mathcal{F}) - \mathrm{ch}_m(\mathcal{S}_Z) \tag{3.9}$$

for each m. Thus torsion-free sheaves $\mathcal{E}$ fail to be vector bundles at singular points of gauge fields, and including the singular locus can reinstate the nontrivial topological quantum numbers that we desired above.

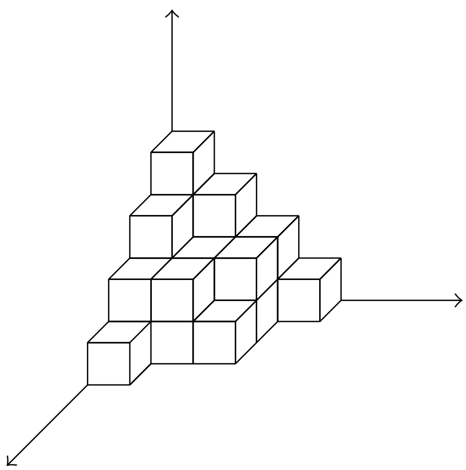
As we will discuss in detail in this section, this construction is realized explicitly by considering a noncommutative gauge theory on the target space $X = \mathbb{C}^3$. We will see that the instanton solutions of gauge theory on a noncommutative deformation $\mathbb{C}^3_\theta$ are described in terms of *ideals* $\mathcal{I}$ in the polynomial ring $\mathbb{C}[z^1, z^2, z^3]$. For generic X, the global object that corresponds locally to an ideal is an *ideal sheaf*, which in each coordinate patch $U_\alpha \subset X$ is described as an ideal $\mathcal{I}_{U_\alpha}$ in the ring $\mathcal{O}_{U_\alpha}$ of holomorphic functions on U_α. More abstractly, an ideal sheaf is a rank one torsion-free sheaf $\mathcal{E}$ with $c_1(\mathcal{E}) = 0$. This is a purely commutative description, since the holomorphic functions on $\mathbb{C}^3$ form a commutative subalgebra of $\mathbb{C}^3_\theta$ for the Moyal deformation that we will consider. Thus the desired singular gauge field configurations will be realized explicitly in terms of noncommutative instantons [23, 24].

3.2. Crystal Melting and Random Plane Partitions

As we will see, the counting of ideal sheaves is in fact equivalent to a combinatorial problem, which provides an intriguing connection between the Kähler quantum gravity model of Section 3.1. and a particular statistical mechanics model [6]. Consider a cubic crystal



located on the lattice $\mathbb{Z}^3_{\geq 0} \subset \mathbb{R}^3$. Suppose that we start heating the crystal at its outermost right corner. As the crystal melts, we remove atoms, depicted symbolically here by boxes, and arrange them into stacks of boxes in the positive octant. Owing to the rules for arranging the boxes according to the order in which they melt, this configuration defines a plane partition or a three-dimensional Young diagram.



Removing each atom from the corner of the crystal contributes a factor $q = \mathrm{e}^{-\mu/T}$ to the Boltzmann weight, where μ is the chemical potential and T is the temperature.

Let us define more precisely the combinatorial object that we have constructed, which generalizes the usual notion of partition and Young tableau. A *plane partition* is a semiinfinite rectangular array of nonnegative integers

$$\pi = \begin{matrix} \pi_{1,1} & \pi_{1,2} & \pi_{1,3} & \cdots \\ \pi_{2,1} & \pi_{2,2} & \pi_{2,3} & \cdots \\ \pi_{3,1} & \pi_{3,2} & \pi_{3,3} & \cdots \\ \vdots & \vdots & \vdots & \end{matrix} \tag{3.10}$$

such that $\pi_{i,j} \geq \pi_{i+1,j}$, and $\pi_{i,j} \geq \pi_{i,j+1}$ for all $i, j \geq 1$. We may regard a partition π as a three-dimensional Young diagram, in which we pile $\pi_{i,j}$ cubes vertically at the (i, j)th position in the plane as depicted above. The volume of a plane partition

$$|\pi| = \sum_{i,j\geq 1} \pi_{i,j} \tag{3.11}$$

is the total number of cubes. The diagonal slices of π are the partitions $(\pi_{i,i+m})_{i\geq 1}$, $m \geq 0$, obtained by cutting the three-dimensional Young diagram with planes, and they represent a sequence of ordinary partitions (Young tableaux) $\lambda = (\lambda_1, \lambda_2, \ldots)$, with $\lambda_i \geq \lambda_{i+1}$ for all $i \geq 1$. Here $\lambda_i \geq 0$ is the length of the ith row of the Young diagram, viewed as a collection of unit squares, and only finitely many λ_i are nonzero.

The counting problem for random plane partitions can be solved explicitly in closed form. For this, we consider the statistical mechanics in a canonical ensemble in which each plane partition π has energy proportional to its volume $|\pi|$. The corresponding partition function then gives the generating function for plane partitions

$$\begin{aligned} Z &:= \sum_{\pi} q^{|\pi|} \\ &= \sum_{N=0}^{\infty} pp(N) q^N \\ &= \prod_{n=1}^{\infty} \frac{1}{(1-q^n)^n} =: M(q), \end{aligned} \tag{3.12}$$

where $pp(N)$ is the number of plane partitions π with $|\pi| = N$ boxes. The function $M(q)$ is called the *MacMahon function*.

3.3. Six-Dimensional Cohomological Gauge Theory

We will now describe a $U(1)$ gauge theory formulation of the above statistical models [7, 8, 24]. If we gauge-fix the residual symmetry of the quantized Kähler gravity action (3.5), we obtain the action

$$S = \frac{1}{2} \int_X \left(\mathrm{d}_A \Phi \wedge \star \mathrm{d}_A \overline{\Phi} + \left| F^{2,0} \right|^2 + \left| F^{1,1} \right|^2 \right) + \frac{1}{2} \int_X \left(F \wedge F \wedge \omega_0 + \frac{g_s}{3} F \wedge F \wedge F \right), \tag{3.13}$$

where $\mathrm{d}_A = \mathrm{d} - \mathrm{i}\,[A, -]$ is the gauge covariant derivative acting on the complex scalar field Φ, $\star$ denotes the Hodge operator with respect to the Kähler metric of X, and $F = \mathrm{d}A$ is the curvature two-form which has the Kähler decomposition $F = F^{2,0} + F^{1,1} + F^{0,2}$. The field theory defined by this action arises in three (related) instances such as:

(1) a topological twist of maximally supersymmetric Yang-Mills theory in six dimensions,

(2) the dimensional reduction of supersymmetric Yang-Mills theory in ten dimensions on X,

(3) the low-energy effective field theory on a D6-brane wrapping X in Type IIA string theory, with D2 and D0 brane sources.

The gauge theory has a BRST symmetry [25, 26] and its partition function localizes at the BRST fixed points described by the equations

$$F^{2,0} = 0 = F^{0,2}, \tag{3.14}$$

$$F^{1,1} \wedge \omega_0 \wedge \omega_0 = 0, \tag{3.15}$$

$$\mathrm{d}_A \Phi = 0. \tag{3.16}$$

These equations also describe three (related) quantities:

(i) the *Donaldson-Uhlenbeck-Yau (DUY) equations* expressing Mumford-Takemoto slope stability of holomorphic vector bundles over X with finite characteristic classes,

(ii) BPS solutions in the gauge theory which correspond to (generalized) instantons,

(iii) bound states of D0–D2 branes in a single D6-brane wrapping X.

Recall that (3.14) and (3.15) are a special instance of the Hermitean Yang-Mills equations in which a constant λ is added to the right-hand side of (3.15). These equations arise in compactifications of heterotic string theory. The condition that the compactification preserves at least one unbroken supersymmetry requires $\lambda = 0$. These are the natural BPS conditions on a Kähler manifold (X, ω_0) which generalize the usual self-duality equations in four dimensions.

The localization of the gauge theory partition function Z onto the corresponding instanton moduli space $\mathfrak{M}_X$ can be written symbolically as [24, 26]

$$Z = \int_{\mathfrak{M}_X} e(\mathfrak{N}_X), \tag{3.17}$$

where $e(\mathfrak{N}_X)$ is the Euler class of the obstruction bundle $\mathfrak{N}_X$ whose fibres are spanned by the zero modes of the antighost fields. The zero modes of the fermion fields in the full supersymmetric extension of (3.13) [25, 26] are in correspondence with elements in the cohomology groups of the twisted Dolbeault complex

$$\Omega^{0,0}(X, \mathrm{ad}\ \mathcal{P}) \xrightarrow{\bar{\partial}_A} \Omega^{0,1}(X, \mathrm{ad}\ \mathcal{P}) \xrightarrow{\bar{\partial}_A} \Omega^{0,2}(X, \mathrm{ad}\ \mathcal{P}) \xrightarrow{\bar{\partial}_A} \Omega^{0,3}(X, \mathrm{ad}\ \mathcal{P}) \tag{3.18}$$

with ad $\mathcal{P}$ the adjoint gauge bundle over X. By incorporating the gauge fields, one can rewrite this complex in the form [24]

$$\Omega^{0,0}(X, \mathrm{ad}\ \mathcal{P}) \longrightarrow \begin{array}{c} \Omega^{0,1}(X, \mathrm{ad}\ \mathcal{P}) \\ \oplus \\ \Omega^{0,3}(X, \mathrm{ad}\ \mathcal{P}) \end{array} \longrightarrow \Omega^{0,2}(X, \mathrm{ad}\ \mathcal{P}), \tag{3.19}$$

which describes solutions of the DUY equations up to linearized complex gauge transformations. The morphism $\Omega^{0,3}(X, \mathrm{ad}\ \mathcal{P}) \to \Omega^{0,2}(X, \mathrm{ad}\ \mathcal{P})$ here is responsible for the appearance of the obstruction bundle in (3.17) [24, 26].

In order for the integral (3.17) to be well defined, we need to choose a compactification of $\mathfrak{M}_X$. In light of our earlier discussion, we will take this to be the Gieseker compactification, that is, the moduli space of ideal sheaves on X. The corresponding variety $\mathfrak{M}_X$ stratifies into components $\mathrm{Hilb}_{n,\beta}(X)$ given by the *Hilbert scheme* of points and curves in X, parameterizing isomorphism classes of ideal sheaves $\mathcal{E}$ with $\mathrm{ch}_1(\mathcal{E}) = c_1(\mathcal{E}) = 0$, $\mathrm{ch}_2(\mathcal{E}) = -\beta$, and $\mathrm{ch}_3(\mathcal{E}) = -n$. The partition function (3.17) is the generating function for the number of D0-D2 brane bound states in the D6-brane wrapping X. Mathematically, these are the *Donaldson-Thomas invariants* of X. We will define this moduli space integration, and hence these invariants, more precisely in Section 3.9.

3.4. Localization in Toric Geometry

Toric varieties provide a large class of algebraic varieties in which difficult problems in algebraic geometry can be reduced to combinatorics. Much of this paper will be concerned with these geometries as they possess symmetries which facilitate computations, particularly those involving moduli space integrations. Let us start by recalling some basic notions from toric geometry. Below we give the pertinent definitions specifically in the case of varieties of complex dimension three, the case of immediate interest to us, but they extend to arbitrary dimensions in the obvious ways.

A smooth complex threefold X is called a *toric manifold* if it densely contains a (complex algebraic) torus T^3 and the natural action of T^3 on itself (by translations) extends to the whole of X. Basic examples are the torus T^3 itself, the affine space $\mathbb{C}^3$, and the complex projective space $\mathbb{P}^3$. If in addition X is Calabi-Yau, then X is necessarily noncompact.

One of the great virtues of working with toric varieties X is that their geometry can be completely described by combinatorial data encoded in a *toric diagram*. The toric diagram is a graph consisting of the following ingredients:

(i) a set of *vertices* f which are the fixed points of the T^3-action on X, such that X can be covered by T^3-invariant open charts homeomorphic to $\mathbb{C}^3$,

(ii) a set of *edges* e which are T^3-invariant projective lines $\mathbb{P}^1 \subset X$ joining particular pairs of fixed points f_1, f_2,

(iii) a set of "*gluing rules*" for assembling the $\mathbb{C}^3$ patches together to reconstruct the variety X. In a neighbourhood of each edge e, X looks like the normal bundle over the corresponding $\mathbb{P}^1$. Since this normal bundle is a holomorphic bundle of rank two and every bundle over $\mathbb{P}^1$ is a sum of line bundles (by the Grothendieck-Birkhoff theorem), it is of the form

$$\mathcal{O}_{\mathbb{P}^1}(-m_1) \oplus \mathcal{O}_{\mathbb{P}^1}(-m_2) \tag{3.20}$$

for some integers m_1, m_2. The normal bundle in this way determines the local geometry of X near the edge e via the transition function

$$(w_1, w_2, w_3) \longmapsto \left(w_1^{-1}, w_2 w_1^{-m_1}, w_3 w_1^{-m_2}\right) \tag{3.21}$$

between the corresponding affine patches (going from the north pole to the south pole of the associated $\mathbb{P}^1$). In the Calabi-Yau case, the Chern numbers $c_1(X) = 0$ and $c_1(\mathbb{P}^1) = 2$ imply the condition $m_1 + m_2 = 2$.

For an open toric manifold X, we can exploit the toric symmetries to regularize the infrared singularities on the instanton moduli space $\mathfrak{M}_X$ by "undoing" the T^3-rotations by gauge transformations [5]. In this way we will compute our moduli space integrals by using techniques from equivariant localization, which in the present context will be refered to as *toric localization*. Recall that the bosonic sector of the topologically twisted theory comprises a gauge connection A_i and a complex Higgs field Φ. In particular, the supercharges contain a scalar Q and a vector Q_i. Generically, only Q is conserved and can be used to define the topological twist of the gauge theory. If the threefold X has symmetries then one can also use Q_i. In the generic formulation of the theory, one only considers the scalar topological charge Q and restricts attention to gauge-invariant observables. But in the present case one can also use the linear combination

$$Q_\Omega = Q + \epsilon_a \Omega^a_{ij} x^i Q^j, \tag{3.22}$$

where ϵ^a are the parameters of the isometric action of $T^3 \subset U(3)$ on the Kähler space $\mathbb{C}^3$, and $\Omega^a = \Omega^a_{ij}\, x^j(\partial/\partial x_i)$ are vector fields which generate $SO(6)$ rotations of $\mathbb{C}^3 \cong \mathbb{R}^6$. In this case we also consider observables which are only gauge-invariant up to a rotation. This means that the new observables are equivariant differential forms and the BRST charge Q_Ω can be interpreted as an equivariant differential $\mathrm{d} + \iota_\Omega$ on the space of field configurations, where ι_Ω acts by contraction with the vector field Ω.

This procedure modifies the action and the equations of motion by mixing gauge invariance with rotations. This set of modifications can sometimes be obtained by defining the gauge theory on an appropriate supergravity background called the "Ω-background". In particular, the fixed point equation (3.16) is modified to

$$\mathrm{d}_A \Phi = \iota_\Omega F. \tag{3.23}$$

The set of equations (3.14), (3.15), and (3.23) minimizes the action of the cohomological gauge theory in the Ω-background and describes T^3-invariant instantons (or, as we will see, ideal sheaves). In particular, there is a natural lift of the toric action to the instanton moduli space $\mathfrak{M}_X$. We will henceforth study the gauge theory equivariantly and interpret the truncation of the partition function (3.17) as an equivariant integral over $\mathfrak{M}_X$. This will always mean that we work solely in the Coulomb branch of the gauge theory. Due to the equivariant deformation of the BRST charge, these moduli space integrals can be computed via equivariant localization.

3.5. Equivariant Integration over Moduli Spaces

We now explain the localization formulas that will be used to compute partition functions throughout this paper. Let $\mathfrak{M}$ be a smooth algebraic variety. Then we can define the $\widetilde{T}$-equivariant cohomology $H^\bullet_{\widetilde{T}}(\mathfrak{M},\mathbb{Q})$ as the ordinary cohomology $H^\bullet(\mathfrak{M}_{\widetilde{T}},\mathbb{Q})$ of the Borel-Moore homotopy quotient $\mathfrak{M}_{\widetilde{T}} := (\mathfrak{M}\times E\widetilde{T})/\widetilde{T}$, where $E\widetilde{T} = (\mathbb{C}^\infty \setminus \{0\})^{N+k}$ is a contractible space on which $\widetilde{T} = U(1)^N \times T^k$ acts freely. In the present example of interest, $N=1$ and $k=3$. Given a $\widetilde{T}$-equivariant vector bundle $\mathfrak{E}\rightarrow\mathfrak{M}$, the quotient $\mathfrak{E}_{\widetilde{T}} = (\mathfrak{E}\times E\widetilde{T})/\widetilde{T}$ is a vector bundle over $\mathfrak{M}_{\widetilde{T}} = (\mathfrak{M}\times E\widetilde{T})/\widetilde{T}$. The $\widetilde{T}$-equivariant Euler class of $\mathfrak{E}$ is the invertible element defined by

$$e_{\widetilde{T}}(\mathfrak{E}) := e\big(\mathfrak{E}_{\widetilde{T}}\big) \in H^\bullet_{\widetilde{T}}(\mathfrak{M},\mathbb{Q}), \tag{3.24}$$

where e is the ordinary Euler class for vector bundles (the top Chern class).

Let $B\widetilde{T} := E\widetilde{T}/\widetilde{T} = (\mathbb{P}^\infty)^{k+N}$. Then $E\widetilde{T}\rightarrow B\widetilde{T}$ is a universal principal $\widetilde{T}$-bundle, and there is a fibration $\mathfrak{M}_{\widetilde{T}}\rightarrow B\widetilde{T}$ with fibre $\mathfrak{M}$. Integration in equivariant cohomology is defined as the pushforward $\oint_{\mathfrak{M}}$ of the collapsing map $\mathfrak{M}\rightarrow \mathrm{pt}$, which coincides with integration over the fibres $\mathfrak{M}$ of the bundle $\mathfrak{M}_{\widetilde{T}}\rightarrow B\widetilde{T}$ in ordinary cohomology. Let $p_i : B\widetilde{T}\rightarrow\mathbb{P}^\infty$ for $i=1,\dots,k$ and let $q_l : B\widetilde{T}\rightarrow\mathbb{P}^\infty$ for $l=1,\dots,N$ be the canonical projections onto the ith and lth factors. Introduce equivariant parameters $\epsilon_i = (c_1)_{\widetilde{T}}(p_i^*\mathcal{O}_{\mathbb{P}^\infty}(1))$, with $t_i = \mathrm{e}^{\epsilon_i} = (\mathrm{ch}_{\widetilde{T}})_1(p_i^*\mathcal{O}_{\mathbb{P}^\infty}(1))$ and $a_l = (c_1)_{\widetilde{T}}(q_l^*\mathcal{O}_{\mathbb{P}^\infty}(1))$, with $e_l = \mathrm{e}^{a_l} = (\mathrm{ch}_{\widetilde{T}})_1(q_l^*\mathcal{O}_{\mathbb{P}^\infty}(1))$.

The Atiyah-Bott localization formula in equivariant cohomology states that

$$\oint_{\mathfrak{M}}\alpha = \oint_{\mathfrak{M}^{\widetilde{T}}}\frac{\alpha|_{\mathfrak{M}^{\widetilde{T}}}}{e_{\widetilde{T}}(\mathfrak{N})} \tag{3.25}$$

for any equivariant differential form $\alpha\in H^\bullet_{\widetilde{T}}(\mathfrak{M},\mathbb{Q})$, where the complex vector bundle $\mathfrak{N}\rightarrow\mathfrak{M}^{\widetilde{T}}$ is the normal bundle over the (compact) fixed point submanifold in $\mathfrak{M}$. When $\mathfrak{M}^{\widetilde{T}}$ consists of finitely many isolated points f, this formula is simplified to

$$\oint_{\mathfrak{M}}\alpha = \sum_{f\in\mathfrak{M}^{\widetilde{T}}}\frac{\alpha(f)}{e_{\widetilde{T}}(T_f\mathfrak{M})}. \tag{3.26}$$

Each term in this sum takes values in the polynomial ring

$$H^\bullet_{\widetilde{T}}(f,\mathbb{Q}) = H^\bullet\left(B\widetilde{T},\mathbb{Q}\right) \cong \mathbb{Q}[\epsilon_1,\dots,\epsilon_k,a_1,\dots,a_N] \tag{3.27}$$

in the generators of $\widetilde{T} = U(1)^N \times T^k$. When the manifold $\mathfrak{M}$ is noncompact, integration along the fibre is not a well-defined $\mathbb{Q}$-linear map. Nevertheless, when $\mathfrak{M}^{\widetilde{T}}$ is compact, we can formally *define* the equivariant integral $\oint_{\mathfrak{M}}\alpha$ by the right-hand side of the formula (3.25).

Going back to our example, when $X = \mathbb{C}^3$, one has $\mathrm{ch}_2(\mathcal{E}) = 0$ and the partition function Z is saturated by contributions from isolated, pointlike instantons (D0-branes) by a formal application of the localization formula (3.26). However, these expressions are all rather symbolic, as we are not guaranteed that the algebraic scheme $\mathfrak{M}_X$ is a smooth variety, that is, the instanton moduli space has a well-defined stable tangent bundle with tangent spaces all of the same dimension. However, the variety $\mathfrak{M}_X$ is *generically* smooth and there is a well-defined *virtual* tangent bundle. The moduli space integration (3.26) can then be formally defined by virtual localization in equivariant Chow theory. As discussed in [27], the (stratified components of the) instanton moduli space $\mathfrak{M}_X$ carries a canonical perfect obstruction theory in the sense of [22]. In obstruction theory, the virtual tangent space at a point $[\mathcal{E}] \in \mathfrak{M}_X$ is given by

$$T^{\mathrm{vir}}_{[\mathcal{E}]}\mathfrak{M}_X = \mathrm{Ext}^1_{\mathcal{O}_X}(\mathcal{E},\mathcal{E}) \ominus \mathrm{Ext}^2_{\mathcal{O}_X}(\mathcal{E},\mathcal{E}), \tag{3.28}$$

where $\mathrm{Ext}^1_{\mathcal{O}_X}(\mathcal{E},\mathcal{E})$ is the Zariski tangent space and $\mathrm{Ext}^2_{\mathcal{O}_X}(\mathcal{E},\mathcal{E})$ the obstruction space of $\mathfrak{M}_X$ at $[\mathcal{E}]$. Its dimension is given by the difference of Euler characteristics $\chi(\mathcal{O}_X \otimes \mathcal{O}_X^\vee) - \chi(\mathcal{E} \otimes \mathcal{E}^\vee)$. The kernel of the trace map

$$\mathrm{Ext}^2_{\mathcal{O}_X}(\mathcal{E},\mathcal{E}) \longrightarrow H^2(X,\mathcal{O}_X) \tag{3.29}$$

is the obstruction to smoothness at a point $[\mathcal{E}]$ of the moduli space.

The bundles $\mathfrak{E}_i := \mathrm{Ext}^i_{\mathcal{O}_X}(\mathcal{E},\mathcal{E})$, $i = 1,2$ for $[\mathcal{E}] \in \mathfrak{M}_X$ define a canonical T^k-equivariant perfect obstruction theory $\mathfrak{E}_\bullet = (\mathfrak{E}_1 \to \mathfrak{E}_2)$ (see [22, Section 1]) on the instanton moduli space $\mathfrak{M} = \mathfrak{M}_X$. In this case, one may construct a virtual fundamental class $[\mathfrak{M}]^{\mathrm{vir}}$ and apply a virtual localization formula. The general theory is developed in [22] and requires a T^k-equivariant embedding of $\mathfrak{M}$ in a smooth variety $\mathfrak{Y}$. The existence of such an embedding in the present case follows from the stratification of $\mathfrak{M}_X$ into Hilbert schemes of points and curves. Then one can deduce the localization formula over $\mathfrak{M}$ from the known ambient localization formula over the smooth variety $\mathfrak{Y}$, as above. In this paper we will only need a special case of this general framework, the virtual Bott residue formula.

We can decompose $\mathfrak{E}_i$ into T^k-eigenbundles. The scheme theoretic fixed point locus $\mathfrak{M}^{T^k}$ is the maximal T^k-fixed closed subscheme of $\mathfrak{M}$. It carries a canonical perfect obstruction theory, defined by the T^k-fixed part of the restriction of the complex $\mathfrak{E}_\bullet$ to $\mathfrak{M}^{T^k}$, which may be used to define a virtual structure on $\mathfrak{M}^{T^k}$. The sum of the nonzero T^k-weight spaces of $\mathfrak{E}_\bullet|_{\mathfrak{M}^{T^k}}$ defines the virtual normal bundle $\mathfrak{N}^{\mathrm{vir}}$ to $\mathfrak{M}^{T^k}$. Define the Euler class of a virtual bundle $\mathfrak{A} = \mathfrak{A}_1 \ominus \mathfrak{A}_2$ using formal multiplicativity, that is, as the ratio of the Euler classes of the two bundles, $e(\mathfrak{A}) = e(\mathfrak{A}_1)/e(\mathfrak{A}_2)$. Then the virtual Bott localization formula for the Euler class of a bundle $\mathfrak{A}$ of a rank equal to the virtual dimension of $\mathfrak{M}$ reads [22]

$$\oint_{[\mathfrak{M}]^{\rm vir}} e(\mathfrak{A}) = \oint_{[\mathfrak{M}^{T^k}]^{\rm vir}} \frac{e_{T^k}\left(\mathfrak{A}|_{\mathfrak{M}^{T^k}}\right)}{e_{T^k}(\mathfrak{N}^{\rm vir})}, \tag{3.30}$$

where the integration is again defined via pushforward maps. The equivariant Euler classes on the right-hand side of this formula are invertible in the localized equivariant Chow ring of the scheme $\mathfrak{M}$ given by $CH^{\bullet}_{T^k}(\mathfrak{M})\otimes_{\mathbb{Q}[\epsilon_1,\dots,\epsilon_k]}\mathbb{Q}[\epsilon_1,\dots,\epsilon_k]_{\mathfrak{m}}$, where $\mathbb{Q}[\epsilon_1,\dots,\epsilon_k]_{\mathfrak{m}}$ is the localization of the ring $\mathbb{Q}[\epsilon_1,\dots,\epsilon_k]$ at the maximal ideal $\mathfrak{m}$ generated by $\epsilon_1,\dots,\epsilon_k$.

If $\mathfrak{M}$ is smooth, then $\mathfrak{M}^{T^k}$ is the nonsingular set theoretic fixed point locus, consisting here of finitely many points $[\mathcal{E}]$. However, in general the formula (3.30) must be understood scheme theoretically, here as a sum over T^k-fixed closed subschemes of $\mathfrak{M}$ supported at the points $[\mathcal{E}] \in \mathfrak{M}^{T^k}$ (with $k = 3$). With $\rho^i_{\mathcal{E}} : T^k \rightarrow \mathrm{End}_{\mathbb{C}}(\mathrm{Ext}^i_{\mathcal{O}_X}(\mathcal{E},\mathcal{E}))$, $i = 1,2$, denoting the induced torus actions on the tangent and obstruction bundles on $\mathfrak{M}$, one generically has decompositions

$$\mathrm{Ext}^i_{\mathcal{O}_X}(\mathcal{E},\mathcal{E}) = \mathrm{Ext}^i_1(\mathcal{E},\mathcal{E}) \oplus \ker\left(\rho^i_{\mathcal{E}}\left(T^k\right)\right), \tag{3.31}$$

where $\mathrm{Ext}^i_1(\mathcal{E},\mathcal{E})$ is a T^k-invariant subspace of $\mathrm{Ext}^i_{\mathcal{O}_X}(\mathcal{E},\mathcal{E})$. As demonstrated in [27, Section 4.5], the kernel module in (3.31) vanishes. Hence each subscheme here is just the reduced point $[\mathcal{E}]$ and the T^k-fixed obstruction theory at $[\mathcal{E}]$ is trivial. Under these conditions, the virtual localization formula (3.30) may be written as

$$\oint_{[\mathfrak{M}]^{\rm vir}} e(\mathfrak{A}) = \sum_{[\mathcal{E}]\in\mathfrak{M}^{T^k}} \frac{e_{T^k}(\mathfrak{A}([\mathcal{E}]))}{e_{T^k}\left(T^{\rm vir}_{[\mathcal{E}]}\mathfrak{M}\right)}. \tag{3.32}$$

The right-hand side of this formula again takes values in the polynomial ring $\mathbb{Q}[\epsilon_1,\dots,\epsilon_k]$. When $\mathrm{Ext}^0_{\mathcal{O}_X}(\mathcal{E},\mathcal{E}) = \mathrm{Ext}^2_{\mathcal{O}_X}(\mathcal{E},\mathcal{E}) = 0$ for all $[\mathcal{E}] \in \mathfrak{M}_X$, the moduli space $\mathfrak{M}_X$ is a smooth algebraic variety with the trivial perfect obstruction theory and this equation reduces immediately to the standard localization formula in equivariant cohomology given above. In this paper, we will make the natural choice for the bundle $\mathfrak{A}$, the virtual tangent bundle $T^{\rm vir}\mathfrak{M}$ itself.

3.6. Noncommutative Gauge Theory

To compute the instanton contributions (3.17) to the partition function of the cohomological gauge theory, we have to resolve the small instanton ultraviolet singularities of $\mathfrak{M}_X$. This can be achieved by replacing the space $X = \mathbb{C}^3 \cong \mathbb{R}^6$ with its noncommutative deformation $\mathbb{R}^6_\theta$ defined by letting the coordinate generators x^i, $i = 1,\dots,6$, satisfy the commutation relations of the Weyl algebra

$$\left[x^i, x^j\right] = \mathrm{i}\theta^{ij}, \tag{3.33}$$

where

$$\left(\theta^{ij}\right)=\begin{pmatrix} 0 & \theta_1 & & & & \\ -\theta_1 & 0 & & & & \\ & & 0 & \theta_2 & & \\ & & -\theta_2 & 0 & & \\ & & & & 0 & \theta_3 \\ & & & & -\theta_3 & 0 \end{pmatrix} \tag{3.34}$$

is a constant 6×6 skew-symmetric matrix which we take in Jordan canonical form without loss of generality (by a suitable linear transformation of $\mathbb{R}^6$ if necessary). We will assume that $\theta_1,\theta_2,\theta_3>0$ for simplicity. The noncommutative polynomial algebra

$$\mathcal{A}=\frac{\mathbb{C}\left[x^1,\ x^2,x^3\right]}{\left\langle\left[x^i,x^j\right]-\mathrm{i}\theta^{ij}\right\rangle} \tag{3.35}$$

is regarded as the "algebra of functions" on the noncommutative space $\mathbb{R}^6_\theta$.

We can represent the algebra $\mathcal{A}$ on the standard Fock module

$$\mathcal{H}=\mathbb{C}\left[\alpha_1^\dagger,\alpha_2^\dagger,\alpha_3^\dagger\right]|0,0,0\rangle=\bigoplus_{i,j,k=0}^{\infty}\mathbb{C}\big|i,j,k\big\rangle, \tag{3.36}$$

where the orthonormal basis states $|i,j,k\rangle$ are connected by the action of creation and annihilation operators $\alpha_a^\dagger$ and α_a, $a=1,2,3$. They obey $\alpha_a|0,0,0\rangle=0$ and

$$\left[\alpha_a^\dagger,\alpha_b\right]=\delta_{ab},\qquad [\alpha_a,\alpha_b]=0=\left[\alpha_a^\dagger,\alpha_b^\dagger\right]. \tag{3.37}$$

In the Weyl operator realization with the complex combinations of operators

$$z^a=x^{2a-1}-\mathrm{i}x^{2a}=\sqrt{2\theta_a}\,\alpha_a,\qquad \overline{z}^{\overline{a}}=x^{2a-1}+\mathrm{i}x^{2a}=\sqrt{2\theta_a}\,\alpha_a^\dagger \tag{3.38}$$

for $a=1,2,3$, derivatives of fields are replaced by the inner automorphisms

$$\partial_{z^a}f\longrightarrow\frac{1}{2\theta_a}\delta_{a\overline{b}}\left[\overline{z}^{\overline{b}},f\right], \tag{3.39}$$

while spacetime averages are replaced by traces over $\mathcal{H}$ according to

$$\int_{\mathbb{R}^6}\mathrm{d}^6x\,f(x)\longrightarrow(2\pi)^3\theta_1\theta_2\theta_3\ \mathrm{Tr}_{\mathcal{H}}(f). \tag{3.40}$$

In the noncommutative gauge theory, we introduce the covariant coordinates

$$X^i=x^i+\mathrm{i}\theta^{ij}A_j \tag{3.41}$$

and their complex combinations

$$Z^a = \frac{1}{\sqrt{2\theta_a}} \left(X^{2a-1} + iX^{2a} \right) \tag{3.42}$$

for $a = 1,2,3$. Then the $(1,1)$ and $(2,0)$ components of the curvature two-form can be, respectively, expressed as

$$F_{a\bar{b}} = [Z_a, Z_{\bar{b}}] + \frac{1}{2\theta_a}\delta_{a\bar{b}}, \qquad F_{ab} = [Z_a, Z_b], \tag{3.43}$$

while the covariant derivatives of the Higgs field become

$$(\partial_A)_a \Phi = [Z_a, \Phi]. \tag{3.44}$$

The instanton equations (3.14), (3.15), and (3.23) then become algebraic equations

$$\begin{gathered} \left[Z^a, Z^b\right] = 0, \\ \left[Z^a, Z_a^\dagger\right] = 3, \\ [Z_a, \Phi] = \epsilon_a Z_a. \end{gathered} \tag{3.45}$$

These equations describe BPS bound states of the D0–D6 system in a B-field background, which is necessary for reinstating supersymmetry [28, 29]. In addition, T^3-invariance of the (unique) holomorphic three-form on X imposes the Calabi-Yau condition

$$\epsilon_1 + \epsilon_2 + \epsilon_3 = 0. \tag{3.46}$$

3.7. Instanton Moduli Space

A major technical advantage of introducing the noncommutative deformation is that the instanton moduli space can be constructed explicitly, by solving the noncommutative instanton equations (3.45). First we construct the vacuum solution of the noncommutative gauge theory, with $F = 0$. It is obtained by setting $A = 0$ and is given explicitly by harmonic oscillator algebra

$$Z^a = \alpha_a, \qquad \Phi = \sum_{a=1}^{3} \epsilon_a \alpha_a^\dagger \alpha_a. \tag{3.47}$$

Other solutions are found via the solution generating technique described in, for example, references [30, 31]. For the general solution, fix an integer $n \geq 1$ and let U_n be a partial isometry on the Hilbert space $\mathscr{H}$ which projects out all states $|i, j, k\rangle$ with $i + j + k < n$. Such an operator satisfies the equations

$$U_n^\dagger U_n = 1 - \Pi_n, \qquad U_n U_n^\dagger = 1, \tag{3.48}$$

where

$$\Pi_n = \sum_{i+j+k<n} |i,j,k\rangle\langle i,j,k| \tag{3.49}$$

is a Hermitean projection operator onto a finite-dimensional subspace of $\mathscr{H}$. Then we make the ansatz

$$Z^a = U_n \alpha_a f(N) U_n^\dagger, \qquad \Phi = U_n \left(\sum_{a=1}^{3} \epsilon_a \alpha_a^\dagger \alpha_a \right) U_n^\dagger, \tag{3.50}$$

where f is a real function of the number operator

$$N = \sum_{a=1}^{3} \alpha_a^\dagger \alpha_a. \tag{3.51}$$

Using standard harmonic oscillator algebra, we can write the DUY equations (3.45) as

$$U_n \left(N f^2(N-1) - (N+3) f^2(N) + 3 \right) U_n^\dagger = 0. \tag{3.52}$$

This recursion relation has a unique solution with the initial conditions $f(i) = 0, i = 0, 1, \dots, n-1$, and the finite energy condition $f(r) \to 1$ as $r \to \infty$. It is given by [32]

$$f(N) = \sqrt{1 - \frac{n\ (n+1)(n+2)}{(N+1)(N+2)(N+3)}}(1 - \Pi_n). \tag{3.53}$$

The topological charge of the corresponding noncommutative instanton is

$$\ell(n) = \mathrm{ch}_3(\mathcal{E}) = -\frac{\mathrm{i}}{6}\ \theta_1\theta_2\theta_3\ \mathrm{Tr}_{\mathscr{H}}(F \wedge F \wedge F) = \frac{1}{6} n(n+1)(n+2). \tag{3.54}$$

Thus the instanton number is the number of states in $\mathscr{H}$ with $N < n$, that is, the number of vectors removed by U_n, or equivalently the rank of the projector Π_n.

The partial isometry U_n identifies the full Fock space $\mathscr{H} = \mathbb{C}[\alpha_1^\dagger, \alpha_2^\dagger, \alpha_3^\dagger]|0,0,0\rangle$ with the subspace

$$\mathscr{H}_{\mathcal{I}} = \bigoplus_{f \in \mathcal{I}} f\left(\alpha_1^\dagger, \alpha_2^\dagger, \alpha_3^\dagger\right)|0,0,0\rangle, \tag{3.55}$$

where

$$\mathcal{I} = \mathbb{C}\left\langle w_1^i\ w_2^j\ w_3^k \mid i+j+k \geq n \right\rangle \tag{3.56}$$

is a monomial ideal of codimension $\ell = \ell(n)$ in the polynomial ring $\mathbb{C}[w_1, w_2, w_3]$. The instanton moduli space can thus be identified as the Hilbert scheme

$$\mathfrak{M}_X = \mathrm{Hilb}_{\ell,0}(X) = X^{[\ell]} \tag{3.57}$$

of ℓ points in $X = \mathbb{C}^3$. The Hilbert-Chow morphism

$$X^{[\ell]} \longrightarrow \mathrm{Sym}^{\ell}(X) = \frac{X^{\ell}}{S_{\ell}} \tag{3.58}$$

identifies the Hilbert scheme of points as a crepant resolution of the (coincident point) singularities of the ℓth symmetric product orbifold of X. The ideal $\mathcal{I}$ defines a plane partition π with $|\pi| = \ell$ boxes given by

$$\pi = \left\{ (i,j,k) \mid i,j,k \geq 1, w_1^{i-1} w_2^{j-1} w_3^{k-1} \notin \mathcal{I} \right\}. \tag{3.59}$$

Heuristically, this configuration represents instantons which sit on top of each other at the origin of $\mathbb{C}^3$, and along its coordinate axes where they asymptote to four-dimensional noncommutative instantons at infinity described by ordinary Young tableaux λ.

3.8. Donaldson-Thomas Theory

We can finally compute the instanton contributions to the partition function of the cohomological gauge theory on any toric Calabi-Yau threefold X [7, 8, 24]. Let us start with the case $X = \mathbb{C}^3$. Using (3.54), the contribution of an instanton corresponding to a plane partition π contributes a factor

$$\exp\left(-\frac{\mathrm{i} g_s}{48\pi^3} \mathrm{Tr}_{\mathcal{H}_{\mathcal{I}}}\left(F^3\right)\right) = \mathrm{e}^{-g_s|\pi|} \tag{3.60}$$

to the Boltzmann weight appearing in the functional integral. There is also a measure factor which comes from integrating out the bosonic and fermionic fields in the supersymmetric gauge theory. This yields a ratio of fluctuation determinants

$$\begin{aligned} Z_\pi &= \frac{\det(\mathrm{ad}\,\Phi)\prod_{i<j} \det\left(\mathrm{ad}\,\Phi + \epsilon_i + \epsilon_j\right)}{\det(\mathrm{ad}\,\Phi + \epsilon_1 + \epsilon_2 + \epsilon_3)\prod_{i=1}^{3} \det(\mathrm{ad}\,\Phi + \epsilon_i)} \\ &= \exp\left(-\int_0^\infty \frac{\mathrm{d}t}{t} \frac{\mathrm{ch}_{\mathcal{I}}(t)\mathrm{ch}_{\mathcal{I}}(-t)}{(1-\mathrm{e}^{t\epsilon_1})(1-\mathrm{e}^{t\epsilon_2})(1-\mathrm{e}^{t\epsilon_3})}\right) \end{aligned} \tag{3.61}$$

with the normalized character

$$\begin{aligned}\mathrm{ch}_{\mathcal{J}}(t) &= \prod_{i=1}^{3}\left(1-\mathrm{e}^{t\epsilon_i}\right)\mathrm{Tr}_{\mathcal{H}_{\mathcal{J}}}\left(\mathrm{e}^{t\Phi}\right) \\ &= 1-\prod_{i=1}^{3}\left(1-\mathrm{e}^{t\epsilon_i}\right)\sum_{(i,j,k)\in\pi}\mathrm{e}^{t(\epsilon_1(i-1)+\epsilon_2(j-1)+\epsilon_3(k-1))},\end{aligned} \tag{3.62}$$

where we have used the solution for Φ in (3.50). Using the Calabi-Yau condition $\epsilon_1+\epsilon_2+\epsilon_3=0$, it is easy to see that these determinants cancel up to a sign.

After some computation, one can explicitly determine this sign to get

$$Z_\pi = Z_{\pi=\emptyset}\cdot(-1)^{|\pi|}. \tag{3.63}$$

The contribution $Z_\emptyset$ from the empty partition is the one-loop perturbative contribution to the functional integral, and hence will be dropped. Then the instanton sum for the partition function is given by

$$Z_{\mathrm{DT}}^{\mathbb{C}^3}(q) = \sum_\pi\left(-\mathrm{e}^{-g_s}\right)^{|\pi|} = \sum_\pi q^{|\pi|} \tag{3.64}$$

which is just the MacMahon function $M(q)$ with $q=-\mathrm{e}^{-g_s}$. This is the known formula for the Donaldson-Thomas partition function on $\mathbb{C}^3$.

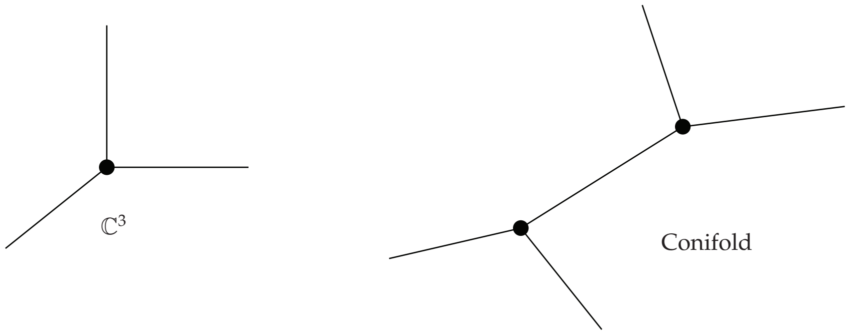
For later use, let us note a convenient resummation formula for this partition function [6]. Using interlacing relations, the sum over plane partitions π can be converted into a triple sum over the Young tableaux obtained from the main diagonal slice $\lambda=(\pi_{i,i})_{i\geq 1}$, together with a sum over pairs of semistandard tableaux of shape λ obtained by putting $m+1$ of them in boxes of the skew diagram associated to the mth diagonal slice for each $m\geq 0$. The partial sum over each semistandard tableaux coincides with the combinatorial definition of the *Schur functions* at a particular value, which can be expressed through the hook formula

$$s_\lambda(q^\rho) = q^{n(\lambda)+|\lambda|/2}\prod_{(i,j)\in\lambda}\frac{1}{1-q^{h(i,j)}}, \tag{3.65}$$

where $n(\lambda)=\sum_i(i-1)\lambda_i$, and $h(i,j)$ is the hook length of the box located at position (i,j) in the Young tableau $\lambda\subset\mathbb{Z}_{\geq 0}^2$. Then the partition function can be rewritten as a sum over *ordinary* partitions

$$Z_{\mathrm{DT}}^{\mathbb{C}^3}(q) = \sum_\lambda s_\lambda(q^\rho)^2. \tag{3.66}$$

This construction can be generalized to *arbitrary* toric Calabi-Yau threefolds X by using the gluing rules of toric geometry. The two simplest such varieties are described by the toric diagrams



with a single vertex representing $X = \mathbb{C}^3$, whose partition function was computed above and is the basic building block for the generating functions on more complicated geometries, and a single line joining two vertices representing the resolved conifold $X = \mathcal{O}_{\mathbb{P}^1}(-1) \oplus \mathcal{O}_{\mathbb{P}^1}(-1)$, where the $\mathbb{P}^1$ contribute to the $F \wedge F \wedge \omega_0$ term of the gauge theory action (3.13). The T^3-invariant noncommutative $U(1)$ instantons on X correspond to ideal sheaves $\mathcal{E}$ and are described by the following combinatorial data:

(1) three-dimensional partitions π_f at each vertex f of the toric diagram, corresponding to monomial ideals $\mathcal{I}_f \subset \mathbb{C}[w_1, w_2, w_3]$,

(2) two-dimensional partitions λ_e at each edge e of the toric diagram, representing the four-dimensional instanton asymptotics of π_f.

This description requires generalizing the calculation on $X = \mathbb{C}^3$ above to compute the *perpendicular partition function* $P_{\lambda,\mu,\nu}(q)$ [6], which is defined to be the generating function for three-dimensional partitions with fixed asymptotics λ, μ, and ν in the three coordinate directions. Such partitions correspond to instantons on $\mathbb{C}^3_\theta$ with nontrivial boundary conditions at infinity along each of the coordinate axes. It can be expressed in terms of *skew Schur functions*, with $Z^{\mathbb{C}^3}_{\mathrm{DT}}(q) = P_{\emptyset,\emptyset,\emptyset}(q)$.

For the example of the resolved conifold, using the gluing rules one easily computes

$$\begin{aligned} Z^{\text{conifold}}_{\mathrm{DT}}(q,Q) &= \sum_{\pi_f} q^{|\pi_f| + \sum_{(i,j)\in\lambda}(i+j+1)} (-1)^{|\lambda|} Q^{|\lambda|} \\ &= \prod_{n=1}^{\infty} \frac{\left(1-(-1)^n q^n Q\right)^n}{\left(1-(-1)^n q^n\right)^{2n}} = M(-q)^2 \prod_{n=1}^{\infty} \left(1-(-q)^n Q\right)^n. \end{aligned} \tag{3.67}$$

More generally, with these rules one finds that the instanton partition function is the generating function

$$Z^{X}_{\mathrm{DT}}(q,\mathbf{Q}) = \sum_{n\in\mathbb{Z}} \sum_{\beta\in H_2(X,\mathbb{Z})} I_{n,\beta}(X) q^n \mathbf{Q}^\beta \tag{3.68}$$

for the Donaldson-Thomas invariants $I_{n,\beta}(X) \in \mathbb{Z}$, which are defined as follows. The moduli variety $\mathrm{Hilb}_{n,\beta}(X)$ of ideal sheaves on X is a projective scheme with a perfect obstruction theory. For general threefolds X, it has virtual dimension [33, Lemma 1]

$$\int_\beta c_1(X) \tag{3.69}$$

which coincides with that of $\mathfrak{M}_g(X,\beta)$ from Section 2.1. In the Calabi-Yau case, the virtual dimension is zero, and the corresponding virtual cycle is

$$[\mathrm{Hilb}_{n,\beta}(X)]^{\mathrm{vir}} \in CH_0\big(\mathrm{Hilb}_{n,\beta}(X)\big). \tag{3.70}$$

Then the Donaldson-Thomas invariants

$$I_{n,\beta}(X) := \int_{[\mathrm{Hilb}_{n,\beta}(X)]^{\mathrm{vir}}} 1 \tag{3.71}$$

count the virtual numbers of ideal sheaves on X with the given Chern character. The right-hand side is defined via equivariant integration, as explained in Section 3.5. The torus action on X lifts to the moduli scheme $\mathrm{Hilb}_{n,\beta}(X)$. The $\widetilde{T}$-fixed locus $\mathrm{Hilb}_{n,\beta}(X)^{\widetilde{T}}$ has a $\widetilde{T}$-equivariant virtual theory with cycle $[\mathrm{Hilb}_{n,\beta}(X)^{\widetilde{T}}]^{\mathrm{vir}} \in CH_0(\mathrm{Hilb}_{n,\beta}(X)^{\widetilde{T}})$ and virtual normal bundle $\mathfrak{N}_X^{\mathrm{vir}}$ in the equivariant K-theory $K^0_{\widetilde{T}}(\mathrm{Hilb}_{n,\beta}(X)^{\widetilde{T}})$. This construction gives precise meaning to the moduli space integral (3.17) via application of the virtual localization formula in equivariant Chow theory, described in Section 3.5.

3.9. Wall-Crossing Formulas

We will now make contact with Section 2. For the present class of threefolds X, there is a gauge theory/string theory duality [27]. This follows from the fact that the perpendicular partition function $P_{\lambda,\mu,\nu}(q)$ is related to the Calabi-Yau crystal formulation of the *topological vertex*

$$C_{\lambda,\mu,\nu}(q) = M(q)^{-1} q^{(1/2)\,(\|\lambda\|^2+\|\mu\|^2+\|\nu\|^2)} P_{\lambda,\mu,\nu}(q) \tag{3.72}$$

with $\|\lambda\|^2 = \sum_i \lambda_i^2$, which are the building blocks for the computation of the generating function for Gromov-Witten invariants using rules analogous to those described in Section 3.8 [14, 15]. Using these relations one can show that the six-dimensional cohomological gauge theory is S-dual to the A-model topological string theory. The respective partition functions are related by

$$Z_{\mathrm{top}}^{X}(g_s,\mathbf{Q}) = M(q)^{-\chi(X)} Z_{\mathrm{DT}}^{X}(q=-\mathrm{e}^{-g_s},\mathbf{Q}), \tag{3.73}$$

where the Euler characteristic $\chi(X)$ of X is the number of vertices in its toric diagram. For the conifold example, the gluing rules for the topological vertex yield [14, 15]

$$Z_{\mathrm{top}}^{\mathrm{conifold}}(g_s,Q) = \sum_{\lambda} C_{\emptyset,\emptyset,\lambda}(q) C_{\emptyset,\emptyset,\lambda}(q) Q^{|\lambda|} = \exp\left(\sum_{n=1}^{\infty}\frac{Q^n}{n}\,\frac{1}{\left(q^{n/2}-q^{-n/2}\right)^2}\right), \tag{3.74}$$

which should be compared with (3.67). This Gromov-Witten/Donaldson-Thomas correspondence is known to hold for arbitrary toric threefolds [34]. The relationship (3.73) can be thought of as a wall-crossing formula, as we now explain.

The relationship (3.73) is in apparent contradiction with the OSV conjecture (2.25) if we wish to interpret the right-hand side as the generating function $Z_{\mathrm{BH}}(1,0,\phi^2,\phi^0)$ of a suitable index for black hole microstates. However, the conjectural relations (2.25) and (3.73) hold in different regimes of validity. The number of BPS particles in four dimensions formed by wrapping supersymmetric bound states of D-branes around holomorphic cycles of X depends on the choice of a stability condition, and the BPS countings for different stability conditions are related by wall-crossing formulas. For example, stability of black holes requires that their chemical potentials μ^I lie in the ranges $Q_0\phi^0 > 0$ and $Q_2^i\phi_i^2 > 0$.

On the other hand, the validity of (3.73) is related to the existence of BPS invariants $B_{g,\beta}(X) \in \mathbb{Z}$ such that the topological string amplitudes have an expansion given by [35, 36]

$$Z_{\mathrm{top}}^X(g_s,\mathbf{Q}) = \sum_{g=0}^{\infty}\sum_{\beta\in H_2(X,\mathbb{Z})}\mathbf{Q}^{\beta}\sum_{\substack{\gamma\in H_2(X,\mathbb{Z})\setminus\{0\}\\ \beta=k\,\gamma}} B_{g,\beta}(X)\frac{1}{k}\left(2\sinh\left(\frac{kg_s}{2}\right)\right)^{2g-2}, \tag{3.75}$$

of which the conifold partition function (3.67) is an explicit case. These are partition functions of D6-D2-D0 brane bound states only for certain Kähler moduli. Analyses of Calabi-Yau compactifications of Type II string theory show that the Hilbert spaces of BPS states jump discontinuously across real codimension one walls in the moduli space of vacua, known as walls of marginal stability. The noncommutative instantons do not account for walls of marginal stability extending to infinity. One should instead apply some sort of stability condition (such as Π-stability) to elements of the bounded derived category of coherent sheaves $\mathrm{D}^b(\mathrm{coh}(X))$ of the given charge, which gives a topological classification of A-model D-branes on X. These issues are discussed in more detail in [37–39].

From a mathematical perspective, we can study this phenomenon by looking at framed moduli spaces, which consist of instantons that are trivial "at infinity". More precisely, we can consider a toric compactification of X obtained by adding a compactification divisor D_∞, and consider sheaves $\mathcal{F}$ with a fixed trivialization on D_∞. The Kähler polarization defined by ω_0 allows us to define the moduli space $\mathfrak{M}_X^{\omega_0} = \mathfrak{M}_X$ of stable sheaves. Then the symbolic definition of the gauge theory partition function (3.17) as a particular Euler characteristic can be made precise in the more local definition of Donaldson-Thomas invariants given by [40].

As a scheme with a perfect obstruction theory, the instanton moduli space $\mathfrak{M}_X$ can be viewed locally as the scheme theoretic critical locus of a holomorphic function, the superpotential W, on a compact manifold $\mathcal{X}$ with the action of a gauge group $\mathcal{G}$. $\mathfrak{M}_X$ has virtual dimension zero, and at nonsingular points, the obstruction sheaf $\mathfrak{N}_X$ on $\mathfrak{M}_X$ coincides with the cotangent bundle. Hence if $\mathfrak{M}_X$ were everywhere nonsingular, then the partition function (3.17) would just compute the signed Euler characteristic $(-1)^{\dim_{\mathbb{C}}(\mathfrak{M}_X)}\,\chi(\mathfrak{M}_X)$. At singular points, however, the invariants differ from these characteristics.

There is a constructible function $\nu : \mathfrak{M}_X \to \mathbb{Z}$ which can be used to define the *weighted* Euler characteristic

$$\chi(\mathfrak{M}_X,\nu) := \sum_{n\in\mathbb{Z}} n\cdot\chi\left(\nu^{-1}(n)\right). \tag{3.76}$$

For sheaves of fixed Chern character, this coincides with the curve-counting invariants $I_{n,\beta}(X)$. At nonsingular points, $\nu = (-1)^{\dim_{\mathbb{C}}(\mathfrak{M}_X)}$, while at singular points it is the more complicated function given by

$$\nu(\mathcal{E}) = (-1)^{\dim_{\mathbb{C}}(\mathcal{X}/\mathcal{G})} \left(1 - \mathrm{MF}_W(\mathcal{E})\right), \tag{3.77}$$

where $\mathrm{MF}_W(\mathcal{E})$ is the Milnor fibre of the superpotential W at the point corresponding to $\mathcal{E}$. The weighted Euler characteristic is a deformation invariant of X.

In this approach, one can use topological Euler characteristics to define $I_{n,\beta}(X)$ as invariants associated to moduli varieties of framed sheaves. Fixing $\beta \in H_2(X,\mathbb{Z})$ and $n \in \mathbb{Z}$, the variety $\mathrm{Hilb}_{n,\beta}(X)$ equivalently parametrizes isomorphism classes of the following objects:

(a) surjections (framings)

$$\mathcal{O}_X \longrightarrow \mathcal{F} \longrightarrow 0 \tag{3.78}$$

with $\mathrm{ch}(\mathcal{F}) = (1,0,\beta,n)$,

(b) stable sheaves $\mathcal{E}$ with $\mathrm{ch}(\mathcal{E}) = (1,0,-\beta,-n)$ and trivial determinant,

(c) subschemes $S \subset X$ of dimension ≤ 1 with curve class $[S] = \beta$ and holomorphic Euler characteristic $\chi(\mathcal{O}_S) = n$.

The equivalences between these three descriptions are described explicitly for $X = \mathbb{C}^3$ in [7, 8].

As we vary the polarization ω_0, the moduli spaces $\mathfrak{M}_X^{\omega_0}(X)$ change and so do the associated counting invariants, leading to a wall-and-chamber structure. The wall-crossing behaviour of the enumerative invariants is studied in [41, 42]. The analog of varying ω_0 for framed sheaves is to consider quotients of the structure sheaf $\mathcal{O}_X$ in different abelian subcategories of the bounded derived category $\mathrm{D}^b(\mathrm{coh}(X))$ of coherent sheaves on X. The analog of wall-crossing gives the Pandharipande-Thomas theory of stable pairs [43, 44] and the BPS invariants above. For this, the quotients of $\mathcal{O}_X$ are the stable pairs $(\mathcal{E},\alpha)$, where $\mathcal{E}$ is a coherent $\mathcal{O}_X$-module of pure dimension one with $\mathrm{ch}_2(\mathcal{E}) = -\beta$ and $\chi(\mathcal{E}) = -n$, and $\alpha : \mathcal{O}_X \to \mathcal{E}$ is a nonzero sheaf map such that $\mathrm{coker}(\alpha)$ is of pure dimension zero, together with le Poitier's δ-stability condition for coherent systems. In this case the change of Donaldson-Thomas invariants is described by the Kontsevich-Soibelman wall-crossing formula [42].

To cast these constructions into the language of noncommutative instantons, a proper definition of noncommutative toric manifolds is desired, beyond the heuristic approach presented above whereby only open $\mathbb{C}^3$ patches are deformed. Isospectral type deformations of toric geometry, and instantons therein, are investigated in [45]. It may also aid in the classification of $U(N)$ noncommutative instantons on $\mathbb{C}^3$ for rank $N > 1$, along the lines of what was done in Section 3.7. (See [7, 8] for some explicit examples.) This appears to be related to the problem of defining a nonabelian version of Donaldson-Thomas theory which counts higher-rank torsion-free sheaves, for which no general, appropriate notion of stability is yet known.

4. D4-Brane Gauge Theory and Euler Characteristics

In this section we will take $Q_6 = 0$ (no D6-branes) and consider N D4-branes wrapping a four-cycle $C \subset X$. In this case the worldvolume gauge theory on the D4-branes is the $\mathcal{N} = 4$ Vafa-Witten topologically twisted $U(N)$ Yang-Mills theory on C, where the topological twist is generically required in order to realize covariantly constant spinors on a curved geometry. When the gauge theory is formulated on an arbitrary toric singularity C in four dimensions, we may regard C as a four-cycle inside the Calabi-Yau threefold $X = K_C$, and we will obtain an explicit description of the instanton moduli spaces and their Euler characteristics. The precise forms of the partition functions will be amenable to checks of the OSV conjecture (2.25), and hence a description of wall-crossing phenomena.

4.1. $\mathcal{N} = 4$ *Supersymmetric Yang-Mills Theory on Kähler Surfaces*

Vafa and Witten [46] introduced a topologically twisted version of $\mathcal{N} = 4$ supersymmetric Yang-Mills theory in four dimensions. The twisting procedure modifies the quantum numbers of the fields in the physical theory in such a way that a particular linear combination of the supercharges becomes a scalar. This scalar supercharge is used to define the cohomological field theory and its observables on an arbitrary four-manifold C. In the following we will only consider the case where C is a connected smooth Kähler manifold with Kähler two-form k_0. When certain conditions are met, the partition function of the twisted gauge theory computes the Euler characteristic of the instanton moduli space.

Let g_{ij} be the Kähler metric of (C, k_0). Then the twisted gauge theory corresponds to the moduli problem associated with the equations

$$\begin{gathered}\sigma_{ij} := F^+_{ij} + \frac{1}{4}\left[B^+_{ik}, B^+_{jl}\right] g^{kl} + \frac{1}{2}\left[\Phi, B^+_{ij}\right] = 0,\\ \kappa_i := \mathrm{d}^j_A B^+_{ij} + (\mathrm{d}_A)_i \Phi = 0,\end{gathered} \tag{4.1}$$

where $F^+ = (1/2)(F - \star F)$ is the self-dual part of the curvature two-form with respect to the Kähler metric. The field space $\mathfrak{W}$ is spanned by a connection A_i on a principal G-bundle $\mathcal{P} \to C$, a scalar field Φ, and a self-dual two-form B^+_{ij}, so that

$$\mathfrak{W} = \mathfrak{A}_{\mathcal{P}} \times \Omega^0(C, \mathrm{ad}\ \mathcal{P}) \times \Omega^{2,+}(C, \mathrm{ad}\ \mathcal{P}), \tag{4.2}$$

where $\mathfrak{A}_{\mathcal{P}}$ denotes the space of connections on $\mathcal{P}$ and ad $\mathcal{P}$ is the adjoint bundle of $\mathcal{P}$. Their superpartners ψ_i, ζ, and $\widetilde{\psi}^+_{ij}$ live in the tangent space to $\mathfrak{W}$. Associated with the equations of motion are two multiplets (χ^+_{ij}, H^+_{ij}) and $(\widetilde{\chi}_i, \widetilde{H}_i)$ which are sections of the bundle

$$\mathfrak{F} = \Omega^{2,+}(C, \mathrm{ad}\ \mathcal{P}) \oplus \Omega^1(C, \mathrm{ad}\ \mathcal{P}). \tag{4.3}$$

Schematically, the action of the topological gauge theory is of the form

$$S = \{Q, \Psi\} + \int_C \mathrm{Tr}(F \wedge F) + \int_C \mathrm{Tr}(F \wedge k_0), \tag{4.4}$$

where Q is the scalar supercharge singled out by the twisting procedure. The gauge fermion Ψ is a suitable functional of the fields which contains the term

$$\int_C \sqrt{g}\ \operatorname{Tr}\left(\chi^+_{ij}\left(H^{+ij} + \sigma^{ij}\right) + \widetilde{\chi}_i\left(\widetilde{H}^i + \kappa^i\right)\right), \tag{4.5}$$

that makes the gauge theory localize onto the solutions of (4.1).

Geometrically, the partition function can be interpreted as a Mathai-Quillen representative of the Thom class of the bundle $\mathfrak{V} = \mathfrak{W}\times_{\mathcal{G}}\mathfrak{F}$, where $\mathcal{G} = \operatorname{Aut}(\mathcal{P})$ is the group of gauge transformations. Its pullback via the sections in (4.1) gives the Euler class of $\mathfrak{V}$. Under favourable circumstances, appropriate vanishing theorems hold [46] which ensure that each solution of the system (4.1) has $\Phi = B^+ = 0$ and corresponds to an instanton, that is, a solution to the self-duality equations $F^+ = 0$. In this case the gauge theory localizes onto the instanton moduli space $\mathfrak{M}_C$ and the Boltzmann weight gives a representative of the Euler class of the tangent bundle $T\mathfrak{M}_C$. Therefore, the partition function computes moduli space integrals of the form

$$\int_{\mathfrak{M}_C} e(T\mathfrak{M}_C) = \chi(\mathfrak{M}_C), \tag{4.6}$$

which gives the Euler characteristic of the instanton moduli space. Since the instanton moduli space is not generally a smooth variety, most of the quantities introduced above can only be defined formally. We will discuss how to define these integrations more precisely later on. In particular, we will allow for nontrivial vacuum expectation values for the Higgs field Φ, in order to define the partition function in the Ω-background as before. We will assume that the vanishing theorems can be extended to this case as well, by replacing the instanton moduli space with its compactification obtained by adding torsion-free sheaves on C as before.

The Euler characteristic of instanton moduli space can be computed through the index of the deformation complex associated with $\mathcal{N} = 4$ topological Yang-Mills theory via (4.1). It has the form [47]

$$\Omega^0(C, \operatorname{ad}\mathcal{P}) \xrightarrow{\ \mathrm{D}\ } \begin{matrix} \Omega^1(C, \operatorname{ad}\mathcal{P}) \\ \oplus \\ \Omega^0(C, \operatorname{ad}\mathcal{P}) \\ \oplus \\ \Omega^{2,+}(C, \operatorname{ad}\mathcal{P}) \end{matrix} \xrightarrow{\ \mathrm{s}\ } \begin{matrix} \Omega^{2,+}(C, \operatorname{ad}\mathcal{P}) \\ \oplus \\ \Omega^1(C, \operatorname{ad}\mathcal{P}) \end{matrix} \tag{4.7}$$

where the first morphism is an infinitesimal gauge transformation

$$\mathrm{D}(\phi) = \begin{pmatrix} \mathrm{d}_A\phi \\ [\Phi, \phi] \\ [B^+, \phi] \end{pmatrix}, \tag{4.8}$$

while the second morphism corresponds to the linearization of the sections (σ_{ij}, κ_i) given by

$$\mathrm{s}\left(\psi, \zeta, \widetilde{\psi}^+\right) = p^+\mathrm{d}_A\psi - \left[\widetilde{\psi}^+, B^+\right] + \left[\widetilde{\psi}^+, \Phi\right] + \left[B^+, \zeta\right] + \mathrm{d}_A\zeta + \left[\psi, \Phi\right] + p^+\mathrm{d}_A^*\widetilde{\psi}^+ + \left[\psi, B^+\right], \tag{4.9}$$

with p^+ giving the projection of a two-form onto its self-dual part. Under the assumption that all solutions of the original system of (4.1) have $\Phi = B^+ = 0$, the complex (4.7) splits into the Atiyah-Hitchin-Singer instanton deformation complex

$$\Omega^0(C, \mathrm{ad}\, \mathcal{P}) \xrightarrow{\mathrm{d}_A} \Omega^1(C, \mathrm{ad}\, \mathcal{P}) \xrightarrow{p^+ \circ \mathrm{d}_A} \Omega^{2,+}(C, \mathrm{ad}\, \mathcal{P}) \tag{4.10}$$

plus

$$\Omega^0(C, \mathrm{ad}\, \mathcal{P}) \oplus \Omega^{2,+}(C, \mathrm{ad}\, \mathcal{P}) \xrightarrow{(\mathrm{d}_A, p^+ \circ \mathrm{d}_A^*)} \Omega^1(C, \mathrm{ad}\, \mathcal{P}) \tag{4.11}$$

which is again the instanton deformation complex. One can compute the index of the original complex (4.7) (assuming the Vafa-Witten vanishing theorems) by computing the index of the two complexes above. However, these are equal and contribute with opposite signs. This means that the Euler characteristic of instanton moduli space receives contributions only from isolated points and simply counts the number of such points. On a toric surface C, this is anticipated from the toric localization formula (3.26) and will be made explicit below.

In the applications to black hole microstate counting, we will consider gauge group $G = U(N)$. The chemical potential $\int_C C_{(2)} \wedge \mathrm{Tr}(F)$ for the D2-branes requires taking the Ramond-Ramond field $C_{(2)}$ proportional to the two-form k^i on C which are dual to the basis two-cycle S_i, in order to get the correct charges. In this case the D0-brane charges

$$Q_0 = \frac{1}{8\pi^2} \int_C \mathrm{Tr}(F \wedge F) \tag{4.12}$$

correspond to the instanton numbers of the gauge bundle $\mathcal{P}$, while the D2-brane charges

$$Q_2^i = \frac{1}{2\pi} \int_{S_i} \mathrm{Tr}(F) \tag{4.13}$$

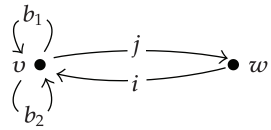
correspond to nontrivial magnetic fluxes $c_1(\mathcal{P}) \neq 0$ through S_i. To compute the macroscopic black hole entropy from the counting of the corresponding BPS states in the gauge theory, we introduce observables associated to these sources and compute their gauge theory expectation values using the localization arguments above to get

$$\begin{aligned} Z_{\mathrm{BH}}\left(N, \phi^2, \phi^0\right) &= \left\langle \exp\left(-\frac{\phi^0}{8\pi^2} \int_C \mathrm{Tr}(F \wedge F) - \frac{\phi_i^2}{2\pi} \int_C k^i \wedge \mathrm{Tr}(F) \right) \right\rangle_{\mathrm{SYM}} \\ &= \sum_{Q_0, Q_2^i} \Omega(Q_0, Q_2; N)\, \mathrm{e}^{-Q_0\, \phi^0 - Q_2^i \phi_i^2}, \end{aligned} \tag{4.14}$$

where $\Omega(Q_0, Q_2; N)$ is the Witten index which computes the Euler characteristic of the moduli space $\mathfrak{M}_{N,Q_2,Q_0}(C)$ of $U(N)$ instantons on C with Chern invariants $c_1(\mathcal{P}) = Q_2 \in H^2(X, \mathbb{Z})$ and $-\mathrm{ch}_2(\mathcal{P}) = Q_0 \nu \in H^4(X, \mathbb{Z})$. Here ν is the generator of $H^4(X, \mathbb{Z}) \cong \mathbb{Z}$ which is Poincaré dual to a point in X.

4.2. Toric Localization and the Instanton Moduli Space

Instantons on $C = \mathbb{C}^2$ can be described as follows. Consider the quiver



with the single relation r specified by the linear combination of paths

$$r = [b_1, b_2] + ij. \tag{4.15}$$

This is called the ADHM quiver $\mathsf{Q}_{\mathrm{ADHM}}$. The *ADHM construction* establishes a one-to-one correspondence between stable framed representations of the quiver $\mathsf{Q}_{\mathrm{ADHM}}$ in the category $\mathrm{Vect}_{\mathbb{C}}$ of finite-dimensional complex vector spaces and framed torsion-free sheaves on the projective plane $\mathbb{P}^2$. In the rank one case, these are equivalent to ideal sheaves on $\mathbb{C}^2$, and the correspondence gives an isomorphism with the Hilbert schemes of points on $\mathbb{C}^2$.

Let V and W be inner product spaces of complex dimensions $k = \mathcal{Q}_0$ and N, respectively. The instanton moduli space $\mathfrak{M}_{N,k}(\mathbb{C}^2)$ can be realized as a hyperKähler quotient by the natural action of $U(k)$ on the variety consisting of linear operators

$$B_1, B_2 \in \mathrm{Hom}_{\mathbb{C}}(V, V), \qquad I \in \mathrm{Hom}_{\mathbb{C}}(W, V), \quad J \in \mathrm{Hom}_{\mathbb{C}}(V, W), \tag{4.16}$$

constrained by the ADHM equations

$$\begin{aligned} \mu_c &= [B_1, B_2] + IJ = 0, \\ \mu_r &= \left[B_1, B_1^\dagger\right] + \left[B_2, B_2^\dagger\right] + II^\dagger - J^\dagger J = 0. \end{aligned} \tag{4.17}$$

On $\mathbb{C}^2$ one can obtain a better compactification of this moduli space by deforming the gauge theory to a noncommutative field theory, as before [2, 3, 48]. This is equivalent to a modification of the hyperKähler quotient that defines the instanton moduli space, obtained by changing the images of the moment maps of (4.17) to

$$\mu_c = 0, \qquad \mu_r = \zeta \ \mathrm{id}_V, \tag{4.18}$$

where $\zeta = \theta_1 + \theta_2$. This quotient gives a compactification of the instanton moduli space $\mathfrak{M}_{N,k}(\mathbb{C}^2)$ obtained by blowing up its singularities.

The classification of toric fixed points is given in [49], by identifying the instanton moduli space $\mathfrak{M}_{1,k}(\mathbb{C}^2)$ with the Hilbert scheme of points $(\mathbb{C}^2)^{[k]}$. The fixed points are point-like instantons which are in one-to-one correspondence with Young tableaux λ having $|\lambda| = k$ boxes. In the more general case of a $U(N)$ gauge theory in the Coulomb branch, one takes N copies of the $U(1)$ theory and the fixed points are classified in terms of N-tuples of Young diagrams $\vec{\lambda} = (\lambda_1, \ldots, \lambda_N)$, called N-coloured Young diagrams. One can show that the fixed points are isolated.

One can describe the local structure of the instanton moduli space by using the ADHM construction at the fixed points of the $\widetilde{T} = U(1)^N \times T^2$ action. Following [49], one introduces a two-dimensional T^2-module $\underline{Q}$ to keep track of the toric action. The operators (B_1, B_2, I, J) corresponding to a fixed point configuration are elements of the $\widetilde{T}$-modules

$$(B_1, B_2) \in \mathrm{Hom}_{\mathbb{C}}(V, V) \otimes \underline{Q}, \qquad I \in \mathrm{Hom}_{\mathbb{C}}(W, V), \qquad J \in \mathrm{Hom}_{\mathbb{C}}(V, W) \otimes \bigwedge^2 \underline{Q}. \tag{4.19}$$

Then the local structure of the instanton moduli space is described by the complex

$$\mathrm{Hom}_{\mathbb{C}}(V, V) \xrightarrow{\sigma} \begin{array}{c} \mathrm{Hom}_{\mathbb{C}}(V, V) \otimes \underline{Q} \\ \oplus \\ \mathrm{Hom}_{\mathbb{C}}(W, V) \\ \oplus \\ \mathrm{Hom}_{\mathbb{C}}(V, W) \otimes \bigwedge^2 \underline{Q} \end{array} \xrightarrow{\tau} \mathrm{Hom}_{\mathbb{C}}(V, V) \otimes \bigwedge^2 \underline{Q} \tag{4.20}$$

which is just a finite-dimensional version of the Atiyah-Hitchin-Singer instanton deformation complex (4.10). The map σ corresponds to infinitesimal (complex) gauge transformations while τ is the linearization of the ADHM constraint $\mu_c = 0$. In general, the complex (4.20) has three nonvanishing cohomology groups. In our case we can safely assume that H^0 and H^2 vanish. The only nonvanishing cohomology H^1 describes field configurations that obey the linearized ADHM constraint $\mu_c = 0$ but are not gauge variations. It is thus a local model for the tangent space to the instanton moduli space at each $\widetilde{T}$-fixed point. Later on we will compute weights of the toric action on the tangent space modelled on (4.20).

The partition function of the $U(1)$ topologically twisted gauge theory on $X = \mathbb{C}^2$ is easily computed. The only nontrivial topological charge is the instanton number $k = -\int_X F \wedge F$ and therefore the partition function has the form

$$Z_{U(1)}^{\mathbb{C}^2}(q) = \sum_{k=0}^{\infty} q^k \chi\left(\mathfrak{M}_{1,k}\left(\mathbb{C}^2\right)\right). \tag{4.21}$$

The expansion parameter can be identified in terms of gauge theory variables $q := \mathrm{e}^{2\pi \mathrm{i} \tau}$ with

$$\tau = \frac{4\pi \mathrm{i}}{g_{\mathrm{YM}}^2} + \frac{\vartheta}{2\pi} \tag{4.22}$$

the complexified gauge coupling, which is related to topological string variables $g_s = g_{\mathrm{YM}}^2/2$ at the attractor point. At a toric fixed point, k is identified as the number of boxes in a partition λ. The Euler classes exactly cancel in the localization formula (3.26), and one is left with the sum over fixed points [50]

$$\chi\left(\mathfrak{M}_{1,k}\left(\mathbb{C}^2\right)\right) = \sum_{\lambda : |\lambda| = k} 1. \tag{4.23}$$

By Euler's formula, one has

$$\widehat{\eta}(q)^{-1} := \prod_{n=1}^{\infty} \frac{1}{1-q^n} = \sum_{N=0}^{\infty} p(N)\, q^N, \tag{4.24}$$

where $p(N)$ is the number of partitions $\lambda = (\lambda_1, \lambda_2, \ldots)$ (ordinary, two-dimensional Young tableaux) of degree $|\lambda| = \sum_i \lambda_i = N$. The function $\widehat{\eta}(q)$ is related to the Dedekind function. It follows that the $U(1)$ partition function

$$Z^{\mathbb{C}^2}_{U(1)}(q) = \widehat{\eta}(q)^{-1} \tag{4.25}$$

is the generating function for two-dimensional Young diagrams.

This construction can be easily generalized to the nonabelian case. The fixed points are now N-coloured Young tableaux $\overrightarrow{\lambda} = (\lambda_1, \ldots, \lambda_N)$ corresponding to a partition of the instanton number $k = (k_1, \ldots, k_N)$. The instanton action is again equal to k, where the additional factor of N arising from the sum over colours $l = 1, \ldots, N$ cancels with the normalization of the $F \wedge F$ term which carries a factor $1/N$ (the inverse of the dual Coxeter number of the gauge group $G = U(N)$). The Euler characteristic of instanton moduli space is now

$$\chi\left(\mathfrak{M}_{N,k}\left(\mathbb{C}^2\right)\right) = \sum_{\overrightarrow{\lambda}:|\overrightarrow{\lambda}|=k} 1 \tag{4.26}$$

with $|\overrightarrow{\lambda}| := \sum_l |\lambda_l|$. The $U(N)$ partition function reduces to N copies of the $U(1)$ partition function

$$Z^{\mathbb{C}^2}_{U(N)}(q) = \left(Z^{\mathbb{C}^2}_{U(1)}(q)\right)^N. \tag{4.27}$$

This factorization follows from the fact that after toric localization, the gauge symmetry $U(N) \to U(1)^N$ is broken to the maximal torus. The Coulomb phase corresponds to well-separated D4-branes, but the topological nature of the gauge theory ensures that the partition function is independent of the Higgs moduli representing the lengths of open strings stretching between D-branes. In the rest of this section, we will extend these constructions to generic toric surfaces X.

4.3. Hirzebruch-Jung Spaces

Our main example will be the most general toric singularity in four dimensions, which defines a class of toric Calabi-Yau (hence open) surfaces known as Hirzebruch-Jung spaces $C = C(p, n)$. They are determined by two relatively prime positive integers p and n with $p > n$. Consider the quotient singularity $\mathbb{C}^2/\Gamma_{(p,n)}$, with the generator of the cyclic group $\Gamma_{(p,n)} \cong \mathbb{Z}_p$ acting on $(z, w) \in \mathbb{C}^2$ as

$$(z\ , w) \longmapsto \left(e^{2\pi i n/p} z, e^{2\pi i/p} w\right). \tag{4.28}$$

This orbifold has an $A_{p,n}$ singularity at the origin of $\mathbb{C}^2$. Then $C(p,n)$ is defined to be the minimal resolution of the $A_{p,n}$ singularity by a chain of ℓ exceptional divisors $S_i \cong \mathbb{P}^1$ whose intersection numbers are summarized in the intersection matrix

$$\mathbf{C} = \begin{pmatrix} -e_1 & 1 & 0 & \dots & 0 \\ 1 & -e_2 & 1 & \dots & 0 \\ 0 & 1 & -e_3 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & -e_\ell \end{pmatrix}, \tag{4.29}$$

which is called a *generalized Cartan matrix*. The divisors thus only intersect transversally with their nearest neighbours in the chain. The self-intersection numbers $e_i \geq 2$ of the spheres $\mathbb{P}^1$ of the blow-up are determined from the continued fraction expansion

$$\frac{p}{n} = e_1 - \cfrac{1}{e_2 - \cfrac{1}{e_3 - \cfrac{1}{\ddots\, e_{\ell-1} - \cfrac{1}{e_\ell}}}}. \tag{4.30}$$

Let us consider two particular well-known instances of these spaces.

4.3.1. Local $\mathbb{P}^1$

Setting $n = 1$, the space $C(p,1)$ can be identified with the total space of the holomorphic line bundle $\mathcal{O}_{\mathbb{P}^1}(-p)$ over $\mathbb{P}^1$ of degree $-p$, with $\ell = 1$ and $e_1 = p$. In this case, $S = \mathbb{P}^1$ is the zero section divisor. In the context of topological string theory, such four cycles appear in the "local" Calabi-Yau threefolds X which are regarded as neighbourhoods of a holomorphically embedded rational curve in a compact Calabi-Yau threefold, that is, as the normal bundle $\mathcal{N} \to \mathbb{P}^1$. Since $\mathcal{N}$ is a holomorphic vector bundle of rank two over $\mathbb{P}^1$ and the Calabi-Yau condition implies $c_1(\mathcal{N}) = -\chi(\mathbb{P}^1) = -2$, it follows that X is the total space of a bundle of the form $\mathcal{O}_{\mathbb{P}^1}(-p) \oplus \mathcal{O}_{\mathbb{P}^1}(p-2) \to \mathbb{P}^1$.

4.3.2. A_{p-1} ALE Space

The complex surface $C = C(p,p-1)$ is an example of an asymptotically locally Euclidean (ALE) space. This means that C carries a scalar flat Kähler metric g such that (C,g) is complete, and there exists a compact set K such that $C \setminus K \cong (\mathbb{R}^4 \setminus \overline{B_R}\,)/\mathbb{Z}_p$. Here $\mathbb{Z}_p \subset O(4)$ acts freely on $\mathbb{R}^4 \setminus \overline{B_R}$ and the metric g approximates the flat Euclidean metric on $\mathbb{R}^4$. Such a coordinate system is called a coordinate system at infinity. We regard $\mathbb{Z}_p \subset U(2)$ acting on $(z,w) \in \mathbb{C}^2 \cong \mathbb{R}^4$ as described above (with $n = p-1$), and the complex structure I on C approximates that on $\mathbb{C}^2 = \mathbb{R}^4$. As the resolution of the Klein singularity $\mathbb{C}^2/\mathbb{Z}_p$, $C(p,p-1)$ contains a chain of $\ell = p-1$ projective lines $\mathbb{P}^1$, each with self-intersection number $e_i = 2$. In this case, the intersection matrix $\mathbf{C}$ coincides with the Cartan matrix of the A_{p-1} Dynkin diagram.

4.4. Instantons on ALE Spaces

We begin by describing in some detail the instanton moduli space in the case of the A_{p-1} ALE spaces, for which a rigorous construction is known. $U(N)$ instantons on ALE spaces are given by the ADHM construction. Since the topological gauge theory is invariant under blow-ups of the surface (using blow-up formulas), one can do the instanton computation on the orbifold $\mathbb{C}^2/\Gamma$ where $\Gamma = \Gamma_{(p,p-1)} \cong \mathbb{Z}_p$. This result is at the heart of the McKay correspondence which provides a one-to-one correspondence between irreducible representations of the orbifold group Γ and tautological bundles over the exceptional divisors of the minimal resolution $C = C(p, p-1)$.

Since C is noncompact, the instanton moduli space $\mathfrak{M}_C$ must be defined with respect to connections which have appropriate asymptotic decay at infinity. We will describe this in more generality later on in terms of framed moduli spaces of torsion-free sheaves. These connections correspond to instantons of finite energy and are asymptotic to flat connections with $F = 0$. In particular, there are solutions which have fractional first Chern class and are related to instantons that asymptote to flat connections with nontrivial holonomy at the boundary of C, which is topologically the Lens space $L(p, p-1) = S^3/\Gamma$. The flat connections are classified by homomorphisms $\rho : \pi_1(C) \to U(N)$, where $\pi_1(C) = \Gamma \cong \mathbb{Z}_p$. The asymptotic connection at infinity is thus labelled by irreducible representations $(k_0, k_1, \ldots, k_{p-1})$ of the orbifold group $\mathbb{Z}_p$, with $\sum_i k_i = N$, and are given explicitly by

$$\rho_k\left(\mathrm{e}^{2\pi \mathrm{i}/p}\right) = \mathrm{e}^{2\pi \mathrm{i} k/p}, \tag{4.31}$$

where $k = 0, 1, \ldots, p-1$.

Starting from the ADHM construction on $\mathbb{C}^2$ outlined in Section 4.2., one constructs its Γ-invariant decomposition. Consider the universal scheme $\mathfrak{Z} \subset C \times \mathbb{C}^2$ given by the correspondence diagram

$$\begin{tikzcd} & \mathfrak{Z} \arrow[dl, "q_1"'] \arrow[dr, "q_2"] & \\ C & & \mathbb{C}^2 \end{tikzcd} \tag{4.32}$$

The tautological bundle on C is defined by

$$\mathcal{R} := q_{1*}\mathcal{O}_{\mathfrak{Z}}. \tag{4.33}$$

Under the action of Γ on $\mathfrak{Z}, \mathcal{R}$ transforms in the regular representation and can thus be decomposed into irreducible representations

$$\mathcal{R} = \bigoplus_{k=0}^{p-1} \mathcal{R}_k \otimes \rho_k, \qquad \mathcal{R}_k = \mathrm{Hom}_\Gamma(\mathcal{R}, \rho_k). \tag{4.34}$$

By the McKay correspondence, the bundles $\mathcal{R}_0 = \mathcal{O}_C, \mathcal{R}_1, \ldots, \mathcal{R}_{p-1}$ form the canonical integral basis of the K-theory group $K^0(C)$ constructed in [51].

In this case, we take $\underline{Q} \cong \mathbb{C}^2$ to be a module on which the regular representation of Γ acts. We also take $\Gamma \subset SU(2)$ so that the determinant representation is trivial as a Γ-module, that is, $\wedge^2 \underline{Q} \otimes_\Gamma \mathcal{R} \cong \mathcal{R}$. The two vector spaces V and W which feature in the ADHM construction have a natural grading under the action of the orbifold group Γ given by

$$V = \bigoplus_{k=0}^{p-1} V_k \otimes \rho_k, \qquad W = \bigoplus_{k=0}^{p-1} W_k \otimes \rho_k. \tag{4.35}$$

The modification of (4.20) is given by [52]

$$\mathrm{Hom}_\Gamma(\mathcal{R}^*, V) \xrightarrow{\sigma} \begin{array}{c} \mathrm{Hom}_\Gamma\left(\mathcal{R}^*, \underline{Q} \otimes_\Gamma V\right) \\ \oplus \\ \mathrm{Hom}_\Gamma(\mathcal{R}^*, W) \end{array} \xrightarrow{\tau} \mathrm{Hom}_\Gamma(\mathcal{R}^*, V), \tag{4.36}$$

and the condition that the sequence (4.36) is a complex is equivalent to the (generalized) ADHM equations. After imposing a certain stability condition, this construction realizes the instanton moduli space as a quiver variety $\mathfrak{M}(V, W)$.

This construction identifies two distinct types of instanton contributions to the ALE partition function, which we consider in turn. As before, after toric localization the gauge symmetry breaks as $U(N) \to U(1)^N$ and the $U(N)$ partition function factorizes as

$$Z_{U(N)}^{\mathrm{ALE}}(q, \mathbf{Q}) = \left(Z_{U(1)}^{\mathrm{ALE}}(q, \mathbf{Q})\right)^N. \tag{4.37}$$

It therefore suffices to focus on the $U(1)$ case in the following.

4.4.1. Regular Instantons

Regular instantons on A_{p-1} live in the regular representation $k_0 = k_1 = \cdots = k_{p-1} = k$ of the orbifold group $\Gamma = \mathbb{Z}_p$. They correspond to D0-branes moving freely on C with p orbifold images away from the orbifold point. For gauge group $U(1)$, the moduli space is given by specifying $K = kp$ points on C up to permutations. Hence the moduli space $\mathfrak{M}_{reg}^{U(1)}(C)$ of regular $U(1)$ instantons on C is isomorphic to the Hilbert scheme $C^{[K]}$. The generating function for the Euler numbers of the instanton moduli spaces can then be computed explicitly by applying Göttsche's formula to get

$$Z_{\mathrm{reg}}^{U(1)}(q) = \sum_{K=0}^{\infty} q^K \chi\left(C^{[K]}\right) = \widehat{\eta}(q)^{-p}. \tag{4.38}$$

The $U(N)$ partition function is the Nth power of this quantity. Heuristically, we may think of this formula as originating by covering C with $p = \chi(C)$ open charts to get p copies of $U(N)$ instantons on $\mathbb{C}^2$, each contributing $\widehat{\eta}(q)^{-N}$. This can be demonstrated rigorously on any toric surface C by a localization computation [53, 54].

4.4.2. Fractional Instantons

To each irreducible representation $(k_0, k_1, \ldots, k_{p-1})$ of the orbifold group $\mathbb{Z}_p$, there corresponds a *fractional instanton* which is stuck at the orbifold points. It has *no* moduli (or orbifold images) and can be regarded as a state in which open strings ending on the same D0-brane are projected out by the action of the orbifold group. They carry magnetic fluxes through the $\mathbb{P}^1$'s of the minimal resolution, and correspond to self-dual $U(1)$ gauge connections with curvatures

$$F = -2\pi \mathrm{i} u_i c_1(\mathcal{R}_i), \tag{4.39}$$

where $u_i \in \mathbb{Z}$ and $\mathcal{R}_i = \mathcal{O}_{\mathbb{P}^1}(e_i)$ are the tautological line bundles. The Chern classes $c_1(\mathcal{R}_i)$, $i = 1, \ldots, p-1$, form a basis of $H^2(C, \mathbb{Z})$. Fractional instantons can thus be thought of as Dirac monopoles on the two-spheres of the orbifold resolution.

The corresponding intersection numbers are given by

$$\int_C c_1(\mathcal{R}_i) \wedge c_1(\mathcal{R}_j) = -\left(\mathbf{C}^{-1}\right)_{ij}, \quad \int_{S_i} c_1(\mathcal{R}_j) = \delta_{ij}. \tag{4.40}$$

Since C is noncompact, the intersection matrix $\mathbf{C}$ is not necessarily unimodular, and the corresponding instanton charges can be fractional. The contribution of fractional instantons to the supersymmetric Yang-Mills action with observables is thus given by

$$S_{\mathrm{frac}} = -\frac{\mathrm{i}\tau}{4\pi} \int_C F \wedge F - \frac{\mathrm{i}\phi_j^2}{2\pi} \int_C F \wedge c_1(\mathcal{R}_j) = -\pi \mathrm{i} \tau \left(\mathbf{C}^{-1}\right)^{ij} u_i u_j + z_i u_i \tag{4.41}$$

with $z_i = (\mathbf{C}^{-1})_{ij}\, \phi_j^2$. Setting $\mathbf{u} := (u_1, \ldots, u_{p-1})$ and identifying $Q_i = \mathrm{e}^{-z_i}$ using the attractor mechanism, we find that the contribution of $U(1)$ fractional instantons to the full partition function is given by a theta-function

$$Z_{\mathrm{frac}}^{U(1)}(q, \mathbf{Q}) = \sum_{\mathbf{u} \in \mathbb{Z}^{p-1}} q^{(1/2)\, \mathbf{u} \cdot \mathbf{C}^{-1} \mathbf{u}} \mathbf{Q}^{\mathbf{u}} \tag{4.42}$$

on a Riemann surface of genus $g = p - 1$ and period matrix $\tau \mathbf{C}^{-1}$.

4.5. Instantons on Local $\mathbb{P}^1$

Let us now discuss what is known beyond the ALE case, in the instance that C is the total space of the holomorphic line bundle $\mathcal{O}_{\mathbb{P}^1}(-p)$ [55]. In this case an ADHM construction is not available. Nevertheless, much of the construction in the ALE case carries through, and by introducing weighted Sobolev norms, one shows that $\mathfrak{M}_C$ is a smooth Kähler manifold with torsion-free homology groups which vanish in odd degrees. Let us consider some explicit examples.

The zero section of $\mathcal{O}_{\mathbb{P}^1}(-p)$, considered as a divisor $S = \mathbb{P}^1$ of C, produces a line bundle $\mathcal{L} \to C$ such that $c_1(\mathcal{L})$ is the generator of $H^2(C, \mathbb{Z}) = \mathbb{Z}$. It has a unique self-dual connection asymptotic to the trivial connection. Let $\underline{\mathbb{C}} = C \times \mathbb{C}$ be the trivial line bundle over C. Set $\mathcal{E} = \underline{\mathbb{C}} \oplus \mathcal{L}$, and let $\mathfrak{M}(\mathcal{E})$ be the moduli space of self-dual connections on $\mathcal{E}$ which are asymptotic to the trivial connection at infinity. Then $\dim_{\mathbb{R}}(\mathfrak{M}(\mathcal{E})) = 2p$. Since $H_1(C, \mathbb{R}) = 0$, using Morse theory one can show that the only nonvanishing homology groups of the instanton moduli space are $H_0(\mathfrak{M}(\mathcal{E}), \mathbb{R}) = H_2(\mathfrak{M}(\mathcal{E}), \mathbb{R}) = \mathbb{R}$.

Alternatively, set $\mathcal{E} = \mathcal{L} \oplus \mathcal{L}^\vee$, and let $\mathfrak{M}_{(k)}(\mathcal{E})$, $k = 0, 1, \dots, p-1$, be the moduli space of self-dual connections asymptotic to $\rho_k \oplus \rho_k^*$. Then for any $p > 2$, $\dim_{\mathbb{R}}(\mathfrak{M}_{(k)}(\mathcal{E})) = 2$, while for $p = 2$ (whereby $\mathcal{O}_{\mathbb{P}^1}(-2)$ coincides with the A_1 ALE space), one has $\dim_{\mathbb{R}}(\mathfrak{M}_{(k)}(\mathcal{E})) = 4$. In particular, for $p = 2$ there is a diffeomorphism $\mathfrak{M}_{(k)}(\mathcal{E}) \cong T^*\mathbb{P}^1$ with $H_0(\mathfrak{M}_{(k)}(\mathcal{E}), \mathbb{R}) = H_2(\mathfrak{M}_{(k)}(\mathcal{E}), \mathbb{R}) = \mathbb{R}$, while for $p = 4$ one has $\mathfrak{M}_{(k)}(\mathcal{E}) \cong B^2$. Instantons on local $\mathbb{P}^1$ will be studied in more generality later on in terms of moduli spaces of framed torsion-free sheaves.

4.6. Wall-Crossing Formulas

In [56] it was suggested that the structure of the instanton partition function on ALE spaces can be extrapolated to give a general result valid for all Hirzebruch-Jung surfaces $C = C(p, n)$. Thus one postulates the form of the $U(N)$ partition function

$$\begin{aligned} Z^C_{U(N)}(q, \mathbf{Q}) &= \left(Z^{U(1)}_{\text{reg}}(q) Z^{U(1)}_{\text{frac}}(q, \mathbf{Q}) \right)^N \\ &= \frac{1}{\hat{\eta}(\tau)^{N\chi(C)}} \sum_{\vec{\mathbf{u}} \in \mathbb{Z}^{N\, b_2(C)}} q^{(1/2)\, \vec{\mathbf{u}} \cdot \mathbf{C}^{-1} \vec{\mathbf{u}}} \mathbf{Q}^{\mathbf{u}}, \end{aligned} \tag{4.43}$$

where

$$\mathbf{u} := \sum_{l=1}^{N} \mathbf{u}_l. \tag{4.44}$$

In the ALE case, one has $\chi(A_{p-1}) = p$ and $b_2(A_{p-1}) = p - 1$. The evidence for this formula is supported by calculations in the reduction to q-deformed two-dimensional Yang-Mills theory [13, 57], which captures the contributions from fractional instantons on the exceptional divisors. It was even conjectured to hold in the case when C is a *compact* toric surface, at least in the region of moduli space where the instanton charges are large. Though the regular and fractional instantons are always readily constructed exactly as in the ALE case, the issue is whether or not this formula takes into account *all* of the instanton contributions, and in which regions of moduli space the factorization into $U(1)$ partition functions holds. Later on we will give more precise meanings to these "asymptotic charge regions" in terms of moduli spaces of torsion-free sheaves.

Curiously, the formula (4.43) coincides with the chiral torus partition function of a conformal field theory in two dimensions with central charge $c = N\ \chi(C)$, that is, of $N\ \chi(C)$ free bosons, with $N\ b_2(C)$ of them being compact. The compact degrees of freedom live in a torus determined by the lattice $H^2(C, \mathbb{Z})$ with the bilinear form $\mathbf{C}$. The appearence of this two-dimensional field theory can be understood in M-theory, wherein the D4-D2-D0 brane quantum mechanics lifts to a $(4,0)$ two-dimensional superconformal field theory on an M5-brane worldvolume [58]. This is reminscent of the recent conjectural relations between four-dimensional superconformal gauge theories and two-dimensional Liouville conformal field theories [59, 60]. When C is an A_{p-1} ALE space and the gauge group is $SU(N)$, there is yet another conformal field theoretic interpretation [46, 55, 61, 62]. In this case, the level N affine $\widehat{\mathfrak{su}(p)}$ Lie algebra acts on the cohomology ring of the instanton moduli space, and the partition function (with appropriate local curvature and signature corrections inserted) coincides with the character of the affine Kac-Moody algebra given by

$$Z^{\mathrm{ALE}}_{SU(N)}(q,\mathbf{Q}) = \sum_{n=0}^{\infty}\sum_{\mathbf{u}\in\mathbb{Z}^{p-1}} \Omega(n,\mathbf{u})q^{n-c/24}\mathbf{Q}^{\mathbf{u}} = \mathrm{Tr}_{\mathscr{H}}\left(q^{L_0-c/24}\mathbf{Q}^{\mathbf{J}_0}\right), \tag{4.45}$$

where $\mathbf{u} = c_1(\mathcal{P})$, $n = c_2(\mathcal{P}) = k + (1/2)\ \mathbf{u}\cdot\mathbf{C}\mathbf{u}$, and $\Omega(n,\mathbf{u})$ are the degeneracies of BPS states with the specified quantum numbers. Here $\mathscr{H}$ is the Hilbert space on which the chiral algebra acts, which can be represented in terms of free fermion or boson conformal field theories with extended symmetry generators $\mathbf{J}_0$ [61]. The $U(1)$ partition function is also expressed as a $\widehat{\mathfrak{u}(1)}_1$ character in [61, 62].

Let us now compare the instanton partition function (4.43) with the black hole partition function. Microscopic black hole entropy formulas for BPS bound states of D0-D2-D4 branes in Type IIA supergravity are readily available on *compact* Calabi-Yau threefolds X [58], in the large volume limit and when contributions from worldsheet instantons are negligible. In this limit we can expand the cycle $[C]$ as in (2.20). Let α_i, $i = 1,\dots,b_2(X)$, be an integral basis of two cocycles for $H^2(X,\mathbb{Z})$ dual to the four-cycle $[C_i]$ with the intersection numbers

$$D_{ijk} = \frac{1}{6}\int_X \alpha_i\wedge\alpha_j\wedge\alpha_k, \qquad c_{2,i} = \int_X \alpha_i\wedge c_2(X), \tag{4.46}$$

and let D^{ij} be the matrix inverse of $D_{ij} = D_{ijk}Q_4^k$. The genus zero topological string amplitude F_0 can be expressed as

$$F_0 = D_{ijk}\frac{X^iX^jX^k}{X^0}. \tag{4.47}$$

With $Q_0 = (1/8\pi^2)\int_C F\wedge F$ and $Q_i = Q_2^i = (\mathbf{C}^{-1})^{ij}u_j$, the black hole entropy is given by [58]

$$S_{\mathrm{BH}}(Q_0,\boldsymbol{Q},\boldsymbol{Q}_4) = 2\pi\sqrt{\left(D_{ijk}Q_4^i\ Q_4^j\ Q_4^k + \frac{1}{6}\ c_{2,i}Q_4^i\right)\left(Q_0 + \frac{1}{12}D^{ij}Q_iQ_j\right)}. \tag{4.48}$$

We would now like to interpret the gauge theory partition function (4.43) as the corresponding black hole partition function $Z_{\mathrm{BH}}(\boldsymbol{Q}_4,\boldsymbol{\phi}^2,\phi^0)$. For this, we expand it as

$$Z^{C}_{U(N)}(q,\mathbf{Q}) = \sum_{Q_0,Q_i}\Omega(Q_0,\boldsymbol{Q},\boldsymbol{Q}_4)\,\mathrm{e}^{-Q_0\phi^0-\boldsymbol{Q}\cdot\boldsymbol{\phi}^2}. \tag{4.49}$$

Then Cardy's formula gives the black hole entropy as

$$S_{\mathrm{BH}}(Q_0, \mathcal{Q}, \mathcal{Q}_4) = \log \Omega(Q_0, \mathcal{Q}, \mathcal{Q}_4). \tag{4.50}$$

This expression agrees [56] with the macroscopic supergravity result (4.48) for the Bekenstein-Hawking-Wald entropy in the large Q_0 limit. Wall-crossing issues in a similar context are discussed in more detail in [63–67].

In parallel to what we did in Section 3.9., in order to explore the pertinent wall-crossing formulas in this instance, we will now consider moduli spaces which parametrize isomorphism classes of the following objects:

(a) surjections (framings) $\mathcal{O}_C \to \mathcal{F} \to 0$ of torsion-free sheaves with $\mathrm{ch}(\mathcal{F}) = (N, d, -k)$,

(b) stable torsion-free sheaves $\mathcal{E}$ on C with $\mathrm{ch}(\mathcal{E}) = (N, d, k)$,

(c) closed subschemes $S \subset C$ of dimension ≤ 1 with dual curve class $[S]^\vee = d$ and holomorphic Euler characteristic $\chi(\mathcal{O}_S) = k$.

In contrast to the six-dimensional situation, in four dimensions the connections between these three classes of objects are somewhat more subtle. We will examine each of them in turn and how they compare with the gauge theory results we have thus far obtained.

4.7. Moduli Spaces of Framed Instantons

We begin with Point (a) at the end of Section 4.6. Let C be a smooth, quasiprojective, open, toric surface. We will assume that C admits a projective compactification $\overline{C}$, that is, $\overline{C}$ is a smooth, compact, projective, toric surface with a smooth divisor $\ell_\infty \subset \overline{C}$ (called the "line at infinity") which is a T^2-invariant $\mathbb{P}^1$ in $\overline{C}$, and such that $C = \overline{C} \setminus \ell_\infty$. We will also require that $\ell_\infty \cdot \ell_\infty > 0$, in addition to $\ell_\infty \cong \mathbb{P}^1$. The difference between the counting of framed instantons on the compact toric surface $\overline{C}$ (with boundary condition at "infinity") and of unframed instantons on the open toric surface C is a universal perturbative contribution, which will be dropped here. Since $\overline{C}$ is compact, these calculations will only capture the contributions from instantons with integer charges on C. We will mention later on how to incorporate the contributions from fractional instantons on C.

Let us fix two numbers $N \in \mathbb{N}$ and $k \in \mathbb{Q}$, and an integer cohomology class $d \in H^2(\overline{C}, \mathbb{Z})$. Let $\mathfrak{M}_{N,d,k}(C)$ be the framed moduli space consisting of isomorphism classes $[\mathcal{E}]$ of torsion-free sheaves $\mathcal{E}$ on $\overline{C}$ such that

(1) $\mathcal{E}$ has the following topological Chern invariants:

$$N = \mathrm{ch}_0(\mathcal{E}) = \mathrm{rank}(\mathcal{E}), \qquad d = \mathrm{ch}_1(\mathcal{E}) = c_1(\mathcal{E}), \qquad k = -\int_{\overline{C}} \mathrm{ch}_2(\mathcal{E}); \tag{4.51}$$

(2) $\mathcal{E}$ is locally-free in a neighbourhood of ℓ_∞, and there is an isomorphism $\mathcal{E}|_{\ell_\infty} \cong \mathcal{O}_{\ell_\infty}^{\oplus N}$ called the "framing at infinity".

These topological conditions imply that the moduli space $\mathfrak{M}_{N,d,k}(C)$ is nonempty only when

$$d|_{\ell_\infty} = N c_1(\mathcal{O}_{\ell_\infty}) = 0, \tag{4.52}$$

so that d defines a class $d \in H^2_{\mathrm{cpt}}(C,\mathbb{Z})$ in the compactly supported cohomology of C. Furthermore, from $\mathrm{ch}_2(\mathcal{E}) = -c_2(\mathcal{E}) + (1/2)c_1(\mathcal{E}) \wedge c_1(\mathcal{E})$, the last relation in (4.51) can be written as

$$k = \int_{\overline{C}} \left(c_2(\mathcal{E}) - \frac{1}{2}\, d \wedge d \right). \tag{4.53}$$

Framed sheaves yield stable pairs [68], analogous to those described in Section 3.9., after suitable choices of polarization on $\overline{C}$ and stability parameter.

We write $\mathcal{E}(-\ell_\infty) := \mathcal{E} \otimes \mathcal{O}_{\overline{C}}(-\ell_\infty)$. At a given point $[\mathcal{E}] \in \mathfrak{M}_{N,d,k}(C)$, the space of reducible connections is $\mathrm{Ext}^0_{\mathcal{O}_{\overline{C}}}(\mathcal{E},\mathcal{E}(-\ell_\infty)) = \mathrm{Hom}_{\mathcal{O}_{\overline{C}}}(\mathcal{E},\mathcal{E}(-\ell_\infty))$, the Zariski tangent space is $\mathrm{Ext}^1_{\mathcal{O}_{\overline{C}}}(\mathcal{E},\mathcal{E}(-\ell_\infty))$, and the obstruction space is $\mathrm{Ext}^2_{\mathcal{O}_{\overline{C}}}(\mathcal{E},\mathcal{E}(-\ell_\infty))$. The cohomology of the instanton deformation complex is greatly simplified by the fact that in this case [69]

$$\mathrm{Ext}^0_{\mathcal{O}_{\overline{C}}}(\mathcal{E},\mathcal{E}(-\ell_\infty)) = \mathrm{Ext}^2_{\mathcal{O}_{\overline{C}}}(\mathcal{E},\mathcal{E}(-\ell_\infty)) = 0. \tag{4.54}$$

Using (4.54) and the Riemann-Roch theorem, one shows [69] that the moduli space $\mathfrak{M}_{N,d,k}(C)$ is a smooth quasiprojective variety of (complex) dimension $2N\ k + d^2$, where $d^2 := \int_C d \wedge d$, whose tangent space at a point $[\mathcal{E}]$ is isomorphic to the vector space $\mathrm{Ext}^1_{\mathcal{O}_{\overline{C}}}(\mathcal{E},\mathcal{E}(-\ell_\infty))$.

We will now describe a natural torus invariant subspace of the instanton moduli space. The T^2-invariance of ℓ_∞ implies that the pullback of the T^2-action on $\overline{C}$ defines an action on $\mathfrak{M}_{N,d,k}(C)$. There is also an action of the diagonal maximal torus T^N of $GL(N,\mathbb{C})$ on the framing. Altogether we get an action of the complex algebraic torus $\widetilde{T} = T^2 \times T^N$ on $\mathfrak{M}_{N,d,k}(C)$ [69]. We are interested in the fixed point set $\mathfrak{M}_{N,d,k}(C)^{\widetilde{T}}$ of this torus action.

An isomorphism class $[\mathcal{E}] \in \mathfrak{M}_{N,d,k}(C)$ is fixed by the T^N-action if and only if it decomposes as

$$\mathcal{E} = \mathcal{E}_1 \oplus \cdots \oplus \mathcal{E}_N, \qquad \mathcal{E}_l \in \mathfrak{M}_{1,d,k}(C) \tag{4.55}$$

such that $\mathcal{E}_l|_{\ell_\infty}$ is mapped to the lth factor $\mathcal{O}_{\ell_\infty}$ of $\mathcal{O}_{\ell_\infty}^{\oplus N}$ under the framing isomorphism. Since the double dual $\mathcal{E}_l^{\vee\vee}$ is a line bundle which is trivial on ℓ_∞, it is equal to $\mathcal{O}_{\overline{C}}(D_l)$ for some divisor $D_l \subset \overline{C}$ disjoint from ℓ_∞. Via the natural injection $\mathcal{E}_l \subset \mathcal{E}_l^{\vee\vee} = \mathcal{O}_{\overline{C}}(D_l)$, the sheaf $\mathcal{E}_l$ is thus equal to $\mathcal{I}_l(D_l) = \mathcal{I}_l \otimes \mathcal{O}_{\overline{C}}(D_l)$ for some ideal sheaf $\mathcal{I}_l$ of a zero-dimensional subscheme $Z_l \subset \overline{C}$, also disjoint from ℓ_∞. If $\mathcal{E}$ is also fixed by the T^2-action, then so are D_l, $\mathcal{I}_l$, and Z_l.

The supports of the T^2-invariant subschemes Z_l and D_l can be represented explicitly in terms of the toric geometry of C. Recall that this is described in terms of an underlying toric graph $\Delta(C)$ whose vertices are in bijective correspondence with the T^2-fixed points in C, and two vertices are joined by an edge e if and only if the corresponding fixed points are connected by a T^2-invariant $\mathbb{P}^1$. Let $V(C)$ and $E(C)$, respectively, denote the set of vertices and edges of $\Delta(C)$. Then since $\Delta(C)$ is a chain, one has $|E(C)| = |V(C)| - 1$. The number of vertices $n = |V(C)|$ is also the number of two-cones in the toric fan of C and it is related to the Euler characteristic $\chi(C)$ of the surface as

$$n = \chi(C) = \chi\left(\overline{C} \right) - 2. \tag{4.56}$$

We denote by p_f the T^2-invariant point in C corresponding to the vertex $f \in V(C)$ and by ℓ_e the T^2-invariant line $\mathbb{P}^1$ corresponding to the edge $e \in E(C)$.

Since the zero-cycles Z_l are not supported on ℓ_∞, they must be contained in the fixed point set $V(C)$ in C. Thus each Z_l is a union of subschemes $Z_l^f, f \in V(C)$ supported at the T^2-fixed points $p_f \in C$. If we choose a local coordinate system $(x, y) \in \mathbb{C}^2$ in a patch U_f around p_f, then the T^2-invariant ideal of Z_l^f in the coordinate ring $\mathbb{C}[x, y]$ of $U_f \cong \mathbb{C}^2$ is generated by the T^2-eigenfunctions with nontrivial characters, which are monomials $x^i y^j$, and hence Z_l^f corresponds to a Young diagram λ_l^f with $|\lambda_l^f|$ boxes (with monomial $x^i y^j$ placed at $(i+1, j+1)$). The ideal is spanned by monomials outside the Young diagram. Likewise, since the T^2-invariant two-cycles D_l are disjoint from the line at infinity, they are supported along the edges $\ell_e \cong \mathbb{P}^1, e \in E(C)$.

It follows that the fixed point set $\mathfrak{M}_{N,d,k}(C)^{\widetilde{T}}$ is parameterized by finitely many N-tuples

$$\left(\overrightarrow{D}, \overrightarrow{\boldsymbol{\lambda}}\right) = ((D_1, \boldsymbol{\lambda}_1), \ \ldots, (D_N, \boldsymbol{\lambda}_N)), \tag{4.57}$$

where

$$D_l \in H_2(C, \mathbb{Z}) \cong \bigoplus_{e \in E(C)} \mathbb{Z}[\ell_e] \tag{4.58}$$

are T^2-invariant divisors in C and

$$\boldsymbol{\lambda}_l = \left(\lambda_l^f\right)_{f \in V(C)} \tag{4.59}$$

is a vector of Young tableaux with $|\boldsymbol{\lambda}_l| := \sum_{f \in V(C)} |\lambda_l^f|$ boxes. The Young tableaux parametrize the contributions from regular pointlike D0-brane instantons (freely moving inside C). The divisors parametrize the contributions from D2-brane instantons (wrapping D_l with appropriate units of magnetic flux). We will see that these two types of contributions factorize completely. We can write the topological invariants of a generic element (4.57) in $\mathfrak{M}_{N,d,k}(C)^{\widetilde{T}}$ in terms of this combinatorial data. Since $c_1(\mathcal{I}_l) = 0$, the constraint $d = c_1(\mathcal{E})$ can be written as

$$d = \sum_{l=1}^{N} c_1(\mathcal{O}_C(D_l)), \tag{4.60}$$

whereas (4.53) becomes

$$k = \sum_{l=1}^{N} |\boldsymbol{\lambda}_l| + \int_C \left(\sum_{l<l'} c_1(\mathcal{O}_C(D_l)) \wedge c_1(\mathcal{O}_C(D_{l'})) - \frac{1}{2}\ d \wedge d \ \right). \tag{4.61}$$

The Young diagram box sum in (4.61) gives the length of the singularity set of the sheaf (4.55).

To make contact with our previous formulas, let us rewrite these constraints in a more explicit parametrization. Define the invertible, symmetric, integer-valued intersection matrix $\mathbf{C} = (\mathbf{C}_{e,e'})_{e,e' \in E(C)}$ between lines of the toric graph $\Delta(C)$ as

$$\mathbf{C}_{e,e'} := \ell_e \cdot \ell_{e'}. \tag{4.62}$$

Each divisor ℓ_e is canonically associated to a T^2-equivariant holomorphic line bundle $\mathcal{L}_e := \mathcal{O}_C(\ell_e)$ whose first Chern class $c_1(\mathcal{L}_e) \in H^2_{\mathrm{cpt}}(C, \mathbb{Z})$ is Poincaré dual to ℓ_{e} with

$$\int_{\ell_e} c_1(\mathcal{L}_{e'}) = \mathbf{C}_{e,e'}, \qquad \int_C c_1(\mathcal{L}_e) \wedge c_1(\mathcal{L}_{e'}) = \mathbf{C}_{e,e'}. \tag{4.63}$$

Note that these intersection numbers differ from (4.40), that is, the basis of line bundles $\mathcal{L}_e$ does *not* coincide with the tautological line bundles $\mathcal{R}_i$. Using (4.58) we can write $c_1(\mathcal{O}_C(D_l)) = \sum_{e \in E(C)} u_l^e c_1(L_e)$ with $u_l^e \in \mathbb{Z}$. Then the constraint (4.60) can be expressed as

$$d = \sum_{l=1}^{N} \sum_{e \in E(C)} u_l^e c_1(\mathcal{L}_e). \tag{4.64}$$

Combined with (4.61) and using (4.63), this gives

$$k = \sum_{l=1}^{N} |\lambda_l| + \sum_{l<l'} u_l^e \mathbf{C}_{e,e'} u_{l'}^{e'} - \frac{1}{2} \sum_{l,l'=1}^{N} u_l^e \mathbf{C}_{e,e'} u_{l'}^{e'} = \sum_{l=1}^{N} |\lambda_l| - \frac{1}{2} \sum_{l=1}^{N} u_l^e \mathbf{C}_{e,e'} u_l^{e'}. \tag{4.65}$$

The Vafa-Witten-Nekrasov partition function on the toric surface C is now defined as the $\widetilde{T}$-equivariant index

$$Z_{U(N)}^{C}(q, \mathbf{Q})^{\mathrm{fr}} = \sum_{k \in \mathbb{Q}} q^k \sum_{d \in H^2_{\mathrm{cpt}}(C, \mathbb{Z})} \mathbf{Q}^d \chi(\mathfrak{M}_{N,d,k}(C)), \tag{4.66}$$

where as before $q = \mathrm{e}^{2\pi \mathrm{i} \tau}$ with τ the complexified gauge coupling constant and

$$\mathbf{Q}^d := \prod_{e \in E(C)} (Q_e)^{(\mathbf{C}^{-1})^{e,e'} \int_{\ell_{e'}} d} \tag{4.67}$$

for a collection of formal variables $Q_e, e \in E(C)$ and a given cohomology class $d \in H^2_{\mathrm{cpt}}(C, \mathbb{Z})$. From the localization formula (3.25), we see that the $\widetilde{T}$-equivariant Euler class cancels in (3.26) at the fixed points. This cancellation in the fluctuation determinants is a consequence of (4.54) which implies that the obstruction bundle is trivial, and there is an isomorphism between the tangent and normal bundles over the fixed point set $\mathfrak{M}_{N,d,k}(C)^{\widetilde{T}}$ in the instanton moduli space. Because of this cancellation, we see that the contribution from each critical point of the gauge theory is independent of the equivariant parameters ϵ_i and a_l. In the sum over critical points, we can replace the sum over $k \in \mathbb{Q}$ using (4.65) by sums over Young diagrams $\vec{\lambda}$ and over the integers $\vec{\mathbf{u}} = (\mathbf{u}_1, \dots, \mathbf{u}_N)$ and $\mathbf{u}_l = (u_l^e)_{e \in E(C)} \in \mathbb{Z}^{n-1}$ representing magnetic fluxes through ℓ_e on D2-branes wrapping the divisors D_l. The sum over $d \in H^2_{\mathrm{cpt}}(C, \mathbb{Z})$ may then be saturated using (4.64).

Putting everything together, and using (4.64) and (4.63) to write

$$\int_{\ell_e} d = \sum_{l=1}^{N} \mathbf{C}_{e,e'} u_l^{e'}, \tag{4.68}$$

the partition function (4.66) thereby becomes

$$Z^C_{U(N)}(q, \mathbf{Q})^{\mathrm{fr}} = \sum_{\vec{\lambda}} \sum_{\vec{\mathbf{u}} \in \mathbb{Z}^{N\,(n-1)}} q^{|\vec{\lambda}| - (1/2)\, \vec{\mathbf{u}} \cdot \mathbf{C} \vec{\mathbf{u}}} \mathbf{Q}^{\mathbf{u}}, \tag{4.69}$$

where we have used (4.44). The sums over Young tableaux decouple, and for each vertex $f \in V(C)$ and each $l = 1, \dots, N$, they produce a factor of the Euler function $\widehat{\eta}(q)^{-1}$. Then we can bring (4.69) into the final form

$$Z^C_{U(N)}(q, \mathbf{Q})^{\mathrm{fr}} = \frac{1}{\widehat{\eta}(q)^{N\chi(C)}} \sum_{\vec{\mathbf{u}} \in \mathbb{Z}^{N\, b_2(C)}} q^{-(1/2)\, \vec{\mathbf{u}} \cdot \mathbf{C} \vec{\mathbf{u}}} \mathbf{Q}^{\mathbf{u}}, \tag{4.70}$$

where we recall that $\chi(C) = |V(C)| = n$ and $b_2(C) = |E(C)| = n - 1$.

This formula differs from (4.43) in that only integral values of the first Chern class are permitted in (4.70). In the case of an A_{p-1} singularity, fractional first Chern classes can be incorporated by constructing instead the moduli space of torsion-free sheaves on the orbifold compactification $\overline{C}^{\mathrm{orb}} = C \cup \widetilde{\ell}_\infty$ of the hyper-Kähler ALE space C, where $\widetilde{\ell}_\infty = \mathbb{P}^1/\Gamma$ [52, 55, 70, 71]. In a neighbourhood of infinity, we can approximate $\overline{C}$ by $\mathbb{P}^2/\Gamma$ with the singularity at the origin resolved. More precisely, we obtain the divisor $\widetilde{\ell}_\infty$ by gluing together the trivial bundle $\mathcal{O}_C$ on C with the line bundle $\mathcal{O}_{\mathbb{P}^2/\Gamma}(1)$ on $\mathbb{P}^2/\Gamma$. The latter bundle has a Γ-equivariant structure such that the map $\mathcal{O}_C \to \mathcal{O}_{\mathbb{P}^2/\Gamma}(1)$ is Γ-equivariant. Let us examine some examples which are covered by the analysis above.

4.7.1. *Affine Plane*

Let $C = \mathbb{C}^2$. Then $\overline{C} = \mathbb{P}^2$ with $\ell_\infty = [0, z_1, z_2] \cong \mathbb{P}^1$ and intersection number $\ell_\infty \cdot \ell_\infty = 1 > 0$. Since $H^2(\mathbb{P}^2, \mathbb{Z}) \cong \mathbb{Z}$ in this case, the constraint (4.52) implies $d = 0$. The instanton moduli space $\mathfrak{M}_{N,k}(\mathbb{C}^2)$ in this case is a smooth variety of complex dimension $2N\,k$. Furthermore, $n = \chi(\mathbb{P}^2) - 2 = 3 - 2 = 1$. The partition function (4.66) computed with fixed $d = 0$ and $n = 1$ is thus given by

$$Z^{\mathbb{C}^2}_{U(N)}(q)^{\mathrm{fr}} = \widehat{\eta}(q)^{-N}, \tag{4.71}$$

which coincides with the instanton partition function on $\mathbb{C}^2$ described in Section 4.2.

4.7.2. Local $\mathbb{P}^1$

Let $C = C_p$, $p > 0$, be the total space of the holomorphic line bundle $\mathcal{O}_{\mathbb{P}^1}(-p)$ of degree $-p$ over $\mathbb{P}^1$. Then $\overline{C} = \mathbb{F}_p := \mathbb{P}(\mathcal{O}_{\mathbb{P}^1}(-p) \oplus \mathcal{O}_{\mathbb{P}^1})$ is the pth Hirzebruch surface. Let $\ell_0 = \mathbb{P}(0 \oplus \mathcal{O}_{\mathbb{P}^1}) \cong \mathbb{P}^1$ and $\ell_\infty = \mathbb{P}(\mathcal{O}_{\mathbb{P}^1}(-p) \oplus 0) \cong \mathbb{P}^1$. Then $\ell_0 \cdot \ell_0 = -p < 0$, $\ell_\infty \cdot \ell_\infty = p > 0$, and $\ell_0 \cdot \ell_\infty = 0$ so that the line ℓ_0 does not pass through the "line at infinity". From the constraint (4.52) and the second Betti number $b_2(C_p) = 1$, it follows that $d = mc_1(\mathcal{O}_{C_p}(\ell_0))$ for some $m \in \mathbb{Z}$. In this case, the instanton moduli space $\mathfrak{M}_{N,m,k}(C_p)$ is a smooth variety of complex dimension $2Nk - pm^2$. The positivity of this dimension is a constraint on the integral classes which ensures that the moduli space is non-empty. One has $\chi(\mathbb{F}_p) = 1 + 2 + 1 = 4$, and hence $n = \chi(\mathbb{F}_p) - 2 = 2$. The partition function (4.70) in this case becomes

$$Z^{C_p}_{U(N)}(q,Q)^{\mathrm{fr}} = \frac{1}{\widehat{\eta}(q)^{2N}} \sum_{\vec{u} \in \mathbb{Z}^N} q^{(p/2)\, \vec{u} \cdot \vec{u}} Q^u \tag{4.72}$$

with $u = u_1 + \cdots + u_N$.

Like the ALE spaces, one can work instead with a stack compactification $\overline{C}^{\mathrm{orb}}$ of the total space of the bundle $\mathcal{O}_{\mathbb{P}^1}(-p)$ obtained by adding a divisor $\widetilde{\ell}_\infty \cong \ell_\infty/\Gamma$. The resulting variety is a toric Deligne-Mumford stack whose coarse space is the Hirzebruch surface $\mathbb{F}_p$ [72]. With this compactification, one produces framed sheaves with fractional first Chern classes $d = (m/p)\ c_1(\mathcal{O}_{C_p}(\ell_0))$, $m \in \mathbb{Z}$, and in this case the partition function takes the form [72]

$$Z^{C_p}_{U(N)}(q,Q)^{\mathrm{orb}} = \left(\frac{\theta_3(v/p \mid \tau/p)}{\widehat{\eta}(q)^2} \right)^N \tag{4.73}$$

anticipated by [13, 56], where $q = \mathrm{e}^{2\pi \mathrm{i} \tau}$, $Q = \mathrm{e}^{2\pi \mathrm{i} v}$, and $\theta_3(v \mid \tau) = \sum_{n \in \mathbb{Z}} q^{(1/2)\, n^2} Q^n$ is a Jacobi elliptic function.

On these spaces, one also has a version of the Hitchin-Kobayashi correspondence which makes contact with the realization of instantons as self-dual connections. Namely, $SU(N)$-instantons on C_p are in one-to-one correspondence with holomorphic bundles $\mathcal{E}$ of rank N on C_p with $c_1(\mathcal{E}) = 0$ together with a framing at infinity [73]. When $p = 2$, these spaces coincide with the A_1 ALE space.

4.7.3. Compact Surfaces

We will now examine situations under which the formula (4.70) holds in the case of *compact* surfaces, providing rigorous justification for some of the conjectural formulas of [56]. Let C be a compact, smooth, projective toric surface. Let c_∞ be a generic point in C which is disjoint from the torus invariant lines ℓ_e, $e \in E(C)$ of C. Let $\mathfrak{M}_{N,d,k}(C, c_\infty)$ be the moduli space of isomorphism classes $[\mathcal{E}]$ of torsion-free coherent sheaves $\mathcal{E}$ on C with topological Chern invariants as in (4.51), together with a "framing" at the point c_∞. When $\mathcal{E}$ is locally free, this framing means a choice of basis for the fibre space $\mathcal{E}_{c_\infty} \cong \mathbb{C}^N$. When $\mathcal{E}$ is not locally free, the framing is defined with respect to a locally free resolution $\mathcal{E}^\bullet \to \mathcal{E} \to 0$.

Let $\sigma : (\widehat{C}, \widehat{\ell}_\infty) \to (C, c_\infty)$ be the blow-up of C at c_∞. Then $\widehat{C}$ is also a compact, smooth, projective toric surface. We will suppose that the exceptional divisor $\widehat{\ell}_\infty = \sigma^{-1}(c_\infty) \cong \mathbb{P}^1$ has positive self-intersection $\widehat{\ell}_\infty \cdot \widehat{\ell}_\infty > 0$. Let $\mathfrak{M}^{\rm fr}_{N,d,k}(\widehat{C})$ be the moduli space of isomorphism classes $[\,\widehat{\mathcal{E}}\,]$ of torsion-free sheaves $\widehat{\mathcal{E}}$ on $\widehat{C}$ which are framed on the line $\widehat{\ell}_\infty$ as above. The blow-up map σ determines mutually inverse sheaf morphisms $[\widehat{\mathcal{E}}] \mapsto [\mathcal{E}] = [\sigma_* \widehat{\mathcal{E}}]$ and $[\mathcal{E}] \mapsto [\widehat{\mathcal{E}}] = [\sigma^* \mathcal{E}]$, and hence an explicit isomorphism

$$\mathfrak{M}_{N,d,k}(C, c_\infty) \cong \mathfrak{M}^{\rm fr}_{N,d,k}\left(\widehat{C}\right). \tag{4.74}$$

It follows from above that the moduli space $\mathfrak{M}_{N,d,k}(C, c_\infty)$ is thus a smooth variety of complex dimension $2N\, k + d^2$, where $d^2 := \int_C d \wedge d$. Using the isomorphism (4.74), the instanton partition function on C can be computed from the blow-up formula

$$Z^{C}_{U(N)}(q, \mathbf{Q})^{\rm fr} = Z^{\widehat{C} \setminus \widehat{\ell}_\infty}_{U(N)}(q, \mathbf{Q})^{\rm fr}, \tag{4.75}$$

where the right-hand side is given by the formula (4.70).

As an explicit example, consider the complex projective plane $C = \mathbb{P}^2$. Let z_∞ be a generic point on $\mathbb{P}^2$ disjoint from the line $\ell_\infty = [0, z_1, z_2]$. Then the instanton moduli space $\mathfrak{M}_{N,d,k}(\mathbb{P}^2, z_\infty)$ is a smooth variety of complex dimension $2N\, k + d^2$, where $d \in \mathbb{Z}$. Let $\sigma : \widehat{\mathbb{P}}^2 \to \mathbb{P}^2$ be the blowup of $\mathbb{P}^2$ at z_∞ with exceptional divisor $\widehat{\ell}_\infty = \sigma^{-1}(z_\infty) \cong \mathbb{P}^1$. Then there are $n = \chi(\widehat{\mathbb{P}}^2 \setminus \widehat{\ell}_\infty) = \chi(\mathbb{P}^2) = 3$ maximal two cones and $b_2(\widehat{\mathbb{P}}^2 \setminus \widehat{\ell}_\infty) = b_2(\mathbb{P}^2) = 1$ edge in the toric graph $\Delta(\widehat{\mathbb{P}}^2 \setminus \widehat{\ell}_\infty) = \Delta(\mathbb{P}^2)$. Since $\ell_\infty \cdot \ell_\infty = 1$ in this case, the partition function (4.75) gives the $U(N)$ formula

$$Z^{\mathbb{P}^2}_{U(N)}(q, Q)^{\rm fr} = \frac{1}{\widehat{\eta}(q)^{3N}} \sum_{\vec{u} \in \mathbb{Z}^N} q^{-(1/2)\, \vec{u} \cdot \vec{u}} Q^u \tag{4.76}$$

with $u = u_1 + \cdots + u_N$. For $N = 1$ this agrees with the $U(1)$ gauge theory partition function on $\mathbb{P}^2$ derived in [54, Section 5.1].

4.8. Stability

Next we address Point (b) at the end of Section 4.6. If one wishes to relax the requirement of framing, then one must carefully analyse stability issues in order to obtain well-defined instanton moduli spaces. Let C be a smooth, quasiprojective toric surface. Again we fix invariants $N \in \mathbb{N}$, $k \in \mathbb{Q}$, and $d \in H^2_{\rm cpt}(C, \mathbb{Z})$. Let $\mathfrak{M}_{N,d,k}(C)$ be the unframed moduli space consisting of isomorphism classes $[\mathcal{E}]$ of semistable torsion-free coherent sheaves $\mathcal{E}$ on C such that $\mathcal{E}$ has topological Chern invariants

$$N = \mathrm{ch}_0(\mathcal{E}) = \mathrm{rank}(\mathcal{E}), \qquad d = \mathrm{ch}_1(\mathcal{E}) = c_1(\mathcal{E}), \qquad k = -\int_C \mathrm{ch}_2(\mathcal{E}). \tag{4.77}$$

The notion of "semistability" considers sheaves with fixed Hilbert polynomial. In this case, a general result of Maruyama [74, 75] constructs the algebraic scheme $\mathfrak{M}_{N,d,k}(C)$ and shows that it is *projective*. In particular, it admits a $\widetilde{T}$-equivariant embedding into a smooth variety. In the case that C is a polarized surface, that is, it admits a smooth distribution in its tangent bundle TC which is integrable and Lagrangian (in an appropriate sense), we can formulate the notion of semistability in terms of the more familiar slope semistability following [76]. Let $\mathcal{L}$ be a fixed ample line bundle on C. For a coherent torsion-free sheaf $\mathcal{E}$ on C, define the polynomial

$$\rho_{\mathcal{E}}(n) = \frac{\chi(\mathcal{E} \otimes \mathcal{L}^{\otimes n})}{\operatorname{rank}(\mathcal{E})} \tag{4.78}$$

for $n \in \mathbb{N}_0$. Then $\mathcal{E}$ is said to be $\mathcal{L}$-semistable if

$$\rho_{\mathcal{F}}(n) \leq \rho_{\mathcal{E}}(n) \tag{4.79}$$

for all $n \gg 0$ whenever $\mathcal{F}$ is a coherent subsheaf of $\mathcal{E}$. In this case, Gieseker [76] constructs $\mathfrak{M}_{N,d,k}(C)$ as a projective *Quot* scheme. The slope of $\mathcal{E}$ is the rational number

$$\mu(\mathcal{E}) = \frac{\deg(\mathcal{E})}{\operatorname{rank}(\mathcal{E})}, \tag{4.80}$$

where the degree of $\mathcal{E}$ is defined using the polarization as

$$\deg(\mathcal{E}) = \int_C c_1(\mathcal{E}) \wedge c_1(\mathcal{L}). \tag{4.81}$$

An application of the Riemann-Roch theorem shows [76]

$$\rho_{\mathcal{E}}(n) = \frac{n}{2} \int_C \left(c_1(\mathcal{L}) \wedge (n c_1(\mathcal{L}) + c_1(C)) + \frac{2\operatorname{ch}_2(\mathcal{E}) + c_1(\mathcal{E}) \wedge c_1(C)}{\operatorname{rank}(\mathcal{E})} \right) + \mu(\mathcal{E}) + \chi(\mathcal{O}_C). \tag{4.82}$$

It follows that $\rho_{\mathcal{E}}(n) \geq \rho_{\mathcal{F}}(n)$ for $n \gg 0$ if and only if $\mu(\mathcal{E}) \geq \mu(\mathcal{F})$, and hence $\mathcal{L}$-semistability is equivalent to the usual quasi-BPS instanton equations in this case. In physics applications, one is interested in instances where C is a Kähler surface. In this case one can use the Kähler polarization and take $c_1(\mathcal{L})$ to be the Kähler two-form k_0 in (4.32).

The following result computes the expected dimension of the instanton moduli space in this case.

Lemma 4.1. *The virtual dimension of* $\mathfrak{M}_{N,d,k}(C)$ *equals* $2N\,k + d^2 - (N^2-1)\,\chi(\mathcal{O}_C)$, *where* $d^2 := \int_C d \wedge d$ *and* $\chi(\mathcal{O}_C)$ *is the holomorphic Euler characteristic of* C.

Proof. As discussed in Section 3.5., the dimension of the virtual tangent space $T^{\mathrm{vir}}_{[\mathcal{E}]}\mathfrak{M}_{N,d,k}(C) = \operatorname{Ext}^1_{\mathcal{O}_C}(\mathcal{E},\mathcal{E}) \ominus \operatorname{Ext}^2_{\mathcal{O}_C}(\mathcal{E},\mathcal{E})$ to the instanton moduli space at a point $[\mathcal{E}] \in \mathfrak{M}_{N,d,k}(C)$ in obstruction theory is the difference of Euler characteristics $\chi(\mathcal{O}_C \otimes \mathcal{O}_C^{\vee}) - \chi(\mathcal{E} \otimes \mathcal{E}^{\vee})$. The latter quantity may be computed for $\mathcal{E}$ locally free by using the Hirzebruch-Riemann-Roch theorem to write

$$\chi\big(\mathcal{E}\otimes\mathcal{E}^\vee\big) = \int_C \operatorname{ch}\big(\mathcal{E}\otimes\mathcal{E}^\vee\big)\wedge \operatorname{Td}(C). \tag{4.83}$$

Let ν be the generator of $H^4(C,\mathbb{Z})\cong\mathbb{Z}$ which is Poincaré dual to a point in C. Then using (4.77) one computes the Chern character

$$\operatorname{ch}\big(\mathcal{E}\otimes\mathcal{E}^\vee\big) = \operatorname{ch}(\mathcal{E})\wedge\operatorname{ch}\big(\mathcal{E}^\vee\big) = (N+d-k\nu)\wedge(N-d-k\nu) = N^2 - d\wedge d - 2Nk\nu. \tag{4.84}$$

The Todd characteristic class of the tangent bundle of C is given in terms of Chern classes of TC as

$$\operatorname{Td}(C) = 1 + \frac{1}{2}c_1(C) + \frac{1}{12}(c_1(C)\wedge c_1(C) + c_2(C)). \tag{4.85}$$

We may thus write (4.83) as

$$\begin{aligned}\chi\big(\mathcal{E}\otimes\mathcal{E}^\vee\big) &= \int_C\left(\frac{N^2}{12}(c_1(C)\wedge c_1(C)+c_2(C)) - d\wedge d - 2Nk\nu\right)\\ &= -2N\ k - d^2 + N^2\int_C \operatorname{Td}(C).\end{aligned} \tag{4.86}$$

From the Hirzebruch-Riemann-Roch formula and $\operatorname{ch}(\mathcal{O}_C)=1$, one has

$$\chi(\mathcal{O}_C) = \int_C \operatorname{Td}(C) = \chi\big(\mathcal{O}_C\otimes\mathcal{O}_C^\vee\big) \tag{4.87}$$

and the result follows when $\mathcal{E}$ is a bundle. When $\mathcal{E}$ is not locally free, we use a locally free resolution $\mathcal{E}^\bullet \to \mathcal{E} \to 0$ along with additivity of the Chern character. □

This result shows that the expected dimension of the instanton moduli space for $N=1$ coincides with the dimension of the framed moduli space $\mathfrak{M}_{1,d,k}(C)$ of Section 4.7. Indeed, torsion-free sheaves of rank one are always stable in the sense explained above. Since any torsion-free sheaf decomposes as the product $\mathcal{E}=\mathcal{I}\otimes\mathcal{L}$ of an ideal sheaf of a zero-dimensional subscheme and a line bundle, the moduli space factorizes into a product

$$\mathfrak{M}_{1,d,k}(C) = \mathfrak{M}_{1,0,k+d^2/2}(C)\times\mathfrak{M}_{1,d,-d^2/2}(C), \tag{4.88}$$

where $\mathfrak{M}_{1,0,K}(C)\cong C^{[K]}$ is the Hilbert scheme of K points on C which is a smooth variety of dimension $2K$, and the (zero-dimensional) Picard group $\mathfrak{M}_{1,d,-d^2/2}(C)\cong\operatorname{Pic}^d(C)$ parameterizes fractional instantons. The factorization of the $U(1)$ gauge theory partition function into contributions from regular and fractional instantons then follows from the multiplicativity of the Euler class under tensor product of (tangent) bundles. For compact toric surfaces C, it is given by the formula (4.70) with $N = 1$. In the case of Hirzebruch-Jung surfaces, the contributions of fractional charges in this case are shown in [54] to arise from the noncompact prime divisors of C via linear equivalences. It amounts to identifying the generators $c_1(\mathcal{R}_i)$ of the Kähler cone in $CH^1(C) \otimes \mathbb{Q}$ with the duals to the exceptional divisors D_i which generate the Mori cone in $CH_1^{\mathrm{cpt}}(C) \otimes \mathbb{Q}$. For the ALE spaces, the moduli space $\mathfrak{M}_{1,0,k+d^2/2}(C)$ coincides with the quiver variety of the ADHM construction (with suitable stability conditions) [77].

The situation for higher rank $N > 1$ is much more complicated. In this case, slope-stability does not seem to properly account for walls of marginal stability extending to infinity which describe wall-crossing behaviour of the partition functions counting D4-D2-D0 brane bound states on Calabi-Yau manifolds X with $h^{1,1}(X) > 1$ [64]. As discussed in Section 3.9., the physical theory is described by moduli spaces of stable objects in the derived category $\mathrm{D}^b(\mathrm{coh}(X))$, as the observed D-brane decays are impossible in the abelian category $\mathrm{coh}(X)$ of coherent sheaves on X.

4.9. Perpendicular Partition Functions and Universal Sheaves

Finally, we come to the last Point (c) at the end of Section 4.6. For $U(1)$ gauge theory, this relationship is analysed in detail in [54]. In this case a four-dimensional analog of the topological vertex formalism can be developed. On the gauge theory side, the vertex contributions should be computed by a version of the perpendicular partition function of Section 3.8., that is, the generating function $P^{m_1,m_2}_{U(N)}(q)$ for instantons on $C = \mathbb{C}^2$ with fixed asymptotics such that $Z^{\mathbb{C}^2}_{U(N)}(q) = P^{0,0}_{U(N)}(q)$, while the edge contributions should be read off from the character of the corresponding universal sheaf. In the remainder of this section, we will describe in detail the instanton moduli space with boundary conditions specified by integers m_1, and m_2, and show that the fixed point loci of the induced $\widetilde{T}$-action are enumerated by two-dimensional Young diagrams with asymptotics m_1, and m_2. This gives the four-dimensional version of the gauge theory gluing rules of Section 3.8. and also a first principle derivation of the empirical vertex rules of [54, Section 4.3].

4.9.1. Instanton Moduli Spaces on $\mathbb{F}_0$

We will first describe how to generate instantons with nontrivial first Chern class using our previous formalism. For this, rather than working with the projective plane $\overline{C} = \mathbb{P}^2$, it is more convenient to work with the "two-point compactification" of $C = \mathbb{C}^2$ given by a product of two projective lines $\mathbb{F}_0 = \mathbb{P}^1_z \times \mathbb{P}^1_w$, where the labels will be used to keep track of each of the two factors. This variety can be identified as the zeroth Hirzebruch surface $\mathbb{F}_0 = \mathbb{P}(\mathcal{O}_{\mathbb{P}^1} \oplus \mathcal{O}_{\mathbb{P}^1})$, that is, the projective compactification of the trivial line bundle $\mathcal{O}_{\mathbb{P}^1} = \mathbb{C} \times \mathbb{P}^1$, which is again a toric surface. For $(t_z, t_w) \in T^2$, the toric action is described by the automorphism F_{t_z,t_w} of $\mathbb{F}_0$ defined by

$$F_{t_z,t_w}([z_0, z_1];\ [w_0, w_1]) = ([z_0, t_z z_1]; [w_0, t_w w_1]). \tag{4.89}$$

Let $p_z : \mathbb{F}_0 \to \mathbb{P}^1_z$ and $p_w : \mathbb{F}_0 \to \mathbb{P}^1_w$ be the canonical projections, and $i_z : \mathbb{P}^1_z \cong \mathbb{P}^1_z \times \mathrm{pt} \hookrightarrow \mathbb{F}_0$ and $i_w : \mathbb{P}^1_w \cong \mathrm{pt} \times \mathbb{P}^1_w \hookrightarrow \mathbb{F}_0$ the natural inclusions. There are now two independent distinguished lines at infinity fixed by this T^2 action, given by the pushforwards $\ell_z = (i_z)_* \mathbb{P}^1_z$ and $\ell_w = (i_w)_* \mathbb{P}^1_w$, for which the respective factors of the action of $T^2 = \mathbb{C}^* \times \mathbb{C}^*$ reduce to the standard $\mathbb{C}^*$-action on $\ell_z, \ell_w \cong \mathbb{P}^1$. These two divisors have intersection products $\ell_z \cdot \ell_z = \ell_w \cdot \ell_w = 0$ and $\ell_z \cdot \ell_w = 1$, and the canonical divisor is $K_{\mathbb{F}_0} = -2\ell_z - 2\ell_w$. They generate the Picard group of $\mathbb{F}_0$, and hence induce a bigrading on line bundles over $\mathbb{F}_0$. For $m_z, m_w \in \mathbb{Z}$, we write

$$\mathcal{O}_{\mathbb{F}_0}(m_z, m_w) := \mathcal{O}_{\mathbb{F}_0}(m_z \ell_z + m_w \ell_w) = p_z^* \mathcal{O}_{\mathbb{P}^1_z}(m_z) \otimes p_w^* \mathcal{O}_{\mathbb{P}^1_w}(m_w). \tag{4.90}$$

By Künneth's theorem, the nontrivial cohomology groups of $\mathbb{F}_0$ are given by

$$H^0(\mathbb{F}_0, \mathbb{Z}) = \mathbb{Z}, \qquad H^2(\mathbb{F}_0, \mathbb{Z}) = \mathbb{Z}[\xi_z] \oplus \mathbb{Z}[\xi_w], \qquad H^4(\mathbb{F}_0, \mathbb{Z}) = \mathbb{Z}[\xi_z \wedge \xi_w], \tag{4.91}$$

where $\xi_z = c_1(\mathcal{O}_{\mathbb{F}_0}(1,0)) = p_z^* c_1(\mathcal{O}_{\mathbb{P}^1_z}(1))$ and $\xi_w = c_1(\mathcal{O}_{\mathbb{F}_0}(0,1)) = p_w^* c_1(\mathcal{O}_{\mathbb{P}^1_w}(1))$.

Let $\mathfrak{M}_{N;(m_z,m_w);k}(\mathbb{F}_0)$ be the moduli space of isomorphism classes $[\mathcal{E}]$ of torsion-free sheaves $\mathcal{E}$ on $\mathbb{F}_0$ with topological Chern invariants as in (4.51), where $d = m_z \xi_z + m_w \xi_w$, and nontrivial framings along the two independent "directions at infinity" are prescribed by two fixed isomorphisms $\mathcal{E}|_{\ell_z} \cong W \otimes \mathcal{O}_{\mathbb{P}^1_z}(m_z)$ and $\mathcal{E}|_{\ell_w} \cong W \otimes \mathcal{O}_{\mathbb{P}^1_w}(m_w)$, where W is a fixed N-dimensional complex vector space. We will say that such a sheaf is "trivialized at infinity" if it is equipped with two isomorphisms $\mathcal{E}|_{\ell_z} \cong W \otimes \mathcal{O}_{\mathbb{P}^1_z}$ and $\mathcal{E}|_{\ell_w} \cong W \otimes \mathcal{O}_{\mathbb{P}^1_w}$. The moduli space of trivialized sheaves is denoted $\mathfrak{M}_{N,k}(\mathbb{F}_0) := \mathfrak{M}_{N;(0,0);k}(\mathbb{F}_0)$. These moduli spaces carry an obvious induced action of the torus $\widetilde{T} = T^2 \times T^N$, and the following formal arguments show that the instanton counting in these moduli spaces is the same as before.

Proposition 4.2. *There is a natural $\widetilde{T}$-equivariant birational equivalence between the moduli spaces*

$$\mathfrak{M}_{N,k}(\mathbb{F}_0) \cong \mathfrak{M}_{N,k}\left(\mathbb{C}^2\right). \tag{4.92}$$

Proof. Represent $\mathbb{F}_0$ as a divisor in $\mathbb{P}^2 \times \mathbb{P}^1$ via the embedding $\mathbb{P}^1_z \hookrightarrow \mathbb{P}^2$, $[z_0, z_1] \mapsto [z_0, z_1, z_1]$. Let $\widehat{\widehat{\mathbb{P}}}^2$ be the surface obtained as the blowup p of a pair of points on the line at infinity $\ell_\infty \subset \mathbb{P}^2$ (see Section 4.7.). Note that ℓ_∞ is disjoint from the image of the line ℓ_z. Then there is a correspondence diagram

$$\begin{tikzcd} & \widehat{\widehat{\mathbb{P}}}^2 \arrow[dl, "p"'] \arrow[dr, "q"] & \\ \mathbb{P}^2 & & \mathbb{F}_0 \end{tikzcd} \tag{4.93}$$

where q is the blowup of the intersection point $\ell_z \cdot \ell_w$ on $\mathbb{F}_0 = \mathbb{P}^1_z \times \mathbb{P}^1_w$. The corresponding Fourier-Mukai functors $q_* p^*$ and $p_* q^*$ determine equivalences on the categories of coherent sheaves over $\mathbb{P}^2$ and $\mathbb{F}_0$, and hence induce mutually inverse rational maps between the corresponding sets of isomorphism classes of trivially framed torsion-free sheaves. $\square$

Proposition 4.3. *For each fixed $(m_z, m_w) \in \mathbb{Z}^2$, there is a natural $\widetilde{T}$-equivariant bijection between the moduli spaces*

$$\mathfrak{M}_{N,k}(\mathbb{F}_0) \cong \mathfrak{M}_{N;(Nm_z,Nm_w);k(m_z,m_w)}(\mathbb{F}_0), \tag{4.94}$$

where

$$k(m_z, m_w) = k - Nm_z m_w. \tag{4.95}$$

Proof. Given $[\mathcal{E}] \in \mathfrak{M}_{N,k}(\mathbb{F}_0)$, define

$$\mathcal{E}(m_z, m_w) = \mathcal{E} \otimes \mathcal{O}_{\mathbb{F}_0}(m_z, m_w). \tag{4.96}$$

By (4.51) (with $d = 0$), one has

$$\mathrm{ch}(\mathcal{E}) = N - k\xi_z \wedge \xi_w, \tag{4.97}$$

and using multiplicativity of the Chern character we compute

$$\begin{aligned} \mathrm{ch}(\mathcal{E}(m_z, m_w)) &= (N - k\xi_z \wedge \xi_w) \wedge (1 + m_z\xi_z) \wedge (1 + m_w\xi_w) \\ &= N + N(m_z\xi_z + m_w\xi_w) + (Nm_z m_w - k)\xi_z \wedge \xi_w. \end{aligned} \tag{4.98}$$

This therefore gives a map $[\mathcal{E}] \mapsto [\mathcal{E}(m_z, m_w)]$ on $\mathfrak{M}_{N,k}(\mathbb{F}_0) \to \mathfrak{M}_{N;(N\, m_z,Nm_w);k(m_z,m_w)}(\mathbb{F}_0)$. An identical calculation shows that the map $\mathcal{E}(m_z, m_w) \mapsto \mathcal{E}(m_z, m_w) \otimes \mathcal{O}_{\mathbb{F}_0}(-m_z, -m_w)$ is its inverse. These maps clearly induce bijections between (m_z, m_w)-framings and trivializations at infinity. □

Proposition 4.2 shows that an instanton gauge bundle $\mathcal{E}$ with trivial asymptotics on $\mathbb{F}_0$ can again be regarded as an element $[\mathcal{E}] \in \mathfrak{M}_{N,k}(\mathbb{C}^2)$. For fixed $m_z, m_w \in \mathbb{Z}$, Proposition 4.3 shows that the counting of torsion-free sheaves in the moduli spaces $\mathfrak{M}_{N,k}(\mathbb{F}_0)$ and $\mathfrak{M}_{N;(Nm_z,Nm_w);k(m_z,m_w)}(\mathbb{F}_0)$ coincides. In other words, the number of instantons with fixed nontrivial asymptotics matches the Young tableau count, as the number of sheaves (4.96) is independent of the integers m_z, m_w. Proposition 4.3 also reproduces the formula [54, Section 4.2.2] for the renormalized volume of Young tableaux (reproduced above for $N = 1$) as the induced shift in instanton charge in (4.95) of the sheaves (4.96).

The bijection of Proposition 4.2 is not necessarily a diffeomorphism. In fact, we have not yet shown that $\mathfrak{M}_{N,k}(\mathbb{F}_0)$ is a smooth variety. Naively, one may try to argue this by using the formalism of [69], as our sheaves are trivialized on the divisor $\ell_\infty = \ell_z + \ell_w$ which has positive self-intersection $\ell_\infty \cdot \ell_\infty = 2$, and $\mathbb{F}_0 \setminus \ell_\infty \cong \mathbb{C}^2$. However, this divisor is not irreducible; in particular, it is not a $\mathbb{P}^1$. Note that formally applying the arguments which led to (4.72) in the limit $p = 0$ would produce the partition function

$$Z_{U(N)}^{\mathbb{F}_0}(q) = \left(Z_{U(N)}^{\mathbb{C}^2}(q)\right)^2, \tag{4.99}$$

as the Hirzebruch surface has $n = \chi(\mathbb{F}_0) - 2 = 2$. Although the generating function (4.99) is "doubled", the instanton counting problems on $\mathbb{F}_0$ and on $\mathbb{C}^2$ are nevertheless identical. A rigorous derivation of the formula (4.99) using the techniques of Section 4.7. follows once we establish that the equivalence of Proposition 4.2 is an isomorphism of underlying smooth varieties. Recall from Section 4.7. that the complex dimension of the instanton moduli space $\mathfrak{M}_{N,k}(\mathbb{C}^2)$ is $2N\ k$.

Proposition 4.4. *The moduli space $\mathfrak{M}_{N,k}(\mathbb{F}_0)$ is a smooth quasiprojective variety of complex dimension $2N\ k$.*

Proof. As usual, the trivialization condition at infinity guarantees Gieseker semistability and hence quasi-projectivity, as discussed in Section 4.8. Using the divisor $\ell_\infty = \ell_z + \ell_w$ of $\mathbb{F}_0$, one constructs the framed moduli space $\mathfrak{M}_{N,k}(\mathbb{F}_0)$ as in [68] with tangent spaces $\mathrm{Ext}^1_{\mathcal{O}_{\mathbb{F}_0}}(\mathcal{E}, \mathcal{E}(-\ell_\infty))$ and obstruction spaces given by $\mathrm{Ext}^2_{\mathcal{O}_{\mathbb{F}_0}}(\mathcal{E}, \mathcal{E}(-\ell_\infty))$. By reference [78, Section 5], one has

$$\mathrm{Ext}^0_{\mathcal{O}_{\mathbb{F}_0}}(\mathcal{E}, \mathcal{E}(-1,-1)) = \mathrm{Ext}^2_{\mathcal{O}_{\mathbb{F}_0}}(\mathcal{E}, \mathcal{E}(-1,-1)) = 0 \tag{4.100}$$

and hence $\mathfrak{M}_{N,k}(\mathbb{F}_0)$ is smooth. The dimension of the tangent spaces is thus equal to minus the Euler characteristic $-\chi(\mathcal{E}, \mathcal{E}(-1,-1))$, which for $\mathcal{E}$ locally free may be calculated by using the Hirzebruch-Riemann-Roch theorem to write

$$\chi(\mathcal{E}, \mathcal{E}(-1,-1)) = \int_{\mathbb{F}_0} \mathrm{ch}\big(\mathcal{E}^\vee \otimes \mathcal{E}(-1,-1)\big) \wedge \mathrm{Td}(\mathbb{F}_0). \tag{4.101}$$

Using (4.97) and (4.98), one computes the Chern character

$$\begin{aligned}\mathrm{ch}\big(\mathcal{E}^\vee \otimes \mathcal{E}(-1,-1)\big) &= \mathrm{ch}\big(\mathcal{E}^\vee\big) \wedge \mathrm{ch}(\mathcal{E}(-1,-1)) \\ &= (N - k\xi_z \wedge \xi_w) \wedge (N - N(\xi_z + \xi_w) + (N-k)\xi_z \wedge \xi_w) \\ &= N^2 - N^2(\xi_z + \xi_w) + \Big(N^2 - 2Nk\Big)\xi_z \wedge \xi_w.\end{aligned} \tag{4.102}$$

The Todd characteristic class and holomorphic Euler characteristic are given by (4.85) and (4.87) with $C = \mathbb{F}_0$, and we may thus write (4.101) as

$$\begin{aligned}\chi(\mathcal{E}, \mathcal{E}(-1,-1)) &= \int_{\mathbb{F}_0} \Bigg(\frac{N^2}{12}\ (c_1(\mathbb{F}_0) \wedge c_1(\mathbb{F}_0) + c_2(\mathbb{F}_0)) - \frac{N^2}{2}\ (\xi_z + \xi_w) \wedge c_1(\mathbb{F}_0) \\ &\qquad + \Big(N^2 - 2N\ k\Big)\xi_z \wedge \xi_w \Bigg) \\ &= N^2\ \chi(\mathcal{O}_{\mathbb{F}_0}) - \frac{N^2}{2} \int_{\mathbb{F}_0} c_1(\mathbb{F}_0) \wedge (\xi_z + \xi_w) + N^2 - 2N\ k.\end{aligned} \tag{4.103}$$

One has $\chi(\mathcal{O}_{\mathbb{F}_0}) = 1$ since $b_2(\mathbb{F}_0) = 2 = h^{1,1}(\mathbb{F}_0)$. Since $\xi_z, \xi_w \in H^2(\mathbb{F}_0, \mathbb{Z})$ are the Poincaré duals of $[\ell_z], [\ell_w] \in H_2(\mathbb{F}_0, \mathbb{Z})$, with $\ell_z, \ell_w \cong \mathbb{P}^1$ and $\ell_z \cdot \ell_z = \ell_w \cdot \ell_w = 0$, the remaining integral in (4.103) may be computed as

$$\int_{\mathbb{F}_0} c_1(\mathbb{F}_0) \wedge (\xi_z + \xi_w) = 2 \int_{\mathbb{P}^1} c_1\left(\mathbb{P}^1\right) = 4, \tag{4.104}$$

where in the last step we used the fact that the holomorphic tangent bundle of $\mathbb{P}^1$ can be identified with $\mathcal{O}_{\mathbb{P}^1}(2)$. Putting everything together, we find

$$\chi(\mathcal{E}, \mathcal{E}(-1,-1)) = -2N\ k \tag{4.105}$$

and the result follows when $\mathcal{E}$ is a bundle. If $\mathcal{E}$ is not locally free, then we simply consider a locally free resolution $\mathcal{E}^\bullet \to \mathcal{E} \to 0$ and use additivity of the Chern characteristic class. □

In a similar vein, the equivalence established in Proposition 4.3 as it stands is only a bijective correspondence. The moduli space for nontrivial asymptotics is an example of the moduli spaces of "decorated sheaves" described in [20, Section 4.B], with the (m_z, m_w)-framed sheaves on $\mathbb{F}_0$ providing examples of semistable framed modules. The moduli space $\mathfrak{M}_{N;(Nm_z,\, Nm_w);k(m_z,m_w)}(\mathbb{F}_0)$ is thus a projective scheme with a universal family, that is, it is fine, and so possesses a universal sheaf. We will now argue that the moduli spaces $\mathfrak{M}_{N;(Nm_z,Nm_w);k(m_z,m_w)}(\mathbb{F}_0)$ are smooth and diffeomorphic to one another for all $m_z, m_w \in \mathbb{Z}$. Then the computations of Section 4.7. can be repeated by integrating over $\mathfrak{M}_{N;(Nm_z,Nm_w);k(m_z,m_w)}(\mathbb{F}_0)$ using (4.95) and Göttsche's formula to get the desired perpendicular partition function

$$P^{m_z,m_w}_{U(N)}(q) = q^{-Nm_z m_w}\, \widehat{\eta}(q)^{-2N}. \tag{4.106}$$

For $N = 1$, this yields a rigorous derivation of the vertex rules of [54].

Let $\widetilde{\mathfrak{M}}(\mathbb{F}_0)$ be the moduli space of isomorphism classes of pairs $\widetilde{\mathcal{E}} = (\mathcal{E}, \alpha)$, where $\mathcal{E}$ is an (m_z, m_w)-framed torsion-free sheaf on $\mathbb{F}_0$ and $\alpha : \mathcal{O}_{\mathbb{F}_0}(m_z, m_w)^{\oplus N} \to \mathcal{E}$ is the surjective morphism induced by the framing. Consider the family of pairs $\widetilde{\mathcal{F}} = (\mathcal{F}, \beta)$ consisting of a coherent sheaf $\mathcal{F}$ on $\mathbb{F}_0$ together with a homomorphism $\beta : \mathcal{O}_{\mathbb{F}_0}(m_z, m_w)^{\oplus r} \to \mathcal{F}$ for some $r \in \mathbb{N}$. These pairs define objects of an abelian category $\mathrm{coh}_f(\mathbb{F}_0)$ with the obvious morphisms between pairs induced by morphisms of coherent sheaves. Any (m_z, m_w)-framed torsion-free sheaf $\mathcal{E}$ on $\mathbb{F}_0$ clearly defines an object of $\mathrm{coh}_f(\mathbb{F}_0)$, with $r = N$ and $\mathrm{ch}(\mathcal{F}) = \mathrm{ch}(\mathcal{E})$.

Given the abelian category $\mathrm{coh}_f(\mathbb{F}_0)$, one can derive functors, and hence compute sheaf cohomology in $\mathrm{coh}_f(\mathbb{F}_0)$. We denote the corresponding Ext-functors by Ext^i_f. Then $\mathrm{Ext}^2_f(\widetilde{\mathcal{E}}, \widetilde{\mathcal{E}})$ is the obstruction space for $\widetilde{\mathfrak{M}}(\mathbb{F}_0)$ at $\widetilde{\mathcal{E}}$, while $\mathrm{Ext}^1_f(\widetilde{\mathcal{E}}, \widetilde{\mathcal{E}})$ is the tangent space to $\widetilde{\mathfrak{M}}(\mathbb{F}_0)$ at $\widetilde{\mathcal{E}}$.

Generally, for any two objects $\widetilde{\mathcal{F}}, \widetilde{\mathcal{G}}$ of $\mathrm{coh}_f(\mathbb{F}_0)$, the group $\mathrm{Ext}^1_f(\widetilde{\mathcal{F}}, \widetilde{\mathcal{G}})$ can be computed in a purely algebraic way in terms of equivalence classes of extensions

$$\begin{tikzcd}
\mathcal{O}_{\mathbb{F}_0}(m_z, m_w)^{\oplus s} \arrow[r] \arrow[d] & \mathcal{O}_{\mathbb{F}_0}(m_z, m_w)^{\oplus(s+r)} \arrow[r] \arrow[d] & \mathcal{O}_{\mathbb{F}_0}(m_z, m_w)^{\oplus r} \arrow[d] \\
\widetilde{\mathcal{G}} \arrow[r] & \widetilde{\mathcal{K}} \arrow[r] & \widetilde{\mathcal{F}}.
\end{tikzcd} \tag{4.107}$$

This definition of $\mathrm{Ext}^1_f(\widetilde{\mathcal{F}}, \widetilde{\mathcal{G}})$ is equivalent to the usual definition in terms of a twisted Dolbeault complex, and in particular it gives the deformation theory of the moduli space $\widetilde{\mathfrak{M}}(\mathbb{F}_0)$. Furthermore, there is a spectral sequence connecting it to the ordinary Ext-groups $\mathrm{Ext}^i_{\mathcal{O}_{\mathbb{F}_0}}(\mathcal{F}, \mathcal{G})$.

The line bundle $\mathcal{O}_{\mathbb{F}_0}(m_z, m_w)$ is a projective $\mathcal{O}_{\mathbb{F}_0}$-module and is therefore flat, whence the tensor product map defined by (4.96) preserves exact sequences such as (4.107). It follows that the tangent spaces $\mathrm{Ext}^1_f(\widetilde{\mathcal{E}}, \widetilde{\mathcal{E}})$ are the same for all $m_z, m_w \in \mathbb{Z}$. Thus by Proposition 4.4, the moduli spaces $\widetilde{\mathfrak{M}}(\mathbb{F}_0)$ are all smooth and diffeomorphic to one another. This is consistent with the fact that stability (either slope stability of torsion-free sheaves [76] or stability of framed modules [20, Section 4.B]) is preserved by twisting with the line bundles $\mathcal{O}_{\mathbb{F}_0}(m_z, m_w)$.

4.9.2. Universal Sheaves

We now construct the universal sheaf for the standard framing on $\mathbb{F}_0$, and compute its $\widetilde{T}$-equivariant character. Since the moduli space $\mathfrak{M}_{N,k}(\mathbb{F}_0)$ is fine [68], there exists a universal sheaf $\mathfrak{E} \to \mathbb{F}_0 \times \mathfrak{M}_{N,k}(\mathbb{F}_0)$, that is, a torsion-free sheaf $\mathfrak{E}$ such that $\mathfrak{E}|_{\mathbb{F}_0 \times [\mathcal{E}]} \cong \mathcal{E}$ for every point $[\mathcal{E}] \in \mathfrak{M}_{N,k}(\mathbb{F}_0)$, unique up to tensor product with a (unique) line bundle. There are two natural vector bundles over $\mathfrak{M}_{N,k}(\mathbb{F}_0)$ associated to a universal sheaf $\mathfrak{E}$. Firstly, there is the framing bundle $W = H^0(\mathbb{F}_0, \mathfrak{E}|_{\ell_\infty})$ of rank N given by the fibre at infinity. Secondly, there is the bundle V of "Dirac zero modes", whose fibre at a given locally free sheaf $[\mathcal{E}] \in \mathfrak{M}_{N,k}(\mathbb{F}_0)$ restricted to $\mathbb{R}^4 \cong \mathbb{C}^2 \cong \mathbb{F}_0 \setminus \ell_\infty$ is the space of L^2-solutions to the Dirac equation on $\mathbb{R}^4$ in the background of the fundamental representation of the instanton gauge field corresponding to $[\mathcal{E}]|_{\mathbb{C}^2}$. It is constructed explicitly as follows. Let π_1 and π_2 be the canonical projections of $\mathbb{F}_0 \times \mathfrak{M}_{N,k}(\mathbb{F}_0)$ onto the first and second factors, respectively. Then the Dirac bundle is

$$V = R^1\pi_{2*}\left(\mathfrak{E} \otimes \pi_1^*\mathcal{O}_{\mathbb{F}_0}(-1,-1)\right), \tag{4.108}$$

where $R^1\pi_{2*}$ is the first right derived functor of the pushforward functor π_{2*}. Its fibre over a point $[\mathcal{E}] \in \mathfrak{M}_{N,k}(\mathbb{F}_0)$ is the vector space $H^1(\mathbb{F}_0, \mathcal{E}(-1,-1))$.

Proposition 4.5. *The Dirac bundle V is a vector bundle of rank k over $\mathfrak{M}_{N,k}(\mathbb{F}_0)$.*

Proof. By reference [78, Section 5], one has

$$H^0(\mathbb{F}_0, \mathcal{E}(-1,-1)) = H^2(\mathbb{F}_0, \mathcal{E}(-1,-1)) = 0 \tag{4.109}$$

and hence V is a vector bundle of rank equal to $-\chi(\mathcal{E}(-1,-1))$. For $\mathcal{E}$ locally free, the Hirzebruch-Riemann-Roch theorem, together with (4.98), (4.85) with $C = \mathbb{F}_0$, (4.104), and $\chi(\mathcal{O}_{\mathbb{F}_0}) = 1$, gives

$$\begin{aligned}
\operatorname{rank}(V) &= -\int_{\mathbb{F}_0} \operatorname{ch}(\mathcal{E}(-1,-1)) \wedge \operatorname{Td}(\mathbb{F}_0) \\
&= -\int_{\mathbb{F}_0} (N - N\ (\xi_z + \xi_w) + (N-k)\xi_z \wedge \xi_w) \\
&\quad \wedge \left(1 + \frac{1}{2}\ c_1(\mathbb{F}_0) + \frac{1}{12}\ (c_1(\mathbb{F}_0) \wedge c_1(\mathbb{F}_0) + c_2(\mathbb{F}_0))\right) \\
&= -N\chi(\mathcal{O}_{\mathbb{F}_0}) + \frac{N}{2} \int_{\mathbb{F}_0} c_1(\mathbb{F}_0) \wedge (\xi_z + \xi_w) + k - N = k.
\end{aligned} \tag{4.110}$$

Again, the statement for generic torsion-free sheaves $[\mathcal{E}] \in \mathfrak{M}_{N,k}(\mathbb{F}_0)$ follows by considering a locally free resolution $\mathcal{E}^\bullet \to \mathcal{E} \to 0$. □

By definition, (4.89) and the construction of Section 4.7., the $\widetilde{T}$-equivariant characters of the framing and Dirac bundles regarded as $\widetilde{T}$-modules, at a fixed point in $\mathfrak{M}_{N,k}(\mathbb{F}_0)^{\widetilde{T}}$ labelled by an N-coloured Young tableau $\vec{\lambda} = (\lambda_1, \ldots, \lambda_N)$, are given by

$$\operatorname{ch}_{\widetilde{T}}(W)\left(\vec{\lambda}\right) = \sum_{l=1}^{N} e_l, \qquad \operatorname{ch}_{\widetilde{T}}(V)\left(\vec{\lambda}\right) = \sum_{l=1}^{N} e_l \sum_{(i,j)\in\lambda_l} t_z^{i-1} t_w^{j-1}, \tag{4.111}$$

where $e_l = \mathrm{e}^{a_l}$. Let us now describe how the ADHM construction is modified in this case. This has been worked out in [78, Section 5]. The linear algebraic data is formally the same as for the analysis of framed instantons on $\mathbb{P}^2$ [49], defined by linear operators

$$B_z \in \operatorname{Hom}(V,V), \qquad B_w \in \operatorname{Hom}(V,V), \qquad I \in \operatorname{Hom}(W,V), \qquad J \in \operatorname{Hom}(V,W). \tag{4.112}$$

Then any framed torsion-free sheaf $[\mathcal{E}] \in \mathfrak{M}_{N,k}(\mathbb{F}_0)$ can be represented as the middle cohomology group of the complex

$$0 \longrightarrow V \otimes \mathcal{O}_{\mathbb{F}_0}(-1,-1) \xrightarrow{\ \sigma\ } \begin{array}{c} V \otimes \mathcal{O}_{\mathbb{F}_0}(-1,0) \\ \oplus \\ V \otimes \mathcal{O}_{\mathbb{F}_0}(0,-1) \\ \oplus \\ W \otimes \mathcal{O}_{\mathbb{F}_0} \end{array} \xrightarrow{\ \tau\ } V \otimes \mathcal{O}_{\mathbb{F}_0} \longrightarrow 0 \tag{4.113}$$

which is exact at the first and last terms, where

$$\sigma = \begin{pmatrix} z_0 B_z - z_1 \\ w_0 B_w - w_1 \\ z_0 w_0 J \end{pmatrix}, \qquad \tau = \begin{pmatrix} -(w_0 B_w - w_1) & z_0 B_z - z_1 & I \end{pmatrix}. \tag{4.114}$$

The virtual $\widetilde{T}$-equivariant bundle defined by the cohomology of the complex (4.113) gives a representative of the isomorphism class of the universal sheaf $\mathfrak{E}$ in the $\widetilde{T}$-equivariant K-theory group $K^0_{\widetilde{T}}(\mathbb{F}_0 \times \mathfrak{M}_{N,k}(\mathbb{F}_0))$ as

$$\mathfrak{E} = (\mathcal{O}_{\mathbb{F}_0} \boxtimes W) \oplus \left(S^- \ominus S^+\right) \boxtimes V, \tag{4.115}$$

where

$$S^+ = \mathcal{O}_{\mathbb{F}_0}(-1,-1) \oplus \mathcal{O}_{\mathbb{F}_0}, \qquad S^- = \mathcal{O}_{\mathbb{F}_0}(-1,0) \oplus \mathcal{O}_{\mathbb{F}_0}(0,-1) \tag{4.116}$$

are $\widetilde{T}$-equivariant bundles over $\mathbb{F}_0$ which, after tensoring with a line bundle of degree one, restrict to the usual positive/negative chirality spinor bundles over $\mathbb{R}^4 \cong \mathbb{C}^2 \cong \mathbb{F}_0 \setminus \ell_\infty$. In the topologically twisted gauge theory, fermion fields become differential forms and these bundles are identified with the bundles of even/odd holomorphic forms over $\mathbb{F}_0$ [79]. By (4.89), the holomorphic line bundles (4.90) have $\widetilde{T}$-equivariant characters

$$\mathrm{ch}_{\widetilde{T}}(\mathcal{O}_{\mathbb{F}_0}(m_z, m_w)) = t_z^{-m_z} t_w^{-m_w}, \tag{4.117}$$

and consequently,

$$\mathrm{ch}_{\widetilde{T}}(S^+) = 1 + t_z\ t_w, \qquad \mathrm{ch}_{\widetilde{T}}\left(S^-\right) = t_z + t_w. \tag{4.118}$$

Using (4.111) and (4.118), we may thus compute the character of the universal sheaf (4.115) at a fixed point in the instanton moduli space as

$$\begin{aligned}\mathrm{ch}_{\widetilde{T}}(\mathfrak{E})\left(\overrightarrow{\lambda}\right) &= \mathrm{ch}_{\widetilde{T}}(W)\left(\overrightarrow{\lambda}\right) + \left(\mathrm{ch}_{\widetilde{T}}\left(S^-\right) - \mathrm{ch}_{\widetilde{T}}(S^+)\right)\mathrm{ch}_{\widetilde{T}}(V)\left(\overrightarrow{\lambda}\right) \\ &= \sum_{l=1}^{N} e_l \left(1 - (1-t_z)(1-t_w) \sum_{(i,j)\in\lambda_l} t_z^{i-1} t_w^{j-1}\right),\end{aligned} \tag{4.119}$$

which coincides with the standard expression for instantons on $\mathbb{R}^4$ [80].

For completeness, let us record here the $\widetilde{T}$-equivariant Chern character of the tangent bundle $T\mathfrak{M}_{N,k}(\mathbb{F}_0)$ over the instanton moduli space at the torus fixed points. As the only nonvanishing cohomology group of (4.20) is a model of the tangent space, it can be computed as the equivariant index of this complex. For this, let us recall some combinatorial definitions. Let λ be a Young diagram. Define the arm and leg lengths of a box $(i,j) \in \lambda$, respectively, by

$$A_\lambda(i,j) = \varrho_i - j, \qquad L_\lambda(i,j) = \varrho_j^t - i, \tag{4.120}$$

where ϱ_i is the length of the ith column of λ and ϱ_j^t is the length of the jth row of λ. At a fixed point $\vec{\lambda}$, the character can then be expressed after some algebra in terms of the characters of the representation as

$$\mathrm{ch}_{\widetilde{T}}(T\mathfrak{M}_{N,k}(\mathbb{F}_0))\left(\vec{\lambda}\right) = \sum_{l,m=1}^{N} e_l e_m^{-1}\left(\sum_{(i,j)\in\lambda_l} t_z^{-L_{\lambda_l}(i,j)} t_w^{A_{\lambda_m}(i,j)+1} + \sum_{(i,j)\in\lambda_m} t_z^{L_{\lambda_l}(i,j)+1} t_w^{-A_{\lambda_m}(i,j)}\right). \quad (4.121)$$

Again this coincides with the standard result [50, 80, 81]. As in [82], after toric localization one has

$$\mathrm{ch}_{\widetilde{T}}(T\mathfrak{M}_{N,k}(\mathbb{F}_0)) = -\oint_{\mathbb{F}_0\times\mathfrak{M}_{N,k}(\mathbb{F}_0)} \mathrm{ch}_{\widetilde{T}}(\mathfrak{E}) \wedge \mathrm{ch}_{\widetilde{T}}\left(\mathfrak{E}^\vee\right) \wedge \mathrm{Td}_{\widetilde{T}}(\mathbb{F}_0) \quad (4.122)$$

at the $\widetilde{T}$-fixed points $\vec{\lambda}$, up to a universal perturbative contribution (the character of $W\otimes W^*$). This expression formally generalizes to generic toric surfaces [83] and is used to construct vertex gluing rules below. From the top Chern class, one can also straightforwardly extract the equivariant Euler classes

$$e_{\widetilde{T}}(T\mathfrak{M}_{N,k}(\mathbb{F}_0))\left(\vec{\lambda}\right) = \prod_{l,m=1}^{N} n_{l,m}^{\vec{\lambda}}(\epsilon_z,\epsilon_w,\vec{a}), \quad (4.123)$$

where

$$\begin{aligned} n_{l,m}^{\vec{\lambda}}(\epsilon_z,\epsilon_w,\vec{a}) = &\prod_{(i,j)\in\lambda_l}\left(a_m - a_l - L_{\lambda_m}(i,j)\epsilon_z + \left(A_{\lambda_l}(i,j)+1\right)\epsilon_w\right) \\ &\times \prod_{(i',j')\in\lambda_m}\left(a_m - a_l + \left(L_{\lambda_l(i',j')}+1\right)\epsilon_z - A_{\lambda_m}(i',j')\epsilon_w\right). \end{aligned} \quad (4.124)$$

4.9.3. Framed Modules on $\mathbb{F}_0$ from Dirac Modules on $\mathbb{P}^1$

For the case of nontrivial (m_z,m_w)-framings, let

$$\iota(m_z,m_w) : \mathfrak{M}_{N,k}(\mathbb{F}_0) \longrightarrow \mathfrak{M}_{N;(Nm_z,Nm_w);\,k(m_z,m_w)}(\mathbb{F}_0) \quad (4.125)$$

be the isomorphism induced by the map $[\mathcal{E}] \mapsto [\mathcal{E}(m_z,m_w)] = [\mathcal{E}\otimes\mathcal{O}_{\mathbb{F}_0}(m_z,m_w)]$, and let $\mathfrak{E}$ be a universal sheaf on $\mathbb{F}_0\times\mathfrak{M}_{N,k}(\mathbb{F}_0)$. Then the torsion-free sheaf

$$\mathfrak{E}(m_z,m_w) = \left(\mathrm{id}_{\mathbb{F}_0}\times\iota(m_z,m_w)^{-1}\right)^*(\mathfrak{E}) \quad (4.126)$$

is a universal sheaf on $\mathbb{F}_0\times\mathfrak{M}_{N;(Nm_z,N\,m_w);k(m_z,m_w)}(\mathbb{F}_0)$. The equivariant character of $\mathfrak{E}(m_z,m_w)$ at the fixed points of the toric action modifies the decompositions (4.111) to allow for "propagators" which appear along the edges $\mathbb{P}^1$ of the toric diagram $\Delta(C)$ of a generic toric surface C. These modifications of the $\widetilde{T}$-equivariant character of the universal sheaf corresponding to the instanton sheaves (4.96) are determined via shifts by the $\mathbb{C}^*$-equivariant character of modules of solutions to the Dirac equation on $\mathbb{P}^1$. We will now describe these modules explicitly. As we have seen, the bundle $\mathcal{O}_{\mathbb{P}^1}(m)$ of degree $m \in \mathbb{Z}$ over the Riemann sphere $\mathbb{P}^1$ is the crucial ingredient in generating instantons with nontrivial framings. After choosing a hermitean metric on the fibres of $\mathcal{O}_{\mathbb{P}^1}(m)$, it is the holomorphic line bundle underlying the standard Dirac monopole line bundle of topological charge m. The form of the Dirac operator $\not{D}_m$ in the background of the corresponding monopole gauge potential is well known and can be conveniently described in the following way [32].

Given homogeneous coordinates $[z_0, z_1]$ on $\mathbb{P}^1$, let $y = z_1/z_0$ denote stereographic coordinates on the northern hemisphere. Then the twisted Dirac operator in the monopole background of magnetic charge $m \in \mathbb{Z}$ is given by

$$\not{D}_m = \begin{pmatrix} 0 & \not{D}_m^- \\ \not{D}_m^+ & 0 \end{pmatrix}, \tag{4.127}$$

where

$$\begin{aligned} \not{D}_m^+ &= \frac{1}{2}\left[\left(1 + y\overline{y}\right)\frac{\partial}{\partial\overline{y}} - \frac{1}{2}\,(m+1)y\right], \\ \not{D}_m^- &= -\frac{1}{2}\left[\left(1 + y\overline{y}\right)\frac{\partial}{\partial y} + \frac{1}{2}\,(m-1)\overline{y}\right]. \end{aligned} \tag{4.128}$$

These operators act on sections of the chiral/antichiral spinor line bundles associated to the twisted holomorphic spinor bundles $\mathcal{O}_{\mathbb{P}^1}(\pm 1) \otimes \mathcal{O}_{\mathbb{P}^1}(m)$. We are interested in the subspaces of zero modes

$$S_m^{\pm} = \ker \not{D}_m^{\pm}. \tag{4.129}$$

The action of $t \in \mathbb{C}^*$ is implemented by the automorphism F_t of $\mathbb{P}^1$ defined by

$$F_t(y, \overline{y}) = \left(ty, t^{-1}\overline{y}\right). \tag{4.130}$$

The irreducible representations $\underline{T}^i \cong \mathbb{C}$ of $\mathbb{C}^*$ are labelled by their weights $i \in \mathbb{Z}$ and are defined by $z \mapsto t \cdot z = t^i\, z$ for $t \in \mathbb{C}^*$ and $z \in \mathbb{C}$. The corresponding $\mathbb{C}^*$-eigenspace decomposition of the modules (4.129) can be described for all $m \in \mathbb{Z}$ as follows.

Theorem 4.6. *The isotopical decompositions of the spinor modules $S_m^{\pm}$ over $\mathbb{P}^1$, as $\mathbb{C}^*$-modules, are given by*

$$\begin{aligned} S_m^+ &= \bigoplus_{i=1}^{|m|} \underline{T}^{\,i-1}, \qquad S_m^- = \{0\} \quad \textit{for } m < 0, \\ S_m^- &= \bigoplus_{i=1}^{m} \underline{T}^{\,-(i-1)}, \qquad S_m^+ = \{0\} \quad \textit{for } m > 0. \end{aligned} \tag{4.131}$$

Proof. The solutions of the Dirac equation are given by L^2-solutions of the differential equations

$$\not{D}_m^{\pm}\psi_m^{\pm} = 0 \tag{4.132}$$

for the spinors $\psi_m^{\pm} \in \ker \not{D}_m^{\pm}$. The line bundle $\mathcal{O}_{\mathbb{P}^1}(m)$ has holomorphic transition function y^m transforming sections from the northern hemisphere to the southern hemisphere of $\mathbb{P}^1$, which after unitary reduction to a hermitean line bundle becomes $(y/\overline{y})^{m/2}$. By using these transition functions, it is easy to see that the only solutions of the equations (4.132) which are regular on both the northern and southern hemispheres are of the form

$$\begin{aligned} \psi_m^{+} &= \frac{1}{(1+y\,\overline{y})^{(|m|-1)/2}}\sum_{i=1}^{|m|}\xi_i y^{i-1}, \qquad \psi_m^{-} = 0 \quad \text{for } m<0, \\ \psi_m^{-} &= \frac{1}{(1+y\,\overline{y})^{(m-1)/2}}\sum_{i=1}^{m}\widetilde{\xi}_i \overline{y}^{\,i-1}, \qquad \psi_m^{+} = 0 \quad \text{for } m>0, \end{aligned} \tag{4.133}$$

with constant coefficients $\xi_i, \widetilde{\xi}_i \in \mathbb{C}$. The result now follows from (4.130). □

We can use Theorem 4.6 to compute the $\mathbb{C}^*$-equivariant characters of the spinor modules (4.129). One finds

$$\mathrm{ch}_{\mathbb{C}^*}\left(S_m^{-}\right) = \sum_{i=1}^{m} t^{-(i-1)} = \frac{1-t^{-m}}{1-t^{-1}} \tag{4.134}$$

for $m > 0$, while

$$\mathrm{ch}_{\mathbb{C}^*}\left(S_m^{+}\right) = \sum_{i=1}^{|m|} t^{i-1} = \frac{1-t^{|m|}}{1-t} \tag{4.135}$$

for $m < 0$. In the nonequivariant limit $t \to 1$, the characters (4.134) and (4.135) reproduce the known index of the Dirac operator, $\mathrm{index}(\not{D}_m) = -m$, in the monopole background. These characters shift the character (4.119). For example, with $m_z, m_w > 0$ and $N = 1$, one has

$$\mathrm{ch}_{\widetilde{T}}(\mathfrak{E}(m_z, m_w))(\lambda) = \mathrm{ch}_{\widetilde{T}}(\mathfrak{E})(\lambda) + (1-t_z)\ \mathrm{ch}_{\mathbb{C}_z^*}\left(S_{m_z}^{-}\right) + (1-t_w)\mathrm{ch}_{\mathbb{C}_w^*}\left(S_{m_w}^{-}\right). \tag{4.136}$$

This completes the derivation of the gluing rules of [54, Section 4.3].

5. D2-Brane Gauge Theory and Gromov-Witten Invariants

In this final section, we study the reduction of the four-dimensional gauge theories of Section 4 on local curves and examine in detail the example of local $\mathbb{P}^1$ which has been extensively described from the point of view of Vafa-Witten theory. We begin with a somewhat heuristic description of how these two-dimensional supersymmetric gauge theories are induced. Then we proceed to a more formal topological field theory formalism which systematically computes topological string amplitudes and Gromov-Witten invariants. Finally, we briefly address wall-crossing issues once again from a physical standpoint.

5.1. q-Deformed Two-Dimensional Yang-Mills Theory

One of the simplest classes of examples are the noncompact local Calabi-Yau threefolds which are fibred over curves. If $\Sigma_g \to X$ is a holomorphically embedded curve of genus g in a Calabi-Yau threefold, then the holomorphic tangent bundle restricts to Σ_g as $T\Sigma_g \oplus \mathcal{N}_{\Sigma_g}$, where the normal bundle $\mathcal{N}_{\Sigma_g}$ is a holomorphic bundle of rank two over Σ_g. By the Calabi-Yau condition $c_1(X) = 0$, one has $c_1(\mathcal{N}_{\Sigma_g}) = -\chi(\Sigma_g) = 2g - 2$. Thus in a neighbourhood of Σ_g, the manifold X looks like the total space of a holomorphic bundle $\mathcal{N} \to \Sigma_g$ of rank two with $c_1(\mathcal{N}) = 2g - 2$. Hence we consider the total space of the bundle

$$X = X_p = \mathcal{O}_{\Sigma_g}(p + 2g - 2) \oplus \mathcal{O}_{\Sigma_g}(-p) \tag{5.1}$$

over Σ_g. Since $\mathcal{O}_{\Sigma_g}(p)$ is a holomorphic line bundle over Σ_g of degree p, one has

$$c_1(X_p) = (p + 2g - 2) + (-p) = 2g - 2 \tag{5.2}$$

as required. For genus $g = 0$, this is just the example of local $\mathbb{P}^1$ which was introduced in Section 4.3.

In our applications to black hole microstate counting, we count bound states of D4-D2-D0 branes in X_p with N D4-branes wrapping the four-cycle C_p which is the total space of the line bundle $\mathcal{O}_{\Sigma_g}(-p) \to \Sigma_g$, together with D2-branes wrapping the base Riemann surface Σ_g (embedded in C and X_p as the zero section). One can then localize the path integral of the $\mathcal{N} = 4$ topological gauge theory on C_p to the $\mathbb{C}^*$-invariant modes along the fibre of $\mathcal{O}_{\Sigma_g}(-p)$. The result is an effective $U(N)$ gauge theory on Σ_g whose action is given by [84]

$$S = \frac{1}{g_s}\int_{\Sigma_g} \operatorname{Tr}(\Phi F) + \frac{\vartheta}{g_s}\int_{\Sigma_g} \operatorname{Tr}(\Phi k) - \frac{p}{2g_s}\int_{\Sigma_g} \operatorname{Tr}\left(\Phi^2 k\right), \tag{5.3}$$

where $\vartheta = g_s\, \phi^2/2\pi$ and the last term is a mass deformation which originates in four dimensions due to the nontriviality of the bundle $\mathcal{O}_{\Sigma_g}(-p)$ [85]. The scalar field Φ is given by the holonomy of the four-dimensional gauge connection at infinity

$$\Phi(z) = \int_{\mathcal{O}_{\Sigma_g}(-p)_z} F(z) = \oint_{S^1_{z,|u|=\infty}} A, \tag{5.4}$$

where $z \in \Sigma_g$ and $u \in \mathcal{O}_{\Sigma_g}(-p)$.

The action (5.3) is just the BF-theory representation of two-dimensional Yang-Mills theory, with the usual F^2-term coming from performing the gaussian integral over Φ in the functional integral. If Φ were arbitrary, then we could use diagonalization techniques to reduce the partition function to the usual heat kernel expansion of two-dimensional Yang-Mills theory. However, the identification of Φ as a holonomy means that it should be treated as a periodic field, and the change from a Lie algebra-valued variable to a group-like variable affects the Jacobian of the path integral measure that arises from diagonalization. This modifies the usual Migdal expansion of two-dimensional gauge theory to give the partition function

$$Z_{U(N)}^{C_p}(q,Q) = \sum_R \dim_q(R)^{2-2g} q^{(p/2)C_2(R)} Q^{C_1(R)}, \tag{5.5}$$

where $q = e^{-g_s}$ as before and $Q = e^{i\vartheta}$. The sum runs over unitary irreducible representations $R = (R_1, \ldots, R_N)$ of the gauge group $U(N)$ with weights R_i, first and second Casimir invariants $C_1(R) = \sum_i R_i$ and $C_2(R) = \sum_i R_i^2$, and quantum dimension

$$\dim_q(R) = \prod_{1 \le i < j \le N} \frac{\left[R_i - R_j + j - i\right]_q}{\left[j - i\right]_q}, \tag{5.6}$$

where the q-number is defined by

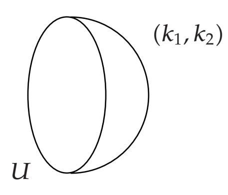
$$[n]_q := \frac{q^{n/2} - q^{-n/2}}{q^{1/2} - q^{-1/2}} \tag{5.7}$$

with $[n]_q = n + O(q-1)$ for $q \to 1$. This reduced two-dimensional gauge theory is a q-deformation of ordinary Yang-Mills theory.

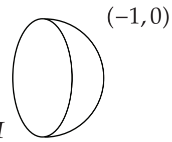
5.2. Topological Field Theory

We will now give a more precise definition of this gauge theory using techniques of two-dimensional topological field theory. This formalism can then be used to compute topological string amplitudes. The idea is that the cutting and pasting of base Riemann surfaces is equivalent to the cutting and pasting of the corresponding local Calabi-Yau threefolds, by either adding or cancelling off D-branes corresponding to the boundaries in the topological vertex formalism [84, 86]. The operations of gluing manifolds satisfy all axioms of a two-dimensional topological field theory. By computing the open string topological A-model amplitudes on a few Calabi-Yau manifolds, we then get *all* others by gluing. In the following we focus on the genus zero case $\Sigma_0 = \mathbb{P}^1$ for simplicity and illustrative purposes.

When the base of the fibration is the sphere, we will only need the basic *Calabi-Yau cap amplitude* corresponding to a disk D, represented symbolically as



where the *levels* $k_1 = e(\mathcal{L}_1)$ and $k_2 = e(\mathcal{L}_2)$ label the degrees in $H^2(D, \partial D)$ of a pair of line bundles $\mathcal{L}_1 \oplus \mathcal{L}_2 \to D$, and the unitary matrix U labels the holonomy of a gauge connection on $\partial D \cong S^1$. Under concatenation the levels add. The basic cap amplitude is defined by



$$Z\left(U \qquad\qquad \right) = \sum_R d_q(R) q^{-k_R/4} \operatorname{ch}_R(U),$$

where $q = \mathrm{e}^{-g_s}$ as before, the sum over R runs through the space of symmetric functions (Young tableaux), that is, irreducible representations of $SU(\infty)$, and $\operatorname{ch}_R(U)$ is the character of the holonomy U in the representation R. The quantity

$$d_q(R) := \prod_{(i,j)\in R} \frac{1}{[h(i,j)]_q} \tag{5.8}$$

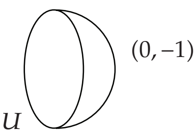
is the quantum dimension of the symmetric group representation corresponding to the Young diagram of R, and

$$k_R = 2 \sum_{(i,j)\in R} (i-j). \tag{5.9}$$

The quantity

$$d_q(R)\; q^{-k_R/4} =: C_{R,\emptyset,\emptyset}(q) = W_{R,\emptyset}(q) \tag{5.10}$$

is the topological vertex amplitude for the A-model topological string theory with a single stack of D-branes on the noncompact Calabi-Yau threefold $X = \mathbb{C}^3$. Similarly, define the cap amplitude



$$Z\left(U \qquad\qquad \right) = \sum_R d_q(R)\; q^{k_R/4} \operatorname{ch}_R(U).$$

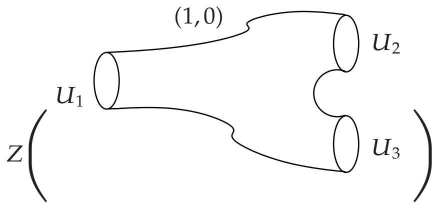
For the gluing rules, we sew two open Riemann surfaces Σ_L and Σ_R together along their common boundary to get the Riemann surface $\Sigma_{L\cup R}$. For this, the orientations of the corresponding boundary circles must be opposite. Reversing the orientation acts on the boundary holonomy as $U \to U^{-1}$, and hence the gluing rules read

$$Z(\Sigma_{L\cup R}) = \int_{SU(\infty)} \mathrm{d}U\, Z(\Sigma_L)(U) Z(\Sigma_R)\left(U^{-1}\right), \tag{5.11}$$

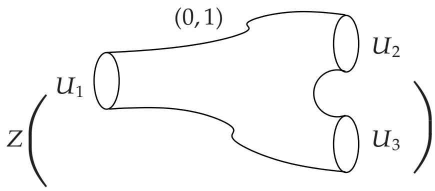
where $\mathrm{d}U$ is the bi-invariant Haar measure and the integrals can be evaluated explicitly by using the orthogonality relations for characters

$$\int_{SU(\infty)} \mathrm{d}U \operatorname{ch}_{R_1}(U) \operatorname{ch}_{R_2}\left(U^{-1}\right) = \delta_{R_1,R_2}. \tag{5.12}$$

We will also need the Calabi-Yau trinion (or "pants") amplitudes



$$= \sum_R \frac{1}{d_q(R)} q^{k_R/4} \prod_{i=1}^{3} \mathrm{ch}_R(U_i),$$



$$= \sum_R \frac{1}{d_q(R)} q^{-k_R/4} \prod_{i=1}^{3} \mathrm{ch}_R(U_i).$$

Let us point out two particularly noteworthy aspects of this construction thus far. Firstly, the construction defines a functor of tensor categories

$$Z : S_{\mathcal{L}_1,\mathcal{L}_2} \longrightarrow \mathrm{Rep}, \tag{5.13}$$

where Rep is the representation category of $SU(\infty)$ and $S_{\mathcal{L}_1,\mathcal{L}_2}$ is the geometric tensor category defined as follows. The objects of $S_{\mathcal{L}_1,\mathcal{L}_2}$ are compact oriented one-manifold Y, that is, disjoint unions of oriented circles. A morphism $Y_1 \rightarrow Y_2$ between two objects of $S_{\mathcal{L}_1,\mathcal{L}_2}$ is a triple (W, Y_1, Y_2), where W is an oriented cobordism between Y_1 and Y_2, that is, a smooth oriented two-manifold with boundary $\partial W = Y_1 \coprod (-Y_2)$, and the complex line bundles $\mathcal{L}_1, \mathcal{L}_2$ over W are trivialized on ∂W. This is analogous to the Baum-Douglas description of D-branes in K-homology [87]. There is a natural notion of equivalence provided by boundary preserving, oriented diffeomorphisms $f : W \rightarrow W'$ with $\mathcal{L}_i \cong f^*\mathcal{L}'_i$, $i = 1, 2$. Composition of morphisms is given by concatenation of cobordism and gluing of bundles along the concatenation using the trivializations. For a connected cobordism W, we can label the isomorphism classes of a pair of line bundles $(\mathcal{L}_1, \mathcal{L}_2)$ by the *levels* (k_1, k_2), where $k_i = e(\mathcal{L}_i) \in H^2(W, \partial W)$. Under concatenation, these levels add.

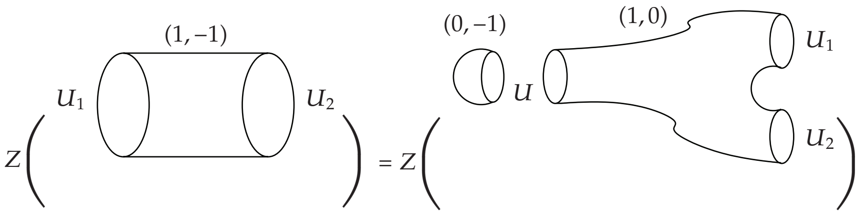
Secondly, the $\mathbb{C}^*$-action on $\mathbb{P}^1$ lifts to an action of the torus $T = (\mathbb{C}^*)^2$ on X_p, via the natural scaling action on the fibres. The Gromov-Witten invariants in this case are *defined* via the virtual localization formula as a residue integral over the T-fixed point locus, along the lines explained in Section 3.5. A stable map to X_p which is T-invariant factors through the zero section. It follows that there is a natural isomorphism of moduli spaces $\mathfrak{M}_g(X_p, \beta)^T = \mathfrak{M}_g(\mathbb{P}^1, \beta)$, with $\beta = d \in \mathbb{Z}$, and consequently

$$\left[\mathfrak{M}_g(X_p, d)^T\right]^{\mathrm{vir}} = \left[\mathfrak{M}_g\left(\mathbb{P}^1, d\right)\right]^{\mathrm{vir}}. \tag{5.14}$$

This equality of virtual fundamental cycles implies that the invariants constructed by integrating over each moduli space coincide. It means that topological string theory on X_p

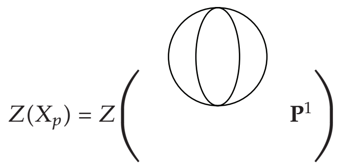
is equivalent to a field theory on $\mathbb{P}^1$, which is just the reduction we argued in Section 5.1. In this case the Gromov-Witten invariants of X_p correspond to degree d Hurwitz numbers of $\mathbb{P}^1$.

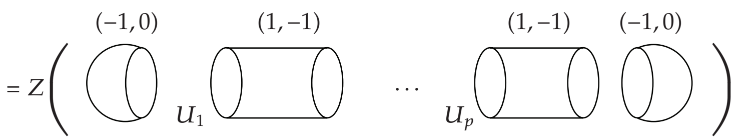
Returning to our computations, we get the annulus (or tube) amplitude by contracting the cap and trinion amplitudes to get



$$= \sum_R q^{k_R/2}\, \mathrm{ch}_R(U_1)\, \mathrm{ch}_R(U_2).$$

Finally, we compute the amplitude of the fibration $X_p \to \mathbb{P}^1$ by using (5.14) and the gluing rules. In order to get the appropriate bundle degrees $(p-2,-p)$, we glue p tubes between two caps to get





(5.15)

$$= \sum_R d_q(R)^2 q^{(p-1)k_R/2}.$$

For $p = 1$, the threefold $X_1 \cong K_{\mathbb{P}^1}^{1/2} \oplus K_{\mathbb{P}^1}^{1/2}$ is the resolved conifold, and this formula is a q-deformation of the classical Hurwitz formula counting unramified covers of the Riemann sphere $\mathbb{P}^1$.

In fact, the quantity $d_q(R)$ is analogously the $N \to \infty$ limit of the quantum dimension $\dim_q(R)$ of the $U(N)$ representation with the same Young tableau. This follows by Schur-Weyl reciprocity which writes $SU(N)$ representations in terms of representations of the symmetric group S_N. Let R be a representation corresponding to a Young diagram λ with row lengths λ_i, $i = 1,\dots,d(\lambda)$, where $\lambda_1 \geq \lambda_2 \geq \cdots$, and $d(\lambda)$ is the number of rows in λ. Set $\mu = q^N$, and define another q-number by

$$[n]_\mu = \frac{\mu^{1/2}\, q^{n/2} - \mu^{-1/2}\, q^{-n/2}}{q^{1/2} - q^{-1/2}}. \tag{5.16}$$

Then the quantum dimension of R can be written as

$$\dim_q(R) = \prod_{1\le i<j\le d(\lambda)} \frac{\left[\lambda_i - \lambda_j + j - i\right]_q}{\left[j-i\right]_q} \prod_{i=1}^{d(\lambda)} \frac{\prod_{v=1-i}^{\lambda_i - i}[v]_\mu}{\prod_{v=1}^{\lambda_i}[v - i + d(\lambda)]_q}. \tag{5.17}$$

This is a Laurent polynomial in $\mu^{\pm 1/2}$ whose coefficients are rational functions of $q^{\pm 1/2}$. The leading power of μ is $|\lambda|/2$, and the coefficient of this power is the rational function of $q^{\pm 1/2}$ given by

$$q^{-k_R/4} \prod_{1\le i<j\le d(\lambda)} \frac{\left[\lambda_i - \lambda_j + j - i\right]_q}{\left[j-i\right]_q} \prod_{i=1}^{d(\lambda)} \prod_{v=1}^{\lambda_i} \frac{1}{[v - i + d(\lambda)]_q} = d_q(R) q^{-k_R/4}. \tag{5.18}$$

As before, this is just the topological vertex amplitude $C_{R,\emptyset,\emptyset}(q)$. The limit $q \to 1$ gives the ordinary dimension of the representation R.

Generalizing the generating function (5.15) thereby gives

$$Z_{q\mathrm{YM}}\left(\mathbb{P}^1\right) = \sum_R \dim_q(R)^2 q^{(p/2)C_2(R)}, \tag{5.19}$$

where now the sum runs over irreducible representations R of $U(N)$ and $C_2(R) = k_R + N(R)$ is the second Casimir invariant of R. By the attractor mechanism, the Kähler modulus of $\mathbb{P}^1$ is given by $t = (N/2)\ (p-2)\ g_s$. In the limit $q \to 1$, this two-dimensional field theory coincides with ordinary Yang-Mills theory on the sphere $\mathbb{P}^1$. Note that this is a weak-coupling limit $g_s \to 0$, with $p g_s = g_{\mathrm{YM}}^2\ A$ and A the area of $\mathbb{P}^1$. Using similar constructions as above, it is possible to formulate this gauge theory directly as a two-dimensional topological field theory, without any reference to the extrinsic bundle structure of the local threefold X_p. Explicit computations of the corresponding Gromov-Witten invariants of X_p can be found in [9–11].

5.3. Wall-Crossing Formulas

Finally, let us briefly comment on the relationship to black hole entropy counting. Reinstating the ϑ-angle as in (5.5), the attractor mechanism fixes the Kähler modulus of $\Sigma_0 = \mathbb{P}^1$ as

$$t = 2\pi \mathrm{i} \frac{X^1}{X^0} = (p-2)\frac{N g_s}{2} - \mathrm{i}\vartheta. \tag{5.20}$$

The large N limit of the partition function (5.5) should possess this modulus. The modulus squared structure anticipated by the OSV relation (2.25) is given in this limit by [84]

$$Z_{q\mathrm{YM}}\left(\mathbb{P}^1\right) = \sum_{r\in\mathbb{Z}} \sum_{R_1,R_2} Z_{R_1,R_2}^{q\mathrm{YM}+}(t + p g_s r)\ Z_{R_1,R_2}^{q\mathrm{YM}-}\left(\bar{t} + p g_s r\right), \tag{5.21}$$

where the sum over r runs through $U(1)$ charges in the local decomposition $U(N) \sim U(1) \times SU(N)$ of the gauge group, the second sum runs through pairs of irreducible

$SU(N)$ representations R_1, R_2, and $Z^{q\mathrm{YM}+}_{R_1,R_2}(t)$ is the perturbative topological string amplitude on X_p with two stacks of D-branes in the fibre. The conjugate amplitude is $Z^{q\mathrm{YM}-}_{R_1,R_2}(t) = (-1)^{C_1(R_1)+C_1(R_2)}\ Z^{q\mathrm{YM}+}_{R_1^t,R_2^t}(t)$. Thus in this case the OSV relation (2.25) is modified to the symbolic form

$$Z_{\mathrm{BH}} = \sum_{r\in\mathbb{Z}} \sum_{b} \left| Z^{(b,r)}_{\mathrm{top}} \right|^2, \tag{5.22}$$

where the first sum runs over Ramond-Ramond fluxes r through the base Riemann surface Σ_g while the second sum is over fibre D-branes which carry additional moduli $\widehat{t}$ measuring their distances from the base Σ_g along the noncompact fibre directions.

Further analysis of the relationships with $\mathcal{N} = 4$ Yang-Mills theory on C_p is studied in [9, 10, 12, 13, 57], where it is observed that the instanton expansions of the two-dimensional and four-dimensional gauge theory partition functions differ by perturbative contributions. These extra factors arise from the noncompactness of the surface C_p. When C_p is the resolution of an $A_{p,n}$ singularity, its boundary is the three-dimensional Lens space $L(p,n) = S^3/\Gamma_{(p,n)}$. The perturbative factors can then be identified as the partition function of Chern-Simons gauge theory on the boundary and arise as a consequence of the fact that the two-dimensional gauge theory implicitly integrates over all boundary conditions. Once these boundary contributions are stripped, the enumeration of instantons in the two-dimensional and four-dimensional gauge theories coincide [13].

Acknowledgments

It is a pleasure to thank M. Cirafici, E. Gasparim, A.-K. Kashani-Poor, A. Maciocia, and A. Sinkovics for enjoyable discussions and collaborations on some of the results presented here. This work was supported in part by Grant ST/G000514/1 "String Theory Scotland" from the UK Science and Technology Facilities Council.

References

[1] H. Ooguri, A. Strominger, and C. Vafa, "Black hole attractors and the topological string," *Physical Review D*, vol. 70, no. 10, Article ID 106007, 13 pages, 2004.

[2] M. R. Douglas and N. A. Nekrasov, "Noncommutative field theory," *Reviews of Modern Physics*, vol. 73, no. 4, pp. 977–1029, 2001.

[3] A. Konechny and A. Schwarz, "Introduction to M(atrix) theory and noncommutative geometry," *Physics Reports*, vol. 360, no. 5-6, pp. 353–465, 2002.

[4] R. J. Szabo, "Quantum field theory on noncommutative spaces," *Physics Reports*, vol. 378, no. 4, pp. 207–299, 2003.

[5] N. A. Nekrasov, "Seiberg-Witten prepotential from instanton counting," *Advances in Theoretical and Mathematical Physics*, vol. 7, no. 5, pp. 831–864, 2003.

[6] A. Okounkov, N. Reshetikhin, and C. Vafa, "Quantum Calabi-Yau and classical crystals," in *The Unity of Mathematics*, vol. 244 of *Progress in Mathematics*, pp. 597–618, Birkhäuser, Boston, Mass, USA, 2006.

[7] M. Cirafici, A. Sinkovics, and R. J. Szabo, "Cohomological gauge theory, quiver matrix models and Donaldson-Thomas theory," *Nuclear Physics B*, vol. 809, no. 3, pp. 452–518, 2009.

[8] M. Cirafici, A. Sinkovics, and R. J. Szabo, "Instantons and Donaldson-Thomas invariants," *Fortschritte der Physik*, vol. 56, no. 7–9, pp. 849–855, 2008.

[9] N. Caporaso, M. Cirafici, L. Griguolo, S. Pasquetti, D. Seminara, and R. J. Szabo, "Topological strings and large N phase transitions. I. Nonchiral expansion of q-deformed Yang-Mills theory," *Journal of High Energy Physics*, vol. 2006, no. 1, article 35, 2006.

[10] N. Caporaso, M. Cirafici, L. Griguolo, S. Pasquetti, D. Seminara, and J. Szabo, "Black-holes, topological strings and large N phase transitions," *Journal of Physics: Conference Series*, vol. 33, no. 1, pp. 13–25, 2006.

[11] N. Caporaso, M. Cirafici, L. Griguolo, S. Pasquetti, D. Seminara, and R. J. Szabo, "Topological strings and large N phase transitions. II. Chiral expansion of q-deformed Yang-Mills theory," *Journal of High Energy Physics*, vol. 2006, no. 1, article 36, 2006.

[12] N. Caporaso, M. Cirafici, L. Griguolo, S. Pasquetti, D. Seminara, and R. J. Szabo, "Topological strings, two-dimensional Yang-Mills theory and Chern-Simons theory on torus bundles," *Advances in Theoretical and Mathematical Physics*, vol. 12, no. 5, pp. 981–1058, 2008.

[13] L. Griguolo, D. Seminara, R. J. Szabo, and A. Tanzini, "Black holes, instanton counting on toric singularities and q-deformed two-dimensional Yang-Mills theory," *Nuclear Physics B*, vol. 772, no. 1-2, pp. 1–24, 2007.

[14] M. Mariño, "Chern-Simons theory and topological strings," *Reviews of Modern Physics*, vol. 77, no. 2, pp. 675–720, 2005.

[15] M. Mariño, *Chern-Simons Theory, Matrix Models and Topological Strings*, vol. 131 of *International Series of Monographs on Physics*, Oxford University Press, Oxford, UK, 2005.

[16] M. Vonk, "A mini-course on topological strings," http://arxiv.org/abs/hep-th/0504147.

[17] W. Fulton, *Introduction to Toric Varieties*, vol. 131 of *Annals of Mathematics Studies*, Princeton University Press, Princeton, NJ, USA, 1993.

[18] D. A. Cox, "Toric varieties and toric resolutions," in *Resolution of Singularities*, vol. 181 of *Progress in Mathematics*, pp. 259–284, Birkhäuser, Basel, Switzerland, 2000.

[19] D. Cox, "What is a toric variety?" in *Topics in Algebraic Geometry and Geometric Modeling*, vol. 334 of *Contemporary Mathematics*, pp. 203–223, American Mathematical Society, Providence, RI, USA, 2003.

[20] D. Huybrechts and M. Lehn, *The Geometry of Moduli Spaces of Sheaves*, Aspects of Mathematics, E31, Friedrich Vieweg & Sohn, Braunschweig, Germany, 1997.

[21] P. S. Aspinwall, "D-branes on Calabi-Yau manifolds," in *Theoretical Advanced Study Institute in Elementary Particle Physics (TASI 2003): Recent Trends in String Theory*, J. M. Maldacena, Ed., pp. 1–152, World Scientific, Hackensack, NJ, USA, 2005.

[22] T. Graber and R. Pandharipande, "Localization of virtual classes," *Inventiones Mathematicae*, vol. 135, no. 2, pp. 487–518, 1999.

[23] H. W. Braden and N. A. Nekrasov, "Space-time foam from non-commutative instantons," *Communications in Mathematical Physics*, vol. 249, no. 3, pp. 431–448, 2004.

[24] A. Iqbal, C. Vafa, N. Nekrasov, and A. Okounkov, "Quantum foam and topological strings," *Journal of High Energy Physics*, vol. 2008, no. 4, article 11, 2008.

[25] L. Baulieu, H. Kanno, and I. M. Singer, "Special quantum field theories in eight and other dimensions," *Communications in Mathematical Physics*, vol. 194, no. 1, pp. 149–175, 1998.

[26] C. Hofman and J.-S. Park, "Cohomological Yang-Mills theories on Kähler 3-folds," *Nuclear Physics B*, vol. 600, no. 1, pp. 133–162, 2001.

[27] D. Maulik, N. Nekrasov, A. Okounkov, and R. Pandharipande, "Gromov-Witten theory and Donaldson-Thomas theory. I," *Compositio Mathematica*, vol. 142, no. 5, pp. 1263–1285, 2006.

[28] M. Mihailescu, I. Y. Park, and T. A. Tran, "D-branes as solitons of an $\mathcal{N} = 1$, $D = 10$ noncommutative gauge theory," *Physical Review D*, vol. 64, no. 4, Article ID 046006, 6 pages, 2001.

[29] E. Witten, "BPS bound states of D0-D6 and D0-D8 systems in a B-field," *Journal of High Energy Physics*, vol. 2002, no. 4, article 12, 2002.

[30] P. Kraus and M. Shigemori, "Non-commutative instantons and the Seiberg-Witten map," *Journal of High Energy Physics*, vol. 2002, no. 6, article 34, 2002.

[31] N. A. Nekrasov, "Lectures on open strings, and noncommutative gauge fields," in *Unity from Duality: Gravity, Gauge Theory and Strings*, C. Bachas, A. Bilal, M. R. Douglas, N. A. Nekrasov, and F. David, Eds., NATO Advanced Study Institute, pp. 477–495, EDP Sciences, Les Ulism, France, 2002.

[32] A. D. Popov and R. J. Szabo, "Quiver gauge theory of non-abelian vortices and noncommutative instantons in higher dimensions," *Journal of Mathematical Physics*, vol. 47, no. 1, Article ID 012306, 42 pages, 2006.

[33] D. Maulik, N. Nekrasov, A. Okounkov, and R. Pandharipande, "Gromov-Witten theory and Donaldson-Thomas theory. II," *Compositio Mathematica*, vol. 142, no. 5, pp. 1286–1304, 2006.

[34] D. Maulik, A. Oblomkov, A. Okounkov, and R. Pandharipande, "Gromov-Witten/Donaldson-Thomas correspondence for toric 3-folds," http://arxiv.org/abs/0809.3976.
[35] R. Gopakumar and C. Vafa, "M-theory and topological strings I," http://arxiv.org/abs/hep-th/9809187.
[36] R. Gopakumar and C. Vafa, "M-theory and topological strings II," http://arxiv.org/abs/hep-th/9812127.
[37] D. L. Jafferis and G. W. Moore, "Wall-crossing in local Calabi-Yau manifolds," http://arxiv4.library.cornell.edu/abs/0810.4909.
[38] W.-y. Chuang and D. L. Jafferis, "Wall crossing of BPS states on the conifold from Seiberg duality and pyramid partitions," *Communications in Mathematical Physics*, vol. 292, no. 1, pp. 285–301, 2009.
[39] M. Aganagic, H. Ooguri, C. Vafa, and M. Yamazaki, "Wall-crossing and M-theory," http://arxiv.org/abs/0908.1194.
[40] K. Behrend, "Donaldson-Thomas type invariants via microlocal geometry," *Annals of Mathematics*, vol. 170, no. 3, pp. 1307–1338, 2009.
[41] D. Joyce, "Configurations in abelian categories. IV. Invariants and changing stability conditions," *Advances in Mathematics*, vol. 217, no. 1, pp. 125–204, 2008.
[42] M. Kontsevich and Y. Soibelman, "Stability structures, motivic Donaldson-Thomas invariants and cluster transformations," http://arxiv.org/abs/0811.2435.
[43] R. Pandharipande and R. P. Thomas, "Curve counting via stable pairs in the derived category," *Inventiones Mathematicae*, vol. 178, no. 2, pp. 407–447, 2009.
[44] R. Pandharipande and R. P. Thomas, "Stable pairs and BPS invariants," *Journal of the American Mathematical Society*, vol. 23, no. 1, pp. 267–297, 2010.
[45] L. Cirio, G. Landi, and R. J. Szabo, "Algebraic deformations of toric varieties I. General constructions," http://arxiv.org/abs/1001.1242.
[46] C. Vafa and E. Witten, "A strong coupling test of S-duality," *Nuclear Physics B*, vol. 431, no. 1-2, pp. 3–77, 1994.
[47] J. M. F. Labastida and C. Lozano, "Mathai-Quillen formulation of twisted $\mathcal{N} = 4$ supersymmetric gauge theories in four dimensions," *Nuclear Physics B*, vol. 502, no. 3, pp. 741–790, 1997.
[48] N. Nekrasov and A. Schwarz, "Instantons on noncommutative $\mathbb{R}^4$, and (2,0) superconformal six-dimensional theory," *Communications in Mathematical Physics*, vol. 198, no. 3, pp. 689–703, 1998.
[49] H. Nakajima, *Lectures on Hilbert Schemes of Points on Surfaces*, vol. 18 of *University Lecture Series*, American Mathematical Society, Providence, RI, USA, 1999.
[50] U. Bruzzo, F. Fucito, J. F. Morales, and A. Tanzini, "Multi-instanton calculus and equivariant cohomology," *Journal of High Energy Physics*, vol. 2003, no. 5, article 54, 2003.
[51] G. Gonzalez-Sprinberg and J.-L. Verdier, "Construction géométrique de la correspondance de McKay," *Annales Scientifiques de l'École Normale Supérieure*, vol. 16, no. 3, pp. 409–449, 1983.
[52] P. B. Kronheimer and H. Nakajima, "Yang-Mills instantons on ALE gravitational instantons," *Mathematische Annalen*, vol. 288, no. 2, pp. 263–307, 1990.
[53] G. Ellingsrud and S. A. Strømme, "On the homology of the Hilbert scheme of points in the plane," *Inventiones Mathematicae*, vol. 87, no. 2, pp. 343–352, 1987.
[54] M. Cirafici, A.-K. Kashani-Poor, and R. J. Szabo, "Crystal melting on toric surfaces," http://arxiv.org/abs/0912.0737.
[55] H. Nakajima, "Sheaves on ALE spaces and quiver varieties," *Moscow Mathematical Journal*, vol. 7, no. 4, pp. 699–722, 2007.
[56] F. Fucito, J. F. Morales, and R. Poghossian, "Instanton on toric singularities and black hole countings," *Journal of High Energy Physics*, vol. 2006, no. 12, article 73, 2006.
[57] G. Bonelli and A. Tanzini, "Topological gauge theories on local spaces and black hole entropy countings," *Advances in Theoretical and Mathematical Physics*, vol. 12, no. 6, pp. 1429–1446, 2008.
[58] J. Maldacena, A. Strominger, and E. Witten, "Black hole entropy in M-theory," *The Journal of High Energy Physics*, vol. 1997, no. 12, article 2, 1997.
[59] L. F. Alday, D. Gaiotto, and Y. Tachikawa, "Liouville correlation functions from four-dimensional gauge theories," *Letters in Mathematical Physics*, vol. 91, no. 2, pp. 167–197, 2010.
[60] N. Wyllard, "$A_{\mathcal{N}-1}$ conformal Toda field theory correlation functions from conformal $\mathcal{N} = 1$ SU($\mathcal{N}$) quiver gauge theories," *Journal of High Energy Physics*, vol. 2009, no. 11, article 2, 2009.
[61] R. Dijkgraaf, L. Hollands, P. Sułkowski, and C. Vafa, "Supersymmetric gauge theories, intersecting branes and free fermions," *Journal of High Energy Physics*, vol. 2008, no. 2, article 106, 2008.
[62] R. Dijkgraaf and P. Sułkowski, "Instantons on ALE spaces and orbifold partitions," *Journal of High Energy Physics*, vol. 2008, no. 3, article 13, 2008.

[63] F. Denef and G. W. Moore, "Split states, entropy enigmas, holes and halos," http://arxiv.org/abs/hep-th/0702146.

[64] D.-E. Diaconescu and G. W. Moore, "Crossing the wall: Branes vs. bundles," http://arxiv.org/abs/0706.3193.

[65] J. de Boer, F. Denef, S. El-Showk, I. Messamah, and D. Van den Bleeken, "Black hole bound states in $AdS_3 \times S^2$," *Journal of High Energy Physics*, vol. 2008, no. 11, article 50, 2008.

[66] E. Andriyash and G. W. Moore, "Ample D4-D2-D0 decay," http://arxiv.org/abs/0806.4960.

[67] D. Gaiotto, G. W. Moore, and A. Neitzke, "Wall-crossing, Hitchin systems, and the WKB approximation," http://arxiv.org/abs/0907.3987.

[68] D. Huybrechts and M. Lehn, "Stable pairs on curves and surfaces," *Journal of Algebraic Geometry*, vol. 4, no. 1, pp. 67–104, 1995.

[69] E. Gasparim and C.-C. M. Liu, "The Nekrasov conjecture for toric surfaces," *Communications in Mathematical Physics*, vol. 293, no. 3, pp. 661–700, 2009.

[70] H. Nakajima, "Moduli spaces of anti-self-dual connections on ALE gravitational instantons," *Inventiones Mathematicae*, vol. 102, no. 2, pp. 267–303, 1990.

[71] H. Nakajima, "Instantons on ALE spaces, quiver varieties, and Kac-Moody algebras," *Duke Mathematical Journal*, vol. 76, no. 2, pp. 365–416, 1994.

[72] U. Bruzzo, R. Poghossian, and A. Tanzini, "Poincaré polynomial of moduli spaces of framed sheaves on (stacky) Hirzebruch surfaces," http://arxiv.org/abs/0909.1458.

[73] E. Gasparim, T. Köppe, and P. Majumdar, "Local holomorphic Euler characteristic and instanton decay," *Pure and Applied Mathematics Quarterly*, vol. 4, no. 2, pp. 363–381, 2008.

[74] M. Maruyama, "Moduli of stable sheaves. II," *Journal of Mathematics of Kyoto University*, vol. 18, no. 3, pp. 557–614, 1978.

[75] M. Maruyama, "On boundedness of families of torsion free sheaves," *Journal of Mathematics of Kyoto University*, vol. 21, no. 4, pp. 673–701, 1981.

[76] D. Gieseker, "On the moduli of vector bundles on an algebraic surface," *Annals of Mathematics*, vol. 106, no. 1, pp. 45–60, 1977.

[77] A. Kuznetsov, "Quiver varieties and Hilbert schemes," *Moscow Mathematical Journal*, vol. 7, no. 4, pp. 673–697, 2007.

[78] V. Baranovsky, V. Ginzburg, and A. Kuznetsov, "Wilson's grassmannian and a noncommutative quadric," *International Mathematics Research Notices*, no. 21, pp. 1155–1197, 2003.

[79] M. Blau and G. Thompson, "Euclidean SYM theories by time reduction and special holonomy manifolds," *Physics Letters. B*, vol. 415, no. 3, pp. 242–252, 1997.

[80] H. Nakajima and K. Yoshioka, "Lectures on instanton counting," in *Algebraic Structures and Moduli Spaces*, vol. 38 of *CRM Proceedings & Lecture Notes*, pp. 31–101, American Mathematical Society, Providence, RI, USA, 2004.

[81] S. Shadchin, "Saddle point equations in Seiberg-Witten theory," *Journal of High Energy Physics*, vol. 2004, no. 10, article 33, 2004.

[82] A. S. Losev, A. V. Marshakov, and N. A. Nekrasov, "Small instantons, little strings and free fermions," in *From Fields to Strings: Circumnavigating Theoretical Physics*, M. Shifman, A. Vainshtein, and J. F. Wheater, Eds., vol. 1, pp. 581–621, World Scientific, Singapore, 2005.

[83] N. A. Nekrasov, "Localizing gauge theories," in *14th International Congress on Mathematical Physics*, J.-C. Zambrini, Ed., pp. 645–654, World Scientific, Hackensack, NJ, USA, 2005.

[84] M. Aganagic, H. Ooguri, N. Saulina, and C. Vafa, "Black holes, q-deformed *2D* Yang-Mills, and non-perturbative topological strings," *Nuclear Physics B*, vol. 715, no. 1-2, pp. 304–348, 2005.

[85] C. Vafa, "Two-dimensional Yang-Mills, black holes and topological strings," http://arxiv.org/abs/hep-th/0406058.

[86] J. Bryan and R. Pandharipande, "The local Gromov-Witten theory of curves," *Journal of the American Mathematical Society*, vol. 21, no. 1, pp. 101–136, 2008.

[87] R. M. G. Reis and R. J. Szabo, "Geometric K-homology of flat D-branes," *Communications in Mathematical Physics*, vol. 266, no. 1, pp. 71–122, 2006.

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 270694, 28 pages
doi:10.1155/2010/270694

Review Article

Recent Developments in Instantons in Noncommutative $\mathbb{R}^4$

Akifumi Sako

Department of General Education, Kushiro National College of Technology Otanoshike-Nishi 2-32-1, Kushiro 084-0916, Japan

Correspondence should be addressed to Akifumi Sako, sako@kushiro-ct.ac.jp

Received 24 January 2010; Accepted 20 April 2010

Academic Editor: K. B. Sinha

We review recent developments in noncommutative deformations of instantons in $\mathbb{R}^4$. In the operator formalism, we study how to make noncommutative instantons by using the ADHM method, and we review the relation between topological charges and noncommutativity. In the ADHM methods, there exist instantons whose commutative limits are singular. We review smooth noncommutative deformations of instantons, spinor zero-modes, the Green's functions, and the ADHM constructions from commutative ones that have no singularities. It is found that the instanton charges of these noncommutative instanton solutions coincide with the instanton charges of commutative instantons before noncommutative deformation. These smooth deformations are the latest developments in noncommutative gauge theories, and we can extend the procedure to other types of solitons. As an example, vortex deformations are studied.

1. Introduction

Instantons in commutative space are one of the most important objects for nonperturbative analysis. We can overview them for example in [1] from the physicist's view points or in [2] from mathematical view points. See for example [3] for recent developments of them. Noncommutative (NC for short) instantons were discovered by Nekrasov and Schwarz [4]. After [4], NC instantons have been investigated by many physicists and mathematicians. However, many enigmas are left until now. Let us focus into instantons of $U(N)$ gauge theories in NC $\mathbb{R}^4$ and understand what is clarified and what is unknown.

Instanton connections in the 4-dim Yang-Mills theory are defined by

$$F^+ = \frac{1}{2}(1 + *)F = 0, \tag{1.1}$$

where F is a curvature 2-form and $*$ is the Hodge star operator. This condition says that curvature is anti-self-dual. In this paper, we call anti-self-dual connections instantons. The choice of anti-self-dual connection or self-dual connection to define instantons is not important to mathematics but just a habit.

NC instanton solutions were discovered by Nekrasov and Schwartz by using the ADHM method [4]. (See also [5] for the original ADHM method.) The ADHM construction which generates the instanton $U(N)$ gauge field requires a pair of the two complex vector spaces $V = \mathbb{C}^k$ and $W = \mathbb{C}^N$. Here $-k$ is an integer called instanton number. Introduce $B_1, B_2 \in \mathrm{Hom}(V,V)$, $I \in \mathrm{Hom}(W,V)$, and $J \in \mathrm{Hom}(V,W)$ which are called ADHM data that satisfy the ADHM equations that we will see soon. In other words, B_1 and B_2 are complex-valued $k \times k$ matrices, and I and $J^\dagger$ are complex-valued $k \times N$ matrices that satisfy (2.13) and (2.14) in Section 2.2. Using these ADHM data, we can construct instanton [6–17]. We call it NC ADHM instanton in the following. The NC ADHM construction is a strong method. A lot of instanton solutions are constructed by using the NC ADHM construction [6–17]. The NC ADHM method also clarifies some important features, for example, topological charge, index theorems, Green's functions, and so on. As a characteristic feature of NC ADHM construction, the NC ADHM instantons can be instantons that have singularities in the commutative limit. On the other hand, we can study NC instantons from a point of view of deformation quantization. Recently, NC instanton that is smoothly deformed from commutative instanton is constructed [18]. The method in [18] makes success in analysis for topological charges, index theorems, and the method derives the ADHM equations from NC instanton and proves a one-to-one correspondence between the ADHM data and NC instantons [19]. We review them in this article.

This paper is organized as follows. In Section 2, we review the NC ADHM instanton and their natures (For example, we investigate topological charges of instantons. We distinguish the terms "instanton number" from "instanton charge". In this article, we define the instanton number by the dimension of some vector space V; on the other hand, the instanton charge is defined by integral of the 2nd Chern class. We will soon see more details.) . In Section 3, we construct an NC instanton solution which is a smooth deformation of the commutative instanton [18]. We study the NC instanton charge, an index theorem, and the correspondence relation with the ADHM construction for the smooth NC deformations of instantons [19]. In Section 4, we apply the method in Section 3 to a gauge theory in $\mathbb{R}^2$, and we make NC vortex solutions which are smooth deformations of commutative vortex solutions [20, 21].

2. Noncommutative ADHM Instantons

In this section, we review the NC ADHM instanton that may have singularities in commutative limit. An NC $U(1)$ instanton is a typical example that has a singularity in commutative limit.

2.1. Notations for the Fock Space Formalism

Let us consider coordinate operators x^μ ($\mu = 1,2,3,4$) satisfying $[x^\mu, x^\nu] = i\theta^{\mu\nu}$, where θ is a skew symmetric real valued matrix and we call $\theta^{\mu\nu}$ NC parameter. We set the noncommutativity of the space to the self-dual case of $\theta^{12} = -\zeta_1$, $\theta^{34} = -\zeta_2$, and the other

$\theta^{\mu\nu} = 0$ for convenience. By transformations of coordinates x^μ, the NC parameters are possible to be put in this form in general. Here we introduce complex coordinate operators

$$z_1 = \frac{1}{\sqrt{2}}\left(x^1 + ix^2\right), \qquad z_2 = \frac{1}{\sqrt{2}}\left(x^3 + ix^4\right). \tag{2.1}$$

Then the commutation relations become

$$[z_1, \overline{z}_1] = -\zeta_1, \quad [z_2, \overline{z}_2] = -\zeta_2, \quad \text{others are zero.} \tag{2.2}$$

We define creation and annihilation operators by

$$c_\alpha^\dagger = \frac{z_\alpha}{\sqrt{\zeta_\alpha}}, \quad c_\alpha = \frac{\overline{z}_\alpha}{\sqrt{\zeta_\alpha}}, \quad (\alpha = 1, 2); \tag{2.3}$$

then they satisfy

$$\left[c_\alpha, c_\alpha^\dagger\right] = 1, \quad \left[c_\alpha, c_\beta\right] = \left[c_\alpha^\dagger, c_\beta^\dagger\right] = 0 \quad (\alpha, \beta = 1, 2). \tag{2.4}$$

The Fock space $\mathscr{H}$ on which the creation and annihilation operators (2.4) act is spanned by the Fock state

$$|n_1, n_2\rangle = \frac{\left(c_1^\dagger\right)^{n_1}\left(c_2^\dagger\right)^{n_2}}{\sqrt{n_1! n_2!}}|0, 0\rangle, \tag{2.5}$$

with

$$\begin{aligned} c_1|n_1, n_2\rangle &= \sqrt{n_1}|n_1 - 1, n_2\rangle, & c_1^\dagger|n_1, n_2\rangle &= \sqrt{n_1 + 1}|n_1 + 1, n_2\rangle, \\ c_2|n_1, n_2\rangle &= \sqrt{n_2}|n_1, n_2 - 1\rangle, & c_2^\dagger|n_1, n_2\rangle &= \sqrt{n_2 + 1}|n_1, n_2 + 1\rangle, \end{aligned} \tag{2.6}$$

where n_1 and n_2 are the occupation number. The number operators are also defined by

$$\widehat{n}_\alpha = c_\alpha^\dagger c_\alpha, \qquad \widehat{N} = \widehat{n}_1 + \widehat{n}_2, \tag{2.7}$$

which act on the Fock states as

$$\widehat{n}_\alpha|n_1, n_2\rangle = n_\alpha|n_1, n_2\rangle, \qquad \widehat{N}|n_1, n_2\rangle = (n_1 + n_2)|n_1, n_2\rangle. \tag{2.8}$$

In the operator representation, derivatives of a function f are defined by

$$\partial_\alpha f(z) = \left[\widehat{\partial}_\alpha, f(z)\right], \qquad \partial_{\overline{\alpha}} f(z) = \left[\widehat{\partial}_{\overline{\alpha}}, f(z)\right], \tag{2.9}$$

where $\hat{\partial}_\alpha = \bar{z}_\alpha/\zeta_\alpha$ and $\hat{\partial}_{\bar{\alpha}} = -z_\alpha/\zeta_\alpha$ which satisfy $[\hat{\partial}_\alpha, \hat{\partial}_{\bar{\alpha}}] = -1/\zeta_\alpha$. The integral on NC $\mathbb{R}^4$ is defined by the standard trace in the operator representation,

$$\int d^4x = \int d^4z = (2\pi)^2\zeta_1\zeta_2 \mathrm{Tr}_{\mathscr{H}}. \tag{2.10}$$

Note that $\mathrm{Tr}_{\mathscr{H}}$ represents the trace over the Fock space whereas the trace over the gauge group is denoted by $\mathrm{tr}_{U(N)}$.

2.2. Noncommutative ADHM Instantons

Let us consider the $U(N)$ Yang-Mills theory on NC $\mathbb{R}^4$. Let M be a projective module over the algebra that is generated by the operator x_μ.

In the NC space, the Yang-Mills connection is defined by $D_\mu\psi = -\psi\hat{\partial}_\mu + \hat{D}_\mu\psi$, where ψ is a matter field in fundamental representation type and $\hat{D}_\mu \in \mathrm{End}(M)$ are anti-Hermitian gauge fields [22–24]. The relation between $\hat{D}_\mu$ and usual gauge connection A_μ is $\hat{D}_\mu = -i\theta_{\mu\nu}x^\nu + A_\mu$, where $\theta_{\mu\nu}$ is an inverse matrix of $\theta^{\mu\nu}$. In our notation of the complex coordinates (2.1) and (2.2), the curvature is given as

$$F_{\alpha\bar{\alpha}} = \frac{1}{\zeta_\alpha} + \left[\hat{D}_\alpha, \hat{D}_{\bar{\alpha}}\right], \quad F_{\alpha\bar{\beta}} = \left[\hat{D}_\alpha, \hat{D}_{\bar{\beta}}\right] \quad (\alpha \neq \beta). \tag{2.11}$$

Note that there is a constant term originated with the noncommutativity in $F_{\alpha\bar{\alpha}}$. Instanton solutions satisfy the antiself-duality condition $F = -*F$. These conditions are rewritten in the complex coordinates as

$$F_{1\bar{1}} = -F_{2\bar{2}}, \qquad F_{12} = F_{\bar{1}\bar{2}} = 0. \tag{2.12}$$

In the commutative spaces, instantons are classified by the topological charge $Q = (1/8\pi^2)\int \mathrm{tr}_{U(N)}F \wedge F$, which is always integer $-k$ and coincide with the opposite sign of dimension of the vector space V in the ADHM methods, and $-k$ is called instanton number. In the NC spaces, the same statement is conjectured, and some partial proofs are given. (See Section 2.4 and see also [18, 25–30].)

In the commutative spaces, the ADHM construction is proposed by Atiyah et al. [5] to construct instantons. Nekrasov and Schwarz first extended this method to NC cases [4]. Here we review briefly on the ADHM construction of $U(N)$ instantons [22, 23].

The first step of ADHM construction on NC $\mathbb{R}^4$ is looking for $B_1, B_2 \in \mathrm{End}(\mathbb{C}^k)$, $I \in \mathrm{Hom}(\mathbb{C}^n, \mathbb{C}^k)$, and $J \in \mathrm{Hom}(\mathbb{C}^k, \mathbb{C}^n)$ which satisfy the deformed ADHM equations

$$\left[B_1, B_1^\dagger\right] + \left[B_2, B_2^\dagger\right] + II^\dagger - J^\dagger J = \zeta_1 + \zeta_2, \tag{2.13}$$

$$[B_1, B_2] + IJ = 0. \tag{2.14}$$

We call $-k$ "instanton number" in this article. In the previous section, we denote V as the vector space $\mathbb{C}^k$. Note that the right-hand side of (2.13) is caused by the noncommutativity of the space $\mathbb{R}^4$. The set of $B_1, B_2, I,$ and J satisfying (2.13) and (2.14) is called ADHM data. Using this ADHM data, we define operator $\mathfrak{D} : \mathbb{C}^k \oplus \mathbb{C}^k \oplus \mathbb{C}^n \to \mathbb{C}^k \oplus \mathbb{C}^k$ by

$$\begin{aligned}
\mathfrak{D}^\dagger &= \begin{pmatrix} \tau \\ \sigma^\dagger \end{pmatrix}, \\
\tau &= (B_2 - z_2, B_1 - z_1, I) = \left(B_2 - \sqrt{\zeta_2}c_2^\dagger, B_1 - \sqrt{\zeta_1}c_1^\dagger, I\right), \\
\sigma^\dagger &= \left(-B_1^\dagger + \overline{z}_1, B_2^\dagger - \overline{z}_2, J^\dagger\right) = \left(-B_1^\dagger + \sqrt{\zeta_1}c_1, B_2^\dagger - \sqrt{\zeta_2}c_2, J^\dagger\right).
\end{aligned} \tag{2.15}$$

The ADHM equations (2.13) and (2.14) are replaced by

$$\tau\tau^\dagger = \sigma^\dagger\sigma \equiv \Box, \qquad \tau\sigma = 0. \tag{2.16}$$

Let us denote by $\Psi : \mathbb{C}^n \to \mathbb{C}^k \oplus \mathbb{C}^k \oplus \mathbb{C}^n$ the solution to the following equation:

$$\mathfrak{D}^\dagger\Psi^a = 0 \quad (a = 1, \ldots, n), \quad \Psi^{\dagger a}\Psi^b = \delta^{ab}. \tag{2.17}$$

Theorem 2.1. *Let Ψ^a be orthonormal zero-modes defined in* (2.17). *Then NC $U(N)$ instanton A_μ with instanton number $-k$ is obtained by*

$$A_\mu = \Psi^\dagger\partial_\mu\Psi = -i\Psi^\dagger\theta_{\mu\nu}[x^\nu, \Psi]. \tag{2.18}$$

Here $\theta_{\mu\nu}$ is inverse of $\theta^{\mu\nu}$, that is, $\theta_{\mu\nu}\theta^{\nu\rho} = \delta_\mu^\rho$.

Proof. The curvature two-form determined by this connection is given as follows.

$$\begin{aligned}
F &= dA + A \wedge A \\
&= d\left(\Psi^\dagger d\Psi\right) + \left(\Psi^\dagger d\Psi\right) \wedge \left(\Psi^\dagger d\Psi\right) \\
&= d\Psi^\dagger \wedge d\Psi - (d\Psi)\Psi\Psi^\dagger \wedge d\Psi \\
&= d\Psi^\dagger\left(1 - \Psi\Psi^\dagger\right) \wedge d\Psi.
\end{aligned} \tag{2.19}$$

Here we use $d\Psi^\dagger\Psi + \Psi^\dagger d\Psi = 0$ that follows from the differentiating of (2.17). Note that

$$\Psi\Psi^\dagger = I - \mathfrak{D}\frac{1}{\mathfrak{D}^\dagger\mathfrak{D}}\mathfrak{D}^\dagger, \tag{2.20}$$

since

$$
\begin{aligned}
I &= (\mathfrak{D}\ \ \Psi)(\mathfrak{D}\ \ \Psi)^{-1}\left((\mathfrak{D}\ \ \Psi)^{\dagger}\right)^{-1}(\mathfrak{D}\ \ \Psi)^{\dagger} \\
&= (\mathfrak{D}\ \ \Psi)\begin{pmatrix} \mathfrak{D}^{\dagger}\mathfrak{D} & 0 \\ 0 & 1 \end{pmatrix}^{-1}(\mathfrak{D}\ \ \Psi)^{\dagger} = \mathfrak{D}\frac{1}{\mathfrak{D}^{\dagger}\mathfrak{D}}\mathfrak{D}^{\dagger} + \Psi\Psi^{\dagger}.
\end{aligned}
\tag{2.21}
$$

From (2.19) and (2.20),

$$
F = d\Psi^{\dagger}\left(\mathfrak{D}\frac{1}{\mathfrak{D}^{\dagger}\mathfrak{D}}\mathfrak{D}^{\dagger}\right)\wedge d\Psi = \Psi^{\dagger}(d\mathfrak{D})\wedge\frac{1}{\mathfrak{D}^{\dagger}\mathfrak{D}}\left(d\mathfrak{D}^{\dagger}\right)\Psi,
\tag{2.22}
$$

where we use $(d\mathfrak{D}^{\dagger})\Psi + \mathfrak{D}^{\dagger}d\Psi = 0$ that follows from differentiating $\mathfrak{D}^{\dagger}\Psi = 0$. If the coordinates (x_1, x_2, x_3, x_4) are renamed (x_2, x_1, x_4, x_3) for convenience, we obtain

$$
\partial_{\mu}\mathfrak{D}^{\dagger} = \frac{1}{\sqrt{2}}(-\overline{\sigma}_{\mu}\ \ 0), \qquad \partial_{\mu}\mathfrak{D} = \frac{1}{\sqrt{2}}\begin{pmatrix} -\sigma_{\mu} \\ 0 \end{pmatrix}.
\tag{2.23}
$$

Here, we define σ_{μ} and $\overline{\sigma}_{\mu}$ by

$$
\begin{aligned}
(\sigma_1, \sigma_2, \sigma_3, \sigma_4) &:= (-i\tau_1, -i\tau_2, -i\tau_3, 1_{2\times 2}), \\
(\overline{\sigma}_1, \overline{\sigma}_2, \overline{\sigma}_3, \overline{\sigma}_4) &:= (i\tau_1, i\tau_2, i\tau_3, 1_{2\times 2}),
\end{aligned}
\tag{2.24}
$$

where τ_i are the Pauli matrices and $1_{2\times 2}$ is an identity matrix of degree 2. Note that $\mathfrak{D}^{\dagger}\mathfrak{D} = \left(\begin{smallmatrix} \square & 0 \\ 0 & \square \end{smallmatrix}\right)$ owing to (2.16), and $\mathfrak{D}^{\dagger}\mathfrak{D}$ and its inverse commute with σ_{μ}. Then we find (2.22) is in proportion to

$$
\sigma_{\mu}\overline{\sigma}_{\nu}dx^{\mu}\wedge dx^{\nu}.
\tag{2.25}
$$

$\sigma_{\mu}\overline{\sigma}_{\nu} - \sigma_{\nu}\overline{\sigma}_{\mu}$ is a component of anti-self-dual two-form, that is easily checked by direct calculations. This fact and (2.22) show that the curvature F is anti-self-dual and the connections given by (2.18) are instantons. □

With the complex coordinate z_{α}, NC instanton connections are given by

$$
\widehat{D}_{\alpha} = \frac{1}{\zeta_{\alpha}}\Psi^{\dagger}\overline{z}_{\alpha}\Psi, \qquad \widehat{D}_{\overline{\alpha}} = -\frac{1}{\zeta_{\alpha}}\Psi^{\dagger}z_{\alpha}\Psi.
\tag{2.26}
$$

One of the most important feature to understand the origin of the instanton charges is existence of zero-modes of $\Psi\Psi^{\dagger}$.

Theorem 2.2 (Zero-mode of $\Psi\Psi^{\dagger}$). *Suppose that Ψ and $\Psi^{\dagger}$ are given as above. The vector $|v\rangle \in (\mathbb{C}^k \oplus \mathbb{C}^k \oplus \mathbb{C}^n)\otimes\mathscr{H}$ satisfying*

$$
\Psi\Psi^{\dagger}|v\rangle = \langle v|\Psi\Psi^{\dagger} = 0, \qquad |v\rangle \neq 0
\tag{2.27}
$$

is said to be a zero-mode of $\Psi\Psi^\dagger$. The zero-modes are given by following three types:

$$|v_1\rangle = \begin{pmatrix} \left(-B_1 + \sqrt{\zeta_1}c_1\right)|u\rangle \\ \left(B_2 - \sqrt{\zeta_2}c_2\right)|u\rangle \\ J|u\rangle \end{pmatrix}, \qquad |v_2\rangle = \begin{pmatrix} \left(B_2^\dagger - \sqrt{\zeta_2}c_2^\dagger\right)|u'\rangle \\ \left(B_1^\dagger - \sqrt{\zeta_1}c_1^\dagger\right)|u'\rangle \\ I^\dagger|u'\rangle \end{pmatrix},$$
$$|v_0\rangle = \begin{pmatrix} \left(\exp\sum_\alpha B_\alpha^\dagger c_\alpha^\dagger\right)|0,0\rangle v_0^i \\ \left(\exp\sum_\alpha B_\alpha^\dagger c_\alpha^\dagger\right)|0,0\rangle v_0^i \\ 0 \end{pmatrix}. \tag{2.28}$$

Here $|u\rangle$ ($|u'\rangle$) is some element of $\mathbb{C}^k \otimes \mathcal{H}$ (i.e., $|u\rangle$ is expressed with the coefficients $u_i^{nm} \in \mathbb{C}$ as $|u\rangle = \sum_i \sum_{n,m} u_i^{nm}|n,m\rangle e_i$, where e_i is a base of k-dim vector space). v_0^i is a element of k-dim vector.

The proof is given in [25]. We will see the fact that zero-modes $|v_0\rangle$ play an essential role, in the following subsections.

2.3. $U(1)$ *N.C. ADHM Multi-Instanton*

One of the most characteristic features of NC instantons is found in regularizations of the singularities. In commutative $\mathbb{R}^4$, we cannot construct a nonsingular $U(1)$ instanton. On the other hand, there exist in NC $\mathbb{R}^4$. Let us see how to construct them as typical NC ADHM instantons.

At the beginning, we review the methods in [23]. Let B_1, B_2, I, J be constant matrices satisfying (2.13) and (2.14). We consider $\zeta = \zeta_1 + \zeta_2 > 0$; then we can put $J = 0$ in general by using a symmetry. B_1 and B_2 are $k \times k$ matrices, and I is $k \times 1$ matrices:

$$B_1 = \begin{array}{c c} & \begin{array}{ccc} 1 & \cdots & k \end{array} \\ \begin{array}{c} 1 \\ \vdots \\ k \end{array} & \begin{pmatrix} B_{11}^1 & \cdots & B_{1k}^1 \\ \vdots & \ddots & \vdots \\ B_{k1}^1 & \cdots & B_{kk}^1 \end{pmatrix} \end{array}, \qquad B_2 = \begin{array}{c c} & \begin{array}{ccc} 1 & \cdots & k \end{array} \\ \begin{array}{c} 1 \\ \vdots \\ k \end{array} & \begin{pmatrix} B_{11}^2 & \cdots & B_{1k}^2 \\ \vdots & \ddots & \vdots \\ B_{k1}^2 & \cdots & B_{kk}^2 \end{pmatrix} \end{array},$$
$$I = \begin{array}{c c} & 1 \\ \begin{array}{c} 1 \\ 2 \\ \vdots \\ k \end{array} & \begin{pmatrix} I_1 \\ I_2 \\ \vdots \\ I_k \end{pmatrix} \end{array}, \qquad J = 0. \tag{2.29}$$

We define β_α, $c_\alpha^\dagger$ and c_α by

$$B_\alpha = \sqrt{\zeta_\alpha}\beta_\alpha \quad (\alpha = 1, 2). \tag{2.30}$$

We introduce $\widehat{\Delta}$ as

$$\widehat{\Delta} = \sum_\alpha \zeta_\alpha \left(\beta_\alpha - c_\alpha^\dagger\right)\left(\beta_\alpha^\dagger - c_\alpha\right), \tag{2.31}$$

and we define a projection operator P as a projection onto 0-eigenstates of $\widehat{\Delta}$ by

$$P = I^\dagger e^{\sum_\alpha \beta_\alpha^\dagger c_\alpha^\dagger} |0,0\rangle G^{-1} \langle 0,0| e^{\sum_\alpha \beta_\alpha c_\alpha} I, \tag{2.32}$$

where

$$G = \langle 0,0| e^{\sum_\alpha \beta_\alpha c_\alpha} I I^\dagger e^{\sum_\alpha \beta_\alpha^\dagger c_\alpha^\dagger} |0,0\rangle. \tag{2.33}$$

We define shift operators S and $S^\dagger$ and a operator Λ by

$$\begin{gathered} SS^\dagger = 1, \qquad S^\dagger S = 1 - P, \\ \Lambda = 1 + I^\dagger \frac{1}{\widehat{\Delta}} I. \end{gathered} \tag{2.34}$$

Theorem 2.3 (Nekrasov). *$U(1)$ instantons are given by*

$$D_\alpha = -\sqrt{\frac{1}{\zeta_\alpha}} S \Lambda^{-1/2} c_\alpha \Lambda^{1/2} S^\dagger, \qquad D_{\overline{\alpha}} = \sqrt{\frac{1}{\zeta_\alpha}} S \Lambda^{1/2} c_\alpha^\dagger \Lambda^{-1/2} S^\dagger. \tag{2.35}$$

Proof. At first, we check that the inverse of $\widehat{\Delta}$ in (2.34) is well defined. $\widehat{\Delta}$ has k zero-modes:

$$e^{\sum_\alpha \beta_\alpha^\dagger c_\alpha^\dagger} |0,0\rangle \otimes e_i \quad (i = 1, \dots, k) \tag{2.36}$$

which satisfy $\widehat{\Delta} e^{\sum_\alpha \beta_\alpha^\dagger c_\alpha^\dagger} |0,0\rangle \otimes e_i = 0$. Here $e_i = (\delta_{1i}, \delta_{2,i}, \dots, \delta_{ki})^t$ is a base of V. Note that $S \cdots S^\dagger = S(1-P) \cdots S^\dagger$. This implies that S removes the zero-modes, and Hilbert spaces $\mathscr{H}$ is projected on to a space that does not include the zero-modes. Therefore, the inverse of Λ exists if it is sandwiched between S and $S^\dagger$ and (2.35) is well defined.

Next, we check that (2.35) is an instanton. Let us see how the equation $\mathfrak{D}^\dagger \Psi = 0$ is solved under orthonormalization condition $\Psi^\dagger \Psi = 1$. $\psi_\pm$ and ξ are introduced as

$$\Psi = \begin{pmatrix} \psi_+ \\ \psi_- \\ \xi \end{pmatrix}, \quad \psi_\pm \in V \otimes \mathscr{H}, \ \xi \in \mathscr{H}. \tag{2.37}$$

The orthonormalization condition is expressed as

$$\varphi_+^\dagger\varphi_+ + \varphi_-^\dagger\varphi_- + \xi^\dagger\xi = 1. \tag{2.38}$$

We put anzats for the solution by

$$\varphi_+ = -\sqrt{\zeta_2}\left(\beta_2^\dagger - c_2\right)v, \qquad \varphi_- = \sqrt{\zeta_1}\left(\beta_1^\dagger - c_1\right)v, \tag{2.39}$$

and substitute them into $\mathfrak{D}^\dagger\Psi = 0$; then we get

$$\widehat{\Delta}v + I\xi = 0. \tag{2.40}$$

The orthonormalization condition is rewritten as

$$v^\dagger\widehat{\Delta}v + \xi^\dagger\xi = 1. \tag{2.41}$$

If there exist the inverse of $\widehat{\Delta}$,

$$v = -\frac{1}{\widehat{\Delta}}I\xi. \tag{2.42}$$

0-eigenstates of $\widehat{\Delta}$ are (2.36) and we define the projection operator to project out the 0-eigenstates by

$$P = \mathrm{I}^\dagger e^{\sum_\alpha \beta_\alpha^\dagger c_\alpha^\dagger}|0,0\rangle G^{-1}\langle 0,0|e^{\sum_\alpha \beta_\alpha c_\alpha}I. \tag{2.43}$$

Shift operators $S, S^\dagger$ satisfying

$$SS^\dagger = 1, \qquad S^\dagger S = 1 - P \tag{2.44}$$

are determined by the definition of P. Then the inverse of $\widehat{\Delta}$ is well defined at the left side of $S^\dagger$ or the right side of S.

Using the orthonormalization condition, we obtain

$$\xi = \Lambda^{-1/2}S^\dagger, \qquad \Lambda = 1 + I^\dagger\frac{1}{\widehat{\Delta}}I. \tag{2.45}$$

Through these processes, Ψ is determined by the ADHM data, and after substituting this Ψ into (2.26) we obtain the instantons.

$$\begin{aligned} D_\alpha &= -\frac{1}{\xi_\alpha}\psi^\dagger \widehat{\overline{z}}_\alpha \psi = -\frac{1}{\xi_\alpha} S\Lambda^{-1/2}\left(I^\dagger \frac{1}{\widehat{\Delta}}\widehat{\Delta}\widehat{\overline{z}}_\alpha \frac{1}{\widehat{\Delta}} I - \widehat{\overline{z}}_\alpha\right)\Lambda^{-1/2}S^\dagger \\ &= -\frac{1}{\sqrt{\xi_\alpha}} S\Lambda^{-1/2} c_\alpha \Lambda^{1/2} S^\dagger. \end{aligned} \tag{2.46}$$

$D_{\overline{\alpha}}$ is given similarly. □

This expression (2.35) is useful, but there exist other issues to get concrete expression of instantons. For example, it is not easy to obtain the explicit expression of $\widehat{\Delta}^{-1}$.

As an example, let us construct an NC $U(1)$ multi-instanton having concrete expressions with the instanton number $-k$ [31, 32], which is made from the ADHM data:

$$\begin{gathered} B_1 = \sum_{l=1}^{k-1}\sqrt{l\zeta}\, e_l e_{l+1}^\dagger = \sqrt{\zeta}\begin{pmatrix} 0 & \sqrt{1} & 0 & \cdots & \cdots & 0 \\ 0 & 0 & \sqrt{2} & 0 & \cdots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & \cdots & 0 & \sqrt{k-2} & 0 \\ 0 & \cdots & & \cdots & 0 & \sqrt{k-1} \\ 0 & \cdots & & & \cdots & 0 \end{pmatrix}, \qquad B_2 = 0, \\ I = \sqrt{k\zeta}\, e_k = \sqrt{\zeta}\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \sqrt{k} \end{pmatrix}, \qquad J = 0, \end{gathered} \tag{2.47}$$

where $\zeta = \zeta_1 + \zeta_2$. It is easy to check that this data satisfies the ADHM equations (2.13) and (2.14), and substituting them into definition of P derives

$$P = \sum_{n_1=0}^{k-1} |n_1, 0\rangle\langle n_1, 0|. \tag{2.48}$$

To construct an instanton, it is necessary to obtain $\widehat{\Delta}$ or Λ. By definition,

$$\begin{gathered} \widehat{\Delta}(k) = \zeta_1 \widehat{n}_1 + \zeta_2 \widehat{n}_2 + \zeta\sum_{i=1}^{k-1} i e_i e_i^\dagger - \sqrt{\zeta_1\zeta}\sum_{i=1}^{k-1}\sqrt{i}\left\{c_1 e_i e_{i+1}^\dagger + c_1^\dagger e_{i+1} e_i^\dagger\right\}, \\ \Lambda(k) = 1 + \zeta k \widehat{\Delta}_{kk}^{-1}(k). \end{gathered} \tag{2.49}$$

$\widehat{\Delta}$ and Λ depend on k, so we denote them $\widehat{\Delta}(k)$ and $\Lambda(k)$, respectively. $\widehat{\Delta}^{-1}_{kk}(k)$ is (k,k) entry of matrix $\widehat{\Delta}^{-1}(k)$. To obtain $\widehat{\Delta}^{-1}_{kk}(k)$, it is enough to calculate the kth row vector of $\widehat{\Delta}^{-1}(k)$. The kth row vector of $\widehat{\Delta}^{-1}(k)$ is determined by $\widehat{\Delta}^{-1}\widehat{\Delta} = 1$. We denote the kth row vector of $\widehat{\Delta}^{-1}(k)$ by $(u_1,\ldots,u_k)$, that is, $\widehat{\Delta}^{-1}_{ki}(k) = u_i$. Then, we obtain the following recurrence equation from the kth row of $\widehat{\Delta}^{-1}\widehat{\Delta} = 1$

$$\begin{gathered} u_2 c_1^{\dagger} - \frac{1}{\sqrt{\widetilde{\theta}}} u_1 \left(1 + \widetilde{\theta}\widehat{n}_1 + \left(1-\widetilde{\theta}\right)\widehat{n}_2\right) = 0, \\ \sqrt{i} u_{i+1} c_1^{\dagger} - \frac{1}{\sqrt{\widetilde{\theta}}} u_i \left(i + \widetilde{\theta}\widehat{n}_1 + \left(1-\widetilde{\theta}\right)\widehat{n}_2\right) + \sqrt{i-1}\, u_{i-1} c_1 = 0 \quad (1 \le i \le k-2), \end{gathered} \tag{2.50}$$

where $\widetilde{\theta} = \zeta_1/(\zeta_1+\zeta_2)$. We change variables as

$$u_i = w_{i-1} \frac{c_1^{\dagger(k-i)}}{\sqrt{(i-1)!}}; \tag{2.51}$$

then we can rewrite the above recurrence relation by w_i as

$$\begin{aligned} & w_1 - \frac{1}{\sqrt{\widetilde{\theta}}}\left(1 + \widetilde{\theta}(\widehat{n}_1 - k + 1) - \left(1-\widetilde{\theta}\right)\widehat{n}_2\right) w_0 = 0, \\ & w_{i+1} - \left\{ i\left(\frac{1}{\sqrt{\widetilde{\theta}}} + \sqrt{\widetilde{\theta}}\right) + \frac{1}{\sqrt{\widetilde{\theta}}}\left(\widetilde{\theta}(\widehat{n}_1 - k) + \left(1-\widetilde{\theta}\right)\widehat{n}_2\right) + \left(\frac{1}{\sqrt{\widetilde{\theta}}} + \sqrt{\widetilde{\theta}}\right) \right\} w_i \\ & \quad + i((i-1)\widehat{n}_1 - k + 2) w_{i-1} = 0, \quad (2 \le i \le k-1). \end{aligned} \tag{2.52}$$

Note that $\widehat{n}_1$ and $\widehat{n}_2$ are commutative to each other, so we can treat them like C-numbers in the following. We introduce an anzats for the generating function $F(t;k)$ by

$$\begin{aligned} F(t;k) &= e^{f(t)}(1-at)^{\alpha} = \sum_{i=0}^{\infty} \frac{w_i}{i!} t^i, \\ f(t) &= \int dt \frac{ct}{(1-at)(1-bt)} \\ &= \frac{c}{2ab}\left\{ \ln\left(1 - (a+b)t + abt^2\right) + \frac{a+b}{\sqrt{D}} \ln\left| \frac{2abt - (a+b) - \sqrt{D}}{2abt - (a+b) + \sqrt{D}} \right| \right\}, \end{aligned} \tag{2.53}$$

where a, b, c, and α are real parameter determined by the request that w_i satisfy (2.52), and $D = (a-b)^2$. From the differentiation of $F(t;k)$, we obtain

$$(1-at)(1-bt)\sum_{i=1}^{\infty} \frac{w_1}{(i-1)!} t^{i-1} = \{-a\alpha + (c + ab\alpha)t\} \sum_{i=0}^{\infty} \frac{w_i}{i!} t^i, \tag{2.54}$$

and we find that w_i satisfy the following relation:

$$w_{i+1} - (i(a+b) - a\alpha)w_i + i(ab(i-1) - c - ab\alpha)w_{i-1} = 0. \tag{2.55}$$

From (2.52) and (2.55), we obtain

$$\begin{gathered} a = \sqrt{\tilde{\theta}} \quad \text{or} \quad \frac{1}{\sqrt{\tilde{\theta}}}, \qquad b = \frac{1}{a}, \\ \alpha = -\frac{h\left(\tilde{\theta}, n_1, n_2\right)}{a}, \qquad c = -n_1 + k - 2 + \frac{h\left(\tilde{\theta}, n_1, n_2\right)}{a}, \end{gathered} \tag{2.56}$$

where

$$h\left(\tilde{\theta}, n_1, n_2\right) = \frac{1}{\sqrt{\tilde{\theta}}}\left\{\tilde{\theta}(n_1 - k) + \left(1 - \tilde{\theta}\right)n_2\right\} + \tilde{\theta} + \frac{1}{\sqrt{\tilde{\theta}}}. \tag{2.57}$$

Thus the generating function $F(t;k)$ is determined as an elementary function for each instanton number $-k$. Using this $F(t;k)$, we obtain w_i, and $\Delta^{-1}_{kk}(k)$ is determined as

$$\Delta^{-1}_{kk}(k) = u_k = \left(1 + \sqrt{\zeta_1 \zeta}\frac{\sqrt{k-1}}{\sqrt{(k-2)!}}w_{k-2}\hat{n}_1\right)(\zeta_1\hat{n}_1 + \zeta_2\hat{n}_2)^{-1}. \tag{2.58}$$

Using them, G, P, S, and Λ are determined as

$$\begin{gathered} G = \zeta k! \sum_{i=1}^{k}\left\{i!(k-i)!\tilde{\theta}^{k-i}\right\}^{-1} e_i e_i^{\dagger}, \\ P = I^{\dagger} e^{\sum_\alpha \beta_\alpha^{\dagger} c_\alpha^{\dagger}} |0,0\rangle G^{-1} \langle 0,0| e^{\sum_\alpha \beta_\alpha c_\alpha} I = \sum_{n_1=0}^{k-1} |n_1, 0\rangle\langle n_1, 0|, \\ S^{\dagger} = \sum_{n_1=0}^{\infty} |n_1 + k, 0\rangle\langle n_1, 0| + \sum_{n_1=0}^{\infty}\sum_{n_2=1}^{\infty} |n_1, n_2\rangle\langle n_1, n_2|, \\ \Lambda = 1 + \zeta k u_k. \end{gathered} \tag{2.59}$$

Finally we obtain instanton gauge fields with instanton number $-k$ as

$$\begin{gathered} D_1 = \sqrt{\frac{1}{\zeta_1}} \sum_{n_1=0}^{\infty}\sum_{n_2=0}^{\infty} |n_1, n_2\rangle\langle n_1 + 1, n_2| d_1(n_1, n_2; k), \\ D_2 = \sqrt{\frac{1}{\zeta_2}}\left\{\sum_{n_1=0}^{\infty} |n_1, 0\rangle\langle n_1 + k, 1| d_2(n_1, 0; k) + \sum_{n_1=0}^{\infty}\sum_{n_2=1}^{\infty} |n_1, n_2\rangle\langle n_1, n_2 + 1| d_2(n_1, n_2; k)\right\}, \end{gathered} \tag{2.60}$$

where

$$d_1(n_1,n_2;k)=\begin{cases}\sqrt{n_1+k+1}\left[\dfrac{\Lambda(n_1+k+1,0)}{\Lambda(n_1+k,0)}\right]^{1/2}, & (n_2=0),\\ \sqrt{n_1+1}\left[\dfrac{\Lambda(n_1+1,n_2)}{\Lambda(n_1,n_2)}\right]^{1/2}, & (n_2\neq 0),\end{cases}$$
$$d_2(n_1,n_2;k)=\begin{cases}\left[\dfrac{\Lambda(n_1+k,1)}{\Lambda(n_1+k,0)}\right]^{1/2}, & (n_2=0),\\ \sqrt{n_2+1}\left[\dfrac{\Lambda(n_1,n_2+1)}{\Lambda(n_1,n_2)}\right]^{1/2}, & (n_2\neq 0).\end{cases} \tag{2.61}$$

Therefore, we obtain NC multi-instanton solutions expressed completely by elementary functions. This solution is one of the examples of the many kinds of the NC multi-instantons discovered until now [6–17].

2.4. Some Aspects

In this section, we overview some facts and important aspects of NC instantons without detailed derivations.

2.4.1. Instanton Charges and Instanton Numbers

Let us see a rough sketch of how to define instanton charges by using characteristic classes. The instanton charge in commutative space is determined $(1/8\pi^2)\int \mathrm{tr}\, F\wedge\star F$ and coincides with the instanton number defined by the dimension of the vector space V in the ADHM construction. A naive definition of the instanton charges in NC $\mathbb{R}^4$ is given by replacement of $\int d^4x$ by $(2\pi)^2\zeta_1\zeta_2\mathrm{Tr}_{\mathscr{H}}$, but it is conditionally convergent in general. In [25, 26], we introduce cut-off N_C for the Fock space and make the instanton charge be a converge series. The region of the initial and final state of the Fock space with the boundary is

$$|n_1,n_2\rangle \quad (n_1=0,\ldots,N_1(n_2),\ n_2=0,\ldots,N_2(n_1)), \tag{2.62}$$

where $N_1(n_2)$ $(N_2(n_1))$ is a function of n_2 (n_1) and we suppose that the length of the boundary is order $N_C\gg k$, that is, $N_1(n_2)\approx N_2(n_1)\approx N_C\gg k$.

Using this cut-off (boundary), we define the instanton charge by

$$Q=\lim_{N_C\to\infty}Q_{N_C},$$
$$Q_{N_C}=\zeta^2\sum_{n_1=0}\sum_{n_2=0}^{N_1(n_1)}\langle n_1,n_2|\left(F_{1\bar{1}}F_{2\bar{2}}-F_{1\bar{2}}F_{2\bar{1}}\right)|n_1,n_2\rangle. \tag{2.63}$$

As described in [25, 26], the regions for summations of intermediate states are shifted. This phenomenon is caused by the existence of the $\Psi\Psi^\dagger$ zero-mode $\langle v_0|$.

The following terms appear in the instanton charge Q_{N_C}:

$$-\mathrm{tr}_{U(N)}\mathrm{Tr}_{N_C}\left(\frac{1}{2}\left[\Psi^\dagger c_2^\dagger \Psi, \Psi^\dagger c_2 \Psi\right] + \frac{1}{2}\left[\Psi^\dagger c_1^\dagger \Psi, \Psi^\dagger c_1 \Psi\right]\right). \tag{2.64}$$

We denote Tr_{N_C} as trace over some finite domain of Fock space characterized by N_C which is the length of the Fock space boundary. Using the Stokes' like theorem in [25], only trace over the boundary is left, then $\mathrm{Tr}_{N_C}[\Psi^\dagger c_2^\dagger \Psi, \Psi^\dagger c_2 \Psi]$ becomes

$$-\mathrm{tr}_{U(N)} \sum_{\text{boundary}} (N_2(n_1) + 1) = -\mathrm{tr}_{U(N)}\mathrm{Tr}_{N_C} 1 - k. \tag{2.65}$$

The same value is obtained from $\mathrm{Tr}_N[\Psi^\dagger c_1^\dagger \Psi, \Psi^\dagger c_1 \Psi]$, too. The first term in (2.65) and the term from the constant curvature in (2.11) cancel out. The second term $-k$ is occurred by zero-modes $|v_0\rangle$. Finally the second term of (2.65) is understood as the source of the instanton charge. The origin of the instanton charge is shift of intermediate states caused by k zero-modes $|v_0\rangle$. After all, we get

$$Q_N = -k + O\left(N^{-1/2}\right), \qquad Q = \lim_{N\to\infty} Q_N = -k. \tag{2.66}$$

Theorem 2.4 (Instanton number). *Consider U(N) gauge theory on NC $\mathbb{R}^4$ with self-dual $\theta^{\mu\nu}$. The instanton charge Q is possible to be defined by limit of converge series and it is identified with the dimension k that appears in the ADHM construction and is called "instanton number".*

The strict proof is given in [25].

Note that the proof of the equivalence between the topological charge defined as the integral of the second Chern class and the instanton number given by the dimension of the vector space in the ADHM construction is not completed in NC space. In [27], Furuuchi shows how to appear zero-modes in the NC ADHM construction, and he shows that zero-modes project out some states in Fock space. In [28, 29], the geometrical origin of the instanton number for NC $U(1)$ gauge theory is clarified. In [25], the identification between the topological charge and the dimension of the vector space in the ADHM construction is shown for a $U(1)$ gauge theory. In [26], this identification is shown when the NC parameter is self-dual for a $U(N)$ gauge theory. In [30], the equivalence between the instanton numbers and instanton charges is shown with no restrictions on the NC parameters, but an NC version of the Osborn's identity (Corrigan's identity) is assumed. Until now, the relation between the instanton numbers and the topological charges in NC spaces had not been clarified completely. Moreover, the calculation in [25, 26] shows that the origin of the instanton number is deeply related to the noncommutativity. These results make us feel anomalous, because the instanton number of course exists for the instanton in the commutative space but $|v_0\rangle$ zero-modes or some counterparts of them do not exist in the commutative space. From these observations, we might wonder if there is a deep disconnection between commutative instantons and NC instantons. To clarify the connection between the NC instantons and commutative instantons, let us consider the smooth NC deformation from the commutative instanton in the next section.

Propagators and the Index Theorems

The zero-modes of the Dirac operator in the ADHM instanton background are studied in [33]. They show that the Atiyah-Singer index of the Dirac operator is equal to the instanton number. In [34], Green functions are constructed for a field in an arbitrary representation of gauge group propagating in NC ADHM instanton backgrounds.

Other Kinds of Solutions

We have reviewed the ADHM method. There are some other methods to construct NC instantons.

In [35], Lechtenfeld and Popov study the NC generalization of 't Hooft's multi-instanton configurations for the $U(2)$ gauge group. They solve the problem in the naive application of Nekrasov and Schwarz method to the 't Hooft instanton solution. The problem originates from the appearance of a source term in the equation in the Corrigan-Fairlie-'t Hooft-Wilczek ansatz. They generalize the method of [36] to naive NC multi-instantons.

In [37], Horváth et al. use the method of dressing transformations, an iterative procedure for generating solutions from a given solution, and they generalize Belavin and Zakharov method to the NC case.

In [38], Hamanaka and Terashima construct NC instantons by using the solution generating technique introduced by Harvey et al. [39].

More details and an embracive list including other kinds of NC space and other kinds of BPS states are found in [40] for example.

Another approach that is smooth deformation of commutative instanton is given in the last few years. We will see it in the next section.

3. Smooth NC Deformation of Instantons

In this section, we construct NC instantons deformed smoothly from commutative instantons, and we study their natures.

We define NC deformations by formal expansions in a deformation parameter $\hbar$. So, let us pay attention to the mathematical meaning of the formal expansion. We introduce our star products by using formal expansions in $\hbar$, as we will see soon. Such products are not closed in the set of all smooth functions in general, so one of the simple ways to define the star products is using formal expansion. The star product is defined by putting some conditions on each order of $\hbar$ expansion to be a smooth bounded function or a square integrable function and so on. Therefore, we have to check their conditions for all quantities represented by using the star product. Someone might wonder how can we manage such difficulties when the Fock space formalism is used. The Fock space formalism itself is regarded as a formal expansion by complex coordinates of $\mathbb{C}^2 \cong \mathbb{R}^4$. For example, an integrable condition of a function in the star product formulation is replaced by a convergence of the corresponding series. Space integrations are replaced by the trace operations (2.10). When we estimate topological charges like instanton charges by mathematically rigorous calculation, we have to use the Stokes' like theorem in the Fock space, as mentioned in Section 2.4. Therefore, the complexities of calculations are essentially same as the ones in star product formalism. One of the merits of using the star product formalism is that it does not require some specific representation. In calculations in the operator formalism, we have to introduce some basis like the Fock basis, but in the star product formalism, we can obtain physical values without introducing any representation.

3.1. Smooth NC Deformations

In this section, to easy understand that NC instantons smoothly connect into commutative instantons, we use a star product formulation. In the previous section, we use an operator formalism. Formally, there is a one-to-one correspondence between the operator formalism and the star product formalism, and the Weyl-transformation connects them with each other. Commutation relations of coordinates are given by

$$[x^\mu, x^\nu]_\star = x^\mu \star x^\nu - x^\nu \star x^\mu = i\theta^{\mu\nu}, \quad \mu, \nu = 1, 2, \ldots, 4, \tag{3.1}$$

where $(\theta^{\mu\nu})$ are a real, x-independent, skew-symmetric matrix entries, called the NC parameters. $\star$ is known as the Moyal product [41]. The Moyal product (or star product) is defined on functions by

$$f(x) \star g(x) := f(x) \exp\left(\frac{i}{2}\overleftarrow{\partial}_\mu \theta^{\mu\nu} \overrightarrow{\partial}_\nu\right) g(x). \tag{3.2}$$

Here $\overleftarrow{\partial}_\mu$ and $\overrightarrow{\partial}_\nu$ are partial derivatives with respect to x^μ for $f(x)$ and to x^ν for $g(x)$, respectively.

The curvature two form F is defined by $F := (1/2)F_{\mu\nu} dx^\mu \wedge_\star dx^\nu = dA + A \wedge_\star A$, where $\wedge_\star$ is defined by $A \wedge_\star A := (1/2)(A_\mu \star A_\nu) dx^\mu \wedge dx^\nu$.

To consider smooth NC deformations, we introduce a parameter $\hbar$ and a fixed constant $\theta_0^{\mu\nu} < \infty$ with $\theta^{\mu\nu} = \hbar \theta_0^{\mu\nu}$. We define the commutative limit by letting $\hbar \to 0$.

Formally we expand the connection as

$$A_\mu = \sum_{l=0}^{\infty} A_\mu^{(l)} \hbar^l. \tag{3.3}$$

Then,

$$\begin{gathered} A_\mu \star A_\nu = \sum_{l,m,n=0}^{\infty} \hbar^{l+m+n} \frac{1}{l!} A_\mu^{(m)} \left(\overline{\Delta}\right)^l A_\mu^{(n)}, \\ \overline{\Delta} \equiv \frac{i}{2} \overleftarrow{\partial}_\mu \theta_0^{\mu\nu} \overrightarrow{\partial}_v. \end{gathered} \tag{3.4}$$

We introduce the self-dual projection operator P by

$$P := \frac{1 + *}{2}; \qquad P_{\mu\nu,\rho\tau} = \frac{1}{4}\left(\delta_{\mu\rho}\delta_{\nu\tau} - \delta_{\nu\rho}\delta_{\mu\tau} + \epsilon_{\mu\nu\rho\tau}\right). \tag{3.5}$$

Then the instanton equation is given as

$$P_{\mu\nu,\rho\tau}F^{\rho\tau} = 0. \tag{3.6}$$

In the NC case, the lth order equation of (3.6) is given by

$$\begin{aligned}
&P^{\mu\nu,\rho\tau}\left(\partial_\rho A_\tau^{(l)} - \partial_\tau A_\rho^{(l)} + i\left[A_\rho^{(l)}, A_\tau^{(0)}\right] + i\left[A_\rho^{(0)}, A_\tau^{(l)}\right] + C_{\rho\tau}^{(l)}\right) = 0,\\
&C_{\rho\tau}^{(l)} := \sum_{(p;m,n)\in I(l)} \hbar^{p+m+n}\frac{1}{p!}\left(A_\rho^{(m)}\left(\overleftarrow{\Delta}\right)^p A_\tau^{(n)} - A_\tau^{(m)}\left(\overleftarrow{\Delta}\right)^p A_\rho^{(n)}\right),\\
&I(l) \equiv \left\{(p;m,n)\in\mathbb{Z}^3 \mid p+m+n=l,\ p,m,n\geq 0,\ m\neq l,\ n\neq l\right\}.
\end{aligned} \tag{3.7}$$

Note that the 0th order is the commutative instanton equation with solution $A_\mu^{(0)}$ being a commutative instanton. The asymptotic behavior of commutative instanton $A_\mu^{(0)}$ is given by

$$A_\mu^{(0)} = gdg^{-1} + O\left(|x|^{-2}\right), \quad gdg^{-1} = O\left(|x|^{-1}\right), \tag{3.8}$$

where $g\in G$ and G is a gauge group. (See, e.g., [2].) We introduce covariant derivatives associated to the commutative instanton connection by

$$D_\mu^{(0)} f := \partial_\mu f + i\left[A_\mu^{(0)}, f\right], \qquad D_{A^{(0)}} f := df + A^{(0)}\wedge f. \tag{3.9}$$

Using this, (3.7) is given by

$$P^{\mu\nu,\rho\tau}\left(D_\rho^{(0)}A_\tau^{(l)} - D_\tau^{(0)}A_\rho^{(l)} + C_{\rho\tau}^{(l)}\right) = 0. \tag{3.10}$$

In the following, we fix a commutative instanton connection $A^{(0)}$. We impose the following gauge fixing condition for $A^{(l)}$ $(l\geq 1)$ [18, 42]

$$A - A^{(0)} = D_{A^{(0)}}^* B, \quad B\in\Omega_+^2, \tag{3.11}$$

where $D_{A^{(0)}}^*$ is defined by

$$\begin{aligned}
\left(D_{A^{(0)}}^*\right)_\rho^{\mu\nu} B_{\mu\nu} &= \delta_\rho^\nu \partial^\mu B_{\mu\nu} - \delta_\rho^\mu \partial^\nu B_{\mu\nu} + i\delta_\rho^\nu\left[A^\mu, B_{\mu\nu}\right] - \delta_\rho^\mu\left[A^\nu, B_{\mu\nu}\right]\\
&= \delta_\rho^\nu D^{(0)\mu} B_{\mu\nu} - \delta_\rho^\mu D^{(0)\nu} B_{\mu\nu}.
\end{aligned} \tag{3.12}$$

We expand B in $\hbar$ as we did with A. Then $A^{(l)} = D_{A^{(0)}}^* B^{(l)}$. In this gauge, using the fact that the $A^{(0)}$ is an anti-self-dual connection, (3.10) simplified to

$$2D_{(0)}^2 B^{(l)\mu\nu} + P^{\mu\nu,\rho\tau}C_{\rho\tau}^{(l)} = 0, \tag{3.13}$$

where

$$D_{(0)}^2 \equiv D_{A^{(0)}}^{\rho} D_{A^{(0)}\rho}. \tag{3.14}$$

We consider the Green's function for $D_{(0)}^2$:

$$D_{(0)}^2 G_0(x,y) = \delta(x-y), \tag{3.15}$$

where $\delta(x-y)$ is a four-dimensional delta function. $G_0(x,y)$ has been constructed in [43] (see also [44, 45]). Using the Green's function, we solve (3.13) as

$$B^{(l)\mu\nu} = -\frac{1}{2}\int_{\mathbb{R}^4} G_0(x,y) P^{\mu\nu,\rho\tau} C_{\rho\tau}^{(l)}(y) d^4 y \tag{3.16}$$

and the NC instanton $A = \sum A^{(l)}\hbar^l$ is given by

$$A^{(l)} = D_{A^{(0)}}^{*} B^{(l)}. \tag{3.17}$$

In the following, we call NC instantons smoothly deformed from commutative instantons SNCD instantons. The asymptotic behavior of Green's function of $D_{(0)}^2$ is important, which is given by

$$G_0(x,y) = O\left(|x-y|^{-2}\right). \tag{3.18}$$

We introduce the notation $O'(|x|^{-m})$ as in [2]. If s is a function of $\mathbb{R}^4$ which is $O(|x|^{-m})$ as $|x| \to \infty$ and $|D_{(0)}^k s| = O(|x|^{-m-k})$, then we denote this natural growth condition by $s = O'(|x|^{-m})$.

Theorem 3.1. *If $C^{(l)} = O'(|x|^{-4})$, then $B^{(k)} = O'(|x|^{-2})$.*

We gave a proof of this theorem in [18].

In our case, $C_{\rho\tau}^{(1)} = O'(x^{-4})$ by (3.8), and so $B^{(1)} = O'(|x|^{-2})$, $A^{(1)} = O'(|x|^{-3})$ as $A^{(l)} = D_{A^{(0)}}^{*} B^{(l)}$. Repeating the argument l times, we get

$$\left|A^{(l)}\right| < O'\left(|x|^{-3+\epsilon}\right), \quad \forall\epsilon > 0. \tag{3.19}$$

3.2. Instanton Charge

The instanton charge is defined by

$$Q_{\hbar} := \frac{1}{8\pi^2}\int \mathrm{tr}_{U(N)} F \wedge \star F. \tag{3.20}$$

We rewrite (3.20) as

$$\frac{1}{8\pi^2}\int \mathrm{tr}_{U(N)} d\left(A\wedge\star dA+\frac{2}{3}A\wedge\star A\wedge\star A\right)+\frac{1}{8\pi^2}\int \mathrm{tr}_{U(N)} P_\star, \tag{3.21}$$

where

$$P_\star=\frac{1}{3}\{F\wedge\star A\wedge\star A+2A\wedge\star F\wedge\star A+A\wedge\star A\wedge\star F+A\wedge\star A\wedge\star A\wedge\star A\}. \tag{3.22}$$

$\int \mathrm{tr}_{U(N)} P_\star$ is 0 in the commutative limit, but it does not vanish in NC space, because the cyclic symmetry of trace operation is broken by the NC deformation.

The terms in $\int \mathrm{tr}_{U(N)} P_\star$ are typically written as

$$\int_{\mathbb{R}^d} \mathrm{tr}_{U(N)}\left(P\wedge\star R-(-1)^{n(4-n)}R\wedge\star P\right), \tag{3.23}$$

where P and R are some n-form and $(4-n)$-form ($n=0,\dots,4$), respectively, and let $P\wedge R$ be $O(\hbar^k)$. The lowest order term in $\hbar$ vanishes because of the cyclic symmetry of the trace, that is, $\int \mathrm{tr}_{U(N)}(P\wedge R-(-1)^{n(4-n)}R\wedge P)=0$. The term of order $\hbar$ is given by

$$\begin{aligned}&\frac{i}{2}\int_{\mathbb{R}^4} \mathrm{tr}_{U(N)}\left\{\hbar\theta_0^{\mu\nu}(\partial_\mu P\wedge\partial_\nu R)\right\}\\&\quad=\frac{i}{2}\int_{\mathbb{R}^4}(n!(4-n)!)\epsilon^{\mu_1\mu_2\mu_3\mu_4}\mathrm{tr}_{U(N)} d\left\{(*\theta)\wedge\left(P_{\mu_1\cdots\mu_n}dR_{\mu_{n+1}\cdots\mu_4}\right)\right\},\end{aligned} \tag{3.24}$$

where $*\theta=\epsilon_{\mu\nu\rho\tau}\theta^{\rho\tau}dx^\mu\wedge dx^\nu/4$. These integrals are zero if $P_{\mu_1\cdots\mu_n}dR_{\mu_{n+1}\cdots\mu_4}$ is $O'(|x|^{-(4-1+\epsilon)})$ ($\epsilon>0$) and this condition is satisfied for SNCD instantons. Similarly, higher-order terms in $\hbar$ in (3.23) can be written as total divergences and hence vanish under the decay hypothesis. This fact and (3.19) imply that $\int \mathrm{tr}_{U(N)} P_\star=0$.

Because of the similar estimation, we found the other terms of $\int \mathrm{tr}_{U(N)} F\wedge\star F-\int \mathrm{tr}_{U(N)} F^{(0)}\wedge F^{(0)}$ vanish, where $F^{(0)}$ is the curvature two form associated to $A^{(0)}$.

Summarizing the above discussions, we get following theorems [18].

Theorem 3.2. *Let $A_\mu^{(0)}$ be a commutative instanton solution in $\mathbb{R}^4$. There exists a formal NC instanton solution $A_\mu=\sum_{l=0}^{\infty}A_\mu^{(l)}\hbar^l$ (SNCD instanton) such that the instanton number $Q_\hbar$ defined by* (3.20) *is independent of the NC parameter $\hbar$:*

$$\frac{1}{8\pi^2}\int \mathrm{tr}_{U(N)} F\wedge\star F=\frac{1}{8\pi^2}\int \mathrm{tr}_{U(N)} F^{(0)}\wedge F^{(0)}. \tag{3.25}$$

3.3. Index of the Dirac Operator and Green's Function

Dirac(-Weyl) operators $\mathfrak{D}_A : \Gamma(S^+ \otimes E) \to \Gamma(S^- \otimes E)$ and $\overline{\mathfrak{D}}_A : \Gamma(S^- \otimes E) \to \Gamma(S^+ \otimes E)$ are defined as

$$\mathfrak{D}_A := \sigma^\mu D_\mu, \qquad \overline{\mathfrak{D}}_A := \overline{\sigma}^\mu D_\mu^\dagger. \tag{3.26}$$

Here, σ_μ and $\overline{\sigma}_\mu$ are defined by (2.24). Consider $\hbar$ expansion of $\psi \in \Gamma(S^+ \otimes E)$ and $\overline{\psi} \in \Gamma(S^- \otimes E)$ as

$$\psi = \sum_{n=0}^{\infty} \hbar^n \psi^{(n)}, \qquad \overline{\psi} = \sum_{n=0}^{\infty} \hbar^n \overline{\psi}^{(n)}. \tag{3.27}$$

In [19], the zero-modes of $\mathfrak{D}_A$ and $\overline{\mathfrak{D}}_A$, which are defined by

$$\mathfrak{D}_A \star \psi = 0, \qquad \overline{\mathfrak{D}}_A \star \overline{\psi} = 0, \tag{3.28}$$

are investigated, and the following theorem is obtained.

Theorem 3.3. *Let $\mathfrak{D}_A$ and $\overline{\mathfrak{D}}_A$ be the Dirac(-Weyl) operators for an SNCD instanton background with its instanton number $-k$. There is no zero-mode for $\mathfrak{D}_A \star \psi = 0$, and there are k zero-modes for $\overline{\mathfrak{D}}_A \star \overline{\psi}_i = 0 \quad (i = 1, \ldots, k)$ that are given as*

$$\overline{\psi}_i = \sum_{n=0}^{\infty} \left(\sum_{j=1}^{k} a^j_{n,i} \eta_j \right) \hbar^n + O'\left(|x|^{-5+\epsilon}\right), \quad \eta_j = O'\left(|x|^{-3}\right), \tag{3.29}$$

where $a^j_{n,i}$ is a constant matrix and η_j is a base of the zero mode of $\overline{\mathfrak{D}}_A^{(0)}$.

Note that it is a well-known fact as an index theorem in commutative space that the dimension of $\ker \overline{\mathfrak{D}}_A^{(0)}$ is equal to k the instanton number (of opposite sign), and there exists k zero-mode η_i $(i = 1, 2, \ldots, k)$. Theorem 3.3 says that zero-modes deformed from the ones in commutative space are obtained, but there is no new zero-mode appearing. Then we get the following theorem [19].

Theorem 3.4. *If* $\mathrm{Ind}\, \not{D}^0 := \dim\ker \mathfrak{D}_A^{(0)} - \dim\ker \overline{\mathfrak{D}}_A^{(0)} = -k$, *then* $\mathrm{Ind}\, \not{D} := \dim\ker \mathfrak{D}_A - \dim\ker \overline{\mathfrak{D}}_A = -k$.

Next, we construct the Green's function of $\Delta_A \equiv D_\mu \star D^\mu$,

$$\Delta_A \star G_A(x, y) = \delta(x - y). \tag{3.30}$$

We expand (3.18) by $\hbar$, for $n > 0$, then $\hbar^n$ order equation is given as

$$\Delta_A^{(0)} G_A^{(n)}(x,y) + \left[\Delta_A \sum_{0\le k<n} \hbar^k G_A^{(k)}(x,y)\right]^{(n)} = 0, \tag{3.31}$$

where $G_A^{(n)}$ is defined by $G_A(x,y) = \sum_{k=0}^{\infty} G_A^{(k)}\hbar^k$. We solve them recursively

$$G_A^{(n)}(x,y) = \int d^4w G_A^{(0)}(x,w)\left[\Delta_A \sum_{0\le k<n} \hbar^k G_A^{(k)}(w,y)\right]^{(n)}. \tag{3.32}$$

Note that $G_A^{(0)}(x,w)$ was constructed in [43–45]. Using property of $G_A^{(0)}(x,w)$ and $A^{(n)}$, we obtain the following decay condition in [19]:

$$G_A^{(n)}(x,y) = O'\left(|x|^{-3}\right). \tag{3.33}$$

3.4. From an Instanton to the ADHM Equations

Let us see how to derive the ADHM equations from an SNCD instanton.

Let $\overline{\varphi}_i$ $(i = 1,\dots,k)$ be orthonormal zero-modes of $\overline{\mathfrak{D}}_A$ and $\overline{\varphi} = (\overline{\varphi}_i)$, which are introduced in Section 3.3.

At first we define T^μ by

$$T^\mu := \int_{\mathbb{R}^4} d^4x \frac{1}{2}\left(x^\mu \star \overline{\varphi}^\dagger \star \overline{\varphi} + \overline{\varphi}^\dagger \star \overline{\varphi} \star x^\mu\right). \tag{3.34}$$

Next we introduce an asymptotically parallel section $g^{-1}S$ of $S^+ \otimes E$ by

$$\widetilde{\varphi} = -\frac{g^{-1}Sx^\dagger}{|x|^4} + O'\left(|x|^{-4}\right), \tag{3.35}$$

where $x^\dagger := \overline{\sigma}_\mu x^\mu$ and $\widetilde{\varphi} :={}^t\overline{\varphi}\sigma_2$. This t means transposing spinor suffixes.

Using various properties and decay conditions of $A^{(n)}, G_A^{(n)}, \overline{\varphi}^{(n)}$, and theorems in the previous subsections, we finally obtain the following theorem.

Theorem 3.5. *Let A^μ be an SNCD instanton and $\overline{\varphi}$ the zero-mode of $\overline{\mathfrak{D}}_A$ determined by A^μ as in Section 3.3. Let T^μ, S be constant matrices defined by* (3.34) *and* (3.35)*, respectively. Then, they satisfy the ADHM equations:*

$$[T^\mu, T^\nu]^+ = \frac{1}{2}\operatorname{tr}\left(S^\dagger S \overline{\sigma}^{\mu\nu}\right) - i\theta^{\mu\nu+} 1_{k\times k}. \tag{3.36}$$

Here $\overline{\sigma}_{\mu\nu} := (1/4)(\overline{\sigma}_\mu \sigma_\nu - \overline{\sigma}_\nu \sigma_\mu)$ and $1_{k\times k}$ is an identity matrix.

Rough sketch of the proof

Let us see the essence of the proof. Let us introduce $\star_x$ as $\star$ associated with variable x. The completeness of $\overline{\varphi}(x)$ is written as

$$\star_x\overline{\varphi}(x)\overline{\varphi}^\dagger(y)\star_y = \star_x\delta(x-y)\star_y - \star_x\mathfrak{D}_A\star_x G_A(x,y)\star_y\overleftarrow{\overline{\mathfrak{D}}}_A\star_y. \tag{3.37}$$

From the definition of the T^μ,

$$T^\mu T^\nu = \int_{\mathbb{R}^4} d^4x \int_{\mathbb{R}^4} d^4y \Big(x^\mu\star_x\overline{\varphi}^\dagger(x)\star_x\overline{\varphi}(x)\Big)\Big(\overline{\varphi}^\dagger(y)\star_y\overline{\varphi}(y)\star_y y^\nu\Big). \tag{3.38}$$

Using Theorem 3.3, (3.37), and integration by parts, (3.38) becomes

$$\begin{aligned} T^\mu T^\nu = &\int_{\mathbb{R}^4} d^4x x^\mu \star \overline{\varphi}^\dagger \star \overline{\varphi} \star x^\nu \\ &+ \int_{S^3} dS_x^\rho \int_{\mathbb{R}^4} d^4y \Big(x^\mu\star_x\overline{\varphi}^\dagger(x)\sigma_\rho\Big)\star_x G_A(x,y)\star_y\overleftarrow{\overline{\mathfrak{D}}}_A\star_y\left(\overline{\varphi}(y)\star_y y^\nu\right) \\ &- \int_{\mathbb{R}^4} d^4x \int_{\mathbb{R}^4} d^4y \Big(\overline{\varphi}^\dagger(x)\sigma^\mu\Big)\star_x G_A(x,y)\star_y\overleftarrow{\overline{\mathfrak{D}}}_A\star_y\left(\overline{\varphi}(y)\star_y y^\nu\right), \end{aligned} \tag{3.39}$$

where $dS_x^\mu = |x|^2 x^\mu d\Omega$ and $d\Omega$ is the solid angle. The first term is deformed as follows.

$$\begin{aligned} &\int_{\mathbb{R}^4} d^4x x^\mu \star \overline{\varphi}^\dagger \star \overline{\varphi} \star x^\nu \\ &\quad = \int_{\mathbb{R}^4} d^4x \Big(\overline{\varphi}^\dagger \star \overline{\varphi} \star x^\nu \star x^\mu + \left[x^\mu, \overline{\varphi}^\dagger \star \overline{\varphi}\right]_\star \star x^\nu + \overline{\varphi}^\dagger \star \overline{\varphi} \star [x^\mu, x^\nu]_\star\Big) \\ &\quad = \int_{\mathbb{R}^4} d^4x \Big(\overline{\varphi}^\dagger \star \overline{\varphi} \star x^\nu \star x^\mu + i\theta^{\mu\rho}\partial_\rho\Big(\overline{\varphi}^\dagger \star \overline{\varphi}\Big) \star x^\nu + i\theta^{\mu\nu}\overline{\varphi}^\dagger \star \overline{\varphi}\Big) \\ &\quad = \int_{\mathbb{R}^4} d^4x \overline{\varphi}^\dagger \star \overline{\varphi} \star x^\nu \star x^\mu. \end{aligned} \tag{3.40}$$

Here $\overline{\psi} = O'(|x|^{-3})$ given in Theorem 3.3 is used in the third equality. By integration by parts again, we get

$$T^{\mu}T^{\nu} = \int_{\mathbb{R}^4} d^4x \overline{\psi}^{\dagger} \star \overline{\psi} \star x^{\nu} \star x^{\mu} \tag{3.41}$$

$$+ \int_{S^3} dS_x^{\rho} \int_{S^3} dS_y^{\tau} \left(x^{\mu} \star_x \overline{\psi}^{\dagger}(x) \sigma_{\rho} \right) \star_x G_A(x,y) \star_y (\overline{\sigma}_{\tau} \overline{\psi}(y) \star_y y^{\nu}) \tag{3.42}$$

$$- \int_{S^3} dS_x^{\rho} \int_{\mathbb{R}^4} d^4y \left(x^{\mu} \star_x \overline{\psi}^{\dagger}(x) \sigma_{\rho} \right) \star_x G_A(x,y) \star_y (\overline{\sigma}^{\nu} \overline{\psi}(y)) \tag{3.43}$$

$$- \int_{\mathbb{R}^4} d^4x \int_{S^3} dS_y^{\tau} \left(\overline{\psi}^{\dagger}(x) \sigma^{\mu} \right) \star_x G_A(x,y) \star_y (\overline{\sigma}^{\tau} \overline{\psi}(y) \star_y y^{\nu}) \tag{3.44}$$

$$+ \int_{\mathbb{R}^4} d^4x \int_{\mathbb{R}^4} d^4y \left(\overline{\psi}^{\dagger}(x) \sigma^{\mu} \right) \star_x G_A(x,y) \star_y (\overline{\sigma}^{\nu} \overline{\psi}(y)). \tag{3.45}$$

Equations (3.42) and (3.44) vanish when $R_y \to \infty$, where R_y is a radius of S_y^3. Equation (3.45) will vanish on the self-dual projection $[T^{\mu}, T^{\nu}]^{+} := P^{\mu\nu,\rho\tau}[T_{\rho}, T_{\tau}]$, because $\sigma^{\mu}\overline{\sigma}^{\nu} - \sigma^{\nu}\overline{\sigma}^{\mu}$ is anti-self-dual with respect to the μ, ν. Thus only (3.41) and (3.43) remain. By the asymptotic behaviors of $\overline{\psi}$ and some calculations, we can prove that (3.43) becomes

$$\frac{1}{8} \operatorname{tr}\left(S^{\dagger} S \overline{\sigma}^{\mu} \sigma^{\nu} \right), \tag{3.46}$$

where the trace tr is taken with respect to the spinor indices. In the $[T^{\mu}, T^{\nu}]^{+}$ combination, (3.41) becomes $-i\theta^{\mu\nu +} = -iP^{\mu\nu,\rho\tau}\theta^{\rho\tau}$. Therefore, we get (3.36). The complete proof is given in [19].

These ADHM equations (3.36) are coincident with the ones provided by Nekrasov and Schwarz [4]. After identification of

$$S^{\dagger} = \begin{pmatrix} I \\ J^{\dagger} \end{pmatrix}, \qquad T^{\mu}\overline{\sigma}_{\mu} = \begin{pmatrix} -B_2 & -B_1 \\ B_1^{\dagger} & -B_2^{\dagger} \end{pmatrix}, \tag{3.47}$$

and setting the NC parameter as in (2.2), we find that (3.36) is identified with (2.13) and (2.14).

Similar to the commutative case, we obtain the following theorem.

Theorem 3.6. *There is a one-to-one correspondence between ADHM data satisfying* (3.36) *and SNCD instantons in NC* $\mathbb{R}^4$.

The proof is given in [19].

4. Smooth NC Deformation of Vortexes

In the previous section, we investigate the smooth deformation of instantons. This method is applicable to gauge theories in other dimensions. In this section we study NC deformation of the vortex solutions [46, 47]. We consider the Abelian-Higgs model in commutative $\mathbb{R}^2$ and deform the Taubes' vortex solutions into NC vortexes [48].

Let coordinates of NC Euclidean space $\mathbb{R}^2$ be $x^\mu,\ \mu = 1,2$, with commutation relations

$$[x^\mu, x^\nu]_\star = i\hbar\epsilon^{\mu\nu} \quad (\mu,\nu = 1,2), \tag{4.1}$$

where $\epsilon^{\mu\nu} = -\epsilon^{\nu\mu} \quad (\epsilon^{12} = 1)$ is an antisymmetric tensor.

The curvature components of the connection A are given by

$$\begin{gathered} F_{zz} = F_{\overline{z}\,\overline{z}} = 0, \\ F_{z\overline{z}} = iF_{12} = \partial_z A_{\overline{z}} - \partial_{\overline{z}} A_z - i[A_z, A_{\overline{z}}]_\star =: iB. \end{gathered} \tag{4.2}$$

Using these complex coordinates, the covariant derivatives of the Higgs fields are

$$\begin{aligned} D\star\phi &= (\partial - iA)\star\phi, \qquad & \overline{D}\star\phi &= \left(\overline{\partial} - i\overline{A}\right)\star\phi, \\ D\star\overline{\phi} &= \partial\overline{\phi} + i\overline{\phi}\star A, \qquad & \overline{D}\star\overline{\phi} &= \overline{\partial}\,\overline{\phi} + i\overline{\phi}\star\overline{A}. \end{aligned} \tag{4.3}$$

The vortex equations are defined by

$$\overline{D}\star\phi = \left(\overline{\partial} - i\overline{A}\right)\star\phi = 0, \qquad B + \phi\star\overline{\phi} - 1 = 0. \tag{4.4}$$

We call solutions of these equations NC vortexes.

The formal expansions of the fields are

$$\phi = \sum_{n=0}^{\infty}\hbar^n\phi_n(z,\overline{z}), \qquad A = \sum_{n=0}^{\infty}\hbar^n A_n(z,\overline{z}). \tag{4.5}$$

The kth order equations for (4.4) are

$$-i\left(\partial\overline{A}_k + \overline{\partial}A_k\right) + \phi_k\overline{\phi}_0 + \phi_0\overline{\phi}_k - \delta_{k0} + C_k(z,\overline{z}) = 0, \tag{4.6}$$

$$\overline{\partial}\phi_k - i\overline{A}_k\phi_0 - i\overline{A}_0\phi_k + D_k(z,\overline{z}) = 0. \tag{4.7}$$

Here $C_k(z,\overline{z})$ is the coefficient of $\hbar^k$ in $-[A,\overline{A}]_\star + \phi\star\overline{\phi} - (\phi_k\overline{\phi}_0 + \phi_0\overline{\phi}_k)$, so $C_k(z,\overline{z})$ is a function of $\{A_i, \overline{A}_j, \phi_m, \overline{\phi}_n \mid 0 \le i,j,m,n \le k-1\}$. Similarly, $D_k(z,\overline{z})$ is the coefficient of $\hbar^k$ in $-i\overline{A}\star\phi - (-i\overline{A}_k\phi_0 - i\overline{A}_0\phi_k)$ and a function of $\{A_i, \overline{A}_j, \phi_m, \overline{\phi}_n \mid 0 \le i,j,m,n \le k-1\}$.

In the case of $k = 0$, (4.6) and (4.7) coincide with the commutative $U(1)$ vortex equations $\overline{D}\phi_0 = (\overline{\partial} - i\overline{A}_0)\phi_0 = 0$ and $B_0 + \phi_0\overline{\phi}_0 - 1 = 0$, where $B_0 = -i(\partial\overline{A}_0 - \overline{\partial}A_0)$. In the following, we consider the case that $A_0, \overline{A}_0$, and ϕ_0 are smooth finite vortex solutions. We call it Taubes' vortex solution.

In the region $\phi_0 \neq 0$, substituting (4.7) into (4.6) for A_k and $\overline{A}_k$, we get

$$\begin{aligned}&\left\{\frac{\partial\phi_0}{\phi_0^2}\left(\overline{\partial}\phi_k - i\overline{A}_0\phi_k + D_k\right) - \frac{1}{\phi_0}\left(\Delta\phi_k - i\partial\overline{A}_0\phi_k - i\overline{A}_0\partial\phi_k + \partial D_k\right)\right\} + \{\text{c.c.}\}\\&\quad + \phi_k\overline{\phi}_0 + \phi_0\overline{\phi}_0 - \delta_{k0} + C_k = 0.\end{aligned} \tag{4.8}$$

Here {c.c.} is the complex conjugate of preceding terms and $\Delta = \partial\overline{\partial}$.

Setting

$$\varphi_k = \frac{\phi_k}{\phi_0} + \frac{\overline{\phi}_k}{\overline{\phi}_0} = 2\,\mathrm{Re}\left(\frac{\phi_k}{\phi_0}\right), \qquad d_k = \frac{D_k}{\phi_0}. \tag{4.9}$$

Equation (4.8) is simplified to

$$\left(-\Delta + |\phi_0|^2\right)\varphi_k = E_k, \tag{4.10}$$

where

$$E_k := -C_k + \partial d_k - \overline{\partial}\,\overline{d}_k. \tag{4.11}$$

To show that there exists a unique NC vortex solution deformed from the Taubes' vortex solution, we consider the stationary Schrödinger equation

$$(-\Delta + V(x))u(x) = f(x) \tag{4.12}$$

in $\mathbb{R}^2$, where $V(x)$ is a real-valued C^∞ function. We impose the following assumptions for $V(x)$.

- $(a1)$ $V(x) \geq 0$, for all $x \subset \mathbb{R}^2$.
- $(a2)$ There exist $K \subset \mathbb{R}^2$ and $\exists c > 0$ such that K is a compact set and for $x \in \mathbb{R}^2 \setminus K$, $V(x) \geq c$.
- $(a3)$ There exist $x_1, \ldots, x_N \in \mathbb{R}^2$ such that $V(x_i) = 0$ and $V(x) > 0$ for $x \notin \{x_1, \ldots, x_N\}$.
- $(a4)$ For any $\alpha = (\alpha_1, \alpha_2) \in \mathbb{Z}_+^2$, There exists a positive constant c such that $|\partial_x^\alpha(V - c)| \leq c$ for any $x \in \mathbb{R}^2$.

Note that the system (4.10) satisfies the assumptions $(a1)$–$(a4)$. We set

$$H_l(n) := \left\{ f \mid \|f\| := \sup_{x\in\mathbb{R}^2}\left(1 + |x|^n\right)\left|\partial_x^\alpha f(x)\right| < \infty \text{ for any } |\alpha| \leq l \right] \tag{4.13}$$

for $n \in \mathbb{Z}_+$. Then we obtain the following theorem.

Theorem 4.1. *Under the assumptions* $(a1)$–$(a4)$, *there exists a unique solution* $u \in H_l(n)$ *of* (4.12) *for any* $f \in H_l(n)$.

This theorem's proof is given by using standard techniques of Green's function [20].

Equation (4.10) is a particular example of (4.12). Theorem 4.1 and some asymptotic analysis derive the following theorem.

Theorem 4.2. *Let* A_0 *and* ϕ_0 *be a Taubes' vortex solution, in other words,* (A_0, ϕ_0) *is a finite and smooth solution of the commutative vortex equations. Then there exists a unique solution* (A, ϕ) *of the NC vortex equations* (4.4) *with* $A|_{\hbar=0} = A_0$, $\phi|_{\hbar=0} = \phi_0$, *and its vortex number is preserved*

$$\frac{1}{2\pi}\int d^2 x B = \frac{1}{2\pi}\int d^2 x B_0. \tag{4.14}$$

The proof is given in [20].

5. Conclusions

We have reviewed developments for the last dozen years in NC instantons in $\mathbb{R}^4$. The ADHM methods made great progress and broke ground to make strict solutions of the NC soliton equations. A lot of kinds of NC instanton solutions have been made by the ADHM method. Using the solutions and ADHM data, many aspects have been investigated. For example, topological charges, Dirac zero-modes, index theorems, and Green's functions in the NC ADHM instanton backgrounds. However, we could not understand the relation with commutative instantons and how instantons deform from commutative ones rigorously. In recent few years, the smooth NC deformation method has been investigated. For the smooth NC deformed instantons, many features are clarified. For example, the instanton charge, the number of the spinor zero-modes, and the index of the Dirac operator in the NC deformed instanton backgrounds coincide with the ones in commutative instanton backgrounds. The ADHM equations are derived from the NC deformed instantons and we find the ADHM equations coincide with the ones by Nekrasov and Schwarz. A one-to-one correspondence between smooth NC deformed instantons and the ADHM data are also obtained. Thus, about instantons in NC $\mathbb{R}^4$, a lot of features have been investigated. The smooth NC deformation method is useful for other dimensional gauge theories. As an example, smooth deformations of vortexes are studied similarly. Their vortex numbers also coincide with the ones in commutative $\mathbb{R}^2$.

We have considered gauge theories in $\mathbb{R}^n$. One of the essences to prove some theorems of NC instantons or NC vortexes is in infinity of size of the space. So, some of the theorems are changed when we consider finite size spaces. For example, topological charges are deformed under NC deformations of the spaces and they depend on the NC parameters in general [18]. The generic investigations of such changes from the point of view of smooth NC deformations are left for future subjects. Most NC instantons or some other solitons in gauge theories are still in deep mist.

Acknowledgment

The author was supported by KAKENHI (20740049, Grant-in-Aid for Young Scientists (B)).

References

[1] M. A. Shifman, *Instantons in Gauge Theories*, World Scientific, Singapore, 1994.

[2] S. K. Donaldson and P. B. Kronheimer, *The Geometry of Four-Manifolds*, Oxford Mathematical Monographs, The Clarendon Press/Oxford University Press, New York, NY, USA, 1990.

[3] N. Nekrasov, *The Physics and Mathematics of Instantons*, International Series of Monographs on Physics, Oxford University Press, Oxford, UK.

[4] N. Nekrasov and A. S. Schwarz, "Instantons on noncommutative R^4, and (2,0) superconformal six-dimensional theory," *Communications in Mathematical Physics*, vol. 198, no. 3, pp. 689–703, 1998.

[5] M. F. Atiyah, N. J. Hitchin, V. G. Drinfeld, and Yu. I. Manin, "Construction of instantons," *Physics Letters A*, vol. 65, no. 3, pp. 185–187, 1978.

[6] K.-Y. Kim, B.-H. Lee, and H. S. Yang, "Comments on instantons on noncommutative $\mathbb{R}^4$," *Journal of the Korean Physical Society*, vol. 41, no. 3, pp. 290–297, 2002.

[7] N. A. Nekrasov, "Noncommutative instantons revisited," *Communications in Mathematical Physics*, vol. 241, no. 1, pp. 143–160, 2003.

[8] K. Furuuchi, "Dp-D($p+4$) in noncommutative Yang-Mills," *Journal of High Energy Physics*, vol. 2001, no. 03, article 033, 2001.

[9] D. H. Correa, G. Lozano, E. F. Moreno, and F. A. Schaposnik, "Comments on the U noncommutative instanton," *Physics Letters B*, vol. 515, no. 1-2, pp. 206–212, 2001.

[10] O. Lechtenfeld and A. D. Popov, "Noncommutative multi-solitons in 2+1 dimensions," *Journal of High Energy Physics*, vol. 2001, no. 11, article 040, 2001.

[11] S. Parvizi, "Noncommutative instantons and the information metric," *Modern Physics Letters A*, vol. 17, no. 6, pp. 341–353, 2002.

[12] Nikita A. Nekrasov, "Lectures on open strings, and noncommutative gauge fields ," Les Houches 2001, Gravity, gauge theories and strings 477–495, http://arxiv.org/abs/hep-th/0203109.

[13] Y. Tian and C.-J. Zhu, "Instantons on general noncommutative $\mathbb{R}^4$," *Communications in theoretical physics*, vol. 38, no. 6, pp. 691–697, 2002.

[14] D. H. Correa, E. F. Moreno, and F. A. Schaposnik, "Some noncommutative multi-instantons from vortices in curved space," *Physics Letters B*, vol. 543, no. 3-4, pp. 235–240, 2002.

[15] F. Franco-Sollova and T. A. Ivanova, "On noncommutative merons and instantons," *Journal of Physics A*, vol. 36, no. 14, pp. 4207–4219, 2003.

[16] Y. Tian and C.-J. Zhu, "Remarks on the noncommutative Atiyah-Drinfeld-Hitchin-Manin construction," *Physical Review D*, vol. 67, no. 4, Article ID 045016, 9 pages, 2003.

[17] M. Hamanaka, "Noncommutative solitons and D-branes ," http://arxiv.org/abs/hep-th/0303256.

[18] Y. Maeda and A. Sako, "Noncommutative deformation of instantons," *Journal of Geometry and Physics*, vol. 58, no. 12, pp. 1784–1791, 2008.

[19] Y. Maeda and A. Sako, "Noncommutative deformation of spinor zero mode and ADHM construction ," http://arxiv.org/abs/0910.3441.

[20] Y. Maeda and A. Sako, "Are vortex numbers preserved?" *Journal of Geometry and Physics*, vol. 58, no. 8, pp. 967–978, 2008.

[21] A. Sako, "Noncommutative deformation of instantons and vortexes," *Journal of Geometry and Symmetry in Physics*, vol. 14, pp. 85–96, 2009.

[22] N. A. Nekrasov, "Noncommutative instantons revisited," *Communications in Mathematical Physics*, vol. 241, no. 1, pp. 143–160, 2003.

[23] N. A. Nekrasov, "Trieste lectures on solitons in noncommutative gauge theories," http://arxiv.org/abs/hep-th/0011095.

[24] D. J. Gross and N. A. Nekrasov, "Solitons in noncommutative gauge theory," *Journal of High Energy Physics*, vol. 2001, no. 03, article 044, 2001.

[25] T. Ishikawa, S. Kuroki, and A. Sako, "Instanton number on noncommutative $\mathbb{R}^4$," *Journal of High Energy Physics*, vol. 2002, no. 08, article 028, 2002.

[26] A. Sako, "Instanton number of noncommutative U(n) gauge theory," *Journal of High Energy Physics*, no. 4, article 023, 2003.

[27] K. Furuuchi, "Instantons on noncommutative $\mathbb{R}^4$ and projection operators," *Progress of Theoretical Physics*, vol. 103, no. 5, pp. 1043–1068, 2000.

[28] K. Furuuchi, ""Topological" charge of U(1) instantons on noncommutative $\mathbb{R}^4$," *Progress of Theoretical Physics Supplement*, no. 144, pp. 79–91, 2001.

[29] K. Furuuchi, "Equivalence of projections as gauge equivalence on noncommutative space," *Communications in Mathematical Physics*, vol. 217, no. 3, pp. 579–593, 2001.

[30] Y. Tian, C.-J. Zhu, and X.-C. Song, "Topological charge of noncommutative ADHM instanton," *Modern Physics Letters A*, vol. 18, no. 24, pp. 1691–1703, 2003.

[31] T. Ishikawa, S. I. Kuroki, and A. Sako, "Elongated U(1) instantons on noncommutative $\mathbb{R}^4$," *Journal of High Energy Physics*, vol. 2001, no. 11, article 068, 2001.

[32] A. Sako, "Supersymmetric gauge theory and geometry (Nonperturbative approach)," Nippon-Hyoron-sha, 2007.

[33] K.-Y. Kim, B.-H. Lee, and H. S. Yang, "Zero modes and the Atiyah-Singer index in noncommutative instantons," *Physical Review D*, vol. 66, no. 2, Article ID 025034, 11 pages, 2002.

[34] B.-H. Lee and H. S. Yang, "Propagators in noncommutative instantons," *Physical Review D*, vol. 66, no. 4, Article ID 045027, 14 pages, 2002.

[35] O. Lechtenfeld and A. D. Popov, "Noncommutative 't Hooft instantons," *Journal of High Energy Physics*, vol. 2002, no. 03, article 040, 2002.

[36] D. H. Correa, G. Lozano, E. F. Moreno, and F. A. Schaposnik, "Comments on the $U(2)$ noncommutative instanton," *Physics Letters B*, vol. 515, no. 1-2, pp. 206–212, 2001.

[37] Z. Horváth, O. Lechtenfeld, and M. Wolf, "Noncommutative instantons via dressing and splitting approaches," *Journal of High Energy Physics*, no. 12, article 060, 2002.

[38] M. Hamanaka and S. Terashima, "On exact noncommutative BPS solitons," *Journal of High Energy Physics*, no. 3, article 034, 2001.

[39] J. A. Harvey, P. Kraus, and F. Larsen, "Exact noncommutative solitons," *Journal of High Energy Physics*, no. 12, aricle 024, 2000.

[40] M. Hamanaka, "Noncommutative solitons and integrable systems," http://arxiv.org/abs/hep-th/0504001.

[41] J. E. Moyal, "Quantum mechanics as a statistical theory," *Proceedings of the Cambridge Philosophical Society*, vol. 45, pp. 99–124, 1949.

[42] D. S. Freed and K. K. Uhlenbeck, *Instantons and Four-Manifolds*, vol. 1 of *Mathematical Sciences Research Institute Publications*, Springer-Verlag, New York, NY, USA, 1984.

[43] E. Corrigan, P. Goddard, and S. Templeton, "Instanton Green's functions and tensor products," *Nuclear Physics B*, vol. 151, no. 1, pp. 93–117, 1979.

[44] E. Corrigan, D. B. Fairlie, S. Templeton, and P. Goddard, "A green's function for the general selfdual gauge field," *Nuclear Physics B*, vol. 140, no. 31, 1978.

[45] N. H. Christ, E. J. Weinberg, and N. K. Stanton, "General self-dual Yang-Mills solutions," *Physical Review D*, vol. 18, no. 6, pp. 2013–2025, 1978.

[46] V. L. Ginzburg and L. D. Landau, "On the theory of superconductivity," *Zhurnal Eksperimental'noi i Teoreticheskoi Fiziki*, vol. 20, pp. 1064–1082, 1950.

[47] L. D. Landau and D. Ter Haar, *English Translation Men of Physics*, vol. 138, Pergamon Oxford, Oxford, UK, 1965.

[48] A. Jaffe and C. Taubes, *Vortices and Monopoles: Structure of Static Gauge Theories*, vol. 2 of *Progress in Physics*, Birkhäauser, Boston, Mass, USA, 1980.

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 146719, 17 pages
doi:10.1155/2010/146719

Research Article

A Lie Algebroid on the Wiener Space

Rémi Léandre

Institut de Mathématiques, Université de Bourgogne, 21000 Dijon, France

Correspondence should be addressed to Rémi Léandre, remi.leandre@u-bourgogne.fr

Received 8 September 2009; Accepted 12 January 2010

Academic Editor: M. N. Hounkonnou

We define a Lie algebroid on the space of smooth 1-forms in the Nualart-Pardoux sense on the Wiener space associated to the stochastic linear Poisson structure on the Wiener space defined Léandre (2009).

1. Introduction

Infinite dimensional Poisson structures play a big role in the theory of infinite dimensional Lie algebras [1], in the theory of integrable system [2], and in field theory [3]. But for instance, in [2], the test functional space where the hydrodynamic Poisson structure acts continuously is not conveniently defined. In [4, 5] we have defined such a test functional space in the case of a linear Poisson bracket of hydrodynamic type. On the other hand, it is very well known [6] that the theories of Lie groupoids and Lie algebroids play a key role in Poisson geometry. It is interesting to study a Lie algebroid for the Poisson structure [4] defined analytically in the framework of [4]. We postpone until later the study the Lie groupoid associated to the same Poisson structure but in the algebraic framework of [5]. The definition of this Lie groupoid in the framework of [4] presents, namely, some difficulties. Moreover some deformation quantizations for symplectic structures in infinite dimensional analysis were recently performed (see the review of Léandre [7] on that). The theory of groupoids is related [8] to Kontsevich deformation quantization [9].

Let us recall what a Lie algebroid is [6, 10–13]. We consider a bundle E on a smooth finite dimensional manifold M. TM is the tangent bundle of M. $\Gamma^\infty(E)$ and $\Gamma^\infty(TM)$ denote the space of smooth section of E and TM. A Lie algebroid on E is given by the following data.

(i) A Lie bracket structure $[\cdot,\cdot]_E$ on $\Gamma^\infty(E)$ has in particular to satisfy the Jacobi relation

$$[[X_1, X_2]_E, X_3]_E + [[X_2, X_3]_E, X_1]_E + [[X_3, X_1]_E, X_2]_E = 0. \tag{1.1}$$

(ii) A smooth fiberwise linear map ρ_E, called the anchor map, from E into TM satisfies the relation

$$[X, fY]_E = f[X, Y]_E + \langle df, \rho_E(X) \rangle Y, \tag{1.2}$$

for any smooth sections X, Y of E and any element f of $C^\infty(M)$, the space of smooth functions on M.

Let us recall the definition of a Poisson structure on M. It is an antisymmetric $\mathbb{R}$-bilinear map $\{\cdot,\cdot\}$ from $C^\infty(M) \times C^\infty(M)$ into $C^\infty(M)$, which is a derivation on each components, vanishes on the constant and satisfies the Jacobi relation

$$\{\{f_1, f_2\}, f_3\} + \{\{f_2, f_3\}, f_1\} + \{\{f_3, f_1\}, f_2\} = 0. \tag{1.3}$$

A Poisson structure is given by a bivector π, that is, an element of $\Gamma^\infty(\Lambda^2 TM)$ as

$$\{f_1, f_2\} = \langle \pi, df_1, df_2 \rangle = i_{df_2} i_{df_1} \pi, \tag{1.4}$$

where df_1 and df_2 can be seen as 1-forms and the dual of $T_x^*(M)$ is $T_x(M)$. $i_\alpha \pi$ denotes the interior product of the bivector π by the 1-form α. If π is a bivector, then we can define a fiberwise linear smooth map from T^*M, the cotangent bundle of M, into TM, called $\tilde{\pi}$.

If α is a smooth section of T^*M, then

$$\tilde{\pi}\alpha = i_\alpha \pi. \tag{1.5}$$

This allows us to define a Lie algebroid structure on T^*M [11, 14–18] as follows.

(i) The Bracket is defined by

$$[\alpha, \beta]_{T^*M} = L_{\tilde{\pi}\alpha}\beta - L_{\tilde{\pi}\beta}\alpha - d(\pi(\alpha, \beta)), \tag{1.6}$$

where L is the usual Lie derivative of a 1-form: if X is a vector field and α a k-form, then $L_X\alpha$ is given by the Cartan formula

$$L_X\alpha = i_X d\alpha + d i_X \alpha. \tag{1.7}$$

(ii) The anchor map ρ_{T^*M} is the map $\tilde{\pi}$.

Infinite dimensional symplectic structures and their related Poisson structure were introduced by Dito and Léandre [19], Léandre [7, 20–22], and Léandre and Obame [23] in the infinite dimensional analysis, motivated by the theory of deformation quantization in infinite dimension. We refer to the review of Léandre on that in [7].

The infinite dimensional Poisson structure is a tool in the theory of integrable system [24]. We refer to the review of Mokhov in [3] and Dubrovin and Novikov in [2] on that. In particular, some partial differential equations of the theory of integrable systems are described as Hamiltonian systems associated to some Poisson structure. For instance, the Gardner-Zakharov-Faddeev (ultralocal) bracket (see [2, pages 52-53])

$$\{x(s), x(t)\} = \delta_0'(s-t) \tag{1.8}$$

is used to describe the KdV equation as a Hamiltonian system.

Another simple Poisson structure of Dubrovin and Novikov [2] is given as follows. We consider the set of smooth paths $t \to (x_i(t))$ into $(\mathbb{R}^m)^*$, the dual of a Lie algebra structure on $\mathbb{R}^m$ with structural constants $c_{i,j}^k$. In such a case,

$$\left\{x_i(s), x_j(t)\right\} = \sum_k \delta_0(s-t) c_{i,j}^k x_k(t). \tag{1.9}$$

It is useful to avoid the presence of the Dirac mass, and Léandre [4] has given an appropriate definition of the Poisson structure of the previous formula in the framework of Malliavin Calculus.

The goal of infinite dimensional analysis is to give a rigorous meaning to some formal considerations of mathematical physics. The formal operations of mathematical physics are defined consistently on some functional spaces. It is very well known, for instance, that the vacuum expectation of some operator algebras [25] is given by formal path integrals on the fields. Infinite dimensional analysis deals in the simplest case where these objects are mathematically well established.

(i) The functional integral side is given by the Malliavin Calculus [26].

(ii) The operator algebra side is given by white noise analysis and quantum probability [27, 28].

Let us recall basically the objects of these Calculi.

(i) The main object of white noise analysis and quantum probability is given by the Bosonic Fock space Fock$(\mathbb{H}_2)$ associated to the Hilbert space $\mathbb{H}_2$ of L^2 maps from $[0,1]$ into $\mathbb{R}$. Fock$(\mathbb{H}_2)$ is constituted of series $\sigma = \sum h^n$ where h^n belongs to $\mathbb{H}_2^{\hat{\otimes} n}$, the symmetric n-tensor product of $\mathbb{H}_2$ such that

$$\|\sigma\|^2 = \sum n! \|h^n\|^2_{\mathbb{H}_2^{\hat{\otimes} n}} < \infty. \tag{1.10}$$

The operator algebra is the algebra of annihilation and creation operator on the Fock space submitted to the canonical commutation relations

$$\begin{gathered}
[a(s), a(t)] = 0, \\
[a^*(s), a^*(t)] = 0, \\
[a(s), a^*(t)] = \delta_0(s-t),
\end{gathered} \tag{1.11}$$

where $a(s)$ is an elementary annihilation operator. $a^*(t)$ is an elementary creation operator. The presence of a Dirac mass leads to the same difficulties as in (1.8) and (1.9) and leads white noise analysis to consider an improvement of the Fock space (called the Hida Fock space) such that these operators act continuously on it. $a(s) + a^*(s)$ is called the white noise and can be interpreted in the measure theory.

(ii) The main object of the Malliavin Calculus is the L^p space of the Wiener measure.

If we consider the Brownian motion $t \to B(t)$ on $\mathbb{R}$ (see part 2 of this work), then we can introduce the Brownian functional associated to σ in the Bosonic Fock space:

$$\psi(\sigma) = \sum \int_{[0,1]^n} h^n(s_1, \ldots, s_n)\delta B(s_1) \cdots \delta B(s_n). \tag{1.12}$$

An element h^n of $\mathbb{H}_2^{\hat{\otimes}n}$ can be realized as a symmetric map from $[0,1]^n$ into $\mathbb{R}$ and $\int_{[0,1]^n} h^n(s_1, \ldots, s_n)\delta B(s_1) \cdots \delta B(s_n)$ denotes a Wiener chaos [28]. This map ψ realizes an isometry between the Fock space and the L^2 of the Wiener measure. The main ingredient of the Malliavin Calculus is to take the derivative (*almost surely defined!*) in the direction of an element $t \to \int_0^t h(s)$; $h \in \mathbb{H}_2$. An element $t \to \int_0^t h(s)ds$ is called an element of the Cameron-Martin space $\mathbb{H}$. This operation can be interpreted as a "nonelementary" annihilation operator on the Fock space. Through this isomorphism, $a(t) + a^*(t)$ can be interpreted as $d/dtB(t)$, the white noise associated to the Brownian motion, which does not exist in the traditional sense because the Brownian motion is *only continuous!* Since there are integrations by parts associated to a derivativation along an element of the Cameron-Martin space of a cylindrical functional, this operation is closable. It is the generalization in infinite dimension of the traditional definition of Sobolev spaces on finite dimensional spaces. But in infinite dimension, we consider Gaussian measures and not Lebesgue measure, which does not exist as a measure in infinite dimension! But since there is no Sobolev imbedding in infinite dimension, functionals which belong to all the Sobolev spaces of the Malliavin Calculus (these functionals are said to be smooth in the Malliavin sense) are in general only *almost surely defined*!

The study of Poisson structures requests that the test functional space where this Poisson structure acts is an *algebra*.

(i) In the case of the Malliavin Calculus, there is a natural way to choose an algebra starting from the considerations of measure theory. With the intersection of all the L^p, $p < \infty$ is indeed an algebra through the Hoelder inequality. We consider the Wiener product on the Wiener space which is the classical product of functionals.

(ii) In white noise analysis, there is on the Fock space another product called the standard Wick product. The traditional product of a Wiener chaos of length n and of length m is not a chaos of length $n + m$ by the help of the Itô formula. It is an infinite dimensional generalization of the fact that the product of two Hermite polynomials in finite dimension is not a Hermite polynomial. The classical Wick product consists to keep in the product of these two chaoses the chaoses of length $n+m$. Reference [29] has defined another Wick product (called the normalized Wick product), which fits well with Stratonovitch chaos. We consider now

$$\psi^{st}(\sigma) = \sum \int_{[0,1]^n} h(s_1, \ldots, s_n)dB(s_1) \cdots dB(s_n). \tag{1.13}$$

We consider this time multiple Stratonovitch integrals. They are the limit when $k \to \infty$ of the classical random multiple integrals $\int_{[0,1]^n} h(s_1, \ldots, s_n) dB^k(s_1) \cdots dB^k(s_n)$ where $t \to B^k(t)$ is the polygonal approximation of the Brownian motion. The Itô-Stratonovitch integral is the classical one. The normalized Wick product of Léandre and Rogers [29] $: \sigma_1 \cdot \sigma_2 :$ of σ_1 and σ_2 belonging to the Hida Fock space is done in order

$$\psi^{st}(: \sigma_1 \cdot \sigma_2 :) = \psi^{st}(\sigma_1)\psi^{st}(\sigma_2), \tag{1.14}$$

using the Itô-Stratonovitch formula [23]. This reflects in an infinite dimensional sense the classical fact that the product of two monomials is still a monomial in a finite dimensional polynomial algebra.

Léandre [4] has given an appropriate definition to the Poisson structure of the previous formula on an algebra of functional on the Wiener space of the Malliavin type. The main difficulty to overcome is that the good understanding of this Poisson structure leads to the study of some anticipative Stratonovitch integrals. The traditional Malliavin Calculus [26] is not suitable to study some anticipative Stratonovitch integrals. This means that if we consider a random element $s \to h(s)$ of $\mathbb{H}_2$ which belongs to all the Sobolev spaces of the Malliavin Calculus, then the anticipative Stratonovitch integral

$$\int_0^1 h(s) dB(s) = \lim_{k \to \infty} \int_0^1 h(s) dB^k(s) \tag{1.15}$$

does not exist. This pathology is not true if we consider a refinement of the Itô integral called the Hitsuda-Skorokhod integral [30]. Let us recall that the first authors who have studied anticipative Stratonovitch integrals are Nualart and Pardoux [30]. Léandre has defined conveniently some Sobolev spaces on the Wiener space such that the map anticipative Stratonovitch integral acts continuously on them [31–37]. Thisat means that if $h \in \mathbb{H}_2$ belongs to all the Sobolev spaces in the Nualart-Pardoux sense, then the anticipative Stratonovitch integral (1.15) exists. The main difference between the classical definition of the Sobolev space in the Malliavin sense and the Sobolev spaces in the Nualart-Pardoux sense is that some regularity on the kernels on the derivatives on the considered functionals is requested. This is a generalization of the following fact: we can define a nonanticipative Itô integral $\int_0^1 h(s) \delta B(s)$ without assuming a lot of regularity on the nonanticipative element h of $\mathbb{H}_2$. But in the case of a nonanticipative Stratonovitch integral, $\int_0^1 h(s) dB(s)$, we have to assume that h is a semi-martingale; this means some regularity on h! The Sobolev spaces of Nualart-Pardoux type were introduced by Léandre in [31–36] in order to study some Sobolev cohomology theories of some loop space endowed with the Brownian bridge measure on a compact Riemannian manifold. So the Poisson structure (1.9) can be defined consistently on the Nualart-Pardoux test algebra [4].

Let us recall that in white noise analysis, the algebraic counterpart of the Malliavin Calculus, the main tool is the Fock space and the algebra of creation and annihilation operators on the Fock space. The Bosonic Fock space is transformed into the L^2 of an infinite dimensional Gaussian measure by the help of the map Wiener chaoses. The Poisson structure (1.9) was defined by Léandre in [5] on the Hida test algebra endowed with the normalized Wick product.

The goal of this paper is to define a Lie algebroid associated to the Poisson structure (1.9) on the Nualart-Pardoux test algebra. The main remark is that the map $\tilde{\pi}$ transforms a 1-form on the Wiener space smooth in the Nualart-Pardoux sense in a generalized vector field on the Wiener space, whose theory was done by Léandre in [32, 35], and not in an ordinary vector field on the Wiener space! Classical vector field on the Wiener space are random elements of the Cameron-Martin space which belongs to all the Sobolev spaces of the Malliavin Calculus. Generalized vectorfield is $H(t) = \int_0^t h(s)dB(s) + \int_0^t h^1(s)ds$ where $h(s)$ is chosen well such that we can define the anticipative Stratonovitch integral $\int_0^t h(s)dB(s)$. In general, we cannot define the derivative of a functional which belongs to all the Sobolev spaces of the Malliavin Calculus along a generalized vector field. But we can do that if the functional belongs to all the Sobolev spaces in the Nualart-Pardoux sense. In this paper, since we consider smooth 1-forms in the Nualart-Pardoux sense, we can define still their interior product by a generalized vector field through our theory of anticipative Stratonovitch integral. So the formulas (1.1) and (1.2) are still true, but almost surely!

2. The Linear Stochastic Poisson Structure

We consider the set of continuous paths $C([0,1];\mathbb{R}^m)$ from $[0,1]$ into $\mathbb{R}^m$ endowed with the uniform topology. A typical path is denoted by $t \to B(t) = (B_i(t))$, on which we consider the Brownian motion measure dP [38].

Let us recall how we construct dP. We consider the Cameron-Martin Hilbert space $\mathbb{H}$ [39] of maps from $[0,1]$ into $\mathbb{R}^m$ such that

$$\int_0^1 \left|h'_s\right|^2 ds = \|h\|^2 < \infty, \tag{2.1}$$

and dP is formally the Gaussian probability measure

$$\frac{1}{Z}\exp\left[-\frac{\|h\|^2}{2}\right]dD(h), \tag{2.2}$$

and dD is the formal Lebesgue measure on $\mathbb{H}$ which does not exist as a measure (we refer to the works of Léandre [40–42], Asada [43], and Pickrell [44] for various approaches to the Lebesgue measure in infinite dimension). Let $\mathbb{H}_1$ be a finite dimensional real Hilbert space ($h_1 \in \mathbb{H}_1$) with Hilbert norm $\|\cdot\|_1$. Let us consider the centered normalized Gaussian measure on $\mathbb{H}_1$. It is classically represented by $\sum e_i N_i$ where N_i are centered normalized independent one-dimensional Gaussian variables and the system of e_i constitutes an orthonormal basis of the real Hilbert space $\mathbb{H}_1$.

It should be tempting to represent dP by using the same procedure. We consider an orthonormal basis e_i of $\mathbb{H}$. The law of the Brownian motion is represented by the *series* $\sum e_i N_i$ where the N_i is a collection of independent centered one-dimensional Gaussian variables. This series *does not converge* in $\mathbb{H}$ but in $C([0,1];\mathbb{R}^m)$ [45]. We refer to the textbook of Kuo [46] for the theory of infinite dimensional Gaussian measures.

Let us consider a functional F on $C([0,1];\mathbb{R}^m)$. Its rth stochastic derivative $\nabla^r F$, according to the framework of the Malliavin Calculus [26, 47, 48], is defined if it exists by

$$\langle \nabla^r F, h_1, \ldots, h_r \rangle = \int_{[0,1]^r} \left\langle \nabla^r F(s_1, \ldots, s_r), h'_{1,s_1}, \ldots, h'_{r,s_r} \right\rangle ds_1 \cdots ds_r, \tag{2.3}$$

where h_i belongs to the Hilbert space of paths $\mathbb{H}$ from $[0,1]$ into $\mathbb{R}^m$ satisfying

$$\int_0^1 \left|h'_s\right|^2 ds = \|h\|^2 < \infty. \tag{2.4}$$

The Sobolev norms of the Malliavin Calculus are defined by the following formula. If F is a Brownian functional, then

$$\left\{ E\left[\left(\int_{[0,1]^r} |\nabla^r F(s_1, \ldots, s_r)|^2 ds_1 \cdots ds_r \right)^{p/2} \right] \right\}^{1/p} = \|F\|_{r,p}. \tag{2.5}$$

The Malliavin test algebra consists of Brownian functionals F all of whose Sobolev norms $\|F\|_{r,p}$ are finite $r, p < \infty$. Let us recall how we construct these Sobolev spaces. Let f be a smooth function from $(\mathbb{R}^m)^d$ into $\mathbb{R}$ with compact support and some times $0 < t_1 < \cdots < t_d \leq 1$. We introduce the cylindrical functional $F(B(\cdot)) = f(B(t_1), \ldots, B(t_d))$. We consider the Gateaux derivative of F along a deterministic direction h of $\mathbb{H}$:

$$\langle \nabla F, h \rangle = \left\langle \sum \frac{\partial}{\partial x_i} f(B(t_1), \ldots, B(t_d)), h_{t_i} \right\rangle. \tag{2.6}$$

There is absolutely no problem to define it. We use the integration by parts formula for a cylindrical functional

$$E[\langle \nabla F, h \rangle] = E\left[F \int_0^1 \langle h'_s, \delta B(s) \rangle \right], \tag{2.7}$$

where $\delta B(s)$ is the Itô differential. The Itô integral is the limit in all the $L^p(dP)$, $p < \infty$ of the sum $\sum \langle h'_{s_i}, B(s_{i+1}) - B(s_i) \rangle$ where $0 < s_1 < \cdots < s_i < s_{i+1} < \cdots < s_{2^n - 1} < 1 = s_{2^n}$ is a dyadic subdivision of $[0,1]$ of length 2^n. The convergence does not pose any problem because h is deterministic. Since we have the integration by parts formula (2.7), we can extend the operation of taking the stochastic derivative of a Brownian functional F consistently, as we establish classically the definition of Sobolev spaces in finite dimension. The main novelty of the Malliavin Calculus with respects of [49–52] motivated by mathematical physics is that the *algebra* of functionals which belong to all the Sobolev spaces of the Malliavin Calculus (these functionals are said to be smooth in the Malliavin sense) is constituted of functionals almost surely defined. The reader interested in the Malliavin Calculus can see the books of [26, 47].

If we consider the same dyadic subdivision as before, we can introduce the polygonal approximation $B^n(t)$ of $B(t)$. Let us consider a "nondeterministic!" map from $[0,1]$ into

$\mathbb{R}^m t \to \beta_t$ which belongs to $L^2([0,1];\mathbb{R}^m)$. We can consider the random ordinary integral $\int_0^1 \langle \beta_t, dB^n(t)\rangle$. It can also be easily defined. In general, the limit may not exist when n tends to infinity, because the Brownian motion is only continuous. If we can pass to the limit, we say that the limit $\int_0^1 \langle \beta_t, dB(t)\rangle$ is an anticipative Stratonovitch integral. Nualart and Pardoux [30] are the first authors who have defined some anticipative Stratonovitch integrals. An appropriate theory was established by Léandre [31–36] in order to understand some Sobolev cohomology theories on the loop space. Let us recall it quickly.

We consider another set of Sobolev norms [31]. We suppose that outside the diagonals of $[0,1]^r$

$$\left\{E\left[\left|\nabla^r F(s_1,\ldots,s_r) - \nabla^r F(s'_1,\ldots,s'_r)\right|^p\right]\right\}^{1/p} \le C_{r,p} \sum \left|s_i - s'_i\right|^{1/2}. \tag{2.8}$$

The smallest $C_{r,p}$ such that the previous inequality is satisfied is called the first Nualart-Pardoux Sobolev norm. The second Nualart-Pardoux Sobolev norm is the smallest $C^1_{r,p}$ such that, for all $(s_i) \in [0,1]^r$,

$$\left\{E\left[\left|\nabla^r F(s_1,\ldots,s_r)\right|^p\right]\right\}^{1/p} \le C^1_{r,p}. \tag{2.9}$$

Definition 2.1. The Nualart-Pardoux test algebra $N.P_{\infty-}$ consists of functionals F whose all Nualart-Pardoux Sobolev norms of first type and second type are finite. The elements of $N.P_{\infty-}$ are said to be smooth in the Nualart-Pardoux sense.

Let us recall that $N.P_{\infty-}$ is an algebra [31].

We can consider a random element of $L^2([0,1];\mathbb{R}^m)$, $t \to \beta_t$. We can consider its rth stochastic derivative

$$\left\langle \nabla^r \beta_t, h_1,\ldots,h_r\right\rangle = \int_{[0,1]^r} \left\langle \nabla^r \beta_t(s_1,\ldots,s_r), h'_{1,s_1},\ldots,h'_{r,s_r}\right\rangle ds_1\cdots ds_r. \tag{2.10}$$

Its first Nualart-Pardoux Sobolev norm $C_{r,p}$ is the smallest number such that outside the diagonals of $[0,1]\times[0,1]^r$

$$\left\{E\left[\left|\nabla^r \beta_t(s_1,\ldots,s_r) - \nabla^r \beta_{t'}(s'_1,\ldots,s'_r)\right|^p\right]\right\}^{1/p} \le C_{r,p}\left(\left|t-t'\right|^{1/2} + \sum \left|s_i - s'_i\right|^{1/2}\right). \tag{2.11}$$

The second type of Nualart-Pardoux Sobolev norm $C^1_{r,p}$ of $\beta_{(\cdot)}$ is the smallest number such that for all $(t,s_1,\ldots,s_r) \in [0,1]\times[0,1]^r$

$$\left\{E\left[\left|\nabla^r \beta_t(s_1,\ldots,s_r)\right|^p\right]\right\}^{1/p} \le C^1_{r,p}. \tag{2.12}$$

Let us recall the theorem of Léandre [31].

Theorem 2.2. *Let β be a random element of $L^2([0,1];\mathbb{R}^m)$ such that all its Nualart-Pardoux Sobolev norms are finite. Then the anticipative Stratonovitch integral*

$$\int_0^1 \langle \beta_t, dB(t)\rangle \tag{2.13}$$

is smooth in the Nualart-Pardoux sense and its Nualart-Pardoux Sobolev norms can be estimated in terms of the Nualart-Pardoux norms of β.

In such a case, $\int_0^1 \langle \beta_t, dB(t)\rangle$ is the limit in all the $L^p(dP)$, $p < \infty$ of $\int_0^1 \langle \beta_t, dB^n(t)\rangle$. Moreover,

$$\left\langle \nabla\left(\int_0^1 \langle \beta_t, dB(t)\rangle\right), \widetilde{h}\right\rangle = \int_0^1 \left\langle \left\langle \nabla\beta_t, \widetilde{h}\right\rangle, dB(t)\right\rangle + \int_0^1 \left\langle \beta_t, \frac{d}{dt}\widetilde{h}_t\right\rangle dt. \tag{2.14}$$

This means that the kernel of the stochastic derivative of $\int_0^1 \langle \beta_t, dB(t)\rangle$ is $\int_0^1 \langle \nabla\beta_t(s), dB(t)\rangle + \beta_s$. Let us explain this formula; in order to take the stochastic derivative of $\int_0^1 \langle \beta_t, dB(t)\rangle$, we do the same formal computations as if the anticipative Stratonovitch integral had been a classical integral; we take first of all derivatives of β_t which lead to the term $\langle \nabla\beta_t, \widetilde{h}\rangle$ and derivatives of dB_t which lead to $d/dt\widetilde{h}_t dt$.

Let us recall the notion of a Poisson bracket $\{\cdot,\cdot\}$. We consider a commutative Frechet unital real algebra endowed with a family of Banach norms $\|\cdot\|_p$. This means that for all p, there exists p^1 such that for all F^1, F^2 in A

$$\left\|F^1F^2\right\|_p \leq C_p\left\|F^1\right\|_{p^1}\left\|F^2\right\|_{p^1}. \tag{2.15}$$

A Poisson Bracket is a bilinear map from $A\times A$ into A, which is a derivation in each argument, vanishes on the unit. The derivation property means that for all F^1, F^2, F^3 in A

$$\left\{F^1F^2, F^3\right\} = F^1\left\{F^2, F^3\right\} + \left\{F^1, F^3\right\}F^2. \tag{2.16}$$

Moreover, it satisfies the following properties: if F^1, F^2, F^3 belong to A, then

$$\begin{gathered}\left\{F^1, F^2\right\} = -\left\{F^2, F^1\right\}, \\ \left\{\left\{F^1, F^2\right\}, F^3\right\} + \left\{\left\{F^2, F^3\right\}, F^1\right\} + \left\{\left\{F^3, F^1\right\}, F^2\right\} = 0.\end{gathered} \tag{2.17}$$

Moreover, for all p, there exist p' and C_p such that

$$\left\|\left\{F^1, F^2\right\}\right\|_p \leq C_p\left\|F^1\right\|_{p'}\left\|F^2\right\|_{p'}. \tag{2.18}$$

In the sequel, we will choose $A = N.P_{\infty-}$. We consider the structural constants $c_{i,j}^k$ of a Lie algebra structure on $(\mathbb{R}^m)^*$. The stochastic gradient ∇F of a functional F can be written $\nabla F = (\nabla F_i)$. Formula (1.9) reads in this framework as

$$\left\{F^1, F^2\right\} = \sum_{i,j,k}\int_0^1 \nabla F_i^1(s)\nabla F_j^2(s)c_{i,j}^k dB_k(s), \tag{2.19}$$

where we consider a Stratonovitch anticipative integral. This defines a Poisson structure on $N.P_{\infty-}$ in our framework [4, Theorem 1].

Remark 2.3. Let us motivate (2.19). Let us consider the Hilbert space $\mathbb{H}_2$ of L^2 maps from $[0,1]$ into $\mathbb{R}^m$. The $c_{i,j}^k$ define a structure of Lie algebra on $\mathbb{R}^{m*}$, and therefore, on $\mathbb{H}_2$. Let us consider two functionals F^1 and F^2 Frechet smooth on $\mathbb{H}_2^*$. Their derivatives are given by kernels

$$\left\langle \nabla F^i, h \right\rangle = \int_0^1 \left\langle \nabla F^i(s), h(s) \right\rangle ds. \tag{2.20}$$

The Lie bracket $[\nabla F^1, \nabla F^2]$ is given by $[\nabla F^1(s), \nabla F^2(s)]$ and the classical Lie-Poisson structure is given by

$$\left\{ F^1, F^2 \right\} = \int_0^1 \left\langle h(s), \left[\nabla F^1(s), \nabla F^2(s) \right] \right\rangle ds. \tag{2.21}$$

These considerations are heuristic because the product of two elements of L^2 is not an element of L^2. If we replace $dB(s)$ by $h(s)ds$ and if we consider the white-noise measure on $\mathbb{H}_2 1/Z \exp[-\|h\|_{\mathbb{H}_2}^2] dD(h)$ instead of the Brownian measure (2.2), then this heuristic formula gives the formula (2.19). This is relevant of the so-called Malliavin transfer principle: a formula becomes almost surely true through the theory of Stratonovitch integrals.

3. The Stochastic Lie Algebroid

Smooth vector fields in the Nualart-Pardoux sense on the Wiener space are functions β_t from $[0,1]$ into $\mathbb{R}^m$ such that

$$\left\{ E\left[\left| \nabla^r \beta_t(s_1, \ldots, s_r) - \nabla^r \beta_{t'}(s'_1, \ldots, s'_r) \right|^p \right] \right\}^{1/p} \leq C_{r,p} \left(\left| t - t' \right|^{1/2} + \sum \left| s_i - s'_i \right|^{1/2} \right) \tag{3.1}$$

on the connected complements of $[0,1] \times [0,1]^r$ where we have removed the diagonals and such that

$$\left\{ E\left[\left| \nabla^r \beta_t(s_1, \ldots, s_r) \right|^p \right] \right\}^{1/p} \leq C_{r,p}^1. \tag{3.2}$$

The infimums of $C_{r,p}$ and of $C_{r,p}^1$ in the previous formula are called the Nualart-Pardoux Sobolev norms of the vector field $\beta_{(\cdot)}$.

Smooth 1-forms in the Nualart-Pardoux sense on the Wiener space are functions $\alpha_{(\cdot)}$ from $[0,1]$ into $\mathbb{R}^{m*}$ such that

$$\left(E\left[\left| \nabla^r \alpha_t(s_1, \ldots, s_r) - \nabla^r \alpha_{t'}(s'_1, \ldots, s'_r) \right|^p \right] \right)^{1/p} \leq C_{r,p} \left(\left| t - t' \right|^{1/2} + \sum \left| s_i - s'_i \right|^{1/2} \right) \tag{3.3}$$

on the connected complements of $[0,1] \times [0,1]^r$ where we have removed the diagonals and such that

$$\left\{E\left[|\nabla^r \alpha_t(s_1, \ldots, s_r)|^p\right]\right\}^{1/p} \leq C^1_{r,p}. \tag{3.4}$$

The infimums of $C_{r,p}$ and of $C^1_{r,p}$ in the previous formula are called the Sobolev norms of the 1-form $\alpha_{(\cdot)}$. The pairing between a 1-form $\alpha_{(\cdot)}$ and a vector field $\beta_{(\cdot)}$ is realized via the formula

$$\langle \alpha_{(\cdot)}, \beta_{(\cdot)} \rangle = \int_0^1 \langle \alpha_t, \beta_t \rangle dt. \tag{3.5}$$

If α^1, α^2 are two smooth 1-forms in the Nualart-Pardoux sense on the Wiener space, then the bivector π associated to the stochastic Poisson structure is given by

$$\pi\left(\alpha^1, \alpha^2\right) = \sum_{i,j,k} \int_0^1 \alpha^1_{i,s} \alpha^2_{j,s} c^k_{i,j} dB_k(s). \tag{3.6}$$

The stochastic bivector π realizes a continuous bilinear map on the space of smooth 1-forms smooth into the space of smooth functionals.

A generalized vector field according to our theory [32, 35] is a a random application from $[0,1]$ into $\mathbb{R}^m$ $\beta^g_{(\cdot)}$ of the form

$$\beta^g_t = \sum_{i,j} \int_0^t \beta_{(i,j,s)} dB_j(s) e_i + \int_0^t \beta_s ds, \tag{3.7}$$

where $\beta_{(i,j,.)}$ and $\beta_{(\cdot)}$ are smooth in the Nualart-Pardoux sense. The Nualart-Pardoux Sobolev norms of a generalized vector field $\beta^g_{(\cdot)}$ are the collection of Nualart-Pardoux norms of $\beta_{(i,j,.)}$ and $\beta_{(\cdot)}$.

We can define a pairing between smooth 1-form and generalized vector fields by using the formula

$$\langle \alpha, \beta^g \rangle = \sum_{i,j} \int_0^1 \alpha_{(i,s)} \beta_{(i,j,s)} dB_j(s) + \int_0^1 \langle \beta_s, \alpha_s \rangle ds. \tag{3.8}$$

This allows us to define $\tilde{\pi}\alpha = i_\alpha \pi$ for a smooth 1-form in the Nualart-Pardoux sense as the generalized vector field

$$\tilde{\pi}\alpha_t = \sum_{i,j,k} \int_0^t \alpha_i(s) c^k_{i,j} dB_k(s) e_j. \tag{3.9}$$

This allows us to put the following definition.

Definition 3.1. If α and β are smooth 1-forms in the Nualart-Pardoux sense, then we define

$$[\alpha, \beta] = L_{\tilde{\pi}\alpha}\beta - L_{\tilde{\pi}\beta}\alpha - d\pi(\alpha, \beta), \tag{3.10}$$

where the Lie derivative is defined as usual by the formula

$$L_{\tilde{\alpha}}\beta = i_{\tilde{\pi}\alpha}d\beta + di_{\tilde{\pi}\alpha}\beta, \tag{3.11}$$

and d is the exterior derivative.

Since $\tilde{\pi}\alpha$ is a generalized vector field, the introduction of the Lie derivative leads to stochastic integral (we refer to [32, 33, 35] for similar constructions). Let us recall some results of [31, 32]. Let $\alpha_{(t_1,t_2)}$ be a map from $[0,1] \times [0,1]$ into $\mathbb{R}^{m*}$ or later into $(\mathbb{R}^{m*})^{\otimes 2}$ such that

$$\begin{aligned} &\left\{E\left[\left|\nabla^r \alpha_{(t_1,t_2)}(s_1,\dots,s_r) - \nabla^r \alpha_{(t'_1,t'_2)}(s'_1,\dots,s'_r)\right|^p\right]\right\}^{1/p} \\ &\quad \le C_{r,p}\left(\left|t_1 - t'_1\right|^{1/2} + \left|t_2 - t'_2\right| + \sum \left|s_i - s'_i\right|^{1/2}\right) \end{aligned} \tag{3.12}$$

on the connected complements of $[0,1] \times [0,1] \times [0,1]^r$ where we have removed the diagonals and such that

$$\left\{E\left[\left|\nabla^r \alpha_{t_1,t_2}(s_1,\dots,s_r)\right|^p\right]\right\}^{1/p} \le C^1_{r,p}. \tag{3.13}$$

The infimums of $C_{r,p}$ and of $C^1_{r,p}$ in the previous formula are called the Nualart-Pardoux Sobolev norms of the $\alpha_{(\cdot)}$. In such case $\int_0^1 \alpha_{(t_1,t_2)} dB(t_1)$ is still smooth in the Nualart-Pardoux sense, with t_1 being included as well as $\iint_0^1 \alpha_{t_1,t_2} dB(t_1) dB(t_2)$.

This allows us to show the following theorem.

Theorem 3.2. $[\cdot,\cdot]$ *is a continuous antisymmetric bilinear application acting on the space of smooth 1-forms in the Nualart-Pardoux sense with values in the set of smooth 1-forms in the Nualart-Pardoux sense.*

Proof of Theorem 3.2. We remark that

$$\pi\left(\alpha^1, \alpha^2\right) = \sum_{i,j,k} \int_0^1 \alpha^1_i(s)\alpha^2_j(s)c^k_{i,j} dB_k(s), \tag{3.14}$$

which is smooth and its Sobolev norms can be estimated in terms of the Sobolev norms of α^i by Theorem 2.2.

Moreover,

$$i_{\tilde{\pi}\alpha^1}\alpha^2 = \sum_{i,j,k} \int_0^1 \alpha^1_i(s)\alpha^2_j(s)c^k_{j,i} dB_k(s), \tag{3.15}$$

which is still smooth. Moreover,

$$d\alpha^2(s,t) = \sum_i \left(\nabla \left(\alpha_i^2 \right)_t (s) - \nabla \left(\alpha_i^2 \right)_s (t) \right) e_i. \tag{3.16}$$

Moreover,

$$i_{\tilde{\pi}\alpha^1} d\alpha^2(t) = \sum_{i,j,k} \left(\int_0^1 \nabla \left(\alpha_i^2 \right)_t (s) \alpha_j^1(s) c_{j,i}^k dB_k(s) - \int_0^1 \nabla \left(\alpha_i^2 \right)_s (t) \alpha_j^1(s) c_{j,i}^k dB_k(s) \right), \tag{3.17}$$

which is smooth by the remark preceding the theorem. □

Theorem 3.3. $[\cdot,\cdot]$ *defines a Lie bracket.*

Let $0 < t_1 < \cdots < t_r = 1$ be a dyadic subdivision of $[0,1]$. We deduce a partition of $[0,1]^r$ in cubes $I_{n,r}$ of volume V_n. If β is a function from $[0,1]^r$ into some linear space, then we put

$$\chi_n \beta(t_1,\ldots,t_r) = V_n^{-1} \sum 1_{I_{n,r}}(t_1,\ldots,t_r) \int_{I_{n,r}} \beta(s_1,\ldots,s_r) ds_1 \cdots ds_r. \tag{3.18}$$

If we consider the polygonal approximation B^n of B and F_n being the σ-algebra associated to B^n, We denote by Π_n the operation of taking the conditional expectation of a functional f by F_n. The results of Léandre [31, Appendix], allow to state the proposition.

Proposition 3.4. *Let one have*

$$\nabla^r \Pi_n F = \Pi_n \chi_n \nabla^r. \tag{3.19}$$

If $\alpha(t_1)$ with values in $\mathbb{R}^{m*}$ is smooth in the Nualart-Pardoux sense, t_1 being included, then the random ordinary integral $\int_0^1 \langle \chi_n \Pi_n \alpha(t), dB_t^n \rangle$ tends in all the Sobolev spaces of the Malliavin Calculus to the anticipative Stratonovitch integral $\int_0^1 \langle \alpha(t), dB_t \rangle$. If $\alpha(t_1,t_2)$ which takes its values in $(\mathbb{R}^{m*})^{\otimes 2}$ is smooth in the Nualart-Pardoux sense, then the double random ordinary integral $\iint_0^1 \langle \chi_n \Pi_n \alpha(t_1,t_2), dB_{t_1}^n, dB_{t_2}^n \rangle$ tends in all the Sobolev spaces of the Malliavin Calculus to the double anticipative Stratonovitch integral $\iint_0^1 \langle \alpha(t_1,t_2), dB_{t_1}, dB_{t_2} \rangle$.

Proof of Theorem 3.3. Let us consider the finite dimensional Gaussian space B_t^n. A 1-form α_t^n is piecewise constant as well as a vector field h_t. If F^n is a functional which depends on B^n only, then $\nabla^r F^n$ is constant on each $I_{n,r}$. We put

$$\left\{ F^{1,n}, F^{2,n} \right\}_n = \sum_{i,j,k} \int_0^1 \nabla_i F_i^{1,n}(s) \nabla_j F^{2,n}(s) c_{i,j}^k dB_k^n(s). \tag{3.20}$$

This defines a Poisson structure on the finite dimensional Gaussian space.

We can define π^n, $[\cdot,\cdot]_n$, and $\tilde{\pi}^n$ according to the line of the introduction. To a 1-form smooth in the Nualart-Pardoux sense α on the total Wiener space, we consider the 1-form $\Pi^n \chi_n \alpha = \alpha^n$ on the finite dimensional Gaussian space. We get

$$\left[\left[\alpha^{1,n},\alpha^{2,n}\right]_n,\alpha^{3,n}\right]_n + \left[\left[\alpha^{2,n},\alpha^{3,n}\right]_n,\alpha^{1,n}\right]_n + \left[\left[\alpha^{3,n},\alpha^{1,n}\right]_n,\alpha^{2,n}\right]_n = 0, \tag{3.21}$$

by doing as in finite dimension. By the previous proposition $[[\alpha^{1,n},\alpha^{2,n}]_n,\alpha^{3,n}]_n$ tends in all this attained L^p to $[[\alpha^1,\alpha^2],\alpha^3]$. Therefore the result.

Let us give the scheme of the proof of this last result. When we write $[[\alpha^{1,n},\alpha^{2,n}]_n,\alpha^{3,n}]_n$, there are a lot of terms which will appear. All these terms will tends separately to the corresponding term in $[[\alpha^1,\alpha^2],\alpha^3]$. Let us treat one of them, which will lead to double anticipative Stratonovitch integral. The other terms will be treated identically. For instance $i_{\tilde{\pi}[\alpha^{1,n},\alpha^{2,n}]_n} d\alpha^{3,n}$ will lead to double Stratonovitch integral which will tend in all the Sobolev spaces of the Malliavin Calculus to $i_{\tilde{\pi}[\alpha^1,\alpha^2]} d\alpha^3$. We can consider in these expressions the term $i_{\tilde{\pi}^n[i_{\tilde{\pi}^n(\alpha^{1,n})}d\alpha^{2,n}]} d\alpha^{3,n}$ which will lead to a double stochastic integral and which will tend to $i_{\tilde{\pi}[i_{\tilde{\pi}(\alpha^1)}d\alpha^2]} d\alpha^3$. But there are two parts in $i_{\tilde{\pi}(\alpha^1)} d\alpha^2$ and $i_{\tilde{\pi}^n(\alpha^{1,n})} d\alpha^{2,n}$. We will consider the parts $\langle \nabla\alpha^2, \tilde{\pi}(\alpha^1)\rangle$ and $\langle \nabla\alpha^{2,n}, \tilde{\pi}^n(\alpha^{1,n})\rangle$ where we take the covariant derivative of the 1-form α^2 in the direction of the generalized vector field $\tilde{\pi}(\alpha^1)$. We will show that $i_{\tilde{\pi}^n[\langle \nabla\alpha^{2,n},\tilde{\pi}^n(\alpha^{1,n})\rangle]} d\alpha^{3,n}$ tends to $i_{\tilde{\pi}[\langle \nabla\alpha^2,\tilde{\pi}(\alpha^1)\rangle]} d\alpha^3$. But in these expressions there are still two parts which can be treated similarly. We will show that the expression $\langle \nabla\alpha^{3,n}, \tilde{\pi}^n[\langle \nabla\alpha^{2,n}, \tilde{\pi}^n(\alpha^{1,n})\rangle]\rangle$ tends to $\langle \nabla\alpha^3, \tilde{\pi}[\langle \nabla\alpha^2, \tilde{\pi}(\alpha^1)\rangle]\rangle$ (we consider the covariant derivative of $\alpha^{3,n}$ in the direction of $\tilde{\pi}^n[\langle \nabla\alpha^{2,n}, \tilde{\pi}^n(\alpha^{1,n})\rangle]$).

But

$$\tilde{\pi}^n\left(\alpha^{1,n}\right)_t = \sum_{i,j,k}\int_0^t \alpha^{1,n}_{s,i} c^k_{i,j} dB^n_k(s) e_j. \tag{3.22}$$

Therefore

$$\left\langle \nabla\alpha^{2,n}_t, \tilde{\pi}^n\left(\alpha^{1,n}\right)\right\rangle = \sum_{i,j,k}\int_0^1 \left\langle \nabla\alpha^{2,n}_t(s), \alpha^{1,n}_{s,i} c^k_{i,j} dB^n_k(s) e_j\right\rangle. \tag{3.23}$$

This implies that

$$\begin{aligned} &\tilde{\pi}^n\left[\left\langle \nabla\alpha^{2,n}(\cdot), \tilde{\pi}^n(\alpha^{1,n})\right\rangle\right]_t \\ &\quad = \sum_{i,j,k,i',j',k'}\int_0^t \left\langle \int_0^1 \left\langle \nabla\alpha^{2,n}_{u,i'}(s), \alpha^{1,n}_{s,i} c^k_{i,j} dB^n_k(s) e_j\right\rangle, c^{k'}_{i',j'} dB^n_{k'}(u) e_{j'}\right\rangle, \end{aligned} \tag{3.24}$$

where $\alpha^{2,n}_t(s) = \sum_{i'} \alpha^{2,n}_{t,i'} e^*_{i'}$ (we consider a 1-form in the t variable).

Therefore

$$\left\langle \nabla\alpha^{3,n}, \tilde{\pi}^n\left[\left\langle \nabla\alpha^{2,n}, \tilde{\pi}^n\left(\alpha^{1,n}\right)\right\rangle\right]\right\rangle$$
$$= \sum_{i,j,k,i',j',k'} \int_0^1 \left\langle \nabla\alpha^{3,n}_{(\cdot),j'}(u), \left\langle \int_0^1 \left\langle \nabla\alpha^{2,n}_{u,i'}(s), \alpha^{1,n}_{s,i} c^k_{i,j} dB^n_k(s) e_j \right\rangle, c^{k'}_{i',j'} dB^n_{k'}(u) e_{j'} \right\rangle \right\rangle. \tag{3.25}$$

By Proposition 3.4, this tends in all the Sobolev spaces of the Malliavin Calculus to

$$\sum_{i,j,k,i',j',k'} \int_0^1 \left\langle \nabla\alpha^{3}_{(\cdot),j'}(u), \left\langle \int_0^1 \left\langle \nabla\alpha^{2}_{u,i'}(s), \alpha^{1}_{s,i} c^k_{i,j} dB_k(s) e_j \right\rangle, c^{k'}_{i',j'} dB_{k'}(u) e_{j'} \right\rangle \right\rangle. \tag{3.26}$$

We recognize in this quantity $\langle \nabla\alpha^3, \tilde{\pi}[\langle \nabla\alpha^2, \tilde{\pi}(\alpha^1)\rangle]\rangle$ where we take the covariant derivative of α^3 in the direction of the generalized vector field $\tilde{\pi}[\langle \nabla\alpha^2, \tilde{\pi}(\alpha^1)\rangle]$ (we had taken the covariant derivative of α^2 in the direction of the generalized vector field $\tilde{\pi}(\alpha^2)$). □

Theorem 3.5. *$\tilde{\pi}$ is an anchor map. This means that for all 1-form α, β which are smooth in the Nualart-Pardoux sense all functional F are smooth in the Nualart-Pardoux sense; one has the relation:*

$$[\alpha, F\beta] = F[\alpha, \beta] + \langle \nabla F, \tilde{\pi}(\alpha)\rangle\beta. \tag{3.27}$$

Proof of Theorem 3.5. We get by classical results in finite dimension

$$[\alpha^n, F^n\beta^n]_n = F^n[\alpha_n, \beta_n]_n + \langle \nabla F^n, \tilde{\pi}^n(\alpha^n)\rangle\beta^n. \tag{3.28}$$

By the results of Proposition 3.4, this tends when $n \to \infty$ to the formula

$$[\alpha, F\beta] = F[\alpha, \beta] + \langle \nabla F, \tilde{\pi}(\alpha)\rangle\beta. \tag{3.29}$$

Therefore the result is attained. □

4. Conclusion

We can summarize that $([\cdot,\cdot], \tilde{\pi})$ realizes a stochastic Lie algebroid acting on the space of smooth 1-forms in the Nualart-Pardoux sense on the Wiener space and functional smooth in the Nualart-Pardoux sense on the Wiener space. $\tilde{\pi}$ takes its values in the space of generalized vector fields.

References

[1] A. A. Kirillov, "Local Lie algebras," *Russian Mathematical Surveys*, vol. 31, no. 4, pp. 55–75, 1976.

[2] B. A. Dubrovin and S. P. Novikov, "Hydrodynamics of weakly deformed soliton lattices. Differential geometry and Hamiltonian theory," *Russian Mathematical Surveys*, vol. 44, no. 6, pp. 35–124, 1989.

[3] O. I. Mokhov, "Differential geometry of symplectic and Poisson structures on loop spaces of smooth manifolds, and integrable systems," in *Loop Spaces and Groups of Diffeomorphisms*, vol. 217 of *Proceedings of the Steklov Institute of Mathematics*, pp. 91–125, Nauka, Moscow, Russia, 1997.

[4] R. Léandre, "A stochastic Poisson structure," in *Symmetry*, vol. 1, pp. 55–63, 2009.

[5] R. Léandre, "A Poisson structure in white-noise analysis," in *XXVIII Workshop of Geometrical Method in Physics*, M. Schlichenmaier and Th. Voronov, Eds., Bialowiesza, Poland, 2009.

[6] A. Weinstein, "Poisson geometry," *Differential Geometry and Its Applications*, vol. 9, no. 1-2, pp. 213–238, 1998.

[7] R. Léandre, "Deformation quantization in infinite dimensional analysis," in *Trends in Stochastic Analysis (Festchrift in honour of H. v. Weizsaecker)*, M. Scheutzow, P. Morters, and J. Blath, Eds., vol. 353 of *London Mathematical Society Lecture Note Series*, pp. 303–325, Cambridge University Press, Cambridge, UK, 2009.

[8] A. S. Cattaneo, B. Dherin, and G. Felder, "Formal symplectic groupoid," *Communications in Mathematical Physics*, vol. 253, no. 3, pp. 645–674, 2005.

[9] M. Kontsevich, "Deformation quantization of Poisson manifolds," *Letters in Mathematical Physics*, vol. 66, no. 3, pp. 157–216, 2003.

[10] A. Cannas da Silva and A. Weinstein, *Geometric Models for Noncommutative Algebras*, vol. 10 of *Berkeley Mathematics Lecture Notes*, American Mathematical Society, Providence, RI, USA, 1999.

[11] Y. Kosmann-Schwarzbach, "Poisson manifolds, Lie algebroids, modular classes: a survey," *SIGMA. Symmetry, Integrability and Geometry. Methods and Applications*, vol. 4, paper 005, 30 pages, 2008.

[12] K. Mackenzie, *Lie Groupoids and Lie Algebroids in Differential Geometry*, vol. 124 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge, UK, 1987.

[13] K. C. H. Mackenzie, *General Theory of Lie Groupoids and Lie Algebroids*, vol. 213 of *London Mathematical Society Lecture Note Series*, Cambridge University Press, Cambridge, UK, 2005.

[14] A. Coste, P. Dazord, and A. Weinstein, "Groupoïdes symplectiques," in *Publications du Département de Mathématiques*, vol. 87, pp. 1–62, Université Claude Bernard, Lyon, France, 1987.

[15] B. Fuchssteiner, "The Lie algebra structure of degenerate Hamiltonian and bi-Hamiltonian systems," *Progress of Theoretical Physics*, vol. 68, no. 4, pp. 1082–1104, 1982.

[16] I. M. Gelfand and I. Ya. Dorfman, "Schouten bracket and Hamiltonian operators," *Functional Analysis and Its Applications*, vol. 14, no. 3, pp. 71–74, 1980 (Russian).

[17] M. V. Karasev, "Analogues of objects of Lie groups for nonlinear Poisson brackets," *Mathematics of the USSR-Izvestiya*, vol. 28, pp. 497–527, 1987 (Russian).

[18] F. Magri and C. Morosi, "A geometrical characterization of integrable Hamiltonian systems through the theory of Poisson-Nijenhuis manifolds," *Quaderno*, vol. 19, 1984.

[19] G. Dito and R. Léandre, "Stochastic Moyal product on the Wiener space," *Journal of Mathematical Physics*, vol. 48, no. 2, Article ID 023509, 8 pages, 2007.

[20] R. Léandre, "Deformation quantization in white noise analysis," *SIGMA. Symmetry, Integrability and Geometry. Methods and Applications*, vol. 3, paper 027, 8 pages, 2007.

[21] R. Léandre, "Fedosov quantization in white noise analysis," *Journal of Nonlinear Mathematical Physics*, vol. 15, supplement 3, pp. 251–263, 2008.

[22] R. Léandre, "Hochschild cohomology theories in white noise analysis," *SIGMA. Symmetry, Integrability and Geometry. Methods and Applications*, vol. 4, paper 066, 13 pages, 2008.

[23] R. Léandre and B. Obame, "Quantization of the Connes algebra," in preparation.

[24] A. Odzijewicz and T. S. Ratiu, "Banach Lie-Poisson spaces and reduction," *Communications in Mathematical Physics*, vol. 243, no. 1, pp. 1–54, 2003.

[25] P. Di Francesco, P. Mathieu, and D. Sénéchal, *Conformal Field Theory*, Springer, Heidelberg, Germany, 1996.

[26] D. Nualart, *The Malliavin Calculus and Related Topics*, Probability and Its Applications, Springer, New York, NY, USA, 1995.

[27] T. Hida, H.-H. Kuo, J. Potthoff, and L. Streit, *White Noise: An Infinite-Dimensional Calculu*, vol. 253 of *Mathematics and Its Applications*, Kluwer Academic Publishers, Dordrecht, The Netherlands, 1993.

[28] P.-A. Meyer, *Quantum Probability for Probabilists*, vol. 1610 of *Lecture Notes in Mathematics*, Springer, Berlin, Germany, 1995.

[29] R. Léandre and A. Rogers, "Equivariant cohomology, Fock space and loop groups," *Journal of Physics A*, vol. 39, no. 38, pp. 11929–11946, 2006.

[30] D. Nualart and É. Pardoux, "Stochastic calculus with anticipating integrands," *Probability Theory and Related Fields*, vol. 78, no. 4, pp. 535–581, 1988.

[31] R. Léandre, "Cohomologie de Bismut-Nualart-Pardoux et cohomologie de Hochschild entière," in *Séminaire de Probabilités, XXX*, J. Azéma, M. Emery, and M. Yor, Eds., vol. 1626 of *Lecture Notes in Mathematics*, pp. 68–99, Springer, Berlin, Germany, 1996.

[32] R. Léandre, "Brownian cohomology of an homogeneous manifold," in *New Trends in Stochastic Analysis (Charingworth, 1994)*, K. D. Elworthy, S. Kusuoka, and I. Shigekawa, Eds., pp. 305–347, World Scientific, River Edge, NJ, USA, 1997.

[33] R. Léandre, "Stochastic cohomology and Hochschild cohomology," in *Development of Infinite-Dimensional Noncommutative Analysis, Kyoto, Japan 1998*, A. Hora, Ed., vol. 1099 of *RIMS Kokyuroku*, pp. 17–26, University Press, Kyoto, Kyoto, Japan, 1999.

[34] R. Léandre, "Cover of the Brownian bridge and stochastic symplectic action," *Reviews in Mathematical Physics*, vol. 12, no. 1, pp. 91–137, 2000.

[35] R. Léandre, "Stochastic Adams theorem for a general compact manifold," *Reviews in Mathematical Physics*, vol. 13, no. 9, pp. 1095–1133, 2001.

[36] R. Léandre, "Stochastic algebraic de Rham complexes," *Acta Applicandae Mathematicae*, vol. 79, no. 3, pp. 217–247, 2003.

[37] R. Léandre and I. A. Volovich, "The stochastic Lévy Laplacian and Yang-Mills equation on manifolds," *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, vol. 4, no. 2, pp. 161–172, 2001.

[38] P. Lévy, *Processus Stochastiques et Mouvement Brownien*, Gauthier-Villars, Paris, France, 3rd edition, 1992.

[39] R. H. Cameron and W. T. Martin, "Transformations of Wiener integrals under translations," *Annals of Mathematics*, vol. 45, pp. 386–396, 1944.

[40] R. Léandre, "Paths integrals in noncommutative geometry," in *Encyclopedia of Mathematical Physics*, J. P. Françoise, G. L. Naber, and T. S. Tsun, Eds., pp. 8–12, Elsevier, Oxford, UK, 2006.

[41] R. Léandre, "Infinite Lebesgue distribution on a current group as an invariant distribution," in *Foundations of Probability and Physics IV, Vaxjoe, Sweden, 2006*, G. Adenier, C. Fuchs, and A. Khrennikov, Eds., vol. 889 of *AIP Conference Proceedings*, pp. 332–335, American Institute of Physics, Melville, NY, USA, 2007.

[42] R. Léandre, "The Lebesgue measure in infinite-dimensional spaces as an infinite-dimensional distribution," *Fundamental'naya i Prikladnaya Matematika*, vol. 13, no. 8, pp. 127–132, 2007 (Russian), translation in *Journal of Mathematical Sciences*, vol. 159, no. 6, pp. 833–836, 2009.

[43] A. Asada, "Regularized calculus: an application of zeta regularization to infinite dimensional geometry and analysis," *International Journal of Geometric Methods in Modern Physics*, vol. 1, no. 1-2, pp. 107–157, 2004.

[44] D. Pickrell, "Invariant measures for unitary groups associated to Kac-Moody Lie algebras," *Memoirs of the American Mathematical Society*, vol. 146, no. 693, p. 125, 2000.

[45] J. Neveu, *Processus Aléatoires Gaussiens*, Séminaire de Mathématiques Supérieures, no. 34, Les Presses de l'Université de Montréal, Montréal, Canada, 1968.

[46] H. H. Kuo, *Gaussian Measures in Banach Spaces*, Lecture Notes in Mathematics, vol. 463, Springer, Berlin, Germany, 1975.

[47] N. Ikeda and S. Watanabe, *Stochastic Differential Equations and Diffusion Processes*, vol. 24 of *North-Holland Mathematical Library*, North-Holland, Amsterdam, The Netherlands, 2nd edition, 1989.

[48] P. Malliavin, "Stochastic calculus of variation and hypoelliptic operators," in *Proceedings of the International Symposium on Stochastic Differential Equations (Research Institute for Mathematical Sciences, Kyoto University, Kyoto, 1976)*, K. Itô, Ed., pp. 195–263, John Wiley & Sons, New York, NY, USA, 1978.

[49] S. Albeverio and R. Høegh-Krohn, "Dirichlet forms and diffusion processes on rigged Hilbert spaces," *Zeitschrift für Wahrscheinlichkeitstheorie und Verwandte*, vol. 40, no. 1, pp. 1–57, 1977.

[50] Y. Berezanskii, "The self adjointness of elliptic operators with an infinite number of variables," *Ukrainian Mathematical Journal*, vol. 27, pp. 729–742, 1975.

[51] L. Gross, "Potential theory on Hilbert space," *Journal of Functional Analysis*, vol. 1, pp. 123–181, 1967.

[52] T. Hida, *Analysis of Brownian Functionals*, Carleton Mathematical Lecture Notes, no. 1, Carleton University, Ottawa, Canada, 1975.

[53] R. Léandre, "Long time behaviour of the Wiener process on a path group," to appear in *Group Theory: Classes, Representations and Connections, and Applications*.

Hindawi Publishing Corporation
Advances in Mathematical Physics
Volume 2010, Article ID 941560, 44 pages
doi:10.1155/2010/941560

Research Article

Renormalization, Isogenies, and Rational Symmetries of Differential Equations

A. Bostan,[1] S. Boukraa,[2] S. Hassani,[3] J.-M. Maillard,[4] J.-A. Weil,[5] N. Zenine,[3] and N. Abarenkova[6]

[1] *INRIA Paris-Rocquencourt, Domaine de Voluceau, B.P. 105, 78153 Le Chesnay Cedex, France*
[2] *LPTHIRM and Département d'Aéronautique, Université de Blida, 09470 Blida, Algeria*
[3] *Centre de Recherche Nucléaire d'Alger, 2 Boulevard. Frantz Fanon, BP 399, 16000 Alger, Algeria*
[4] *LPTMC, UMR 7600 CNRS, Université de Paris, Tour 24, 4ème étage, case 121, 4 Place Jussieu, 75252 Paris Cedex 05, France*
[5] *XLIM, Université de Limoges, 123 avenue Albert Thomas, 87060 Limoges Cedex, France*
[6] *St Petersburg Department of Steklov Institute of Mathematics, 27 Fontanka, 191023 St. Petersburg, Russia*

Correspondence should be addressed to J.-M. Maillard, maillard@lptl.jussieu.fr

Received 21 December 2009; Accepted 17 January 2010

Academic Editor: Richard Kerner

We give an example of infinite-order rational transformation that leaves a linear differential equation covariant. This example can be seen as a nontrivial but still simple illustration of an exact representation of the renormalization group.

1. Introduction

There is no need to underline the success of the renormalization group revisited by Wilson [1, 2] which is nowadays seen as a fundamental symmetry in lattice statistical mechanics or field theory. It contributed to promote 2d conformal field theories and/or scaling limits of second-order phase transition in lattice statistical mechanics.[1] If one does not take into account most of the subtleties of the renormalization group, the simplest sketch of the renormalization group corresponds to Migdal-Kadanoff decimation calculations, where the new coupling constants created at each step of the (real-space) decimation calculations are forced[2] to stay in some (slightly arbitrary) finite-dimensional parameter space. This drastic projection may be justified by the hope that the basin of attraction of the fixed points of the corresponding (renormalization) transformation in the parameter space is "large enough."

One heuristic example is always given because it is one of the very few examples of *exact* renormalization, the renormalization of the one-dimensional Ising model without a magnetic field. It is a straightforward undergraduate exercise to show that performing various decimations summing over every two, three, or N spins, one gets *exact generators of the renormalization group* reading $T_N : t \longrightarrow t^N$, where t is (with standard notations) the high temperature variable $t = \tanh(K)$. It is easy to see that these transformations T_N, depending on the integer N, commute together. Such an *exact symmetry* is associated with a covariance of the partition function per site $Z(t) = C(t) \cdot Z(t^2)$. In this particular case one recovers the (very simple) expression of the partition function per site, $2\cosh(K)$, as an infinite product of the action of (for instance) T_2 on the cofactor $C(t)$. In this very simple case, this corresponds to the using of the identity (valid for $|x| < 1$):

$$\prod_{n=0}^{\infty}\left(1 + x^{2^n}\right) = \frac{1}{1-x}. \tag{1.1}$$

For $T_3 : t \longrightarrow t^3$ one must use the identity

$$\prod_{n=0}^{\infty}\left(1 + x^{3^n} + x^{2\cdot 3^n}\right) = \prod_{n=0}^{\infty}\left(\frac{1 - x^{3^{n+1}}}{1 - x^{3^n}}\right) = \frac{1}{1-x}, \tag{1.2}$$

and for $T_N : t \longrightarrow t^N$ a similar identity where the 3 in the exponents is changed into N.

Another simple heuristic example is the one-dimensional Ising model *with a magnetic field*. Straightforward calculations enable to get an infinite number of exact generators of the corresponding renormalization group, represented as *rational* transformations[3]

$$T_N : (x, z) \longrightarrow T_N(x, z) = (x_N, z_N), \tag{1.3}$$

where the first two transformations T_2 and T_3 read in terms of the two (low-temperature well-suited and fugacity-like) variables $x = e^{4K}$ and $z = e^{2H}$:

$$\begin{aligned} x_2 &= \frac{(x+z)(1+xz)}{x\cdot(1+z)^2}, \qquad z_2 = z\cdot\frac{(1+xz)}{x+z}, \\ x_3 &= x\cdot\frac{\left(z^2x + 2z + 1\right)\left(z^2 + 2z + x\right)}{\left(z^2x + z + xz + x\right)^2}, \qquad z_3 = z\cdot\frac{z^2x + 2z + 1}{z^2 + 2z + x}. \end{aligned} \tag{1.4}$$

One simply verifies that these rational transformations of two (complex) variables commute. This can be checked by formal calculations for T_N and T_M for any N and M less than 30, and one can easily verify a fundamental property expected for renormalization group generators:

$$T_N \cdot T_M = T_M \cdot T_N = T_{NM}, \tag{1.5}$$

where the "dot" denotes the *composition* of two transformations. The infinite number of these rational transformations of two (complex) variables (1.3) are thus a *rational representation of the positive integers together with their product*. Such rational transformations can be studied "per se" as discrete dynamical systems, the iteration of any of these various exact generators corresponding to an orbit of the renormalization group.

Of course these two examples of exact representation of the renormalization group are extremely degenerate since they correspond to one-dimensional models.[4] Migdal-Kadanoff decimation will quite systematically yield *rational*[5] transformations similar to (1.3) in two, or more, variables.[6] Consequently, they are never (except "academical" self-similar models) exact representations of the renormalization group. The purpose of this paper is to provide simple (but *nontrivial*) examples of *exact* renormalization transformations that are not degenerate like the previous transformations on one-dimensional models.[7] In several papers [3, 4] for Yang-Baxter integrable models with a canonical genus one parametrization [5, 6] (elliptic functions of modulus k), we underlined that the *exact* generators of the renormalization group must necessarily identify with the various isogenies which amount to multiplying or dividing τ, the ratio of the two periods of the elliptic curves, by an integer. The simplest example is the *Landen transformation* [4] which corresponds to multiplying (*or dividing* because of the modular group symmetry $\tau \leftrightarrow 1/\tau$), the ratio of the two periods is

$$k \longrightarrow k_L = \frac{2\sqrt{k}}{1+k}, \quad \tau \longleftrightarrow 2\tau. \tag{1.6}$$

The other transformations[8] correspond to $\tau \leftrightarrow N \cdot \tau$, for various integers N. In the (transcendental) variable τ, it is clear that they satisfy relations like (1.5). However, in the natural variables of the model ($e^K, \tanh(K), k = \sinh(2K)$, not transcendental variables like τ), these transformations are *algebraic* transformations corresponding in fact to the *fundamental modular curves*. For instance, (1.6) corresponds to the *genus zero fundamental modular curve*

$$\begin{aligned} &j^2 \cdot j'^2 - (j+j') \cdot \left(j^2 + 1487 \cdot jj' + j'^2\right) + 3 \cdot 15^3 \cdot \left(16j^2 - 4027jj' + 16j'^2\right) \\ &\quad - 12 \cdot 30^6 \cdot (j+j') + 8 \cdot 30^9 = 0, \end{aligned} \tag{1.7}$$

or

$$\begin{aligned} &5^9 v^3 u^3 - 12 \cdot 5^6 u^2 v^2 \cdot (u+v) + 375uv \cdot \left(16u^2 + 16v^2 - 4027vu\right) \\ &\quad - 64(v+u) \cdot \left(v^2 + 1487vu + u^2\right) + 2^{12} \cdot 3^3 \cdot uv = 0, \end{aligned} \tag{1.8}$$

which relates the two Hauptmoduls $u = 12^3/j(k)$, $v = 12^3/j(k_L)$:

$$j(k) = 256 \cdot \frac{\left(1-k^2+k^4\right)^3}{k^4 \cdot \left(1-k^2\right)^2}, \qquad j(k_L) = 16 \cdot \frac{\left(1+14k^2+k^4\right)^3}{\left(1-k^2\right)^4 \cdot k^2}. \tag{1.9}$$

One verifies easily that (1.7) is verified with $j = j(k)$ and $j' = j(k_L)$.

The selected values of k, the modulus of elliptic functions, $k = 0, 1$, are actually *fixed points of the Landen transformations*. The Kramers-Wannier duality $k \leftrightarrow 1/k$ maps $k = 0$ onto $k = \infty$. For the Ising (resp. Baxter) model these selected values of k correspond to the three selected subcases of the model ($T = \infty$, $T = 0$, and the critical temperature $T = T_c$), for which the elliptic parametrization of the model degenerates into a rational parametrization [4]. We have the same property for all the other algebraic modular curves corresponding to $\tau \leftrightarrow N\cdot\tau$. This is certainly the main property most physicists expect for an exact representation of a *generator of the renormalization group*, namely, that it maps a generic point of the parameter space onto the critical manifold (fixed points). Modular transformations are, in fact, the only transformations to be compatible with all the other symmetries of the Ising (resp. Baxter) model like, for instance, the gauge transformations, some extended $sl(2) \times sl(2) \times sl(2) \times sl(2)$ symmetry [7], and so forth. It has also been underlined in [3, 4] that seeing (1.6) as a transformation on *complex variables* (instead of real variables) provides two other complex fixed points which actually correspond to *complex multiplication* for the elliptic curve, and are, actually, fundamental new singularities[9] discovered on the $\chi^{(3)}$ linear ODE [8–10]. In general, this underlines the deep relation between the renormalization group and the theory of elliptic curves in a deep sense, namely, *isogenies of elliptic curves, Hauptmoduls,*[10] *modular curves and modular forms*.

Note that an algebraic transformation like (1.6) or (1.8) cannot be obtained from any *local* Migdal-Kadanoff transformation which naturally yields *rational* transformations; an exact renormalization group transformation like (1.6) can only be deduced from *nonlocal* decimations. The emergence of modular transformations as representations of exact generators of the renormalization group explains, in a quite subtle way, the difficult problem of how renormalization group transformations can be compatible with *reversibility*[11] (iteration forward and backwards). An algebraic modular transformation (1.8) corresponds to $\tau \rightarrow 2\tau$ *and* $\tau \rightarrow \tau/2$ *in the same time, as a consequence of the modular group symmetry* $\tau \leftrightarrow 1/\tau$.

A simple rational parametrization[12] of the genus zero modular curve (1.8) reads:

$$u = 1728\frac{z}{(z+16)^3}, \qquad v = 1728\frac{z^2}{(z+256)^3} = u\left(\frac{2^{12}}{z}\right). \tag{1.10}$$

Note that the previously mentioned reversibility is also associated with the fact that the modular curve (1.8) is invariant by $u \leftrightarrow v$, and, within the previous rational parametrization (1.10), with the fact that permuting u and v corresponds[13] to the Atkin-Lehner involution $z \leftrightarrow 2^{12}/z$.

For many Yang-Baxter integrable models of lattice statistical mechanics the physical quantities (partition function per site, correlation functions, etc.) are solutions of selected[14] linear differential equations. For instance, the partition function per site of the square (resp. triangular, etc.) Ising model is an integral of an elliptic integral of the third kind. It would be too complicated to show the precise covariance of these physical quantities with respect to (algebraic) modular transformations like (1.8). Instead, let us give, here, an illustration of the nontrivial action of the renormalization group on some elliptic function that actually occurs in the 2D Ising model: a weight-one modular form. This modular form actually, and remarkably, emerged [11] in a second-order linear differential operator factor denoted Z_2 occurring [8] for $\chi^{(3)}$, and that the reader can think as a physical quantity solution of a particular linear ODE replacing the too complicated integral of an elliptic integral of the third kind. Let us consider the second-order linear differential operator (D_z denotes d/dz):

$$\alpha = D_z^2 + \frac{(z^2+56z+1024)}{z\cdot(z+16)(z+64)}\cdot D_z - \frac{240}{z\cdot(z+16)^2(z+64)}, \tag{1.11}$$

which has the (modular form) solution

$$
\begin{aligned}
& {}_2F_1\left(\left[\frac{1}{12}, \frac{5}{12}\right], [1]; 1728\frac{z}{(z+16)^3}\right) \\
& \quad = 2\cdot\left(\frac{z+256}{z+16}\right)^{-1/4} \cdot {}_2F_1\left(\left[\frac{1}{12}, \frac{5}{12}\right], [1]; 1728\frac{z^2}{(z+256)^3}\right).
\end{aligned}
\tag{1.12}
$$

Do note that the two pull-backs in the arguments of the *same* hypergeometric function are *actually related by the modular curve relation* (1.8) (see (1.10)). The covariance (1.12) is thus the very expression of a modular form property with respect to a modular transformation ($\tau \leftrightarrow 2\tau$) corresponding to the modular transformation (1.8).

The hypergeometric function at the rhs of (1.12) is solution of the second-order linear differential operator

$$
\beta = D_z^2 + \frac{z^2+416z+16384}{(z+256)(z+64)z}\cdot D_z - \frac{60}{(z+64)(z+256)^2}, \tag{1.13}
$$

which is the transformed of operator α by the Atkin-Lehner duality $z \leftrightarrow 2^{12}/z$, and, also, a conjugation of α:

$$
\beta = \left(\frac{z+16}{z+256}\right)^{-1/4}\cdot\alpha\cdot\left(\frac{z+16}{z+256}\right)^{1/4}. \tag{1.14}
$$

Along this line we can also recall that the (modular form) function[15]

$$
F(j) = j^{-1/12}\cdot {}_2F_1\left(\left[\frac{1}{12}, \frac{5}{12}\right], [1]; \frac{12^3}{j}\right), \tag{1.15}
$$

verifies:

$$
F\left(\frac{(z+16)^3}{z}\right) = 2\cdot z^{-1/12}\cdot F\left(\frac{(z+256)^3}{z^2}\right). \tag{1.16}
$$

A relation like (1.12) is a straight generalization of the covariance we had in the one-dimensional model $Z(t) = C(t)\cdot Z(t^2)$, which basically amounts to seeing the partition function per site as some "automorphic function" with respect to the renormalization group, with the simple renormalization group transformation $t \longrightarrow t^2$ being replaced by the algebraic modular transformation (1.8) corresponding to $\tau \leftrightarrow 2\tau$ (i.e., the Landen transformation (1.6)).

We have here all the ingredients for seeing the identification of exact algebraic representations of the renormalization group with the modular curves structures we tried so many times to promote (preaching in the desert) in various papers [3, 4]. However, even if there are no difficulties, just subtleties, these Ising-Baxter examples of exact algebraic representations of the renormalization group already require some serious knowledge of the modular curves, modular forms, and Hauptmoduls in the theory of elliptic curves, mixed with the subtleties naturally associated with the various branches of such algebraic (multivalued) transformations.

The purpose of this paper is to present another elliptic hypergeometric function and other much simpler (Gauss hypergeometric) second-order linear differential operators covariant by infinite-order rational transformations.

The replacement of *algebraic (modular) transformations* by simple *rational* transformations will enable us to display a complete *explicit description of an exact representation of the renormalization group* that any graduate student can completely dominate.

2. Infinite Number of Rational Symmetries on a Gauss Hypergeometric ODE

Keeping in mind modular form expressions like (1.12), let us recall a particular Gauss hypergeometric function introduced by Vidunas in [12]

$$\begin{aligned} {}_2F_1\left(\left[\frac{1}{2},\frac{1}{4}\right],\left[\frac{5}{4}\right];z\right) &= \frac{1}{4}\cdot z^{-1/4}\cdot\int_0^z t^{-3/4}(1-t)^{-1/2}dt \\ &= (1-z)^{-1/2}\cdot {}_2F_1\left(\left[\frac{1}{2},\frac{1}{4}\right],\left[\frac{5}{4}\right];\frac{-4z}{(1-z)^2}\right). \end{aligned} \tag{2.1}$$

This hypergeometric function corresponds to the integral of a holomorphic form on a *genus-one* curve $P(y,t)=0$:

$$\frac{dt}{y}, \quad \text{with:} \quad y^4 - t^3\cdot(1-t)^2 = 0. \tag{2.2}$$

Note that the function

$$\mathcal{F}(z) = z^{1/4}\cdot {}_2F_1\left(\left[\frac{1}{2},\frac{1}{4}\right],\left[\frac{5}{4}\right];z\right), \tag{2.3}$$

which is exactly an integral of an algebraic function, has an extremely simple covariance property with respect to the *infinite-order rational* transformation $z \to -4z/(1-z)^2$:

$$\mathcal{F}\left(\frac{-4z}{(1-z)^2}\right) = (-4)^{1/4}\cdot\mathcal{F}(z). \tag{2.4}$$

The occurrence of this specific infinite-order transformation is reminiscent of Kummer's quadratic relation

$${}_2F_1([a,b],[1+a-b];z) = (1-z)^{-a}\cdot {}_2F_1\left(\left[\frac{a}{2},\frac{1+a}{2}-b\right],[1+a-b];-\frac{4z}{(1-z)^2}\right), \tag{2.5}$$

but it is crucial to note that, relation (2.4) does not relate two different functions, but is an "automorphy" relation on the *same function*.

It is clear from the previous paragraph that we want to see such functions as "ideal" examples of physical functions covariant by an exact (here, rational) generator of the renormalization group. The function (2.3) is actually solution of the second-order linear differential operator:

$$\Omega = D_z^2 + \frac{1}{4}\frac{3-5z}{z\cdot(1-z)}\cdot D_z = \omega_1 \cdot D_z, \quad \text{with}$$
$$\omega_1 = D_z + \frac{1}{4}\frac{3-5z}{z\cdot(1-z)} = D_z + \frac{1}{4}\cdot\frac{d\ln\left(z^3(1-z)^2\right)}{dz}. \tag{2.6}$$

From the previous expression of ω_1 involving a log derivative of a rational function it is obvious that this second-order linear differential operator has two solutions, the constant function and an integral of an algebraic function. Since these two solutions behave very simply under the infinite-order rational transformation $z \to -4z/(1-z)^2$, it is totally and utterly natural to see how the linear differential operator Ω transforms under the rational change of variable $z \to R(z) = -4z/(1-z)^2$ (which amounts to seeing how the two-order-one operators ω_1 and D_z transform). It is a straightforward calculation to see that introducing the cofactor $C(z)$ which is the inverse of the derivative of $R(z)$

$$C(z) = -\frac{1}{4}\cdot\frac{(1-z)^3}{1+z}, \qquad \frac{1}{C(z)} = \frac{dR(z)}{dz}, \tag{2.7}$$

D_z and ω_1, respectively, transform under the rational change of variable $z \to R(z) = -4z/(1-z)^2$ as

$$D_z \longrightarrow C(z)\cdot D_z, \qquad \omega_1 \longrightarrow (\omega_1)^{(R)} = C(z)^2\cdot\omega_1\cdot\frac{1}{C(z)}, \quad \text{yielding:} \quad \Omega \longrightarrow C(z)^2\cdot\Omega. \tag{2.8}$$

Since $z \to -4z/(1-z)^2$ is of infinite-order, the second-order linear differential operator (2.6) has *an infinite number of rational symmetries* (isogenies):

$$z \longrightarrow \frac{-4z}{(1-z)^2} \longrightarrow 16\cdot\frac{(1-z)^2\cdot z}{(1+z)^4} \longrightarrow -64\cdot\frac{(1-z)^2(1+z)^4 z}{(1-6z+z^2)^4} \longrightarrow \cdots. \tag{2.9}$$

Once we have found a second-order linear differential operator (written in a unitary or monic form) Ω, covariant by the infinite-order rational transformation $z \to -4z/(1-z)^2$, it is natural to seek for higher-order linear differential operators also covariant by $z \to -4z/(1-z)^2$. One easily verifies that the successive symmetric powers of Ω are (of course) also covariant. The symmetric square of Ω,

$$D_z^3 + \frac{3}{4}\frac{3-5z}{(1-z)z}\cdot D_z^2 + \frac{3}{8}\frac{1-5z}{(1-z)z^2}\cdot D_z, \tag{2.10}$$

factorizes in simple order-one operators

$$\left(D_z + \frac{2}{4}\frac{3-5z}{(1-z)z}\right)\cdot\left(D_z + \frac{1}{4}\frac{3-5z}{(1-z)z}\right)\cdot D_z, \tag{2.11}$$

and, more generally, the symmetric Nth power[16] of Ω reads

$$\left(D_z + \frac{N}{4}\frac{3-5z}{z(1-z)}\right)\cdot\left(D_z + \frac{N-1}{4}\frac{3-5z}{z(1-z)}\right)\cdots\left(D_z + \frac{1}{4}\frac{3-5z}{z(1-z)}\right)\cdot D_z. \tag{2.12}$$

The covariance of such expressions is the straight consequence of the fact that the order-one factors

$$\omega_k = D_z + \frac{k}{4}\frac{3-5z}{z\cdot(1-z)}, \quad k = 0,1,\ldots,N, \tag{2.13}$$

transform very simply under $z \to -4z/(1-z)^2$:

$$\omega_k \longrightarrow (\omega_k)^{(R)} = (C(z))^{k+1}\cdot\omega_k\cdot(C(z))^{-k}. \tag{2.14}$$

More generally, let us consider a rational transformation $z \to R(z)$, the corresponding cofactor $C(z) = 1/R'(z)$, and the order-one operator $\omega_1 = D_z + A(z)$. We have the identity

$$C(z)\cdot D_z\cdot\left(\frac{1}{C(z)}\right) = D_z - \frac{d\ln(C(z))}{dz}. \tag{2.15}$$

The change of variable $z \to R(z)$ on ω_1 reads

$$D_z + A(z) \longrightarrow C(z)\cdot D_z + A(R(z)) = C(z)\cdot(D_z + B(z)). \tag{2.16}$$

We want to impose that this rhs expression can be written (see (2.8)) as

$$C(z)^2\cdot(D_z + A(z))\cdot\frac{1}{C(z)}, \tag{2.17}$$

which, because of (2.15), occurs if

$$B(z) = A(z) - \frac{d\ln(C(z))}{dz}, \tag{2.18}$$

yielding a "Rota-Baxter-like" [13, 14] functional equation on $A(z)$ and $R(z)$

$$\left(\frac{dR(z)}{dz}\right)^2\cdot A(R(z)) = \frac{dR(z)}{dz}\cdot A(z) + \frac{d^2R(z)}{dz^2}. \tag{2.19}$$

Remark 2.1. Coming back to the initial Gauss hypergeometric differential operator the covariance of Ω becomes a conjugation. Let us start with the Gauss hypergeometric differential operator for (2.1):

$$H_0 = 8z \cdot (1-z) \cdot D_z^2 + 2 \cdot (5-7z) \cdot D_z - 1. \tag{2.20}$$

It is transformed by $z \longrightarrow R(z) = -4z/(1-z)^2$ into

$$H_1 = 8z \cdot (1-z) \cdot D_z^2 - 2(3z-5) \cdot D_z + \frac{4}{1-z} = (1-z)^{1/2} \cdot H_0 \cdot (1-z)^{-1/2}, \tag{2.21}$$

then by $z \longrightarrow R(R(z)) = R_2(z) = 16z(1-z)^2/(1+z)^4$ into

$$\begin{aligned} H_2 &= 8z \cdot (1-z) \cdot D_z^2 - 2\frac{(3z-1)(z+5)}{z+1} \cdot D_z + 16\frac{z-1}{(z+1)^2} \\ &= \left(\frac{z+1}{\sqrt{z-1}}\right) \cdot H_0 \cdot \left(\frac{z+1}{\sqrt{z-1}}\right)^{-1}, \end{aligned} \tag{2.22}$$

and more generally for $z \longrightarrow R_N = R(R(R \cdots (R(z) \cdots)$

$$H_N = C_N \cdot H_0 \cdot C_N^{-1}, \quad \text{where: } C_N = z^{1/4} \cdot R_N^{-1/4}. \tag{2.23}$$

2.1. A Few Remarks on the "Rota-Baxter-Like" Functional Equation

The functional equation[17](2.19) is the (necessary and sufficient) condition for $\Omega = (D_z + A(z)) \cdot D_z$ to be covariant by $z \longrightarrow R(z)$.

Using the chain rule formula of derivatives of composed functions:

$$\begin{aligned} \frac{dR(R(z))}{dz} &= \frac{dR(z)}{dz} \cdot \left[\frac{dR(z)}{dz}(R(z))\right], \\ \frac{d^2R(R(z))}{dz^2} &= \frac{d^2R(z)}{dz^2} \cdot \left[\frac{dR(z)}{dz}(R(z))\right] + \left(\frac{dR(z)}{dz}\right)^2 \cdot \left[\frac{d^2R(z)}{dz^2}(R(z))\right], \end{aligned} \tag{2.24}$$

one can show that, for $A(z)$ fixed, the "Rota-Baxter-like" functional equation (2.19) is invariant by the composition of $R(z)$ by itself $R(z) \longrightarrow R(R(z)), R(R(R(z))), \ldots$. This result can be generalized to any composition of various $R(z)$'s satisfying (2.19). This is in agreement with the fact that (2.19) is the condition for $\Omega = (D_z + A(z)) \cdot D_z$ to be covariant by $z \longrightarrow R(z)$ it must be invariant by composition of $R(z)$'s (for $A(z)$ fixed).

Note that we have not used here the fact that for globally nilpotent [11] operators, $A(z)$ and $B(z)$ are necessarily log derivatives of Nth roots of rational functions.

For $R(z) = -4z/(1-z)^2$:

$$A(z) = \frac{1}{4} \cdot \frac{d \ln(a(z))}{dz}, \qquad B(z) = \frac{1}{4} \cdot \frac{d \ln(b(z))}{dz},$$
$$a(z) = (1-z)^2 \cdot z^3, \qquad b(z) = z^3 \cdot \frac{(1+z)^4}{(1-z)^{10}}. \tag{2.25}$$

The existence of the underlying $a(z)$ in (2.25) consequence of a global nilpotence of the order-one differential operator, can however be seen in the following remark on the zeros of the lhs and rhs terms in the functional equation (2.19). When $R(z)$ is a rational function (e.g., $-4z/(1-z)^2$ or any of its iterates $R^{(n)}(z)$), the lhs and rhs of (2.19) are rational expressions. The zeros are roots of the numerators of these rational expressions. Because of (2.25) the functional equation (2.19) can be rewritten (after dividing by $R'(z)$) as

$$\left(\frac{dR(z)}{dz}\right) \cdot A(R(z)) = A(z) + \frac{d}{dz}\left(\ln\left(\frac{dR(z)}{dz}\right)\right) = \frac{1}{4} \cdot \frac{d}{dz}\left(\ln\left(a(z) \cdot \left(\frac{dR(z)}{dz}\right)^4\right)\right). \tag{2.26}$$

One easily verifies, in our example, that the zeros of the rhs of (2.26) come from the zeros of $A(R(z))$ (and not from the zeros of $R'(z)$ in the lhs of (2.26)). The zeros of the log-derivative rhs of (2.26) correspond to $a(z) \cdot R'(z)^4 = \rho$, where ρ is a constant to be found. Let us consider for $R(z)$ the nth iterates of $-4z/(1-z)^2$ that we denote $R^{(n)}(z)$. A straightforward calculation shows that the zeros of $A(R^{(n)}(z))$ or $a'(R^{(n)}(z))$ (where $a'(z)$ denotes the derivative of $a(z)$ namely, $(z-1)(5z-3) \cdot z^2$) actually correspond to the general closed formula:

$$5^5 \cdot a(z) \cdot \left(\frac{dR^{(n)}(z)}{dz}\right)^4 - 4 \cdot 3^3 \cdot (-4)^n = 0. \tag{2.27}$$

More precisely the zeros of $5 \cdot R^{(n)}(z) - 3$ verify (2.27), or, in other words, the numerator of $5R^{(n)}(z) - 3$ divides the numerator of the lhs of (2.27).

In another case for $T(z)$ given by (2.45), which also verifies (2.19) (see below), the relation (2.27) is replaced by

$$5^5 \cdot a(z) \cdot \left(\frac{dT^{(n)}(z)}{dz}\right)^4 - 4 \cdot 3^3 \cdot (-7 - 24i)^n = 0. \tag{2.28}$$

More generally for a rational function $\rho(x)$, obtained by an arbitrary composition of $-4z/(1-z)^2$ and $T(z)$, we would have

$$5^5 \cdot a(z) \cdot \left(\frac{d\rho(z)}{dz}\right)^4 - 4 \cdot 3^3 \cdot \lambda^n = 0. \tag{2.29}$$

where λ corresponds to

$$\rho(x) = \lambda \cdot z + \cdots, \quad \lambda = \left[\frac{d\rho(z)}{dz}\right]_{z=0}. \tag{2.30}$$

2.2. Symmetries of Ω, Solutions to the "Rota-Baxter-Like" Functional Equation

Let us now analyse all the symmetries of the linear differential operator $\Omega = (D_z + A(z)) \cdot D_z$ by analyzing all the solutions of (2.19) for a given $A(z)$. For simplicity we will restrict to $A(z) = (3 - 5z)/z/(1 - z)/4$ which corresponds to $R(z) = -4z/(z - 1)^2$ and all its iterates (2.9). Let us first seek for other (more general) solutions that are *analytic at* $z = 0$:

$$R(z) = a_1 \cdot z + a_2 \cdot z^2 + a_3 \cdot z^3 + \cdots . \tag{2.31}$$

It is a straightforward calculation to get, order by order from (2.19), the successive coefficients a_n in (2.31) as polynomial expressions (with rational coefficients) of the first coefficient a_1 with

$$\begin{gathered} a_2 = -\frac{2}{5} \cdot a_1 \cdot (a_1 - 1), \qquad a_3 = \frac{1}{75} \cdot a_1 \cdot (a_1 - 1) \cdot (7a_1 - 17), \\ a_4 = -\frac{2}{4875} \cdot a_1 \cdot (a_1 - 1) \cdot \left(41a_1^2 - 232a_1 + 366\right), \ldots, \\ a_n = -\frac{n}{5} \cdot a_1 \cdot (a_1 - 1) \cdot \frac{P_n(a_1)}{P_n(-4)}, \end{gathered} \tag{2.32}$$

where $P_n(a_1)$ is a polynomial with integer coefficients of degree $n - 2$. Since we have here a series depending on one parameter a_1 we will denote it $R_{a_1}(z)$. This is a quite remarkable series depending on one parameter.[18] One can easily verify that this series actually reduces (as it should!) to the successive iterates (2.9) of $-4z/(1 - z)^2$ for $a_1 = (-4)^n$. In other words this one-parameter family of "functions" actually reduces to rational functions for an infinite number of integer values $a_1 = (-4)^n$.

Furthermore, one can also verify a quite essential property we expect for a representation of the renormalization group, namely, that two $R_{a_1}(z)$ for different values of a_1 commute, the result corresponding to the product of these two a_1:

$$R_{a_1}(R_{b_1}(z)) = R_{b_1}(R_{a_1}(z)) = R_{a_1 \cdot b_1}(z). \tag{2.33}$$

The neutral element must necessarily correspond to $a_1 = 1$ which is actually the identity transformation $R_1(z) = z$. We have an "absorbing" element corresponding to $a_1 = 0$, namely, $R_0(z) = 0$. Performing the inverse of $R_{a_1}(z)$ (with respect to the composition of functions) amounts to changing a_1 into its inverse $1/a_1$. Let us explore some "reversibility" property of our exact representation of a renormalization group with the inverse of the rational transformations (2.9). The inverse of $R_{-4}(z) = -4z/(1 - z)^2$ must correspond to $a_1 = -1/4$:

$$R_{-1/4}(z) = -\frac{1}{4} \cdot z - \frac{1}{8} z^2 - \frac{5}{64} z^3 - \frac{7}{128} z^4 - \frac{21}{512} z^5 + \cdots . \tag{2.34}$$

However, a straight calculation of the inverse of $R_{-4}(z) = -4z/(1 - z)^2$ gives a multivalued function, or if one prefers, two functions

$$\begin{aligned} S_{-1/4}^{(1)}(z) &= \frac{z - 2 + 2\sqrt{1 - z}}{z} = -\frac{1}{4} \cdot z - \frac{1}{8} z^2 + \cdots , \\ S_{-1/4}^{(2)}(z) &= \frac{z - 2 - 2\sqrt{1 - z}}{z} = -\frac{4}{z} + 2 + \frac{1}{4} z + \frac{1}{8} z^2 + \cdots , \end{aligned} \tag{2.35}$$

which are the two roots of the simple quadratic relation ($R_{-4}(z') = z$):

$$z'^2 - 2 \cdot \left(1 - \frac{2}{z}\right) \cdot z' + 1 = 0, \tag{2.36}$$

where it is clear that the product of these two functions is equal to +1. The radius of convergence of $S^{(1)}_{-1/4}(z)$ is 1.

Because of our choice to seek for functions analytical at $z = 0$ our renormalization group representation "chooses" the unique root that is analytical at $z = 0$, namely, $S^{(1)}_{-1/4}(z)$. For the next iterate of $R_{-4}(z) = -4z/(1-z)^2$ in (2.9) the inverse transformation corresponds to the roots of the polynomial equation of degree four ($R_{16}(z') = z$):

$$z'^4 + \left(4 - \frac{16}{z}\right) \cdot z'^3 + \left(6 + \frac{32}{z}\right) \cdot z'^2 + \left(4 - \frac{16}{z}\right) \cdot z' + 1 = 0, \tag{2.37}$$

which yields four roots, one of which is analytical at $z = 0$ and corresponds to $a_1 = 1/(-4)^2$ in our one-parameter family of (renormalization) transformations:

$$S^{(1)}_{1/16}(z) = \frac{1}{16}z + \frac{3}{128}z^2 + \frac{53}{4096}z^3 + \frac{277}{32768}z^4 + \frac{3181}{524288}z^5 + \cdots, \tag{2.38}$$

its (multiplicative) inverse $S^{(2)}_{1/16}(z) = 1/S^{(1)}_{1/16}(z)$:

$$S^{(2)}_{1/16}(z) = \frac{16}{z} - 6 - \frac{17}{16}z - \frac{67}{128}z^2 - \frac{1333}{4096}z^3 - \frac{7445}{32768}z^4 + \cdots, \tag{2.39}$$

and two (formal) Puiseux series ($u = \pm\sqrt{z}$):

$$S^{(3)}_{1/16}(z) = 1 + u + \frac{1}{2}u^2 + \frac{3}{8}u^3 + \frac{1}{4}u^4 + \frac{27}{128}u^5 + \frac{5}{32}u^6 + \cdots. \tag{2.40}$$

Many of these results are better understood when one keeps in mind that there is a special transformation $J : z \leftrightarrow 1/z$ which is *also a R-solution of* (2.19) and verifies many compatibility relations with these transformations (Id denotes the identity transformation $R_0(z)$):

$$\begin{gathered} R_{-4} \cdot J = R_{-4}, \qquad S^{(2)}_{-1/4} \cdot R_{-4} = J, \qquad R_{-4} \cdot S^{(1)}_{-1/4} = S^{(1)}_{-1/4} \cdot R_{-4} = Id, \\ S^{(1)}_{1/16}(z) = S^{(1)}_{-1/4} \cdot S^{(1)}_{-1/4}, \qquad S^{(2)}_{1/16}(z) = S^{(1)}_{-1/4} \cdot S^{(2)}_{-1/4}, \\ J \cdot S^{(1)}_{-1/4} = S^{(2)}_{-1/4}, \qquad J \cdot S^{(2)}_{-1/4} = S^{(1)}_{-1/4}, \ldots, \end{gathered} \tag{2.41}$$

where the dot corresponds, here, to the composition of functions. These symmetries of the linear differential operator Ω correspond to isogenies of the elliptic curve (2.2).

It is clear that we have another one-parameter family corresponding to $J \cdot R_{a_1}$ with an expansion of the form

$$
\begin{aligned}
J \cdot R_{a_1} = {} & \frac{b_1}{z} - \frac{2}{5} \cdot (b_1 - 1) - \frac{1}{15} \cdot \frac{b_1^2 - 1}{b_1} \cdot z - \frac{2}{975} \cdot \frac{(b_1 - 1)(4b_1 + 1)(4b_1 + 3)}{b_1^2} \cdot z^2 \\
& - \frac{1}{248625} \cdot \frac{(b_1 - 1)(4b_1 + 1)\left(1268b_1^2 + 951b_1 + 91\right)}{b_1^3} \cdot z^3 \\
& - \frac{2}{2071875} \cdot \frac{(b_1 - 1)(4b_1 + 1)\left(3688b_1^3 + 2766b_1^2 + 404b_1 + 17\right)}{b_1^4} \cdot z^4 + \cdots .
\end{aligned}
\tag{2.42}
$$

For $b_1 = -1/4$, $b_1 = (-1/4)^2$, $b_1 = (-1/4)^3$, this family reduces to the (multiplicative) inverse of the successive rational functions displayed in (2.9)

$$
-\frac{1}{4} \cdot \frac{(1-z)^2}{z} \longrightarrow \frac{1}{16} \cdot \frac{(1+z)^4}{(1-z)^2 \cdot z} \longrightarrow -\frac{1}{64} \cdot \frac{\left(1 - 6z + z^2\right)^4}{(1-z)^2(1+z)^4 \cdot z} \longrightarrow \cdots ,
\tag{2.43}
$$

which can also be written as:

$$
\begin{aligned}
& -\frac{1}{4} \cdot \left(z + \frac{1}{z}\right) + \frac{1}{2}, \qquad \frac{1}{16} \cdot \left(z + \frac{1}{z}\right) + \frac{3}{8} + \frac{z}{(1-z)^2}, \\
& -\frac{1}{64} \cdot \left(z + \frac{1}{z}\right) + \frac{13}{32} - \frac{z}{4} \cdot \frac{17 - 60z + 102z^2 - 60z^3 + 17z^4}{(1-z)^2(1+z)^4}, \\
& \frac{1}{256} \cdot \left(z + \frac{1}{z}\right) + \frac{51}{128} + \frac{z}{16} \cdot \frac{17 - 60z + 102z^2 - 60z^3 + 17z^4}{(1-z)^2(1+z)^4} + 16\frac{z \cdot (1-z)^2(1+z)^4}{\left(z^2 - 6z + 1\right)^4}, \\
& -\frac{1}{1024} \cdot \left(z + \frac{1}{z}\right) + \frac{205}{512} - \frac{z}{164} \cdot \frac{17 - 60z + 102z^2 - 60z^3 + 17z^4}{(1-z)^2(1+z)^4} \\
& \quad - 4\frac{z \cdot (1-z)^2(1+z)^4}{\left(z^2 - 6z + 1\right)^4} - 64\frac{z \cdot (1-z)^2(1+z)^4\left(z^2 - 6z + 1\right)^4}{\left(1 + 20z - 26z^2 + 20z^3 + z^4\right)^4}, \ldots, \\
& \frac{1}{(-4)^n} \cdot \left(z + \frac{1}{z}\right) + \frac{2}{54^n}\left(4^n - (-1)^n\right) \\
& \quad + \frac{z}{(-4)^{n-2}} \cdot \frac{17 - 60z + 102z^2 - 60z^3 + 17z^4}{(1-z)^2(1+z)^4} + \frac{z}{(-4)^{n-6}} \cdot \frac{(1-z)^2(1+z)^4}{\left(z^2 - 6z + 1\right)^4} \\
& \quad + \frac{z}{(-4)^{n-8}} \cdot \frac{(1-z)^2(1+z)^4\left(z^2 - 6z + 1\right)^4}{\left(1 + 20z - 26z^2 + 20z^3 + z^4\right)^4} + \cdots ,
\end{aligned}
\tag{2.44}
$$

where we discover some "additive structure" of these successive rational functions.

In fact, due to the specificity of this elliptic curve (occurrence of complex multiplication), we have another remarkable rational transformation solution of (2.19), preserving covariantly Ω. Let us introduce the rational transformation (i denotes $\sqrt{-1}$):

$$T(z) = z \cdot \left(\frac{z-(1+2i)}{1-(1+2i)\cdot z}\right)^4, \tag{2.45}$$

we also have the remarkable covariance [12]:

$${}_2F_1\left(\left[\frac{1}{2},\frac{1}{4}\right],\left[\frac{5}{4}\right];z\right) = \frac{1-z/(1+2i)}{1-(1+2i)z} \cdot {}_2F_1\left(\left[\frac{1}{2},\frac{1}{4}\right],\left[\frac{5}{4}\right];T(z)\right), \tag{2.46}$$

which can be rewritten in a simpler way on (2.3) (see (2.4)).

It is a straightforward matter to see that $T(z)$ actually belongs to the $R_{a_1}(z)$ one-parameter family:

$$\begin{gathered} T(z) = R_{a_1}(z) = -(7+24i)\cdot z + \cdots, \quad a_1 = -25\cdot\rho, \\ \rho = \frac{(7+24i)}{25}, \qquad |\rho| = 1. \end{gathered} \tag{2.47}$$

As far as the reduction of (2.32) to a rational function is concerned, it is straightforward to see that:

$$\begin{aligned} &(1-z)^2\cdot(1+z)^4\cdot R_{a_1}(z) \\ &\quad = a_1\cdot z + \cdots - \frac{2}{175746796875}\cdot a_1\cdot(a_1-1)\cdot(a_1+4)\cdot(a_1-16)\cdot P_8(a_1)\cdot z^8 + \cdots \\ &\qquad - \frac{1}{N(n)}\cdot a_1\cdot(a_1-1)\cdot(a_1+4)\cdot(a_1-16)\cdot P_n(a_1)\cdot z^n + \cdots, \end{aligned} \tag{2.48}$$

where $N(n)$ is a large integer growing with n, and P_n is a polynomial with integer coefficients of degree $n-4$, or

$$\begin{aligned} &(1-(1+2i)\cdot z)^4\cdot R_{a_1}(z) \\ &\quad = a_1\cdot z + \cdots - \frac{4}{1243125}\cdot a_1\cdot(a_1-1)\cdot(a_1+7+24i)\cdot(P_6(a_1)+iQ_6(a_1))\cdot z^6 + \cdots \\ &\qquad + \frac{1}{N(n)}\cdot a_1\cdot(a_1-1)\cdot(a_1+7+24i)\cdot(P_n(a_1)+iQ_n(a_1))\cdot z^n + \cdots, \end{aligned} \tag{2.49}$$

where P_n and Q_n are two polynomials with integer coefficients of degree, respectively, $n-3$ and $n-4$.

Similar calculations can be performed for $T^*(z)$ defined by

$$T^*(z) = z\cdot\left(\frac{z-(1-2i)}{(1-2i)z-1}\right)^4, \tag{2.50}$$

for which we also have the covariance

$$ {}_2F_1\left(\left[\frac{1}{2},\frac{1}{4}\right],\left[\frac{5}{4}\right];z\right) = \frac{1-z/(1-2i)}{1-(1-2i)z}\cdot {}_2F_1\left(\left[\frac{1}{2},\frac{1}{4}\right],\left[\frac{5}{4}\right];T^*(z)\right). \quad (2.51) $$

It is a simple calculation to check that any iterate of $T(z)$ (resp. $T^*(z)$) is actually a solution of (2.19) and corresponds to $R_{a_1}(z)$ for the infinite number of values $a_1 = (-7-24i)^N$ (resp. $(-7+24i)^N$). Furthermore, one verifies, as it should (see (2.33)), that the three rational functions $R_{-4}(z)$, $T(z)$, and $T^*(z)$ commute. It is also a straightforward calculation to see that the rational function built from any composition of $R_{-4}(z)$, $T(z)$, and $T^*(z)$ is actually a solution of (2.19). We thus have a *triple infinity* of values of a_1, namely $a_1 = (-4)^M\cdot(-7-24i)^N\cdot(-7+24i)^P$ *for any integer* M, N *and* P, for which $R_{a_1}(z)$ reduces to rational functions. We are in fact describing (some subset of) the isogenies of the elliptic curve (2.2), and identifying these isogenies with a discrete subset of the renormalization group. Conversely, a functional equation like (2.19) can be seen as a way to extend the n-fold composition of a rational function $R(z)$ (namely $R(R(\cdots R(z)\cdots)))$ to *n any complex number*.

2.3. Revisiting the One-Parameter Family of Solutions of the "Rota-Baxter-Like" Functional Equation

This extension can be revisited as follows. Keeping in mind the well-known example of the parametrization of the standard map $z \to 4z\cdot(1-z)$ with $z = \sin^2(\theta)$, yielding $\theta \to 2\theta$, let us seek for a (transcendental) parametrization $z = P(u)$ such that

$$ R_{-4}(P(u)) = P(-4u) \quad \text{or}:\ R_{-4} = P\cdot H_{-4}\cdot P^{-1}, \quad (2.52) $$

where H_{a_1} denotes the scaling transformation $z \to a_1\cdot z$ (here $H_{-4} : z \to -4\cdot z$) and P^{-1} denotes the inverse transformation of P (for the composition). One can easily find such a (transcendental) parametrization order by order

$$ \begin{aligned} P(z) = {}& z - \frac{2}{5}z^2 + \frac{7}{75}z^3 - \frac{82}{4875}z^4 + \frac{1078}{414375}z^5 \\ & - \frac{452}{1243125}z^6 + \frac{57311}{1212046875}z^7 - \frac{1023946}{175746796875}z^8 + \cdots, \end{aligned} \quad (2.53) $$

and similarly for its inverse (for the composition) transformation

$$ \begin{aligned} Q(z) = P^{-1}(z) = {}& z + \frac{2}{5}z^2 + \frac{17}{75}z^3 + \frac{244}{1625}z^4 + \frac{45043}{414375}z^5 \\ & + \frac{2302}{27625}z^6 + \frac{128941}{1939275}z^7 + \frac{15365176}{281194875}z^8 + \cdots. \end{aligned} \quad (2.54) $$

This approach is reminiscent of the conjugation introduced in Siegel's theorem [15–17]. It is a straightforward matter to see (order by order) that one actually has

$$ R_{a_1}(P(u)) = P(a_1\cdot u) \quad \text{or}:\ R_{a_1} = P\cdot H_{a_1}\cdot P^{-1}. \quad (2.55) $$

The structure of the (one-parameter) renormalization group and the extension of the composition of n times a rational function $R(z)$ (namely, $R(R(\cdots R(z)\cdots)))$ to n *any complex number* become a straight consequence of this relation. Along this line one can define some "infinitesimal composition" ($\epsilon \simeq 0$):

$$R_{1+\epsilon}(z) = P \cdot H_{1+\epsilon} \cdot P^{-1}(z) = z + \epsilon \cdot F(z) + \cdots , \tag{2.56}$$

where one can find, order by order, the "infinitesimal composition" function $F(z)$:

$$\begin{aligned} F(z) = z - \frac{2}{5}z^2 - \frac{2}{15}z^3 - \frac{14}{195}z^4 - \frac{154}{3315}z^5 - \frac{22}{663}z^6 \\ - \frac{418}{16575}z^7 - \frac{9614}{480675}z^8 - \frac{2622}{160225}z^9 + \cdots . \end{aligned} \tag{2.57}$$

It is straightforward to see, from (2.33), that the function $F(z)$ satisfies the following functional equations involving a rational function $R(z)$ (in the one-parameter family $R_{a_1}(z)$):

$$\begin{gathered} \frac{dR(z)}{dz} \cdot F(z) = F(R(z)), \qquad \frac{dR^{(n)}(z)}{dz} \cdot F(z) = F\left(R^{(n)}(z)\right), \quad \text{where:} \\ R^{(n)}(z) = R(R(\cdots R(z))\cdots). \end{gathered} \tag{2.58}$$

$F\ (z)$ cannot be a rational or algebraic function. Let us consider the fixed points of $R^{(n)}(z)$. Generically $dR^{(n)}(z)/dz$ is not equal to 0 or ∞ at any of these fixed points. Therefore one must have $F(z) = 0$ or $F(z) = \infty$ for the infinite set of these fixed points: $F(z)$ cannot be a rational or algebraic function, it is a transcendental function, and similarly for the parametrization function $P(z)$. In fact, let us introduce the function

$$\begin{aligned} G(z) &= (1-z) \cdot F(z), \\ G(z) &= z - \frac{7}{5}z^2 + \frac{4}{15}z^3 + \frac{4}{65}z^4 + \frac{28}{1105}z^5 + \frac{44}{3315}z^6 + \frac{44}{5525}z^7 \\ &\quad + \frac{836}{160225}z^8 + \frac{1748}{480675}z^9 + \cdots + g_n \cdot z^n + \cdots . \end{aligned} \tag{2.59}$$

One actually finds that the successive g_n satisfies the very simple (hypergeometric function) relation:

$$\frac{g_{n+1}}{g_n} = \frac{4n-9}{4n+1}. \tag{2.60}$$

The function $G(z)$ is actually the hypergeometric function solution of the homogeneous operator

$$D_z^2 + \frac{1}{4}\frac{13z-3}{z\cdot(1-z)} \cdot D_z + \frac{3}{4}\frac{6z^2-3z+1}{(1-z)^2 \cdot z^2}, \tag{2.61}$$

or of the inhomogeneous ODE

$$4z \cdot (1-z) \cdot \frac{dG(z)}{dz} + (9z-3) \cdot G(z) - z \cdot (1-z)^2 = 0. \tag{2.62}$$

One deduces the expression of $F(z)$ as a hypergeometric function

$$F(z) = z \cdot (1-z)^{1/2} \cdot {}_2F_1\left(\left[\frac{1}{4}, \frac{1}{2}\right], \left[\frac{5}{4}\right]; z\right) = \left.\frac{\partial R_{a_1}}{\partial a_1}\right|_{a_1=1}. \tag{2.63}$$

Finally we get the linear differential operator annihilating $F(z)$

$$\Omega_F = D_z^2 + \frac{1}{4} \cdot \frac{5z-3}{z(1-z)} \cdot D_z + \frac{1}{4} \cdot \frac{3-6z+5z^2}{(1-z)^2 z^2} = D_z \cdot \left(D_z - \frac{1}{4} \cdot \frac{3-5z}{z \cdot (1-z)}\right), \tag{2.64}$$

which is, in fact, nothing but Ω^* the *adjoint of* linear differential operator Ω (see (2.6)). One easily checks[19] that the second-order differential equation $\Omega_F(y(z)) = 0$ transforms under the change of variable $z \longrightarrow -4z/(1-z)^2$ into the second-order differential equation $\Omega_F^{(R)}(y(z)) = 0$ with $\Omega_F^{(R)} = C(z)^2 \cdot \omega_F^{(R)}$ where the unitary (monic) operator $\omega_F^{(R)}$ is the conjugate of Ω_F:

$$\begin{aligned} \omega_F^{(R)} &= D_z^2 - \frac{1}{4} \cdot \frac{11z^2+30z+3}{z \cdot (1-z)(1+z)} \cdot D_z + \frac{1}{4} \cdot \frac{3+12z+50z^2+12z^3+3z^4}{z^2 \cdot (1-z)^2(1+z)^2} \\ &= \left(\frac{1}{C(z)}\right) \cdot D_z \cdot \left(D_z - \frac{1}{4} \cdot \frac{3-5z}{z \cdot (1-z)}\right) \cdot C(z) \\ &= \left(\frac{1}{C(z)}\right) \cdot \Omega_F \cdot C(z) = \left(\frac{1}{C(z)}\right) \cdot \Omega^* \cdot C(z) \end{aligned} \tag{2.65}$$

with $C(z) = 1/R'(z)$ and the "dot" denotes the composition of operators. Actually, the factors in the adjoint Ω^* transform under the change of variable $z \longrightarrow -4z/(1-z)^2$ as follows[20]:

$$D_z \longrightarrow C(z) \cdot D_z, \qquad \omega_1^* = \longrightarrow \left(\omega_1^*\right)^{(R)} = \omega_1^* \cdot C(z), \qquad \Omega^* \longrightarrow \Omega_F^{(R)} = C(z) \cdot \Omega^* \cdot C(z) \tag{2.66}$$

which is precisely the transformation we need to match with (2.58) and see the ODE $\Omega^*(F(z)) = 0$ compatible with the change of variable $z \longrightarrow -4z/(1-z)^2$:

$$\begin{aligned} \Omega^*(F(z)) = 0 &\longrightarrow (C(z) \cdot \Omega^* \cdot C(z))(F(R(z))) \\ &= (C(z) \cdot \Omega^* \cdot C(z))\left(R'(z) \cdot F(z)\right) = C(z) \cdot \Omega^*(F(z)) = 0. \end{aligned} \tag{2.67}$$

This is, in fact, a quite general result that will be seen to be valid in a more general (higher genus) framework (see (2.148), (2.150) in what follows).

Not surprisingly one can deduce from (2.33) and the previous results, in particular (2.63), the following results for $R_{a_1}(z)$:

$$-4 \cdot \frac{\partial R_{a_1}}{\partial a_1}\bigg|_{a_1=-4} = F(R(z)), \qquad (-4)^n \cdot \frac{\partial R_{a_1}}{\partial a_1}\bigg|_{a_1=(-4)^n} = F\Big(R^{(n)}(z)\Big), \tag{2.68}$$

where $R(z) = -4z/(1-z)^2$ and $R^{(n)}(z)$ denotes $R(R(\cdots R(R(z))))$. Of course we have similar relation for $T(z)$, -4 being replaced by $-7-24i$. Therefore the partial derivative $\partial R_{a_1}/\partial a_1$ that can be expressed in terms of hypergeometric functions for for a *double infinity* of values of a_1, namely, $a_1 = (-4)^M \times (-7-24i)^N$.

One can, of course, check, order by order, that (2.58) is actually verified for any function in the one-parameter family $R_{a_1}(z)$:

$$\frac{dR_{a_1}(z)}{dz} \cdot F(z) = F(R_{a_1}(z)), \tag{2.69}$$

which corresponds to an infinitesimal version of (2.33).

From (2.56) one simply deduces

$$z \cdot \frac{dP(z)}{dz} = F(P(z)), \tag{2.70}$$

that we can check, order by order from (2.53), the series expansion of $P(z)$, and from (2.57) the series expansion of $F(z)$, but also

$$\frac{dQ(z)}{dz} \cdot F(z) = Q(z) \tag{2.71}$$

that we can, check order by order, from (2.54), the series expansion of $Q(z) = P^{-1}(z)$ and from (2.57). We now deduce that the log-derivative of the "well-suited change of variable" $Q(z)$ is nothing but the (multiplicative) inverse of a hypergeometric function $F(z)$:

$$\frac{d\ln(Q(z))}{dz} = \frac{1}{F(z)}, \qquad Q(z) = \lambda \cdot \exp\left(\int^z \frac{dz}{F(z)}\right). \tag{2.72}$$

The function $Q(z)$ is solution of the *nonlinear* differential equation

$$\begin{aligned} &-4z^2 \cdot (1-z)^2 \cdot \left(Q \cdot Q^{(1)} \cdot Q^{(3)} + \left(Q^{(1)}\right)^2 \cdot Q^{(2)} - 2Q \cdot \left(Q^{(2)}\right)^2\right) \\ &\quad + z \cdot (3-5z)(1-z) \cdot Q^{(1)} \cdot \left(Q \cdot Q^{(2)} - \left(Q^{(1)}\right)^2\right) \\ &\quad + \left(5z^2 - 6z + 3\right) \cdot Q \cdot \left(Q^{(1)}\right)^2 = 0, \end{aligned} \tag{2.73}$$

where the $Q^{(n)}$'s denote the nth derivative of $Q(z)$. At first sight $Q(z)$ would be a *nonholonomic* function, however, remarkably, it is a *holonomic* function solution of an order-five operator which factorizes as follows:

$$\begin{aligned}\Omega_Q = &\left(D_z + \frac{3-5z}{(1-z)\cdot z}\right)\cdot\left(D_z + \frac{3}{4}\cdot\frac{3-5z}{(1-z)\cdot z}\right)\cdot\\ &\times\left(D_z + \frac{2}{4}\cdot\frac{3-5z}{(1-z)\cdot z}\right)\cdot\left(D_z + \frac{1}{4}\cdot\frac{3-5z}{(1-z)\cdot z}\right)\cdot D_z,\end{aligned} \tag{2.74}$$

yielding the exact expression of $Q(z)$ in terms of hypergeometric functions:

$$Q(z) = z\cdot\left({}_2F_1\left(\left[\frac{1}{2},\frac{1}{4}\right],\left[\frac{5}{4}\right];z\right)\right)^4 = \frac{z}{1-z}\cdot\left({}_2F_1\left(\left[\frac{1}{4},\frac{3}{4}\right],\left[\frac{5}{4}\right];-\frac{z}{1-z}\right)\right)^4, \tag{2.75}$$

that is, the fourth power of (2.3), with the differential operator (2.74) being the symmetric fourth power of Ω. From (2.3) we immediately get the covariance of $Q(z)$:

$$Q\left(-\frac{4z}{(1-z)^2}\right) = -4\cdot Q(z), \tag{2.76}$$

and, more generally, $Q(R_{a_1}) = a_1\cdot Q(z)$. Since $Q(z)$ and $F(z)$ are expressed in terms of the same hypergeometric function, the relation (2.71) must be an identity on that hypergeometric function. This is actually the case. This hypergeometric function verifies the inhomogeneous equation:

$$4\cdot z\cdot\frac{d\mathcal{H}(z)}{dz} + \mathcal{H}(z) - (1-z)^{-1/2} = 0, \tag{2.77}$$

where

$$\mathcal{H}(z) = {}_2F_1\left(\left[\frac{1}{2},\frac{1}{4}\right],\left[\frac{5}{4}\right];z\right). \tag{2.78}$$

Recalling $Q(P(z)) = z$, one has the following functional relation on $P(z)$:

$$P(z)\cdot{}_2F_1\left(\left[\frac{1}{4},\frac{1}{2}\right],\left[\frac{5}{4}\right];P(z)\right)^4 = z. \tag{2.79}$$

Noting that $Q(z^4)^{1/4} = \mathcal{F}(z^4)$ (see (2.3)) can be expressed in term of an incomplete elliptic integral of the first kind of argument $\sqrt{-1}$

$$z\cdot{}_2F_1\left(\left[\frac{1}{4},\frac{1}{2}\right],\left[\frac{5}{4}\right];z^4\right) = \mathrm{Elliptic}\,F\left(z,\sqrt{-1}\right), \tag{2.80}$$

one can find that (2.79) rewrites on $P(z)$ as

$$\text{Elliptic}\, F\left(P(z)^{1/4}, \sqrt{-1}\right) = z^{1/4}, \tag{2.81}$$

from which we deduce that the function $P(z)$ is nothing but a *Jacobi elliptic function* [21]

$$P(z) = \left(\mathrm{sn}(z^{1/4}, \sqrt{-1})\right)^4. \tag{2.82}$$

In Appendix B we display a set of "Painlevé-like" ODEs[22] verified by $P(z)$. From the simple nonlinear ODE on the Jacobi elliptic sinus, namely, $S'' + 2 \cdot S^3 = 0$, and the exact expression of $P(z)$ in term of Jacobi elliptic sinus, one can deduce other *nonlinear* ODEs verified by the *nonholonomic* function $P(z)$ ($P^{(1)} = dP(z)/dz$, $P^{(2)} = d^2P(z)/dz^2$):

$$z^{3/2} \cdot \left(P^{(1)}\right)^2 - (1 - P) \cdot P^{3/2} = 0, \tag{2.83}$$

$$P^{(2)} - \frac{3}{4} \cdot \frac{\left(P^{(1)}\right)^2}{P} + \frac{3}{4} \cdot \frac{P^{(1)}}{z} + \frac{1}{2} \cdot \frac{P^{3/2}}{z^{3/2}} = 0. \tag{2.84}$$

2.4. Singularities of the Jacobi Elliptic Function $P(z)$

Most of the results of this section, and to some extent, of the next one, are straight consequences of the exact closed expression of $P(z)$ in terms of an elliptic function. Following the pedagogical approach of this paper we will rather follow a heuristic approach not taking into account the exact result (2.82), to display simple methods and ideas that can be used beyond exact results on a specific example.

From a diff-Padé analysis of the series expansion of $P(z)$, we got the sixty (closest to $z = 0$) singularities. In particular we got that $P(z)$ has a radius of convergence $R \simeq 11.81704500807\cdots$ corresponding to the following (closest to $z = 0$) singularity $z = z_s$ of $P(z)$:

$$\begin{aligned} z_s &= -11.817045008077115768316337283432582087420697\cdots \\ &= (-4) \cdot {}_2F_1\left(\left[\frac{1}{4}, \frac{1}{2}\right], \left[\frac{5}{4}\right]; 1\right)^4 = -\frac{1}{16} \cdot \frac{\pi^6}{\Gamma(3/4)^8}. \end{aligned} \tag{2.85}$$

This singularity corresponds to a pole of order four: $P(z) \simeq (z - z_s)^{-4}$. The function $P(z)$ has many other singularities:

$$\begin{aligned} &3^4 \cdot z_s, (161 \pm 240i) \cdot z_s, (-7 \pm 24i) \cdot z_s, (-119 \pm 120i) \cdot z_s, \cdots \\ &5^4 \cdot z_s, (41 \pm 840i) \cdot z_s, (-527 \pm 336i) \cdot z_s, (-1519 \pm 720i) \cdot z_s, \cdots \\ &7^4 \cdot z_s, (1241 \pm 2520i) \cdot z_s, (-567 \pm 1944i) \cdot z_s, (-3479 \pm 1320i) \cdot z_s, \cdots \end{aligned} \tag{2.86}$$

In fact, introducing x and y the real and imaginary part of these singularities in z_s units, one finds out that they correspond to the double infinity of points

$$x = \left(m_1^2 - 2m_1m_2 - m_2^2\right) \cdot \left(m_1^2 + 2m_1m_2 - m_2^2\right), \\ y = 4 \cdot m_1m_2 \cdot (m_2 - m_1) \cdot (m_2 + m_1), \tag{2.87}$$

where m_1 and m_2 are two integers, and they all lie on the intersection of an infinite number of genus zero curves indexed by the fourth power of an integer $M = m^4$ ($m = m_1$ or $m = m_2$):

$$2^{12}M^4 - 2^{11} \cdot x \cdot M^3 - 2^7 \cdot \left(17y^2 + 14x^2\right) \cdot M^2 - 2^5 \cdot x \cdot \left(8x^2 + 7y^2\right) \cdot M + y^4 = 0. \tag{2.88}$$

The parametrization (2.87) describes not only the poles of $P(z)$ when $m_1 + m_2$ is odd, but also the zeros of $P(z)$ when $m_1 + m_2$ is even. This (infinite) proliferation of singularities confirms the nonholonomic character of $P(z)$.

These results are simply inherited from (2.82). The zeros and poles of the elliptic sinus $sn(z, i)$ correspond to two lattice of periods. Denoting $\mathcal{K}_1$ and $\mathcal{K}_2$ the two periods of the elliptic curve, the location of the poles and zeros reads, respectively,

$$P_{n_1,n_2} = 2n_1 \cdot \mathcal{K}_1 + (2n_2 + 1) \cdot \mathcal{K}_2, \\ Z_{n_1,n_2} = 2n_1 \cdot \mathcal{K}_1 + 2n_2 \cdot \mathcal{K}_2, \\ \mathcal{K}_1 = \frac{\pi^{3/2}}{2^{3/2}} \cdot \frac{1}{\Gamma(3/4)^2}, \quad \mathcal{K}_2 = \left(1 - \sqrt{-1}\right) \cdot K_1 \tag{2.89}$$

making crystal clear the fact that we have complex multiplication for this elliptic curve. The formula (2.87) just amount to saying that the poles and zeros of $sn(z^{1/4}, i)$ are located at $P^4_{n_1,n_2}$ and $Z^4_{n_1,n_2}$:

$$P^4_{n_1,n_2} = -\frac{z_s}{4} \cdot ((2n_1 + 2n_2 + 1) + i \cdot (2n_2 + 1))^4, \\ Z^4_{n_1,n_2} = -\frac{z_s}{4} \cdot ((2n_1 + 2n_2 + 1) + i \cdot 2n_2)^4. \tag{2.90}$$

The correspondence with (2.87) is $m_1 = n_1 + 2n_2 + 1$, $m_2 = -n_1$ for the poles and $m_1 = n_1 + 2n_2$, $m_2 = -n_1$ for the zeros.

Remark 2.2. let us consider the $a_1 \to \infty$ limit of the one-parameter series R_{a_1} (see (2.31), (2.32)) rewriting $R_{a_1}(z)$ as $\widetilde{R}_{b_1}(u)$

$$\widetilde{R}_{b_1}(u) = R_{a_1}(z), \quad \text{with}: \ b_1 = \frac{1}{a_1}, \ u = \frac{z}{b_1}. \tag{2.91}$$

In the $a_1 \to \infty$ limit, that is the $b_1 \to 0$ limit, one easily verifies, order by order in u, that $\widetilde{R}_{b_1}(u)$ becomes *exactly* the transcendental parametrization function (2.53):

$$\widetilde{R}_{b_1}(u) \longrightarrow P(u) \quad \text{when } b_1 \longrightarrow 0. \tag{2.92}$$

For $a_1 = (-4)^n$ ($n \to \infty$), one finds that the radius of convergence of the $R_{a_1}(z)$ series becomes in the $n \to \infty$ limit $R_n \simeq z_s/4^n$, in agreement[23] with (2.91).

2.5. $P(z)$ ***and an Infinite Number of Rational Transformations: The Sky Is the Limit***

Note that some nonlinear ODEs associated with $P(z)$ and displayed in Appendix B, namely (B.3) and (B.10), and the functional equation (2.79), are invariant by the change of variable $(P(z), z) \to (-4P(z)/(1-P(z))^2, -4z)$. In fact (B.3), (2.79), and (B.10) are invariant by $(P(z), z) \to (-4P(z)/(1-P(z))^2, -4z)$, but also $(P(z), z) \to (-(1-P(z))^2/4/P(z), -z/4)$, and also by $(P(z), z) \to (1/P(z), z)$.

The function $P(z)$ satisfies the functional equation:

$$P(-4 \cdot z) = -\frac{4P(z)}{1-P(z)^2}, \tag{2.93}$$

but also

$$\begin{aligned} P((-7-24i)\cdot z) &= T(P(z)), \\ P((-7+24i)\cdot z) &= T^*(P(z)), \end{aligned} \tag{2.94}$$

and, more generally, as can be checked order by order on series expansions

$$P(a_1 \cdot z) = R_{a_1}(P(z)). \tag{2.95}$$

For example, considering the "good" branch (2.35) for the inverse of $-4z/(1-z)^2$, namely $S^{(1)}_{-1/4}(z)$, we can even check, order by order, on the series expansions of $P(z)$ and $S^{(1)}_{-1/4}(z)$ the functional relation

$$S^{(1)}_{-1/4}(P(z)) = P\left(-\frac{z}{4}\right), \tag{2.96}$$

valid for $|P(z)| < 1$ *since the radius of convergence of* $S^{(1)}_{-1/4}(z)$ *is* 1.

Recalling the functional equations (2.94) it is natural to say that if $P(z)$ is singular at $z = z_s$, then, for almost all the rational functions, in particular $T(z)$ (resp. $T^*(z)$) the $T(P(z))$ is also singular $z = z_s$, and thus, from (2.94), $P(z)$ is also singular at $z = (-7 \pm 24i) \cdot z_s$. It is thus extremely natural to see the emergence of the infinite number of singularities in (2.87) of the form $z = (N_1 + i \cdot N_2) \cdot z_s$, as a consequence of (2.95) together with a reduction of the one-parameter series $R_{a_1}(z)$ to a rational function for an infinite number of selected values of a_1, namely the $N_1 + i \cdot N_2$ in (2.87). This is actually the case for all the values displayed in (2.87). For instance, for $a_1 = 3^4 = 81$ we get the following simple rational function:

$$R_{81}(z) = z \cdot \left(\frac{z^2 + 6z - 3}{3z^2 - 6z - 1} \right)^4, \tag{2.97}$$

for which it is straightforward to verify that this rational transformation commutes with $T(z)$, $T^*(z)$, $-4z/(1-z)^2$, and is a solution of the Rota-Baxter-like functional equation (2.19). The case $a_1 = 5^4 = 625$ in (2.87), is even simpler, since it just requires to compose $T(z)$ and $T^*(z)$

$$\begin{aligned} R_{625}(z) &= T(T^*(z)) = T^*(T(z)) \\ &= z \cdot \left(\frac{z^2 - 2z + 5}{5z^2 - 2z + 1} \right)^4 \cdot \left(\frac{1 - 12z - 26z^2 + 52z^3 + z^4}{1 + 52z - 26z^2 - 12z^3 + z^4} \right)^4, \end{aligned} \tag{2.98}$$

which, again verifies (2.19) and commutes with all the other rational functions, in particular (2.97). We also obtained the rational function corresponding to $a_1 = 7^4 = 2401$, namely:

$$R_{2401}(z) = z \cdot \left(\frac{N_{2401}(z)}{D_{2401}(z)} \right)^4, \tag{2.99}$$

with:

$$N_{2401}(z) = z^{12} \cdot D_{2401}\left(\frac{1}{z}\right), \tag{2.100}$$

$$\begin{aligned} D_{2401}(z) = {} & 1 + 196z - 1302z^2 + 14756z^3 - 15673z^4 - 42168z^5 \\ & + 111916z^6 - 82264z^7 + 35231z^8 - 19852z^9 \\ & + 2954z^{10} + 308z^{11} - 7z^{12}. \end{aligned} \tag{2.101}$$

The polynomial $N_{2401}(z)$ satisfies many functional equations, like, for instance (with $R_{-4}(z) = -4z/(1-z^2)$):

$$4^{12} \cdot D_{2401}\left(\frac{1}{R_{-4}(z)}\right) = D_{2401}(z) \cdot D_{2401}\left(\frac{1}{z}\right) \tag{2.102}$$

and also

$$(1-z)^{49} \cdot D_{2401}(R_{-4}(z))^2 = D_{2401}(z)^4 - z^{49} \cdot D_{2401}\left(\frac{1}{z}\right)^4. \tag{2.103}$$

We also obtained the rational function corresponding to $a_1 = 11^4 = 14641$, namely,

$$R_{14641}(z) = z \cdot \left(\frac{N_{14641}(z)}{D_{14641}(z)}\right)^4, \tag{2.104}$$

with:

$$N_{14641}(z) = z^{30} \cdot D_{14641}\left(\frac{1}{z}\right), \tag{2.105}$$

$$\begin{aligned} D_{14641}(z) = 1 &+ 1210z - 33033z^2 + 2923492z^3 + 5093605z^4 \\ &- 385382514z^5 + 3974726283z^6 - 14323974808z^7 \\ &+ 57392757037z^8 - 291359180310z^9 + 948497199067z^{10} \\ &- 1642552094436z^{11} + 1084042069649z^{12} + 1890240552750z^{13} \\ &- 6610669151537z^{14} + 9712525647792z^{15} - 8608181312269z^{16} \\ &+ 5384207244702z^{17} - 3223489742187z^{18} + 2175830922716z^{19} \\ &- 1197743580033z^{20} + 387221579866z^{21} - 50897017743z^{22} \\ &- 7864445336z^{23} + 5391243935z^{24} - 815789634z^{25} \\ &+ 28366041z^{26} - 5092956z^{27} + 207691z^{28} + 2794z^{29} - 11z^{30}, \end{aligned} \tag{2.106}$$

and, of course, one can verify that $R_{14641}(z)$ actually commutes with R_{-4}, R_{81}, R_{625}, $R_{2401}(z)$, and is a solution of the Rota-Baxter-like functional equation (2.19). Similarly to $R_{2401}(z)$ (see (2.102), (2.103)), we also have the functional equations:

$$4^{30} \cdot D_{14641}\left(\frac{1}{R_{-4}(z)}\right) = D_{14641}(z) \cdot D_{14641}\left(\frac{1}{z}\right), \tag{2.107}$$

and also

$$(1-z)^{(4\cdot 30+1)} \cdot D_{14641}(R_{-4}(z))^2 = D_{14641}(z)^4 - z^{(4\cdot 30+1)} \cdot D_{14641}\left(\frac{1}{z}\right)^4. \tag{2.108}$$

Next we obtained the rational function corresponding to $a_1 = 13^4 = 28561$, which verifies (2.19) namely,

$$\begin{aligned} &R_{28561}(z) = z \cdot \left(\frac{N_{28561}(z)}{D_{28561}(z)}\right)^4, \quad \text{with } N_{28561}(z) = z^{42} \cdot D_{28561}\left(\frac{1}{z}\right), \\ &N_{28561}(z) = z^{42} \cdot D_{28561}\left(\frac{1}{z}\right), \end{aligned} \tag{2.109}$$

$$
\begin{aligned}
D_{28561}(z) &= \left(1 - 22z + 235z^2 - 228z^3 + 39z^4 + 26z^5 + 13z^6\right) \cdot D^{(36)}_{28561}(z) \\
D^{(36)}_{28561}(z) &= 1 + 2388z - 61098z^2 + 19225300z^3 + 606593049z^4 \\
&\quad - 1543922656z^5 + 7856476560z^6 - 221753896032z^7 + 1621753072244z^8 \\
&\quad - 4542779886736z^9 + 2731418674664z^{10} + 36717669656304z^{11} \\
&\quad - 200879613202428z^{12} + 547249607666784z^{13} - 934179604482832z^{14} \\
&\quad + 1235038888776160z^{15} - 1788854212778642z^{16} + 3018407750933816z^{17} \\
&\quad - 4349780716415868z^{18} + 4419228090228152z^{19} - 2899766501472914z^{20} \\
&\quad + 931940880451552z^{21} + 413258559018224z^{22} - 857795672629664z^{23} \\
&\quad + 659989056851972z^{24} - 304241349909008z^{25} + 87636987790824z^{26} \\
&\quad - 14593362219920z^{27} + 1073204980340z^{28} + 45138167200z^{29} \\
&\quad - 23660433008z^{30} + 2028597792z^{31} - 29540327z^{32} + 3238420z^{33} \\
&\quad - 73386z^{34} - 492z^{35} + z^{36}.
\end{aligned}
\tag{2.110}
$$

We get similar results, mutatis mutandis, than the ones previously obtained (commutation, functional equations like (2.107), (2.108), etc.), namely,

$$
4^{42} \cdot D_{28561}\left(\frac{1}{R_{-4}(z)}\right) = D_{28561}(z) \cdot D_{28561}\left(\frac{1}{z}\right), \ldots. \tag{2.111}
$$

The "palindromic" nature of (2.97) (2.98), (2.99), (2.104) and (2.109) (see (2.100), (2.105)), (2.109)) corresponds to the fact that these rational transformations commute with J:

$$
\frac{1}{R_{81}(z)} = R_{81}\left(\frac{1}{z}\right), \quad \frac{1}{R_{625}(z)} = R_{625}\left(\frac{1}{z}\right), \ldots. \tag{2.112}
$$

In fact, more generally, we have $R_{N^4}(1/z) = 1/R_{N^4}(z)$ for N any odd integer ($N = 9, 21, \ldots$) and $R_{N^4}(1/z) = R_{N^4}(z)$ for any even integer N.

From (2.87) one can reasonnably conjecture that the fourth power of *any integer* will provide a new example of $R_{a_1}(z)$ being a rational function. The simple nontrivial example corresponds to the already found rational function

$$
R_{16}(z) = 16 \cdot \frac{z \cdot (1 - z)^2}{(z + 1)^4}. \tag{2.113}
$$

We already have explicit rational functions for all values of a_1 of the form N^4 for $N = 2, 3, \ldots, 16$ and of course, we can in principle, build explicit rational functions for all the N's product of the previous integers. Along this line it is worth noticing that the coefficients of the series $R_{a_1}(z)$ are all integers when a_1 is the fourth power of any integer.

We are thus starting to build an infinite number of (elementary) *commuting* rational transformations, *any* composition of these (infinite number of) rational transformations giving rational transformations satisfying (2.19) and preserving the linear differential operator Ω. This set of rational transformations is a pretty large set! Actually this set of rational transformations corresponds to the isogenies of the underlying elliptic function.

The proliferation of the singularities of $P(z)$ corresponds to this (pretty large) set of rational transformations. Recalling (2.96), the previous singularity argument is *not valid* [24] for the (well-suited) inverse transformations ($S^{(1)}_{-1/4}(z), \ldots$) of these rational transformations because (2.96) requires $|P(z)| < 1$ (corresponding to the radius of convergence of $S^{(1)}_{-1/4}(z)$) and the singularity $z = z_s$ corresponds precisely to "hit" the value $P(z) = 1$.

2.6. Other Examples of Selected Gauss Hypergeometric ODEs

For heuristic reasons we have focused on $A(z) = (3-5z)/z/(1-z)/4$, but of course, one can find many other examples and try to generalize these examples.

For instance, introducing

$$A(z) = \frac{1}{6} \cdot \frac{d \ln\left((1-z)^3 z^5\right)}{dz} = \frac{1}{6} \cdot \frac{5-8z}{(1-z)z} \tag{2.114}$$

the rational transformation

$$R(z) = -27 \cdot \frac{z}{(1-4z)^3}, \tag{2.115}$$

verifies the "Rota-Baxter-like" functional relation (2.19). This example corresponds to the following covariance [12] on a Gauss hypergeometric integral (of the $c = 1 + b$ type, see below):

$$\begin{aligned} {}_2F_1\left(\left[\frac{1}{2}, \frac{1}{6}\right], \left[\frac{7}{6}\right]; z\right) &= \frac{z^{-1/6}}{6} \cdot \int_0^z t^{-5/6}(1-t)^{-1/2} \cdot dt \\ &= (1-4z)^{-1/2} \cdot {}_2F_1\left(\left[\frac{1}{2}, \frac{1}{6}\right], \left[\frac{7}{6}\right]; -27\frac{z}{(1-4z)^3}\right) \end{aligned} \tag{2.116}$$

which is associated with the elliptic curve

$$y^6 - (1-t)^3 \cdot t^5 = 0. \tag{2.117}$$

Another example (of the $c = 1 + a$ type, see below) is

$$A(z) = \frac{1}{3} \cdot \frac{d \ln\left((1-z)^2 z^2\right)}{dz} = \frac{2}{3} \cdot \frac{1-2z}{(1-z)z}, \tag{2.118}$$

where the rational transformation

$$R(z) = \frac{z \cdot (z-2)^3}{(1-2z)^3} = -8z - 36z^2 - 126z^3 - 387z^4 + \cdots \tag{2.119}$$

verifies the "Rota-Baxter-like" functional relation (2.19). This example corresponds to the following covariance [12] on a Gauss hypergeometric integral:

$$\begin{aligned} {}_2F_1\left(\left[\frac{1}{3}, \frac{2}{3}\right], \left[\frac{4}{3}\right]; z\right) &= \frac{z^{-1/3}}{3} \cdot \int_0^z t^{-2/3}(1-t)^{-2/3} \cdot dt \\ &= \frac{1}{2} \cdot \frac{2-z}{1-2z} \cdot {}_2F_1\left(\left[\frac{1}{3}, \frac{2}{3}\right], \left[\frac{4}{3}\right]; \frac{z \cdot (z-2)^3}{(1-2z)^3}\right), \end{aligned} \tag{2.120}$$

which is associated with the elliptic curve

$$y^3 - (1-t)^2 \cdot t^2 = 0. \tag{2.121}$$

Note that, similarly to the main example of the paper, there exist many rational transformations[25] satisfying (2.19) that cannot be reduced to iterates of (2.119), for instance,

$$\begin{aligned} T(z) = -27 \cdot \frac{z \cdot (1-z)\left(z^2 - z + 1\right)^3}{\left(z^3 + 3z^2 - 6z + 1\right)^3} &= -27z - 378z^2 \\ &- 3888z^3 - 34074z^4 - 271620z^5 - 2032209z^6 + \cdots . \end{aligned} \tag{2.122}$$

One verifies immediately that (2.122) actually verifies (2.19) with (2.118). Not surprisingly, the two rational transformations (2.119) and (2.122) commute.

Another simple example with rational symmetries corresponds to $\Omega = (D_z + A(z)) \cdot D_z$ with

$$A(z) = -\frac{1}{2} \cdot \frac{3z-1}{z(1-z)} = \frac{1}{2} \cdot \frac{d \ln\left(z \cdot (1-z)^2\right)}{dz}. \tag{2.123}$$

It has the simple (genus zero) hypergeometric solution[26]:

$$\mathcal{F}(z) = z^{1/2} \cdot {}_2F_1\left(\left[1, \frac{1}{2}\right], \left[\frac{3}{2}\right]; z\right) = \operatorname{arctanh}\left(z^{1/2}\right). \tag{2.124}$$

The linear differential operator Ω is covariant under the change of variable $z \rightarrow 1/z$ and $z \rightarrow R(z)$, where[27]

$$R(z) = \frac{4z}{(1+z)^2}. \tag{2.125}$$

One can easily check that (2.123) and (2.125) satisfy the functional equation (2.19). One also verifies that (2.123) and $z \to 1/z$ or the iterates of (2.125) satisfy the functional equation (2.19). The solution of the adjoint operator are $(1-z)\cdot z^{1/2}$ and

$$\begin{aligned} F(z) &= z\cdot(1-z)\cdot {}_2F_1\left(\left[1,\frac{1}{2}\right],\left[\frac{3}{2}\right];z\right) \\ &= z^{1/2}\cdot(1-z)\cdot \operatorname{arctanh}\left(z^{1/2}\right) = z-\frac{2}{3}z^2-\frac{2}{15}z^3 \\ &\quad -\frac{2}{35}z^4-\frac{2}{63}z^5-\frac{2}{99}z^6-\frac{2}{143}z^7+\cdots. \end{aligned} \tag{2.126}$$

One verifies, again, that (2.126) and (2.125) commute, (2.126) corresponding to the "infinitesimal composition" of (2.125) (see (2.56)).

A first natural generalization amounts to keeping the remarkable factorization (2.6) which will, in fact, reduce the covariance of a second-order operator to the covariance of a first-order operator.[28] Such a situation occurs for Gauss hypergeometric functions ${}_2F_1([a,b],[1+a];z)$ solution of the (a,b)-symmetric linear differential operator

$$z\cdot(1-z)\cdot D_z^2+(c-(a+b+1)\cdot z)\cdot D_z-a\cdot b, \tag{2.127}$$

as soon as[29] $c=1+a$. For instance

$$\mathcal{F}(z)=z^a\cdot {}_2F_1([a,b],[1+a];z), \tag{2.128}$$

is an integral of a simple algebraic function and is solution with the constant function of the second-order operator

$$\begin{aligned} \Omega &= \left(D_z+\frac{(a-b-1)z+1-a}{z\cdot(1-z)}\right)\cdot D_z \\ &= \left(D_z+\frac{d\ln\left((1-z)^b\cdot z^{1-a}\right)}{dz}\right)\cdot D_z, \end{aligned} \tag{2.129}$$

yielding a new $A(z)$:

$$A(z)=\frac{(1-a)+(a-b-1)z}{(1-z)\cdot z}=\frac{1-a}{z}-\frac{b}{1-z}. \tag{2.130}$$

The adjoint of (2.129) has the simple solution $z^{1-a}\cdot(1-z)^b$:

$$F(z)=z\cdot(1-z)^b\cdot {}_2F_1([a,b],[1+a];z). \tag{2.131}$$

Due to the (a, b)-symmetry of (2.127) we have a similar result for $c = 1+b$. The function $\mathcal{F}(z) = z^b \cdot {}_2F_1([a,b],[1+b];z)$ is solution of (2.129) where a and b have been permuted:

$$\left(D_z + \frac{(b-a-1)z + 1 - b}{z \cdot (1-z)}\right) \cdot D_z, \tag{2.132}$$

yielding another $A(z)$

$$A(z) = \frac{(1-b) + (b-a-1)z}{(1-z) \cdot z}, \tag{2.133}$$

The adjoint of (2.132) has the solution $(1-z)^a \cdot z^{1-b}$ together with the hypergeometric function:

$$F(z) = z \cdot (1-z)^a \cdot {}_2F_1([a,b],[1+b];z), \tag{2.134}$$

where one recovers the previous result (2.126).

We are seeking for (Gauss hypergeometric) second-order differential equations[30] with an *infinite number* of (hopefully rational, if not algebraic) symmetries: this is another way to say that we are not looking for generic Gauss hypergeometric differential equations, but Gauss hypergeometric differential equations *related to elliptic curves*, and thus having an infinite set of such isogenies. We are necessarily in the framework where the two parameters a and b of the Gauss hypergeometric are *rational numbers* in order to have *integral of algebraic functions* (yielding globally nilpotent [11] second-order differential operators). Let us denote by D the common denominator of the two rational numbers $a = N_a/D$ and $b = N_b/D$, function (2.128) is associated to a period of the algebraic curve

$$y^D = (1-t)^{N_b} \cdot t^{D-N_a}. \tag{2.135}$$

We just need to restrict to triplets of integers (N_a, N_b, D) such that the previous curve is an elliptic curve.

Let us give an example (of the $c = 1 + b$ type) that *does not* correspond to a genus one curve, with

$${}_2F_1\left(\left[\frac{1}{3}, \frac{1}{6}\right], \left[\frac{7}{6}\right]; z\right) = \frac{1}{6} \cdot z^{-1/6} \cdot \int_0^z t^{-5/6}(1-t)^{-1/3} \cdot dt, \tag{2.136}$$

which corresponds to the *genus two* curve:

$$y^6 - (1-t)^2 \cdot t^5 = 0. \tag{2.137}$$

Again one introduces $A(z)$

$$A(z) = \frac{1}{6} \cdot \frac{d \ln\left((1-z)^2 z^5\right)}{dz} = \frac{1}{6} \cdot \frac{5-7z}{z \cdot (1-z)} \tag{2.138}$$

and seeks for $R(z)$ as series expansions analytical at $z = 0$. One gets actually, order by order, a one-parameter family

$$\begin{aligned} R_{a_1}(z) = {} & a_1 \cdot z - \frac{2}{7} a_1 \cdot (a_1 - 1) \cdot z^2 \\ & + \frac{1}{637} a_1 \cdot (a_1 - 1) \cdot (17a_1 - 87) \cdot z^3 \\ & - \frac{2}{84721} a_1 \cdot (a_1 - 1) \cdot \left(113a_1^2 - 856a_1 + 3438\right) \cdot z^4 \\ & - \frac{1}{38548055} a_1 \cdot (a_1 - 1) \cdot \left(3674a_1^3 + 121194a_1^2 - 552261a_1 + 2095059\right) \cdot z^5 + \cdots \\ & + \frac{1+\epsilon_n}{N(n)} \cdot a_1 \cdot (a_1 - 1) \cdot P_n(a_1) \cdot z^n + \cdots, \end{aligned} \tag{2.139}$$

where $\epsilon_n = 0$ for n odd and $\epsilon_n = 1$ for n even, and $N(n)$ is a (large) integer depending on n, and $P_n(a_1)$ is a polynomial with integer coefficients of degree $n-2$. One easily verifies, order by order, that one gets a one-parameter family of transformations commuting for different values of the parameter:

$$R_{a_1}(R_{b_1}(z)) = R_{b_1}(R_{a_1}(z)) = R_{a_1 b_1}(z). \tag{2.140}$$

As far as the "algorithmic complexity" of this series (2.139) is concerned it is worth noticing that the degree growth [18] of the series coefficients is actually *linear* and not exponential as we could expect [19] at first sight. Even if this series is transcendental, it is not a "wild" series.

Seeking for selected values of a_1 such that the previous series (2.139) reduces to a rational function one can try to reproduce the simple calculations (2.48), (2.49), but unfortunately "shooting in the dark" because we have no hint of a well-suited denominator (if any!) like the polynomials in the lhs of (2.48), (2.49).

It is also worth noticing that if we slightly change $A(z)$ into

$$A(z) = \frac{1}{N} \cdot \frac{d\ln\left((1-z)^2 z^5\right)}{dz} = \frac{1}{N} \cdot \frac{5-7z}{z\cdot(1-z)}, \tag{2.141}$$

the algebraic curve (2.137) becomes $y^N - (1-t)^2 \cdot t^5 = 0$ which has, for instance genus five for $N = 11$, but genus zero for $N = 7$. For any of these cases of (2.141) one can easily get, order by order, a one-parameter series R_{a_1} totally similar to (2.139) with, again, polynomials $P_n(a_1)$ of degree $n-2$.

The first coefficient a_2 is in general

$$a_2 = -\frac{2}{2N-5} \cdot a_1 \cdot (a_1 - 1). \tag{2.142}$$

For the genus zero case, $N = 7$

$$a_2 = -\frac{2}{9} \cdot a_1 \cdot (a_1 - 1), \quad a_3 = -\frac{1}{1296} \cdot a_1 \cdot (a_1 - 1) \cdot (127 - a_1),$$
$$a_4 = -\frac{1}{134136} \cdot a_1 \cdot (a_1 - 1) \cdot \left(254a_1^2 + 185a_1 + 7499\right), \ldots, \tag{2.143}$$
$$a_n = -\frac{1}{N(n)} \cdot a_1 \cdot (a_1 - 1) \cdot P_n(a_1),$$

which corresponds to the solution

$$\frac{2}{7} \cdot \int_0^z z^{-5/7} \cdot (1 - z)^{-2/7} \cdot dt = z^{2/7} \cdot {}_2F_1\left(\left[\frac{2}{7}, \frac{2}{7}\right], \left[\frac{9}{7}\right]; z\right). \tag{2.144}$$

Using the parametrization of the *genus zero curve*

$$y = -\frac{(u+1)^2 \cdot u^5}{(u+1)^7 - u^7}, \qquad t = -\frac{u^7}{(u+1)^7 - u^7}, \tag{2.145}$$

one can actually perform the integration (2.144) of dt/y and get an alternative form of the hypergeometric function (2.144):

$$\int_0^z z^{-5/7} \cdot (1 - z)^{-2/7} \cdot dt = \int_0^u \rho(u) \cdot du = \int_0^v \frac{v}{1 - v^7} \cdot dv,$$
$$\text{where:}\ z = -\frac{u^7}{(u+1)^7 - u^7}, \ \rho(u) = \frac{(u+1)^4 \cdot u}{(u+1)^7 - u^7}, \tag{2.146}$$
$$\text{and:}\ v = \frac{u}{1 + u}, \ z = \frac{v^7}{v^7 - 1}.$$

Except transformations like $v \rightarrow \omega \cdot v$ (with $\omega^7 = 1$) which have no impact on z, it seems difficult to find rational symmetries in this genus zero case.

For $N = 11$ (genus five) the first successive coefficients read:

$$a_2 = -\frac{2}{17} \cdot a_1 \cdot (a_1 - 1),$$
$$a_3 = -\frac{1}{8092} \cdot a_1 \cdot (a_1 - 1) \cdot (143a_1 + 367),$$
$$a_4 = -\frac{1}{206346} \cdot a_1 \cdot (a_1 - 1) \cdot \left(1186a_1^2 + 2473a_1 + 5011\right), \ldots, \tag{2.147}$$
$$a_n = -\frac{1}{N(n)} \cdot a_1 \cdot (a_1 - 1) \cdot P_n(a_1).$$

The "infinitesimal composition" function $F(z)$ (see (2.56), (2.57), and (2.58)) reads,

$$F(z) = \frac{\partial R_{a_1}}{\partial a_1}\bigg|_{a_1=1} = z - \frac{2}{17}z^2 - \frac{15}{238}z^3 - \frac{5}{119}z^4 - \frac{37}{1190}z^5 - \frac{888}{36295}z^6 - \frac{2183}{108885}z^7 - \frac{4366}{258213}z^8 - \frac{58941}{4045337}z^9 - \frac{1807524}{141586795}z^{10} - \frac{46543743}{4106017055}z^{11} - \frac{5305986702}{521464165985}z^{12} + \cdots \tag{2.148}$$

and again we can actually check that this is actually the series expansion of the hypergeometric function

$$z \cdot (1-z) \cdot {}_2F_1\left(\left[1, \frac{15}{11}\right], \left[\frac{17}{11}\right]; z\right), \tag{2.149}$$

solution of Ω^* the adjoint of the Ω linear differential operator corresponding to this (genus 5) $N = 11$ case:

$$\Omega^* = D_z \cdot \left(D_z + \frac{1}{11} \cdot \frac{7z-5}{z \cdot (1-z)}\right). \tag{2.150}$$

We have similar results for (2.128), (2.129), (2.130). As far as these one-parameter families of transformations R_{a_1} are concerned, the only difference between the generic cases corresponding to *arbitrary genus* and *genus one* cases like (2.118) is that in the generic higher genus case, only a *finite number* of values of the parameter a_1 can correspond to rational functions. Note that this higher genus result generalizes to the arbitrary genus Gauss hypergeometric functions (2.128) and associated operators (2.129) and function (2.130). In this general case one can also get order by order a one-parameter family of transformations R_{a_1} satisfying a commutation relation (2.141).

Note that $R(z) = 1/z$ is *actually a solution of* (2.19) *for this genus-two example* (2.139). Along this line of selected $R(z)$ solutions of (2.19) many interesting subcases of this general case (2.128), (2.129), (2.130) are given in Appendix C.1.

In our previous genus-one examples, with this close identification between the *renormalization group and the isogenies of elliptic curves*, we saw that, in order to obtain linear differential operators covariant by an infinite number of transformations (rational or algebraic), we must restrict our second-order Gauss hypergeometric differential operator to Gauss hypergeometric *associated to elliptic curves* (see Appendices C and D). Beyond this framework we still have one-parameter families (see (2.141)) but we cannot expect an infinite number of rational (and probably algebraic) transformations to be particular cases of such families of transcendental transformations.

3. Conclusion

We have shown that several selected Gauss hypergeometric linear differential operators associated to elliptic curves, and factorised into order-one linear differential operators, actually present an *infinite number of rational symmetries* that actually identify with the *isogenies of the associated elliptic curves* that are perfect illustrations of *exact representations of the renormalization group*. We actually displayed all these calculations, results, and structures because they are perfect examples of exact renormalization transformations. For more realistic models (corresponding to Yang-Baxter models with elliptic parametrizations), the previous calculations and structures become more involved and subtle, the previous rational transformations being replaced by algebraic transformations corresponding to *modular curves*. For instance, in our models of lattice statistical mechanics (or enumerative combinatorics, etc.), we are often getting globally nilpotent linear differential operators [11] of quite high orders [20–24] that, in fact, factor into globally nilpotent operators of smaller orders[31] which, for Yang-Baxter integrable models with a *canonical elliptic parametrization*, must necessarily " *be associated with elliptic curves.*" Appendix D provides some calculations showing that the integral for $\chi^{(2)}$, the two-particle contribution of the susceptibility of the Ising model [25–27], is clearly and straightforwardly associated with an elliptic curve.

We wanted to highlight the importance of *explicit constructions* in answering difficult or subtle questions.

All the calculations displayed in this paper are elementary calculations given explicitly for heuristic reasons. The simple calculations (in particular with the introduction of a simple Rota-Baxter like functional equation) should be seen as some undergraduate training to more realistic renormalization calculations that will require a serious knowledge of fundamental modular curves, *modular forms*, Hauptmoduls, Gauss-Manin or Picard-Fuchs structures [28, 29] and, beyond, some knowledge of mirror symmetries [30–34] of Calabi-Yau manifolds, these *mirror symmetries* generalizing[32] the Hauptmodul structure for elliptic curves.

Appendices

A. Comment on the Rota-Baxter-Like Functional Equation (2.19)

We saw, several times, that the Rota-Baxter-like functional equation (2.19) is such that for a given $A(z)$ one gets a one-parameter family of analytical functions $R(z)$ obtained order by order by series expansion (see (2.32), (2.139)). Conversely for a given $R(z)$, for instance, $R(z) = -4z/(1-z)^2$, let us see if $R(z)$ can come from a unique $A(z)$. Assume that there are two $A(z)$ satisfying (2.26) with the same $R(z) = -4z/(1-z)^2$. We will denote $\delta(z)$ the difference of these two $A(z)$, and we will also introduce $\Delta(z) = z \cdot \delta(z)$. It is a straightforward calculation to see that $\Delta(z)$ verifies

$$\Delta(z) = \frac{1+z}{1-z} \cdot \Delta\left(\frac{-4z}{(1-z)^2}\right), \tag{A.1}$$

which has, beyond $\Delta(z) = 0$, at least one solution analytical at $z = 0$ that we can get order by order:

$$\Delta(z) = 1 + \frac{2}{5}z + \frac{22}{75}z^2 + \frac{394}{1625}z^3 + \frac{262634}{1243125}z^4 + \cdots . \tag{A.2}$$

It is straightforward to show from (A.1), from similar arguments we introduced for (2.57) on the functional equations (2.58) that $\Delta(z)$ is a transcendental function.

B. Miscellaneous Nonlinear ODEs on $P(z)$

From (2.70) one can get

$$F'(P(z)) = 1 + z \cdot \frac{P^{(2)}}{P^{(1)}},$$
$$F''(P(z)) = \frac{P^{(2)}}{\left(P^{(1)}\right)^2} + z \cdot \frac{P^{(3)}}{\left(P^{(1)}\right)^2} - z \cdot \frac{\left(P^{(2)}\right)^2}{\left(P^{(1)}\right)^3}, \tag{B.1}$$

and from (2.64), the linear second-order ODE on $F(z)$, one deduces the third-order nonlinear ODE[33] on the (at first sight *nonholonomic*) function $P(z)$:

$$\begin{aligned} & z \cdot \left(5P^2 - 6P + 3\right) \cdot \left(P^{(1)}\right)^4 - P \cdot (5P - 3) \cdot (P - 1) \cdot \left(P^{(1)}\right)^3 \\ & \quad - z \cdot (P - 1) \cdot P \cdot (5P - 3) \cdot P^{(2)} \cdot \left(P^{(1)}\right)^2 \\ & \quad + 4P^2 \cdot (P - 1)^2 \cdot \left(P^{(2)} + z \cdot P^{(3)}\right) \cdot P^{(1)} \\ & \quad - 4z \cdot \left(P^{(2)}\right)^2 \cdot P^2 \cdot (P - 1)^2 = 0, \end{aligned} \tag{B.2}$$

where the $P^{(n)}$'s denote the nth derivative of $P(z)$. This third order nonlinear ODE has a rescaling symmetry $z \longrightarrow \rho \cdot z$, for any ρ, and, also, an interesting symmetry, namely an invariance by $z \longrightarrow z^{\alpha}$, for any[34] value of α.

In a second step, using differential algebra tools, and, more specifically, the fact that $P(Q(z)) = Q(P(z)) = z$ together with the linear ODE for $Q(z)$, one finds the simpler second-order nonlinear ODE for $P(z)$:

$$P^{(2)} - \frac{1}{4} \cdot \frac{5P - 3}{(P - 1) \cdot P} \cdot \left(P^{(1)}\right)^2 + \frac{3}{4} \cdot \frac{1}{z} \cdot P^{(1)} = 0 \tag{B.3}$$

or

$$P^{(2)} - \left(\frac{3}{4} \cdot \frac{1}{P} + \frac{1}{2} \cdot \frac{1}{P - 1}\right) \cdot \left(P^{(1)}\right)^2 + \frac{3}{4} \cdot \frac{1}{z} \cdot P^{(1)} = 0. \tag{B.4}$$

Note that, more generally, the second-order nonlinear ODE

$$P^{(2)} - \left(\frac{3}{4} \cdot \frac{1}{P} + \frac{1}{2} \cdot \frac{1}{P - 1}\right) \cdot \left(P^{(1)}\right)^2 + \frac{\eta}{z} \cdot P^{(1)} = 0, \tag{B.5}$$

yields (B.2) for any value of the constant η. The change of variable z $\longrightarrow$ z^{α}, changes the parameter η into $1 + \alpha \cdot (\eta - 1)$. In particular the involution $z \leftrightarrow 1/z$ changes $\eta = 3/4$ into $\eta = 5/4$.

This nonlinear ODE, looking like *Painlevé V*, is actually *invariant* by the change of variable $P \rightarrow -4P/(1-P)^2$. It is, also, invariant by any rescaling $z \rightarrow \lambda z$, like the particular degenerate[35] subcase of *Painlevé V*

$$y'' - \left(\frac{1}{2y} + \frac{1}{y-1}\right) \cdot y'^2 + \frac{1}{z} \cdot y' = 0. \tag{B.6}$$

With (2.70) we recover the "Gauss-Manin" idea of Painlevé functions being seen as deformations of elliptic functions:

$$z \cdot \frac{dP(z)}{dz} = P(z) \cdot (1-P(z))^{1/2} \cdot {}_2F_1\left(\left[\frac{1}{4}, \frac{1}{2}\right], \left[\frac{5}{4}\right]; P(z)\right). \tag{B.7}$$

or

$$-2z \cdot \frac{d \operatorname{arctanh}\left((1-P(z))^{1/2}\right)}{dz} = {}_2F_1\left(\left[\frac{1}{4}, \frac{1}{2}\right], \left[\frac{5}{4}\right]; P(z)\right). \tag{B.8}$$

In fact, recalling $Q(P(z)) = z$, one also has the relation

$$P(z) \cdot {}_2F_1\left(\left[\frac{1}{4}, \frac{1}{2}\right], \left[\frac{5}{4}\right]; P(z)\right)^4 = z, \tag{B.9}$$

yielding with (B.7) the simple *nonlinear order-one* differential equation

$$z^3 \cdot (P')^4 - (1-P)^2 \cdot P^3 = 0, \tag{B.10}$$

already seen with (2.83), and that we can write in a separate way:

$$\frac{dP}{(1-P)^{1/2} \cdot P^{3/4}} = \frac{dz}{z^{3/4}}. \tag{B.11}$$

Note that $P(z^{4\cdot(1-\eta)})$ is actually solution of (B.5).

Equation (B.10) has (B.9) as a solution but in general the Puiseux series solutions $P_\mu(z)$ of the functional equation (μ is a constant):

$$\begin{gathered}
P_\mu(z)^{1/4} \cdot {}_2F_1\left(\left[\frac{1}{4}, \frac{1}{2}\right], \left[\frac{5}{4}\right]; P_A(z)\right) = \mu + z^{1/4} \quad \text{or} \\
P_\mu(z) = P\left(\left(\mu + z^{1/4}\right)^4\right), \\
P_\mu(z) = P\left(\mu^4\right) + 4 \cdot \mu^3 \cdot P'\left(\mu^4\right) \cdot z^{1/4} + \cdots .
\end{gathered} \tag{B.12}$$

It is a straightforward exercise of differential algebra to see that the order-one nonlinear differential equation (B.10) implies (B.3). In particular not only (B.9) is solution of (B.3) but also all the Puiseux series solutions (B.12) of (B.10). More generally the solutions of the functional equation:

$$P_{\mu,\lambda}(z)^{1/4} \cdot {}_2F_1\left(\left[\frac{1}{4}, \frac{1}{2}\right], \left[\frac{5}{4}\right]; P_{\mu,\lambda}(z)\right) = \mu + \lambda \cdot z^{1/4} \tag{B.13}$$

verify (B.3). This corresponds to the fact that

$$z^3 \cdot (P')^4 - \lambda^4 \cdot (1-P)^2 \cdot P^3 = 0, \tag{B.14}$$

yields (B.2) which is scaling symmetric ($z \to \rho \cdot z$) when (B.10) is not. More generally

$$z^{4\eta} \cdot (P')^4 - \lambda^4 \cdot (1-P)^2 \cdot P^3 = 0 \tag{B.15}$$

yields (B.2) for any value of the parameters η and λ. Finally, one also has that the solution of the functional equation

$$P_\eta(z)^{1/4} \cdot {}_2F_1\left(\left[\frac{1}{4}, \frac{1}{2}\right], \left[\frac{5}{4}\right]; P_\eta(z)\right) = \mu + \lambda \cdot z^{1-\eta} \tag{B.16}$$

is solution of (B.2), but also of (B.5) and even of (B.15).

Equation (B.5) with $\eta = 1/2$ (instead of $\eta = 3/2$ in (B.3)) has a solution, analytical at $z = 0$:

$$1 + x + \frac{1}{2}x^2 + \frac{7}{40}x^3 + \frac{1}{20}x^4 + \frac{121}{9600}x^5 + \frac{7}{2400}x^6 + \frac{211}{332800}x^7 + \frac{41}{312000}x^8 + \cdots. \tag{B.17}$$

This series has a singularity at $-1/4 \cdot z_s^2$, where z_s is given by (2.85). The radius of convergence of (B.17) corresponds to this singularity, namely, $R = 1/4 \cdot z_s^2$. This singularity result can be understood from the fact that, at $\eta = 1/2$, $P(z^2)$ is actually solution of (B.5).

In fact, we have the following solutions of (B.5) for various selected values of η. For $\eta = 0$, $P(z^4)$ is solution of (B.5). For $\eta = 2/3$, $P(z^{4/3})$ is solution of (B.5), and, more generally, $P(z^{4\cdot(1-\eta)})$ is solution of (B.5).

C. Gauss Hypergeometric ODEs Related to Elliptic Curves

It is not necessary to recall the close connection between Gauss hypergeometric functions and elliptic curves, or even modular curves [35, 36] and Hauptmoduls. This is very clear on the Goursat-type relation

$$\begin{aligned} &{}_2F_1\left(\left[2a, \frac{2a+1}{3}\right], \left[\frac{4a+2}{3}\right]; x\right) \\ &\quad = \left(1 - x + x^2\right)^{-a} \cdot {}_2F_1\left(\left[\frac{a}{3}, \frac{a+1}{3}\right], \left[\frac{4a+5}{6}\right]; \frac{27}{4} \cdot \frac{(x-1)^2 \cdot x^2}{(1-x+x^2)^3}\right), \end{aligned} \tag{C.1}$$

which generalizes the simpler quadratic Gauss relation:

$$ {}_2F_1\left([a,b],\left[\frac{a+b+1}{2}\right];x\right) = {}_2F_1\left(\left[\frac{a}{2},\frac{b}{2}\right],\left[\frac{a+b+1}{2}\right];4x(1-x)\right). \tag{C.2} $$

On (C.1) one recognizes (the inverse of) the *Klein modular invariant* [36] for the pull-back of the hypergeometric function on the rhs.

Many values of $[[a,b],[c]]$ are known to correspond to elliptic curves like $[[1/2,1/2],[1]]$ (complete elliptic integrals of the first and second kind) or modular forms: $[[1/12,5/12],[1]]$, $[[2/3,2/3],[1]]$, $[[2/3,2/3],[3/2]]$, and they can even be simply related:

$$ \left(\frac{z+27}{27}\right)^{1/3}\cdot {}_2F_1\left(\left[\frac{2}{3},\frac{2}{3}\right],[1];-\frac{1}{27}z\right) = \mu(z)\cdot {}_2F_1\left(\left[\frac{1}{12},\frac{5}{12}\right],[1];1728\frac{z}{(z+27)(z+3)^3}\right), \tag{C.3} $$

where:

$$ \mu(z) = \left(\frac{(z+27)(z+3)^3}{729}\right)^{-1/12}. \tag{C.4} $$

Once we have a hypergeometric function corresponding to an elliptic curve for some values of (a,b,c), one can find other values of (a,b,c) also corresponding to elliptic curves

$$ {}_2F_1([a,b],[c];x) \longrightarrow x^{1-c}\cdot {}_2F_1([1+a-c,1+b-c],[2-c];x). \tag{C.5} $$

In order to provide simple examples of linear differential ODEs we will restrict ourselves (just for heuristic reasons) to Gauss hypergeometric second-order differential equations.

Let us recall the *Euler integral representation* of the Gauss hypergeometric functions:

$$ \begin{aligned} {}_2F_1([a,b],[c];z) &= \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\cdot\int_0^1 \frac{dw}{w}w^b\cdot(1-w)^{c-1-b}\cdot(1-zw)^{-a} \\ &= \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}\cdot\int_0^1 \frac{dw}{w}w^a\cdot(1-w)^{c-1-a}\cdot(1-zw)^{-b} \\ &= \frac{\Gamma(c)}{\Gamma(a)\Gamma(c-a)}\cdot z^{-a}\int_0^z \frac{du}{u}u^a\cdot\left(1-\frac{u}{z}\right)^{c-1-a}\cdot(1-u)^{-b}. \end{aligned} \tag{C.6} $$

On the last line of (C.6), the selected role of $c = 1+a$ is quite clear.

Recall that the corresponding second-order differential operator is *invariant under the permutation of a and b* which is not obvious[37] on the Euler integral representations of the hypergeometric functions (this amounts to permuting 0 and ∞). The permutation of a and b is always floating around in this paper.

When the three parameters a, b and c of the Gauss hypergeometric functions are rational numbers we have integrals of *algebraic functions* and, therefore, we know [11, 37–40] that the corresponding second-order differential operator is necessarily *globally nilpotent* [11, 37–40]. Let us restrict to a, b, and c being *rational numbers* $a = N_a/D$, $b = N_b/D$ and $c = N_c/D$, where D is the common denominator of these three rational numbers. The Gauss hypergeometric functions are naturally associated to the *pencil of algebraic curves*

$$y^D = (1-u)^{N_b} \cdot u^{D-N_a} \cdot \left(1 - \frac{u}{z}\right)^{-N_c+D+N_a}. \tag{C.7}$$

Recalling the main example of the paper, one associates with ${}_2F_1([1/4, 1/2], [5/4]; z)$

$$\begin{aligned} {}_2F_1\left(\left[\frac{1}{2}, \frac{1}{4}\right], \left[\frac{5}{4}\right]; z\right) &= \frac{\Gamma(5/4)}{\Gamma(1/2)\Gamma(3/4)} \cdot \int_0^1 \frac{dw}{w} \cdot w^{1/2} \cdot (1-w)^{-1/4} \cdot (1-zw)^{-1/4} \\ &= \frac{\Gamma(5/4)}{\Gamma(1/2)\Gamma(3/4)} \cdot z^{-1/2} \cdot \int_0^z u^{-1/2} \cdot \left(1 - \frac{u}{z}\right)^{-1/4} \cdot (1-u)^{-1/4} \cdot du \end{aligned} \tag{C.8}$$

the z-pencil of elliptic curves[38]

$$y^4 - u^2 \cdot (1-u) \cdot \left(1 - \frac{u}{z}\right) = 0, \tag{C.9}$$

where we associated (see (2.2)) to ${}_2F_1([1/2, 1/4], [5/4]; z)$ the elliptic curve

$$y^4 - u^3 \cdot (1-u)^2 = 0. \tag{C.10}$$

C.1. Miscellaneous Examples

In the more general (2.128), (2.129), (2.130), (resp. (2.132), (2.133)) framework, one can find many interesting subcases.

(i) The previous $R(z) = 1/z$ involution is solution of the functional relation (2.19) when $a = 2b$ if $c = 1 + b$, or $b = 2a$ if $c = 1 + a$.

(ii) The involution $R(z) = 1-z$ is solution of the functional relation (2.19) when $a+b = 1$ if $c = 1 + b$, or $c = 1 + a$.

(iii) The infinite-order transformation:

$$R(z) = t \cdot \frac{z}{1 + (t-1) \cdot z}, \qquad R^{(n)}(z) = t^n \cdot \frac{z}{1 + (t^n - 1) \cdot z}, \tag{C.11}$$

is solution of the functional relation (2.19) when $a = 1 + b$ if $c = 1 + b$, or $b = 1 + ac = 1 + a$.

(iv) The scaling transformation $R(z) = t \cdot z$ is solution of the functional relation (2.19) when $a = 0$ and $c = 1 + b$ (resp., $b = 0$ and $c = 1 + a$).

(v) We also have a quite degenerate situation for $b = 1$ or $a = 1$ when $c = 2$ with the infinite-order transformation

$$R(z) = 1 - t \cdot (1-z), \qquad R^{(n)}(z) = 1 - t^n \cdot (1-z), \tag{C.12}$$

solution of (2.19).

(vi) The two order-three transformations

$$R(z) = \frac{z-1}{z}, \qquad R(R(z)) = \frac{1}{1-z}, \tag{C.13}$$

are solutions of the functional relation (2.19) for $a = 2/3, b = 1/3, c = 4/3$, or $a = 1/3, b = 1/3, c = 4/3$.

D. Ising Model Susceptibility: $\tilde{\chi}^{(2)}$ and Elliptic Curves

The two-particle contribution of the susceptibility of the Ising model [25–27] is given by a double integral. This double integral on two angles $\tilde{\chi}^{(2)}$ reduces to a simple integral[39] (because the two angles are opposite):

$$\tilde{\chi}^{(2)} = \int_0^{\pi} d\theta \cdot y^2 \cdot \frac{1+x^2}{1-x^2} \cdot \left(\frac{x \cdot \sin(\theta)}{1-x^2}\right)^2, \tag{D.1}$$

where

$$x = A - B, \qquad A = \frac{1}{2w} - \cos(\theta), \qquad B^2 = A^2 - 1, \qquad y^2 = \frac{1}{A^2 - 1}. \tag{D.2}$$

Denoting $C = \cos(\theta)$ we can rewrite the integral $\chi^{(2)}$ as

$$\tilde{\chi}^{(2)} = \int_0^1 \frac{dC}{(1-C^2)^{1/2}} \cdot x^2 \cdot y^2 \cdot \frac{1+x^2}{(1-x^2)^3}, \tag{D.3}$$

that we want to see as:

$$\int_0^1 \frac{dC}{z} = \int_0^w \frac{dq}{Z}. \tag{D.4}$$

The variable z reads:

$$\frac{1}{z} - \frac{1}{(1-C^2)^{1/2}} \cdot x^2 \cdot y^2 \cdot \frac{1+x^2}{(1-x^2)^3} = 0, \tag{D.5}$$

which after simplifications gives

$$A^2\left(C^2 - 1\right) \cdot z^2 + \left(A^2 - 1\right)^5 = 0, \tag{D.6}$$

that is

$$\left(\frac{1}{2w} - C\right)^2 \cdot \left(C^2 - 1\right) \cdot z^2 + \left(\left(\frac{1}{2w} - C\right)^2 - 1\right)^5 = 0. \tag{D.7}$$

In terms of the variable $q = w \cdot C$ one can rewrite (see (D.4)) the integral (D.3) as an incomplete integral:

$$256 \cdot (1-2q)^2\left(q^2 - w^2\right) \cdot Z^2 w^4 + (2q - 1 + 2w)^5 (2q - 1 - 2w)^5 = 0. \tag{D.8}$$

This w-pencil of algebraic curves is actually a *w-pencil of genus one curves*, seen as algebraic curves in Z and q.

Acknowledgments

J. M. Maillard thanks D. Bertrand, A. Enge, M. Hindry, D. Loeffler, J. Nekovár, J. Oesterlé, and J. Watkins for fruitful discussions on modular curves, modular forms, and modular functions. He also thanks D. Mouhanna for large discussions on the renormalization group and A. Ramani and N. Witte for detailed discussions on the Painlevé equations.

Endnotes

1. The renormalization group approach of important problems like first-order phase transitions, commensurate-incommensurate phase transitions, or off-critical problems is more problematic.
2. In contrast with functional renormalization group [41–43].
3. One simply verifies that these transformations reduce to the previous $T_N : t \rightarrow t^N$ in the $z = 1$ limit (no magnetic field).
4. For instance the fixed points of (1.3) are not isolated fixed points but lie on (an infinite number) of genus zero curves.
5. In well-suited Boltzmann weight variables like x and z in (1.3), and not in (bad) variables like K, the coupling constants or the temperature.
6. Such representations of the renormalization group are not *exact* representations (the exact transformation acts in an infinite number of parameters) but some authors tried to define "improved" renormalization transformations imposing the compatibility (commutation) of the renormalization transformations with some known exact symmetries of the model (Kramers-Wannier duality, gauge symmetries...).
7. For which the partition function or other physical quantities are algebraic functions.
8. See for instance (2.18) in [44].
9. Suggesting an understanding [4, 45] of the quite rich structure of infinite number of the singularities of the $\chi^{(n)}$'s in the complex plane from a Hauptmodul approach [4, 45]. Furthermore the notion of *Heegner numbers* is closely linked to the isogenies mentioned here [4]. An exact value of the j-function $j(\tau)$ corresponding one of the first Heegner number is, e.g., $j(1 + i) = 12^3$.
10. It should be recalled that the mirror symmetry found with Calabi-Yau manifolds [30–34] can be seen as higher-order generalizations of Hauptmoduls. We thus have already generalizations of this identification of the renormalization and modular structure when one is not restricted to elliptic curves anymore.
11. The fact that the renormalization group must be reversible has apparently been totally forgotten by most of the authors who just see a semigroup corresponding to forward iterations converging to the critical points (resp. manifolds).
12. Corresponding to Atkin-Lehner polynomials and Weber's functions.

13. Conversely, and more precisely, writing $1728z^2/(z+256)^3 = 1728z'/(z'+16)^3$ gives the Atkin-Lehner [46] involution $z \cdot z' = 2^{12}$, together with the quadratic relation $z^2 - zz'^2 - 48zz' - 4096z' = 0$
14. They are not only Fuchsian, the corresponding linear differential operators are globally nilpotent or$G$-operators [11].
15. Where j is typically the j-function [44, 47].
16. Such formula is actually valid for $\Omega_A = (D_z + A(z)) \cdot D_z$ for any $A(z)$. Denoting $\mathcal{S}_N$ symmetric Nth power of Ω_A one has $\mathcal{S}_N = (D_z + A(z)) \cdot \mathcal{S}_{N-1}$.
17. The Rota-Baxter relation of weight Θ reads: $R(x)R(y) + \Theta R(xy) \quad R(R(x)y + xR(y))$.
18. For $A(z)$ given we get a one-parameter family of $R(z)$ solution of (2.19). Conversely, for $R(z)$ given one can ask if there are several $A(z)$ such that (2.19) is verified. This is sketched in Appendix A.
19. Using the command "dchange" with PDEtools in Maple.
20. Note that the result for ω_1^* is nothing but transformation (2.14) on ω_k for $k = -1$. Also note that the two transformations, performing the change of variable $z \longrightarrow -4z/(1-z)^2$ and taking the adjoint, *do not commute*: $(\omega_1^*)^{(R)} \neq ((\omega_1)^{(R)})^*$.
21. Denoted $JacobiSN$ in Maple: $P(z) = (JacobiSN(z^{1/4}, I))^4$.
22. As a (nonholonomic) elliptic function $P(z)$ provides elementary examples [48] of nonlinear ODEs with the Painlevé property (like the Weierstrass P-function).
23. It is the absolute value of the inverse of the image of the n-th iterate of $S_{-1/4}^{(1)}$ of -1.
24. If this previous singularity argument was valid we would have had singularities as close as possible to $z = 0$ (namely, $z_s/(-4)^n$), yielding a zero radius of convergence. Similarly combining $T^*(z)$ and the inverse of $T(z)$ we would have obtained an infinite number of singularities on the circle of radius $|z_s|$.
25. Note a (small) misprint in formula (64) page 174 of Vidunas [12].
26. Of the $c = 1 + b$ type (see below).
27. The change of variable (2.125) can be parametrized with hyperbolic tangents: $z \longrightarrow z'$ with $z = \tanh(u)^2$,$z' = \tanh(2u)^2$. Note that $z \longrightarrow 4 \cdot z/(1-z)^2$ is parametrized by $z = \tan(u)^2 z' = \tan(2u)^2$ but $z \longrightarrow -4 \cdot z/(1-z)^2$ is not parametrized by trigonometric functions.
28. Thus avoiding the full complexity (and subtleties) of the covariance of ODEs by algebraic transformations like modular transformations (1.8).
29. See for instance (C.6) in Appendix C.
30. More generally in our models of lattice statistical mechanics (or enumerative combinatorics etc.) we are seeking for (high order) globally nilpotent [11] operators that, in fact, factor into globally nilpotent operators of smaller order, which, for Yang-Baxter integrable models with a canonical elliptic parametrization, must necessarily "be associated with elliptic curves". Appendix D provides some calculations showing that the integral for $\chi^{(2)}$, the two-particle contribution of the susceptibility of the Ising model [25–27] is clearly, and straightforwardly, associated with an elliptic curve.
31. Experimentally [21] and as could be expected from Dwork's conjecture [11], one often finds for these small order factors hypergeometric second-order operators and sometimes selected Heun functions [49] (or their symmetric products).

32. For instance equation (1.9) of [31]. Do note that the periods of certain K3 families (and hence the original Calabi-Yau family) can be described by the squares of the periods of the elliptic curves [31]. The mirror maps of some K3 surface families are always reciprocals of some McKay-Thompson series associated to the Monstruous Moonshine list of Conway and Norton, with the mirror maps of these examples being always automorphic functions for *genus zero* [32, 33].
33. Using differential algebra tools one can verify that (2.84) implies (B.2).
34. Beyond diffeomorphisms of the circle: the parameter α can be a complex number.
35. Having the movable-poles solutions: $(\alpha^\beta + z^\beta)^2/(\alpha^\beta - z^\beta)^2$.
36. Taking for$x$ the elliptic lambda function.
37. For instance for $2F_1([1/4, 1/2], [5/4]; z)$ it changes an Euler integral with $\Gamma(5/4)/\Gamma(1/4)\Gamma(1) = 1/4$ into an Euler integral with $\Gamma(5/4)/\Gamma(3/4)\Gamma(1/2) = (1/4) \cdot ((2\pi)^{1/2}/\Gamma(3/4)^2)$.
38. The algebraic curves (C.9) are genus one curves for any value of z, except $z = 1$, where the curve becomes the union of two rational curves $(u^2 - u + y^2)(u^2 - u - y^2) = 0$.
39. The prefactors in front of the integrals are not relevant for our discussion here.

References

[1] http://www.physics.fsu.edu/courses/Spring05/phy6938-02/decimation.pdf.

[2] M. E. Fisher, “Renormalization group theory: its basis and formulation in statistical physics,” *Reviews of Modern Physics*, vol. 70, no. 2, pp. 653–681, 1998.

[3] J.-M. Maillard and S. Boukraa, “Modular invariance in lattice statistical mechanics,” *Fondation Louis de Broglie. Annales*, vol. 26, no. 2, pp. 287–328, 2001.

[4] S. Boukraa, S. Hassani, J.-M. Maillard, and N. Zenine, “Singularities of n-fold integrals of the Ising class and the theory of elliptic curves,” *Journal of Physics A*, vol. 40, no. 39, pp. 11713–11748, 2007.

[5] J.-M. Maillard, “Automorphisms of algebraic varieties and Yang-Baxter equations,” *Journal of Mathematical Physics*, vol. 27, no. 11, pp. 2776–2781, 1986.

[6] S. Boukraa and J.-M. Maillard, “Let’s Baxterise,” *Journal of Statistical Physics*, vol. 102, no. 3-4, pp. 641–700, 2001.

[7] M. P. Bellon, J.-M. Maillard, and C.-M. Viallet, “Quasi integrability of the sixteen-vertex model,” *Physics Letters. B*, vol. 281, no. 3-4, pp. 315–319, 1992.

[8] N. Zenine, S. Boukraa, S. Hassani, and J.-M. Maillard, “The Fuchsian differential equation of the square lattice Ising model $X^{(3)}$ susceptibility,” *Journal of Physics. A*, vol. 37, no. 41, pp. 9651–9668, 2004.

[9] N. Zenine, S. Boukraa, S. Hassani, and J.-M. Maillard, “Square lattice Ising model susceptibility: series expansion method and differential equation for $X^{(3)}$,” *Journal of Physics. A*, vol. 38, no. 9, pp. 1875–1899, 2005.

[10] N. Zenine, S. Boukraa, S. Hassani, and J.-M. Maillard, “Square lattice Ising model susceptibility: connection matrices and singular behaviour of $X^{(3)}$ and $X^{(4)}$,” *Journal of Physics. A*, vol. 38, no. 43, pp. 9439–9474, 2005.

[11] A. Bostan, S. Boukraa, S. Hassani, J.-M. Maillard, J.-A. Weil, and N. Zenine, “Globally nilpotent differential operators and the square Ising model,” *Journal of Physics A*, vol. 42, no. 12, pp. 125206–125256, 2009.

[12] R. Vidunas, “Algebraic transformations of Gauss hypergeometric functions,” *Funkcialaj Ekvacioj*, vol. 52, no. 2, pp. 139–180, 2009.

[13] K. Ebrahimi-Fard and L. Guo, “Rota-Baxter algebras in renormalization of perturbative quantum field theory,” in *Physics Letters B*, vol. 281, pp. 315–319, 1992.

[14] K. Ebrahimi-Fard, D. Manchon, and F. Patras, “A noncommutative Bohnenblust-Spitzer identity for Rota-Baxter algebras solves Bogoliubov’s recursion,” *Journal of Noncommutative Geometry*, vol. 3, no. 2, pp. 181–222, 2009.

[15] C. L. Siegel, "Iteration of analytic functions," *Annals of Mathematics*, vol. 43, pp. 607–612, 1942.

[16] R. Pérez-Marco, "Solution complète au problème de Siegel de linéarisation d'une application holomorphe au voisinage d'un point fixe (d'après Yoccoz)," *Astérisque*, vol. 206, no. 3, pp. 273–310, 1992, Séminaire Bourbaki, 1991/1992, Exp. No. 753.

[17] S. Boukraa, J.-M. Maillard, and G. Rollet, "Almost integrable mappings," *International Journal of Modern Physics B*, vol. 8, no. 1-2, pp. 137–174, 1994.

[18] N. Abarenkova, J.-Ch. Anglès d'Auriac, S. Boukraa, and J.-M. Maillard, "Growth-complexity spectrum of some discrete dynamical systems," *Physica D*, vol. 130, no. 1-2, pp. 27–42, 1999.

[19] J.-Ch. Anglès d'Auriac, S. Boukraa, and J.-M. Maillard, "Functional relations in lattice statistical mechanics, enumerative combinatorics, and discrete dynamical systems," *Annals of Combinatorics*, vol. 3, no. 2–4, pp. 131–158, 1999.

[20] A. Bostan, S. Boukraa, A. J. Guttmann, et al., "High order Fuchsian equations for the square lattice Ising model: $\tilde{X}^{(5)}$," *Journal of Physics A*, vol. 42, no. 27, pp. 275209–275240, 2009.

[21] S. Boukraa, A. J. Guttmann, S. Hassani, et al., "Experimental mathematics on the magnetic susceptibility of the square lattice Ising model," *Journal of Physics A*, vol. 41, no. 45, pp. 455202–455252, 2009.

[22] N. Zenine, S. Boukraa, S. Hassani, and J.-M. Maillard, "Ising model susceptibility: the Fuchsian differential equation for $X^{(4)}$ and its factorization properties," *Journal of Physics A*, vol. 38, no. 19, pp. 4149–4173, 2005.

[23] S. Boukraa, S. Hassani, J.-M. Maillard, B. M. McCoy, W. P. Orrick, and N. Zenine, "Holonomy of the Ising model form factors," *Journal of Physics A*, vol. 40, no. 1, pp. 75–111, 2007.

[24] S. Boukraa, S. Hassani, J.-M. Maillard, B. M. McCoy, and N. Zenine, "The diagonal Ising susceptibility," *Journal of Physics A*, vol. 40, no. 29, pp. 8219–8236, 2007.

[25] T. T. Wu, B. M. McCoy, C. A. Tracy, and E. Barouch, "Spin-spin correlation functions for the two-dimensional Ising model: exact theory in the scaling region," *Physical Review B*, vol. 13, no. 1, pp. 316–374, 1976.

[26] B. Nickel, "On the singularity structure of the 2D Ising model susceptibility," *Journal of Physics. A*, vol. 32, no. 21, pp. 3889–3906, 1999.

[27] B. Nickel, "Addendum to: "on the singularity structure of the 2D Ising model susceptibility"," *Journal of Physics. A*, vol. 33, no. 8, pp. 1693–1711, 2000.

[28] S. Boukraa, S. Hassani, J.-M. Maillard, B. M. McCoy, J.-A. Weil, and N. Zenine, "Painlevé versus Fuchs," *Journal of Physics A*, vol. 39, no. 39, pp. 12245–12263, 2006.

[29] S. Boukraa, S. Hassani, J.-M. Maillard, B. M. McCoy, J.-A. Weil, and N. Zenine, "Fuchs versus Painlevé," *Journal of Physics A*, vol. 40, no. 11, pp. 12589–12605, 2007.

[30] P. Candelas, X. C. de la Ossa, P. S. Green, and L. Parkes, "A pair of Calabi-Yau manifolds as an exactly soluble superconformal theory," *Nuclear Physics. B*, vol. 359, no. 1, pp. 21–74, 1991.

[31] B. H. Lian and S.-T. Yau, "Mirror maps, modular relations and hypergeometric series. II," *Nuclear Physics B*, vol. 46, no. 1–3, pp. 248–262, 1996.

[32] C. F. Doran, "Picard-Fuchs uniformization and modularity of the mirror map," *Communications in Mathematical Physics*, vol. 212, no. 3, pp. 625–647, 2000.

[33] C. F. Doran, "Picard-Fuchs uniformization: modularity of the mirror map and mirror-moonshine," in *The Arithmetic and Geometry of Algebraic Cycles*, vol. 24 of *CRM Proc. Lecture Notes*, pp. 257–281, American Mathematical Society, Providence, RI, USA, 2000.

[34] C. Krattenthaler and T. Rivoal, "On the integrality of the Taylor coefficients of mirror maps," *Duke Mathematical Journal*, vol. 151, no. 2, pp. 175–218, 2010.

[35] P. F. Stiller, "Classical automorphic forms and hypergeometric functions," *Journal of Number Theory*, vol. 28, no. 2, pp. 219–232, 1988.

[36] W. Zudilin, "The hypergeometric equation and Ramanujan functions," *Ramanujan Journal*, vol. 7, no. 4, pp. 435–447, 2003.

[37] Y. André, *G-Functions and Geometry*, Aspects of Mathematics, E13, Friedrich Vieweg & Sohn, Braunschweig, Germany, 1989.

[38] Y. André and F. Baldassarri, "Geometric theory of G-functions," in *Arithmetic Geometry (Cortona, 1994)*, Symposia in Pure Mathematics, 37, pp. 1–22, Cambridge University Press, Cambridge, UK, 1997.

[39] Yves André, "Arithmetic Gevrey series and transcendence. A survey," *Journal de Théorie des Nombres de Bordeaux*, vol. 15, no. 1, pp. 1–10, 2003.

[40] Yves André, "Sur la conjecture des p-courbures de Grothendieck-Katz et un problème de Dwork," in *Geometric Aspects of Dwork Theory*, A. Adolphson, F. Baldassarri, P. Berthelot, N. Katz, and F. Loeser, Eds., vol. 1-2, pp. 55–112, Walter de Gruyter, Berlin, Germany, 1991.

[41] D. S. Fisher, "Random fields, random anisotropies, nonlinear models, and dimensional reduction," *Physical Review B*, vol. 31, no. 11, pp. 7233–7251, 1985.

[42] J. Berges, N. Tetradis, and C. Wetterich, "Non-perturbative renormalization flow in quantum field theory and statistical physics," *Physics Reports*, vol. 363, no. 4–6, pp. 223–386, 2002.

[43] B. Delamotte, D. Mouhanna, and M. Tissier, "Nonperturbative renormalization-group approach to frustrated magnets," *Physical Review B*, vol. 69, no. 13, 53 pages, 2004.

[44] B. C. Berndt and H. H. Chan, "Ramanujan and the modular j-invariant," *Canadian Mathematical Bulletin*, vol. 42, no. 4, pp. 427–440, 1999.

[45] S. Boukraa, S. Hassani, J.-M. Maillard, and N. Zenine, "Landau singularities and singularities of holonomic integrals of the Ising class," *Journal of Physics. A*, vol. 40, no. 11, pp. 2583–2614, 2007.

[46] H. H. Chan and M. L. Lang, "Ramanujan's modular equations and Atkin-Lehner involutions," *Israel Journal of Mathematics*, vol. 103, pp. 1–16, 1998.

[47] "j-Function," http://mathworld.wolfram.com/j-Function.html.

[48] E. L. Ince, *Ordinary Differential Equations*, Dover, New York, NY, USA, 1944.

[49] R. S. Maier, "On reducing the Heun equation to the hypergeometric equation," *Journal of Differential Equations*, vol. 213, no. 1, pp. 171–203, 2005.