q-DEFORMATIONS OF TWO-DIMENSIONAL YANG-MILLS THEORY: CLASSIFICATION, CATEGORIFICATION AND REFINEMENT

RICHARD J. SZABO AND MIGUEL TIERZ

Abstract. We characterise the quantum group gauge symmetries underlying q-deformations of two-dimensional Yang-Mills theory by studying their relationships with the matrix models that appear in Chern-Simons theory and six-dimensional $\mathcal{N}=2$ gauge theories, together with their refinements and supersymmetric extensions. We develop uniqueness results for quantum deformations and refinements of gauge theories in two dimensions, and describe several potential analytic and geometric realisations of them. We reconstruct standard q-deformed Yang-Mills amplitudes via gluing rules in the representation category of the quantum group associated to the gauge group, whose numerical invariants are the usual characters in the Grothendieck group of the category. We apply this formalism to compute refinements of q-deformed amplitudes in terms of generalised characters, and relate them to refined Chern-Simons matrix models and generalized unitary matrix integrals in the quantum β-ensemble which compute refined topological string amplitudes. We also describe applications of our results to gauge theories in five and seven dimensions, and to the dual superconformal field theories in four dimensions which descend from the $\mathcal{N}=(2,0)$ six-dimensional superconformal theory.

Contents

Date: May 2013 EMPG–13–03 .

1. Introduction and summary

Yang-Mills theory in two dimensions has been vigorously studied over the years because of the analytical tractability of the quantum gauge theory (see [1] for a review), originally pointed out by Migdal [2] and further developed in [3, 4]; an exact lattice gauge theory formalism leads to the heat kernel expansion for the partition function and correlators on any compact, connected and oriented Riemann surface Σ_h of genus h. In this paper we are concerned with the q-deformation of two-dimensional Yang-Mills theory, studied originally in [5, 6] and further developed in [7, 8, 9, 10, 11, 12, 13]. In [7] it was shown that the partition function of topologically twisted $\mathcal{N}=4$ Yang-Mills theory with gauge group $U(N)$ on the ruled Riemann surface $\mathcal{O}(-p)\to\Sigma_h$ reduces to that of q-deformed $U(N)$ Yang-Mills theory on the base Riemann surface Σ_h; this relationship was further clarified and extended in [11] to show how the q-deformed gauge theory captures the counting of instantons on Hirzebruch-Jung spaces. From this result it is anticipated that q-deformed Yang-Mills theory provides a non-perturbative completion of topological string theory on the rank two Calabi-Yau fibration $\mathcal{O}(p+2h-2)\oplus\mathcal{O}(-p)$ over Σ_h. These studies have been reinvigorated in the past few years with the discovery that for $p=1$ the two-dimensional gauge theory also computes the partition function of a strongly-coupled $\mathcal{N}=2$ gauge theory on $S^1\times S^3$ [14, 15, 16]; this duality is conjecturally realised within the putative six-dimensional $\mathcal{N}=(2,0)$ superconformal theory on $S^3\times S^1\times\Sigma_h$ in which the four-dimensional gauge theories are specified by the Riemann surface Σ_h [17]. A refinement of this two-dimensional gauge theory first appeared in its topological BF-theory form in [18] as the dual to a four-dimensional $\mathcal{N}=2$ gauge theory on $S^3\times S^1$ with two superconformal fugacities; in [19] the full (non-topological) refined gauge theory was derived from counting BPS states in refined topological string theory on the local Calabi-Yau threefold $\mathcal{O}(p+2h-2)\oplus\mathcal{O}(-p)\to\Sigma_h$ with a non-selfdual graviphoton background.

This paper is devoted to an in-depth investigation of quantum deformations and refinements of two-dimensional Yang-Mills theory, and their deep and rich connections with a multitude of gauge theories in higher dimensions. We lay emphasis on understanding the precise quantum group structure of the gauge symmetries underlying these models, and a complete characterization of quantum deformations and refinements of gauge theories in two dimensions. We clarify the relationships between the discrete matrix models that appear in q-deformed Yang-Mills theory and the continuous matrix models of Chern-Simons gauge theory using the moment problem, and extend these equivalences to refined settings. We also relate certain topological versions of two-dimensional Yang-Mills amplitudes to gauge theory partition functions in all dimensions ranging from three to seven by reformulating them in terms of (refined) unitary matrix integrals; in particular, we use this connection to reproduce (refined) topological string amplitudes from five-dimensional $\mathcal{N}=1$ gauge theory partition functions.

As we review in §2, the partition function of q-deformed Yang-Mills theory on Σ_h is a simple variation of the Migdal formula (given explicitly in (2.5) below), which involves quantum dimensions rather than ordinary dimensions of representations of the gauge group. This q-deformed gauge theory bears the same relation to q-deformed representation theory as ordinary Yang-Mills theory does to ordinary representation theory. It can be regarded as an analytic continuation of Chern-Simons gauge theory on a Seifert fibration of degree p over the Riemann surface Σ_h. For genus $h=0$, the Seifert manifold is the three-sphere S^3 for $p=1$ (regarded as the Hopf fibration $S^3 \to \Sigma_0 = S^2$) and the lens space $L(p,1) = S^3/\mathbb{Z}_p$ for $p>1$. In this particular q-deformation the quadratic Casimir eigenvalues are ordinary integers because they arise from a sum over torus bundles on Σ_h just as in the ordinary case, which do not undergo any deformation themselves; this is not very natural from the point of view of quantum group theory. Another somewhat non-canonical feature is the dependence of the deformation parameter q on the Yang-Mills coupling constant. Alternatively, the approaches of [5, 6] deform the gauge symmetry to a quantum group, and thereby avoid some of these pitfalls. The role of quantum group gauge symmetries in the complete solution of q-deformed two-dimensional Yang-Mills theory is also stressed in [20]. Insofar as the q-deformed gauge theory is an analytic continuation of Chern-Simons theory, which has a well understood relation to quantum groups, it is natural to explore the relation between the field theories based on an explicit quantum deformation of the gauge group and the field theories with undeformed gauge group where the quantum group symmetries seem to emerge implicitly.

As a first step towards understanding this relationship, in §3 we relate the discrete matrix models that underlie q-deformed Yang-Mills amplitudes to the continuous Chern-Simons matrix models; matrix model techniques and relations permeate this paper and are a driving force in much of our analysis. In particular, we establish a new relationship with the Stieltjes-Wigert matrix model for all $p \in \mathbb{Z}_{\geq 0}$ and a new dual formulation of $U(N)$ Chern-Simons theory as a $U(\infty)$ matrix model; this unitary matrix model is related explicitly to the BF-theory limit of two-dimensional Yang-Mills theory and also to the $U(\infty)$ matrix models describing $\mathcal{N}=2$ gauge theories in six dimensions that we consider in §4.

As a next step, we develop uniqueness results for quantum deformations of two-dimensional Yang-Mills theory in §4. We formulate the q-deformation of the heat kernel expansion by quantum dimensions together with a q-deformed Casimir invariant, and show that it is equivalent to ordinary Yang-Mills theory. We further show that q-deformation of the Boltzmann weight in Migdal's partition function is qualitatively equivalent to deformation by quantum dimensions. As a consequence, both Klimčík's partition function for q-deformed Yang-Mills theory [6] and the partition function for crystal melting with external potentials are equivalent to that of ordinary generalized two-dimensional Yang-Mills theory. We extend these correspondences to $\mathcal{N}=2$ gauge theories in six dimensions by rewriting their partition functions as unitary matrix integrals and applying the strong Szegő limit theorem for Toeplitz determinants to evaluate them in closed form. Our findings are somewhat in line with the arguments of Brzeziński and Majid [21] that gauge theories with quantum group gauge symmetries should be defined on *quantum* spaces in order to get something that is different from ordinary Yang-Mills theory. We elaborate on this observation and consider the diagonalisation technique of [7, 10] in the context of Yang-Mills theory on the standard Podleś quantum sphere S_q^2 [22]; this approach effectively abelianizes the non-abelian gauge theory so that its partition function (and correlation functions) can be calculated explicitly and straightforwardly by summing the resulting abelian field theory over all isomorphism classes of torus bundles. One advantage of this approach is that it computes quantum fluctuations straightforwardly from Gaussian path integrals of free abelian fields (and hence is essentially rigorous), and the non-abelian nature of the original field theory is reflected in the determinants which arise as Jacobians in the diagonalization procedure. When applied to torus bundles over S_q^2, this calculation gives a putative derivation of the heat kernel expansion involving q-deformed Casimir eigenvalues and moreover explains why quantum group gauge symmetries appear as ordinary gauge symmetries in this framework.

In §5 we reformulate the standard two-dimensional topological field theory construction of q-deformed Yang-Mills amplitudes [7] in terms of a functor whose target is the semisimple ribbon category of representations of the quantum universal enveloping algebra associated to the gauge group $U(N)$. This construction "categorifies" the usual building blocks of Yang-Mills amplitudes in terms of $U(N)$ characters, which now appear as numerical invariants in the corresponding Grothendieck group of the ribbon category. A related modular tensor category is used in [23] to reformulate the three-dimensional topological field theory construction of (refined) Chern-Simons theory, which is usually based on the finite category of integrable representations of $U(N)$. Our ribbon category can be regarded as a certain completion of this category (which we describe explicitly) involving direct sums of infinitely many simple objects, the irreducible representations of the quantum group. In this way we are able to build Yang-Mills amplitudes as numerical invariants of this category in a way in which the quantum group gauge symmetry is manifest simply by construction, and which moreover exhibits interesting relationships amongst different correlators. For completeness and convenience of exposition, we review all category theory concepts and results that we use in this paper.

One of the main advantages of our categorical reformulation is that it straightforwardly allows for various generalizations, e.g. to gauge groups other than $U(N)$. In §6 we use this general treatment to construct refinements of q-deformed Yang-Mills amplitudes; they are analytic continuations of the correlators in the refined Chern-Simons theory originally defined in [24]. Now the numerical invariants arising from integration of morphisms in the ribbon category are given by generalised characters. We extend the discrete/continuous matrix model equivalence to show that the refined q-deformed Yang-Mills partition function for all $p \in \mathbb{Z}_{\geq 0}$ is also equivalent to the refined Chern-Simons matrix model derived in [24] from a refined counting of BPS states of spinning M2-branes in certain M-theory compactifications, which is a quantum β-deformation of the usual Stieltjes-Wigert matrix ensemble. By applying the strong Szegő limit theorem to evaluate generalized unitary matrix models in the quantum β-ensemble, we relate various refinements of the two-dimensional gauge theories, together with their supersymmetric extensions, to supersymmetric gauge theories in five, six and seven dimensions; we also describe some uniqueness results for refinements of two-dimensional gauge theories. In particular, building on constructions from the unrefined case [25], we expect that our refined $U(\infty)$ matrix models could be relevant to Macdonald refined stochastic processes [26]. We conclude by applying some of the formalism of generalised characters to discuss possible physical derivations of refined q-deformed Yang-Mills amplitudes directly from the perspective of two-dimensional quantum field theory.

Three appendices at the end of the paper include some of the more technical details which are used in the main text: Appendix A summarises pertinent aspects of the quantum universal enveloping algebra of the Lie algebra of the gauge group $U(N)$ that we use, Appendix B describes techniques for evaluating Toeplitz determinants that are used in our analysis of unitary matrix models throughout the paper, and Appendix C contains more intricate details concerning our category theory constructions.

ACKNOWLEDGMENTS

We thank Mark Adler, Jacob Christiansen, David Evans, Jim Howie, Kurt Johansson, Anatoly Konechny, Giovanni Landi, Mark Lawson, Sanjaye Ramgoolam, Ingo Runkel, Christoph Schweigert, Simon Willerton and Masahito Yamazaki for helpful discussions and correspondence. MT thanks Jørgen Andersen for the warm hospitality during his visit to the Centre for Quantum Geometry of Moduli Spaces, Aarhus University, Denmark, where part of this work was carried out. The work of RJS was partially supported by the Consolidated Grant ST/J000310/1 from the UK Science and Technology Facilities Council, and by Grant RPG-404 from the Leverhulme Trust. The work of MT is supported by a Juan de la Cierva Fellowship.

2. Two-dimensional Yang-Mills theory and its q-deformations

In this section we review some computational aspects and exact analytical expressions for the quantum partition function of two-dimensional Yang-Mills theory. We then survey some of the various q-deformed versions which have appeared in the literature.

2.1. **Ordinary gauge theory.**

The action of Yang-Mills theory with compact gauge group G on an arbitrary connected and oriented Riemann surface Σ with unit area form $\mathrm{d}\mu$ is given by

$$S_{\rm YM}[A]=\frac{1}{4g_s}\,\big(F_A,F_A\big):=-\frac{1}{2g_s}\int_\Sigma \mathrm{d}\mu\ \ {\rm Tr}\,\big(F_A^2\big)\ , \tag{2.1}$$

where the positive parameter g_s plays the role of the coupling constant, $F_A\in\Omega^2(\Sigma,\mathfrak{g})$ is the curvature of a gauge connection $A\in\Omega^1(\Sigma,\mathfrak{g})$ on a trivial principal G-bundle over Σ, and ${\rm Tr}$ is an invariant quadratic form on the Lie algebra $\mathfrak{g}$ of G. Note that the definition (2.1) does not depend on any choice of metric on the surface, but only on a choice of measure $\mathrm{d}\mu$ on Σ which represents the generator of $H^2(\Sigma,\mathbb{Z})=\mathbb{Z}$. In this paper we are mainly interested in the case of a unitary gauge group $G=U(N)$; then ${\rm Tr}$ refers to the trace in the fundamental representation of G. We denote by $\mathcal{A}=\Omega^1(\Sigma,\mathfrak{g})$ the affine space of gauge connections and by $\mathcal{G}=\Omega^0(\Sigma,G)$ the group of gauge transformations.

The quantum gauge theory is defined by the path integral

$$Z_{\rm YM}(g_s;\Sigma):=\frac{1}{{\rm vol}(\mathcal{G})}\,\Big(\,\frac{1}{2\pi\,g_s}\,\Big)^{\dim\mathcal{G}/2}\ \int_{\mathcal{A}}\mathscr{D}\mu[A]\ \exp\big(-S_{\rm YM}[A]\big)\ , \tag{2.2}$$

where $\mathscr{D}\mu[A]$ is the translation-invariant Riemannian measure induced by the metric $(-,-)$ on $\mathcal{A}$. We similarly define an invariant metric on $\mathcal{G}$ (by the same formula), which formally determines the volume of $\mathcal{G}$.

We can also write this partition function as

$$Z_{\rm YM}(g_s;\Sigma)=\frac{1}{{\rm vol}(\mathcal{G})}\int_{\mathcal{A}}\mathscr{D}\mu[A]\ \int_{\Omega^0(\Sigma,\mathfrak{g})}\mathscr{D}\mu[\phi]\ \exp\big(-S_{\rm BF}[\phi,A]\big)\ , \tag{2.3}$$

where

$$S_{\rm BF}[\phi,A]=\ {\rm i}\ \langle F_A,\phi\rangle+\frac{g_s}{2}\,(\phi\,,\,\phi)=-\int_\Sigma\,{\rm Tr}\,\Big({\rm i}\,\phi\,F_A-\frac{g_s}{2}\,\phi^2\ \mathrm{d}\mu\Big) \tag{2.4}$$

is the first order form of the Yang-Mills action functional; in the weak coupling limit $g_s=0$ this is the action for topological BF-theory on the Riemann surface Σ and in this case (2.3) computes the symplectic volume of the moduli space of flat G-connections on Σ. The Euclidean measure $\mathscr{D}\mu[\phi]$ on the Lie algebra of $\mathcal{G}$ is determined by the same invariant form that we use to define the volume ${\rm vol}(\mathcal{G})$. We may regard the curvature $F_A\in\Omega^2(\Sigma,\mathfrak{g})$ as an element of the dual of the Lie algebra of $\mathcal{G}$; then $\langle-,-\rangle$ denotes the pairing between the Lie algebra of $\mathcal{G}$ and its dual. The equality between (2.2) and (2.3) follows from the functional Gaussian integration over the scalar field $\phi \in \Omega^0(\Sigma, \mathfrak{g})$.

The Yang-Mills partition function can be defined and evaluated rigorously via a combinatorial formalism obtained through lattice regularization [4, 1]. This leads to the combinatorial heat kernel expansion for the partition function on a compact Riemann surface Σ_h of genus h which is given by Migdal's formula

$$\mathcal{Z}_{\rm M}(g_s;\Sigma_h) = \Big(\,\frac{{\rm vol}(G)}{(2\pi)^{\dim G}}\,\Big)^{2h-2}~\sum_\lambda\,(\dim\lambda)^{2-2h}~\exp\Big(-\frac{g_s}{2}\,C_2(\lambda)\Big)~, \tag{2.5}$$

where the sum runs over all isomorphism classes of irreducible unitary representations of the gauge group G, $\dim\lambda$ is the dimension of the representation λ, and $C_2(\lambda)$ is the quadratic Casimir invariant of λ associated with the invariant quadratic form ${\rm Tr}$ on the Lie algebra $\mathfrak{g}$. Here ${\rm vol}(G)$ is the volume of G determined by ${\rm Tr}$; we will usually drop this volume factor (as well as other normalization constants) in the following. The formula (2.5) can also be derived in an operator formalism from canonical quantization of the continuum gauge theory defined by the action (2.4) [27]. For $G=U(N)$ the representations can be identified with N-component partitions $\lambda=(\lambda_1,\dots,\lambda_N)$ (equivalently Young diagrams), which are the highest weights. The dimension and quadratic Casimir eigenvalue of the corresponding representation are then given by the explicit formulas

$$\dim\lambda=\prod_{i<j}\,\frac{\lambda_i-\lambda_j+j-i}{j-i} \qquad \text{and} \qquad C_2(\lambda)=\langle\lambda,\lambda+2\rho\rangle=\sum_{i=1}^N\,\big(\lambda_i^2+(N+1-2i)\,\lambda_i\big)~,$$

where $\langle\lambda,\mu\rangle:=\sum_i\,\lambda_i\,\mu_i$ is the invariant bilinear pairing induced by ${\rm Tr}$ and

$$\rho_i=\frac{N-2i+1}{2}$$

are the components of the Weyl vector ρ of G (the half-sum of positive roots).

In the following we will also exploit the relationship with the specialization of the Schur polynomials $s_\lambda(x)$ in N variables $x=(x_1,\dots,x_N)$, which form an orthonormal basis on the ring of symmetric polynomials with respect to the Hall inner product [28]; they can be defined as the character of the $\mathfrak{g}$-module λ evaluated on a matrix X whose eigenvalues are $x_1,\dots,x_N$, i.e. $s_\lambda(x)={\rm Tr}_{\,\lambda}(X)$. Then the Weyl dimension formula for the irreducible representation of G with highest weight λ can be expressed as

$$\dim\lambda=s_\lambda(1,\dots,1)~.$$

Rather than working directly with (2.5), it will sometimes be more practical to change summation variables $\mu_i=\lambda_i+N-i$ for $i=1,\dots,N$ and use the associated discrete Gaussian matrix model [29]

$$\mathcal{Z}_{\rm M}(g_s;\Sigma_h)=\sum_{\mu\in\mathbb{Z}^N}~\prod_{i<j}\,(\mu_i-\mu_j)^{2-2h}~\exp\Big(-\frac{g_s}{2}\,\sum_{i=1}^N\,\mu_i^2\Big)~, \tag{2.6}$$

which agrees with (2.5) up to a standard area-dependent renormalization $\mu_i\to\mu_i-\frac{N-1}{2}$. The dimensions in (2.5) lead to a Vandermonde determinant and the Casimir invariants give the Gaussian potential of the matrix model, but in terms of the discrete eigenvalues $\mu=(\mu_1,\dots,\mu_N)\in\mathbb{Z}^N$.

The expansion (2.6) can also be derived directly from (2.3) by the technique of diagonalization [30]. As we will make reference to it later on, let us briefly review this calculation. The crux of the technique is the classical Weyl integral formula for Lie algebras: Let $f(\phi)$ be a conjugation invariant function on the Lie algebra $\mathfrak{g}$,

$$f\big(g^{-1}\,\phi\,g\big)=f(\phi)\qquad\text{for}\quad g\in G\ ,$$

and assume that it is integrable with respect to the normalized invariant Haar measure $\mathrm{d}\phi$ on $\mathfrak{g}$. The Lie algebra element $\phi\in\mathfrak{g}$ can be conjugated by an element $U\in G$ into the Cartan subalgebra $\mathfrak{h}\subset\mathfrak{g}$; then $f(\phi)$ depends only on the eigenvalues $(\phi_1,\dots,\phi_N)\in\mathbb{R}^N$. Integrating over U gives a factor of the order of the Weyl group $W=\mathfrak{S}_N$ of $G=U(N)$, which is the residual group of conjugation symmetries of $\mathfrak{h}$ acting by permutations of the eigenvalues. Then the Weyl integral formula reads as

$$\int_{\mathfrak{g}}\mathrm{d}\phi\ f(\phi)=\frac{1}{N!}\int_{\mathbb{R}^N}\ \prod_{i=1}^N\mathrm{d}\phi_i\ \prod_{j<k}(\phi_j-\phi_k)^2\ f(\phi_1,\dots,\phi_N)\ , \tag{2.7}$$

where the Vandermonde determinant factor arises as the Jacobian for the change of variables induced by conjugation into the Cartan subalgebra.

Let us now apply the diagonalization formula (2.7) to the BF-form of the partition function (2.3). For this, we observe that the functional

$$F[\phi]:=\frac{1}{\mathrm{vol}(\mathcal{G})}\int_{\mathcal{A}}\mathscr{D}\mu[A]\ \exp\big(-S_{\rm BF}[\phi,A]\big) \tag{2.8}$$

for fixed $\phi\in\Omega^0(\Sigma,\mathfrak{g})$ is conjugation invariant. We can thus use a local gauge transformation U to impose the torus gauge condition where the scalar field $\phi\in\Omega^0(\Sigma_h,\mathfrak{g})$ takes values in the Cartan subalgebra $\mathfrak{h}$ of the Lie algebra $\mathfrak{g}=\mathfrak{h}\oplus\mathfrak{k}$. However, while the diagonalised field $\phi=(\phi_1,\dots,\phi_N)\in\Omega^0(\Sigma_h,\mathfrak{h})$ is globally-defined and smooth, there are obstructions to finding smooth gauge functions $U\in\mathcal{G}$ globally; these obstructions are parametrized by the set of all isomorphism classes of non-trivial torus bundles $\mathcal{L}_n\to\Sigma_h$ with structure group the maximal torus $T=U(1)^N$, which arise as restrictions of the original trivial principal G-bundle over Σ_h [31]. The $\mathfrak{h}$-component $A^{\mathfrak{h}}$ of the gauge connection $A=(A^{\mathfrak{h}},A^{\mathfrak{k}})\in\Omega^1(\Sigma_h,\mathfrak{g})$ is a gauge field on a principal torus bundle $\mathcal{L}_n\to\Sigma_h$, while the $\mathfrak{k}$-components $A^{\mathfrak{k}}$ are sections of the associated vector bundle $\mathcal{L}_n\times_T\mathfrak{k}$. The path integral version of the Weyl integral formula (2.7) should then include a sum over contributions from connections on all isomorphism classes of T-bundles; we denote the affine space of abelian gauge connections on a principal T-bundle $\mathcal{L}_n\to\Sigma_h$ by $\mathcal{A}_n$. The maximal torus bundles $\mathcal{L}_n\to\Sigma_h$ are parametrized by their first Chern classes $c_1(\mathcal{L}_n)=n=(n_1,\dots,n_N)\in\mathbb{Z}^N$; by Chern-Weil theory one has

$$\frac{1}{2\pi}\int_{\Sigma_h}\mathrm{d}A_i^{\mathfrak{h}}=n_i$$

for $A^{\mathfrak{h}}=(A_1^{\mathfrak{h}},\dots,A_N^{\mathfrak{h}})\in\mathcal{A}_n$. Then the Weyl integral formula for the partition function (2.3) gives

$$\begin{aligned}Z_{\rm YM}(g_s;\Sigma_h)\ &=\ \frac{1}{\mathrm{vol}(\mathcal{G})}\sum_{n\in\mathbb{Z}^N}\int_{\mathcal{A}_n}\mathscr{D}\mu[A^{\mathfrak{h}}]\int_{\Omega^1(\Sigma_h,\mathcal{L}_n\times_T\mathfrak{k})}\mathscr{D}\mu[A^{\mathfrak{k}}]\\ &\qquad\times\int_{\Omega^0(\Sigma_h,\mathbb{R}^N)}\ \prod_{i=1}^N\mathscr{D}\mu[\phi_i]\ \Big[\ \prod_{j<k}(\phi_j-\phi_k)^2\Big]\ \exp\big(-S_{\rm BF}[\phi,A^{\mathfrak{h}},A^{\mathfrak{k}}]\big)\ ,\end{aligned}$$

where

$$S_{\rm BF}[\phi,A^{\mathfrak{h}},A^{\mathfrak{k}}]=\sum_{i=1}^N\ \int_{\Sigma_h}\Big(-\,\mathrm{i}\,\phi_i\,\mathrm{d}A_i^{\mathfrak{h}}+\frac{g_s}{2}\,\phi_i^2\,\mathrm{d}\mu\Big)+\sum_{\alpha\in\mathrm{Ad}(G)}\ \int_{\Sigma_h}\langle\alpha,\phi\rangle\,A_\alpha^{\mathfrak{k}}\wedge A_{-\alpha}^{\mathfrak{k}} \tag{2.9}$$

and $\alpha \in \mathrm{Ad}(G)$ are the roots of the Lie algebra $\mathfrak{g}$. Integrating over $A^{\mathfrak{k}}_\alpha$ gives an inverse functional determinant induced by a one-form in $\Omega^1(\Sigma_h)$. As demonstrated in [30], by using the Hodge decomposition of forms on Σ_h one finds that only harmonic forms contribute and yield Vandermonde determinant factors $\prod_{i<j}(\phi_i - \phi_j)^{-2h}$, as there are $2h$ linearly independent harmonic one-forms on the Riemann surface Σ_h. Altogether this reduces the path integral (2.3) to that of an abelian gauge theory based on the maximal torus $T = U(1)^N$. Any gauge field $A^{\mathfrak{h}} \in \mathcal{A}_n$ can be decomposed as $A^{\mathfrak{h}} = a^{\mathfrak{h}} + \tilde{A}^{\mathfrak{h}}$, where the monopole connections $a^{\mathfrak{h}}$ obey $\mathrm{d}a^{\mathfrak{h}}_i = 2\pi\, n_i\, \mathrm{d}\mu$ and integrating over $\tilde{A}^{\mathfrak{h}}_i \in \Omega^1(\Sigma_h)$ gives delta-function constraints

$$\mathrm{d}\phi_i = 0$$

which imply that the scalar fields ϕ_i are constant on Σ_h for each $i = 1, \dots, N$. The complete path integral is now reduced to a sum of finite-dimensional integrals

$$Z_{\mathrm{YM}}(g_s; \Sigma_h) = \sum_{n\in\mathbb{Z}^N} \int_{\mathbb{R}^N} \prod_{i=1}^N \mathrm{d}\phi_i \ \ \mathrm{e}^{\,2\pi\,\mathrm{i}\, n_i\, \phi_i - \frac{g_s}{2}\,\phi_i^2} \ \prod_{j<k} (\phi_j - \phi_k)^{2-2h} \ .$$

The sum over torus bundles $n \in \mathbb{Z}^N$ gives periodic delta-functions via the Poisson resummation formula

$$\sum_{n_i=-\infty}^{\infty} \mathrm{e}^{\,2\pi\,\mathrm{i}\, n_i\, \phi_i} = \sum_{\mu_i=-\infty}^{\infty} \delta(\phi_i - \mu_i) \ ,$$

and we arrive finally at the discrete Gaussian matrix model (2.6).

2.2. q-deformed gauge theories.

An interesting variant of the model described above is the "q-deformation" of two-dimensional Yang-Mills theory. Formally, it arises when one considers the partition function (2.3) but now with the Gaussian integral taken over the domain $\phi \in \Omega^0(\Sigma, G)$. The partition function of q-deformed Yang-Mills theory on Σ_h can again be computed by diagonalization [7, 10] and it results in a simple variation of (2.5) given by the expansion

$$\mathscr{Z}^{(p)}_{\mathrm{M}}(q; \Sigma_h) = \sum_\lambda (\dim_q \lambda)^{2-2h} \exp\Big(-\frac{p\, g_s}{2}\, C_2(\lambda)\Big) \ , \tag{2.10}$$

but now involving the *quantum* dimensions

$$\dim_q \lambda = s_\lambda\big(q^\rho\big) = q^{|\lambda|/2}\ s_\lambda\big(1, q, \dots, q^{N-1}\big) = \prod_{i<j} \frac{[\lambda_i - \lambda_j + j - i]_q}{[j-i]_q}$$

of the representations of G given by the Weyl character formula, where $p \in \mathbb{Z}_{>0}$, the deformation parameter is $q := \mathrm{e}^{\,-g_s}$ and we have defined $q^\rho := (q^{\rho_1}, \dots, q^{\rho_N})$. Here $|\lambda| := \sum_i \lambda_i$ is the number of boxes in the Young diagram corresponding to the partition λ, and the symmetric q-number

$$[x]_q := \frac{q^{x/2} - q^{-x/2}}{q^{1/2} - q^{-1/2}} \tag{2.11}$$

is defined for $q \neq 1$ and any $x \in \mathbb{R}$; note that

$$[n]_q = \sum_{i=1}^n q^{\frac12\,(n-2i+1)} \qquad \text{for} \quad n \in \mathbb{Z}_{>0} \ .$$

One can analytically continue this expression to arbitrary values $q \in \mathbb{C}$. When $q = \mathrm{e}^{\,2\pi\,\mathrm{i}/(k+N)}$ is a root of unity, the series (2.10) acquires an affine Weyl symmetry and should be truncated to the Weyl alcove consisting of integrable representations of G at level k; the resulting sum is then the partition function of Chern-Simons gauge theory at level $k \in \mathbb{Z}$ on a circle bundle of degree p over the Riemann surface Σ_h. For $p = 1$ it is related to the G_k-WZW model on

Σ_h. For $p=0$ the gauge theory is a q-deformation of two-dimensional BF-theory and the partition function reproduces the Verlinde formula for the dimension of the Hilbert space of G_k-WZW conformal blocks on Σ_h; this q-deformed BF-theory is the gauged G/G WZW model, or equivalently Chern-Simons theory on the trivial circle bundle $\Sigma_h \times S^1$, and it may be regarded as a non-linear deformation of ordinary Yang-Mills theory. For generic p the partition function (2.10) computes certain intersection indices on the moduli space of Yang-Mills connections on Σ_h, including those which are not flat. The connections between these two-dimensional and three-dimensional field theories are also analysed in [32].

The discrete matrix model corresponding to (2.10) involves a q-deformed Vandermonde determinant and an ordinary Gaussian potential; it is given by [7]

$$\mathscr{Z}_{\rm M}^{(p)}(q;\Sigma_h) = \sum_{\mu\in\mathbb{Z}^N}\ \prod_{i<j}\,[\mu_i-\mu_j]_q^{2-2h}\ \exp\Big(-\frac{p\,g_s}{2}\,\sum_{i=1}^N\mu_i^2\Big)\ . \tag{2.12}$$

The ordinary Yang-Mills partition functions $\mathscr{Z}_{\rm M}(\tilde g_s;\Sigma_h)$ from (2.5) and (2.6) are respectively recovered from (2.10) and (2.12) in the double scaling limit $p\to\infty$, $g_s\to 0$ (so that $q\to 1$) with the renormalized coupling constant $\tilde g_s := p\,g_s$ fixed.

In the formulation of [7, 10], the q-deformation arises from the change of integration measure for the compact scalar field ϕ, or alternatively, from the point of view of diagonalization, from a restricted sum over torus bundles to those which are of torsion class $\mathbb{Z}_p$; it can be understood in terms of the difference between the Weyl integral formulas for Lie *groups* and for Lie *algebras*. For this, let $f(U)$ be a conjugation invariant function on the Lie group G,

$$f\big(g^{-1}\,U\,g\big) = f(U) \qquad \text{for} \quad g\in G\ ,$$

and assume that it is integrable with respect to the normalized invariant Haar measure ${\rm d}U$ on G. We can again conjugate $U\in G=U(N)$ into the maximal torus $T=U(1)^N$; then $f(U)$ depends only on the eigenvalues $(u_1,\dots,u_N)\in(S^1)^N$. We parametrize these eigenvalues as $u_i={\rm e}^{\,{\rm i}\,\phi_i}$ with $\phi_i\in[0,2\pi)$ for $i=1,\dots,N$. Then the Weyl integral formula relates the measure on $G/{\rm Ad}(G)=T/W$ coming from the Haar measure on G with the Haar measure on T, and it modifies (2.7) to

$$\int_G {\rm d}U\ f(U) = \frac{1}{N!}\,\int_{[0,2\pi)^N}\ \prod_{i=1}^N\frac{{\rm d}\phi_i}{2\pi}\ \prod_{j<k}4\sin^2\Big(\frac{\phi_j-\phi_k}{2}\Big)\ f(\phi_1,\dots,\phi_N) \tag{2.13}$$

where $f(\phi_1,\dots,\phi_N):=f({\rm e}^{\,{\rm i}\,\phi_1},\dots,{\rm e}^{\,{\rm i}\,\phi_N})$; here the determinant of the conjugation map on G is the Weyl determinant. We now follow the same steps that led to the formula (2.6) from (2.7). The path integral is given by

$$\begin{aligned} Z_{\rm YM}^{(p)}(q;\Sigma_h) \ &= \ \frac{1}{{\rm vol}(\mathcal{G})}\,\sum_{n\in\mathbb{Z}^N}\ \int_{\mathcal{A}_n}\mathscr{D}\mu[A^{\mathfrak{h}}]\ \int_{\Omega^1(\Sigma_h,\mathcal{L}_n\times_T\mathfrak{k})}\mathscr{D}\mu[A^{\mathfrak{k}}] \\ &\qquad \times\ \int_{\Omega^0(\Sigma_h,(\mathbb{R}/2\pi\,\mathbb{Z})^N)}\ \prod_{i=1}^N\mathscr{D}\mu[\phi_i]\ \Big[\ \prod_{i<j}4\sin^2\Big(\frac{\phi_j-\phi_k}{2}\Big)\Big] \\ &\qquad \times\ \exp\big(-S_{\rm BF}^{(p)}[\phi,A^{\mathfrak{h}},A^{\mathfrak{k}}]\big)\ , \end{aligned}$$

where the action (2.9) is now modified to

$$S_{\rm BF}^{(p)}[\phi,A^{\mathfrak{h}},A^{\mathfrak{k}}] = \frac{1}{g_s}\,\sum_{i=1}^N\ \int_{\Sigma_h}\Big(-{\rm i}\,\phi_i\,{\rm d}A_i^{\mathfrak{h}}+\frac{p}{2}\,\phi_i^2\,{\rm d}\mu\Big)+\sum_{\alpha\in{\rm Ad}(G)}\ \int_{\Sigma_h}\big(1-{\rm e}^{\,{\rm i}\,\langle\alpha,\phi\rangle}\big)\,A_\alpha^{\mathfrak{k}}\wedge A_{-\alpha}^{\mathfrak{k}}$$

and we have rescaled $g_s \to p\, g_s$ and $\phi \to \phi/g_s$ in the original BF-theory action (2.4). The sum over torus bundles in this way yields the q-deformed discrete Gaussian matrix model (2.12).

A similar formula for the partition function is derived in the lattice formulation of [5], wherein the gauge algebra is taken to be the quantum universal enveloping algebra $\mathcal{U}_q(\mathfrak{g})$ (see Appendix A); invariance of the lattice gauge fields under the coaction of $\mathcal{U}_q(\mathfrak{g})$ implies that the algebra of gauge fields coincides with the noncommutative exchange algebra of gauge fields which arises in Hamiltonian quantization of lattice Chern-Simons theory [33]. By using a suitable q-analog of the usual lattice Yang-Mills-Haar measure, the partition function for the resulting q-deformation of Yang-Mills theory is given by a similar series over q-weights.

In [6], Klimčík considers a general class of "Poisson-Lie Yang-Mills theories" which are obtained by an isotropic gauging of the WZW model on a double $D(G)$ of the gauge group G; these field theories are gauge invariant with respect to two mutually commuting actions of $\mathcal{G}$, and their zero coupling limits are Poisson sigma-models on Σ. When $D(G) = T^*G \cong G \rtimes \mathfrak{g}^*$ is the cotangent bundle of G, the partially gauge-fixed action is that of the standard Yang-Mills theory (2.4) and the associated Poisson sigma-model is BF-theory on Σ. When $D(G) = G \times G$, the associated Poisson sigma-model is the G/G WZW model. Our main interest here is the Lu-Weinstein-Soibelman Drinfel'd double $D(G) = G^{\mathbb{C}}$, with the complexification of G regarded as a real group. In this case there is an extra parameter $\hbar^{-1}$ which multiplies the invariant non-degenerate form $\mathfrak{Im}\ \mathrm{Tr}$ on the Lie algebra of $D(G)$, and we set $q := \mathrm{e}^{-4\pi\hbar}$ (now independently of the coupling g_s); this defines a one-parameter deformation of the standard Yang-Mills theory, which is recovered in the limit $q \to 1$. The associated Poisson sigma-models now include the standard ones whose perturbation series compute the Kontsevich formality maps for global deformation quantization [34]. Klimčík finds a suitable extension of the diagonalization technique to accomodate the double $D(G) = G^{\mathbb{C}}$, which follows from a generalization of the Weyl integral formula (2.13) based on an alternative form of the Cartan decomposition of elements of $G^{\mathbb{C}}$. Proceeding as before to write the path integral as a sum over contributions from all torus bundles on Σ_h, one derives the combinatorial expansion [6, eq. (128)]

$$\mathcal{Z}_{\mathrm{K}}(g_s, q; \Sigma_h) = \sum_\lambda (\dim_q \lambda)^{2-2h} \exp\Big(-\frac{g_s}{2} \sum_{i=1}^N \big([\lambda_i + \rho_i]_q^2 - [\rho_i]_q^2 \big) \Big) . \tag{2.14}$$

When $q = 1$ this is just the heat kernel expansion (2.5) of ordinary Yang-Mills theory. For $g_s = 0$ and q a root of unity, the truncation of this series to integrable highest weights gives the Verlinde formula.

Brzeziński and Majid [35] also consider a version of q-deformed Yang-Mills theory based on a general construction of quantum group-valued connections on quantum principal bundles with Hopf algebra fibre. However, its quantization is not clear, as it is not obvious how to define the path integral on a space of connections taking values in $\mathcal{U}_q(\mathfrak{g})$ such that the ordinary Yang-Mills partition function is recovered in the classical limit $q \to 1$. We shall return to the issue of properly implementing the quantum group gauge symmetry in §6.7.

3. Continuous matrix models

In this section we will explicitly relate the discrete matrix models (2.12) for q-deformed Yang-Mills theory on the sphere $\Sigma_0 = S^2$ to the continuous Stieltjes-Wigert matrix models for Chern-Simons gauge theory on the lens space $L(p,1) = S^3/\mathbb{Z}_p$ for all $p \in \mathbb{Z}_{>0}$.

The Stieltjes-Wigert matrix model is characterized by certain mathematical features which are not present in classical random matrix ensembles [36]. In particular, it is defined by a weight function on $\mathbb{R}_{>0}$ of log-normal type

$$\omega_{\rm SW}(x;s) := \frac{s}{\sqrt{\pi}}\ {\rm e}^{\,-s\log^2 x} \qquad \text{with} \quad s\in\mathbb{R}_{>0}\ , \tag{3.1}$$

which leads to an indeterminate moment problem [37]. This implies that the Stieltjes-Wigert orthogonal polynomials that solve the Chern-Simons matrix model [38] are not dense in the Hilbert space $L^2\big(\mathbb{R}_{>0}\,,\,\omega_{\rm SW}(x;s)\,{\rm d}x\big)$ [38, 39]; one consequence of this feature is the exact discretization of the matrix model [39]. The Stieltjes-Wigert polynomials depend on the q-parameter $q={\rm e}^{\,-1/2s^2}$.

Recall from (2.6) that the partition function of ordinary two-dimensional Yang-Mills theory on S^2 can be expressed as a discrete Gaussian matrix model

$$\mathscr{Z}_{\rm M}(g_s;S^2)=\sum_{u\in\mathbb{Z}^N}\ \exp\Big(-\frac{g_s}{2}\,\sum_{i=1}^N\,u_i^2\Big)\ \prod_{j<k}\,(u_j-u_k)^2\ .$$

Unlike the continuous case, which is a Gaussian unitary ensemble that is solved with Hermite polynomials [36], the discrete Gaussian weight does not have a closed system of orthogonal polynomials associated to it [29]. In fact, the large N phase transition of the gauge theory is related to the discrepancy between the discrete and continuous matrix models [29].

In marked contrast, the orthogonal polynomials for the discrete matrix model of Chern-Simons theory are, just as in the continuous case, the Stieltjes-Wigert polynomials [38]; this is related to the fact that the orthogonality measure for the Stieltjes-Wigert polynomials is not uniquely determined, and also to the absence of a large N phase transition in the q-deformed gauge theory in this case [9]. We demonstrate this in detail below, showing explicitly the equivalence between the discrete and continuous versions of the Chern-Simons matrix model. This calculation fills in the details of the derivation in [39, eqs. (25)–(26)] and gives explicitly the corresponding normalization constants. It specifically relates q-deformed Yang-Mills theory on S^2 for $p=1$, which is characterized by a discrete matrix model, with Chern-Simons theory on S^3, which was originally described by a continuous matrix model [40]. The moment problem will also be employed below to study the analogous relationship in the case $p\neq 1$; we show that in this instance the discrete matrix model is related to certain correlators in the Stieltjes-Wigert ensemble. We further comment on some aspects of the equivalent representation of the Stieltjes-Wigert matrix model as a unitary matrix model, as this correspondence will be exploited in our later considerations; in contrast to the weight functions defined on $\mathbb{R}$, weight functions on the unit circle S^1 are always moment-determined [36].

3.1. $L(1,1)$ **matrix model.**

We begin with the discrete matrix model (2.12) for the partition function (2.10) on $\Sigma_0=S^2$ for $p=1$ and rewrite it as a continuous matrix model. We have

$$\begin{aligned}\mathscr{Z}_{\rm M}^{(1)}(q;S^2)&=\sum_{u\in\mathbb{Z}^N}\ \exp\Big(-\frac{g_s}{2}\,\sum_{i=1}^N\,u_i^2\Big)\ \prod_{j<k}\,4\,\sinh^2\Big(\frac{g_s}{2}\,(u_j-u_k)\Big)\\ &=\sum_{u\in\mathbb{Z}^N}\ \exp\Big(-\frac{g_s}{2}\,\sum_{i=1}^N\,u_i^2\Big)\ \exp\Big((N-1)\,g_s\,\sum_{i=1}^N\,u_i\Big)\ \prod_{j<k}\,\big({\rm e}^{\,-g_s\,u_j}-{\rm e}^{\,-g_s\,u_k}\big)^2\\ &=\sum_{u\in\mathbb{Z}^N}\ q^{\frac12\,\sum_i\,u_i^2}\,(\sigma\,q)^{\sum_i\,u_i}\ \prod_{j<k}\,(q^{u_j}-q^{u_k})^2\end{aligned}$$

where we have introduced $\sigma := \mathrm{e}^{N\,g_s} = q^{-N}$ with $q := \mathrm{e}^{-g_s}$ as usual. This gives

$$\begin{aligned} \mathcal{Z}_{\mathrm{M}}^{(1)}(q;S^2) &= \sum_{u\in\mathbb{Z}^N}\ \prod_{i=1}^N \sigma^{u_i}\, q^{\frac12\, u_i^2+u_i}\ \prod_{j<k}\, \big(q^{u_j}-q^{u_k}\big)^2 \\ &= \sigma^{N\,(1-N)} \int_{\mathbb{R}^N_{>0}}\ \prod_{i=1}^N \mathrm{d}x_i\ \sum_{n_i=-\infty}^\infty \sigma^{n_i}\, q^{\frac12\, n_i^2+n_i}\, \delta(x_i-\sigma\, q^{n_i})\ \prod_{j<k}\, (x_j-x_k)^2 \\ &= \sigma^{N\,(1-N)}\, q^{N/2}\, M(q,\sigma)^N \int_{\mathbb{R}^N_{>0}}\ \prod_{i=1}^N \mathrm{d}x_i\ w_d(x_i;q,\sigma)\ \prod_{j<k}\, (x_j-x_k)^2 \ , \end{aligned} \tag{3.2}$$

where the normalization $M(q,\sigma)$ has a triple product form

$$M(q,\sigma) := \big(-\sigma\, q^{3/2}\,;\,q\big)_\infty\, \big(-\sigma^{-1}\, q^{-1/2}\,;\,q\big)_\infty\, \big(q\,;\,q\big)_\infty \ , \tag{3.3}$$

and for $a,q\in\mathbb{C}$ with $|q|<1$ and $k\in\mathbb{Z}_{>0}\cup\{\infty\}$ we use the standard hypergeometric notation for the q-shifted factorial

$$(a;q)_k := \prod_{n=0}^{k-1}\, \big(1-a\, q^n\big) \qquad \text{and} \qquad (a;q)_0 := 1 \ .$$

In (3.2) we have introduced the two-parameter family of discrete measures on $\mathbb{R}_{>0}$ given by

$$w_d(x;q,\sigma) := \frac{1}{\sqrt{q}\, M(q,\sigma)}\ \sum_{n=-\infty}^\infty \sigma^n\, q^{\frac12\, n^2+n}\, \delta(x-\sigma\, q^n) \ . \tag{3.4}$$

The family $\big(w_d(x;q,\sigma)\big)_{\sigma>0}$ is completely determined by the values $\sigma\in(q,1]$, and translation invariance of the sums leads to the discrete scaling symmetry [41]

$$M(q,q\,\sigma) = \frac{M(q,\sigma)}{\sigma\,\sqrt{q}} \qquad \text{and} \qquad w_d(x;q,q\,\sigma) = w_d(x;q,\sigma) \ . \tag{3.5}$$

The indeterminacy of the moment problem implies that the discrete measure (3.4) is equivalent to the continuous distribution (3.1) [42, 43, 44]. Note that even though there are infinitely many discrete measures $w_d(x;q,\sigma)$ which are equivalent to $\omega_{\rm SW}(x;\sigma)$, in (3.2) we take $\sigma=q^{-N}$, which by (3.5) is equivalent to the choice $\sigma=1$. The equivalence between $w_d(x;q,1)$ and $\omega_{\rm SW}\big(x;\frac{1}{\sqrt{2g_s}}\big)$ enables us to write

$$\mathcal{Z}_{\mathrm{M}}^{(1)}(q;S^2) = \sigma^{N\,(1-N)}\, q^{N/2}\, M(q,\sigma)^N \int_{\mathbb{R}^N_{>0}}\ \prod_{i=1}^N \mathrm{d}x_i\ \omega_{\rm SW}\big(x_i\,;\,\tfrac{1}{\sqrt{2g_s}}\big)\ \prod_{j<k}\, (x_j-x_k)^2 \ , \tag{3.6}$$

which up to overall normalization is the partition function $Z_N(q)$ for Chern-Simons gauge theory on the three-sphere S^3 [38].

In this way one arrives at a rather simple relation between the discrete and continuous Stieltjes-Wigert ensembles given by

$$\begin{aligned} Z_N(q) \ &:= \ \Big(\frac{g_s}{2\pi}\Big)^{-N/2} \int_{\mathbb{R}^N}\ \prod_{i=1}^N \frac{\mathrm{d}u_i}{2\pi}\ \mathrm{e}^{-u_i^2/2g_s}\ \prod_{j<k}\, 4\,\sinh^2\Big(\frac{u_j-u_k}{2}\Big) \\ &= \ \bigg(\frac{q^{-\frac12\,(1-2N+3N^2)}}{\big(-q^{\frac32-N}\,;q\big)_\infty\, \big(-q^{N-\frac12}\,;\,q\big)_\infty\, \big(q\,;\,q\big)_\infty}\bigg)^N \\ &\quad\ \times \sum_{n\in\mathbb{Z}^N} \exp\Big(-\frac{g_s}{2}\, \sum_{i=1}^N n_i^2\Big)\ \prod_{j<k}\, 4\,\sinh^2\Big(\frac{g_s}{2}\,(n_j-n_k)\Big) \end{aligned} \tag{3.7}$$

where $u_i = \log x_i$. This relates the matrix model that arises in q-deformed Yang-Mills theory on S^2 with $p = 1$ (in the second line of (3.7)) to the matrix model in Chern-Simons theory on S^3 (in the first line of (3.7)). It strengthens the analogous relationship derived in [45] where the integration in (3.7) was interpreted as a Jackson q-integral.

3.2. $L(p,1)$ matrix model.

We now consider the case of general degree $p \in \mathbb{Z}_{>0}$. Then the discrete matrix model (2.12) for q-deformed Yang-Mills theory on S^2 is given by

$$\mathcal{Z}_{\rm M}^{(p)}(q;S^2) = \sum_{u\in\mathbb{Z}^N} \exp\Big(-\frac{p\,g_s}{2}\sum_{i=1}^N u_i^2\Big)\prod_{j<k} 4\sinh^2\Big(\frac{g_s}{2}\,(u_j-u_k)\Big)\ . \tag{3.8}$$

In the continuous limit $g_s \to 0$ with the rescaling $u_i \to g_s\,u_i$, this expression becomes

$$\lim_{q\to 1}\ \mathcal{Z}_{\rm M}^{(p)}(q;S^2) = \int_{\mathbb{R}^N}\prod_{i=1}^N {\rm d}u_i\ {\rm e}^{-\frac{p}{2g_s}\,u_i^2}\ \prod_{j<k} 4\sinh^2\Big(\frac{u_j-u_k}{2}\Big)\ . \tag{3.9}$$

By (3.7) the model defined by the partition function (3.9) is equivalent to the discrete matrix model

$$\mathcal{Z}_{\rm SW}^{(p)}(q;S^2) := \sum_{v\in\mathbb{Z}^N} \exp\Big(-\frac{g_s}{2p}\sum_{i=1}^N v_i^2\Big)\prod_{j<k} 4\sinh^2\Big(\frac{g_s}{2p}\,(v_j-v_k)\Big)\ .$$

By making the change of variables $v_i = p\,u_i$ we obtain

$$\mathcal{Z}_{\rm SW}^{(p)}(q;S^2) = \sum_{u\in(\mathbb{Z}/p)^N} \exp\Big(-\frac{p\,g_s}{2}\sum_{i=1}^N u_i^2\Big)\prod_{j<k} 4\sinh^2\Big(\frac{g_s}{2}\,(u_j-u_k)\Big)\ . \tag{3.10}$$

This shows that, when $p > 1$, the q-deformed gauge theory is actually a subsector of the discrete version of the Stieltjes-Wigert ensemble. This observation suggests a decomposition into sectors $\mathcal{Z}_{\rm SW}^{(p)}[n](q;S^2)$ labelled by elements $n_i = 0,1,\dots,p-1$ for $i = 1,\dots,N$ of the cyclic group $\mathbb{Z}_p$ in each of which the summation variables are restricted to $u_i \in \mathbb{Z} + \frac{n_i}{p}$. Then

$$\mathcal{Z}_{\rm SW}^{(p)}(q;S^2) = \sum_{n\in\mathbb{Z}_p^N} \mathcal{Z}_{\rm SW}^{(p)}[n](q;S^2)\ .$$

The different torsion sectors n_i are interlaced with one another through the hyperbolic sine function in (3.10); the model (3.8) is recovered as the trivial sector $n_i = 0$ for $i = 1,\dots,N$. Therefore, to study the discrepancy between the continuous matrix model (3.9) and the discrete matrix model (3.8), we will write the latter as a projection from the full ensemble

$$\mathcal{Z}_{\rm M}^{(p)}(q;S^2) = \mathcal{Z}_{\rm SW}^{(p)}[0](q;S^2)\ .$$

For this, we assume that $q = \zeta_k$ is a k-th root of unity, so that $q^k = 1$ (this is, in particular, the case relevant for the connection with Chern-Simons gauge theory). Consider the polynomial

$$\Gamma_{k,p}(z) = \frac{z^{k\,p}-1}{z^k-1} = \sum_{m=0}^{p-1} z^{k\,m}\ . \tag{3.11}$$

This is an equally spaced polynomial which arises in the study of finite fields [46] (here $\mathbb{Z}_p$); for $k = 1$ it is known as the "all in one polynomial". One then has

$$\Gamma_{k,p}(q^u) = p\ \delta_{n,0} \qquad \text{for} \quad u \in \mathbb{Z} + \tfrac{n}{p}\ .$$

Thus by introducing the multivariable polynomial extension of (3.11) given by

$$\Gamma_{k,p}(z_1,\dots,z_N)=\frac{1}{p^N}\,\prod_{i=1}^N\,\Gamma_{k,p}(z_i)\ ,$$

we obtain

$$\mathcal{Z}^{(p)}_{\rm SW}[0](q=\zeta_k;S^2)=\big\langle\Gamma_{k,p}(q^{u_1},\dots,q^{u_N})\big\rangle_{\rm SW}$$

where the average is taken in the full discrete Stieltjes-Wigert matrix model (3.10). Since $\Gamma_{k,p}(z_1,\dots,z_N)$ is a polynomial, only integer moments are involved in computing this correlator and the discrete/continuous equivalence, due to the indeterminate moment problem, applies to it as well. Hence we can write the average in the continuous matrix model to get

$$\mathcal{Z}^{(p)}_{\rm SW}[0](q=\zeta_k;S^2)=\int_{\mathbb{R}^N}\ \prod_{i=1}^N{\rm d}u_i\ \,{\rm e}^{\,-\frac{p}{2g_s}\,u_i^2}\ \frac{{\rm e}^{\,k\,p\,u_i}-1}{{\rm e}^{\,k\,u_i}-1}\ \prod_{j<k}\,4\,\sinh^2\Big(\,\frac{u_j-u_k}{2}\,\Big)\ ,$$

or alternatively by setting $u_i=\log x_i$ we may write

$$\mathcal{Z}^{(p)}_{\rm SW}[0](q=\zeta_k;S^2)=\int_{\mathbb{R}^N_{>0}}\ \prod_{i=1}^N{\rm d}x_i\ \omega_{\rm SW}\big(x_i\,;\,\sqrt{\tfrac{p}{2g_s}}\,\big)\,\Gamma_{k,p}(x_i)\ \prod_{j<k}\,(x_j-x_k)^2\ .$$

For $p=1$ these expressions respectively recover the continuous Stieltjes-Wigert matrix model representations (3.7) and (3.6).

3.3. **Unitary matrix models.**

For later reference, we briefly discuss the unitary matrix model that describes the $U(N)$ Chern-Simons gauge theory on S^3 [47]. It is given by the partition function

$$Z_N(q)=\int_{[0,2\pi)^N}\ \prod_{i=1}^N\,\frac{{\rm d}\phi_i}{2\pi}\ \Theta\big(\,{\rm e}^{\,{\rm i}\,\phi_i}\,;\,q\big)\ \prod_{j<k}\,\big|\,{\rm e}^{\,{\rm i}\,\phi_j}-{\rm e}^{\,{\rm i}\,\phi_k}\big|^2\ , \tag{3.12}$$

where the weight function of the matrix model is the Jacobi elliptic function

$$\Theta(z;q):=\sum_{n=-\infty}^\infty\ q^{n^2/2}\,z^n\ . \tag{3.13}$$

This theta-function can be written in a product form by using the Jacobi triple product identity

$$\Theta(z;q)=\big(q\,;\,q\big)_\infty\,\big(\sqrt{q}\,z\,;\,q\big)_\infty\,\big(\sqrt{q}\,z^{-1}\,;\,q\big)_\infty\ . \tag{3.14}$$

By the Heine-Szegő identity (see Appendix B), one can write the unitary matrix integral (3.12) as a Toeplitz determinant and derive the exact analytical expression [25]

$$Z_N(q)=\prod_{k=1}^{N-1}\,\big(1-q^k\big)^{N-k} \tag{3.15}$$

for the partition function (3.7) of the Stieltjes-Wigert matrix model.

The $U(N)$ Chern-Simons theory has a dual formulation as a $U(\infty)$ matrix model with the product form (3.14) of the theta-function truncated at N: It can be written as

$$Z_N(q)=(q;q)_\infty^N\ \int_{[0,2\pi)^\infty}\ \prod_{i=1}^\infty\,\frac{{\rm d}\phi_i}{2\pi}\ \frac{\Theta_N\big(\,{\rm e}^{\,{\rm i}\,\phi_i}\,;\,q\big)}{(q;q)_N}\ \prod_{j<k}\,\big|\,{\rm e}^{\,{\rm i}\,\phi_j}-{\rm e}^{\,{\rm i}\,\phi_k}\big|^2\ , \tag{3.16}$$

where the infinite-dimensional integration is defined as the $N\to\infty$ limit of the finite N eigenvalue model (see Appendix B for a discussion of the related convergence issues); the weight function of the matrix model is now a truncated theta-function which satisfies a finite version of the Jacobi triple product identity (see e.g. [48])

$$\Theta_N(z;q) = (q;q)_N \sum_{n=-N}^{N} \begin{bmatrix} 2N \\ n+N \end{bmatrix}_q q^{n^2/2}\, z^n = \big(q\,;\,q\big)_N \big(\sqrt{q}\,z\,;q\big)_N \big(\sqrt{q}\,z^{-1}\,;\,q\big)_N \tag{3.17}$$

and

$$\begin{bmatrix} n \\ k \end{bmatrix}_q := \frac{(q;q)_n}{(q;q)_k\,(q;q)_{n-k}}$$

is the q-binomial coefficient defined for $n,k \in \mathbb{Z}_{\geq 0}$ with $n \geq k$; the weights $(q;q)_n$ here have an algebraic interpretation as the polynomials which compute the number of flags in an n-dimensional vector space over a field with q elements. To prove the formula (3.16), we use the Cauchy-Binet formula for the normalization of the Schur measure [28]

$$\sum_\lambda s_\lambda(x)\,s_\lambda(y) = \prod_{i,j\geq 1} \frac{1}{1-x_i\,y_j}\ , \tag{3.18}$$

together with the Gessel identity (see e.g. [25]) which writes the left-hand side of (3.18) as a Toeplitz determinant. This gives

$$\prod_{i,j=1}^{N} \frac{1}{1-x_i\,y_j} = \int_{[0,2\pi)^\infty} \prod_{i=1}^{\infty} \frac{\mathrm{d}\phi_i}{2\pi} \prod_{j=1}^{N} \big(1+x_j\ \mathrm{e}^{\,\mathrm{i}\,\phi_i}\big)\big(1+y_j\ \mathrm{e}^{\,-\,\mathrm{i}\,\phi_i}\big) \prod_{k<l} \big|\,\mathrm{e}^{\,\mathrm{i}\,\phi_k} - \mathrm{e}^{\,\mathrm{i}\,\phi_l}\big|^2\ .$$

At the principal specialization $x_i = y_i = q^{i-\frac12}$ for $i = 1,\dots,N$, this identity shows that (3.16) is equal to (3.15).

This result can also be checked by explicit computation, using the Selberg integral: The $U(k)$ version of the matrix integral (3.16) can be computed as [49]

$$\begin{aligned} \int_{[0,2\pi)^k} \prod_{i=1}^{k} \frac{\mathrm{d}\phi_i}{2\pi}\ \frac{\Theta_n\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\,q\big)}{(q;q)_n} \prod_{j<k} \big|\,\mathrm{e}^{\,\mathrm{i}\,\phi_j} - \mathrm{e}^{\,\mathrm{i}\,\phi_k}\big|^2 \\ = \prod_{i=0}^{k-1} \frac{(q;q)_{i+2n}\,(q;q)_i}{(q;q)_{i+n}^2} = \prod_{j=0}^{n-1} \frac{(q;q)_j\,(q;q)_{k+j+n}}{(q;q)_{j+n}\,(q;q)_{k+j}}\ . \end{aligned} \tag{3.19}$$

The two equivalent expressions in (3.19) make manifest the duality described above: Using the property $(q;q)_n = (1-q^{n-1})\,(q;q)_{n-1}$, with either the limit $k\to\infty$, $n=N$ or the dual limit $k=N$, $n\to\infty$ we arrive at the partition function for $U(N)$ Chern-Simons theory on S^3.

The principal specialization of the Cauchy-Binet formula (3.18) also demonstrates that, up to area-dependent renormalization, the topological limit $p=0$ of the partition function (2.10) for q-deformed Yang-Mills theory on S^2 is related to the unitary matrix models discussed here via

$$\mathcal{Z}^{(0)}(q;S^2) = \sum_\lambda \big(\dim_q \lambda\big)^2 = \frac{Z_N(q)}{(q;q)_\infty^N}\ . \tag{3.20}$$

4. Classification

In this section we analyse the effect of a full q-deformation of the Migdal partition function (2.5), which is a simple and natural modification from the perspective of quantum group theory. By "full q-deformation" we mean a q-deformation of the dimensions, of the Casimir operator, and of the exponential function, the three main ingredients that figure into the combinatorial formula (2.5). This problem is partly inspired by the form of Klimčík's partition function (2.14), which involves q-deformed Casimir eigenvalues $[C_2(\lambda)]_q$ and is obtained from the partition function (2.10) by a full q-deformation of the heat kernel action. Another source of inspiration comes from the observation of [20] that the generic usage of quantum dimensions in the combinatorial quantization of two-dimensional Yang-Mills theory necessitates a modification of the usual Migdal gluing rules.

We shall now argue that (2.10) is essentially equivalent to a full q-deformation, and that the analogous combinatorial expansion involving a quantum dimension together with q-deformed Casimir invariants, as in (2.14), yields ordinary two-dimensional Yang-Mills theory. We will also classify the qualitative effects of partial q-deformations of the partition function (2.5). For example, we shall find that incorporating a q-exponential function instead of an ordinary Boltzmann weight in the heat kernel expansion has a very similar effect to using quantum dimensions instead of ordinary dimensions; we suggest a geometric interpretation of the gauge theory involving q-exponentials on a quantum torus.

In §4.4 we give a concrete application of these equivalences to the problem of crystal melting with an external potential [50], whose partition function is the generating function for correlators in a five-dimensional $\mathcal{N}=1$ supersymmetric gauge theory; it can be interpreted in terms of q-deformed two-dimensional BF-theory and certain supersymmetric extensions. We will show that the potentials considered in this model can be regarded as q-deformed Casimir operators, and then apply the analysis of §4.1 which shows that the heat kernel expansion involving quantum dimensions together with a q-deformed Casimir operator is essentially equivalent to ordinary generalized two-dimensional Yang-Mills theory [51, 52] which is defined by the partition function

$$\mathcal{Z}_{\rm M}^{\rm gen}(g_s,t;\Sigma_h)=\sum_\lambda\,(\dim\lambda)^{2-2h}\,\exp\Big(-\frac{g_s}{2}\,\sum_{k=1}^\infty t_k\,C_k(\lambda)\Big)\ , \tag{4.1}$$

where $t=(t_1,t_2,\dots)$ is a set of coupling constants and

$$C_k(\lambda)=\sum_{i=1}^N(\lambda_i-i+1)^k\,\prod_{j\neq i}\Big(1-\frac{1}{\lambda_i-\lambda_j+j-i}\Big)$$

is the k-th Casimir operator eigenvalue in the representation λ. The addition of higher Casimir operators in (4.1) corresponds to the deformation of the BF-theory type action (2.4) by adding the operators $\sum_k\,t_k\,{\rm Tr}\,\big(\phi^k\big)$ in the ring of invariant polynomials $\mathfrak{S}(\mathfrak{g}^*)^G$ on the Lie algebra $\mathfrak{g}$; we suggest a geometric interpretation of the gauge theory involving q-Casimirs on a quantum sphere. We further demonstrate that this gives a new way of relating these gauge theories to unitary matrix models which are large N limits of the unitary matrix models for $U(N)$ Chern-Simons gauge theory that we considered in §3.3.

4.1. **Klimčík deformations.**

While it is possible to work directly with the heat kernel expansions, in this section we shall find it more convenient to use the partition functions of the associated discrete matrix models. The combinatorial series with both quantum dimensions and a q-deformed Casimir invariant leads to a matrix model of the form

$$\mathcal{Z}_{\rm K}(g_s,q;\Sigma_h)=\sum_{n\in\mathbb{Z}^N}\,\prod_{i<j}\,(q^{n_i}-q^{n_j})^{2-2h}\,\exp\Big(-\frac{p\,g_s}{2}\,\sum_{i=1}^N\,[n_i]_q^2\Big)\ , \tag{4.2}$$

which is the type of matrix model that follows from (2.14). In analogy with the solution of the Chern-Simons matrix model [40] in terms of a Stieltjes-Wigert matrix model [38], we consider the change of variables

$$n_i = \frac{\log(m_i+1)}{\log q} \tag{4.3}$$

for $i = 1, \dots, N$ which maps the q-deformed Vandermonde determinant into the standard one

$$\prod_{i<j} \left(q^{n_i} - q^{n_j}\right)^{2-2h} = \prod_{i<j} (m_i - m_j)^{2-2h} \quad ,$$

and also the q-deformed Gaussian potential becomes the standard one

$$\sum_{i=1}^N [n_i]_q^2 = \frac{1}{(1-q)^2} \sum_{i=1}^N m_i^2$$

where here we have used the asymmetric q-number $[n]_q = (1-q^n)/(1-q)$. Then the partition function (4.2) becomes

$$\mathcal{Z}_{\rm K}(\tilde g_s;\Sigma_h) = \sum_{m\in\mathbb{Z}^N_{\geq 0}} \ \prod_{i<j} (m_i - m_j)^{2-2h} \ \exp\Big(-\frac{p\,\tilde g_s}{2}\,\sum_{i=1}^N m_i^2\Big) \ , \tag{4.4}$$

which is essentially the discrete Gaussian matrix model that describes ordinary Yang-Mills theory on the Riemann surface Σ_h with a renormalized coupling constant

$$\tilde g_s = \frac{g_s}{(1-q)^2} \ .$$

Recall from §3 that without q-deformation of the Casimir eigenvalues, the relevant matrix model at genus zero is a discrete Hermitian Stieltjes-Wigert ensemble, which is equivalent to the continuous Stieltjes-Wigert matrix model that describes Chern-Simons gauge theory on S^3. In addition, due to the change of variables (4.3), the range of the eigenvalues in the resulting Gaussian matrix model (4.4) is restricted to the positive weight lattice $\mathbb{Z}^N_{\geq 0}$, similarly to the domain of the weight function for the Stieltjes-Wigert ensemble (3.1). However, the summand of the matrix model is invariant under Weyl reflections of m_i, so that the summations can be extended from the Weyl alcove to $m \in \mathbb{Z}^N$; this situation also arises when the unitary gauge group is replaced by orthogonal or symplectic groups, and the same extension to the whole weight lattice is carried out in that case in [53]. The important feature here is that the resulting potential is not of the Stieltjes-Wigert type $\log^2 x$, which as we have seen corresponds to q-deformed Yang-Mills theory. The latter model is uniquely characterised by the self-similarity property of its discrete orthogonality measure in (3.5), known as the q-Pearson equation [39].

The q-deformed Vandermonde determinant in (4.2) is directly related to the more common form used in (3.8), which is also the form that follows from (2.12), via

$$\prod_{i<j} \left(q^{n_i} - q^{n_j}\right)^{2-2h} = \prod_{i=1}^N q^{(N-1)\,(1-h)\,n_i} \ \prod_{i<j} \Big(2\,\sinh\Big(\frac{g_s}{2}\,(n_i - n_j)\Big)\Big)^{2-2h} \ .$$

Hence depending on the particular q-deformation employed in the dimensions and in the Casimir invariants, the compensation may not be exact, in which case the resulting discrete matrix model will not be exactly Gaussian. However, it will in any case always be well described by a polynomial potential; since any polynomial in the representation weights λ_i can be written as a linear combination of Casimir invariants $C_k(\lambda)$, the matrix model will correspond to a deformation via the addition of higher Casimir operators in ordinary (generalized) two-dimensional Yang-Mills theory (4.1).

We conclude that q-deformation of both the dimensions and the Casimir operators together compensate each other, and the resulting heat kernel expansion is equivalent to its undeformed version, i.e. to that of ordinary (generalized) two-dimensional Yang-Mills theory.

This calculation also illustrates the effect of having only the Casimir eigenvalues q-deformed in (4.2). With $q = \mathrm{e}^{-g_s}$ it leads to a matrix model with an exponential potential

$$\mathscr{Z}_{\rm Cas}(q;\Sigma_h) = \sum_{n\in\mathbb{Z}^N}\ \prod_{i<j}\,(n_i-n_j)^{2-2h}\ \exp\Big(-\frac{p\,\tilde g_s}{2}\,\sum_{i=1}^N\,\big(1-\mathrm{e}^{-g_s\,n_i}\big)^2\Big)\ . \tag{4.5}$$

4.2. **Quantum torus deformations.**

We will now demonstrate that a similar cancellation occurs when we use a q-deformation of the exponential function together with a q-deformed Casimir operator in the heat kernel expansion. The relevant matrix model is defined by the partition function

$$\mathscr{Z}_{\rm Exp}(g_s,q;\Sigma_h) = \sum_{n\in\mathbb{Z}^N}\ \prod_{i<j}\,(n_i-n_j)^{2-2h}\ {\rm Exp}_q\Big(-\frac{p\,g_s}{2}\,\sum_{i=1}^N\,[n_i]_q^2\Big)\ , \tag{4.6}$$

where the q-deformed exponential function is defined by

$${\rm Exp}_q(z) := (-z;q)_\infty = \sum_{n=0}^\infty\,\frac{q^{n\,(n-1)/2}\,z^n}{(q;q)_n}$$

for $z\in\mathbb{C}$. It has an asymptotic expansion given by [54]

$$\begin{aligned}{\rm Exp}_q(x) &= \frac{1}{\big(-q\,x^{-1}\,;\,x\big)_\infty}\,\exp\Big(\tfrac12\,\log x-\frac{1}{\log q}\,\big(\tfrac{\pi^2}{6}+\log^2x\big)-\tfrac{1}{12}\,\log q\Big)\\ &\qquad\times\,\exp\Big(\,\sum_{k=1}^\infty\,\frac{\cos\big(2\pi\,k\log x/\log q\big)}{k\sinh\big(\pi^2\,k/\log q\big)}\Big)\end{aligned}$$

for $x\in\mathbb{R}_{>0}$, from which it follows that the q-exponential series has an asymptotic behaviour ${\rm Exp}_q(x)\sim\mathrm{e}^{-\log^2x}$ for $x\to\infty$. An analogous result holds for the dual q-exponential

$${\rm exp}_q(z) := \frac{1}{(z;q)_\infty} = \sum_{n=0}^\infty\,\frac{z^n}{(q;q)_n} = {\rm Exp}_q(-z)^{-1} \tag{4.7}$$

defined for $|z|<1$; for x varying inside compact subsets of $\mathbb{R}$, both ${\rm Exp}_q\big((1-q)\,x\big)$ and ${\rm exp}_q\big((1-q)\,x\big)$ converge uniformly to $\mathrm{e}^{\,x}$ as $q\to1$. Hence we find that using a q-exponential function for the Boltzmann weight of the theory leads rather directly, without any changes of variables, to a matrix model of the Stieltjes-Wigert type.

This result agrees with the fact that the log-normal distribution (3.1) is equivalent to the weight function [45]

$$w_{\rm Exp}(x;q) = {\rm Exp}_{q^{-1}}(q\,x)\ {\rm exp}_q\big(-(q\,x)^{-1}\big)\ .$$

It is also analogous to the equivalent formulation of the Chern-Simons matrix model as a unitary matrix integral (3.12) whose weight function is the Jacobi elliptic function (3.13). This theta-function is itself a q-exponential series, and by the Jacobi triple product identity (3.14) it can be written in terms of the q-exponentials (4.7) as [54]

$$\begin{aligned}\frac{\Theta(x;q)}{(q;q)_\infty} &= {\rm exp}_q\big(-\sqrt q\,x\big)\ {\rm exp}_q\big(-\sqrt q\,x^{-1}\big)\\ &= \exp\Big(-\frac{1}{\log q}\,\big(\tfrac{\pi^2}{12}+\log^2x\big)+\tfrac{1}{12}\,\log q+\sum_{k=1}^\infty\,(-1)^k\,\frac{\cos\big(2\pi\,k\log x/\log q\big)}{k\sinh\big(\pi^2\,k/\log q\big)}\Big)\end{aligned} \tag{4.8}$$

for $x \in \mathbb{R}_{>0}$. The infinite series in (4.8) is a q-periodic function which has no effect on the Stieltjes-Wigert matrix model, again due to the moment problem [39].

On the other hand, the q-deformation of the Casimir operator in (4.6) essentially consists in using the q-numbers $[n_i]_q$, and hence in the weights $\mathrm{Exp}_q\big([n_i]_q^2\big)$ the q-deformation of the exponential function undoes the q-deformation of its argument, leaving an ordinary Boltzmann weight. Let us look more closely at an explicit example of such an exact cancellation. Consider the q-deformation of the Gaussian distribution given by the orthogonality measure for the continuous q-Hermite polynomials. An explicit form is given by [55]

$$f_q(x) = \frac{2\,q^{1/16}}{\sqrt{\pi\log q^{-1}}}\ \exp\Big(\,\frac{4}{\log q}\ \log^2\big(x+\sqrt{x^2+1}\,\big)\Big) = C(q)\ \exp\Big(\,\frac{4}{\log q}\ \mathrm{arcsinh}^2\,x\Big)$$

for $x\in\mathbb{R}$. If we now substitute the q-variable $[x]_q = \sinh(g_s\,x/2)$ (where as in [7] we use here the symmetric q-number without its proper normalization) we get

$$f_q\big([x]_q\big) = C(q)\ \mathrm{e}^{\,(\log q)\,x^2}\ , \tag{4.9}$$

and therefore the q-deformation of the Gaussian distribution cancels out the q-deformation of its argument.

Let us now consider the effect of using a q-deformed exponential alone instead of a quantum dimension in (2.10). This does not exactly lead to the same situation as above. We already know that (2.12) leads to a Stieltjes-Wigert type ensemble, i.e. a Hermitian Stieltjes-Wigert ensemble for genus $h=0$. On the other hand, using a q-exponential of the theta-function type (4.8) in the heat kernel expansion (2.5) instead of the usual Boltzmann weight leads to a discrete matrix model of the form

$$\mathcal{Z}_\Theta(g_s;\Sigma_h) = \sum_{m\in\mathbb{Z}_{>0}^N}\ \prod_{i<j}\ (m_i-m_j)^{2-2h}\ \exp\Big(-\frac{p\,g_s}{2}\ \log^2\Big(\sum_{i=1}^N m_i^2\Big)\Big)\ , \tag{4.10}$$

which is not of the random matrix theory type because it involves a potential of the form $\log^2(m_1^2+\cdots+m_N^2)$ instead of the Stieltjes-Wigert potential $\log^2(m_1)+\cdots+\log^2(m_N)$. However, the confining properties of the two potentials are qualitatively the same, and in fact the q-exponential matrix model (4.10) can be bounded from below by the Stieltjes-Wigert matrix model (2.12) using the inequality

$$\log^2\Big(\sum_{i=1}^N m_i^2\Big) \geq \frac{4}{N}\ \sum_{i=1}^N\ \log^2(m_i)\ .$$

This inequality follows directly from Jensen's inequality.

A geometric interpretation of the deformation of the gauge theory by q-exponentials could be realised in the following way. The transcendental function

$$\Psi_q(z) = \exp_q\big(-\sqrt{q}\,z\,\big)$$

is called the quantum dilogarithm function (see e.g. [56]); it is related to the classical Euler dilogarithm $\mathrm{Li}_2(x) = \sum_{n\in\mathbb{Z}_{>0}}\frac{x^n}{n^2}$ for $|x|<1$ by the asymptotic expansion

$$\Psi_q(x) = \exp\Big(-\frac{\mathrm{Li}_2(-x)}{\log q}\Big)\big(1+\mathcal{O}(\log q)\big) \qquad \text{for} \quad q\to 1^-\ .$$

Consider the quantum torus algebra $\Omega^0(T_q^2)$, which is the associative noncommutative algebra over $\mathbb{C}$ generated by two operators $\hat u$ and $\hat v$ which satisfy the Weyl algebra

$$\hat u\,\hat v = q\ \hat v\,\hat u\ .$$

It can be represented by q-difference operators on $\Omega^0(\mathbb{C})$ by taking $\hat{u} = z$ to be multiplication by $z \in \mathbb{C}$ and $\hat{v} = \exp\big(-\log(q)\, z\, \frac{\mathrm{d}}{\mathrm{d}z}\big)$. Then the quantum dilogarithm function with operator arguments satisfies the relations [57]

$$\Psi_q(\hat{u}+\hat{v}) = \Psi_q(\hat{v})\,\Psi_q(\hat{u}) \qquad \text{and} \qquad \Psi_q\big(\hat{v}+\sqrt{q}\,\hat{v}\,\hat{u}+\hat{u}\big) = \Psi_q(\hat{u})\,\Psi_q(\hat{v}) \ .$$

These relations suggest that the gauge theory with q-deformed Boltzmann weights may be systematically treated as a gauge theory on the quantum torus T_q^2; the role of this quantum torus algebra will be elucidated within a q-deformed Hamiltonian framework in §6.7. The path integral for gauge theory on the noncommutative torus (for which $q \in S^1$) is defined and studied in [58]; the analog of the Migdal expansion in this context is developed in [59] and related to generalized Yang-Mills theory on the torus $\Sigma_1 = T^2 = S^1 \times S^1$ with infinitely many higher Casimir operators.

4.3. **Other q-deformations.**

Altogether there are six possible quantum deformations of the heat kernel expansion (2.5) of two-dimensional Yang-Mills theory, in addition to the standard one (2.10), incorporating at least one q-deformation of either the dimensions, the Boltzmann weight or the Casimir operator. The different choices can be succinctly summarized in the following table:

Theory	dimensions	exponential	Casimir
q-Yang-Mills*	deformed	deformed	deformed
Yang-Mills	deformed	standard	deformed
Double q-deformation	deformed	deformed	standard
q-Yang-Mills*	standard	deformed	standard
Yang-Mills*	standard	deformed	deformed
Exponential potential	standard	standard	deformed

As we have seen, the replacement of the standard Boltzmann weight with a q-exponential does not lead to a random matrix theory form, although the confining potential is qualitatively similar — we emphasize this with an asterix label on the respective gauge theory in this table; in §4.2 we proposed an interpretation of the matrix model (4.10) at genus one in terms of a noncommutative gauge theory on a quantum deformation of the torus $\Sigma_1 = T^2 = S^1 \times S^1$.

We have already thoroughly discussed the cases leading to two-dimensional Yang-Mills theory (2.5) (and (4.1)) and its q-deformation (2.10); in §5 we will provide a more precise explanation of why the standard q-deformed gauge theory (2.10) is singled out.

There are two cases in this table that are not known to have an interpretation in gauge theory. The first case corresponds to a combinatorial expansion with quantum dimensions together with a q-deformed Boltzmann weight: This yields a rather complicated double q-deformation involving intricate combinations of $\log^2$ functions, which appears difficult to treat analytically. The second case is easier to describe: It corresponds to a pure Casimir q-deformation and leads to the matrix model (4.5); in §4.5 we give a gauge theory interpretation of the matrix model (4.5) at genus zero in terms of noncommutative gauge theory on a quantum deformation of the sphere $\Sigma_0 = S^2$. In §6.7 we shall see how the heat kernel action with q-deformed Casimir operator eigenvalues arises in a Hamiltonian formalism with quantum group gauge symmetry.

4.4. **Five-dimensional gauge theory.**

The melting crystal partition function with an external potential is given by [50]

$$\mathcal{Z}_p(q,t) = \sum_\lambda s_\lambda(q^\rho)^2 \, \exp\big(\Phi(q,t;\lambda,p)\big) \ , \tag{4.11}$$

where the sum is over arbitrary sequences of partitions $\lambda = (\lambda_1, \lambda_2, \dots)$, and $s_\lambda(q^\rho) = \dim_q \lambda$ denotes the Schur function $s_\lambda(x)$, involving an infinite number of variables, with the principal specialization $x_i = q^{i-\frac{1}{2}}$, $i \geq 1$. The potential is a function on charged partitions (λ, p), $p \in \mathbb{Z}+\frac{1}{2}$, depending on a further set of coupling constants $t = (t_1, t_2, \dots)$ and is given by

$$\Phi(q,t;\lambda,p) = \sum_{k=1}^\infty t_k \, \Phi_k(q;\lambda,p)$$

with

$$\Phi_k(q;\lambda,p) = \sum_{i=1}^\infty q^{k\,(p+\lambda_i-i+1)} - \sum_{i=1}^\infty q^{k\,(p-i+1)} + q^k \, \Gamma_{k,p}(q) \ , \tag{4.12}$$

where $\Gamma_{k,p}(z)$ is the equally spaced polynomial (3.11). Note that (4.12) is a polynomial in q. The important feature of these potentials is that they are essentially q-deformed Casimir operator eigenvalues, as we now explain.

The partition function (4.11) has an interpretation in a five-dimensional $\mathcal{N}=1$ supersymmetric gauge theory compactified on a circle of radius $R = g_s$, where $q = \mathrm{e}^{-g_s}$ [60]; in this case the potential (4.12) corresponds to the Wilson line operator $\mathrm{Tr}\,\big(\mathcal{P}\exp \mathrm{i} \oint_{S^1} A\big)^k$ on a loop with winding number k around the compactification circle S^1. It possesses an affine Lie algebra symmetry based on the quantum torus algebra of §4.2 which in this setting is interpreted as providing a realisation of the trigonometric basis for $\mathfrak{sl}(\infty)$ [50]. Its reduction to four dimensions gives the Nekrasov partition function [61] for $\mathcal{N}=2$ noncommutative $U(1)$ gauge theory deformed by higher Casimir operators

$$\mathcal{Z}_p^{4D}(t) = \sum_\lambda (\dim\lambda)^2 \, \exp\Big(\sum_{k=1}^\infty \frac{t_k}{k+1} \, \mathsf{ch}_{k+1}(p,\lambda)\Big) \ ,$$

where the Chern polynomials $\mathsf{ch}_k(p,\lambda) = \mathrm{Tr}\,\big(\varphi_p^k\big)$ can be expressed in terms of Casimir operators of $U(\infty)$ with order $k-1$ and lower; here φ_p is a complex scalar field in the $\mathcal{N}=2$ vector multiplet with vacuum expectation value $\langle\, \mathrm{Tr}\,\varphi_p\rangle = p$ which arises from dimensional reduction of the five-dimensional gauge field A. They can be computed from the generating function [62]

$$\sum_{i=1}^\infty \big(\mathrm{e}^{\,u\,(p+\lambda_i-i+1)} - \mathrm{e}^{\,u\,(p+\lambda_i-i)}\big) = \sum_{k=0}^\infty \mathsf{ch}_k(p,\lambda) \, \frac{u^k}{k!} \tag{4.13}$$

which is the Chern character of the universal sheaf at a fixed point of the instanton moduli space, parametrized by the partition λ, of the four-dimensional supersymmetric gauge theory. For example, for the first two relevant polynomials we find explicitly

$$\mathsf{ch}_2(p,\lambda) = p^2 + 2\,C_1(\lambda) \qquad \text{and} \qquad \mathsf{ch}_3(p,\lambda) = p^3 + 6\,p\,C_1(\lambda) + 3\,C_2(\lambda) \ ,$$

where $C_1(\lambda) = |\lambda|$ is the linear Casimir operator and the shifted symmetric polynomial of second order

$$C_2(\lambda) = \frac{1}{2}\sum_{i=1}^\infty \Big(\big(\lambda_i - i + \tfrac{1}{2}\big)^2 - \big(-i+\tfrac{1}{2}\big)^2\Big) \tag{4.14}$$

is the quadratic Casimir operator. In general, one can also write the Chern characters as

$$\begin{aligned} \mathsf{ch}_k(p,\lambda) &= p^k + \sum_{i=1}^\infty \big((p+\lambda_i-i+1)^k-(p-i+1)^k\big) \\ &\quad - \sum_{i=1}^\infty \big((p+\lambda_i-i)^k-(p-i)^k\big) \end{aligned} \tag{4.15}$$

for any $k\in\mathbb{Z}_{\geq0}$.

From (4.13) and (4.15) it follows that the external potential perturbation (4.12) of the melting crystal model may be regarded as a q-deformation of the higher Casimir $C_{k-1}(\lambda)$, or more precisely of the Chern character $\mathsf{ch}_k(p,\lambda)$. In particular, the partition function (4.11) may be regarded as a deformation of the Klimčík partition function (2.14) by higher order q-Casimirs. As it involves quantum dimensions together with q-deformed Casimir operators, we may treat it along the lines of §4.1. By the same arguments that led to (4.4) from (4.2), we conclude that the melting crystal model (4.11) is equivalent to $U(\infty)$ generalized two-dimensional Yang-Mills theory on the sphere S^2; the infinite rank of the gauge group owes to its origin in a noncommutative gauge theory in four dimensions (see e.g. [63, 64]).

When $t=0$, the expansion (4.11) is also the Donaldson-Thomas partition function of an $\mathcal{N}=2$ noncommutative $U(1)$ gauge theory on $\mathbb{C}^3$ [65, 66]. This cohomological gauge theory is equivalent to the q-deformed $U(\infty)$ BF-theory on S^2 which can be obtained as the $N\to\infty$ limit of (3.20), and hence it has a representation as a $U(\infty)$ unitary matrix model with partition function

$$\mathcal{Z}^{6D}(q) := \mathcal{Z}_p(q,0) = \int_{[0,2\pi)^\infty} \prod_{i=1}^\infty \frac{\mathrm{d}\phi_i}{2\pi}\, \frac{\Theta\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\,q\big)}{(q;q)_\infty}\, \prod_{j<k} \big|\,\mathrm{e}^{\,\mathrm{i}\,\phi_j}-\mathrm{e}^{\,\mathrm{i}\,\phi_k}\big|^2 \ . \tag{4.16}$$

Gauge theories on more general local toric Calabi-Yau threefolds with no compact divisors lead to more complicated matrix models characterized by weight functions which are combinations of theta-functions, as studied in [67, 25, 68] (see [69, 70] for reviews). In these works it is shown how $U(\infty)$ matrix models with combinations of theta-functions as their weight functions possess partition functions which are corresponding combinations of MacMahon functions. For later reference, here we will redo these computations using the strong Szegő limit theorem for Toeplitz determinants (see Appendix B).

Let

$$f(z;q) = \frac{\Theta(z;q)}{(q;q)_\infty} = \prod_{n=1}^\infty \big(1+q^{n-1/2}\,z\big)\,\big(1+q^{n-1/2}\,z^{-1}\big) \tag{4.17}$$

be the weight function of the matrix integral (4.16), where we have used the Jacobi triple product identity (3.14). To apply the Szegő theorem, we need to determine the Fourier coefficients $[\log f]_k$, $k\in\mathbb{Z}$ of

$$\begin{aligned} \log f(z;q) &= \sum_{n=1}^\infty \Big(\log\big(1+q^{n-1/2}\,z\big)+\log\big(1+q^{n-1/2}\,z^{-1}\big)\Big) \\ &= \sum_{n=1}^\infty\,\sum_{k\neq0} \frac{(-1)^{k+1}}{k}\,q^{(n-1/2)\,k}\,z^k \;=\; \sum_{k\neq0} \frac{(-1)^{k+1}\,q^{k/2}}{k\,\big(1-q^k\big)}\,z^k \ . \end{aligned}$$

It follows that

$$[\log f]_k = [\log f]_{-k} = \frac{(-1)^{k+1}\,q^{k/2}}{k\,\big(1-q^k\big)} \ , \tag{4.18}$$

and by the Szegő theorem one has

$$\begin{aligned}\log \mathcal{Z}^{6D}(q) &= \sum_{k=1}^{\infty} k\ [\log f]_k\ [\log f]_{-k} \\ &= \sum_{k=1}^{\infty} \frac{q^k}{k\left(1-q^k\right)^2} \\ &= \sum_{k=1}^{\infty}\sum_{n=1}^{\infty} n\,\frac{q^{k\,n}}{k} \ = \ -\sum_{n=1}^{\infty} n\,\log\left(1-q^n\right)\ .\end{aligned} \tag{4.19}$$

Thus the partition function (4.16) evaluates to

$$\mathcal{Z}^{6D}(q) = M(q)\ ,$$

where $M(q)$ is the MacMahon function

$$M(q) = \prod_{n=1}^{\infty} \frac{1}{\left(1-q^n\right)^n}\ . \tag{4.20}$$

This expression agrees with the $N \to \infty$ limit of (3.20).

We can generalise this calculation to the family of $U(\infty)$ matrix models

$$\mathcal{Z}_L^{6D}(\alpha;q,Q) := \int_{[0,2\pi)^\infty} \prod_{i=1}^{\infty} \frac{\mathrm{d}\phi_i}{2\pi}\ F_L\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\,\alpha\,;\,q,Q\big)\ \prod_{j<k} \big|\,\mathrm{e}^{\,\mathrm{i}\,\phi_j}-\mathrm{e}^{\,\mathrm{i}\,\phi_k}\big|^2$$

with weight functions

$$F_L(z;\alpha;q,Q) := \prod_{a=1}^{L} \frac{\Theta(Q_a\,z;q)^{\alpha_a}}{(q;q)_\infty^{\alpha_a}} \tag{4.21}$$

where $\alpha_a \in \mathbb{C}$ for $a = 1,\dots,L$. With this product of theta-functions as a symbol, the Fourier coefficients (4.18) generalize to

$$[\log F_L]_k = \frac{(-1)^{k+1}\ q^{k/2}}{k\left(1-q^k\right)}\ \sum_{a=1}^{L} \alpha_a\,Q_a^k\ , \tag{4.22}$$

and hence the partition function is given by

$$\log \mathcal{Z}_L^{6D}(\alpha;q,Q) = \sum_{k=1}^{\infty} \frac{q^k}{k\left(1-q^k\right)^2}\ \sum_{a,b=1}^{L} \alpha_a\,\alpha_b\,Q_a^k\,Q_b^{-k}\ .$$

Proceeding analogously as before, it is then straightforward to obtain that this expression evaluates to a product of generalised MacMahon functions

$$\mathcal{Z}_L^{6D}(\alpha;q,Q) = \prod_{a=1}^{L} M(q)^{\alpha_a^2}\ \prod_{b\neq c} M\big(Q_b\,Q_c^{-1},q\big)^{\alpha_b\,\alpha_c}\ , \tag{4.23}$$

where

$$M(Q,q) = \prod_{n=1}^{\infty} \frac{1}{\left(1-Q\,q^n\right)^n}\ .$$

Note the symmetry of the expression (4.23) under the interchange $Q_b \leftrightarrow Q_c$ for any $b,c = 1,\dots,L$, despite the absence of this symmetry in the original weight function (4.21).

For example, when $L = 1$ and $2\alpha_1^2 = \chi$ is the topological Euler characteristic of a Calabi-Yau threefold, the partition function (4.23) computes the constant map contributions to the generating function for Gromov-Witten invariants [71, 72]. The corresponding $U(\infty)$ matrix models

$$\mathcal{Z}_1^{6D}\big(\pm\sqrt{\tfrac{\chi}{2}}\,;\,q\big) = M(q)^{\chi/2} = \int_{[0,2\pi)^\infty} \prod_{i=1}^\infty \frac{\mathrm{d}\phi_i}{2\pi}\,\Big(\frac{\Theta\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\,q\big)}{(q;q)_\infty}\Big)^{\pm\sqrt{\frac{\chi}{2}}}\ \prod_{j<k}\big|\,\mathrm{e}^{\,\mathrm{i}\,\phi_j}-\mathrm{e}^{\,\mathrm{i}\,\phi_k}\big|^2 \tag{4.24}$$

then give an explicit realization of the proposal of [73, 74] to study topological recursion in an associated matrix model.

Since the theta-function (3.13) is holomorphic, the Szegő theorem suffices to determine the asymptotics of the $U(\infty)$ gauge theory partition function (4.16) (see Appendix B). In fact, in our context the Szegő theorem gives exact results. This follows from the fact that for the partition function (4.16) the statement of the theorem is equivalent to the Cauchy-Binet formula (3.18) written in Miwa variables

$$\sum_\lambda s_\lambda(x)\,s_\lambda(y) = \exp\Big(\,\sum_{k\geq1} k\,s_k\,t_k\Big)\ , \tag{4.25}$$

where

$$s_k = \frac1k\,\sum_{i\geq1} x_i^k \qquad\text{and}\qquad t_k = \frac1k\,\sum_{i\geq1} y_i^k$$

are power sums of the sets of variables x and y. In the present case the Miwa variables s and t are precisely the Fourier coefficients (4.18) of the potential of the matrix model (4.16), and in fact the equality (4.16) arises from the normalization of the Schur measure (3.18) via Gessel's identity [25].

To generalise this argument to more complex combinations of theta-functions, we consider the analog of the expression (3.18) involving supersymmetric Schur polynomials $\mathrm{HS}_\lambda(x|z)$ [75] (also known as Schur-Littlewood or hook-Schur polynomials [76]). They are defined by

$$\mathrm{HS}_\lambda(x|z) = \sum_{\mu,\nu} N_{\mu\nu}{}^\lambda\,s_\mu(x)\,s_{\nu'}(z)$$

where $N_{\mu\nu}{}^\lambda\in\mathbb{Z}_{\geq0}$ are the Littlewood-Richardson coefficients defined by expressing the ring structure on the space of symmetric polynomials in the basis of Schur functions as

$$s_\mu(x)\,s_\nu(x) = \sum_\lambda N_{\mu\nu}{}^\lambda\ s_\lambda(x)\ , \tag{4.26}$$

and ν' denotes the conjugate partition to ν. The analogous Cauchy-Binet identity is [77, 76]

$$\sum_\lambda \mathrm{HS}_\lambda(x|z)\ \mathrm{HS}_\lambda(y|w) = \prod_{i,j\geq1}\frac{(1+x_i\,w_j)\,(1+y_i\,z_j)}{(1-x_i\,y_j)\,(1-z_i\,w_j)}\ , \tag{4.27}$$

which we note is symmetric under interchange $(x,y)\leftrightarrow(z,w)$. There is also an extension of the Gessel identity which leads to a unitary matrix model description [77]

$$\begin{aligned}\sum_\lambda \mathrm{HS}_\lambda(x|z)\ \mathrm{HS}_\lambda(y|w) \ &=\ \int_{[0,2\pi)^\infty}\prod_{i=1}^\infty\frac{\mathrm{d}\phi_i}{2\pi}\,\prod_{j\geq1}\frac{\big(1+x_j\ \mathrm{e}^{\,\mathrm{i}\,\phi_i}\big)\,\big(1+y_j\ \mathrm{e}^{-\mathrm{i}\,\phi_i}\big)}{\big(1-z_j\ \mathrm{e}^{\,\mathrm{i}\,\phi_i}\big)\,\big(1-w_j\ \mathrm{e}^{-\mathrm{i}\,\phi_i}\big)}\\ &\qquad\times\prod_{k<l}\big|\,\mathrm{e}^{\,\mathrm{i}\,\phi_k}-\mathrm{e}^{\,\mathrm{i}\,\phi_l}\big|^2\ .\end{aligned} \tag{4.28}$$

After specialisation of the variables $x_i = Q_1\,q^{i-\frac12}$, $y_i = Q_1^{-1}\,q^{i-\frac12}$, $z_i = -Q_2\,q^{i-\frac12}$ and $w_i = -Q_2^{-1}\,q^{i-\frac12}$, this result shows that the formula (4.23) with $K=2$ and $\alpha_1=\pm1=-\alpha_2$ is exact, with

$$\begin{aligned}\mathcal{Z}_2^{6D}(\pm 1, \mp 1; q, Q_1, Q_2) &= \frac{M(q)^2}{M\big(Q_1\, Q_2^{-1}\,,\, q\big)\, M\big(Q_1^{-1}\, Q_2\,,\, q\big)} \\ &= \int_{[0,2\pi)^\infty} \prod_{i=1}^\infty \frac{\mathrm{d}\phi_i}{2\pi}\, \Big(\frac{\Theta\big(Q_1\, \mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\, q\big)}{\Theta\big(Q_2\, \mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\, q\big)}\Big)^{\pm 1} \prod_{j<k} \big|\, \mathrm{e}^{\,\mathrm{i}\,\phi_j} - \mathrm{e}^{\,\mathrm{i}\,\phi_k}\big|^2 \ .\end{aligned}$$

When $Q_1 = 1$, $Q_2 = Q$ this is the partition function of the $\mathcal{N} = 2$ gauge theory on the noncommutative conifold [67, 25]; in the context of five-dimensional gauge theories, the parameter $Q = R^2\,\Lambda^2$ where Λ is the dynamical scale which is identified as the coupling of the four-dimensional $U(1)$ gauge theory obtained in the reduction limit. Moreover, by the generalised Cauchy-Binet formula (4.27) we may identify this partition function as that of a *supersymmetric* extension of the topological q-deformed two-dimensional gauge theory (3.20) given by

$$\mathcal{Z}_2^{6D}(\pm 1, \mp 1; q, Q_1, Q_2) = \sum_\lambda \mathrm{HS}_\lambda\left(Q_1\, q^\rho\,\big|\, - Q_2\, q^\rho\right)\, \mathrm{HS}_\lambda\left(Q_1^{-1}\, q^\rho\,\big|\, - Q_2^{-1}\, q^\rho\right)\ .$$

In the general case, since the weight function (4.21) is of Szegő class and the Toeplitz determinant is of infinite dimension, the formula (4.23) is exact. That the Szegő class condition is satisfied can be easily checked in a number of ways; in our case it suffices to notice the exponential decay of the Fourier coefficients (4.22), see Appendix B. In fact, a closely related exact formula exists even in the finite-dimensional case, which is valid for weight functions which are arbitrary rational functions (Appendix B). In our case we consider theta-functions in (4.21) which involve infinite products, but by truncating the products at some finite order m, making the corresponding weight $F_L^{(m)}$ a rational function, we immediately find a direct relationship between the two cases: The effect of this truncation on the Fourier coefficients in (4.22) is simply

$$\big[\log F_L^{(m)}\big]_k = \big[\log F_L\big]_k\, \big(1 - q^{k\,m}\big)\ ,$$

which follows from the truncation of the corresponding geometric series. The case where each product in (4.21) has a different truncation follows manifestly in the same manner. The behaviour is then of the same type, with the exponential decay of the Fourier coefficients maintained and with a simplification of the final result in the non-rational case $m \to \infty$; this is due to the exponential decay of the coefficients $q^{n-1/2}$ in the products (4.17). Since this simplification is exact for the infinite-dimensional Toeplitz determinant, it would be interesting to study the non-rational case also in the finite-dimensional setting of Appendix B.

4.5. **Torus bundles on the quantum sphere.**

In Klimčík's partition function (2.14), the q-deformed Casimir eigenvalues arise as a result of the group-valued fields which appear in the action of the gauged WZW theory based on the Drinfel'd double $D(G) = G^{\mathbb{C}}$. We shall now propose a natural geometrical interpretation of this q-deformation at genus zero based on noncommutative gauge theory.

Let us first recall the reduction onto torus bundles on S^2 when we evaluate the path integral using the diagonalization techniques from §2. It involves a splitting of the original (trivial) vector bundle associated with the given G-bundle into line bundles $\bigoplus_i \mathcal{L}_{n_i}$ parametrized by sequences of integers $n_1, \dots, n_N$. The corresponding gauge potential is $A = \sum_i a_{n_i}$, where a_{n_i} is the monopole potential of first Chern class $n_i \in \mathbb{Z}$ and the i-th block is an abelian connection on the bundle $\mathcal{L}_{n_i} = \mathcal{L}^{\otimes n_i}$. Here $\mathcal{L} \to S^2$ is the standard $SU(2)$-equivariant monopole line bundle of degree one which is classified by the Hopf fibration $S^3 \to S^2$ and whose isomorphism class is the generator of $H^1(S^2,U(1))\cong H^2(S^2,\mathbb{Z})\cong\mathbb{Z}$. Since $f_n:=\mathrm{d}a_n=2\pi\,n\,\mathrm{d}\mu$, the curvature of this connection is given by

$$F_A=\mathrm{d}A=\sum_{i=1}^N 2\pi\,n_i\,\mathrm{d}\mu$$

where $\mathrm{d}\mu$ is the unit area form on S^2.

We shall now demonstrate how this construction can be extended to the standard Podleś quantum sphere S^2_q [22], via a q-deformation of the Hopf fibration which is the well-known quantum principal $U(1)$-bundle over S^2_q whose total space S^3_q is the manifold of the quantum group $SU_q(2)$ [35]. With $q\in(0,1)$, the algebra $\Omega^0(SU_q(2))$ is the $*$-algebra generated by elements a,c with relations defined by requiring that the matrix

$$U=\begin{pmatrix} a & -q\,c^* \\ c & a^* \end{pmatrix}$$

is unitary: $U\,U^*=U^*\,U=1$. It has the usual Hopf algebra structure defined by the coproduct $\Delta(U)=U\otimes U$, the antipode $S(U)=U^*$, and the counit $\varepsilon(U)=1$. The quantum universal enveloping algebra $\mathcal{U}_q(\mathfrak{su}(2))$, with its standard Hopf $*$-algebra structure described in Appendix A, naturally acts on this algebra: There is a bilinear dual pairing $\langle -,-\rangle_q$ between $\mathcal{U}_q(\mathfrak{su}(2))$ and $\Omega^0(SU_q(2))$ which defines canonical left and right $\mathcal{U}_q(\mathfrak{su}(2))$-module structures on $\Omega^0(SU_q(2))$ such that

$$\langle g,g'\triangleright f\rangle_q:=\langle g\,g',f\rangle_q \qquad \text{and} \qquad \langle g,f\triangleleft g'\,\rangle_q:=\langle g'\,g,f\rangle_q$$

for all $g,g'\in\mathcal{U}_q(\mathfrak{su}(2))$ and $f\in\Omega^0(SU_q(2))$.

Consider now the action of the abelian circle group $U(1)\cong S^1=\{z\in\mathbb{C}\mid |z|=1\}$ on the algebra $\Omega^0(SU_q(2))$ given by the automorphism

$$\alpha_z(a)=z\,a \qquad \text{and} \qquad \alpha_z(c)=z\,c$$

extended as a $*$-algebra map. The algebra $\Omega^0(S^2_q)$ of the standard Podleś sphere S^2_q [22] is the corresponding fixed-point subalgebra

$$\Omega^0(S^2_q)=\Omega^0\big(SU_q(2)\big)^{U(1)}=\big\{f\in\Omega^0\big(SU_q(2)\big)\;\big|\;\alpha_z(f)=f\big\}\ .$$

The algebra inclusion $\Omega^0(S^2_q)\hookrightarrow\Omega^0(SU_q(2))$ is a quantum principal bundle which can be endowed with compatible differential calculi [35].

There are natural finitely-generated projective $\Omega^0(S^2_q)$-bimodules of rank one associated to irreducible one-dimensional representations of $U(1)$ of weight $n\in\mathbb{Z}$ by

$$\mathcal{L}_n=\big\{f\in\Omega^0\big(SU_q(2)\big)\;\big|\;\alpha_z(f)=(z^*)^n\,f\big\}\ ,$$

which we regard as sections of $SU_q(2)$-equivariant line bundles over the quantum sphere S^2_q with monopole charges $n\in\mathbb{Z}$. The left action of the group-like element K on $\Omega^0(SU_q(2))$ gives a dual presentation of these line bundles as

$$\mathcal{L}_n=\big\{f\in\Omega^0\big(SU_q(2)\big)\;\big|\;K\triangleright f=q^{n/2}\,f\big\}\ ;$$

this presentation can be understood via the identification $K=q^{H/2}$ from Appendix A, with H the generator of $U(1)\subset SU_q(2)$.

For the canonical left-covariant two-dimensional calculus on $\Omega^0(S^2_q)$, there are natural gauge potentials $a_n\in\mathsf{Hom}_{\Omega^0(S^2_q)}\big(\mathcal{L}_n,\mathcal{L}_n\otimes_{\Omega^0(S^2_q)}\Omega^1(SU_q(2))\big)$ with curvatures [78, §2.3]

$$f_n=\mathrm{d}a_n=2\pi\,q^{\frac12\,(n+1)}\,[n]_q\ \beta$$

in $\mathsf{Hom}_{\Omega^0(S^2_q)}\big(\mathcal{L}_n,\mathcal{L}_n\otimes_{\Omega^0(S^2_q)}\Omega^2(S^2_q)\big)$, where β is the natural generator for the free $\Omega^0(S^2_q)$-bimodule $\Omega^2(S^2_q)$. Inspired by the reduction of the usual Yang-Mills gauge theory on S^2 to an abelian gauge theory based on torus bundles, we wish to integrate these gauge field curvatures over the quantum sphere S^2_q. This requires the introduction of a "twisted" integral which is a linear functional $\int_{S^2_q} : \Omega^2(S^2_q) \to \mathbb{C}$ defined by restriction of the Haar state on the algebra $\Omega^0(SU_q(2))$ (see e.g. [78, §2.7]); one has

$$\int_{S^2_q} (f \cdot f'\,)\ \beta = \int_{S^2_q} \big((f' \triangleleft K^2) \cdot f\big)\ \beta \tag{4.29}$$

for $f, f' \in \Omega^0(S^2_q)$. This is the unique functional which is invariant under the (quantum adjoint) action of $\mathcal{U}_q(\mathfrak{su}(2))$. Using the normalization $\int_{S^2_q} \beta = 1$ for the Haar state on $\Omega^0(S^2_q)$, one finds that the integral of the curvature f_n of the canonical gauge field a_n on S^2_q is given by

$$\frac{1}{2\pi} \int_{S^2_q} \operatorname{Tr}_q(f_n) = q^{1/2}\, [n]_q\ , \tag{4.30}$$

where $\operatorname{Tr}_q(M) = \operatorname{Tr}(q^{\langle \rho, H\rangle}\, M) \in \Omega^0(S^2_q)$, for an $\Omega^0(S^2_q)$-valued matrix M of dimension $|n| + 1$, is the quantum trace with the "twisted" cyclicity $\operatorname{Tr}_q(M_1\, M_2) = \operatorname{Tr}_q\big((M_2 \triangleleft K^2)\, M_1\big)$, see Appendix A. The q-integer (4.30) has a natural geometric interpretation: It is the q-index of the standard Dirac operator in the Hopf algebraic $SU_q(2)$-equivariant K-theory of S^2_q, i.e. the difference between the quantum dimensions of its kernel and cokernel computed using the quantum trace Tr_q above. For further details, see e.g. [78, §2.7].

It follows that abelianised gauge theory based on the corresponding torus bundles $\bigoplus_i \mathcal{L}_{n_i}$ over S^2_q can reproduce Klimčík's q-deformed heat kernel expansion (2.14) via a putative extension of the diagonalisation technique to this quantum homogeneous space. The problem now is that it is not clear how to even define the classical gauge theory on the quantum sphere, because it is not obvious in what sense gauge theories on S^2_q are actually gauge-invariant for $q \neq 1$: The integral $\int_{S^2_q}$ is a quantum trace and the twisted cyclicity property (4.29) breaks the usual gauge symmetry. This problem is discussed in [79], where path integral quantization for field theories on the quantum two-sphere is considered. It is shown there that the group of gauge transformations $\mathcal{G}$ is generated by a real sector of a quotient of a Hopf algebra, and the space of gauge fields $\mathcal{A}$ is a subspace of one-forms valued in a Hopf module algebra of differential forms on S^2_q; however, the role of these gauge symmetries in the quantum field theory is not clear. On the other hand, it is shown in [78] that gauge transformations act trivially on the line bundles $\mathcal{L}_n$. Hence by *defining* the quantum gauge theory within the approach of abelianization sketched here, we arrive at a concrete gauge-invariant definition of Yang-Mills theory on S^2_q which could cure the problems in the quantization of gauge fields on S^2_q observed in [79] (see also [35]).

5. Categorification

In §4 we argued that the particular q-deformation (2.10) is, at least qualitatively, essentially the only non-trivial quantum deformation of the standard Yang-Mills partition function (2.5). The purpose of this section is to provide a more precise and intrinsic characterization of this argument which will also demonstrate the appearence of a quantum group gauge symmetry. For this, we will use the formalism of two-dimensional topological quantum field theory to construct Yang-Mills amplitudes; for ordinary Yang-Mills theory this construction is described in [4, 1], and in [80, 7, 81] for its q-deformation. In this setting the amplitudes give a representation of a certain geometric category $\mathscr{S}$ in the linear category $\mathscr{R} = \mathscr{V}ect$ of complex vector spaces, whose gluing properties are concisely formulated as a functor of tensor categories. In the following we will instead take $\mathscr{R} = \mathscr{R}ep(\mathcal{U}_q(\mathfrak{g}))$ to be the category of representations of the quantum group $\mathcal{U}_q(\mathfrak{g})$; our description parallels the argument of [20] that quantum characters of $\mathcal{U}_q(\mathfrak{g})$ are required to capture the general solution of q-deformed Yang-Mills theory defined via gluing rules. Then $\mathscr{R}$ has the structure of a semisimple ribbon category endowed with certain additional canonical objects and morphisms. Most notably it contains a "ribbon element", and altogether these structures categorify the basic building blocks of Yang-Mills amplitudes. The corresponding numerical invariants are "generalized characters". By integrating to the standard characters, corresponding to morphisms with target the trivial object $V = \mathbb{C}$ of $\mathscr{R}ep(\mathcal{U}_q(\mathfrak{g}))$, we find that the standard q-propagator (with undeformed Casimir elements) is a fundamental object. More general characters will be considered in §6.

5.1. Semisimple ribbon categories.

We begin by briefly reviewing the relevant category theory that we need to abstract and axiomatize the gluing data of two-dimensional gauge theory into a certain abelian tensor category. For further details, see e.g. [82, 83].

The categories $\mathscr{C}$ that we are interested in are known as *ribbon categories*, together with some additional structure; they are defined as follows. First of all, $\mathscr{C}$ is a monoidal (or tensor) category. This means that $\mathscr{C}$ is equipped with a covariant exterior product bifunctor $\otimes : \mathscr{C} \times \mathscr{C} \to \mathscr{C}$ and a unit object $\mathbb{1} \in \mathsf{Ob}(\mathscr{C})$ together with three natural functorial isomorphisms

$$(X \otimes Y) \otimes Z = X \otimes (Y \otimes Z) \qquad \text{and} \qquad \mathbb{1} \otimes X = X = X \otimes \mathbb{1} \tag{5.1}$$

for all objects $X, Y, Z \in \mathsf{Ob}(\mathscr{C})$, called the associativity and unity relations. Throughout we exploit Mac Lane's coherence theorem (see e.g. [82]) which states that any monoidal category $\mathscr{C}$ is equivalent to a strict monoidal category in which these isomorphisms become equalities; this equivalence typically also preserves additional structures on the categories [84], and hence we take all isomorphisms to be equalities in what follows. The isomorphisms (5.1) satisfy the pentagon relations, which state that the five ways of bracketing the exterior products of four objects commute, and also the triangle relations which state that the associativity constraint with $Y = \mathbb{1}$ is compatible with the unity relations. Given morphisms $f \in \mathsf{Hom}_{\mathscr{C}}(X, Y)$ and $g \in \mathsf{Hom}_{\mathscr{C}}(Z, W)$, their exterior product is $f \otimes g \in \mathsf{Hom}_{\mathscr{C}}(X \otimes Z, Y \otimes W)$.

We call $\mathscr{C}$ braided when there are natural bifunctor isomorphisms

$$B_{X,Y} \in \mathsf{Hom}_{\mathscr{C}}(X \otimes Y, Y \otimes X)$$

for any $X, Y \in \mathsf{Ob}(\mathscr{C})$, called commutativity relations. The braiding $B_{X,Y}$ satisfies the hexagon relations which give two conditions, one expressing $B_{X \otimes Y, Z}$ in terms of associativity relations $\mathrm{id}_X \otimes B_{Y,Z}$ and $B_{Z,X} \otimes \mathrm{id}_Y$, and a similar one for $B_{X, Y \otimes Z}$. In this paper we will only work with categories $\mathscr{C}$ which are $\mathbb{C}$-linear and abelian, which means that the morphism spaces $\mathsf{Hom}_{\mathscr{C}}(X, Y)$ are vector spaces over $\mathbb{C}$ and there is an associative $\mathbb{C}$-bilinear composition product $\mathsf{Hom}_{\mathscr{C}}(X, Y) \times \mathsf{Hom}_{\mathscr{C}}(Y, Z) \to \mathsf{Hom}_{\mathscr{C}}(X, Z)$, denoted $(f, g) \mapsto g \circ f$. For any exact sequence $X \to Y \to Z$ of objects of $\mathscr{C}$ and any morphism $f \in \mathsf{Hom}_{\mathscr{C}}(Y, Z)$, there is a kernel morphism denoted by $\mathrm{Ker}(f) \in \mathsf{Hom}_{\mathscr{C}}(X, Y)$ with $f \circ \mathrm{Ker}(f) = 0$, and similarly there are cokernel morphisms. There are also direct sums $X_1 \oplus \cdots \oplus X_n$ for any finite collection of objects $X_i \in \mathsf{Ob}(\mathscr{C})$. Basic examples are the category $\mathscr{V}ect$ of vector spaces over $\mathbb{C}$ and linear maps, and more generally the representation category $\mathscr{R}ep(\mathcal{A})$ of modules over suitable associative $\mathbb{C}$-algebras $\mathcal{A}$ and intertwining operators, both with the usual tensor product.

For an object $X \in \mathsf{Ob}(\mathscr{C})$, a right dual to X is an object $X^\vee$ with two morphisms

$$e_X \,:\, X^\vee \otimes X \longrightarrow \mathbb{1} \qquad \text{and} \qquad i_X \,:\, \mathbb{1} \longrightarrow X \otimes X^\vee$$

which obey the composition laws

$$(\mathrm{id}_X \otimes e_X) \circ (i_X \otimes \mathrm{id}_X) = \mathrm{id}_X \qquad \text{and} \qquad (e_X \otimes \mathrm{id}_{X^\vee}) \circ (\mathrm{id}_{X^\vee} \otimes i_X) = \mathrm{id}_{X^\vee} \ . \tag{5.2}$$

Similarly, one defines left duals ${}^\vee X \in \mathsf{Ob}(\mathscr{C})$. Dual objects are canonically defined (when they exist), and duality canonically extends to a contravariant functor $(-)^\vee : \mathscr{C} \to \mathscr{C}^{\rm op}$. (Here $\mathscr{C}^{\rm op}$ is the opposite or dual category to $\mathscr{C}$ with $\mathsf{Ob}(\mathscr{C}^{\rm op}) = \mathsf{Ob}(\mathscr{C})$ and $\mathsf{Hom}_{\mathscr{C}^{\rm op}}(X,Y) = \mathsf{Hom}_{\mathscr{C}}(Y,X)$ for all $X,Y \in \mathsf{Ob}(\mathscr{C})$.) If $X,Y,Z \in \mathsf{Ob}(\mathscr{C})$ and Y has dual $Y^\vee$, then there exist canonical isomorphisms [82]

(5.3)

$$\mathsf{Hom}_{\mathscr{C}}(X\otimes Y,Z) = \mathsf{Hom}_{\mathscr{C}}(X,Z\otimes Y^\vee) \qquad \text{and} \qquad \mathsf{Hom}_{\mathscr{C}}(X,Y\otimes Z) = \mathsf{Hom}_{\mathscr{C}}(Y^\vee \otimes X,Z) \ .$$

We henceforth work only with objects that have dual objects.

An object $U \in \mathsf{Ob}(\mathscr{C})$ is simple if any injection of objects $V \hookrightarrow U$ is either 0 or an isomorphism. The category $\mathscr{C}$ is semisimple if any object X is isomorphic to a direct sum

$$X = \bigoplus_{i\in I} n_i\, U_i \ ,$$

where U_i are simple objects, I is the set of isomorphism classes of simple objects in $\mathscr{C}$, and $n_i \in \mathbb{Z}_{\geq 0}$ with $n_i \neq 0$ for only finitely many $i \in I$. We assume that the set I includes the tensor unit as $\mathbb{1} = U_0$. One has $\mathsf{Hom}_{\mathscr{C}}(U_i,U_j) = 0$ for $i \neq j$, whereas $\mathsf{End}_{\mathscr{C}}(U_i) := \mathsf{Hom}_{\mathscr{C}}(U_i,U_i) = \mathbb{C}\ {\rm id}_{U_i}$. Using (5.3) we can define fusion coefficients ${N_{ij}}^k \in \mathbb{Z}_{\geq 0}$ by

$$U_i \otimes U_j = \bigoplus_{k\in I} {N_{ij}}^k\ U_k \qquad \text{with} \quad {N_{ij}}^k = \dim\ \mathsf{Hom}_{\mathscr{C}}(U_i\otimes U_j\otimes U_k^\vee, \mathbb{1}) \ .$$

If the tensor product multiplicity ${N_{ij}}^k \neq 0$, then we may choose a basis $(f_l)_{l=1,\dots,{N_{ij}}^k}$ of $\mathsf{Hom}_{\mathscr{C}}(U_i\otimes U_j, U_k)$ and a dual basis $(f_l^\vee)_{l=1,\dots,{N_{ij}}^k}$ of $\mathsf{Hom}_{\mathscr{C}}(U_k, U_i\otimes U_j)$ such that the composition product $f_l \circ f_m^\vee = 0$ for $l \neq m$ and $f_l \circ f_l^\vee$ is proportional to ${\rm id}_{U_k}$ for each $l,m = 1,\dots,{N_{ij}}^k$.

We also assume the existence of a twist θ, defined to be a system of functorial isomorphisms $\theta_X \in \mathsf{End}_{\mathscr{C}}(X)$ for all $X \in \mathsf{Ob}(\mathscr{C})$ satisfying the compatibility condition

$$\theta_{X\otimes Y} = B_{Y,X} \circ (\theta_Y \otimes \theta_X) \circ B_{X,Y} \ , \tag{5.4}$$

together with a compatible duality; in particular, this additional structure ensures that every left dual of an object is also a right dual and vice versa. On simple objects we define scalars $\theta_i \in \mathbb{C}^*$ by

$$\theta_{U_i} = \theta_{U_i^\vee} = \theta_i\ {\rm id}_{U_i} \ ,$$

with $\theta_0 = 1$.

For any object $X \in \mathsf{Ob}(\mathscr{C})$, and for any endomorphism $f \in \mathsf{End}_{\mathscr{C}}(X)$, we define its categorical or Markov trace $\mathsf{Tr}_X\,(f) \in \mathsf{End}_{\mathbb{C}}(\mathbb{1}) \cong \mathbb{C}$ by

$$\mathsf{Tr}_X\,(f) = e_{X^\vee} \circ \big((\psi_X^{-1} \circ \theta_X) \otimes {\rm id}_{X^\vee}\big) \circ (f \otimes {\rm id}_{X^\vee}) \circ i_X \tag{5.5}$$

where

$$\psi_X = ({\rm id}_X \otimes e_{X^\vee}) \circ \big({\rm id}_X \otimes B^{-1}_{X^{\vee\vee},X^\vee}\big) \circ (i_X \otimes {\rm id}_{X^{\vee\vee}})$$

is a functorial isomorphism in $\mathsf{Hom}_{\mathscr{C}}(X^{\vee\vee},X)$. Taking $f = {\rm id}_X$ defines the categorical dimension of the object X as

$$\mathsf{dim}(X) := \mathsf{Tr}_X({\rm id}_X) = e_{X^\vee} \circ \big((\psi_X^{-1} \circ \theta_X) \otimes {\rm id}_{X^\vee}\big) \circ i_X \ . \tag{5.6}$$

The braiding is taken to be maximally nondegenerate with respect to the semisimple structure, in the sense that

$$S_{ij} := \mathsf{Tr}_{U_i\otimes U_j}\big(B_{U_j,U_i} \circ B_{U_i,U_j}\big) \tag{5.7}$$

for $i,j \in I$ is a symmetric invertible matrix. Comparing (5.7), (5.5) and (5.6) with (5.4) we have

$$S_{i0} = \mathsf{dim}(U_i) \qquad \text{and} \qquad S_{ij} = \frac{1}{\theta_i\,\theta_j} \sum_{k\in I} {N_{ij}}^k\, \theta_k\, \mathsf{dim}(U_k) \ .$$

If $\mathscr{C}$ contains only a finite number of isomorphism classes of simple objects, i.e. I is a finite set, then this structure makes it into a *modular tensor category*.

We choose square roots $\nu_i \in \mathbb{C}^*$ such that $\nu_i^2 = \mathsf{dim}(U_i)$ for $i \in I$. Then the trivalent basis morphisms introduced above may be normalized so that

$$f_l \circ f_m^\vee = \delta_{l,m} \, \frac{\nu_i \, \nu_j}{\nu_k} \, \mathrm{id}_{U_k} \ ,$$

and hence they satisfy

$$\mathrm{id}_{U_i \otimes U_j} = \sum_{k \in I \,:\, N_{ij}{}^k \neq 0} \frac{\nu_k}{\nu_i \, \nu_j} \, f_l^\vee \circ f_l \qquad \text{and} \qquad \mathsf{Tr}_{U_k}\big(f_l \circ f_l^\vee\big) = \nu_i \, \nu_j \, \nu_k$$

for each $l = 1, \dots, N_{ij}{}^k$. This choice of basis will be exploited in our gluing constructions later on.

As we explain in Appendix C, for our purposes we may assume that, as an abelian category, $\mathscr{C}$ arises in the form of representation categories $\mathscr{R} = \mathscr{R}ep(\mathcal{A})$ of associative algebras; however, in general there is no natural monoidal structure. What is needed here is a strengthening of this result proven by Ostrik [85, §4.1]: If $\mathscr{C}$ is a semisimple rigid monoidal category enriched over $\mathscr{V}ect$ with finitely many simple objects, then $\mathscr{C} \cong \mathscr{R}ep(\mathcal{A})$ where $\mathcal{A}$ is a (weak) Hopf algebra. Then the objects $\mathsf{Ob}(\mathscr{R})$ are the representations of the algebra $\mathcal{A}$, while the morphisms $\mathsf{Hom}_{\mathscr{R}}(V, W)$ are intertwiners $f : V \to W$ between $\mathcal{A}$-modules V, W, i.e. equivariant maps $f(a \triangleright v) = a \triangleright f(v)$ for $a \in \mathcal{A}$, $v \in V$. The simple objects U_i, $i \in I$ are the irreducible representations of $\mathcal{A}$, the monoidal structure $\otimes$ is the tensor product of $\mathcal{A}$-modules, and the tensor unit $\mathbb{1}$ is the trivial representation $U_0 \cong \mathbb{C}$. The category $\mathscr{R}$ is then evidently rigid, i.e. left and right duals exist for every object V, which in this case coincide with the usual vector space dual $V^\vee = {}^\vee V = V^*$ regarded as a conjugate representation with $V^{**} \cong V$. The linear map $e_V : V^* \otimes V \to \mathbb{C}$ is the evaluation $e_V(\varphi \otimes_{\mathcal{A}} v) = \varphi(v)$, while $i_V : \mathbb{C} \to V \otimes V^* \cong \mathsf{End}_{\mathcal{A}}(V)$ is the coevaluation given by $i_V(1) = \mathrm{id}_V$. For a generic noncommutative Hopf algebra $\mathcal{A}$, the tensor products $V \otimes W$ and $W \otimes V$ of $\mathcal{A}$-modules need not be isomorphic. However, if $\mathcal{A}$ is a quasitriangular Hopf algebra, then there is also a natural non-trivial isomorphism $V \otimes W \to W \otimes V$ of $\mathcal{A}$-modules, which defines a braiding on the category $\mathscr{R}$. These facts will all be exploited in our ensuing constructions.

The corresponding Grothendieck group $K_0(\mathscr{R})$ is generated by the isomorphism classes $[U_i]$ of simple objects modulo the subgroup generated by the elements $[U] + [W] - [V]$ whenever $0 \to U \to V \to W \to 0$ is an exact sequence in the category $\mathscr{R}$; in particular, with this definition we have $[U \oplus V] = [U] + [V]$. In most instances the Grothendieck group also has a product structure defined by $[U \otimes V] = [U] \cdot [V]$. Then $K_0(\mathscr{R})$ is the representation ring with structure constants $N_{ij}{}^k$ and unit element $[\mathbb{1}]$, which is isomorphic to the commutative ring of characters on $\mathcal{A}$. The tensor category $\mathscr{R}$ may be thought of as a *categorification* of the character ring $K_0(\mathscr{R})$; inequivalent categorifications are typically classified by group cohomology. For example, when $\mathscr{R} = \mathscr{R}ep(SU(N)_k)$ is the modular tensor category of integrable representations of $G = SU(N)$ at level $k \in \mathbb{Z}$, there are N monoidal categories with Grothendieck ring isomorphic to $K_0(\mathscr{R}ep(SU(N)_k))$: the representation category $\mathscr{R}ep(SU(N)_k)$ itself, and certain twists of $\mathscr{R}ep(SU(N)_k)$ induced by the natural $\mathbb{Z}_N$-grading on the fusion algebra of $K_0(\mathscr{R}ep(SU(N)_k))$ [85]. However, in the following we will typically deal with cases where this product is not well-defined on $K_0(\mathscr{R})$.

When $\mathcal{A}$ is additionally a ribbon Hopf algebra, the tensor product and braiding in the representation category $\mathscr{R} = \mathscr{R}ep(\mathcal{A})$ endows $\mathscr{R}$ with the structure of a semisimple ribbon category. The twist coefficients $\theta_V \in \mathsf{End}_{\mathscr{R}}(V)$ provide a "Casimir operator" or "ribbon element", while the braiding operator (5.7) gives a representation of the "modular S-matrix". Together with the fusion coefficients ${N_{ij}}^k$ of the representation ring, these are all the data that we need to fully specify the gauge theory; the gluing constraints are determined by the pentagon and hexagon relations. Thus by abstracting the gluing data into the semisimple ribbon category $\mathscr{R}$, we completely recover the gauge theory.

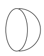
5.2. **Representations of the geometric surface category.**

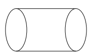
The geometric surface category $\mathscr{S}$ is the tensor category of oriented surfaces with area. Its objects $\mathsf{Ob}(\mathscr{S})$ are collections of disjoint oriented circles S^1, its morphisms $\mathsf{Hom}_{\mathscr{S}}$ are oriented 2-bordisms given by surfaces with area between the circles, i.e. a morphism from an object B_1 to an object B_2 is an oriented surface C with area and boundary $\partial C = B_1 \sqcup (-B_2)$, and the monoidal structure on the category is given by disjoint union of circles. Composition is defined by concatenation of bordisms such that the source and target objects have opposite orientation, and areas are additive. The identity object is the empty one-manifold $\varnothing$; a morphism from $\varnothing$ to itself is given by a compact oriented Riemann surface Σ. The category $\mathscr{S}$ has a duality structure given by the canonical (up to homotopy) orientation reversing diffeomorphism of S^1 and declaring that the cobordisms in (5.2) are equal to the tube bordism defined below.

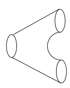
A representation of the category $\mathscr{S}$ is a functor of monoidal categories with duality

$$\mathscr{F}_{\mathcal{A}} \, : \, \mathscr{S} \longrightarrow \mathscr{R}ep(\mathcal{A}) \tag{5.8}$$

for a suitable representation category $\mathscr{R} = \mathscr{R}ep(\mathcal{A})$. We require $\mathscr{F}_{\mathcal{A}}$ to be invariant with respect to boundary and area preserving oriented diffeomorphisms. Because our categories have duality, it suffices to specify the images of the basic cobordisms with one, two and three boundary circles

 $\in \ \mathsf{Hom}_{\mathscr{S}}(S^1, \varnothing)$,

 $\in \ \mathsf{Hom}_{\mathscr{S}}(S^1, S^1)$,

(5.9)  $\in \ \mathsf{Hom}_{\mathscr{S}}(S^1, S^1 \sqcup S^1)$.

Other morphisms (cobordisms) are then obtained by taking compositions (concatenations) and tensor products (disjoint unions) of the morphisms (5.9), and using the duality cobordisms e_{S^1} and i_{S^1}.

In order for (5.8) to be a well-defined functor, we demand that it preserves the unit objects, so that $\mathscr{F}_{\mathcal{A}}(\varnothing) = U_0 \cong \mathbb{C}$, that it takes identity morphisms to identity morphisms, so that $\mathscr{F}_{\mathcal{A}}(\mathrm{id}_B) = \mathrm{id}_{\mathscr{F}_{\mathcal{A}}(B)}$ for all $B \in \mathsf{Ob}(\mathscr{S})$, and that it be compatible with the monoidal structures, so that $\mathscr{F}_{\mathcal{A}}(B_1 \sqcup B_2) = \mathscr{F}_{\mathcal{A}}(B_1) \otimes \mathscr{F}_{\mathcal{A}}(B_2)$ for $B_1, B_2 \in \mathsf{Ob}(\mathscr{S})$, which defines the gluing laws. Since S^1 is an injective cogenerator for the category $\mathscr{S}$ and the collection of simple objects $(U_i)_{i\in I}$ is a family of cogenerators for $\mathscr{R} = \mathscr{R}ep(\mathcal{A})$ (see Appendix C), we define

$$\mathscr{F}_{\mathcal{A}}(S^1) = K = \bigoplus_{i\in I} U_i \tag{5.10}$$

and let

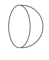
$$\mathscr{F}_{\mathcal{A}}(S^1 \sqcup \cdots \sqcup S^1) = K \otimes \cdots \otimes K \ .$$

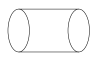
For a disconnected cobordism $C = C_1 \sqcup \cdots \sqcup C_n$, we define

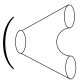
$$\mathscr{F}_{\mathcal{A}}(C) = \mathscr{F}_{\mathcal{A}}(C_1) \otimes \cdots \otimes \mathscr{F}_{\mathcal{A}}(C_n) \ .$$

Using duality in $\mathscr{R}$, we will identity morphisms in $\mathsf{Hom}_{\mathscr{R}}(K\otimes\cdots\otimes K, \mathbb{C})$ with objects $K\otimes\cdots\otimes K$. The representation categories we shall consider are not modular as they contain infinitely many isomorphism classes of simple objects. Thus one should strictly speaking reformulate all of our constructions in a suitable "completion" of $\mathscr{R}ep(\mathcal{A})$, i.e. an ind-category whose objects are infinite direct sums of objects of $\mathscr{R}ep(\mathcal{A})$ and their tensor products. We shall not delve into such technical issues here; see Appendix C for details.

With these identifications made, we define the images of the basic cobordisms (5.9) by their equivalence classes in a suitable completion of the Grothendieck group $K_0(\mathscr{R})$ as

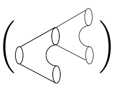
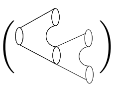
(5.11) $\Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)\Big] = \sum_{i\in I} \mathsf{dim}(U_i)\ \theta_i\ [U_i]\ ,$

(5.12) $\Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)\Big] = \sum_{i\in I} \theta_i\ [U_i\otimes U_i^*]\ ,$

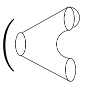
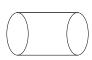
(5.13) $\Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)\Big] = \sum_{i\in I} \frac{\theta_i}{\mathsf{dim}(U_i)}\ [U_i\otimes U_i^*\otimes U_i^*]\ .$

These definitions make sense in any semisimple ribbon category $\mathscr{R}$, see e.g. [86]; in particular, the basic object (5.11) is called the *Gaussian* of the category $\mathscr{R}$. These elements are subjected to the basic gluing axioms of two-dimensional topological quantum field theory, which in the present case can be stated as follows:

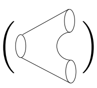
(1) The commutativity constraint implies that (5.13) carries an action of the symmetric group $\mathfrak{S}_3$ which comes from area-preserving diffeomorphisms permuting the three boundaries of the pair of pants bordism.
(2) The associativity constraint implies that

$\Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)\Big] = \Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)\Big]$

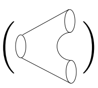
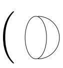
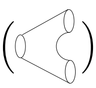
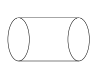
similarly carries an action of $\mathfrak{S}_4$.
(3) The unit constraint comes from capping any of the three boundary circles of the pants bordism and it implies that

$\Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)\Big] = \Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)\Big]\ .$

Here and in the following we suppress the additions of areas in all gluing laws for notational simplicity. It follows from these axioms that

$\Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)^{\otimes(n-2)}\Big]$

carries an $\mathfrak{S}_n$-action for all $n>0$, where we formally define

$\Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)^{\otimes(-1)}\Big] := \Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)\Big]$ and $\Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)^{\otimes(0)}\Big] := \Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)\Big]\ .$

There are two applications of these gluing rules that we are primarily interested in.

First, let us consider the case of a compact oriented Riemann surface Σ_h of genus h, which as a morphism $\varnothing\to\varnothing$ of the category $\mathscr{S}$ can be decomposed into basic cobordims (5.9) as follows:

- For $h=0$ we cut the sphere $\Sigma_0=S^2$ along a circle S^1 into the connected sum of two caps.
- For $h=1$ we cut the torus $\Sigma_1=S^1\times S^1$ along two circles into the connected sum of two tubes.
- For genus $h>1$ we cut Σ_h on $3h-3$ circles into a connected sum of $2h-2$ pants.

We then apply the functor (5.8) using the duality evaluation $e_K : K^* \otimes K \to \mathbb{C}$ and the fact that $\mathsf{Hom}_{\mathscr{R}}(U_i \otimes U_j^*, \mathbb{C}) = \mathsf{Hom}_{\mathscr{R}}(U_i, U_j) \cong \delta_{ij}\ \mathbb{C}$. This results in the quantity $[\mathscr{F}_{\mathcal{A}}(\Sigma_h)] = \mathcal{Z}(\mathcal{A};\Sigma_h)\ [\mathbb{1}]$, where we may generally write

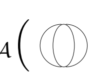
$$\mathcal{Z}(\mathcal{A};\Sigma_h) = \sum_{i\in I} \mathsf{dim}(U_i)^{2-2h}\ \theta_i \tag{5.14}$$

for all $h \geq 0$.

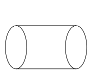
The second class of surfaces we are interested in involves a chain of ℓ Riemann spheres, which arises when we regard a generic lens space $L(p,p'\,) = S^3/\Gamma_{p,p'}$ as a Seifert fibration over the two-sphere (see [87, 11]); here $(p,p'\,)$ are coprime integers with $p > p' > 0$. The base is described by a projective line $\mathbb{P}^1$ with an arbitrarily chosen marked point at which the coordinate neighbourhood is modelled on $\mathbb{C}/\mathbb{Z}_p$, with the cyclic group acting on the local chart coordinate z as $z \mapsto \mathrm{e}^{\,2\pi\,\mathrm{i}/p}\,z$. We construct a line V-bundle over this $\mathbb{P}^1$ orbifold such that the local trivialization over the orbifold point is modelled by $\mathbb{C}^2/\Gamma_{p,p'}$, where $\Gamma_{p,p'} \cong \mathbb{Z}_p$ acts on the local coordinates (z,w) of the base and fibre as $(z,w) \mapsto (\,\mathrm{e}^{\,2\pi\,\mathrm{i}/p}\,z,\ \mathrm{e}^{\,2\pi\,\mathrm{i}\,p'/p}\,w)$. This identifies the lens space $L(p,p'\,)$ as the total space of the associated unit circle bundle. In the base model with a chain of ℓ spheres, the projective lines intersect only once with their nearest neighbours to the right and to the left along the necklace, and the degrees $e_a \geq 2$, $a = 1,\dots,\ell$ are obtained by expanding the rational number $\frac{p}{p'} > 1$ in a simple continued fraction

$$\frac{p}{p'} = [e_1,\dots,e_\ell] := e_1 - \cfrac{1}{e_2 - \cfrac{1}{e_3 - \cfrac{1}{\ddots\ e_{\ell-1} - \cfrac{1}{e_\ell}}}}$$

with e_1 the smallest integer $> \frac{p}{p'}$, and so on. For example, for $p' = 1$ there is only $\ell = 1$ sphere with degree $e_1 = p$; the case $p' = p-1$, where the length of the chain is $\ell = p-1$ and each degree is $e_a = 2$, is considered in [87]. To glue the neighbouring spheres together, we braid the composition of two cap bordisms, corresponding to open disks in the base and fibre directions of the Seifert fibration [87], in $\mathscr{S}$ together using the braiding symmetry (5.7) to define the class

(5.15) $\Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)^{B}\Big] = \sum_{i,j\in I} S_{ij}\ \theta_i\,\theta_j\ [U_i^*\otimes U_j]$

in $K_0(\mathscr{R})$, which carries an action of $\mathfrak{S}_2$. To account for the non-trivial degrees e_a, following [80, 87, 11, 81] we glue two caps in this way at the ends of e_a tubes with classes

$\Big[\mathscr{F}_{\mathcal{A}}\Big($  $\Big)^{\otimes e_a}\Big] \qquad \text{for} \quad a = 1,\dots,\ell\ .$

More precisely, in these constructions one should replace the geometric surface category $\mathscr{S}$ with the category $\mathscr{S}^{L_1,L_2}$ of 2-cobordisms C endowed with line bundles L_1 and L_2 which are trivialized over the boundary components of ∂C, as in [80], but for brevity we do not write this explicitly. In this way we arrive at the partition function

$$\mathcal{Z}(\mathcal{A};S^2,p,p'\,) = \sum_{i_1,\dots,i_\ell\in I} S_{0i_1}\,S_{i_1i_2}\cdots S_{i_{\ell-1}i_\ell}\,S_{i_\ell 0}\ \theta_{i_1}^{e_1}\cdots\theta_{i_\ell}^{e_\ell}\ . \tag{5.16}$$

In particular, for $p' = 1$ we have

$$\mathcal{Z}(\mathcal{A};S^2,p,1) = \sum_{i\in I} \mathsf{dim}(U_i)^2\ \theta_i^p \tag{5.17}$$

which coincides with (5.14) at $p = 1$ and $h = 0$; this calculation easily generalizes to a degree p circle bundle over an arbitrary Riemann surface Σ_h and amounts to replacing the categorical dimension factors $\mathsf{dim}(U_i)^2$ with $\mathsf{dim}(U_i)^{2-2h}$ in the formula (5.17). This formalism could help in explicitly identifying the suitable two-dimensional gauge theory duals to the four-dimensional $\mathcal{N} = 2$ gauge theories on $S^1 \times L(p,p')$ considered in [88, 89]. The construction presented here can also be extended to generic Seifert fibrations over Riemann surfaces Σ_h.

5.3. **Constructing q-deformed Yang-Mills amplitudes.**

We will now specialise the construction of this section to the quantum group $\mathcal{A} = \mathcal{U}_q(\mathfrak{g})$ associated to the Lie algebra $\mathfrak{g}$ of the gauge group $G = U(N)$; then $\mathcal{A}$ has the structure of a quasitriangular Hopf algebra which is described in Appendix A. Let us briefly summarise the structure of the semisimple ribbon category $\mathscr{R} = \mathscr{R}ep(\mathcal{A})$ in this instance, restricting to representations which admit a decomposition into weight spaces; see e.g. [82] for further details.

As a tensor category, $\mathscr{R}$ is equivalent to the category of finite-dimensional representations of $G = U(N)$. In particular, the isomorphism classes of irreducible representations are again labelled by partitions $\lambda = (\lambda_1, \dots, \lambda_N)$, $\lambda_i \geq \lambda_{i+1} \geq 0$; the unit object is the vacuum module corresponding to the empty Young diagram with $\lambda = 0$. The dual of an object V is the dual vector space V^* with the left $\mathcal{A}$-module structure

$$(a \triangleright \varphi)(v) := \varphi\big(S(a) \triangleright v\big)$$

for $a \in \mathcal{A}$, $\varphi \in V^*$ and $v \in V$, where S is the antipode of $\mathcal{U}_q(\mathfrak{g})$. The braiding on the category $\mathscr{R}$ is given by the functorial isomorphisms

$$B_{V,W} = P \circ (R \triangleright) \, : \, V \otimes W \ \longrightarrow \ W \otimes V \ ,$$

where R is the universal R-matrix for $\mathcal{U}_q(\mathfrak{g})$ and P is the trivial "flip" braiding $P(v \otimes w) = w \otimes v$ for $v \in V$, $w \in W$; this yields a non-trivial isomorphism between the tensor product $\mathcal{U}_q(\mathfrak{g})$-modules $V \otimes W$ and $W \otimes V$.

To define the twist, we introduce functorial isomorphisms $q^{\langle \rho, H \rangle} : V \to V^{**}$ which act as multiplication by $q^{\langle \rho, \lambda \rangle}$ on the representation λ, and set

$$\theta = q^{\langle \rho, H \rangle} \ u^{-1} \ ,$$

where $u \in \mathcal{U}_q(\mathfrak{g})$ is Drinfel'd's element. The twist θ is a central element, and it determines the (universal) Casimir operator for the quantum group $\mathcal{U}_q(\mathfrak{g})$: From the explicit formulas in Appendix A, it follows that u^{-1} acts as multiplication by $q^{\frac12 \langle \lambda, \lambda \rangle}$ on the irreducible representation λ and hence

$$\theta_\lambda = q^{\frac12 \langle \lambda, \lambda + 2\rho \rangle} = q^{\frac12 C_2(\lambda)} \ .$$

The functorial isomorphisms $\psi_V : V^{**} \to V$ are given by $\psi_V(x) = u^{-1} \triangleright x$ for $x \in V^{**}$, and hence the categorical trace of any $\mathcal{A}$-module endomorphism $f : V \to V$ is the quantum trace

$$\mathsf{Tr}_V(f) = \operatorname{Tr}_V\big(q^{\langle \rho, H \rangle} \, f\big) \ .$$

In particular, the categorical dimension of a highest weight $\mathcal{U}_q(\mathfrak{g})$-module U_λ coincides with the quantum dimension

$$\mathsf{dim}(U_\lambda) = s_\lambda\big(q^\rho\big) = \dim_q \lambda \ .$$

The partition function (5.14) is therefore given by

$$\mathcal{Z}\big(\mathcal{U}_q(\mathfrak{g})\, ; \, \Sigma_h\big) = \sum_\lambda \, s_\lambda\big(q^\rho\big)^{2-2h} \ q^{\frac12 C_2(\lambda)} \ ,$$

which coincides with (2.10) for an S^1-fibration over Σ_h of degree $p = 1$ when we identify $q = \mathrm{e}^{-g_s}$.

To compute the partition function (5.16), we need the braiding symmetry (5.7) which can be expressed in terms of specializations of Schur functions as

$$S_{\lambda\mu} = s_\lambda\big(q^\rho\big)\ s_\mu\big(q^{\lambda+\rho}\big) = \Big(\prod_{k=1}^{N-1} \frac{1}{\big(q^{-k/2}-q^{k/2}\big)^{N-k}}\Big) \sum_{w\in\mathfrak{S}_N} (-1)^{|w|}\, q^{\langle w(\lambda+\rho),\mu+\rho\rangle}\ .$$

This operator is an analytic continuation of the modular S-matrix of the $U(N)$ WZW model in the Verlinde basis [11]; the second equality here follows from the Weyl character formula and it makes manifest the symmetry $S_{\lambda\mu} = S_{\mu\lambda}$. Note that the fusion coefficients $N_{\lambda\mu}{}^\nu \in \mathbb{Z}_{\geq 0}$ of the category $\mathscr{R}$ are the Littlewood-Richardson coefficients for the fusion of $U(N)$ representations and are represented through Schur functions by (4.26). It follows that

$$\begin{aligned} \mathcal{Z}\big(\mathcal{U}_q(\mathfrak{g})\,;\,S^2,p,p'\big) \ &= \sum_{\lambda_1,\dots,\lambda_\ell} q^{\frac12\,(e_1\,C_2(\lambda_1)+\cdots+e_\ell\,C_2(\lambda_\ell))}\ s_{\lambda_1}\big(q^\rho\big)^2\, s_{\lambda_2}\big(q^\rho\big)\cdots s_{\lambda_\ell}\big(q^\rho\big) \\ &\qquad \times\ s_{\lambda_2}\big(q^{\lambda_1+\rho}\big)\cdots s_{\lambda_\ell}\big(q^{\lambda_{\ell-1}+\rho}\big)\ . \end{aligned}$$

In particular, at $p'=1$ the partition function (5.17) coincides with the partition function (2.10) at $h=0$ for q-deformed Yang-Mills theory on the sphere.

Note that in our case, where q is not a root of unity, the representation category $\mathscr{R}ep(\mathcal{U}_q(\mathfrak{g}))$ is not modular and should be dealt with in the setting of ind-categories as discussed in §5.2. Alternatively, for $q=\zeta_k$ a primitive k-th root of unity, one can work with a "reduced" version of $\mathcal{U}_q(\mathfrak{g})$ defined by imposing the additional relations $E_i^r = 0 = F_i^r$ and $K_i^r = 1$ on the generators, where $r=\frac{k}{2}$ for k even and $r=k$ when k is odd. This is a finite-dimensional Hopf algebra whose ribbon category of finite-dimensional integrable representations has finitely many isomorphism classes of simple objects; however, this category is not semisimple. See [86, 82] for the construction of a modular tensor category of representations of the quantum group $\mathcal{U}_q(\mathfrak{g})$ for q a root of unity.

5.4. **Disk amplitudes.**

From (5.11) it follows that the Gaussian for the category $\mathscr{R} = \mathscr{R}ep(\mathcal{U}_q(\mathfrak{g}))$ is given by

$$\Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{g})}\big(\bigcirc\big)\Big] = \sum_\lambda s_\lambda\big(q^\rho\big)\ q^{\frac12\,C_2(\lambda)}\ [U_\lambda]\ .$$

On the other hand, we can also consider the category of representations of the quantum group $\mathcal{A}=\mathcal{U}_q(\mathfrak{h})$ corresponding to the Cartan subalgebra $\mathfrak{h}\subset\mathfrak{g}$, i.e. the Lie algebra of the maximal abelian subgroup $T=U(1)^N\subset U(N)$. The irreducible representations U_n are now parametrized by the weight lattice $n\in\mathbb{Z}^N$ and are all one-dimensional, so that $\mathsf{dim}(U_n)=1$. The duality $U_n^* = U_{-n}$ and fusion rules

$$U_n\otimes U_m = U_{n+m}$$

furnish the abelian group $\mathbb{Z}^N$. From the explicit expression in Appendix A it follows that the universal R-matrix acts as multiplication by $q^{\frac12\,\langle n,m\rangle}$ on $U_n\otimes U_m$, and hence the braiding is given by

$$B_{U_n,U_m}(v\otimes w) = q^{\frac12\,\langle n,m\rangle}\ w\otimes v\ .$$

The twist eigenvalue of a simple object $U_n\in\mathsf{Ob}(\mathscr{R}ep(\mathcal{U}_q(\mathfrak{h})))$ is

$$\theta_n = q^{\frac12\,\langle n,n\rangle}\ ,$$

while the braiding symmetry is given by

$$S_{nm} = q^{\langle n,m\rangle}\ .$$

Whence the Gaussian (5.11) for this category is given by

$$\Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{h})}\big(\bigcirc\big)\Big] = \sum_{n\in\mathbb{Z}^N} q^{\frac12\,\sum_i\, n_i^2}\,[U_n]\ .$$

In the approach of Etingof and Kirillov [90], the Kostant identity plays a prominent role. It can be stated as the relationship between Gaussians in the Grothendieck rings of the two ribbon categories considered here given by

$$\Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{g})}\big(\bigcirc\big)\Big] = \frac{1}{Z_N(q)}\ \Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{h})}\big(\bigcirc\big)\Big]\ , \tag{5.18}$$

where the normalization $Z_N(q)$ is the Chern-Simons partition function (3.15) on S^3. The identity (5.18) may be regarded as a categorification of the Weyl integral formula (2.13) which relates the Haar measure for integration of G-invariant functions on $G=U(N)$ with the Haar measure for integration of symmetric functions on its maximal torus $T=U(1)^N$, when we identify classes in the Grothendieck rings with characters. It is identical to the result of [91] that the character expansion of the Villain lattice action gives the propagator of q-deformed two-dimensional Yang-Mills theory. The Kostant identity is also derived in [92, App. C] as an expression for the disk amplitude in terms of a theta-function on the weight lattice $\mathbb{Z}^N$, within the framework of the standard gluing rules for q-deformed two-dimensional Yang-Mills theory.

Evaluating both sides of (5.18) at the principal specialization $x=q^\rho$ of the respective characters, we arrive at a simple representation of the partition function (2.10) of q-deformed Yang-Mills theory on S^2 with $p=1$ given by

$$\mathcal{Z}_{\rm M}^{(1)}(q;S^2) = \mathcal{Z}\big(\mathcal{U}_q(\mathfrak{g})\,;\,S^2\big) = \frac{1}{Z_N(q)}\ \prod_{j=1}^N\ \Theta\big(q^{\frac12\,(N+1-2j)}\,;\,q\big)\ , \tag{5.19}$$

where $\Theta(z;q)$ is the Jacobi theta-function (3.13). Using standard modular transformation properties of $\Theta(z;q)$, one shows that (5.19) computes the fractional instanton contributions to the partition function of $\mathcal{N}=4$ gauge theory on the surface $\mathcal{O}(-1)\to\mathbb{P}^1$ (the blow-up of $\mathbb{C}^2$ at a point) [11].

Nevertheless, the two categorifications of the gauge theory are inequivalent: The partition function (5.16) for the category $\mathscr{R}ep(\mathcal{U}_q(\mathfrak{h}))$ is given by the theta-function

$$\mathcal{Z}\big(\mathcal{U}_q(\mathfrak{h})\,;\,S^2,p,p'\,\big) = \sum_{\boldsymbol{n}\in\mathbb{Z}^{\ell\,N}}\ q^{\frac12\,\langle\boldsymbol{n},\boldsymbol{C}\cdot\boldsymbol{n}\rangle}\ , \tag{5.20}$$

where

$$\langle\boldsymbol{n},\boldsymbol{C}\cdot\boldsymbol{n}\rangle = \sum_{a,b=1}^\ell\ C_{ab}\,\langle n^a,n^b\rangle \qquad \text{for} \quad n^a,n^b\in\mathbb{Z}^N\ ,$$

and

$$C=(C_{ab}) = \begin{pmatrix} e_1 & 1 & 0 & \cdots & 0 \\ 1 & e_2 & 1 & \cdots & 0 \\ 0 & 1 & e_3 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & e_\ell \end{pmatrix}$$

is the intersection form of the chain of ℓ base spheres. For $p'=p-1$, so that $\ell=p-1$ and $e_a=2$ for $a=1,\dots,\ell$, the matrix C is essentially the Cartan matrix of the A_{p-1} Dynkin diagram and (5.20) the N-th power of the theta-function on the weight lattice of the Lie algebra $\mathfrak{sl}(p)$.

5.5. Combinatorial Hopf algebra structure.

Our categorical framework points to another way of understanding the Hopf algebraic structure of two-dimensional Yang-Mills amplitudes, at least in the infinite rank limit that was considered in §4.4; we briefly describe this structure now, partly to set the stage for the constructions of §5.6 and §6. For this, we first note that it suffices to work with ordinary representations of the gauge group $G = U(N)$. Let $\mathcal{U}(\mathfrak{g})$ be the ordinary universal enveloping algebra of the Lie algebra $\mathfrak{g}$, and let $\mathcal{U}(\mathfrak{g})[[g_s]]$ be its g_s-adelic completion in the parameter g_s, with $q := \mathrm{e}^{-g_s}$, consisting of formal power series in g_s with coefficients in $\mathcal{U}(\mathfrak{g})$. It is well-known [93, Prop. 3.16] that there is an algebra isomorphism

$$\varphi : \mathcal{U}_q(\mathfrak{g}) \longrightarrow \mathcal{U}(\mathfrak{g})$$

and an invertible twisting cochain $\mathcal{F} = 1 \otimes 1 + \mathcal{O}(g_s)$ which relates the underlying Hopf algebras of $\mathcal{U}_q(\mathfrak{g})$ and $\mathcal{U}(\mathfrak{g})[[g_s]]$; the cochain $\mathcal{F} \in \mathcal{U}(\mathfrak{g})[[g_s]] \otimes \mathcal{U}(\mathfrak{g})[[g_s]]$ induces a non-trivial coassociator $\Phi \in \mathcal{U}(\mathfrak{g})[[g_s]] \otimes \mathcal{U}(\mathfrak{g})[[g_s]] \otimes \mathcal{U}(\mathfrak{g})[[g_s]]$ in this correspondence. At the level of representation categories, this induces a functorial equivalence between $\mathscr{R}ep(\mathcal{U}_q(\mathfrak{g}))$ and Drinfel'd's category of $\mathfrak{g}$-modules with the usual tensor product but with non-trivial associativity isomorphisms induced by Φ [94].

We will further describe our Yang-Mills amplitudes in the representation categories of symmetric groups. For this, we use Frobenius-Schur duality which establishes a bijection between certain representations of $\mathfrak{S}_n$ and of $U(N)$ (see e.g. [1]); this duality will also play a role in §6 when we study the corresponding refined gauge theory amplitudes. Let $V_{\rm fund} = \mathbb{C}^N$ be the fundamental representation of $G = U(N)$. Then the n-th symmetric power $V_{\rm fund}^{\odot n}$ for $n \in \mathbb{Z}_{>0}$ carries, in addition to the G-action, an action of the symmetric group $\mathfrak{S}_n$ by permuting factors. The actions of G and $\mathfrak{S}_n$ commute, so $V_{\rm fund}^{\odot n}$ is a representation of $G \times \mathfrak{S}_n$ which is completely reducible to the form

$$V_{\rm fund}^{\odot n} \cong \bigoplus_\lambda U_\lambda \otimes u_\lambda \ ,$$

where U_λ is the irreducible representation of $U(N)$ corresponding to a partition $\lambda = (\lambda_1, \dots, \lambda_N)$ with $\sum_i \lambda_i = n$, and u_λ is the representation of $\mathfrak{S}_n$ corresponding to the Weyl character $s_\lambda(x)$; this yields a *Frobenius-Schur correspondence* between the representations U_λ and u_λ. A q-deformation of this duality to the quantum group gauge symmetry underlying q-deformed Yang-Mills amplitudes is described in [20], wherein the symmetric group $\mathfrak{S}_n$ is deformed to a Hecke algebra.

This correspondence yields a combinatorial Hopf algebra structure on $U(\infty)$ Yang-Mills amplitudes in the following way. The ring $\mathfrak{Sym}$ of symmetric functions carries a Hopf algebra structure with product map $\mu : \mathfrak{Sym} \otimes \mathfrak{Sym} \to \mathfrak{Sym}$ defined by the Littelwood-Richardson expansion of Schur functions (4.26); its adjoint with respect to the Hall inner product is a coproduct $\Delta : \mathfrak{Sym} \to \mathfrak{Sym} \otimes \mathfrak{Sym}$ making $\mathfrak{Sym}$ into a coalgebra with

$$\Delta(s_\lambda) = \sum_{\mu,\nu} N_{\mu\nu}{}^\lambda \ s_\mu \otimes s_\nu \ .$$

The crucial property to check in this claim is coassociativity: $(\Delta \otimes 1)\,\Delta = (1 \otimes \Delta)\,\Delta$. With the inclusions of symmetric groups $\mathfrak{S}_n \times \mathfrak{S}_m \hookrightarrow \mathfrak{S}_{n+m}$, this is done by interpreting the product μ in $\mathfrak{Sym}$ as induction from representations $\mathfrak{Sym} \otimes \mathfrak{Sym}$ of $\mathfrak{S}_n \times \mathfrak{S}_m$ to $\mathfrak{S}_{n+m}$, and the coproduct Δ as restriction of representations $\mathfrak{Sym}$ of $\mathfrak{S}_{n+m}$ to $\mathfrak{S}_n \times \mathfrak{S}_m$. The coassociativity condition is then equivalent to the commutativity of the induction and restriction morphisms of representations, which follows by Mackey theory. The unit and counit are given respectively by

$$\eta(1) = s_0 \qquad \text{and} \qquad \varepsilon(s_\lambda) = \delta_{\lambda,0} \ ,$$

while the antipode

$$S(s_\lambda) = (-1)^{|\lambda|}\, s_{\lambda'}$$

of the Hopf algebra corresponds to the well-known involution ω of the ring of symmetric functions [28], see e.g. [95, 96]; that the antipode S here is an involution is a consequence of bicommutativity of the Hopf algebra structure on $\mathfrak{Sym}$.

Geissinger [97] shows that the induction and restriction functors associated to these morphisms induce in this way a bialgebra structure on the direct sum of Grothendieck groups of representation categories of $\mathfrak{S}_n$-modules over all $n \geq 0$, which corresponds to the infinite-rank case of the corresponding Yang-Mills amplitudes. These combinatorial Hopf algebras are primarily associated with Grothendieck groups, but they can also be reformulated in terms of the underlying representation categories in a similar vein to the constructions of this section; via Frobenius-Schur duality, the ring $\mathfrak{Sym}$ is thus interpreted as the self-dual Grothendieck Hopf algebra on $\bigoplus_{n\geq 0} \mathbb{C}[\mathfrak{S}_n]$, where $\mathbb{C}[\mathfrak{S}_n]$ is the group ring of the symmetric group $\mathfrak{S}_n$. In particular, Bump and Gamburd [76] show that the coassociativity of this Hopf algebra is equivalent to the generalized Cauchy-Binet formula (4.27) for the supersymmetric Schur polynomials, and hence the Hopf algebraic structure is intimately tied to the underlying supersymmetric structure of $U(\infty)$ Yang-Mills amplitudes discussed in §4.4. Of course, by virtue of the identification of Schur functions as characters of irreducible representations of $U(\infty)$, this structure makes the Grothendieck group $K_0(\mathscr{R}ep(U(\infty)))$ into a graded self-dual, bicommutative Hopf algebra (see e.g. [96]), and ultimately also the representation category $\mathscr{R}ep(\mathcal{U}_q(\mathfrak{g}))$ itself in the $N \to \infty$ limit by replacing $s_\lambda \mapsto U_\lambda$ in the structure maps above. This Hopf algebra structure is exploited in §5.6 below.

5.6. **Defect operators and module categories.**

For completeness, and also in preparation for our discussion of refinement in §6, we conclude this section by describing how the computations of correlators in q-deformed two-dimensional Yang-Mills theory fit into our categorical framework. Correlation functions are associated with insertions of defect operators in partition functions; generalized defect observables are constructed via fundamental boundary observables in representations of the geometric surface category $\mathscr{S}$. For each such boundary observable, we choose a basepoint on each boundary and associated a simple object $U_i \in \mathsf{Ob}(\mathscr{C})$, $i \in I$, to that boundary. There are two perspectives one can take in constructing defect observables: one through the gluing rules of this section, and one through considerations of the Frobenius algebra (5.10) of isomorphism classes of simple objects of the category $\mathscr{C}$. We will begin here with the former perspective which can be used to immediately write down explicit formulas for the correlators, as it is this point of view that will be taken in §6. There are three classes of defect operators which are of relevance for the construction of generic two-dimensional Yang-Mills amplitudes.

Firstly, there is the extension of the representations considered in §5.2 to Riemann surfaces with boundaries, which as morphisms in $\mathscr{S}$ live in $\mathsf{Hom}_{\mathscr{S}}(\varnothing, S^1 \sqcup \cdots \sqcup S^1)$. They are easily constructed from the basic amplitudes with one, two and three boundaries given in (5.11)–(5.13) respectively. For a Riemann surface of genus h having b punctures with boundary conditions fixed to the holonomy eigenvalues $u_1, \dots, u_b \in (S^1)^N$ of the gauge connection around the boundary circles, by evaluation of the corresponding characters at these holonomies we find that the general partition function for the usual q-deformed gauge theory is given by

$$\mathcal{B}_{u_1,\dots,u_b}\big(\mathcal{U}_q(\mathfrak{g})\,;\,\Sigma_h, p, 1\big) = \sum_\lambda s_\lambda\big(q^\rho\big)^{2-2h-b}\ s_\lambda(u_1)\cdots s_\lambda(u_b)\ q^{\frac{p}{2}\,C_2(\lambda)}\ .$$

These amplitudes naturally reflect the monoidal structure of the representation category $\mathscr{R}$, and also the Hopf algebra structure from §5.5. These boundary observables are considered in [20, 98] and related to the four-dimensional $\mathcal{N}=2$ superconformal index in [14].

Secondly, we can consider more general closed defect observables which correspond to closed non-selfintersecting loops on the surface Σ_h. They are given by Wilson loops on Σ_h equipped with a choice of marked point on the loop which corresponds to an insertion of a defect operator in the partition function; Wilson loop observables in q-deformed Yang-Mills theory are considered in [5, 20]. Consider, for example, a correlation function of a single Wilson loop operator $\mathrm{Tr}_\lambda\big(\mathcal{P}\exp \mathrm{i} \oint_C A\big)$ in the representation λ on a non-selfintersecting oriented closed curve $C\subset\Sigma_h$ which divides the Riemann surface Σ_h into inner and outer faces of genera h_1 and h_2, with $h=h_1+h_2$. This correlator is determined in terms of the fusion coefficients $N_{\lambda_1\lambda}{}^{\lambda_2}$ of the representation category $\mathscr{R}$ to be

$$\mathcal{W}_\lambda\big(\mathcal{U}_q(\mathfrak{g})\,;\,\Sigma_h,p,1\big)=\sum_{\lambda_1,\lambda_2} N_{\lambda_1\lambda}{}^{\lambda_2}\ s_{\lambda_1}\big(q^\rho\big)^{1-2h_1}\, s_{\lambda_2}\big(q^\rho\big)^{1-2h_2}\ q^{\frac p2\,C_2(\lambda_1)}\,q^{\frac p2\,C_2(\lambda_2)}\ ,$$

where the representations λ_1 and λ_2 label the inner and outer faces respectively.

Thirdly, we can consider Wilson loop observables in Chern-Simons theory on Seifert manifolds which wrap around the S^1 fibre; they are studied in [99]. In two-dimensional q-deformed Yang-Mills theory they correspond to defect holonomy punctures given by correlators of the gauge-invariant operators $\mathrm{Tr}_\lambda\exp(\mathrm{i}\,\phi)$ inserted on the base Riemann surface Σ_h, which represent the holonomy of the Chern-Simons gauge connection around the S^1 fibre; these defect punctures were considered in [7, 92] and shown in [98] to correspond to insertions of supersymmetric surface operators in the four-dimensional $\mathcal{N}=2$ superconformal index. The amplitude for n defect punctures in representations $\lambda_1,\dots,\lambda_n$ on Σ_h can also be written entirely in terms of data in the underlying semisimple ribbon category $\mathscr{R}$, and one has

$$\begin{aligned}\mathcal{P}_{\lambda_1,\dots,\lambda_n}\big(\mathcal{U}_q(\mathfrak{g})\,;\,\Sigma_h,p,1\big) &= \sum_\lambda (\dim_q\lambda)^{2-2h-n}\ S_{\lambda\lambda_1}\cdots S_{\lambda\lambda_n}\ q^{\frac p2\,C_2(\lambda)}\\ &= \sum_\lambda s_\lambda\big(q^\rho\big)^{2-2h}\ s_{\lambda_1}\big(q^{\lambda+\rho}\big)\cdots s_{\lambda_n}\big(q^{\lambda+\rho}\big)\ q^{\frac p2\,C_2(\lambda)}\\ &= \sum_\lambda s_\lambda\big(q^\rho\big)^{2-2h}\ q^{\frac p2\,C_2(\lambda)}\\ &\qquad\times\sum_{\mu_1,\dots,\mu_{n-1}} N_{\lambda_1\lambda_2}{}^{\mu_1}\,N_{\mu_1\lambda_3}{}^{\mu_2}\cdots N_{\mu_{n-2}\lambda_n}{}^{\mu_{n-1}}\ s_{\mu_{n-1}}\big(q^{\lambda+\rho}\big)\end{aligned}\tag{5.21}$$

where we used (4.26) in the last line.

The independence of correlators such as (5.21) on the insertion points $x_i\in\Sigma_h$ of the operators $\mathrm{Tr}_{\lambda_i}\exp(\mathrm{i}\,\phi(x_i))$ can be understood by appealing to an alternative description of these formulas in terms of a functor from the geometric surface category $\mathscr{S}$ to a module category of Frobenius algebras. Defect punctures correspond to left A-modules and non-selfintersecting Wilson line defects to A-bimodules over a suitable Frobenius algebra object A in the semisimple ribbon category $\mathscr{C}$ (see e.g. [83]); for the representation category $\mathscr{C}=\mathscr{R}=\mathscr{R}ep(\mathcal{A})$ this is the Frobenius algebra $A=K$ of conjugacy classes of $\mathcal{A}=\mathcal{U}_q(\mathfrak{g})$ in (5.10). In the remainder of this section we briefly explain how the construction of correlators of defect operators fits into our categorical framework in this way.

Again we begin by sketching the relevant structures involved from category theory. Recall that a Frobenius algebra A is an associative, unital algebra over $\mathbb{C}$ equipped with a linear function $\varepsilon: A\to\mathbb{C}$ such that the bilinear pairing defined by $(a,b):=\varepsilon(a\,b)$ for $a,b\in A$ is nondegenerate. The basic example is the algebra $A=\mathbb{M}_n$ of $n\times n$ matrices over $\mathbb{C}$, with $\varepsilon(a)=\mathrm{Tr}\,(a)$. More generally, for any Frobenius algebra (A,ε) we can enrich the algebra $\mathbb{M}_n(A) := \mathbb{M}_n \otimes A$ with the Frobenius structure $\varepsilon \circ \mathrm{Tr} : \mathbb{M}_n(A) \to \mathbb{C}$. By Wedderburn's theorem, a finite-dimensional simple algebra is isomorphic to a matrix algebra over a division ring, and hence every finite-dimensional semisimple algebra admits a Frobenius structure. Another example is provided by the complex de Rham cohomology ring $A = H^*(X)$ of a compact n-dimensional complex manifold X. This is an algebra under the wedge product of differential forms on X. Then the integration $\int_X : H^*(X) \to \mathbb{C}$ over X provides a Frobenius structure ε on A. This example illustrates a geometric way in which to think of these algebras. A semisimple Frobenius algebra A is always the algebra of $\mathbb{C}$-valued functions on the set $X = \mathrm{Spec}(A)$ of minimal ideals of A, equipped with a "volume form" ε which assigns a measure ε_x to each point $x \in X$. The Frobenius structure ε thus provides an "integration" (or trace) over the "space" X. When A is finite-dimensional, X is just a finite set of points. Frobenius algebras play a well-known role in two-dimensional topological field theory, see e.g. [100] for an introduction.

This notion extends to the monoidal categories $\mathscr{C}$ we are interested in. An associative, unital algebra in $\mathscr{C}$ is an object $A \in \mathsf{Ob}(\mathscr{C})$ together with a "multiplication" $\mu \in \mathsf{Hom}_{\mathscr{C}}(A \otimes A, A)$ and a "unit" $\eta \in \mathsf{Hom}_{\mathscr{C}}(\mathbb{1}, A)$ satisfying the associativity condition

$$\mu \circ (\mu \otimes \mathrm{id}_A) = \mu \circ (\mathrm{id}_A \otimes \mu)$$

and the unit condition

$$\mu \circ (\eta \otimes \mathrm{id}_A) = \mathrm{id}_A = \mu \circ (\mathrm{id}_A \otimes \eta) \ .$$

An algebra in $\mathscr{C} = \mathscr{V}ect$, with the usual tensor product of vector spaces, is precisely an associative $\mathbb{C}$-algebra. An algebra in the dual category $\mathscr{C}^{\mathrm{op}} = \mathscr{V}ect^{\mathrm{op}}$, with the usual tensor product of vector spaces, is precisely a coassociative $\mathbb{C}$-coalgebra.

Suppose that A is an algebra in $\mathscr{C}$. Then A is said to be a Frobenius algebra in the monoidal category $\mathscr{C}$ if there exists morphisms $\varepsilon \in \mathsf{Hom}_{\mathscr{C}}(A, \mathbb{1})$ and $\Delta \in \mathsf{Hom}_{\mathscr{C}}(A, A \otimes A)$ such that (A, ε, Δ) is a coalgebra and

$$(\mathrm{id}_A \otimes \mu) \circ (\Delta \otimes \mathrm{id}_A) = \Delta \circ \mu = (\mu \otimes \mathrm{id}_A) \circ (\mathrm{id}_A \otimes \Delta) \ .$$

This means that $\Delta : A \to A \otimes A$ is a morphism of A-bimodules. If in addition $\mathscr{C}$ is braided, then A is commutative when

$$\mu \circ B_{A,A} = \mu$$

on $A \otimes A \to A$. This is equivalent to the cocommutativity condition

$$B_{A,A} \circ \Delta = \Delta$$

on $A \to A \otimes A$. We say that A is haploid if $\dim \mathsf{Hom}_{\mathscr{C}}(\mathbb{1}, A) = 1$.

There are two other conditions that we will need below. In the category $\mathscr{C}$, there are two canonical coevaluations

$$d_A \in \mathsf{Hom}_{\mathscr{C}}(\mathbb{1}, A \otimes A^\vee) \qquad \text{and} \qquad \tilde{d}_A \in \mathsf{Hom}_{\mathscr{C}}(\mathbb{1}, A^\vee \otimes A) \ .$$

Then A is symmetric if the natural isomorphisms of A-bimodules

$$\Phi_1 := \big((\varepsilon \circ \mu) \otimes \mathrm{id}_{A^\vee}\big) \circ (\mathrm{id}_A \otimes d_A) \qquad \text{and} \qquad \Phi_2 := \big(\mathrm{id}_{A^\vee} \otimes (\varepsilon \circ \mu)\big) \circ (\tilde{d}_A \otimes \mathrm{id}_A)$$

in $\mathsf{Hom}_{\mathscr{C}}(A, A^\vee)$ coincide, $\Phi := \Phi_1 = \Phi_2$. Haploid algebras are symmetric. The algebra object A is special if Δ is a right inverse of μ and $\varepsilon \circ \eta = \mathsf{dim}(A)\ \mathrm{id}_{\mathbb{1}}$, where $\mathsf{dim}(A)$ is the categorical dimension of $A \in \mathsf{Ob}(\mathscr{C})$. This generalizes the notion of separable algebras over $\mathbb{C}$.

For the representation category $\mathscr{C} = \mathscr{R} = \mathscr{R}ep(\mathcal{U}_q(\mathfrak{g}))$, we take the cogenerator $A = K$ from (5.10) and induce a Frobenius structure on it from the combinatorial Hopf algebra structure of

§5.5. Using $\mathsf{Hom}_{\mathscr{R}}(U_\lambda, A) \cong \mathbb{C}$, the multiplication $\mu : A \otimes A \to A$ is provided by the image of the pants bordism (5.13) regarded as a morphism in the dual category $\mathscr{R}^{\rm op}$; explicitly

$$\mu(u_\mu \otimes u_\nu) = \sum_\lambda N_{\mu\nu}{}^\lambda \, {\rm id}_{U_\lambda}(u_\mu \otimes u_\nu) \tag{5.22}$$

for all simple subobjects U_μ, U_ν, U_λ of A and all vectors $u_\mu \in U_\mu$, $u_\nu \in U_\nu$. Let us explain the meaning of this formula. Let $e_\mu \in \mathsf{Hom}_{\mathscr{R}}(U_\mu, A)$ be the basis of canonical inclusions of simple subobjects, and let $e^\mu \in \mathsf{Hom}_{\mathscr{R}}(A, U_\mu)$ be the dual basis of projections. Then the composition $e^\lambda \circ \mu \circ (e_\mu \otimes e_\nu)$ is an element of $\mathsf{Hom}_{\mathscr{R}}(U_\mu \otimes U_\nu, U_\lambda) \cong N_{\mu\nu}{}^\lambda \, \mathbb{C} \, {\rm id}_{U_\lambda}$. The associativity of the product (5.22) follows easily from the associativity relations for the fusion coefficients $N_{\mu\nu}{}^\lambda$; this multiplication makes A into a (braided) noncommutative algebra. The comultiplication is given by

$$\Delta(u_\lambda) = \sum_{\mu,\nu} N_{\mu\nu}{}^\lambda \, {\rm id}_{U_\mu \otimes U_\nu}(u_\lambda) \tag{5.23}$$

where $u_\lambda \in U_\lambda$. This endows A with the structure of a noncocommutative coalgebra. The Frobenius counit $\varepsilon : A \to \mathbb{C}$ is provided by the image of the cap bordism (5.11); it is dual to the unit η of A which is just the vacuum representation. Nondegeneracy of the inner product $\varepsilon \circ \mu$ follows from the gluing laws, which equate it to the image of the tube bordism (5.12). The Frobenius algebra object A obtained in this way is symmetric but not special. Choosing different defects corresponds to selecting A-modules $\mathtt{M}$; the original defect operator corresponds to the trivial A-module $\mathtt{A}$ acting on itself via its multiplication morphism μ. We shall now study the Morita equivalence class of the Frobenius algebra A, i.e. the A-module structures provided by all other boundary defect operators.

Let A be an algebra in the semisimple monoidal category $\mathscr{C}$. A (left) A-module is a pair $\mathtt{M} = (M, \varrho)$, where $M \in \mathsf{Ob}(\mathscr{C})$ and $\varrho \in \mathsf{Hom}_{\mathscr{C}}(A \otimes M, M)$ with the relations

$$\varrho \circ (\mu \otimes {\rm id}_M) = \varrho \circ ({\rm id}_A \otimes \varrho) \qquad \text{and} \qquad \varrho \circ (\eta \otimes {\rm id}_M) = {\rm id}_M \ .$$

The space of morphisms between two left A-modules $\mathtt{M}_1$ and $\mathtt{M}_2$ forms a $\mathbb{C}$-linear subspace of $\mathsf{Hom}_{\mathscr{C}}(M_1, M_2)$ denoted $\mathsf{Hom}_A(\mathtt{M}_1, \mathtt{M}_2)$. Let $\mathscr{M}od_{\mathscr{C}}(A)$ be the category of left modules over the Frobenius algebra A in the ribbon category $\mathscr{C}$; its objects are the associated boundary defect fields. When A is a special Frobenius algebra, then semisimplicity of $\mathscr{C}$ implies that the category $\mathscr{M}od_{\mathscr{C}}(A)$ is semisimple [101, Prop. 5.24]. Similarly one defines A-bimodules (see e.g. [83, Def. 4.5]), but it suffices to consider left A-modules by [85, Rem. 12]: In any braided tensor category $\mathscr{C}$, an A-bimodule can equivalently be regarded as a left $A \otimes A^{\rm op}$-module. This means that it suffices to focus our attention to insertions of boundary defects on the Riemann surface Σ_h.

The category $\mathscr{M}od_{\mathscr{C}}(A)$ is not a monoidal category, but it carries the structure of a module category. For this, let $\mathtt{M} = (M, \varrho)$ be any left A-module with object $M \in \mathsf{Ob}(\mathscr{C})$ and morphism $\varrho \in \mathsf{Hom}_{\mathscr{C}}(A \otimes M, M)$, and let $X \in \mathsf{Ob}(\mathscr{C})$ be any object of $\mathscr{C}$. Then $\mathtt{M} \otimes X := (M \otimes X, \varrho \otimes {\rm id}_X)$ has the natural structure of a left A-module. For any $X, Y \in \mathsf{Ob}(\mathscr{C})$, the associativity isomorphism

$$M \otimes (X \otimes Y) \ \xrightarrow{\ \approx\ } \ (M \otimes X) \otimes Y$$

yields a morphism of A-modules

$$\mathtt{M} \otimes (X \otimes Y) \ \longrightarrow \ (\mathtt{M} \otimes X) \otimes Y$$

in $\mathsf{Hom}_A(\mathtt{M} \otimes (X \otimes Y), (\mathtt{M} \otimes X) \otimes Y)$. This endows $\mathscr{M}od_{\mathscr{C}}(A)$ with the structure of a module category over $\mathscr{C}$, i.e. the "mixed" tensor functor

$$\otimes \ : \ \mathscr{M}od_{\mathscr{C}}(A) \times \mathscr{C} \ \longrightarrow \ \mathscr{M}od_{\mathscr{C}}(A)$$

is an exact bifunctor with associativity and unit conditions generalizing the triangle and pentagon axioms.

For a given defect operator, specified by a fixed non-zero A-module $\mathtt{M} \in \mathsf{Ob}(\mathscr{M}od_{\mathscr{C}}(A))$, we demonstrate in Appendix C how to construct a canonically defined algebra object $A_M = {}^\vee\mathtt{M} \otimes_A \mathtt{M}$ of $\mathscr{C}$ such that $A_A = A$; given two defect operators $\mathtt{M}_1, \mathtt{M}_2 \in \mathsf{Ob}(\mathscr{M}od_{\mathscr{C}}(A))$, we also prove that the module categories of A_{M_1} and A_{M_2} are equivalent. Thus starting from a single defect operator we get a (symmetric) Frobenius algebra A. Different boundary defects generically produce distinct Frobenius algebras, but any two such Frobenius algebras are Morita equivalent. Morita equivalent Frobenius algebras give rise to equivalent correlation functions in the two-dimensional gauge theory. For example, this explains the feature that the correlators of defect holonomy operators are independent of their insertion points on Σ_h. In Appendix C we give an explicit description of the module category $\mathscr{M}od_{\mathscr{C}}(A)$ by constructing a $\mathbb{C}$-algebra $\mathcal{A}$ such that $\mathscr{R}ep(\mathcal{A}) = \mathscr{M}od_{\mathscr{C}}(A)$; this functorial equivalence endows the representation category $\mathscr{R}ep(\mathcal{A})$ with the structure of a module category

$$\otimes \, : \, \mathscr{R}ep(\mathcal{A}) \times \mathscr{C} \longrightarrow \mathscr{R}ep(\mathcal{A})$$

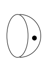
over the tensor category $\mathscr{C}$, i.e. for $\mathtt{V} \in \mathscr{R}ep(\mathcal{A})$, there are natural isomorphisms $(\mathtt{V} \otimes X) \otimes Y \cong \mathtt{V} \otimes (X \otimes Y)$ for all $X, Y \in \mathsf{Ob}(\mathscr{C})$ and $\mathtt{V} \otimes U_0 \cong \mathtt{V}$.

6. Refinement

An immediate spinoff from the categorical reformulation of §5 is that one can also integrate to generalized characters associated to morphisms with target an arbitrary representation V of G. In particular, when V is a symmetric power of the fundamental representation of $G = U(N)$, we unleash a *refinement* of the q-deformed Yang-Mills amplitudes which leads to a two-parameter deformation of the heat kernel expansion (2.5). This refined q-deformed Yang-Mills theory was considered recently in [19], and it is an analytic continuation, in the sense explained in §2.2, of the refinement of Chern-Simons theory on Seifert three-manifolds constructed in [24, 23] which computes the Poincaré polynomials of knot homology (see also [102]). The topological version of this gauge theory is the refined q-deformed BF-theory on the Riemann surface Σ_h considered in [18], or equivalently refined Chern-Simons theory on $\Sigma_h \times S^1$, which is identified as the two-dimensional topological field theory computing the topologically twisted partition function of an $\mathcal{N} = 2$ superconformal field theory on $S^1 \times S^3$, i.e. the $\mathcal{N} = 2$ superconformal index in four dimensions; the non-topological version identifies the partition function of an $\mathcal{N} = 2$ supersymmetric non-linear sigma-model on $S^1 \times S^3$ with the propagator of the refined two-dimensional gauge theory [15].

6.1. Constructing refined q-deformed Yang-Mills amplitudes.

To construct refinements of the q-deformed two-dimensional gauge theory, we need to generalize our categorification slightly, as alluded to in §5.6. For this, we enlarge the morphisms of the source category to include two-dimensional surfaces with marked points. We then modify the functor (5.8) by prescribing additional data at each marked point given by a fixed finite-dimensional module V over the quantum group $\mathcal{U}_q(\mathfrak{g})$, which may be interpreted as the insertion of a defect holonomy puncture in the representation V of $G = U(N)$ at each marked point, i.e. the holonomy of the gauge fields around the marked point is the representation V; the corresponding Yang-Mills amplitude defines a wavefunction in the Hilbert space associated to the boundary. Thus, for example, the basic class (5.11) is correspondingly modified in this case to

$$\Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{g})}\Big(\;$$  $$\;\Big)\Big] = \sum_\lambda \mathsf{dim}(U_\lambda)\ \theta_\lambda\ [U_\lambda \otimes V]\ .$$

We interpret this refinement as an augmentation of the usual Grothendieck group to contain "vector-valued" characters: Given a non-zero intertwining operator $\Phi_\lambda : U_\lambda \to U_\lambda \otimes V$ for $\mathcal{U}_q(\mathfrak{g})$, the quantity $\mathrm{Tr}_{\lambda}(\Phi_\lambda\, X)$ for $X \in G$ is a G-equivariant function on the Lie group G with values in the representation V, called a generalized character in [103, 86]. When $V = \mathbb{C}$ is the trivial module we recover the usual characters and the monoidal functor from §5. More precisely, since conjugacy classes in G are the same as orbits of the Weyl group $W = \mathfrak{S}_N$ in the maximal torus $T = U(1)^N$, the generalized characters are uniquely defined by their values on T and take values in the weight 0 subspace $V^{(0)}$ in the usual weight decomposition of the representation V of G. The space of intertwining operators $\Phi_\lambda : U_\lambda \to U_\lambda \otimes V$ is isomorphic to the space $\mathsf{Hom}_{\mathscr{R}}(U_\lambda, U_\lambda \otimes V) \cong (U_\lambda^* \otimes U_\lambda \otimes V)^{\mathcal{U}_q(\mathfrak{g})}$.

We shall take $V = V_{\rm fund}^{\odot(\beta-1)\,N}$ for fixed $\beta \in \mathbb{Z}_{>0}$ to be the q-deformation of the $(\beta-1)\,N$-th symmetric power of the fundamental representation $V_{\rm fund} = \mathbb{C}^N$ of G; it is isomorphic to the space of homogeneous polynomials in N variables $x_1, \dots, x_N$ of degree $(\beta-1)\,N$. In this case all the weight subspaces of V are one-dimensional, and in particular $V^{(0)} \cong \mathbb{C}$. Since the tensor product multiplicities for $\mathcal{U}_q(\mathfrak{g})$ are the same as those for $\mathfrak{g}$, there is a non-zero $\mathcal{U}_q(\mathfrak{g})$-homomorphism $\Phi_\lambda : U_\lambda \to U_\lambda \otimes V_{\rm fund}^{\odot(\beta-1)\,N}$ if and only if $\lambda = \mu + (\beta-1)\,\rho$ for a highest weight μ; in this instance $\Phi^\bullet_\mu := \Phi_{\mu+(\beta-1)\,\rho}$ is unique up to normalization.

Etingof and Kirillov show that in this case the generalized characters $\mathrm{Tr}_{\lambda+(\beta-1)\,\rho}(\Phi^\bullet_\lambda\, X)$ are given by the monic form $M_\lambda(x;q,t)$ of the Macdonald polynomials in the eigenvalues $x = (x_1, \dots, x_N)$ of the matrix X at $t = q^\beta$ (see [103, Thm. 2.2]). Recall [28, Chap. VI] that $M_\lambda(x;q,t)$ can be defined algebraically as the unique symmetric polynomials satisfying the following two conditions:

(i) Triangular decomposition in dominance order with respect to the basis of monomial symmetric polynomials:

$$M_\lambda(x;q,t) = m_\lambda(x) + \sum_{\mu<\lambda} v_{\lambda,\mu}(q,t)\ m_\mu(x)\ ,$$

where $v_{\lambda,\mu}(q,t)$ are rational functions of q,t, the sum runs over N-component partitions μ such that $|\mu| = |\lambda|$ and $\mu_1 + \dots + \mu_i < \lambda_1 + \dots + \lambda_i$ for all $i = 1, \dots, N$, and

$$m_\lambda(x) = \sum_{w \in \mathfrak{S}_N(\lambda)} x_1^{w(1)} \cdots x_N^{w(N)}$$

with $\mathfrak{S}_N(\lambda)$ the set of distinct permutations of $\lambda_1, \dots, \lambda_N$.

(ii) Orthogonality:

$$\langle M_\lambda, M_\mu\rangle_{q,t} = 0 \qquad \text{for} \quad \lambda \neq \mu\ ,$$

where the inner product here is defined in the basis of power sum symmetric polynomials $p_\lambda = p_{\lambda_1} \cdots p_{\lambda_N}$, with $p_n(x) := m_{(n)}(x) = x_1^n + \dots + x_N^n$ for $n > 0$ and $p_0(x) := 1$, as

$$\langle p_\lambda, p_\mu\rangle_{q,t} = z_\lambda\ \delta_{\lambda,\mu} \prod_{i=1}^{\ell(\lambda)} \frac{1-q^{\lambda_i}}{1-t^{\lambda_i}} \tag{6.1}$$

with the length $\ell(\lambda)$ the number of non-zero parts of the partition $\lambda = (1^{m_1}\, 2^{m_2} \cdots)$ and

$$z_\lambda = \prod_{j\geq 1} j^{m_j}\, m_j!\ .$$

The Macdonald polynomials emcompass the various symmetric polynomials which play a role in this paper:

- The symmetrized monomials $m_\lambda(x) = M_\lambda(x;q,1)$ are obtained in the limit $t=1$ (independently of q).
- The Schur polynomials $s_\lambda(x) = M_\lambda(x;q,q)$ are obtained as the limit $t=q$ of the Macdonald polynomials (also independently of q); in this case the inner product (6.1) reduces to the Hall inner product $\langle p_\lambda, p_\mu\rangle = z_\lambda\, \delta_{\lambda,\mu}$ such that $\langle s_\lambda, s_\mu\rangle = \delta_{\lambda,\mu}$.
- The Hall-Littlewood polynomials $P_\lambda(x;t) = M_\lambda(x;0,t)$ are obtained in the limit $q=0$; they interpolate between the Schur polynomials $s_\lambda(x)$ at $t=0$ and the monomial symmetric polynomials $m_\lambda(x)$ at $t=1$.
- The Jack polynomials $J_\lambda(x;\alpha^{-1}) = \lim_{q\to 1} M_\lambda(x;q,q^\alpha)$ are obtained at $t=q^\alpha$ with $q\to 1$ for $\alpha\in\mathbb{C}$; they are a one-parameter deformation of the Schur polynomials with $s_\lambda(x) = J_\lambda(x;1)$, and the inner product (6.1) in this case is the Jack inner product $\langle p_\lambda, p_\mu\rangle_\alpha = z_\lambda\, \alpha^{-\ell(\lambda)}\ \delta_{\lambda,\mu}$.

The Macdonald inner product (6.1) on the space of symmetric functions can be defined analytically for functions on $T=(S^1)^N$ by the torus scalar product

$$\langle f,g\rangle_{q,t} := \frac{1}{N!}\int_{[0,2\pi)^N}\ \prod_{i=1}^N \frac{\mathrm{d}\phi_i}{2\pi}\ \Delta_{q,t}\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi}\big)\ f\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi}\big)\, g\big(\,\mathrm{e}^{\,-\mathrm{i}\,\phi}\big)\ , \tag{6.2}$$

where

$$\Delta_{q,t}(z) := \prod_{i\neq j}\frac{\big(z_i\, z_j^{-1}\ ;\ q\big)_\infty}{\big(t\, z_i\, z_j^{-1}\ ;\ q\big)_\infty} \tag{6.3}$$

is the Macdonald measure for $z=(z_1,\dots,z_N)\in T$ and $0\le t\le 1$. For $t=q^\beta$, $\beta\in\mathbb{Z}_{>0}$, with $z_i = \mathrm{e}^{\,\mathrm{i}\,\phi_i}$ and $\phi_i\in[0,2\pi)$ for $i=1,\dots,N$, we can write

$$\Delta_{q,t}\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi}\big) = \prod_{m=0}^{\beta-1}\ \prod_{i\neq j}\Big(1-q^m\ \mathrm{e}^{\,\mathrm{i}\,(\phi_i-\phi_j)}\Big)$$

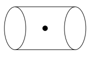
which is the q-analog of the β-th power of the Weyl determinant on $G=U(N)$. Then the norm of the Macdonald polynomials $M_\lambda(x;q,t)$ is given by

$$\|M_\lambda\|^2_{q,t} := \langle M_\lambda, M_\lambda\rangle_{q,t} = \prod_{i<j}\frac{\big(q^{\lambda_i-\lambda_j}\, t^{j-i}\ ;\ q\big)_\infty\ \big(q^{\lambda_i-\lambda_j+1}\, t^{j-i}\ ;\ q\big)_\infty}{\big(q^{\lambda_i-\lambda_j}\, t^{j-i+1}\ ;\ q\big)_\infty\ \big(q^{\lambda_i-\lambda_j+1}\, t^{j-i-1}\ ;\ q\big)_\infty}\ ,$$

which for $t=q^\beta$, $\beta\in\mathbb{Z}_{>0}$ yields Macdonald's inner product identity

$$\|M_\lambda\|^2_{q,t} = \prod_{m=0}^{\beta-1}\ \prod_{i<j}\frac{\big[\lambda_i-\lambda_j+\beta\,(j-i)+m\big]_q}{\big[\lambda_i-\lambda_j+\beta\,(j-i)-m\big]_q}\ .$$

To compute the corresponding refinements of the data for the semisimple ribbon category $\mathscr{R}=\mathscr{R}ep(\mathcal{U}_q(\mathfrak{g}))$, we first consider the punctured tube amplitude which is the modification of (5.12) in this case given by

(6.4) $$\Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{g})}\Big($$  $$\Big)\Big] = \sum_\lambda \theta_\lambda\ \big[U_\lambda\otimes U_\lambda^*\otimes V_{\rm fund}^{\odot(\beta-1)\,N}\big] = \sum_\lambda \frac{\theta_\lambda^\bullet}{\|M_\lambda\|_{q,t}}\ [\Phi_\lambda^\bullet]$$

where $\theta^\bullet_\lambda := \theta_{\lambda+(\beta-1)\,\rho}$, we used $\bigoplus_\lambda \mathsf{Hom}_{\mathscr{R}}\big(U_\lambda\,,\,U_\lambda\otimes V_{\rm fund}^{\odot(\beta-1)\,N}\big) = \bigoplus_\lambda \mathbb{C}\,\Phi_\lambda^\bullet/\|M_\lambda\|_{q,t}$ together with the canonical duality isomorphisms in the category $\mathscr{R}$, and we accounted for the non-trivial normalisation of the intertwining operators $\Phi^\bullet_\lambda$ [86]. It follows that the twist $\theta^\bullet$ in this basis is given by [86]

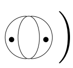
$$\theta^\bullet_\lambda = q^{\frac12\,\langle\lambda,\lambda\rangle}\ t^{\langle\lambda,\rho\rangle}\ , \tag{6.5}$$

where we used

$$C_2\big(\lambda+(\beta-1)\,\rho\big) = \langle\lambda+\beta\,\rho,\lambda+\beta\,\rho\rangle - \langle\rho,\rho\rangle$$

and normalized $\theta^\bullet$ so that $\theta^\bullet_0=1$. Here and in the following we should strictly speaking take $t=q^\beta$ with $\beta\in\mathbb{Z}_{>0}$; however, in [104, 103] it is shown how to analytically continue many of these formulas and results to arbitrary values $\beta\in\mathbb{C}$, and hence we shall often write formulas for arbitrary, algebraically independent refinement parameters $t\in\mathbb{C}$ with $|t|\le q$ as well. In this case $V=V_{\rm fund}^{\odot(\beta-1)\,N}$ is regarded formally as a representation of $\mathcal{U}_q(\mathfrak{g})$ with "highest weight" $(\beta-1)\,N\,\omega_1$ where ω_1 is the fundamental weight.

Next, computing as we did in (6.4), we consider a sphere with two marked points which leads to the modification of the amplitude (5.15) given by

$$\Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{g})}\Big($$

$$\Big)^B\Big] = \sum_{\lambda,\mu} S^\bullet_{\lambda\mu}\ \frac{\theta^\bullet_\lambda\,\theta^\bullet_\mu}{\|M_\lambda\|_{q,t}\,\|M_\mu\|_{q,t}}\ [\Phi^{\bullet}_\lambda{}^*\otimes\Phi^\bullet_\mu]$$

where we have further used the canonical identifications $\mathsf{Hom}_{\mathscr{R}}(\mathbb{1},U_\lambda\otimes U_\lambda^*)\cong\mathbb{C}$ in the category $\mathscr{R}$. The braiding symmetry is computed in this basis in [86, Thm. 5.4] and with suitable normalization it can be expressed in terms of specializations of the Macdonald functions as

$$S^\bullet_{\lambda\mu} = M_\lambda\big(t^\rho\,;\,q,t\big)\ M_\mu\big(t^\rho\,q^\lambda\,;\,q,t\big)\ .$$

Symmetry $S^\bullet_{\lambda\mu}=S^\bullet_{\mu\lambda}$ is now a consequence of the self-duality identity for Macdonald polynomials given by [28, Chap. VI]

$$\frac{M_\lambda\big(q^{\mu_1}\,t^{N-1},q^{\mu_2}\,t^{N-2},\dots,q^{\mu_N}\,;\,q,t\big)}{M_\lambda\big(t^{N-1},t^{N-2},\dots,1\,;\,q,t\big)} = \frac{M_\mu\big(q^{\lambda_1}\,t^{N-1},q^{\lambda_2}\,t^{N-2},\dots,q^{\lambda_N}\,;\,q,t\big)}{M_\mu\big(t^{N-1},t^{N-2},\dots,1\,;\,q,t\big)}\ .$$

In particular, the categorical dimensions in this basis are given by Macdonald's special value identity

$$\begin{aligned}\dim_{q,t}\lambda &:= S^\bullet_{\lambda 0}\\ &= M_\lambda\big(t^\rho\,;\,q,t\big)\\ &= t^{|\lambda|/2}\ M_\lambda\big(1,t,\dots,t^{N-1}\,;\,q,t\big)\ =\ \prod_{i=1}^N t^{(i-\frac12)\,\lambda_i}\ \prod_{j<k}\frac{\big(t\,q^{k-j}\,;\,q\big)_{\lambda_j-\lambda_k}}{\big(q^{k-j}\,;\,q\big)_{\lambda_j-\lambda_k}}\ .\end{aligned}$$

When $t=q^\beta$ with $\beta\in\mathbb{Z}_{>0}$ we can rewrite this formula as [103]

$$\dim_{q,t}\lambda = \prod_{i<j}\ \prod_{m=0}^{\beta-1}\frac{\big[\lambda_i-\lambda_j+\beta\,(j-i)+m\big]_q}{\big[\beta\,(j-i)+m\big]_q}\ ,$$

which exhibits the refined categorical dimension as the q-analog of the β-th power of the quantum dimension $\dim_q\lambda$. Finally, the structure constants in this basis are no longer integer-valued but instead are rational functions $N^\bullet_{\lambda\mu}{}^\nu(q,t)$ of q,t which appear as generalized Littlewood-Richardson coefficients

$$M_\lambda(x;q,t)\,M_\mu(x;q,t) = \sum_\nu N^\bullet_{\lambda\mu}{}^\nu(q,t)\ M_\nu(x;q,t)\ .$$

We thus find that the partition function (5.14) is given by

$$\mathcal{Z}^\bullet\big(\mathcal{U}_q(\mathfrak{g})\,;\,\Sigma_h\big)=\sum_\lambda\,\bigg(\,\frac{M_\lambda\big(t^\rho\,;\,q,t\big)}{\|M_\lambda\|_{q,t}}\,\bigg)^{2-2h}\ q^{\frac12\,\langle\lambda,\lambda\rangle}\,t^{\langle\lambda,\rho\rangle}\ , \tag{6.6}$$

while (5.16) in this case becomes

$$\begin{aligned}\mathcal{Z}^\bullet\big(\mathcal{U}_q(\mathfrak{g})\,;\,S^2,p,p'\,\big)\ &=\ \sum_{\lambda_1,\dots,\lambda_\ell}\frac{q^{\frac12\,(e_1\,\langle\lambda_1,\lambda_1\rangle+\cdots+e_\ell\,\langle\lambda_\ell,\lambda_\ell\rangle)}\,t^{\langle e_1\,\lambda_1+\cdots+e_\ell\,\lambda_\ell,\rho\rangle}}{\|M_{\lambda_1}\|_{q,t}^2\,\|M_{\lambda_2}\|_{q,t}\cdots\|M_{\lambda_{\ell-1}}\|_{q,t}\,\|M_{\lambda_\ell}\|_{q,t}^2}\\&\qquad\times\,M_{\lambda_1}\big(t^\rho\,;\,q,t\big)^2\,M_{\lambda_2}\big(t^\rho\,;\,q,t\big)\cdots M_{\lambda_\ell}\big(t^\rho\,;\,q,t\big)\\&\qquad\times\,M_{\lambda_2}\big(t^\rho\,q^{\lambda_1}\,;\,q,t\big)\cdots M_{\lambda_\ell}\big(t^\rho\,q^{\lambda_{\ell-1}}\,;\,q,t\big)\ .\end{aligned}\tag{6.7}$$

It is also straightforward to write down defect observables along the lines of §5.6: In the conventions of §5.6, for the correlator of a single Wilson loop in the representation λ we find

$$\begin{aligned}\mathcal{W}_\lambda^\bullet\big(\mathcal{U}_q(\mathfrak{g})\,;\Sigma_h,p,1\big)\ &=\ \sum_{\lambda_1,\lambda_2}N^\bullet_{\lambda_1\lambda}{}^{\lambda_2}\,\bigg(\frac{M_{\lambda_1}\big(t^\rho\,;\,q,t\big)}{\|M_{\lambda_1}\|_{q,t}}\bigg)^{1-2h_1}\bigg(\frac{M_{\lambda_2}\big(t^\rho\,;\,q,t\big)}{\|M_{\lambda_2}\|_{q,t}}\bigg)^{1-2h_2}\\&\qquad\times\,q^{\frac p2\,(\langle\lambda_1,\lambda_1\rangle+\langle\lambda_2,\lambda_2\rangle)}\,t^{p\,\langle\lambda_1+\lambda_2,\rho\rangle}\ ,\end{aligned}$$

while the general amplitude for a Riemann surface of genus h with b boundaries fixed at holonomy eigenvalues $u_1,\dots,u_b\in(S^1)^N$ and n defect punctures labelled by irreducible representations $\lambda_1,\dots,\lambda_n$ is given by

$$\begin{aligned}\mathcal{O}_{\substack{u_1,\dots,u_b\\\lambda_1,\dots,\lambda_n}}\big(\mathcal{U}_q(\mathfrak{g})\,;\,\Sigma_h,p,1\big)\ &=\ \sum_\lambda\bigg(\frac{\dim_{q,t}\lambda}{\|M_\lambda\|_{q,t}}\bigg)^{2-2h-b-n}\ q^{\frac p2\,\langle\lambda,\lambda\rangle}\,t^{p\,\langle\lambda,\rho\rangle}\,\prod_{i=1}^b\frac{M_\lambda(u_i;q,t)}{\|M_\lambda\|_{q,t}}\\&\qquad\times\,\prod_{j=1}^n\frac{S^\bullet_{\lambda\lambda_j}}{\|M_\lambda\|_{q,t}\,\|M_{\lambda_j}\|_{q,t}}\\&=\ \sum_\lambda\frac{M_\lambda\big(t^\rho\,;\,q,t\big)^{2-2h-b}}{\|M_\lambda\|_{q,t}^{2-2h}}\ q^{\frac p2\,\langle\lambda,\lambda\rangle}\,t^{p\,\langle\lambda,\rho\rangle}\,\prod_{i=1}^bM_\lambda(u_i;q,t)\\&\qquad\times\,\prod_{j=1}^n\frac{M_{\lambda_j}\big(t^\rho\,q^\lambda\,;\,q,t\big)}{\|M_{\lambda_j}\|_{q,t}}\ .\end{aligned}$$

6.2. **Refined $L(p,1)$ matrix models.**

Let us consider the partition function (6.7) for refined q-deformed Yang-Mills theory on S^2 at $p'=1$ and with $t=q^\beta$ for $\beta\in\mathbb{Z}_{>0}$. Setting $\mu_i=\lambda_i+\beta\,(N-i)$ for $i=1,\dots,N$ yields a β-deformation of the discrete Gaussian matrix model (2.12) at $h=0$ given by

$$\begin{aligned}\mathcal{Z}^\bullet_{(p)}(q,t;S^2)\ &:=\ \mathcal{Z}^\bullet\big(\mathcal{U}_q(\mathfrak{g})\,;\,S^2,p,1\big)\\&=\ \sum_{\mu\in\mathbb{Z}^N}\ \prod_{i<j}\ \prod_{m=0}^{\beta-1}\,[\mu_i-\mu_j+m]_q\,[\mu_i-\mu_j-m]_q\ q^{\frac p2\,\sum_i\mu_i^2}\ .\end{aligned}$$

For $p=1$, by performing analogous steps to those of §3.1 we may rewrite this partition function in the form

$$\begin{aligned}
\mathcal{Z}^{\bullet}_{(1)}(q,t;S^2) &= \sum_{u\in\mathbb{Z}^N} q^{\frac12\sum_i u_i^2} \prod_{j\neq k}\prod_{m=0}^{\beta-1}\Big(q^{\frac12(u_j-u_k+m)}-q^{\frac12(u_k-u_j-m)}\Big)\\
&= q^{\frac12\beta(\beta-1)}\sum_{u\in\mathbb{Z}^N} q^{\frac12\sum_i u_i^2}(\sigma\,q)^{\beta\sum_i u_i}\prod_{j\neq k}\prod_{m=0}^{\beta-1}\big(q^{u_j}-q^{-m}\,q^{u_k}\big)\\
&= q^{\frac12\beta(\beta-1)}\,\sigma^{\beta N(1-\beta N)}\int_{\mathbb{R}^N_{>0}}\prod_{i=1}^N \mathrm{d}x_i\;x_i^{\beta-1}\sum_{n_i=-\infty}^{\infty}\sigma^{\beta n_i}\,q^{\frac12 n_i^2+n_i}\,\delta(x_i-\sigma^\beta\,q^{n_i})\\
&\qquad\qquad\times\prod_{m=0}^{\beta-1}\prod_{j\neq k}\big(x_j-q^{-m}\,x_k\big)\\
&= q^{\frac12\beta(\beta-1)+\frac12 N}\,\sigma^{\beta N(1-\beta N)}\,M(q,\sigma^\beta)^N\\
&\quad\times\int_{\mathbb{R}^N_{>0}}\prod_{i=1}^N\mathrm{d}x_i\;w_d(x_i;q,\sigma^\beta)\,\prod_{j<k}(x_j-x_k)^2\;P(x_1,\dots,x_N;q,t)\ ,
\end{aligned}$$

where as before $\sigma=q^{-N}$, the discrete measure $w_d(x;q,\sigma)$ is given by (3.4), and we have introduced the polynomial

$$P(x_1,\dots,x_N;q,t)=\prod_{m=1}^{\beta-1}\prod_{i=1}^N x_i\prod_{j\neq k}\big(x_j-q^{-m}\,x_k\big)$$

with the convention $P(x_1,\dots,x_N;q,q):=1$. The specialization of the family of measures (3.4) at $\sigma^\beta=q^{-N\beta}$ is again equivalent to the specialization $\sigma=1$ due to the discrete scaling symmetry (3.5). Arguing as in §3, since the β-deformation here amounts to a polynomial average in the matrix model, only integer moments are involved in its computation, and the discrete/continuous equivalence of the indeterminate moment problem for the Stieltjes-Wigert matrix model applies to it as well. With $q:=\mathrm{e}^{-g_s}$ as previously, we can therefore rewrite the refined partition function as a correlator in the continuous matrix model

$$\begin{aligned}
\mathcal{Z}^{\bullet}_{(1)}(q,t;S^2) &= q^{\frac12\beta(\beta-1)+\frac12 N}\,\sigma^{\beta N(1-\beta N)}\,M(q,\sigma^\beta)^N\\
&\quad\times\int_{\mathbb{R}^N_{>0}}\prod_{i=1}^N\mathrm{d}x_i\;\omega_{\rm SW}\big(x_i\,;\,\tfrac{1}{\sqrt{2g_s}}\big)\,\prod_{j<k}(x_j-x_k)^2\;P(x_1,\dots,x_N;q,t)
\end{aligned}$$

defined by the Stieltjes-Wigert distribution (3.1). Setting $u_i=\log(x_i)-\frac14\,g_s\,(\beta-1)$, we may alternatively rewrite this matrix model average as

$$\begin{aligned}
\mathcal{Z}^{\bullet}_{(1)}(q,t;S^2) &= C_N(q,t)\,\Big(\frac{g_s}{2\pi}\Big)^{-N/2}\\
&\quad\times\int_{\mathbb{R}^N}\prod_{i=1}^N\frac{\mathrm{d}u_i}{2\pi}\;\mathrm{e}^{-u_i^2/2g_s}\prod_{m=0}^{\beta-1}\prod_{j\neq k}\Big(\mathrm{e}^{\frac12(u_j-u_k)}-q^{-m}\,\mathrm{e}^{\frac12(u_k-u_j)}\Big)
\end{aligned}\tag{6.8}$$

where

$$C_N(q,t)=q^{-N\beta N^2(1-\beta N)}\,q^{\frac12 N+\frac18(\beta-1)(3\beta+1)}\,\big(-q^{\frac32-\beta N}\,;q\big)_\infty^N\,\big(-q^{\beta N-\frac12}\,;\,q\big)_\infty^N\,\big(q\,;\,q\big)_\infty^N\ .$$

Up to overall normalization, this is just the partition function $Z_N(q,t):=\mathcal{Z}^{\bullet}_{(1)}(q,t;S^2)/C_N(q,t)$ of the β-deformed matrix model for refined Chern-Simons theory on S^3 which was considered in [24]; the product over m in (6.8) yields the q-analog of the β-th power of the square of the Vandermonde determinant in this case.

For $p>1$, completely analogous calculations to those of §3.2 can be similarly carried out. Dropping all normalization constants and setting $q=\mathrm{e}^{-g_s}$, the partition function

$$\begin{aligned}\mathcal{Z}^{\bullet}_{(p)}(q,t;S^2) &= \sum_{u\in\mathbb{Z}^N}\exp\Big(-\frac{p\,g_s}{2}\sum_{i=1}^N u_i^2\Big)\prod_{j\neq k}\prod_{m=0}^{\beta-1}\Big(\mathrm{e}^{\frac{g_s}{2}(u_j-u_k)}-q^{-m}\,\mathrm{e}^{\frac{g_s}{2}(u_k-u_j)}\Big)\\ &= \mathcal{Z}^{(p)}_{\mathrm{SW}}[0](q,t;S^2)\end{aligned}$$

is the $\mathbb{Z}_p$-invariant projection of the partition function of the corresponding full β-deformed Stieltjes-Wigert ensemble

$$\begin{aligned}\mathcal{Z}^{(p)}_{\mathrm{SW}}(q,t;S^2) &= \int_{\mathbb{R}^N}\prod_{i=1}^N \mathrm{d}u_i\ \mathrm{e}^{-\frac{p}{2g_s}u_i^2}\prod_{j\neq k}\prod_{m=0}^{\beta-1}\Big(\mathrm{e}^{\frac12(u_j-u_k)}-q^{-m}\,\mathrm{e}^{\frac12(u_k-u_j)}\Big)\\ &= \sum_{u\in(\mathbb{Z}/p)^N}\exp\Big(-\frac{p\,g_s}{2}\sum_{i=1}^N u_i^2\Big)\prod_{j\neq k}\prod_{m=0}^{\beta-1}\Big(\mathrm{e}^{\frac{g_s}{2}(u_j-u_k)}-q^{-m}\,\mathrm{e}^{\frac{g_s}{2}(u_k-u_j)}\Big)\end{aligned}$$

considered above. In particular, for $q=\zeta_k$ a primitive k-th root of unity one finds the continuous matrix model representation

$$\mathcal{Z}^{\bullet}_{(p)}(q=\zeta_k,t;S^2)=\int_{\mathbb{R}^N}\prod_{i=1}^N\frac{\mathrm{d}u_i}{2\pi}\ \mathrm{e}^{-u_i^2/2g_s}\ \Gamma_{k,p}(\mathrm{e}^{u_i})\prod_{m=0}^{\beta-1}\prod_{j\neq k}\Big(\mathrm{e}^{\frac12(u_j-u_k)}-q^{-m}\,\mathrm{e}^{\frac12(u_k-u_j)}\Big)$$

which coincides with (6.8) for $p=1$.

6.3. **Refined unitary matrix model.**

In [24] it was shown that the matrix integral (6.8) has a representation as a $U(N)$ unitary matrix model obtained by substituting the usual Haar measure of the matrix model (3.12) with the Macdonald measure $\Delta_{q,t}(\mathrm{e}^{\,\mathrm{i}\phi})$ from (6.3), so that

$$Z_N(q,t)=\int_{[0,2\pi)^N}\prod_{i=1}^N\frac{\mathrm{d}\phi_i}{2\pi}\ \Theta\big(\mathrm{e}^{\,\mathrm{i}\phi_i};q\big)\ \Delta_{q,t}\big(\mathrm{e}^{\,\mathrm{i}\phi}\big)\ . \tag{6.9}$$

Below we describe the relevance of such types of matrix integrals to the refined gauge theories considered in [24, 14, 18, 23].

The partition function of the unitary matrix model (6.9) can be computed analytically from the explicit evaluation [49]

$$\int_{[0,2\pi)^k}\prod_{i=1}^k\frac{\mathrm{d}\phi_i}{2\pi}\ \frac{\Theta_n\big(\mathrm{e}^{\,\mathrm{i}\phi_i};q\big)}{(q;q)_n}\ \Delta_{q,t}\big(\mathrm{e}^{\,\mathrm{i}\phi}\big)=\prod_{i=0}^{k-1}\frac{\big(t^i\,q^{n+1};q\big)_n}{\big(t^i\,q;q\big)_n} \tag{6.10}$$

generalising (3.19). This expression reduces to the two formulas in (3.19) in the limit $t=q^\beta$ with $\beta=1$. In the limit $k=N$, $n\to\infty$, it evaluates (6.9) as

$$Z_N(q,t)=\prod_{k=1}^{N-1}\frac{(t^k\,q;q)_\infty}{(q;q)_\infty}\ . \tag{6.11}$$

When $t=q^\beta$ for $\beta\in\mathbb{Z}_{>0}$, the expression (6.11) can be written as

$$Z_N(q,t)=\prod_{k=1}^{N-1}\prod_{m=0}^{\beta-1}\big(1-q^{\beta\,k+m}\big)^{N-k}\ ,$$

which coincides with the exact analytical expression for the partition function of the β-deformed Stieltjes-Wigert matrix model for refined Chern-Simons theory on S^3 that we considered in §6.2 [24]. However, there are no known determinantal expressions for Macdonald polynomials which yield a suitable analog of Gessel's identity for the generalization of the Cauchy-Binet summation formula (3.18); already the limiting case of Jack polynomials involves Toeplitz hyperdeterminants [105]. We are not aware of any analog of the matrix model expression (3.20) for the topological refined q-deformed two-dimensional gauge theory considered in [18] ($p = 0$ in (5.17)); we return to this point below.

6.4. **Refined q-deformed BF-theory.**

We can also consider refined versions of the six-dimensional $\mathcal{N} = 2$ gauge theories that we studied in §4.4, whose partition functions compute refined (or motivic) Donaldson-Thomas invariants [106]. We start from the refinement of the $U(\infty)$ matrix model (4.16) with partition function

$$\mathcal{Z}^{6D}(q,t) := \int_{[0,2\pi)^\infty} \prod_{i=1}^\infty \frac{\mathrm{d}\phi_i}{2\pi} \, \frac{\Theta\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\,q\big)}{(q;q)_\infty} \, \Delta_{q,t}\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi}\big) \ . \tag{6.12}$$

The weight function is given again by (4.17) and there is an extension of the Szegő limit theorem to this generalized case which is spelled out in Appendix B. Using the Fourier coefficients (4.18) and the limit formula (B.3), we compute

$$\begin{aligned} \log \mathcal{Z}^{6D}(q,t) &= \sum_{k=1}^\infty k \, [\log f]_k \, [\log f]_{-k} \, \frac{1-q^k}{1-t^k} \\ &= \sum_{k=1}^\infty \frac{q^k}{k\,\big(1-q^k\big)\,\big(1-t^k\big)} \\ &= \sum_{k=1}^\infty \, \sum_{n,m=1}^\infty \frac{q^{k\,n}\,t^{k\,(m-1)}}{k} \ = \ - \sum_{n,m=1}^\infty \log\big(1-q^n\,t^{m-1}\big) \ . \end{aligned} \tag{6.13}$$

Thus the partition function evaluates explicitly to

$$\mathcal{Z}^{6D}(q,t) = M(q,t) \ ,$$

where $M(q,t)$ is exactly the refined MacMahon function [107]

$$M(q,t) = \prod_{n,m=1}^\infty \frac{1}{1-q^n\,t^{m-1}} \ ;$$

this formula also follows from the $k,n \to \infty$ limit of (6.10). Whence the change of integration measure from the Haar measure to the Macdonald measure produces the refined version of the $\mathcal{N} = 2$ gauge theory partition function, which in this case is the MacMahon function. Employing the bilinear sum identity

$$\sum_\lambda \frac{1}{\|M_\lambda\|^2_{q,t}} \, M_\lambda(x;q,t) \, M_\lambda(y;q,t) = \prod_{i,j\geq 1} \frac{(t\,x_i\,y_j;q)_\infty}{(x_i\,y_j;q)_\infty}$$

for the Macdonald functions, which generalizes the Cauchy-Binet formula (3.18), at the specialisation $x = y = t^\rho$ shows that the partition function (6.12) can also be expanded in terms of refined quantum dimensions as

$$\mathcal{Z}^{6D}(q,t) = \sum_\lambda \bigg(\frac{\dim_{q,t}\lambda}{\|M_\lambda\|_{q,t}}\bigg)^2 \ .$$

This is just the $U(\infty)$ version of the topological (q,t)-deformed two-dimensional gauge theory which was derived in [18] from a four-dimensional superconformal index.

We can also apply the Macdonald measure substitution to define the family of refined partition functions

$$\mathscr{Z}_L^{6D}(\alpha;q,t,Q) := \int_{[0,2\pi)^\infty} \prod_{i=1}^\infty \frac{\mathrm{d}\phi_i}{2\pi}\ F_L\big(\mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\,\alpha\,;\,q,Q\big)\ \Delta_{q,t}\big(\mathrm{e}^{\,\mathrm{i}\,\phi}\big)$$

with weight functions (4.21). This modifies the result (4.23) to

$$\mathscr{Z}_L^{6D}(\alpha;q,t,Q) = \prod_{a=1}^L M(q,t)^{\alpha_a^2}\ \prod_{b\neq c} M\big(Q_b\,Q_c^{-1},q,t\big)^{\alpha_b\,\alpha_c}\ ,$$

where $M(Q,q,t)$ is the refined generalised MacMahon function

$$M(Q,q,t) = \prod_{n,m=1}^\infty \frac{1}{1-Q\,q^n\,t^{m-1}}\ .$$

The derivation presented here is an alternative and equivalent method for obtaining refined partition functions based on refinements of the weight function of the matrix model [108]. Indeed, the form of the generalized strong Szegő theorem (B.3) suggests the possibility of interpreting the additional factor $\big(1-q^k\big)\,/\,\big(1-t^k\big)$ as a generalization of the Fourier coefficients $[\log f]_k$, and therefore one can construct the corresponding $U(\infty)$ matrix model with the usual Haar measure but with a generalized weight function. These are the types of matrix models that were considered in [108], where a "refined" theta-function

$$\widetilde{\Theta}(z;q,t) = \prod_{n=1}^\infty \big(1+q^{n-1/2}\,z\big)\,\big(1+t^{n-1/2}\,z^{-1}\big) \tag{6.14}$$

was used as weight function. We can see the relationship explicitly from the ordinary Szegő limit theorem: One has

$$\big[\log\widetilde{\Theta}\,\big]_k = \frac{(-1)^{k+1}\ q^{k/2}}{k\,\big(1-q^k\big)} \qquad \text{and} \qquad \big[\log\widetilde{\Theta}\,\big]_{-k} = \frac{(-1)^{k+1}\ t^{k/2}}{k\,\big(1-t^k\big)}$$

for $k>0$, and hence

$$\log\widetilde{\mathscr{Z}}^{\,6D}(q,t) = \sum_{k=1}^\infty k\,\big[\log\widetilde{\Theta}\,\big]_k\,\big[\log\widetilde{\Theta}\,\big]_{-k} = \sum_{k=1}^\infty \frac{q^{k/2}\,t^{k/2}}{k\,\big(1-q^k\big)\,\big(1-t^k\big)}\ .$$

Written in this form, the refined partition function $\widetilde{\mathscr{Z}}^{\,6D}(q,t)$ is realised as an elliptic gamma-function [109]. By Gessel's identity, the unitary matrix model with weight function (6.14) follows from a $U(\infty)$ two-dimensional gauge theory with partition function

$$\widetilde{\mathscr{Z}}^{\,6D}(q,t) = \sum_\lambda s_\lambda\big(q^\rho\big)\,s_\lambda\big(t^\rho\big) \tag{6.15}$$

involving two deformation parameters q,t. Whence the refined six-dimensional gauge theory is equivalent to the (q,t)-deformed topological Yang-Mills theory on S^2 in (6.15). This result shows moreover that the selection made in [108] for a refined theta-function (6.14) is equivalent to that given by the refined topological vertex introduced in [107], i.e. application of Gessel's identity to the expressions in [107] directly gives the matrix models of [108]. The partition function (6.15) also coincides with the perturbative part of Nekrasov's partition function for five-dimensional gauge theory [110], with the variables q,t parametrizing the Ω-background.

For the case of $\mathcal{N}=2$ gauge theory on the noncommutative conifold discussed in [108], this double deformation also gives the correct result if one generalizes the Schur polynomials to the supersymmetric Schur polynomials considered in §4.4: The weight function used in the refined matrix model of [108] is reproduced in (4.28) with the specialisation $x_i = q^{i-\frac{1}{2}}$, $y_i = t^{i-\frac{1}{2}}$, $z_i = -Q\, q^{i-\frac{1}{2}}$ and $w_i = -Q^{-1}\, t^{i-\frac{1}{2}}$, which yields the refined generating function for the noncommutative conifold

$$\begin{aligned} \mathcal{Z}_2^{6D}(1,-1;q,t,1,Q) &= \frac{M(q,t)^2}{M(Q,q,t)\, M(Q^{-1},q,t)} \\ &= \int_{[0,2\pi)^\infty} \prod_{i=1}^\infty \frac{\mathrm{d}\phi_i}{2\pi}\, \frac{\widetilde{\Theta}\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\,q,t\big)}{\widetilde{\Theta}\big(Q^{-1}\,\mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\,q,t\big)}\, \prod_{j<k} \big|\,\mathrm{e}^{\,\mathrm{i}\,\phi_j} - \mathrm{e}^{\,\mathrm{i}\,\phi_k}\big|^2 \ . \end{aligned}$$

This expression coincides with the partition function of a supersymmetric extension of the refined topological gauge theory (6.15) given by

$$\mathcal{Z}_2^{6D}(1,-1;q,t,1,Q) = \sum_\lambda \mathrm{HS}_\lambda\left(q^\rho\,\big|\, -Q\, q^\rho\right)\, \mathrm{HS}_\lambda\left(t^\rho\,\big|\, -Q^{-1}\, t^\rho\right)\ .$$

Notice that the expansion based on hook-Schur polynomials also allows for a richer specialization, with up to four deformation parameters; below we consider examples of such multiple refinements.

Thus, at this level of refinement, one may equivalently use either Schur polynomials or Macdonald polynomials. This observation can also be applied to provide a new impetus on the partition function (4.24). It is possible to obtain a linear relationship between the respective powers of the theta-function as a weight function of the matrix model and of the MacMahon partition function, if one considers β-ensembles instead of unitary ensembles; these ensembles naturally arise when one refines to the Jack polynomials $J_\lambda(x;\alpha^{-1})$ instead of Schur polynomials $s_\lambda(x) = J_\lambda(x;1)$. They also satisfy a Cauchy identity

$$\sum_\lambda \frac{1}{\|J_\lambda\|_\alpha^2}\ J_\lambda(x;\alpha^{-1})\, J_\lambda(y;\alpha^{-1}) = \prod_{i,j\geq 1}\ \frac{1}{(1-x_i\, y_j)^\alpha}\ , \tag{6.16}$$

where the norms $\|J_\lambda\|_\alpha^2$ are rational functions of the parameter α. The same results that lead to unitary matrix models from the expansion into Schur functions apply to the Jack functions as well, by replacing Toeplitz determinants with Toeplitz hyperdeterminants and unitary ensembles with β-ensembles (here with $\beta = 2\alpha$); this follows from the Heine-Szegő identity for Toeplitz hyperdeterminants [105]. Hence we may replace the matrix model representation (4.24) with

$$\mathcal{Z}_1^{6D}\big(\sqrt{\tfrac{\chi}{2}}\,;\,q\big) = \sum_\lambda\, \left(\frac{J_\lambda\big(\,q^\rho\,;\,\frac{2}{\chi}\,\big)}{\|J_\lambda\|_{\chi/2}}\right)^2 = \int_{[0,2\pi)^\infty}\, \prod_{i=1}^\infty \frac{\mathrm{d}\phi_i}{2\pi}\, \frac{\Theta\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\,q\big)}{(q;q)_\infty}\, \prod_{j<k} \big|\,\mathrm{e}^{\,\mathrm{i}\,\phi_j} - \mathrm{e}^{\,\mathrm{i}\,\phi_k}\big|^\chi\ , \tag{6.17}$$

and both matrix models represent the partition function $M(q)^{\chi/2}$.

These two types of refinements have been previously considered in the context of the refined topological vertex: Awata and Kanno introduced a refinement of the topological vertex based on Macdonald polynomials [111, 112], whereas the refinement of Iqbal, Kozcaz and Vafa [107] is based on different specializations of the Schur polynomials as in (6.15). The Macdonald refinement of two-dimensional Yang-Mills theory was originally presented in [18]; the analysis presented above is the first consideration and comparison of both refinements in the simpler setting of two-dimensional gauge theory.

6.5. **Higher refinement.**

With the Macdonald polynomials one can easily obtain more general refinements by giving a different set of parameters (q_a, t_a) to each polynomial. We have thus far considered the (related) cases of gauge theories associated to matrix models involving a refined theta-function (6.14) with the ordinary Haar measure, and an ordinary theta-function (3.14) with the Macdonald measure. We will now combine these two refinements and use distinct refinement parameters. The partition function is

$$\mathcal{Z}^{6D}(q_1,t_1;q_2,t_2) := \int_{[0,2\pi)^\infty} \prod_{i=1}^\infty \frac{\mathrm{d}\phi_i}{2\pi} \ \widetilde{\Theta}\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi_i}\,;\,q_1,t_1\big) \ \Delta_{q_2,t_2}\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi}\big) \ ,$$

and by the strong Szegő limit theorem we have

$$\begin{aligned} \log \mathcal{Z}^{6D}(q_1,t_1;q_2,t_2) &= \sum_{k=1}^\infty k\,\big[\log\widetilde{\Theta}\,\big]_k\,\big[\log\widetilde{\Theta}\,\big]_{-k}\,\frac{1-q_2^k}{1-t_2^k} \\ &= \sum_{k=1}^\infty \frac{q_1^{k/2}\,t_1^{k/2}\,\big(1-q_2^k\big)}{k\,\big(1-q_1^k\big)\,\big(1-t_1^k\big)\,\big(1-t_2^k\big)} \\ &= \sum_{k=1}^\infty \frac{1-q_2^k}{k}\ \sum_{n,m,l=1}^\infty q_1^{k\,(n-1/2)}\,t_1^{k\,(m-1/2)}\,t_2^{k\,(l-1)} \\ &= -\sum_{n,m,l=1}^\infty \Big(\log\big(1-q_1^{n-1/2}\,t_1^{m-1/2}\,t_2^{l-1}\big) \\ &\qquad\qquad - \log\big(1-q_2\,q_1^{n-1/2}\,t_1^{m-1/2}\,t_2^{l-1}\big)\Big) \ . \end{aligned}$$

It follows that

$$\mathcal{Z}^{6D}(q_1,t_1;q_2,t_2) = \prod_{n,m,l=1}^\infty \frac{1-q_2\,q_1^{n-1/2}\,t_1^{m-1/2}\,t_2^{l-1}}{1-q_1^{n-1/2}\,t_1^{m-1/2}\,t_2^{l-1}} \ .$$

Note that $\mathcal{Z}^{6D}(q,t;q,q)=\widetilde{\mathcal{Z}}^{\,6D}(q,t)$ is the elliptic gamma-function which corresponds to the perturbative part of the $\mathcal{N}=1$ gauge theory partition function in five dimensions. On the other hand, in the Hall-Littlewood limit $q_2\to0$ of the Macdonald measure [28], we get

$$\mathcal{Z}^{6D}(q_1,t_1;0,t_2) = \prod_{n,m,l=1}^\infty \frac{1}{1-q_1^{n-1/2}\,t_1^{m-1/2}\,t_2^{l-1}} \ .$$

This is the perturbative part of Nekrasov's partition function for seven-dimensional gauge theory compactified on a circle [110]. The Hall-Littlewood limit of the two-dimensional topological field theory underlying the computation of $\mathcal{N}=2$ superconformal indices in four dimensions is also studied in [18].

Notice also that in the limit of coincident deformation parameters $(q_1,t_1)=(q,t)=(q_2,t_2)$ we have

$$\log\mathcal{Z}^{6D}(q,t;q,t) = \sum_{k=1}^\infty \frac{q^{k/2}\,t^{k/2}}{k\,\big(1-t^k\big)^2} \ , \tag{6.18}$$

which is essentially the partition function corresponding to a single t-deformation equivalent to a q-deformation. The additional factor $q^{k/2}$ appearing in (6.18) is inconsequential as it can be removed if one uses, instead of the symmetric product form of the refined theta-function (6.14), the asymmetric product form

$$\widetilde{\Theta}\,'(z;q,t) = \prod_{n=1}^\infty \big(1+q^{n-1}\,z\big)\,\big(1+t^n\,z^{-1}\big) \ . \tag{6.19}$$

In this case the gauge theory partition function evaluates to the result (4.19) with deformation parameter t instead of q. Hence this limiting case is equivalent to the matrix model (4.16) whose weight function is the usual Jacobi theta-function (3.14) with the standard Haar measure; this is just another example of the cancellation of deformations/refinements that we first encountered in §4. Both theta-functions (6.14) and (6.19) give exactly the same result in the unrefined limit $q = t$.

6.6. **Refined disk amplitudes.**

Let us now consider the refined version of the Kostant identity (5.18). The modification of the Gaussian (5.11) of the category $\mathscr{R} = \mathscr{R}ep(\mathcal{U}_q(\mathfrak{g}))$ is given by

$$\Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{g})}\big(\bigcirc\!\!\!\!\cdot\,\big)\Big] = \sum_\lambda \frac{M_\lambda\big(t^\rho\,;\,q,t\big)}{\|M_\lambda\|_{q,t}}\ q^{\frac12\,\langle\lambda,\lambda\rangle}\ t^{\langle\lambda,\rho\rangle}\ \big[U_\lambda\otimes V_{\rm fund}^{\odot(\beta-1)\,N}\big]\ .$$

We can also consider generalized characters as functions on the Cartan subalgebra $\mathfrak{h}\subset\mathfrak{g}$ with values in the representation $V_{\rm fund}^{\odot(\beta-1)\,N}$ by restriction. For this, the sole modification of the ribbon category data that we require is that of the twist, for which we only need to incorporate the appropriate shift in weights $n\in\mathbb{Z}^N$. Using

$$\big\langle n+(\beta-1)\,\rho\,,\,n+(\beta-1)\,\rho\big\rangle = \big\langle n\,,\,n+2(\beta-1)\,\rho\big\rangle + (\beta-1)^2\,\big\langle\rho\,,\,\rho\big\rangle\ ,$$

after suitable normalization we find the refinement of the twist coefficients

$$\theta_n^\bullet = q^{\frac12\,\langle n,n\rangle}\ \Big(\frac{t}{q}\Big)^{\langle n,\rho\rangle}\ .$$

The corresponding modification of the Gaussian is thus given by

$$\Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{h})}\big(\bigcirc\!\!\!\!\cdot\,\big)\Big] = \sum_{n\in\mathbb{Z}^N} q^{\frac12\,\sum_i\,(n_i^2-(N+1-2i)\,n_i)}\ t^{\frac12\,\sum_i\,(N+1-2i)\,n_i}\ \big[U_n\otimes V_{\rm fund}^{\odot(\beta-1)\,N}\big]\ ,$$

where here the representation $V_{\rm fund}^{\odot(\beta-1)\,N}$ is regarded as a $\mathcal{U}_q(\mathfrak{h})$-module by restriction. Hence the modification of the Kostant identity (5.18) is given by [113]

$$\Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{g})}\big(\bigcirc\!\!\!\!\cdot\,\big)\Big] = \frac{1}{Z_N(q,t)}\ \Big[\mathscr{F}_{\mathcal{U}_q(\mathfrak{h})}\big(\bigcirc\!\!\!\!\cdot\,\big)\Big]\ , \tag{6.20}$$

where the normalization $Z_N(q,t)$ is the refined Chern-Simons partition function (6.11) on S^3. As we discuss below, the identity (6.20) may be regarded as a categorification of the generalization of the Weyl integral formula for G-equivariant functions on the Lie group G with values in the representation $V_{\rm fund}^{\odot(\beta-1)\,N}$, which was considered in [114].

Evaluating the generalized characters on both sides of (6.20) at the specialization $x = t^\rho$ yields a simple expression for the partition function (6.6) of the refined q-deformed gauge theory on the sphere S^2 given by

$$\mathcal{Z}^\bullet\big(\mathcal{U}_q(\mathfrak{g})\,;\,S^2\big) = \frac{1}{Z_N(q,t)}\ \prod_{j=1}^N \Theta\big(t^{N+1-2j}\,q^{-\frac12\,(N+1-2j)}\,;\,q\big)\ .$$

A similar formula is also derived in [19]; by using standard modular properties of the Jacobi theta-function $\Theta(z;q)$, it is related there to the Hirzebruch χ_y-genus of the moduli space of instantons on the toric surface $\mathcal{O}(-1)\to\mathbb{P}^1$, which arises in five-dimensional supersymmetric gauge theory; this can be thought of as a categorification of the usual Euler characteristic invariants computed by the $\mathcal{N}=4$ Vafa-Witten gauge theory.

6.7. Quantum gauge theory perspective.

We conclude by describing how to interpret the constructions of (q,t)-deformed Yang-Mills amplitudes of this section from a field theory point of view; this is also discussed in a similar vein but from different perspectives than ours in [24, 19]. For this, we will describe the physical states of refined q-deformed Yang-Mills theory in more detail through an operator formalism. We take the classical field theory to be unchanged as in [24, 19]. The Hilbert space of any topological field theory in two dimensions is based on a circle S^1. As in the case of the ordinary (unrefined and undeformed) two-dimensional Yang-Mills theory [27], canonical quantization of the field theory with action (2.4) shows that ϕ and A are canonically conjugate variables. If $C=S^1$ is an initial value circle in the Riemann surface $\Sigma^\bullet$ with a marked point, then the Hilbert space $\mathcal{H}_C$ obtained by canonical quantization on $\Sigma^\bullet$ consists of functionals of A in Schrödinger polarization, with A taken to be multiplication operators and ϕ acting as the functional derivative $\phi(x)=-\,\mathrm{i}\,\frac{\delta}{\delta A(x)}$. By gauge invariance, such functionals are the wavefunctions $\psi(U)$ depending only on the boundary holonomy $U=\mathcal{P}\exp\,\mathrm{i}\,\oint_C A\in G$ which are valued in the finite-dimensional unitary representation $V=V_{\rm fund}^{\odot(\beta-1)\,N}$ of $G=U(N)$. In particular, they are not conjugation invariant class functions of U, but rather define elements of the vector space $\Omega^0(G,V)^G$ of G-equivariant V-valued functions, i.e. $\psi(g\,U\,g^{-1})=g\triangleright\psi(U)$ for all $g,U\in G$. To define amplitudes of such states, we note that since V is a unitary representation, it has a natural G-invariant inner product $\langle -|-\rangle_V$, i.e. $\langle g\triangleright v|g\triangleright w\rangle_V=\langle v|w\rangle_V$ for all $g\in G$ and $v,w\in V$. Hence we can define an inner product on $\Omega^0(G,V)^G$ by

$$\langle\psi|\chi\rangle:=\int_G \mathrm{d}U\ \big\langle\psi(U)\,\big|\,\chi(U)\big\rangle_V\ . \tag{6.21}$$

This inner product defines the gluing rules for states associated to surfaces with boundary C. By the isomorphism $G/\mathrm{Ad(G)}=T/W$, we have

$$\Omega^0(G,V)^G=\Omega^0\big(T,V^{(0)}\big)^W\ . \tag{6.22}$$

Since the weight zero subspace $V^{(0)}\cong\mathbb{C}$ in the case at hand, this isomorphism identifies equivariant functions on G with symmetric functions on the maximal torus $T=U(1)^N$. As the scalar function $U\mapsto\langle\psi(U)|\chi(U)\rangle_V$ on G is conjugation invariant, we can apply the usual Weyl integral formula (2.13) to write the inner product (6.21) as

$$\langle\psi|\chi\rangle=\frac{1}{N!}\int_{[0,2\pi)^N}\ \prod_{i=1}^N\frac{\mathrm{d}\phi_i}{2\pi}\ \prod_{j<k}4\sin^2\Big(\frac{\phi_j-\phi_k}{2}\Big)\ \big\langle\psi(\phi)\,\big|\,\chi(\phi)\big\rangle_V\ . \tag{6.23}$$

Thus the Hilbert space of the quantum gauge theory

$$\mathcal{H}_C=L^2(G,V)^G$$

is the L^2-completion of the vector space $\Omega^0(G,V)^G$ with respect to the norm induced by the inner product (6.21). A natural basis for $\mathcal{H}_C$ is provided by generalised characters, and there is an analog of the Peter-Weyl theorem for generalised characters which decomposes the Hilbert space under the action of $G\times G$ into an orthogonal direct sum of finite-dimensional G-modules as [114]

$$\mathcal{H}_C=\bigoplus_\lambda\ \big(U_\lambda^*\otimes U_\lambda\otimes V\big)^G\ , \tag{6.24}$$

where the Hilbert space direct sum runs over the orthogonal subspaces $\mathsf{Hom}_G(U_\lambda,U_\lambda\otimes V)$ of intertwining operators Φ_λ with respect to the inner product (6.21). A generalization of the Weyl orthogonality theorem is proven in [104], whereby it is shown that the spaces of intertwining operators $\mathsf{Hom}_{\mathscr{R}}(U_\lambda,U_\lambda\otimes V)$ for the quantum group $\mathcal{U}_q(\mathfrak{g})$ form a mutually orthogonal system with respect to the inner product (6.23); this suggests an analog of the decomposition (6.24) of the Hilbert space of physical states into generalized characters $(U_\lambda^* \otimes U_\lambda \otimes V)^{\mathcal{U}_q(\mathfrak{g})}$. As discussed previously, when the highest-weight modules U_λ are regarded as quantum group representations, the basis of generalised characters for $\mathcal{H}_C$ coincide with the Macdonald polynomials $M_\lambda(u;q,t)$ in the holonomy eigenvalues $u \in (S^1)^N$, regarded as elements of $L^2\big(T, V^{(0)}\big)^W$ under the isomorphism (6.22), which are orthogonal with respect to the inner product (6.2). It follows that any physical state wavefunction ψ has an expansion in generalised characters as

$$\psi(u) = \sum_\lambda c_\lambda(q,t)\, M_\lambda(u;q,t)$$

where $c_\lambda(q,t) \in \mathbb{C}$.

Canonical quantization shows that the Hamiltonian of the gauge theory with action (2.4) is quadratic and given by

$$H = -\frac{g_s}{2} \oint_C \mathrm{d}\sigma \ \ \mathrm{Tr}\, \phi^2$$

where $\sigma \in S^1$ is the local coordinate of the initial value circle $C \subset \Sigma^\bullet$. On the Hilbert space $\mathcal{H}_C$, this operator acts via $\mathrm{Tr}\, \phi^2 = -\ \mathrm{Tr}\, \big(\frac{\delta}{\delta A}\big)^2 = -\ \mathrm{Tr}\, \big(U\, \frac{\partial}{\partial U}\big)^2$, and hence as usual the Hamiltonian operator is proportional to the Laplace-Beltrami operator on the group manifold of G. It is shown by [114] that every conjugation invariant scalar differential operator on G defines an operator on $\Omega^0(G,V)^G$ acting as a scalar on $\mathsf{Hom}_G(U_\lambda, U_\lambda \otimes V)$ for every irreducible representation λ, and using the isomorphism (6.22) this action can be rewritten in terms of differential operators on T with coefficients in $\mathsf{End}_T\,\big(V^{(0)}\big)$. In particular, the Laplace-Beltrami operator acts diagonally on the space $L^2(G,V)^G$ in the generalised characters M_λ with eigenvalue $C_2(\lambda)$ [114], and hence the Hamiltonian is diagonalised in this basis as the quadratic Casimir operator

$$H\, M_\lambda = \tfrac{g_s}{2}\, C_2(\lambda)\, M_\lambda\ . \tag{6.25}$$

Whence the quantum amplitudes, computed as matrix elements of the operator $\exp(-\tau\, H)$ between external states in $\mathcal{H}_C$, involves the standard heat kernel $\mathrm{e}^{\,-\tau\, g_s\, C_2(\lambda)/2}$ of the gauge group G. With $q = \mathrm{e}^{\,-g_s}$ and $t = q^\beta$, this heat kernel coincides (up to area-dependent renormalization ambiguities) with the twist eigenvalues (6.5) which are used in the building blocks of refined q-deformed Yang-Mills amplitudes.

The same spectrum (6.25) and Hilbert space (6.24) were obtained in a similar fashion in ordinary two-dimensional Yang-Mills theory by Gorsky and Nekrasov in [115, §1.3]; they compute the path integral for Yang-Mills theory on a cylinder cut by a Wilson line in the representation $V = V_{\mathrm{fund}}^{\odot(\beta-1)\,N}$ along the temporal direction. They further conjecture that the analogous calculation in the gauged G/G WZW model should be expressible through representations of the quantum group $\mathcal{U}_q(\mathfrak{g})$, with the presence of the Wilson line in the representation V yielding a deformation of the usual gauge group characters, i.e. the Schur polynomials, to Macdonald polynomials. Given the relation of this two-dimensional topological gauge theory to Chern-Simons theory on the corresponding trivial circle bundle, Iqbal and Kozcaz suggest in [23] that this deformed WZW theory could be related to refined Chern-Simons theory; an analogous relationship in the unrefined case is described in [116]. The Hamiltonian analysis presented here further agrees with the results of [15], where the decomposition (6.24) and the standard heat kernel (6.25) on G also appear in the refined two-dimensional Yang-Mills propagator that computes the partition function of an $\mathcal{N} = 2$ non-linear sigma-model on $S^1 \times S^3$.

This construction also illustrates how the refined partition functions $Z_{\mathrm{YM}}^{(p)}(q,t;\Sigma_h^\bullet)$ may be derived directly in the path integral formalism through the technique of diagonalisation. For fixed $\phi \in \Omega^0(\Sigma_h^\bullet, G)$, it follows from our description of the physical states of the refined gauge theory that we should extend the functional (2.8) to a $\mathcal{G}$-equivariant functional valued in $V = V_{\rm fund}^{\odot(\beta-1)\,N}$; in the diagonalisation formula that follows from (6.22), we should then employ the Macdonald inner product (6.2) appropriate to generalised characters associated with $\mathcal{U}_q(\mathfrak{g})$-modules. The path integral is thus given by the sum over torus bundles

$$\begin{aligned} Z_{\rm YM}^{(p)}(q,t;\Sigma_h^\bullet) &= \frac{1}{{\rm vol}(\mathcal{G})}\sum_{n\in\mathbb{Z}^N}\int_{\mathcal{A}_n}\mathscr{D}\mu[A^{\mathfrak{h}}]\int_{\Omega^1(\Sigma_h^\bullet,\mathcal{L}_n\times_T\mathfrak{k})}\mathscr{D}\mu[A^{\mathfrak{k}}] \\ &\quad\times\int_{\Omega^0(\Sigma_h^\bullet,(\mathbb{R}/2\pi\,\mathbb{Z})^N)}\prod_{i=1}^N\mathscr{D}\mu[\phi_i]\,\big[\Delta_{q,t}(\phi)\big]\,\exp\big(-S_{\rm BF}^{\bullet\,(p)}[\phi,A^{\mathfrak{h}},A^{\mathfrak{k}}]\big)\ , \end{aligned}$$

where

$$\begin{aligned} S_{\rm BF}^{\bullet\,(p)}[\phi,A^{\mathfrak{h}},A^{\mathfrak{k}}] &= \frac{1}{g_s}\sum_{i=1}^N\int_{\Sigma_h^\bullet}\Big(-\,{\rm i}\,\phi_i\,{\rm d}A_i^{\mathfrak{h}}+\frac{p}{2}\,\phi_i^2\,{\rm d}\mu\Big) \\ &\quad+\sum_{\alpha\in{\rm Ad}(G)}\int_{\Sigma_h^\bullet}\prod_{m=0}^{\beta-1}\Big(1-{\rm e}^{-m\,g_s}\,{\rm e}^{\,{\rm i}\,\langle\alpha,\phi\rangle}\Big)\,A_\alpha^{\mathfrak{k}}\wedge A_{-\alpha}^{\mathfrak{k}}\ . \end{aligned}$$

Proceeding as in §2, this sum yields the appropriate β-deformation

$$Z_{\rm YM}^{(p)}(q,t;\Sigma_h^\bullet)=\sum_{\mu\in\mathbb{Z}^N}\prod_{i<j}\prod_{m=0}^{\beta-1}[\mu_i-\mu_j+m]_q^{1-h}\,[\mu_i-\mu_j-m]_q^{1-h}\,\exp\Big(-\frac{p\,g_s}{2}\sum_{i=1}^N\mu_i^2\Big)$$

of the q-deformed discrete Gaussian matrix model (2.12). However, a complete Lagrangian or Hamiltonian description of the refined two-dimensional Yang-Mills theory, and of the related refinement of Chern-Simons theory on Seifert three-manifolds, is currently lacking, and it would be interesting to find a more precise and physical derivation of these quantum amplitudes from first principles. In particular, it is not clear at present how to properly incorporate the quantum group gauge symmetry based on $\mathcal{U}_q(\mathfrak{g})$ into the definition of the quantized gauge theory.

The construction presented here further elucidates the physical meaning of the class of deformed gauge theories introduced in §4.1, and in particular of Klimčík's partition function (2.14). It consists in using a Hamiltonian framework based on replacing differential operators with q-difference operators on $\Omega^0(T)$, generalizing the quantum torus deformation of §4.2, and G-modules with representations of the quantum group $\mathcal{U}_q(\mathfrak{g})$. Consider the operators $\hat u_i = z_i$ acting as multiplication by $z_i\in S^1$ and the q-difference operators $\hat v_i=\exp\big(-\log(q)\,z_i\,\frac{\partial}{\partial z_i}\big)$ for $i=1,\dots,N$. They obey the quantum N-torus algebra relations

$$\hat u_i\,\hat v_j = q\ \delta_{ij}\ \hat v_j\,\hat u_i$$

for $i,j=1,\dots,N$. It follows by [104, Prop. 6.2] that corresponding to the central quantum Casimir element $\mathtt{Cas}_q$ for $\mathcal{U}_q(\mathfrak{g})$ (see Appendix A) there exists a unique q-difference operator $\hat H_{\mathtt{Cas}_q}(\hat u,\hat v)$ whose spectrum consists of the twist eigenvalues θ_λ from §5.3, such that the generalized characters M_λ satisfy the difference equation

$$\hat H_{\mathtt{Cas}_q}(\hat u,\hat v)\,M_\lambda = q^{\frac12\,C_2(\lambda)}\,M_\lambda\ .$$

This spectrum generalizes (6.25) and the corresponding "heat kernel" yields q-deformed Casimir eigenvalues as in §4.1 and (2.14).

Appendix A. Quantum groups

The quantum universal enveloping algebra $\mathcal{U}_q(\mathfrak{g})$ for $\mathfrak{g} = \mathfrak{sl}(N)$ is defined as the unital algebra over $\mathbb{C}$ generated by elements $E_i, F_i, K_i^{\pm 1}$, $i = 1, \dots, N-1$, with the defining relations

$$K_i\,K_i^{-1} = 1 = K_i^{-1}\,K_i \qquad \text{and} \qquad \big[K_i\,,\,K_j^{\pm 1}\big] = 0\ ,$$
$$K_i\,E_i = q\ E_i\,K_i \qquad \text{and} \qquad K_i\,F_i = q^{-1}\ F_i\,K_i\ ,$$
$$K_i\,E_{i\pm1} = q^{-1/2}\ E_{i\pm1}\,K_i \qquad \text{and} \qquad K_i\,F_{i\pm1} = q^{1/2}\ F_{i\pm1}\,K_i\ ,$$
$$[K_i, E_j] = 0 \qquad \text{and} \qquad [K_i, F_j] = 0 \qquad \text{for} \quad j \neq i, i\pm1\ ,$$
$$[E_i, F_j] \;=\; \delta_{ij}\,\frac{K_i - K_i^{-1}}{q^{1/2} - q^{-1/2}}\ ,$$

together with the Serre relations

$$E_i^2\,E_j - \big(q^{1/2} + q^{-1/2}\big)\,E_i\,E_j\,E_i + E_j\,E_i^2 = 0 \qquad \text{for} \quad j = i\pm1\ ,$$
$$F_i^2\,F_j - \big(q^{1/2} + q^{-1/2}\big)\,F_i\,F_j\,F_i + F_j\,F_i^2 = 0 \qquad \text{for} \quad j = i\pm1\ ,$$
$$[E_i, E_j] = 0 = [F_i, F_j] \qquad \text{for} \quad j \neq i, i\pm1\ .$$

A vector space basis for $\mathcal{U}_q(\mathfrak{g})$ is given by the set $\{E_i^{n_i}\,K_i^{m_i}\,F_i^{l_i} \mid n_i, l_i \in \mathbb{Z}_{\geq 0}\ ,\ m_i \in \mathbb{Z}\ ,\ i = 1, \dots, N-1\}$. We usually take the deformation parameter $q \in \mathbb{R}$ to lie in the interval $0 < q < 1$ without loss of generality. The quantum Casimir element for $\mathcal{U}_q(\mathfrak{g})$ is given by

$$\mathtt{Cas}_q := \sum_{i=1}^{N-1}\,\Big(E_i\,F_i + \frac{q^{-1/2}\,K_i + q^{1/2}\,K_i^{-1}}{\big(q^{1/2} - q^{-1/2}\big)^2}\Big) = \sum_{i=1}^{N-1}\,\Big(F_i\,E_i + \frac{q^{1/2}\,K_i + q^{-1/2}\,K_i^{-1}}{\big(q^{1/2} - q^{-1/2}\big)^2}\Big)\ .$$

The element $\mathtt{Cas}_q$ is central in the algebra $\mathcal{U}_q(\mathfrak{g})$.

A Hopf algebra structure on $\mathcal{U}_q(\mathfrak{g})$ is provided by the coproduct $\Delta : \mathcal{U}_q(\mathfrak{g}) \to \mathcal{U}_q(\mathfrak{g}) \otimes \mathcal{U}_q(\mathfrak{g})$, the counit $\varepsilon : \mathcal{U}_q(\mathfrak{g}) \to \mathbb{C}$ and the antipode $S : \mathcal{U}_q(\mathfrak{g}) \to \mathcal{U}_q(\mathfrak{g})$ given on generators by

$$\Delta\big(K_i^{\pm1}\big) = K_i^{\pm1} \otimes K_i^{\pm1}\ , \qquad \varepsilon\big(K_i^{\pm1}\big) = 1 \qquad \text{and} \qquad S\big(K_i^{\pm1}\big) = K_i^{\mp1}\ ,$$
$$\Delta(E_i) = E_i \otimes K_i + 1 \otimes E_i\ , \qquad \varepsilon(E_i) = 0 \qquad \text{and} \qquad S(E_i) = -E_i\,K_i^{-1}\ ,$$
$$\Delta(F_i) = F_i \otimes 1 + K_i^{-1} \otimes F_i\ , \qquad \varepsilon(F_i) = 0 \qquad \text{and} \qquad S(F_i) = -K_i\,F_i\ ,$$

and extended as (anti-)algebra homomorphisms. The $*$-structure is the anti-algebra morphism given on generators by

$$K_i^* = K_i\ , \qquad E_i^* = F_i \qquad \text{and} \qquad F_i^* = E_i\ .$$

The Hopf algebra $\mathcal{U}_q(\mathfrak{g})$ has a quasitriangular structure defined by a universal R-matrix, which is an invertible element R in a certain completed tensor product algebra $\mathcal{U}_q(\mathfrak{g})\,\widehat{\otimes}\,\mathcal{U}_q(\mathfrak{g})$ that intertwines the coproduct Δ and the opposite coproduct $\Delta^{\rm op} := P \circ \Delta$ where P is the flip isomorphism $P(a \otimes b) = b \otimes a$ for $a, b \in \mathcal{U}_q(\mathfrak{g})$. It has the form

$$R = q^{\frac12\,\sum_i\,H_i \otimes H_i}\ R^\vee \qquad \text{with} \quad R^\vee \in \mathcal{U}_q^+(\mathfrak{g})\,\widehat{\otimes}\,\mathcal{U}_q^-(\mathfrak{g})\ ,$$

where H_i, $i = 1, \dots, N-1$ is an orthonormal basis for the Cartan subalgebra $\mathfrak{h} \subset \mathfrak{g}$ with respect to the invariant bilinear form $\mathrm{Tr}\,|_{\mathfrak{h}}$; we formally identify the group-like generators $K_i = q^{H_i/2}$. Here $\mathcal{U}_q^+(\mathfrak{g})$ (resp. $\mathcal{U}_q^-(\mathfrak{g})$) is the subalgebra of the quantum enveloping algebra $\mathcal{U}_q(\mathfrak{g})$ generated by $K_i^{\pm1}, E_i$ (resp. $K_i^{\pm1}, F_i$). The element $R^\vee$ satisfies

$$(\varepsilon \otimes 1)\big(R^\vee\big) = 1 \otimes 1 = (1 \otimes \varepsilon)\big(R^\vee\big)\ .$$

We use the standard Sweedler notation

$$R = R_{(1)} \otimes R_{(2)}$$

with implicit summation.

From the quasitriangular structure R, one constructs Drinfel'd's element u in a certain completion of $\mathcal{U}_q(\mathfrak{g})$ as

$$u := S\big(R_{(2)}\big)\, R_{(1)} \ .$$

It is invertible and has the property that $u\, S(u)$ is a central element with $S^2 = S \circ S$ acting as an inner automorphism

$$S^2(a) = u\, a\, u^{-1}$$

for all $a \in \mathcal{U}_q(\mathfrak{g})$, and moreover

$$\begin{aligned} R_{(2)}\, u\, R_{(1)} &= S\big(R_{(2)}\big)\, u\, S\big(R_{(1)}\big) = 1 \ , \\ \Delta(u) = (u \otimes u)\, \big(R_{(2)}\, R_{(1)} \otimes R_{(1)}\, R_{(2)}\big)^{-1} &= \big(R_{(2)}\, R_{(1)} \otimes R_{(1)}\, R_{(2)}\big)^{-1}\, (u \otimes u) \ . \end{aligned}$$

Appendix B. Toeplitz determinants

Let $f(z)$ be a complex-valued function on $\mathbb{C}$ with Laurent series expansion $f(z) = \sum_{k\in\mathbb{Z}} f_k\, z^k$, and let $T_N(f) = (f_{i-j})_{i,j=1,\dots,N}$ be the associated Toeplitz operator of dimension N and symbol f. By the Heine-Szegő identity, the corresponding Toeplitz determinant is the partition function of a $U(N)$ unitary matrix model

$$Z_N[f] := \det T_N(f) = \int_{[0,2\pi)^N} \prod_{i=1}^N \frac{\mathrm{d}\phi_i}{2\pi}\ f\big(\,\mathrm{e}^{\,\mathrm{i}\,\phi_i}\big) \prod_{j<k} \big|\,\mathrm{e}^{\,\mathrm{i}\,\phi_j} - \mathrm{e}^{\,\mathrm{i}\,\phi_k}\big|^2 \ . \tag{B.1}$$

Let $[\log f]_k$, $k \in \mathbb{Z}$ denote the coefficients in the Fourier series expansion on the unit circle S^1 of the logarithm of the symbol,

$$\log f(z) = \sum_{k=-\infty}^{\infty} [\log f]_k\ z^k \ ,$$

and suppose that they obey the absolute summability conditions

$$\sum_{k=-\infty}^{\infty} \big|\,[\log f]_k\,\big| < \infty \qquad \text{and} \qquad \sum_{k=-\infty}^{\infty} k\,\big|\,[\log f]_k\,\big|^2 < \infty \ .$$

Let $G(f) = \exp([\log f]_0)$ denote the geometric mean of the symbol f.

Then the strong Szegő limit theorem for Toeplitz determinants states [117, 118]

$$\lim_{N\to\infty} \frac{\det T_N(f)}{G(f)^N} = \exp\Big(\, \sum_{k=1}^{\infty} k\ [\log f]_k\ [\log f]_{-k}\,\Big) \ .$$

By the Heine-Szegő identity (B.1), the Szegő theorem is not only a statement about Toeplitz determinants but also about unitary matrix models. In particular, the strong Szegő theorem gives the partition function $Z_\infty[f]$ of a $U(\infty)$ unitary matrix model, defined as the $N \to \infty$ limit of (B.1). The strong Szegő theorem generally involves an exponentially small error term $O\big(\,\mathrm{e}^{\,-B\,N}\big)$ [118]; if $\log f(z)$ is real-valued and analytic in a neighbourhood of the unit circle $S^1 \subset \mathbb{C}$, then the error term is simply $O\big(\,\mathrm{e}^{\,-B\,N}\big)$, i.e. there are no $\frac{1}{N}$ corrections [118]. The analyticity condition holds if the Fourier coefficients $[\log f]_k$ have exponential decay as $k \to \infty$, which is precisely the case studied in this paper; although we have considered complex powers α_a in (4.21), a simple argument shows that the reality condition can also be relaxed [119]. If the function $f(z)$ is holomorphic, then the Szegő theorem is enough to determine the asymptotics.

For the more general case of a symbol with zeroes or poles, one has to use the more refined Fisher-Hartwig asymptotics.

The Heine-Szegő identity for generalized Toeplitz determinants reads as [49]

$$Z_N[f;q,t] := \det \widetilde{T}_N(f;q,t) = \int_{[0,2\pi)^N} \prod_{i=1}^N \frac{\mathrm{d}\phi_i}{2\pi} \, f\big(\mathrm{e}^{\,\mathrm{i}\,\phi_i}\big) \, \Delta_{q,t}\big(\mathrm{e}^{\,\mathrm{i}\,\phi}\big) \ , \tag{B.2}$$

where

$$\Delta_{q,t}(z) = \prod_{i<j} \left| \frac{\big(z_i\,z_j^{-1}\,;\,q\big)_\infty}{\big(t\,z_i\,z_j^{-1}\,;\,q\big)_\infty} \right|^2$$

is the Macdonald measure [28]. This class of integrals has been considered for a long time in the context of the Selberg integral [120, 121]. Under the same conditions on the Fourier coefficients of the logarithm of the symbol f, there is an extension of the Szegő theorem to this generalized case which reads as [49]

$$\lim_{N\to\infty} \frac{\det \widetilde{T}_N(f;q,t)}{G(f)^N} = \exp\Big(\sum_{k=1}^\infty k\ [\log f]_k\ [\log f]_{-k}\ \frac{1-q^k}{1-t^k}\Big) \ . \tag{B.3}$$

This is a statement about the $N\to\infty$ limit of a matrix model in the refined ensemble rather than the unitary ensemble, obtained as in (B.2) by replacing the usual $U(N)$ Haar measure in the matrix integral with the Macdonald measure.

Finite-dimensional Toeplitz determinants can be computed with the method of Day [122]; this problem has been revisited recently in the context of random matrix theory applications to L-functions in number theory. For this, let $R_1,R_2\in\mathbb{R}$ with $0\le R_1<R_2$. Consider complex polynomials

$$D(z) = \prod_{j=1}^k (z-\delta_j) \ , \qquad F(z) = \prod_{j=1}^h \big(1-\rho_j^{-1}\,z\big) \qquad \text{and} \qquad G(z) = \prod_{j=1}^p (z-r_j)$$

where the zeroes δ_i satisfy $|\delta_i|\le R_1$ for $i=1,\dots,k$, the zeroes ρ_j satisfy $|\rho_j|\ge R_2$ for $j=1,\dots,h$, and the zeroes $r_1,\dots,r_p$ are distinct. Let $f(z)=G(z)/F(z)\,D(z)$ on the annulus $\{z\in\mathbb{C}\mid R_1<|z|<R_2\}$. If $p=k+m$ with $m\ge h$, then

$$\det T_N(f) = (-1)^{m\,(N+1)} \sum_{\substack{I\subset\{1,\dots,k+m\}\\ |I|=m}} \ \prod_{i\in I}\ \prod_{s=1}^k\ \prod_{j\in\overline{I}}\ \prod_{t=1}^h r_i^{N+1}\, \frac{(r_i-\delta_s)\,(\rho_t-r_j)}{(r_i-r_j)\,(\rho_t-\delta_s)}$$

where $\overline{I}:=\{1,\dots,k+m\}\setminus I$.

Appendix C. Embedding theorems for abelian categories

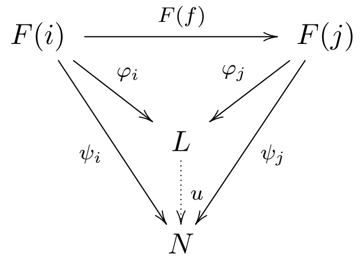
C.1. Ind-completions.

In category theory the notions of disjoint unions and direct sums are generalised to *colimits*, which are diagrams indexed by discrete categories. We describe here how to construct a category of all countable direct colimits in an abelian category which contains all necessary features justifying formal sums over infinite sets of simple objects; it is obtained by formally adjoining directed colimits. For further details and properties of the construction, see [123].

We begin by summarizing some of the basic notions that we need, beginning with that of filtered categories, which generalise the notion of directed set to category theory. A non-empty category I is *filtered* when:

(1) For every pair of objects $i, i' \in \mathsf{Ob}(I)$, there exists an object $k \in \mathsf{Ob}(I)$ and morphisms $(f : i \to k) \in \mathsf{Hom}_I(i, k)$ and $(f' : i' \to k) \in \mathsf{Hom}_I(i', k)$.
(2) For every pair of morphisms $(u, v : i \to j) \in \mathsf{Hom}_I(i, j)$, there exists an object $k \in \mathsf{Ob}(I)$ and a morphism $(w : j \to k) \in \mathsf{Hom}_I(j, k)$ such that $w \circ u = w \circ v$.

The actual objects and morphisms in I are largely irrelevant; only the ways in which they are interrelated above matters. Let $\mathscr{C}$ be a small abelian category enriched over $\mathscr{V}ect$, and I a filtered category. A functor $F : I \to \mathscr{C}$ is called a "diagram of type I"; a diagram can be thought of as indexing a collection of objects and morphisms of $\mathscr{C}$ patterned on the directed index category I. Diagrams are also called *direct systems.* A *co-cone* of a diagram $F : I \to \mathscr{C}$ is an object $N \in \mathsf{Ob}(\mathscr{C})$ together with a family of morphisms $(\psi_i : F(i) \to N) \in \mathsf{Hom}_{\mathscr{C}}(F(i), N)$, $i \in \mathsf{Ob}(I)$, such that $\psi_j \circ F(f) = \psi_i$ for all morphisms $(f : i \to j) \in \mathsf{Hom}_I(i, j)$. A *filtered colimit* of a diagram $F : I \to \mathscr{C}$ is co-cone (L, φ) of F which is universal: For any other co-cone (N, ψ), there exists a unique mediating morphism $u \in \mathsf{Hom}_{\mathscr{C}}(L, N)$ such that the diagrams



commute for all $i, j \in \mathsf{Ob}(I)$. Colimits are also called "direct limits" or "inductive limits"; if a diagram F has a colimit then it is unique up to unique isomorphism. Similarly, one defines limits by taking colimits in the corresponding dual categories.

Let $\mathscr{C}^\vee$ be the cocomplete functor category of $\mathbb{C}$-linear contravariant functors from $\mathscr{C}$ to the category $\mathscr{V}ect$ of complex vector spaces and linear transformations. There are two ways in which we can adjoin filtered colimits to the category $\mathscr{C}$: Either formally by regarding objects of the cocompletion as filtered diagrams in $\mathscr{C}$, or concretely as objects in $\mathscr{C}^\vee$ which are expressible as filtered colimits of representable functors. The Yoneda embedding $X \mapsto \mathsf{Hom}_{\mathscr{C}}(-, X)$ for $X \in \mathsf{Ob}(\mathscr{C})$ is a fully faithful exact functor sending $\mathscr{C} \hookrightarrow \mathscr{C}^\vee$; we identify $\mathscr{C}$ with this full subcategory of $\mathscr{C}^\vee$ in what follows. It has the properties

$$
\begin{aligned}
\mathsf{Hom}_{\mathscr{C}^\vee}\big(\mathsf{Hom}_{\mathscr{C}}(-, X)\,,\, F\big) &= F(X)\ , \\
\mathsf{Hom}_{\mathscr{C}^\vee}\big(\mathsf{Hom}_{\mathscr{C}}(-, X)\,,\, \mathsf{Hom}_{\mathscr{C}}(-, Y)\big) &= \mathsf{Hom}_{\mathscr{C}}(X, Y)
\end{aligned} \tag{C.1}
$$

for $F \in \mathsf{Ob}(\mathscr{C}^\vee)$ and $X, Y \in \mathsf{Ob}(\mathscr{C})$. An ind-object or formal inductive limit over $\mathscr{C}$ is a diagram $\underline{X} : I \to \mathscr{C}$, where I is a partially ordered directed set regarded as a small filtered category. We write $\underline{X} = (X_i)_{i \in \mathsf{Ob}(I)}$ where $X_i := \underline{X}(i) \in \mathsf{Ob}(\mathscr{C})$. Since the $\mathsf{Hom}_{\mathscr{C}}$ functor preserves all limits in $\mathscr{C}$, it relates colimits in $\mathscr{C}$ to colimits in $\mathscr{V}ect$ and the ind-object $\underline{X} = (X_i)_{i \in \mathsf{Ob}(I)}$ is uniquely isomorphic to the object of $\mathscr{C}^\vee$ of the form

$$
Y \longmapsto \mathsf{Hom}_{\mathscr{C}^\vee}\big(\mathsf{Hom}_{\mathscr{C}}(-, Y)\,,\, \underline{X}\big) = \varinjlim_{i \in \mathsf{Ob}(I)} \mathsf{Hom}_{\mathscr{C}}(Y, X_i)\ , \tag{C.2}
$$

which is natural in Y and respects colimiting cones; here the inductive limit is taken in the category of vector spaces. This notion of an ind-object is a refinement of the notion of colimit.

We write $\mathscr{I}nd(\mathscr{C})$ for the category whose objects are ind-objects over $\mathscr{C}$. Two ind-objects $\underline{X} : I \to \mathscr{C}$ and $\underline{Y} : J \to \mathscr{C}$ determine a functor $\mathsf{Hom}_{\mathscr{C}}(X_i, Y_j) : I^{\rm op} \times J \to \mathscr{V}ect$; hence by Yoneda's lemma and the construction of colimits in $\mathscr{C}^\vee$, we can use (C.1) and (C.2) to define the morphisms of $\mathscr{I}nd(\mathscr{C})$ by

$$\mathsf{Hom}_{\mathscr{I}nd(\mathscr{C})}\big(\,\underline{X}\,,\,\underline{Y}\,\big) = \varinjlim_{i\in\mathsf{Ob}(I)}\ \varprojlim_{j\in\mathsf{Ob}(J)}\ \mathsf{Hom}_{\mathscr{C}}(X_i,Y_j)\ ,$$

where again the limit and colimit are taken in the category $\mathscr{V}ect$. Then $\mathscr{I}nd(\mathscr{C})$ is a full abelian subcategory of $\mathscr{C}^\vee$, and the embedding $\mathscr{I}nd(\mathscr{C})\hookrightarrow\mathscr{C}^\vee$ preserves all limits. In fact, the functor $\underline{X}\mapsto\mathsf{Hom}_{\mathscr{I}nd(\mathscr{C})}(-,\underline{X}\,)$ is a fully faithful left-exact equivalence between $\mathscr{I}nd(\mathscr{C})$ and the abelian category $\mathscr{C}^{\vee,\mathrm{add}}$ of additive functors $\mathscr{C}^{\mathrm{op}}\to\mathscr{V}ect$. The ind-category $\mathscr{I}nd(\mathscr{C})$ is $\mathbb{C}$-linear and it has a concrete description in many cases; for example, if $\mathscr{C}$ is the category of finite-dimensional vector spaces, then the ind-category $\mathscr{I}nd(\mathscr{C})=\mathscr{V}ect$ may be identified with the category of all vector spaces $\mathscr{V}ect$ itself. The category $\mathscr{I}nd(\mathscr{C})$ has the usual nice properties which can be found in [123]; in particular, by (C.2) it is equivalent to the full subcategory of $\mathscr{C}^\vee$ consisting of functors which are filtered colimits of representable functors.

The Yoneda embedding identifies $\mathscr{C}$ with a full subcategory of its ind-category $\mathscr{I}nd(\mathscr{C})$. Then the objects $\mathsf{Ob}(\mathscr{C})$ are identified with the constant ind-objects: The map

$$\mathscr{Y}\,:\,\mathscr{C}\ \longrightarrow\ \mathscr{I}nd(\mathscr{C})\ ,\quad \mathscr{Y}(X)=(X_i)_{i\in I_\emptyset}\qquad\text{with}\quad I_\emptyset=\{\emptyset\}\ ,\ X_\emptyset=X\ , \tag{C.3}$$

is a natural $\mathbb{C}$-linear fully faithful exact functor. It preserves all limits which exist in $\mathscr{C}$. The category $\mathscr{I}nd(\mathscr{C})$ is cocomplete, i.e. it has all colimits, and in fact the full embedding (C.3) makes it into a *cocompletion* of the category $\mathscr{C}$: Every object of $\mathscr{I}nd(\mathscr{C})$ is a colimit of objects in the image of $\mathscr{Y}$. Note that the Yoneda embedding $\mathscr{C}\hookrightarrow\mathscr{C}^\vee$ is also a cocompletion of $\mathscr{C}$. What uniquely characterizes the ind-completion among all cocompletions of $\mathscr{C}$ is that all objects X in the image of (C.3) are finitely presentable in $\mathscr{I}nd(\mathscr{C})$, i.e. the functor $\mathsf{Hom}_{\mathscr{C}}(X,-):\mathscr{I}nd(\mathscr{C})\to\mathscr{V}ect$ preserves partially ordered directed colimits.

If $\mathscr{C}$ is in addition a monoidal category, then its tensor structure extends to $\mathscr{I}nd(\mathscr{C})$ in the following way. Let $\underline{X}=(X_i)_{i\in\mathsf{Ob}(I)}$ and $\underline{Y}=(Y_j)_{j\in\mathsf{Ob}(J)}$ be ind-objects over $\mathscr{C}$. Using the exterior product bifunctor $\otimes:\mathscr{C}\times\mathscr{C}\to\mathscr{C}$ we define $\otimes:\mathscr{I}nd(\mathscr{C})\times\mathscr{I}nd(\mathscr{C})\to\mathscr{I}nd(\mathscr{C})$ by the formula $\underline{X}\otimes\underline{Y}:=(X_i\otimes Y_j)_{i\in\mathsf{Ob}(I),j\in\mathsf{Ob}(J)}$. Given another ind-object $\underline{Z}=(Z_k)_{k\in\mathsf{Ob}(K)}$, there are functorial isomorphisms $(X_i\otimes Y_j)\otimes Z_k\to X_i\otimes(Y_j\otimes Z_k)$ which induce an isomorphism between the ind-objects $(\,\underline{X}\otimes\underline{Y}\,)\otimes\underline{Z}:=((X_i\otimes Y_j)\otimes Z_k)_{i\in\mathsf{Ob}(I),j\in\mathsf{Ob}(J),k\in\mathsf{Ob}(K)}$ and $\underline{X}\otimes(\,\underline{Y}\otimes\underline{Z}\,):=(X_i\otimes(Y_j\otimes Z_k))_{i\in\mathsf{Ob}(I),j\in\mathsf{Ob}(J),k\in\mathsf{Ob}(K)}$. One can extend other structures on the category $\mathscr{C}$ to its ind-completion $\mathscr{I}nd(\mathscr{C})$ in similar ways; in particular, in this manner it is possible to define notions of ind-algebras, ind-modules, and so on.

C.2. **Morita equivalence.**

Let $\mathscr{M}od_{\mathscr{C}}(A)$ be the module category over an algebra object A of a semisimple monoidal category $\mathscr{C}$. Let $\mathtt{M}_1,\mathtt{M}_2\in\mathsf{Ob}(\mathscr{M}od_{\mathscr{C}}(A))$. Since the functor $X\mapsto\mathsf{Hom}_A(\mathtt{M}_1\otimes X,\mathtt{M}_2)$ is (right) exact, it has a left adjoint $\mathscr{H}om_{\mathscr{C}}(M_1,M_2)$ called the internal Hom from $\mathtt{M}_1$ to $\mathtt{M}_2$. This is an ind-object of $\mathscr{C}$ representing this functor, and it is an internal version of the "space of morphisms from $\mathtt{M}_1$ to $\mathtt{M}_2$". When both categories $\mathscr{C}$ and $\mathscr{M}od_{\mathscr{C}}(A)$ have finitely-many simple objects, then $\mathscr{H}om_{\mathscr{C}}(M_1,M_2)\in\mathsf{Ob}(\mathscr{C})$ is an object of $\mathscr{C}$ which is uniquely defined up to isomorphism by Yoneda's lemma, and so $\mathscr{H}om_{\mathscr{C}}(-,-)$ is a bifunctor. Then the internal Hom is defined by the relation

$$\mathsf{Hom}_{\mathscr{C}}\big(X\,,\,\mathscr{H}om_{\mathscr{C}}(M_1,M_2)\big):=\mathsf{Hom}_A(\mathtt{M}_1\otimes X,\mathtt{M}_2) \tag{C.4}$$

for all objects $X\in\mathsf{Ob}(\mathscr{C})$.

By (C.4) there is an isomorphism

$$\mathsf{Hom}_{\mathscr{C}}\big(\mathscr{H}om_{\mathscr{C}}(M_1,M_2)\,,\,\mathscr{H}om_{\mathscr{C}}(M_1,M_2)\big)=\mathsf{Hom}_A\big(\mathtt{M}_1\otimes\mathscr{H}om_{\mathscr{C}}(M_1,M_2)\,,\,\mathtt{M}_2\big)\ .$$

We define a canonical evaluation morphism

$$\mathrm{ev}_{M_1,M_2} \, : \, \mathtt{M}_1 \otimes \mathscr{H}om_{\mathscr{C}}(M_1,M_2) \ \longrightarrow \ \mathtt{M}_2$$

as the image of $\mathrm{id}_{\mathscr{H}om_{\mathscr{C}}(M_1,M_2)}$ under this isomorphism. Now let $\mathtt{M}_1,\mathtt{M}_2,\mathtt{M}_3 \in \mathsf{Ob}(\mathscr{M}od_{\mathscr{C}}(A))$. Then the sequence of morphisms

$$\begin{aligned} \mathtt{M}_1 \otimes \big(\mathscr{H}om_{\mathscr{C}}(M_1,M_2) \otimes \mathscr{H}om_{\mathscr{C}}(M_2,M_3)\big) \ &= \ \big(\mathtt{M}_1 \otimes \mathscr{H}om_{\mathscr{C}}(M_1,M_2)\big) \otimes \mathscr{H}om_{\mathscr{C}}(M_2,M_3) \\ &\xrightarrow{\mathrm{ev}_{M_1,M_2}\otimes \mathrm{id}_{\mathscr{H}om_{\mathscr{C}}(M_2,M_3)}} \ \mathtt{M}_2 \otimes \mathscr{H}om_{\mathscr{C}}(M_2,M_3) \ \xrightarrow{\mathrm{ev}_{M_2,M_3}} \ \mathtt{M}_3 \end{aligned}$$

defines a canonical composition morphism

$$\mathscr{H}om_{\mathscr{C}}(M_1,M_2) \otimes \mathscr{H}om_{\mathscr{C}}(M_2,M_3) \ \longrightarrow \ \mathscr{H}om_{\mathscr{C}}(M_1,M_3) \ . \tag{C.5}$$

This multiplication is associative and compatible with isomorphisms involving the internal Hom.

Let us now fix a particular non-zero A-module $\mathtt{M} \in \mathsf{Ob}(\mathscr{M}od_{\mathscr{C}}(A))$. Then the multiplication morphism (C.5) defines an algebra structure on the object

$$A_M := \mathscr{H}om_{\mathscr{C}}(M,M) \tag{C.6}$$

of the category $\mathscr{C}$. If $\mathscr{M}od_{\mathscr{C}}(A)$ is a semisimple indecomposable module category, then $A_M \in \mathsf{Ob}(\mathscr{C})$ is a semisimple indecomposable algebra.

Define a functor $\mathscr{F} : \mathscr{M}od_{\mathscr{C}}(A) \to \mathscr{C}$ by

$$\mathtt{N} \ \longmapsto \ \mathscr{H}om_{\mathscr{C}}(M,N) \qquad \text{for} \quad \mathtt{N} \in \mathsf{Ob}\big(\mathscr{M}od_{\mathscr{C}}(A)\big) \ . \tag{C.7}$$

The multiplication morphism (C.5) defines the structure of a left A_M-module on $\mathscr{H}om_{\mathscr{C}}(M,N)$, and hence (C.7) restricts to a functor $\mathscr{F} : \mathscr{M}od_{\mathscr{C}}(A) \to \mathscr{M}od_{\mathscr{C}}(A_M)$. Furthermore, for $X,Y \in \mathsf{Ob}(\mathscr{C})$ one has, by (C.4) together with the canonical isomorphisms (5.3) and $X = ({}^\vee X)^\vee$, the isomorphisms

$$\begin{aligned} \mathsf{Hom}_{\mathscr{C}}\big(Y\,,\,\mathscr{H}om_{\mathscr{C}}(M,N\otimes X)\big) \ &= \ \mathsf{Hom}_A(\mathtt{M}\otimes Y, \mathtt{N}\otimes X) \\ &= \ \mathsf{Hom}_A\big((\mathtt{M}\otimes Y)\otimes {}^\vee X\,,\,\mathtt{N}\big) \\ &= \ \mathsf{Hom}_A\big(\mathtt{M}\otimes (Y\otimes {}^\vee X\,)\,,\,\mathtt{N}\big) \\ &= \ \mathsf{Hom}_{\mathscr{C}}\big(Y\otimes {}^\vee X\,,\,\mathscr{H}om_{\mathscr{C}}(M,N)\big) \\ &= \ \mathsf{Hom}_{\mathscr{C}}\big(Y\,,\,\mathscr{H}om_{\mathscr{C}}(M,N)\otimes X\big) \ , \end{aligned}$$

and hence the canonical isomorphism

$$\mathscr{H}om_{\mathscr{C}}(M,N\otimes X) = \mathscr{H}om_{\mathscr{C}}(M,N)\otimes X \tag{C.8}$$

for all $X \in \mathsf{Ob}(\mathscr{C})$. This isomorphism defines a structure of a module functor on the functor $\mathscr{F} : \mathscr{M}od_{\mathscr{C}}(A) \to \mathscr{M}od_{\mathscr{C}}(A_M)$ given by (C.7), by compatibility of the multiplication (C.5) with (C.8). Using standard homological algebra, one then proves [85] that this functor is an equivalence of module categories.

It follows that for any two A-modules $\mathtt{M}_1,\mathtt{M}_2 \in \mathsf{Ob}(\mathscr{M}od_{\mathscr{C}}(A))$, the module categories of A_{M_1} and A_{M_2} are equivalent. Hence the algebras A_{M_1} and A_{M_2} in $\mathscr{C}$ are *Morita equivalent*. By the explicit construction (C.7) of the functor $\mathscr{F}$ and (C.5), the Morita equivalence bimodules are given explicitly by the objects $\mathscr{H}om_{\mathscr{C}}(M_1,M_2)$ and $\mathscr{H}om_{\mathscr{C}}(M_2,M_1)$ of $\mathscr{C}$.

Consider the case where $\mathtt{M} = \mathtt{A} := (A,\mu)$ is the trivial A-bimodule. That the algebra object A is haploid is equivalent to the statement that A is simple as a bimodule $\mathtt{A}$ over itself [124]. Then $A_A = \mathscr{H}om_{\mathscr{C}}(A,A)$. The unit morphism $\eta \in \mathsf{Hom}_{\mathscr{C}}(\mathbb{1},A)$ defines a canonical isomorphism

$\mathsf{Hom}_A(\mathtt{A},\mathtt{M}) = \mathsf{Hom}_{\mathscr{C}}(\mathbb{1}, M)$ for any A-module $\mathtt{M}$. Then for any object $X \in \mathsf{Ob}(\mathscr{C})$ one has the canonical isomorphism

$$\mathsf{Hom}_A(\mathtt{A} \otimes X, \mathtt{M}) = \mathsf{Hom}_A(\mathtt{A}, \mathtt{M} \otimes X^\vee) = \mathsf{Hom}_{\mathscr{C}}(\mathbb{1}, M \otimes X^\vee) = \mathsf{Hom}_{\mathscr{C}}(X, M) \ . \tag{C.9}$$

This isomorphism is generated by the reciprocity map $f \mapsto \varrho \circ (\mathrm{id}_A \otimes f)$ for $f \in \mathsf{Hom}_{\mathscr{C}}(X, M)$. Using duality one similarly has a canonical isomorphism

$$\mathsf{Hom}_A(\mathtt{M}, \mathtt{A} \otimes X) = \mathsf{Hom}_{\mathscr{C}}(M, X) \tag{C.10}$$

generated by $f \mapsto (\mathrm{id}_A \otimes (f \circ \varrho)) \circ ((\Delta \circ \eta) \otimes \mathrm{id}_M)$ for $f \in \mathsf{Hom}_{\mathscr{C}}(M, X)$. From the definition (C.4) it then follows that there is a natural isomorphism

$$\mathscr{H}om_{\mathscr{C}}(A, M) = M$$

and in particular

$$A_A = \mathscr{H}om_{\mathscr{C}}(A, A) = A \ ,$$

as expected.

It follows from these properties of the internal Hom that there is a natural identification $\mathscr{H}om_{\mathscr{C}}(M_1, M_2) = {}^\vee\mathtt{M}_1 \otimes_A \mathtt{M}_2$. In particular, $A_M = {}^\vee\mathtt{M} \otimes_A \mathtt{M}$, and the Morita equivalence bimodules above are ${}^\vee\mathtt{M}_1 \otimes_{A_{M_1}} \mathtt{M}_2 \cong A_{M_2}$ and ${}^\vee\mathtt{M}_2 \otimes_{A_{M_2}} \mathtt{M}_1 \cong A_{M_1}$. (See [101, Def. 5.2] for the definition of the tensor product $\mathtt{N} \otimes_A \mathtt{M}$ of a right A-module $\mathtt{N}$ and a left A-module $\mathtt{M}$.) This is useful in explicit calculations, and it works whenever both $\mathscr{C}$ and the module category $\mathscr{M}od_{\mathscr{C}}(A)$ are semisimple with finitely-many simple objects, so that the internal Hom is always an object of $\mathscr{C}$. In more general cases, one needs to work with ind-objects of the category $\mathscr{C}$ as discussed above. In particular, if $\mathscr{C}$ has infinitely many simple objects, then one can establish a similar Morita equivalence result by working with ind-algebras and ind-modules [85].

C.3. **Freyd-Mitchell embedding theorem.**

It is a classical result in category theory that every small abelian category is equivalent to a full subcategory of the representation category of modules over some ring $\mathcal{A}$. If the category in question $\mathscr{C}$ is enriched over $\mathscr{V}ect$, then $\mathcal{A}$ can be taken to be an algebra. In the semisimple case, with $\mathscr{C}$ containing a finite number N of simple objects, we can easily construct an equivalent category of modules, simply with as many irreducible representations as the category at hand. By Wedderburn theory, finite-dimensional semisimple algebras are essentially direct sums of full matrix algebras: The algebra $\mathbb{M}_n$ of $n \times n$ complex matrices contains no non-trivial minimal ideals. We can thus obtain N irreducible representations by taking $\mathcal{A}$ to be a direct sum of full matrix algebras over $\mathbb{C}$ as

$$\mathcal{A} = \bigoplus_{i=0}^{N-1} \mathbb{M}_{n_i} \ .$$

Each summand has only one irreducible representation, an n_i-dimensional complex vector space $\mathbb{C}^{n_i}$. This construction is not canonical since we are free to choose the dimensions n_i as we wish.

The ring in question is formed by taking the endomorphism ring $\mathcal{A} = \mathsf{End}_{\mathscr{M}}(K)$ of some object K which is an injective cogenerator in some functor category $\mathscr{M}$. This means that every object injects in $K^{\oplus n}$ for suitable n (depending on the object). In the semisimple case, with $\mathscr{C}$ containing finitely many isomorphism classes of simple objects, we can actually apply the argument inside the category $\mathscr{C}$ itself, as then the sum K of simple objects is an injective cogenerator: Every object is a sum of simple objects, and so will inject into a sufficiently big sum of copies of K. In more general cases, e.g. when $\mathscr{C}$ has infinitely many simple objects, one needs to work with ind-objects of the category $\mathscr{C}$ and pass to an associated A_∞-category $\mathscr{M}$; we shall return to this point later on. For the moment, we briefly work through some of the details of the construction of the ring $\mathcal{A}$; for more details, see [125, §4.4] and [126, §IV.4].

We begin with some standard category theory definitions. We call an object $K \in \mathsf{Ob}(\mathscr{M})$ in an abelian category $\mathscr{M}$ injective if the contravariant functor

$$\mathscr{H} := \mathsf{Hom}_{\mathscr{M}}(-,K) \, : \, \mathscr{M} \longrightarrow \mathscr{V}ect$$

is coexact, i.e. it carries exact sequences into exact sequences (with the arrows reversed). We say that K is a cogenerator if the functor $\mathscr{H}$ is an embedding, i.e. for $X,Y \in \mathsf{Ob}(\mathscr{M})$, the map $\mathsf{Hom}_{\mathscr{M}}(X,Y) \to \mathsf{Hom}_{\mathscr{V}ect}(\mathscr{H}(X),\mathscr{H}(Y))$ is injective. A collection of objects $(K_i)_{i\in I}$ of $\mathscr{M}$ is called a family of cogenerators for $\mathscr{M}$ if for every $X,Y \in \mathsf{Ob}(\mathscr{M})$ and every non-zero morphism $\alpha \in \mathsf{Hom}_{\mathscr{M}}(X,Y)$ there exists $\kappa_i \in \mathsf{Hom}_{\mathscr{M}}(Y,K_i)$ such that $\kappa_i \circ \alpha \neq 0$. Note that K is a cogenerator if (K) is a family of cogenerators for $\mathscr{M}$. If

$$K = \bigoplus_{i\in I} K_i$$

and $\mathsf{Hom}_{\mathscr{M}}(X,K_i) \neq \emptyset$ for all $i \in I$ and $X \in \mathsf{Ob}(\mathscr{M})$, then K is a cogenerator for $\mathscr{M}$ if and only if $(K_i)_{i\in I}$ is a family of cogenerators for $\mathscr{M}$. Below we will use the fact that it suffices to construct ind-objects on cogenerators [123].

Now let $\mathcal{A}$ be an associative algebra over $\mathbb{C}$, and let $\mathscr{R}ep(\mathcal{A})$ denote the representation category of left $\mathcal{A}$-modules. Then $\mathcal{A}$ is an injective cogenerator for $\mathscr{R}ep(\mathcal{A})$: The contravariant functor

$$\mathsf{Hom}_{\mathscr{R}ep(\mathcal{A})}(-,\mathcal{A}) \, : \, \mathscr{R}ep(\mathcal{A}) \longrightarrow \mathscr{V}ect \ ,$$

with $\mathcal{A}$ regarded as the trivial left $\mathcal{A}$-module, is the “forgetful” functor assigning to each $\mathcal{A}$-module V the underlying complex vector space (forgetting that V is an $\mathcal{A}$-module). Thus any category equivalent to $\mathscr{R}ep(\mathcal{A})$ also has an injective cogenerator. We now show that the converse is also true, i.e. if $\mathscr{M}$ is an abelian category enriched over $\mathscr{V}ect$ which possesses an injective cogenerator K, then $\mathscr{M}$ is equivalent to the representation category $\mathscr{R}ep(\mathcal{A})$ for some algebra $\mathcal{A}$.

Let

$$\mathcal{A} := \mathsf{End}_{\mathscr{M}}(K) = \mathsf{Hom}_{\mathscr{M}}(K,K) \tag{C.11}$$

be the $\mathbb{C}$-vector space of endomorphisms of $K \in \mathsf{Ob}(\mathscr{M})$. For every $X \in \mathsf{Ob}(\mathscr{M})$, the $\mathbb{C}$-vector space $\mathsf{Hom}_{\mathscr{M}}(X,K)$ has a canonical $\mathcal{A}$-module structure: For $\alpha \in \mathsf{Hom}_{\mathscr{M}}(X,K)$ and $\rho \in \mathcal{A}$, define $\rho \triangleright \alpha \in \mathsf{Hom}_{\mathscr{M}}(X,K)$ to be the composition $\rho \circ \alpha$. We may thereby define the functor

$$\mathscr{T} \, : \, \mathscr{M} \longrightarrow \mathscr{R}ep(\mathcal{A}) \ , \qquad \mathscr{T}(X) = \mathsf{Hom}_{\mathscr{M}}(X,K) \ . \tag{C.12}$$

Using the fact that K is an injective cogenerator, one shows that $\mathscr{T}$ is an equivalence of abelian categories, such that the map $\mathsf{Hom}_{\mathscr{M}}(X,Y) \to \mathsf{Hom}_{\mathscr{R}ep(\mathcal{A})}(\mathscr{T}(X),\mathscr{T}(Y))$ induced by $\mathscr{T}$ is an isomorphism whenever X is finitely-generated. By Yoneda's lemma, $\mathscr{T}$ is a left-exact functor, and the assignment $X \mapsto \mathsf{Hom}_{\mathscr{M}}(X,-)$ yields a duality between $\mathscr{R}ep(\mathcal{A})$ and a subcategory of the category $\mathscr{M}^\vee$ of left-exact functors $\mathscr{M} \to \mathscr{V}ect$.

If the category $\mathscr{C}$ is semisimple, abelian and finite with simple objects U_i, $i \in I$, then we may take $\mathscr{M} = \mathscr{C}$ and

$$K = \bigoplus_{i\in I} U_i \tag{C.13}$$

is an injective cogenerator for $\mathscr{C}$. Hence the associative $\mathbb{C}$-algebra (C.11) and the equivalence of abelian categories (C.12) are given explicitly by

$$\mathcal{A} = \bigoplus_{i,j\in I} \mathsf{Hom}_{\mathscr{C}}(U_i,U_j) \cong \mathbb{C}^{|I|} \qquad \text{and} \qquad \mathscr{T}(X) = \bigoplus_{i\in I} \mathsf{Hom}_{\mathscr{C}}(X,U_i) \tag{C.14}$$

for $X \in \mathsf{Ob}(\mathscr{C})$. As mentioned before, the algebra $\mathcal{A}$ is a non-canonical object as it is only defined up to Morita equivalence.

If the category $\mathscr{C}$ is not semisimple, then one requires the general construction of the Freyd-Mitchell embedding theorem [125, 126]. For this, let $\mathscr{M} = \mathscr{C}^\vee$ be the category of left-exact functors $\mathscr{C} \to \mathscr{V}ect$, and construct an injective cogenerator K with endomorphism algebra (C.11). Define a functor $\mathscr{T} : \mathscr{C} \to \mathscr{R}ep(\mathcal{A})$ by $\mathscr{T}(X) := \mathsf{Hom}_{\mathscr{C}^\vee}\big(\mathsf{Hom}_{\mathscr{C}}(X,-), K\big)$ for $X \in \mathsf{Ob}(\mathscr{C})$. Then $\mathscr{T}$ is a fully faithful exact functor which yields an equivalence between $\mathscr{C}$ and a subcategory of $\mathscr{R}ep(\mathcal{A})$. If $\mathscr{C}$ contains infinitely many simple objects U_i, $i \in I$, then the cogenerator K may be taken to be the ind-object $\underline{U} = (U_i)_{i\in I}$; in this instance the functor $\mathscr{T}$ coincides with the presentation of $\underline{U}$ given by (C.2).

We now consider the case where $\mathscr{M} = \mathscr{M}od_{\mathscr{C}}(A)$ is the module category of an algebra object $A \in \mathsf{Ob}(\mathscr{C})$ in a tensor category. For $U \in \mathsf{Ob}(\mathscr{C})$, the induced A-module is the (left) A-module $\mathtt{A} \otimes U = (A \otimes U, \mu \otimes \mathrm{id}_U)$. This defines an induction functor $\mathscr{C} \to \mathscr{M}od_{\mathscr{C}}(A)$. There is also a restriction functor $\mathscr{M}od_{\mathscr{C}}(A) \to \mathscr{C}$ given by the forgetful map $\mathtt{M} \mapsto M$ on objects. Induced A-modules have the useful computational properties [101, Lem. 5.8]

$$\mathtt{N} \otimes_A (\mathtt{A} \otimes X) \cong \mathtt{N} \otimes X \tag{C.15}$$

as A-modules, for every right A-module $\mathtt{N}$ and every $X \in \mathsf{Ob}(\mathscr{C})$, and also

$$(\mathtt{A} \otimes X) \otimes_A (\mathtt{A} \otimes Y) = \mathtt{A} \otimes (X \otimes Y) \tag{C.16}$$

for all $X, Y \in \mathsf{Ob}(\mathscr{C})$. When A is a Frobenius algebra, every module $\mathtt{M}$ over A is a submodule of an induced module [101, Lem. 5.23], because there is an injection sending $\mathtt{M} \mapsto \mathtt{A} \otimes M$. If the tensor category $\mathscr{C}$ is semisimple, abelian and finite with simple objects U_i, $i \in I$, and the Frobenius algebra A is special, then

$$\mathtt{K} = \mathtt{A} \otimes \Big(\bigoplus_{i\in I} U_i\Big) \tag{C.17}$$

is an injective cogenerator for $\mathscr{M}od_{\mathscr{C}}(A)$. Then the endomorphism ring $\mathcal{A} = \mathsf{End}_A(\mathtt{K})$ of the A-module $\mathtt{K}$ is the (semisimple) associative $\mathbb{C}$-algebra we are looking for; the equivalence of abelian categories given by (C.12), $\mathscr{M} = \mathscr{M}od_{\mathscr{C}}(A) = \mathscr{R}ep(\mathcal{A})$, can then be used to induce the structure of a module category over $\mathscr{C}$ on $\mathscr{R}ep(\mathcal{A})$. Using (C.9) and (C.10), we can write the equivalence between abelian categories explicitly in terms of morphism spaces in the *original* tensor category $\mathscr{C}$ as

$$\mathcal{A} = \bigoplus_{i,j\in I} \mathsf{Hom}_{\mathscr{C}}(U_i, A \otimes U_j) \qquad \text{and} \qquad \mathscr{T}(\mathtt{M}) = \bigoplus_{i\in I} \mathsf{Hom}_{\mathscr{C}}(M, U_i) \tag{C.18}$$

for $\mathtt{M} \in \mathsf{Ob}(\mathscr{M}od_{\mathscr{C}}(A))$. If the Frobenius algebra A is not special, then the category $\mathscr{M}od_{\mathscr{C}}(A)$ need not be semisimple; in these instances we need the general construction of the Freyd-Mitchell embedding theorem combined with the ind-object treatment as above.

References

[1] S. Cordes, G. W. Moore and S. Ramgoolam, "Lectures on $2D$ Yang-Mills theory, equivariant cohomology and topological field theories," Nucl. Phys. Proc. Suppl. **41** (1995) 184–244 [arXiv:hep-th/9411210].

[2] A. A. Migdal, "Recursion equations in gauge field theories," Sov. Phys. JETP **42** (1975) 413–418 [Zh. Eksp. Teor. Fiz. **69** (1975) 810–822].

[3] B. E. Rusakov, "Loop averages and partition functions in $U(N)$ gauge theory on two-dimensional manifolds," Mod. Phys. Lett. A **5** (1990) 693–703.

[4] E. Witten, "On quantum gauge theories in two dimensions," Commun. Math. Phys. **141** (1991) 153–209.

[5] E. Buffenoir and P. Roche, "Two-dimensional lattice gauge theory based on a quantum group," Commun. Math. Phys. **170** (1995) 669–698 [arXiv:hep-th/9405126].

[6] C. Klimčík, "The formulae of Kontsevich and Verlinde from the perspective of the Drinfel'd double," Commun. Math. Phys. **217** (2001) 203–228 [arXiv:hep-th/9911239].

[7] M. Aganagic, H. Ooguri, N. Saulina and C. Vafa, "Black holes, q-deformed $2D$ Yang-Mills, and non-perturbative topological strings," Nucl. Phys. B **715** (2005) 304–348 [arXiv:hep-th/0411280].

[8] C. Beasley and E. Witten, "Non-abelian localization for Chern-Simons theory," J. Diff. Geom. **70** (2005) 183–323 [arXiv:hep-th/0503126].

[9] N. Caporaso, M. Cirafici, L. Griguolo, S. Pasquetti, D. Seminara and R. J. Szabo, "Topological strings and large N phase transitions I: Non-chiral expansion of q-deformed Yang-Mills theory," J. High Energy Phys. **0601** (2006) 035 [arXiv:hep-th/0509041].

[10] M. Blau and G. Thompson, "Chern-Simons theory on S^1-bundles: Abelianisation and q-deformed Yang-Mills theory," J. High Energy Phys. **0605** (2006) 003 [arXiv:hep-th/0601068]

[11] 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," Nucl. Phys. B **772** (2007) 1–24 [arXiv:hep-th/0610155].

[12] G. Thompson, "Intersection pairings on spaces of connections and Chern-Simons theory on Seifert manifolds," AMS/IP Studies Adv. Math. **50** (2011) 317–330 [arXiv:1001.2885 [math.DG]].

[13] K. Ohta and Y. Yoshida, "Nonabelian localization for supersymmetric Yang-Mills-Chern-Simons theories on Seifert manifold," Phys. Rev. D **86** (2012) 105018 [arXiv:1205.0046 [hep-th]].

[14] A. Gadde, L. Rastelli, S. S. Razamat and W. Yan, "The $4D$ superconformal index from q-deformed $2D$ Yang-Mills," Phys. Rev. Lett. **106** (2011) 241602 [arXiv:1104.3850 [hep-th]].

[15] Y. Tachikawa, "$4D$ partition function on $S^1 \times S^3$ and $2D$ Yang-Mills with non-zero area," Progr. Theor. Exp. Phys. **2013** (2013) 013B01 [arXiv:1207.3497 [hep-th]].

[16] Y. Fukuda, T. Kawano and N. Matsumiya, "$5D$ SYM and $2D$ q-deformed YM," Nucl. Phys. B **869** (2013) 493–522 [arXiv:1210.2855 [hep-th]].

[17] D. Gaiotto, "$\mathcal{N} = 2$ dualities," J. High Energy Phys. **1208** (2012) 034 [arXiv:0904.2715 [hep-th]].

[18] A. Gadde, L. Rastelli, S. S. Razamat and W. Yan, "Gauge theories and Macdonald polynomials," Commun. Math. Phys. **319** (2013) 147–193 [arXiv:1110.3740 [hep-th]].

[19] M. Aganagic and K. Schaeffer, "Refined black hole ensembles and topological strings," J. High Energy Phys. **1301** (2013) 060 [arXiv:1210.1865 [hep-th]].

[20] S. de Haro, S. Ramgoolam and A. Torrielli, "Large N expansion of q-deformed two-dimensional Yang-Mills theory and Hecke algebras," Commun. Math. Phys. **273** (2007) 317–355 [arXiv:hep-th/0603056].

[21] T. Brzeziński and S. Majid, "Quantum group gauge theory on classical spaces," Phys. Lett. B **298** (1993) 339–343 [arXiv:hep-th/9210024].

[22] P. Podleś, "Quantum spheres," Lett. Math. Phys. **14** (1987) 193–202.

[23] A. Iqbal and C. Kozcaz, "Refined Hopf link revisited," J. High Energy Phys. **1204** (2012) 046 [arXiv:1111.0525 [hep-th]].

[24] M. Aganagic and S. Shakirov, "Knot homology from refined Chern-Simons theory," [arXiv:1105.5117 [hep-th]].

[25] R. J. Szabo and M. Tierz, "Matrix models and stochastic growth in Donaldson-Thomas theory," J. Math. Phys. **53** (2012) 103502 [arXiv:1005.5643 [hep-th]].

[26] A. Borodin and I. Corwin, "Macdonald processes," [arXiv:1111.4408 [math.PR]]

[27] E. Witten, "Two-dimensional gauge theories revisited," J. Geom. Phys. **9** (1992) 303–368 [arXiv:hep-th/9204083].

[28] I. G. Macdonald, *Symmetric Functions and Hall Polynomials* (Clarendon Press, Oxford, 1995).

[29] D. J. Gross and A. Matytsin, "Instanton induced large N phase transitions in two-dimensional and four-dimensional QCD," Nucl. Phys. B **429** (1994) 50–74 [arXiv:hep-th/9404004].

[30] M. Blau and G. Thompson, "Derivation of the Verlinde formula from Chern-Simons theory and the G/G model," Nucl. Phys. B **408** (1993) 345–390 [arXiv:hep-th/9305010]; "Lectures on $2D$ gauge theories: Topological aspects and path integral techniques," ICTP Ser. Theor. Phys. **10** (1994) 175–244 [arXiv:hep-th/9310144].

[31] M. Blau and G. Thompson, "On diagonalization in $\mathsf{Map}(M,G)$," Commun. Math. Phys. **171** (1995) 639–660 [arXiv:hep-th/9402097].

[32] S. G. Naculich and H. J. Schnitzer, "Level-rank duality of the $U(N)$ WZW model, Chern-Simons theory, and $2D$ q-YM theory," J. High Energy Phys. **0706** (2007) 023 [arXiv:hep-th/0703089].

[33] A. Y. Alekseev, H. Grosse and V. Schomerus, "Combinatorial quantization of the Hamiltonian Chern-Simons theory," Commun. Math. Phys. **172** (1995) 317–358 [arXiv:hep-th/9403066].

[34] A. S. Cattaneo and G. Felder, "A path integral approach to the Kontsevich quantization formula," Commun. Math. Phys. **212** (2000) 591–611 [arXiv:math.QA/9902090].

[35] T. Brzeziński and S. Majid, "Quantum group gauge theory on quantum spaces," Commun. Math. Phys. **157** (1993) 591–638 [arXiv:hep-th/9208007].

[36] P. J. Forrester, *Log-Gases and Random Matrices* (Princeton University Press, 2010).

[37] N. I. Akhiezer, *The Classical Moment Problem* (Hafner Publishing Co., New York, 1965); J. A. Shohat and J. D. Tamarkin, "The problem of moments," Amer. Math. Soc. Surv. **1** (1943) 1–143.

[38] M. Tierz, "Soft matrix models and Chern-Simons partition functions," Mod. Phys. Lett. A **19** (2004) 1365–1378 [arXiv:hep-th/0212128].

[39] S. de Haro and M. Tierz, "Discrete and oscillatory matrix models in Chern-Simons theory," Nucl. Phys. B **731** (2005) 225–241 [arXiv:hep-th/0501123].

[40] M. Mariño, "Chern-Simons theory, matrix integrals, and perturbative three-manifold invariants," Commun. Math. Phys. **253** (2004) 25–49 [arXiv:hep-th/0207096].

[41] C. Berg, "From discrete to absolutely continuous solutions of indeterminate moment problems," Arab. J. Math. Sci. **4** (1998) 1–18.

[42] T. S. Chihara, "A characterization and a class of distribution functions for the Stieltjes-Wigert polynomials," Canad. Math. Bull. **13** (1970) 529–532.

[43] T. S. Chihara, "On generalized Stieltjes-Wigert and related polynomials," J. Comput. Appl. Math. **5** (1979) 291–297.

[44] J. S. Christiansen, "The moment problem associated with the Stieltjes-Wigert polynomials," J. Math. Anal. Appl. **277** (2003) 218–245.

[45] R. J. Szabo and M. Tierz, "Chern-Simons matrix models, two-dimensional Yang-Mills theory and the Sutherland model," J. Phys. A **43** (2010) 265401 [arXiv:1003.1228 [hep-th]].

[46] R. Lidl and H. Niederreiter, *Introduction to Finite Fields and Their Applications* (Cambridge University Press, 1986).

[47] T. Okuda, "Derivation of Calabi-Yau crystals from Chern-Simons gauge theory," J. High Energy Phys. **0503** (2005) 047 [arXiv:hep-th/0409270].

[48] D. Foata and G. N. Han, "The triple, quintuple and septuple product identities revisited," Sém. Lothar. Combin. **42** (1999) 1–12.

[49] S. Matsumoto, "Moments of characteristic polynomials for compact symmetric spaces and Jack polynomials," J. Phys. A **40** (2007) 13567–13586 [arXiv:math.PR/0608751].

[50] T. Nakatsu and K. Takasaki, "Melting crystal, quantum torus and Toda hierarchy," Commun. Math. Phys. **285** (2009) 445–468 [arXiv:0710.5339 [hep-th]].

[51] M. R. Douglas, K. Li and M. Staudacher, "Generalized two-dimensional QCD," Nucl. Phys. B **420** (1994) 118–140 [arXiv:hep-th/9401062].

[52] O. Ganor, J. Sonnenschein and S. Yankielowicz, "The string theory approach to generalized $2D$ Yang-Mills theory," Nucl. Phys. B **434** (1995) 139–178 [arXiv:hep-th/9407114].

[53] M. J. Crescimanno, S. G. Naculich and H. J. Schnitzer, "Evaluation of the free energy of two-dimensional Yang-Mills theory," Phys. Rev. D **54** (1996) 1809–1813 [arXiv:hep-th/9601104].

[54] A. B. Olde Daalhuis, "Asymptotic expansions for q-gamma, q-exponential, and q-Bessel functions," J. Math. Anal. Appl. **186** (1994) 896–913.

[55] N. M. Atakishiyev, A. Frank and K. B. Wolf, "A simple difference realization of the Heisenberg q-algebra," J. Math. Phys. **35** (1994) 3253–3260.

[56] A. N. Kirillov, "Dilogarithm identities," Progr. Theor. Phys. Suppl. **118** (1995) 61–142 [arXiv:hep-th/9408113].

[57] L. D. Faddeev and A. Y. Volkov, "Abelian current algebra and the Virasoro algebra on the lattice," Phys. Lett. B **315** (1993) 311–318 [arXiv:hep-th/9307048].

[58] L. D. Paniak and R. J. Szabo, "Instanton expansion of noncommutative gauge theory in two dimensions," Commun. Math. Phys. **243** (2003) 343–387 [arXiv:hep-th/0203166].

[59] L. D. Paniak and R. J. Szabo, "Open Wilson lines and group theory of noncommutative Yang-Mills theory in two dimensions," J. High Energy Phys. **0305** (2003) 029 [arXiv:hep-th/0302162].

[60] T. Maeda, T. Nakatsu, K. Takasaki and T. Tamakoshi, "Five-dimensional supersymmetric Yang-Mills theories and random plane partitions," J. High Energy Phys. **0503** (2005) 056 [arXiv:hep-th/0412327].

[61] N. A. Nekrasov, "Seiberg-Witten prepotential from instanton counting," Adv. Theor. Math. Phys. **7** (2004) 831–864 [arXiv:hep-th/0206161].

[62] A. Marshakov and N. A. Nekrasov, "Extended Seiberg-Witten theory and integrable hierarchy," J. High Energy Phys. **0701** (2007) 104 [arXiv:hep-th/0612019].

[63] F. Lizzi, R. J. Szabo and A. Zampini, "Geometry of the gauge algebra in noncommutative Yang-Mills theory," J. High Energy Phys. **0108** (2001) 032 [arXiv:hep-th/0107115].

[64] R. J. Szabo, "Quantum field theory on noncommutative spaces," Phys. Rept. **378** (2003) 207–299 [arXiv:hep-th/0109162].

[65] A. Iqbal, N. A. Nekrasov, A. Okounkov and C. Vafa, "Quantum foam and topological strings," J. High Energy Phys. **0804** (2008) 011 [arXiv:hep-th/0312022].

[66] M. Cirafici, A. Sinkovics and R. J. Szabo, "Cohomological gauge theory, quiver matrix models and Donaldson-Thomas theory," Nucl. Phys. B **809** (2009) 452–518 [arXiv:0803.4188 [hep-th]].

[67] H. Ooguri, P. Sulkowski and M. Yamazaki, "Wall crossing as seen by matrix models," Commun. Math. Phys. **307** (2011) 429–462 [arXiv:1005.1293 [hep-th]].

[68] P. Sulkowski, "Wall-crossing, open BPS counting and matrix models," J. High Energy Phys. **1103** (2011) 089 [Erratum-ibid. **1104** (2011) 046] [arXiv:1011.5269 [hep-th]].

[69] M. Yamazaki, "Geometry and combinatorics of crystal melting," RIMS Kokyuroku Bessatsu B **28** (2011) 193–213 [arXiv:1102.0776 [math-ph]].

[70] P. Sulkowski, "BPS states, crystals and matrices," Adv. High Energy Phys. **2011** (2011) 357016 [arXiv:1106.4873 [hep-th]].

[71] C. Faber and R. Pandharipande, "Hodge integrals and Gromov-Witten theory," Invent. Math. **139** (2000) 174–199 [arXiv:math.AG/9810173].

[72] K. Behrend and B. Fantechi, "Symmetric obstruction theories and Hilbert schemes of points on threefolds," Algebra Number Theory **2** (2008) 313–345 [arXiv:math.AG/0512556].

[73] V. Bouchard, A. Catuneanu, O. Marchal and P. Sulkowski, "The remodeling conjecture and the Faber-Pandharipande formula," Lett. Math. Phys. **103** (2013) 59–77 [arXiv:1108.2689 [math.AG]].

[74] V. Bouchard and P. Sulkowski, "Topological recursion and mirror curves," [arXiv:1105.2052 [hep-th]].

[75] A. Berele and A. Regev, "Hook Young diagrams with applications to combinatorics and to representations of Lie superalgebras," Adv. Math. **64** (1987) 118–175.

[76] D. Bump and A. Gamburd, "On the averages of characteristic polynomials from classical groups," Commun. Math. Phys. **265** (2006) 227–274 [arXiv:math-ph/0502043].

[77] P. J. Forrester and A. Gamburd, "Counting formulas associated with some random matrix averages," J. Combin. Theory A **113** (2006) 934–951 [arXiv:math.CO/0503002].

[78] G. Landi and R. J. Szabo, "Dimensional reduction over the quantum sphere and non-abelian q-vortices," Commun. Math. Phys. **308** (2011) 365–413 [arXiv:1003.2100 [hep-th]].

[79] H. Grosse, J. Madore and H. Steinacker, "Field theory on the q-deformed fuzzy sphere 2. Quantization," J. Geom. Phys. **43** (2002) 205–240 [arXiv:hep-th/0103164].

[80] J. Bryan and R. Pandharipande, "The local Gromov-Witten theory of curves," J. Amer. Math. Soc. **21** (2008) 101–136 [arXiv:math.AG/0411037].

[81] R. J. Szabo, "Instantons, topological strings and enumerative geometry," Adv. Math. Phys. **2010** (2010) 107857 [arXiv:0912.1509 [hep-th]].

[82] B. Bakalov and A. A. Kirillov, Jr., "Tensor categories and modular functors," Amer. Math. Soc. Univ. Lect. Ser. **21** (2001) 1–216.

[83] J. Fuchs, I. Runkel and C. Schweigert, "TFT construction of RCFT correlators I: Partition functions," Nucl. Phys. B **646** (2002) 353–497 [arXiv:hep-th/0204148].

[84] S.-H. Ng and P. Schauenburg, "Higher Frobenius-Schur indicators for pivotal categories," Contemp. Math. **441** (2007) 63–90 [arXiv:math.QA/0503167].

[85] V. Ostrik, "Module categories, weak Hopf algebras and modular invariants," Transform. Groups **8** (2003) 177–206 [arXiv:math.QA/0111139].

[86] A. A. Kirillov, Jr., "On an inner product in modular tensor categories," J. Amer. Math. Soc. **9** (1996) 1135–1169 [arXiv:q-alg/9508017].

[87] M. Aganagic, D. Jafferis and N. Saulina, "Branes, black holes and topological strings on toric Calabi-Yau manifolds," J. High Energy Phys. **0612** (2006) 018 [arXiv:hep-th/0512245].

[88] F. Benini, T. Nishioka and M. Yamazaki, "$4D$ index to $3D$ index and $2D$ TQFT," Phys. Rev. D **86** (2012) 065015 [arXiv:1109.0283 [hep-th]].

[89] L. F. Alday, M. Bullimore and M. Fluder, "On S-duality of the superconformal index on lens spaces and $2D$ TQFT," [arXiv:1301.7486 [hep-th]].

[90] P. I. Etingof and A. A. Kirillov, Jr., "On Cherednik-Macdonald-Mehta identities," Amer. Math. Soc. **4** (1998) 43–47 [arXiv:q-alg/9712051].

[91] M. Romo and M. Tierz, "Unitary Chern-Simons matrix model and the Villain lattice action," Phys. Rev. D **86** (2012) 045027 [arXiv:1103.2421 [hep-th]].

[92] M. Aganagic, A. Neitzke and C. Vafa, "BPS microstates and the open topological string wavefunction," Adv. Theor. Math. Phys. **10** (2006) 603–656 [arXiv:hep-th/0504054].

[93] V. G. Drinfel'd, "Quasi-Hopf algebras," Leningrad Math. J. **1** (1991) 1419–1457.

[94] D. Kazhdan and G. Lusztig, "Tensor structures arising from affine Lie algebras IV," J. Amer. Math. Soc. **7** (1994) 383–453.

[95] C. Mazza and C. Weibel, "Schur-finiteness in λ-rings," J. Algebra **374** (2013) 66–78 [arXiv:1011.1444 [math.KT]].

[96] B. Fauser, P. D. Jarvis and R. C. King, "The Hopf algebra structure of the character rings of classical groups," J. Phys. A **46** (2013) 035205 [arXiv:0704.2029 [math.RT]].

[97] L. Geissinger, "Hopf algebras of symmetric functions and class functions," Lect. Notes Math. **579** (1977) 168–181.

[98] L. F. Alday, M. Bullimore, M. Fluder and L. Hollands, "Surface defects, the superconformal index and q-deformed Yang-Mills," [arXiv:1303.4460 [hep-th]].

[99] C. Beasley, "Localization for Wilson loops in Chern-Simons theory," [arXiv:0911.2687 [hep-th]].

[100] J. Kock, *Frobenius Algebras and 2D Topological Quantum Field Theory* (Cambridge University Press, 2003).

[101] J. Fuchs and C. Schweigert, "Category theory for conformal boundary conditions," Fields Inst. Commun. **39** (2003) 25–71 [arXiv:math.CT/0106050].

[102] H. Fuji, S. Gukov, P. Sulkowski and H. Awata, "Volume conjecture: Refined and categorified," [arXiv:1203.2182 [hep-th]].

[103] P. I. Etingof and A. A. Kirillov, Jr., "Representation theoretic proof of the inner product and symmetry identities for Macdonald's polynomials," Compos. Math. **102** (1996) 179–202 [arXiv:hep-th/9410169].

[104] P. I. Etingof and A. A. Kirillov, Jr., "Macdonald's polynomials and representations of quantum groups," Math. Res. Lett. **1** (1994) 279–296 [arXiv:hep-th/9312103].

[105] S. Matsumoto, "Hyperdeterminant expressions for Jack functions of rectangular shape," J. Algebra **320** (2008) 612–632 [arXiv:math.CO/0603033].

[106] M. Cirafici, A. Sinkovics and R. J. Szabo, "Instanton counting and wall-crossing for orbifold quivers," Ann. Henri Poincaré **14** (2013) 1001–1041 [arXiv:1108.3922 [hep-th]].

[107] A. Iqbal, C. Kozcaz and C. Vafa, "The refined topological vertex," J. High Energy Phys. **0910** (2009) 069 [arXiv:hep-th/0701156].

[108] P. Sulkowski, "Refined matrix models from BPS counting," Phys. Rev. D **83** (2011) 085021 [arXiv:1012.3228 [hep-th]].

[109] V. P. Spiridonov, "Essays on the theory of elliptic hypergeometric functions," Russian Math. Surveys **63** (2008) 405–472 [arXiv:0805.3135 [math.CA]].

[110] N. A. Nekrasov, "Instanton partition functions and M-theory," Japan. J. Math. **4** (2009) 63–93.

[111] H. Awata and H. Kanno, "Instanton counting, Macdonald functions and the moduli space of D-branes," J. High Energy Phys. **0505** (2005) 039 [arXiv:hep-th/0502061].

[112] H. Awata and H. Kanno, "Refined BPS state counting from Nekrasov's formula and Macdonald functions," Int. J. Mod. Phys. A **24** (2009) 2253–2306 [arXiv:0805.0191 [hep-th]].

[113] I. Cherednik, "Difference Macdonald-Mehta conjecture," Int. Math. Res. Not. **10** (1997) 449–467 [arXiv:q-alg/9702022].

[114] P. I. Etingof, I. B. Frenkel and A. A. Kirillov, Jr., "Spherical functions on affine Lie groups," Duke Math. J. **80** (1995) 59–90 [arXiv:hep-th/9407047].

[115] A. Gorsky and N. A. Nekrasov, "Relativistic Calogero-Moser model as gauged WZW theory," Nucl. Phys. B **436** (1995) 582–608 [arXiv:hep-th/9401017].

[116] S. de Haro, "Chern-Simons theory in lens spaces from 2D Yang-Mills on the cylinder," J. High Energy Phys. **0408** (2004) 041 [arXiv:hep-th/0407139].

[117] G. Szegő, "On certain Hermitian forms associated with the Fourier series of a positive function," Commun. Sém. Math. Univ. Lund **1952** (1952) 228–237.

[118] B. Simon, *Szegő's Theorem and its Descendants: Spectral Theory for L^2 Perturbations of Orthogonal Polynomials* (Princeton Univeristy Press, 2010).

[119] K. Johansson, "On Szegő's asymptotic formula for Toeplitz determinants and generalizations," Bull. Sci. Math. **112** (1988) 257–304.

[120] P. J. Forrester and S. O. Warnaar, "The importance of the Selberg integral," Bull. Amer. Math. Soc. **45** (2008) 489–534 [arXiv:0710.3981 [math.CA]].

[121] S. O. Warnaar, "q-Selberg integrals and Macdonald polynomials," Ramanujan J. **10** (2005) 237–268.

[122] K. M. Day, "Toeplitz matrices generated by the Laurent series expansion of an arbitrary rational function," Trans. Amer. Math. Soc. **206** (1975) 224–245.

[123] M. Kashiwara and P. Shapira, *Categories and Sheaves* (Springer-Verlag, 2006).

[124] L. Kong and I. Runkel, "Cardy algebras and sewing constraints. I," Commun. Math. Phys. **292** (2009) 871–912 [arXiv:0807.3356 [math.QA]].

[125] P. J. Freyd, *Abelian Categories* (Harper and Row, 1964).

[126] B. Mitchell, *Theory of Categories* (Academic Press, 1965).

Department of Mathematics
Heriot-Watt University
Colin Maclaurin Building, Riccarton, Edinburgh EH14 4AS, UK
Maxwell Institute for Mathematical Sciences, Edinburgh, UK
The Tait Institute, Edinburgh, UK
E-mail address: R.J.Szabo@hw.ac.uk

Departamento de Análisis Matemático, Facultad de Ciencias Matemáticas
Universidad Complutense de Madrid
Plaza de Ciencias 3, Ciudad Universitaria, 28040 Madrid, Spain
E-mail address: tierz@mat.ucm.es