Modular Structures in Superconformal Field Theories

by

MADHUSUDHAN RAMAN PHYS10201205002

The Institute of Mathematical Sciences, Chennai

A thesis submitted to the

Board of Studies in Physical Sciences

In partial fulfilment of requirements

For the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE



July, 2018

Modular Structures in Superconformal Field Theories

by

MADHUSUDHAN RAMAN PHYS10201205002

The Institute of Mathematical Sciences, Chennai

A thesis submitted to the

Board of Studies in Physical Sciences

In partial fulfilment of requirements

For the Degree of

DOCTOR OF PHILOSOPHY

of

HOMI BHABHA NATIONAL INSTITUTE



July, 2018

Homi Bhabha National Institute

Recommendations of the Viva Voce Board

As members of the Viva Voce Board, we certify that we have read the dissertation prepared by Mr. Madhusudhan Raman titled "Modular Structures in Superconformal Field Theories" and recommend that it maybe accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy.

	Date:
Chair - Balachandran Sathiapalan	
	Date:
Guide/Convener - Sujay K. Ashok	
	Date:
Examiner - Sunil Mukhi	
	Date:
Member 1 - Nemani V. Suryanarayana	
	Date:
Member 2 - Partha Mukhopadhyay	
	Date:

Member 3 - Ganesh Ramachandran

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copies of the dissertation to HBNI.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it may be accepted as fulfilling the dissertation requirement.

Date:

Place: IMSc, Chennai

Guide: Sujay K. Ashok

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfilment of requirements for an advanced degree at Homi Bhabha National Institute (HBNI), and is deposited in the Library to be made available to borrowers under the rules of the HBNI.

Brief quotations from this dissertation are allowable without special permission, provided that accurate acknowledgment of the source is made. Requests for permission for extended quotation from, or for reproduction of, this manuscript in whole or in part may be granted by the Competent Authority of HBNI when in his or her judgment the proposed use of the material is in the interests of scholarship. In all other instances, however, permission must be obtained from the author.

DECLARATION

I, hereby declare that the investigation presented in the thesis has been carried out by me. The work is original and has not been submitted earlier as a whole or in part for a degree / diploma at this or any other Institution / University.

- S-duality, triangle groups and modular anomalies in N = 2 SQCD Sujay K. Ashok, Eleonora Dell'Aquila, Alberto Lerda, Madhusudhan Raman JHEP 04 (2016) 118, arXiv:1601.01827 [hep-th]
- Chiral observables and S-duality in N = 2^{*} U(N) gauge theories
 Sujay K. Ashok, Marco Billò, Eleonora Dell'Aquila, Marialuisa Frau, Alberto Lerda, Micha Moskovic, Madhusudhan Raman
 JHEP 11 (2016) 020, arXiv:1607.08327 [hep-th]

1. Modular anomaly equations and S-duality in $\mathcal{N}=2$ conformal SQCD

Sujay K. Ashok, Marco Billò, Eleonora Dell'Aquila, Marialuisa Frau, Alberto Lerda, Madhusudhan Raman JHEP **10** (2015) 091, arXiv:1507.07476 [hep-th]

2. Exact WKB Analysis of $\mathcal{N} = 2$ Gauge Theories,

Sujay K. Ashok, Dileep P. Jatkar, Renjan R. John, Madhusudhan Raman, Jan Troost

JHEP 07 2016 115, arXiv:1604.05520 [hep-th]

3. Horizon tunneling revisited: the case of higher dimensional black holes,

Madhusudhan Raman JHEP **12** (2017) 144, arXiv:1711.00287 [hep-th]

4. Aspects of Hecke Symmetry I: Ramanujan Identities and Inversion Formulas,

Madhusudhan Raman

arXiv:1803.10224 [math.NT]

If the pages of this book contain some successful verse, the reader must excuse me the discourtesy of having usurped it first. Our nothingness differs little; it is a trivial and chance circumstance that you should be the reader of these exercises and I their author.

JORGE LUIS BORGES

ACKNOWLEDGMENTS

At our first meeting I remember asking Sujay what the prerequisites were for working with him, and his answer was simple: "Confidence." By example, he taught me to be unafraid of the dense and difficult parts of our subject. I have constantly attempted to emulate his work ethic, admired his enthusiasm and boundless energy, and enjoyed his friendship. I thank him for his encouragement, all the freedoms he accorded me, and the things he taught me about quantum fields and strings.

My interactions with Alok have had a profound impact on me, both personally and professionally. I am deeply grateful to Alok for his abiding friendship, his counsel, his kindness, his book recommendations, and the many hours we spent talking. I thank Dileep for his entertaining and light-hearted conversation, his playful and irreverent approach to physics and mathematics, and our enjoyable collaborations.

This thesis owes much to my collaborators: Alberto, for his keen eye and the clarity of his drafts; Eleonora, for her formidable computer programs that taught me by example how to code and how to localize quantum field theories, and without which this thesis would not have been possible; Jan, for his very helpful and detailed correspondence; Marco, for many illuminating conference calls; Marialuisa, for her lucid scanned notes; and Micha, for teaching me some new and beautiful things about spectral curves. I am especially grateful to Renjan our discussions, for his humour and friendship, and for his excellent company on our travels together.

I would remiss if I didn't acknowledge the participants of various seminars and journal clubs in Chennai, and the annual string theory meetings which did much to keep me motivated. I am also thankful to the string theory group at the Tata Institute for Fundamental Research, for arranging an enjoyable three-month-long visit in my final year. I thank Bala Sathiapalan, Nemani Suryanarayana, Partha Mukhopadhyay, and Ganesh Ramachandran for serving on my doctoral committee, and for checking in on me periodically. I am also grateful to Prof. Sunil Mukhi and Prof. Francesco Fucito for patiently reviewing this thesis.

I am thankful for the friendship of an stellar cast of colourful characters: Anup, Ashutosh, Nitin, Prathyush, and Soumyadeep for memorable, fun-filled evenings at Thoraipakkam; Anantha, Ramanathan, and Varuni, for many evenings spent together, and countless conversations; Chandrashekar, Prathamesh, and Sridhar for the hilarity that ensues when they are in the same room; Madhushree, Thomas, and Srividya for their kindness and humour; Kunal, for his unfailing friendship and sage advice; and Mukunth, for enlightening conversations about science and society, for being an excellent roommate, and for encouraging me to write.

To Aditi I owe the completion of this thesis, and I am grateful for her affection, her sense of humour, and the wonderful times we have spent together.

I thank the administrative staff at IMSc, especially Ms. Indra and Ms. Prema, for help during various stages of my PhD. I am indebted to the canteen workers and housekeeping staff at IMSc, who work tirelessly to keep us well-fed, and our offices and classrooms neat and tidy.

Finally, I have benefited enormously from the love and support from my parents and my sister. This thesis is dedicated to them, with affection.

ABSTRACT

We study the modular structures underlying $\mathcal{N} = 2$ superconformal gauge theories in four dimensions. Using constraints from a nonperturbative duality symmetry called S-duality, we show that observables of interest may be resummed into quasimodular forms of the S-duality group. We study these gauge theories coupled to different kinds of matter.

Adjoint Matter: We study chiral observables in U(N) gauge theories and show that they may be resummed into quasimodular forms of the nonperturbative S-duality group. We do this using a number of complimentary approaches: explicitly evaluating the period integrals; invoking a correspondence between these gauge theories and an integrable model called the elliptic Calogero-Moser system; and microscopically evaluating nonperturbative contributions to chiral observables using the machinery of equivariant localization.

Fundamental Matter: We study SQCD theories with gauge group SU(N) with $N_f = 2N$ fundamental hypermultiplets. When the flavours are massless, we focus on the period matrix of these theories in a \mathbb{Z}_N -symmetric locus on the Coulomb moduli space: the special vacuum. We clarify the underlying modular structure, in particular to understand the manner in which the S-duality group acts on the renormalized couplings, and show that this action is consistent with more recent studies of S-duality that focus on the bare couplings of the gauge theory. We also study massive hypermultiplet configurations that respect the \mathbb{Z}_N symmetry of the special vacuum. Here, we find that the modular structure of the massless theory will is deformed; more specifically, we find that the renormalized couplings admit semiclassical expansions with mass-dependent coefficients. We use constraints from S-duality to derive modular anomaly equations, which are then used to solve for the mass-dependent coefficients order-by-order in the mass expansion.

List of Tables

4.1	Relevant parameters for the j_{λ_k} functions, their <i>q</i> -expansions, and the	
	weak-coupling expansion of the bare coupling q_0 defined in (4.96).	120

4.2	The correspondence between λ_k and p_k according to (4.102) in the	
	arithmetic cases	21

List of Figures

4.1	The fundamental domain F' of $\Gamma^*(\lambda_k)$. The point τ_k^A is the fixed
	point of the S, τ_k^B is the fixed point of ST^{-1} , while τ_k^C is the fixed
	point of T

Contents

0	Synopsis			xxix	
	0.1	Backg	round	. xxix	
		0.1.1	Seiberg-Witten Theory	. XXX	
		0.1.2	Integrable Systems and a Correspondence	. xxxi	
		0.1.3	Equivariant Localization	. xxxiii	
	0.2	Chiral	Rings	. XXXV	
		0.2.1	Curves and Differentials	. xxxvi	
		0.2.2	Localization	. xxxvii	
	0.3	Funda	mental Matters	. xxxviii	
		0.3.1	Massless SQCD and Duality Groups	. xl	
		0.3.2	Massive Hypermultiplets and Modular Anomalies	. xli	
	0.4	Plan (Of Thesis	. xliii	

1	Inst	anton	s and Supersymmetry	1
	1.1	Instan	tons	2
		1.1.1	ADHM Construction	4
	1.2	Super	symmetry	9
		1.2.1	Field Representations	10
		1.2.2	Actions	13
		1.2.3	1-Loop Corrections	16
2	Seil	oerg-W	itten, Calogero-Moser, and Nekrasov	19
	2.1	Seiber	g-Witten Theory	19
		2.1.1	Moduli Space	19
		2.1.2	Electric-Magnetic Duality	21
		2.1.3	Monodromies, Monopoles, and Dyons	23
		2.1.4	Solution	27
	2.2	Calog	ero-Moser System	29
		2.2.1	A Brief Interlude: Integrable Systems	29
		2.2.2	Calogero-Moser Systems	31
		2.2.3	Lax Pairs for Elliptic Calogero-Moser Systems	32
	2.3	Equiv	ariant Localization	33
		2.3.1	Topological Twisting	33

		2.3.2 The Ω Background \ldots	37
		2.3.3 The Prepotential	45
3	Chi	al Rings	47
	3.1	Introduction	47
	3.2	Review: $\mathcal{N} = 2^* U(N)$ SYM Theories	51
	3.3	Seiberg-Witten Curves	55
		3.3.1 The Donagi-Witten Curve	56
		3.3.2 The D'Hoker-Phong Curve	61
		3.3.3 Comparing Curves	63
	3.4	Period Integrals and Modular Anomalies	66
		3.4.1 Period Integrals	66
		3.4.2 A Modular Anomaly Equation	73
	3.5	Chiral Observables from Localization	77
	3.6	1-Instanton Results	84
4	Fun	amental Matters	87
	4.1	Introduction	88
Μ	lassle	s $\mathcal{N} = 2$ SQCD and Duality Groups	93
	4.2	Massless SQCD and the Special Vacuum	93

	4.2.1	The Prepotential
	4.2.2	The Special Vacuum
	4.2.3	Observables
4.3	The S-	-duality Group
	4.3.1	S-Action on Period Integrals and Gauge Coupling 104
4.4	The A	rithmetic Theories
	4.4.1	N=2
	4.4.2	N = 3
	4.4.3	N = 4
	4.4.4	$N = 6 \dots \dots \dots \dots \dots \dots \dots \dots \dots $
4.5	S-Dua	lity and j -Invariants for Arithmetic Theories
	4.5.1	The S-duality Group of Conformal SU(2) Gauge Theory \dots 114
	4.5.2	Renormalized Couplings and <i>j</i> -Invariants
4.6	$\mathrm{SU}(N)$) Theories and Triangle Groups
	4.6.1	A Short Digression on Triangle Groups
	4.6.2	N = 5
	4.6.3	Generalization to Higher N
	4.6.4	Relation to Earlier Work

Massive $\mathcal{N} = 2$ SQCD and Modular Anomaly Equations 134

	4.7	Mass Deformations and Observables	. 135
	4.8	S-Duality in Massive SQCD	. 137
		4.8.1 The Modular Anomaly Equation	. 139
		4.8.2 Coupled Modular Anomaly Equations	. 141
	4.9	Resummation: $N = 4$ and $N = 6$. 144
		$4.9.1 N = 4 \dots \dots \dots \dots \dots \dots \dots \dots \dots $. 144
		4.9.2 $N = 6$. 150
5	Out	look	154
	5.1	Chiral Rings	. 154
	5.2	Fundamental Matters	. 156
A	Eise	enstein Series and Elliptic Functions	160
в	Gen	neralized Donagi-Witten Polynomials	166
С	Moo	dular Covariance and the D'Hoker-Phong Curve	168
D	Con	agruence Subgroups	170
E	Mas	ssless $SU(7)$ SQCD	177

Synopsis

The subject of our investigations in this thesis are $\mathcal{N} = 2$ supersymmetric theories in four dimensions. The study of these theories was pioneered by Seiberg and Witten [1, 2], who identified the complexified gauge coupling with the modular parameter of a torus, thus drawing upon results from the classical geometry of genus-1 surfaces to furnish a solution to the low-energy effective action. Later, Nekrasov and Okounkov [3, 4] brought the techniques of equivariant localization to bear on the problem of evaluating integrals over instanton moduli spaces, allowing for a microscopic check of the Seiberg-Witten solution, in addition to paving the way for more sophisticated developments. This thesis will focus on the modular structures underlying these gauge theories, in particular the resummation of various observables into quasimodular forms of the S-duality group.

0.1 Background

Here we discuss the background material relevant to our thesis: Seiberg-Witten theory, its correspondence with integrable systems, and equivariant localization.

0.1.1 Seiberg-Witten Theory

At low energies, a SU(N) gauge theory with eight supercharges on \mathbb{R}^4 is Higgsed down to its maximal torus, and one ends up with a U(1)^{N-1} theory whose vacuum is specified by a set $\{a_1, \dots, a_{N-1}\}$; these are vacuum expectation values of the adjoint scalar in the $\mathcal{N} = 2$ vector multiplet. The low-energy effective action for these theories is completely specified by a single holomorphic function F called the *prepotential*. The nonrenormalization theorems of Seiberg [5] may be invoked to show that perturbative corrections to the classical prepotential truncate at 1-loop. One *solves* an $\mathcal{N} = 2$ theory by computing nonperturbative contributions to the prepotential.

Seiberg-Witten theory offers a geometric approach to this problem. Broadly, to an $\mathcal{N} = 2$ theory one associates an algebraic curve and an associated 1-form called the Seiberg-Witten differential, denoted λ_{SW} . To such a surface, one may associate a canonical, symplectically paired basis of cycles (α_k, β_k) ; the insight of Seiberg and Witten was to make the identifications

$$a_k = \oint_{\alpha_k} \lambda_{\rm SW}$$
 and $a_k^{\rm D} = \oint_{\beta_k} \lambda_{\rm SW}$. (1)

Here, $a_k^{\rm D}$ are related to the prepotential as

$$a_k^{\rm D} = \frac{\partial F}{\partial a_k} \ . \tag{2}$$

These identifications have the happy consequence that the period matrix of the algebraic curve

$$\tau_{ij} = \frac{\partial a_i^{\rm D}}{\partial a_j} = \frac{\partial^2 F}{\partial a_i \,\partial a_j} \,, \tag{3}$$

automatically satisfies $\text{Im } \tau_{ij} > 0$, as is required for the kinetic energy term for the gauge fields in the action to be positive definite.

When evaluated, the integrals (1) are functions of $\{u_2, \dots, u_N\}$, gauge invariant coordinates on the Coulomb moduli space. The strategy of Seiberg-Witten (who studied the case r = 1) was to evaluate a(u) and invert it, i.e. obtain u(a). Once this inversion is plugged into a^{D} , we use (2) and integrate to determine the prepotential.

The Seiberg-Witten solution has been generalized to accommodate other gauge algebras, fundamental and adjoint matter, and even quiver gauge theories that have multiple gauge groups and matters charged under them.

0.1.2 Integrable Systems and a Correspondence

The geometric nature of the Seiberg-Witten solution allows for a correspondence between $\mathcal{N} = 2$ gauge theories and integrable systems, as chronicled in [6]. In this section, we briefly discuss this correspondence.

A mechanical system is said to be *integrable* if there exist as many integrals of motion (i.e. conserved quantities) as there are degrees of freedom. More precisely, if the mechanical system under investigation has n degrees of freedom, we require that there exist n functionally independent conserved quantities $I_k(x, p)$; this is embodied in the conditions

$$\{I_k, H\} = 0 \text{ and } \{I_k, I_\ell\} = 0.$$
 (4)

A more fruitful (and fully equivalent) formulation of the criterion of integrability may be formulated as follows: an integrable system is said to have a *Lax pair* if one can find a pair of $N \times N$ matrix-valued functions on phase space (L, M) such that the equation

$$\dot{L} = [L, M] , \qquad (5)$$

is equivalent to Hamilton's equations of motion. It may be verified that once a Lax pair is found, the integrals of motion are given by

$$I_k = \operatorname{Tr} L^k . (6)$$

In many examples of integrable systems, it is possible to find a family of Lax pairs labelled by a complex parameter called a *spectral parameter*. That is, the equation

$$\dot{L}(z) = [L(z), M(z)],$$
(7)

is equivalent to Hamilton's equations of motion for all values of z. To such a Lax pair with a spectral parameter, we may naturally associate a spectral curve

$$\Gamma = \{ (k, z) \in \mathbb{C} \times \mathbb{C} : \det [kI - L(z)] = 0 \} , \qquad (8)$$

and a differential 1-form

$$\mathrm{d}\lambda = k\,\mathrm{d}z\;.\tag{9}$$

The correspondence between $\mathcal{N} = 2$ supersymmetric gauge theories and integrable systems goes as follows:

- The $\mathcal{N} = 2$ supersymmetric gauge theory data defines an integrable system. We will be interested in SU(N) Yang-Mills theory with $\mathcal{N} = 2$ supersymmetry and a massive adjoint hypermultiplet, which maps to the elliptic Calogero-Moser system, a system of N particles on a line interacting pairwise via a potential given by the Weierstrass \wp -function.
- The Seiberg-Witten curve maps to the spectral curve of the integrable system,
- The Seiberg-Witten differential maps to the differential 1-form on the spectral

curve of the integrable system.

• The gauge invariant Coulomb moduli correspond to the conserved integrals of motion in the integrable system.

In this dissertation, we will concern ourselves with the study of chiral observables in these gauge theories, which correspond in turn to integrals of motion in the integrable system.

0.1.3 Equivariant Localization

As we have discussed, the solution to an $\mathcal{N} = 2$ gauge theory is given by computing all non-perturbative contributions to the prepotential. This amounts, equivalently, to computing the effective action. This is difficult to do via the usual instanton calculus [7], and in order to make progress beyong low instanton numbers, we must adopt the methods of Nekrasov [3, 4].

Nekrasov key insights may be broken down into a few key steps. We start with a topological twist [8]. Here, we start with the global symmetry of the gauge theory

$$\operatorname{SU}(2)_l \times \operatorname{SU}(2)_r \times \operatorname{SU}(2)_R ,$$
 (10)

where $\mathrm{SU}(2)_l$ and $\mathrm{SU}(2)_r$ correspond to a decomposition of the Euclidean rotation group on \mathbb{R}^4 . The group $\mathrm{SU}(2)_R$ is the *R*-symmetry of the $\mathcal{N} = 2$ supersymmetry algebra. Topological twisting essentially redefines the (Euclidean) Lorentz group; let $\mathrm{SU}(2)_d = \operatorname{diag} \mathrm{SU}(2)_r \times \mathrm{SU}(2)_R$, the diagonal subgroup, and let the (Euclidean) Lorentz group be identified with

$$\mathrm{SU}(2)_l \times \mathrm{SU}(2)_d$$
 (11)

which in turn yields a scalar supercharge Q. This scalar supercharge is then used to write the action in a Q-exact form, which in turn (using techniques from topological field theory) allows one to localize onto the space of self-dual solutions to the Yang-Mills field equations. Since this space is finite-dimensional, in principle this results in an enormous simplification.

Self-dual solutions to the Yang-Mills field equations are called instantons, described by the ADHM construction [9]. The ADHM moduli space, however, is non-compact, and consequently integrals over these moduli spaces are divergent. In order to cure these divergences, one introduces the Ω -deformation [3] as a regulator. This perspective allows us to take advantage of the techniques of equivariant localization to compute integrals over the k-instanton moduli space exactly.

Finally, in order to make contact with the prepotential of the undeformed super Yang-Mills theory, Nekrasov and Okounkov [4] proved that

$$F = \lim_{\Omega \to 0} \log Z_{\text{Nekrasov}} , \qquad (12)$$

where Z_{Nekrasov} is called the Nekrasov partition function, which in turn may be expressed in terms of contour integrals. There exists a recipe for constructing Nekrasov integrands for every k-instanton moduli space, that accounts for any matter present in the theory as well.

We will use Nekrasov localization as a microscopic check of our results in both the study of chiral rings in theories with adjoint matter, as well as observables in SQCD theories.

0.2 Chiral Rings

In this chapter we focus on $\mathcal{N} = 2^*$ theories. Besides the gauge vector multiplet, they contain an adjoint hypermultiplet of mass m that interpolates between the $\mathcal{N} = 4$ SYM theories (when $m \to 0$) and the pure $\mathcal{N} = 2$ SYM theories (when $m \to \infty$). The $\mathcal{N} = 2^*$ theories inherit from the $\mathcal{N} = 4$ models an interesting action of the S-duality group; in particular, their prepotentials satisfy modular anomaly equations first discussed in [10] and developed further in [11, 12, 13]. This approach has led to a very efficient way of determining the mass expansion of the prepotential in terms of: *i*) quasi-modular functions of the gauge coupling and *ii*) the vacuum expectation values a_u of the scalar field Φ of the gauge multiplet such that only particular combinations, defined purely in terms of sums over the root lattice of the corresponding Lie algebra, appear. These results have been checked against explicit computations using equivariant localization.

In this work, we take the first steps towards showing that similar modular structures also exist for other observables of $\mathcal{N} = 2^*$ gauge theories. We choose to work with U(N) gauge groups, and consider the quantum expectation values

$$\langle \operatorname{Tr} \Phi^n \rangle$$
 . (13)

A priori, it is not obvious that these chiral observables exhibit modular behaviour. However, we show that it is always possible to find combinations that transform as modular forms of definite weight under the non-perturbative duality group $SL(2, \mathbb{Z})$. These combinations have a natural interpretation as modular-covariant coordinates on the Coulomb moduli space, and can be analysed using two different techniques: *i*) the SW approach via curves and differentials, and *ii*) equivariant localization combined with the constraints arising from S-duality.
0.2.1 Curves and Differentials

For $\mathcal{N} = 2^*$ theories there are many distinct forms of the SW curve that capture different properties of the chiral observables. In one approach, due to Donagi and Witten [14, 15], the SW curve has coefficients A_n that have a natural interpretation as modular-covariant coordinates on the Coulomb moduli space. Thus, this approach provides us with a natural setting to study the elliptic and modular properties of the observables (3.1). Another form of the SW curve was found by using the relation with integrable systems [16]. For the $\mathcal{N} = 2^*$ theory, the relevant curve was proposed by D'Hoker and Phong [17, 6], who used the close relation between the gauge theory and the elliptic Calogero-Moser system [18]. In this second formulation, the coefficients of the spectral curve of the integrable system are interpreted as symmetric polynomials built out of the quantum chiral ring elements:

$$W_n = \sum_{u_1 < \dots < u_n} e_{u_1} \cdots e_{u_n} . \tag{14}$$

The e_u are interpreted as the quantum-corrected vacuum expectation values of the scalar field Φ and, at weak coupling, they have the following form

$$e_u = a_u + O(q). \tag{15}$$

We review and relate these two descriptions of the SW curve. This comparison will lead to interesting relationships between the coefficients of the respective curves, of the form

$$W_n = \sum_{\ell=0}^{[n/2]} (-1)^{\ell} \binom{N-n+2\ell}{2\ell} (2\ell-1)!! \left(\frac{m^2 E_2}{12}\right)^{\ell} A_{n-2\ell} .$$
(16)

Along the way, we will find it necessary to modify the analysis of [14] in a subtle but important way. It is clearly desirable to work with chiral observables that in the classical limit coincide with the symmetric polynomials built out of the vacuum expectation values a_u . As we discuss, this can be done in two ways. The first is to compute the period integrals in the Donagi-Witten form of the curve as a series expansion in the mass m of the adjoint hypermultiplet. Inverting this expansion order by order in m gives us an expression for the A_n in terms of the a_u . The second way is to postulate that the A_n have a definite modular weight under the S-duality group, and use the well-understood action of S-duality to derive a modular anomaly equation that recursively determines them up to modular pieces.

$$\frac{\partial A_n^{(\ell)}}{\partial E_2} + \frac{1}{12} \sum_{k=0}^{\ell} \frac{\partial A_n^{(k)}}{\partial a} \cdot \frac{\partial f_{\ell-k}}{\partial a} = 0 .$$
 (17)

In this derivation, we see that it is crucial that the prepotential and hence the dual periods of the $\mathcal{N} = 2^*$ theory are known in terms of quasi-modular forms. In both ways it turns out that the chiral observables can be expressed in terms of quasimodular forms and of particular functions of the a_u involving only sums over the weight and root lattices of the Lie algebra $\mathfrak{u}(N)$, generalizing those appearing in the prepotential.

0.2.2 Localization

Next, we test our findings against explicit microscopic computations of the observables (3.1) using equivariant localization techniques. We find that the chiral observables computed using localization can be matched with those obtained from the SW curves by a redefinition of the chiral ring elements. Such a redefinition contains only a finite number of terms and is exact both in the mass of the hypermultiplet and in the gauge coupling. It is well known that the localization results for the chiral observables do not, in general, satisfy the classical chiral ring relations [19, 20, 21]. Strikingly, we show that the redefinition of the chiral ring elements which allow the matching of the two sets of results can be interpreted as a judicious choice of coordinates on the Coulomb moduli space in which the *classical* chiral ring relations are naturally satisfied.

Finally, we focus on the 1-instanton contributions and, just as it was done for the prepotential in [11, 12], we manage to resum the mass expansion to obtain an exact expression involving only sums over roots and weights of the corresponding Lie algebra:

$$\left\langle \operatorname{Tr} \Phi^{n} \right\rangle \Big|_{k=1} = -n(n-1) q \, m^{2} \sum_{\lambda \in \mathcal{W}} (\lambda \cdot A)^{n-2} \left[1 - \sum_{\alpha \in \Psi_{\lambda}} \frac{m^{2}}{(\alpha \cdot A)^{2}} \prod_{\beta \in \Psi_{\alpha}} \left(1 + \frac{m}{\beta \cdot A} \right) \right].$$
(18)

0.3 Fundamental Matters

There has been much progress in understanding conformally invariant $\mathcal{N} = 2$ supersymmetric gauge theories in four dimensions, especially following the seminal work of Gaiotto [22]. In that work, the four-dimensional $\mathcal{N} = 2$ theories were realized as compactifications of the six-dimensional (2,0) theory on a punctured Riemann surface Σ . One of the important results of this approach was to identify the complex structure moduli space of Σ with the space of gauge couplings modulo the action of the S-duality group. For linear quiver gauge theories in the weak coupling limit, the Riemann surface degenerates into a collection of three-punctured spheres connected by long thin tubes, and the sewing parameters are identified with the bare coupling constants of the superconformal gauge theory.

This approach is fruitfully contrasted with the original solution of $\mathcal{N} = 2$ gauge theories due to Seiberg and Witten [1, 2], where the quantum effective action on the Coulomb branch is obtained from an algebraic curve describing a Riemann surface,

and an associated holomorphic differential. A natural question to ask in this context is whether the non-perturbative S-duality group can be used to solve for the effective action. For $\mathcal{N} = 2^*$ theories (i.e. mass deformed $\mathcal{N} = 4$ theories) with unitary gauge groups it has been shown [10, 23, 24, 25] that the constraints coming from S-duality take the form of a modular anomaly equation whose solution allows one to reconstruct the prepotential on the Coulomb branch order by order in the mass of the adjoint hypermultiplet to all orders in the gauge coupling. To achieve this result one has to organize the low-energy effective prepotential as a semi-classical expansion in inverse powers of the vacuum expectation values of the scalar fields in the gauge vector multiplet and realize that the coefficients of this expansion satisfy a recursion relation whose solution can be written in terms of quasi-modular forms of $PSL(2,\mathbb{Z})$ acting on the bare gauge coupling. These modular forms resum the instanton series and therefore provide an exact result. It is of particular importance that $\mathcal{N} = 2^{\star}$ theories are characterized by the absence of any renormalization of the coupling constant, even non-perturbatively; thus, the bare coupling is the only coupling that is present in the effective theory. This procedure has been applied also to $\mathcal{N} = 2^{\star}$ theories with arbitrary gauge groups in [11, 12], where it has been observed that for non-simply laced algebras the effective prepotential is expressed in terms of quasi-modular forms of congruence subgroups of $PSL(2,\mathbb{Z})$.

In this work we study $\mathcal{N} = 2$ gauge theories with gauge group SU(N) and 2N fundamental flavours, generalizing the analysis of the SU(3) gauge theory with six flavours recently presented in [26]. When all flavours are massless, these SQCD theories are superconformal. However, unlike the case of $\mathcal{N} = 2^*$ theories, the bare gauge coupling in $\mathcal{N} = 2$ SQCD is renormalized by quantum corrections which arise from a finite 1-loop contribution as well as from an infinite series of non-perturbative contributions due to instantons. In general these corrections are different for the various U(1) factors and thus one expects to find several effective couplings in the low-energy theory. This chapter is divided into two parts.

0.3.1 Massless SQCD and Duality Groups

In the first part, we work in the conformal limit with all flavour masses set to zero, and calculate various observables of the effective theory such as the prepotential, the period integrals and the period matrix, using equivariant localization. In particular, we work in a special locus of the Coulomb branch which possesses a \mathbb{Z}_N symmetry and which we call the *special vacuum* [27]. In this special vacuum, the period matrix has fewer independent components than it does at a generic point of the moduli space. More precisely, when all quantum corrections are taken into account there are $\left[\frac{N}{2}\right]$ distinct matrix structures which correspond to $\left[\frac{N}{2}\right]$ renormalized coupling constants in the effective theory.¹

$$\Omega = \underbrace{\tau_1 \ \mathcal{M}_1 + \tau_2 \ \mathcal{M}_2 + \cdots}_{\left[\frac{N}{2}\right] \text{ terms}}$$
(19)

Of course, one could in principle use any basis of matrices \mathcal{M}_k to write Ω , but a particularly insightful choice is the one that "diagonalizes" the action of the Sduality group. In such a basis, under S-duality each \mathcal{M}_k stays invariant and each τ_k transforms as

$$\tau_k \to -\frac{1}{\lambda_k \, \tau_k} \tag{20}$$

for some positive λ_k . We conjecture that the spectrum of λ_k is given by

$$\lambda_k = 4\sin^2\left(\frac{k\,\pi}{N}\right) \,\,,\tag{21}$$

and find complete agreement with explicit localization computations. Note that for $N \in \{2, 3, 4, 6\}$ all the λ_k 's take integer values. We call these cases *arithmetic*. If instead $N \notin \{2, 3, 4, 6\}$, then the λ_k 's are not necessarily integer. We refer to the latter as the *non-arithmetic* cases. Of course, at leading order such renormalized couplings

¹Here $[\cdot]$ denotes the floor function.

are all equal to the bare coupling, but when 1-loop and instanton corrections are taken into account, they begin to differ from one another. Given that the S-duality group naturally acts on the bare coupling, an obvious question to ask is how Sduality is realized on the various parameters of the quantum theory. The answer we provide is that on each individual effective coupling S-duality acts as a generalized triangle group (see for example [28]). Moreover, using this insight, we propose a non-perturbatively *exact* relation between the bare coupling $(2\pi i \tau_0 = \log q_0)$ and the renormalized ones (τ_k) that takes a universal form in terms of the *j*-invariants of the triangle groups.

$$q_0 = \frac{\sqrt{j_{\lambda_k}(\tau_k) - d_{\lambda_k}^{-1}} - \sqrt{j_{\lambda_k}(\tau_k)}}{\sqrt{j_{\lambda_k}(\tau_k) - d_{\lambda_k}^{-1}} + \sqrt{j_{\lambda_k}(\tau_k)}}$$
(22)

We perform several successful checks of this proposal by comparing the instanton contributions predicted by the exact relation with the explicit results obtained from multi-instanton localization. As a further evidence in favour of our proposal, we show that the action of S-duality on the renormalized couplings is fully consistent with the action on the bare coupling as obtained from Gaiotto's analysis [22]. We believe that our results, and in particular the exact relation we propose, can play an important role in the study of these SQCD theories at strong coupling [29]. This is because the *j*-invariants have a well-understood behaviour near those cusp points where the coupling constants become large and the usual weak-coupling expansion cannot be used.

0.3.2 Massive Hypermultiplets and Modular Anomalies

In the second part of the chapter we move away from the conformal limit by giving a mass to the fundamental flavour hypermultiplets. For generic masses the \mathbb{Z}_N symmetry of the special vacuum is broken; to avoid this, we restrict our analysis to \mathbb{Z}_N -symmetric mass configurations so that the modular structure uncovered in the massless limit gets deformed in a natural and smooth manner. In particular, with these \mathbb{Z}_N -symmetric mass configurations we find that the $\left[\frac{N}{2}\right]$ matrix structures of the massless theories are preserved, while the $\left[\frac{N}{2}\right]$ effective couplings simply receive further contributions proportional to the hypermultiplet masses. Building on earlier literature [30, 31], this analysis was already carried out for the SU(2) theory in [23, 24], where it was shown that the prepotential can be written in terms of quasimodular forms of the modular group PSL(2, \mathbb{Z}). Moreover, after expanding the prepotential in powers of the flavour masses, it was realized that the coefficients of this expansion satisfy a modular anomaly equation that takes the form of a recursion relation, similar to that of the $\mathcal{N} = 2^*$ case. These results have been recently extended to the SU(3) theory with six massive flavours in [26], where it has been shown that the prepotential, the dual periods and the period matrix are constrained by S-duality to obey again a recursion relation that can be written as a modular anomaly equation, which takes the form

$$\frac{\partial g_n}{\partial \widetilde{E}_2} = \frac{(3n+1)}{24} \sum_{\ell < n} g_\ell g_{n-\ell-1} , \qquad (23)$$

where the g_n are coefficients in the semi-classical (large-*a*) expansion of the dual period integrals. In this case, the solutions of this equation are quasi-modular forms of $\Gamma_1(3)$, which is a subgroup of the S-duality group that is also a congruence subgroup of PSL(2, \mathbb{Z}). Here we further extend these results to the general SU(*N*) theory with 2*N* massive flavours and show that the constraints arising from S-duality can always be written as a recursion relation for any *N*. However, beyond this step, the analysis crucially depends on the arithmetic properties of the S-duality group. It turns out that for N = 2, 3, 4 and 6, the S-duality group acting on each quantum coupling always has a subgroup which is a congruence subgroup of PSL(2, \mathbb{Z}). For these theories, which we call arithmetic, the discussion proceeds along the same

lines described in [26] for the SU(3) theory, with one important modification: in the higher rank cases, the S-duality constraints are written as *coupled* modular anomaly equations.

$$\frac{\partial g_n^{(k)}}{\partial \widetilde{E}_2^{(\lambda_k)}} = \frac{1}{12} \sum_{m=0}^{n-1} \frac{(Nm+N-1)(N(n-m)-1)}{Nn+N-1} g_m^{(k)} g_m^{(k)} g_{n-m-1}^{(k)} ,$$

$$\frac{\partial g_n^{(k)}}{\partial \widetilde{E}_2^{(\lambda_1)}} = \frac{1}{12} \sum_{m=0}^{n-1} \frac{(Nm+N)(Nm+N-1)}{Nn+N-1} g_m^{(k)} g_{n-m-1} .$$
(24)

These coupled equations are nevertheless integrable and their solutions turn out to be polynomials in meromorphic quasi-modular forms of congruence subgroups of $PSL(2,\mathbb{Z})$.

0.4 Plan Of Thesis

The goal of this thesis is to discuss the modular structure of various observables in gauge theories with $\mathcal{N} = 2$ supersymmetry. It will consist of the following chapters.

- Chapter 1 will provide a general introduction to N = 2 supersymmetric gauge theories and instantons.
- Chapter 2 will review Seiberg-Witten theory, the elliptic Calogero-Moser system, and equivariant localization.
- Chapter 3 will study the modular properties of chiral observables in supersymmetric gauge theories with adjoint matter.
- Chapter 4 will study the modular structure of various observables in SQCD theories.
- Chapter 5 will conclude with a discussion of the results as well as open problems.

Instantons and Supersymmetry

Consider an interacting quantum system with a coupling constant g^2 . At weak coupling, an observable of interest may be computed perturbatively in g^2 . By this, we mean that an observable E admits an expansion of the form

$$E(g) = c_0 + c_1 g^2 + c_2 g^4 + \cdots, \qquad (1.1)$$

and the machinery of quantum theory teaches us how to compute the quantities c_k , which may subsequently be compared against experiment. In the presence of multiple vacua, we must account for transitions between them via quantum mechanical tunneling. At weak coupling $g^2 \ll 1$, these effects may be ignored, but at strong coupling they dominate and must be taken into account as well. In this chapter, we will encounter an example of tunneling effects in gauge theories, called *instantons*. For example, the 1-instanton contribution to the partition function of an SU(N) gauge theory on \mathbb{R}^4 is proportional to

$$\exp\left\{-\frac{8\pi}{g^2}\right\} \ . \tag{1.2}$$

As we see above, the instanton—which tunnels between gauge-inequivalent vacua contributes to the partition function in a manner that is *non-perturbative* in the coupling constant; that is, in the limit $g^2 \rightarrow 0$, these effects cannot be represented as a Taylor series in powers of g^2 . Section 1.1 will discuss instantons solutions to the Yang-Mills field equations, and the manner in which they may be constructed.

Since the bulk of this thesis is devoted to the study of $\mathcal{N} = 2$ supersymmetric gauge theories, in Section 1.2 we discuss the required background in supersymmetry.

1.1 Instantons

In gauge theories, instantons are defined as gauge field configurations that are finiteaction solutions to the classical equations of motion. They describe quantum mechanical tunneling between gauge-inequivalent vacua.

Consider Yang-Mills theory with gauge group SU(N) on four-dimensional Euclidean flat space, with the action

$$S_{\rm YM} = -\frac{1}{2g^2} \int d^4x \, {\rm Tr} \, F^2 \,, \qquad (1.3)$$

where $F = F_{\mu\nu} = F^a_{\mu\nu}T_a$, the generators T_a of the Lie algebra of SU(N) are traceless anti-Hermitian matrices satisfying the algebra

$$[T_a, T_b] = f_{abc} T_c , \qquad (1.4)$$

and with the convention $\text{Tr} T_a T_b = -\frac{1}{2} \delta_{ab}$, the action is positive definite. The classical equations of motion read

$$\nabla_{\mu}F_{\mu\nu} = 0 , \qquad (1.5)$$

where ∇_{μ} is the gauge covariant derivative, defined as $\nabla_{\mu} = \partial_{\mu} + A_{\mu}$, and the field

strength in terms of the gauge covariant derivative is

$$F_{\mu\nu} = [\nabla_{\mu}, \nabla_{\nu}] \quad , \tag{1.6}$$

$$=\partial_{\mu}A_{\nu} - \partial_{\nu}A_{\mu} + [A_{\mu}, A_{\nu}] \quad . \tag{1.7}$$

We are interested in solutions to the above equation that have finite action, and in order to achieve this we will require that the field strength vanishes at infinity; it is easy to show that if the gauge fields are asymptotically pure gauge, i.e.

$$A_{\mu} \to U^{-1} \partial_{\mu} U$$
 as $|x|^2 \to \infty$, (1.8)

for some $U \in SU(N)$, then the field strength vanishes at infinity. Gauge fields with these boundary conditions may be classified according to their *instanton number*, defined as

$$k = -\frac{1}{16\pi^2} \int \mathrm{d}^4 x \,\mathrm{Tr}\, F_{\mu\nu} \widetilde{F}_{\mu\nu} \,\,, \qquad (1.9)$$

where $\tilde{F}_{\mu\nu} = \frac{1}{2} \epsilon_{\mu\nu\rho\sigma} F_{\rho\sigma}$. The integrand above can be written as the divergence of a topological current, and consequently reduces to an integral over the 3-sphere at infinity. The instanton number k counts the number of times the gauge group wraps this 3-sphere, and is a gauge invariant quantity that breaks up the space of gauge fields into different topological sectors. Within each of these sectors, we will now show that the field configurations that minimize the action have either self-dual (+) or anti-self-dual (-) field strengths:

$$F_{\mu\nu} = \pm \widetilde{F}_{\mu\nu} \ . \tag{1.10}$$

We do this by writing the action as a square of a sum plus the instanton number:

$$S = -\frac{1}{2g^2} \int d^4x \,\mathrm{Tr}\,F^2 \,, \tag{1.11}$$

$$= -\frac{1}{4g^2} \int \mathrm{d}^4 x \,\mathrm{Tr}\,(F \mp \widetilde{F})^2 \mp \frac{1}{2g^2} \int \mathrm{d}^4 x \,\mathrm{Tr}\,F\widetilde{F} \,\,, \tag{1.12}$$

$$\geq \mp \frac{1}{2g^2} \int \mathrm{d}^4 x \,\mathrm{Tr} \, F \widetilde{F} \,\,, \tag{1.13}$$

$$\geq \pm \frac{8\pi^2}{g^2} |k| \quad . \tag{1.14}$$

It is easy to see that the bound is saturated by (anti-)self-dual field strengths. Further, these (anti-)self-dual field strengths minimize the action, and thus automatically satisfy the classical equations of motion.

A more detailed discussion of instantons can be found in [32]. We now turn to the question of how to systematically construct all solutions to (1.10).

1.1.1 ADHM Construction

It is well-known that the group of Euclidean rotations SO(4) admits a decomposition into $SU(2)_l \times SU(2)_r$. We can intertwine the SO(4) and the $SU(2)_l \times SU(2)_r$ using the Pauli matrices:

$$x_{\alpha\dot{\beta}} = x^{\mu} \left(\sigma_{\mu}\right)_{\alpha\dot{\beta}} \quad , \tag{1.15}$$

where α, β, \cdots and $\dot{\alpha}, \dot{\beta}, \cdots$ are the SU(2)_l and SU(2)_r indices respectively, and run over {1,2}. Our conventions for the σ_{μ} matrices are

$$\sigma_{\mu} = (1, -i\tau_1, -i\tau_2, -i\tau_3) , \qquad (1.16)$$

where the τ_i are the familiar Pauli matrices. It will be useful to keep in mind that the σ_{μ} matrices form a representation of the algebra of quaternions. The goal in this section is to construct self-dual field strengths, which we will do using the ADHM construction [9] (see also [33, 34]). For gauge group SU(N) and instanton number k, start with a matrix

$$\Delta_{\dot{\alpha}} = \mathbf{A}_{\dot{\alpha}} + \mathbf{B}^{\alpha} x_{\alpha \dot{\alpha}} , \qquad (1.17)$$

which is a $(N + 2k) \times 2k$ complex matrix with maximal rank. It is important for what follows that the above matrix is linear in the spacetime coordinate x, that we work with a quaternionic representation of x, and that the matrix $\Delta_{\dot{\alpha}}$ satisfy a factorization constraint we will soon introduce. Next, consider the $(2k + N) \times N$ matrix v(x) that satisfies the constraint

$$\left(\Delta^{\dagger}\right)^{\dot{\alpha}} v = 0 , \qquad (1.18)$$

and the normalization condition

$$v^{\dagger}v = 1_N . \tag{1.19}$$

With this data, the gauge field is given by

$$A_{\mu} = v^{\dagger} \partial_{\mu} v . \qquad (1.20)$$

We will now show that the field strength associated to this connection is in fact selfdual. For this construction to work, we need to impose the factorization condition

$$\left(\Delta^{\dagger}\right)^{\dot{\alpha}} \Delta_{\dot{\beta}} = \delta^{\dot{\alpha}}_{\dot{\beta}} f^{-1} , \qquad (1.21)$$

where f is an invertible $k \times k$ matrix that commutes with quaternions. We can now construct a projection operator orthogonal to v

$$P = 1_{2k+N} - vv^{\dagger} , \qquad (1.22)$$

which admits a more fruitful representation as

$$P = \Delta_{\dot{\alpha}} f \left(\Delta^{\dagger} \right)^{\alpha} . \tag{1.23}$$

One can check that with either of these representations, $P^2 = P$, and Pv = 0 due (respectively) to the factorization condition (1.21) and the constraint (1.18). Given this projection operator and the definition of the gauge field in (1.20), we can write the connection in terms of this projection operator as

$$F_{\mu\nu} = \partial_{\mu}v^{\dagger} P \,\partial_{\nu}v - (\mu \leftrightarrow \nu) \quad . \tag{1.24}$$

The conceit now is to use the aforementioned linearity of $\Delta_{\dot{\alpha}}$, i.e. shifting all the derivatives from v to $\Delta_{\dot{\alpha}}$ using the derivative of (1.18). Finally, since f is diagonal in the spinor indices, we find that the corresponding field strength may be written as

$$F_{\mu\nu} = 4i \left(\sigma_{\mu\nu}\right)_{\alpha}^{\ \beta} v^{\dagger} \mathbf{B}^{\alpha} f \, \mathbf{B}^{\dagger}_{\beta} v , \qquad (1.25)$$

which is manifestly self-dual, due to the appearance of the self-dual tensor $(\sigma_{\mu\nu})^{\beta}_{\alpha}$, defined as

$$\left(\sigma_{\mu\nu}\right)_{\alpha}^{\ \beta} = \frac{1}{4} \left(\sigma_{\mu}\bar{\sigma}_{\nu} - \sigma_{\nu}\bar{\sigma}_{\mu}\right)_{\alpha}^{\ \beta} , \qquad (1.26)$$

and $\bar{\sigma}_{\mu} = (\sigma_{\mu})^{\dagger}$. A similar procedure may be used to construct anti-self-dual solutions to the Yang-Mills field equations.

The factorization condition (1.21) in terms of the submatrices **A** and **B** takes the form

$$2\left(\mathbf{A}^{\dagger}\right)^{\dot{\alpha}}\mathbf{A}_{\dot{\beta}} = \delta^{\dot{\alpha}}_{\dot{\beta}}\left(\mathbf{A}^{\dagger}\right)^{\dot{\gamma}}\mathbf{A}_{\dot{\gamma}} , \qquad (1.27)$$

$$2\left(\mathbf{B}^{\dagger}\right)_{\alpha}\mathbf{B}^{\beta} = \delta^{\beta}_{\alpha}\left(\mathbf{B}^{\dagger}\right)_{\gamma}\mathbf{B}^{\gamma} , \qquad (1.28)$$

$$\left(\mathbf{B}^{\dagger}\right)_{\alpha}\mathbf{A}_{\dot{\beta}} = \left(\mathbf{A}^{\dagger}\right)_{\dot{\beta}}\mathbf{B}_{\alpha} . \tag{1.29}$$

There is a large amount of gauge freedom in the choice of matrix Δ , i.e. the above conditions are left invariant under the transformations

$$\Delta_{\dot{\alpha}} \mapsto U \Delta_{\dot{\alpha}} M$$
 and $v \mapsto U v$, (1.30)

where U is an $(N + 2k) \times (N + 2k)$ unitary matrix and M is invertible. This gauge freedom will allow us to send the matrices **A** and **B** into what are called their "canonical" forms:

$$\mathbf{B} = \begin{pmatrix} 0\\ 1_k \otimes 1_2 \end{pmatrix} , \qquad (1.31)$$

and the free data is contained in the matrices \mathbf{A} and v which can be written as

$$\mathbf{A} = (\mathbf{A}_{1}, \mathbf{A}_{2}) = \begin{pmatrix} S_{1} & S_{2} \\ X^{\mu} \otimes \sigma_{\mu} \end{pmatrix} \quad \text{and} \quad v = \begin{pmatrix} T \\ Q_{\alpha} \end{pmatrix} . \quad (1.32)$$

Here, each of the matrices have spacetime transformation properties:

- $S_{\dot{\alpha}}$ is a right-handed spinor,
- X^{μ} is a vector,
- T is a scalar, and
- Q_{α} is a left-handed spinor.

The gauge transformations (1.30) have the following effects on the above data:

$$S_{\dot{\alpha}} \mapsto U_N S_{\dot{\alpha}} U_k^{-1} , \qquad (1.33)$$

$$X^{\mu} \mapsto U_k X^{\mu} U_k^{-1} , \qquad (1.34)$$

$$T \mapsto U_N T$$
, (1.35)

$$Q_{\alpha} \mapsto U_k Q_{\alpha} , \qquad (1.36)$$

where $U_k \in U(k)$ and $U_N \in U(N)$ respectively. We can summarize this gauge freedom in the form of a quiver diagram, but let us try and understand the impact of the factorization condition (1.21) on the canonical forms presented above. In addition to the requirement of hermiticity for X^{μ} , the following *ADHM equations* are implied

$$\mu_i = \left(\mathbf{A}^{\dagger}\right)^{\dot{\alpha}} \left(\tau_i\right)_{\dot{\alpha}}^{\dot{\beta}} \mathbf{A}_{\dot{\beta}} = 0 \ . \tag{1.37}$$

These relations may be massaged into a more familiar form: identify

$$I = S_{2}^{\dagger} , \qquad (1.38)$$

$$J = S_{i}$$
, (1.39)

$$B_1 = X^0 - iX^3 , \qquad (1.40)$$

$$B_2 = -iX^1 + X^2 , (1.41)$$

in terms of which the ADHM equations may be written as

$$\mu_{\mathbb{R}} = -\mu^3 = II^{\dagger} - J^{\dagger}J + \left[B_1, B_1^{\dagger}\right] + \left[B_2, B_2^{\dagger}\right] = 0 , \qquad (1.42)$$

$$\mu_{\mathbb{C}} = \frac{1}{2}(\mu^1 - i\mu^2) = IJ + [B_1, B_2] = 0.$$
(1.43)

The general solution to the ADHM equations is not known except for some cases with small values of the instanton number [7], and because of this it will be useful to restate the ADHM construction as follows: consider the space of maps (I, J, B_1, B_2) , each a linear operator such that

$$I: \mathbb{C}^N \to \mathbb{C}^k , \qquad (1.44)$$

$$J: \mathbb{C}^k \to \mathbb{C}^N , \qquad (1.45)$$

$$B_{1,2}: \mathbb{C}^k \to \mathbb{C}^k \ . \tag{1.46}$$

The k-instanton moduli space \mathcal{M}_k may be thought of as the space of all such linear

operators modulo identifications via the gauge transformations (1.33)–(1.36). This characterization of the instanton moduli space will serve us well in Section 2.3.2, where we discuss the manner in which integrals over \mathcal{M}_k are computed. We now turn to the discussion of supersymmetric gauge theories, and the manner in which they may be constructed.

1.2 Supersymmetry

Quantum field theories on Minkowski space usually take as the spacetime symmetry group the Poincaré group, composed of translations and Lorentz rotations generated by momentum and angular momentum respectively. Supersymmetry expands this list of spacetime symmetries to include spin- $\frac{1}{2}$ generators: the left- and right-handed Weyl spinor supercharges Q_{α}^{I} and $\bar{Q}_{\dot{\beta},J}$. These supercharges commute with the momentum and transform as spinors under the Lorentz group. The indices $I, J \in$ $\{1, \dots, \mathcal{N}\}$ and correspond to the \mathcal{N} supersymmetries—this is at times referred to as the \mathcal{N} -extended supersymmetry algebra. In this thesis, we will focus on the case $\mathcal{N} = 2$, but for the purposes of this chapter we will also discuss the case $\mathcal{N} = 1$.

The centrally extended $\mathcal{N} = 2$ supersymmetry algebra is given by

$$\left\{Q^{I}_{\alpha}, \bar{Q}_{\dot{\beta},J}\right\} = 2\left(\sigma^{\mu}\right)_{\alpha\dot{\beta}} P_{\mu}\delta^{I}_{J} , \qquad (1.47)$$

$$\left\{Q^{I}_{\alpha}, Q^{J}_{\beta}\right\} = 2\sqrt{2}\epsilon_{\alpha\beta}\epsilon^{IJ}Z , \qquad (1.48)$$

$$\left\{\bar{Q}_{\dot{\alpha},I},\bar{Q}_{\dot{\beta},J}\right\} = 2\sqrt{2}\epsilon_{\dot{\alpha}\dot{\beta}}\epsilon_{IJ}Z \ . \tag{1.49}$$

Here, the central charge of the supersymmetry algebra is Z. It is straightforward to verify that the $\mathcal{N} = 1$ supersymmetry algebra does not admit a central extension.

1.2.1 Field Representations

It is natural to expect fields to sit in representations of the supersymmetry algebra, just as in Poincaré invariant quantum field theory, fields sit in representations of the Lorentz algebra. Since the generators of supersymmetry transformations mix bosonic and fermionic fields—as exemplified in the commutator (1.47)—fields in supersymmetric theories are grouped into *supermultiplets*. A convenient way to package this information is to collect all the fields in a supermultiplet into a *superfield* that lives on an enlarged spacetime that has both bosonic and fermionic coordinates—this is called *superspace*. The action of the supersymmetry generators has a natural interpretation in terms of superspace. We will begin our discussion with the $\mathcal{N} = 1$ superspace, where in addition to the usual spacetime coordinates x^{μ} , we have two anticommuting coordinates θ_{α} and $\bar{\theta}_{\dot{\beta}}$.

In general, a superfield will have more component fields than is required to obtain an irreducible representation of the supersymmetry algebra, so we have to impose supersymmetry-invariant constraints on these superfields. Different choices of constraints give rise to different supermultiplets, and we will now see a couple of relevant examples.

Chiral Multiplet

The physical fields that constitute the $\mathcal{N} = 1$ chiral multiplet are a complex scalar and Weyl spinor. One can package these fields into the chiral superfield Φ which admits an expansion in the anticommuting Grassmann coordinates as

$$\Phi(y,\theta) = \phi(y) + \sqrt{2}\,\theta\psi(y) + \theta^2\,F(y) , \qquad (1.50)$$

where Φ depends on superspace coordinate θ and the chiral combination $y^{\mu} = x^{\mu} + i\theta\sigma^{\mu}\bar{\theta}^{,1}$ This special dependence on chiral combinations of superspace coordinates may be encoded in the constraint

$$\bar{D}_{\dot{\alpha}}\Phi = 0 , \qquad (1.51)$$

where $\bar{D}_{\dot{\alpha}}$ is a supercovariant derivative

$$\bar{D}_{\dot{\alpha}} = -\frac{\partial}{\partial \bar{\theta}^{\dot{\alpha}}} - \mathrm{i} \left(\sigma^{\mu}\right)_{\alpha \dot{\alpha}} \theta^{\alpha} \frac{\partial}{\partial x^{\mu}} \,. \tag{1.52}$$

The additional field F is an auxiliary field that is required in order to allow the supersymmetry algebra to close off-shell. The anti-chiral multiplet may be defined analogously, as

$$D_{\alpha}\bar{\Phi} = 0 , \qquad (1.53)$$

and D_{α} is defined in much the same way as (1.52).

Any function of (anti-)chiral superfields is also an (anti-)chiral superfield. A superpotential is any function of only chiral superfields, and admits the following expansion: for a collection of chiral superfields $\{\Phi_i\}$, we have

$$W(\Phi_i) = W + 2\sqrt{2} \frac{\partial W}{\partial \phi_i} \theta \psi + \theta^2 \left(\frac{\partial W}{\partial \phi_i} F_i - \frac{1}{2} \frac{\partial^2 W}{\partial \phi_i \partial \phi_j} \psi_i \psi_j \right) . \tag{1.54}$$

As we can see, superpotentials are responsible for coupling the various fields within the chiral supermultiplet.

¹We will be suppressing Weyl indices in most of this text. A summation convention is in effect: when reading $\theta \psi$, for example, it stands for $\theta^{\alpha} \psi_{\alpha}$. As another example involving both dotted and undotted spinors, $\theta \sigma^{\mu} \bar{\theta}$ stands for $\theta^{\alpha} (\sigma^{\mu})_{\alpha \dot{\alpha}} \bar{\theta}^{\dot{\alpha}}$.

Vector Multiplet

It is often of physical interest to study theories with gauge fields. Start with a superfield V that is real, i.e.

$$V^{\dagger} = V . \qquad (1.55)$$

In the Wess-Zumino gauge, a vector superfield admits the following expansion into its component fields:

$$V = -\theta \sigma^{\mu} \bar{\theta} A_{\mu} + i\theta^2 \bar{\theta} \bar{\lambda} - i\bar{\theta}^2 \theta \lambda + \frac{1}{2} \theta^2 \bar{\theta}^2 D , \qquad (1.56)$$

where A_{μ} is a non-abelian gauge field, λ and $\overline{\lambda}$ are Weyl fermions (typically called gauginos) and D is an auxiliary field. Since these fields belong to the same multiplet, it follows that all of them sit in the adjoint representation of the gauge group. One can check that the gauge field strength may be defined as

$$W_{\alpha} = \frac{1}{8} \bar{D}^2 \left(e^{2V} D_{\alpha} e^{-2V} \right) , \qquad (1.57)$$

which when expanded in terms of its component fields has the form (we are suppressing the index α for convenience)

$$W = -i\lambda + \theta D - \frac{i}{2}\sigma^{\mu}\bar{\sigma}^{\nu}\theta F_{\mu\nu} + \theta^{2}\sigma^{\mu}\nabla_{\mu}\bar{\lambda} , \qquad (1.58)$$

where $F_{\mu\nu}$ is a nonabelian field strength and ∇_{μ} is a gauge covariant derivative.

1.2.2 Actions

Consider the following Lagrangian with $\mathcal{N} = 1$ supersymmetry:

$$\mathcal{L} = \frac{1}{8\pi} \operatorname{Im} \operatorname{Tr} \left[\tau \int d^2 \theta \ W^{\alpha} W_{\alpha} \right] + \int d^2 \theta \ d^2 \bar{\theta} \ \Phi^{\dagger} e^{-2V} \Phi + \int d^2 \theta \ W(\Phi) + \int d^2 \bar{\theta} \ \bar{W}(\Phi^{\dagger}) \ .$$
(1.59)

In the above equation, τ is the complexified Yang-Mills coupling

$$\tau = \frac{\theta}{2\pi} + i\frac{4\pi}{g^2} . \qquad (1.60)$$

It is instructive to expand the above Lagrangian out in terms of the component fields to see the kinds of interactions between the gauge fields, fermions, and scalars. As we will soon explain, the requirement of $\mathcal{N} = 2$ supersymmetry forces the superpotential W to be zero. The final Lagrangian then reduces to one with just the first two lines of (1.59), which when expanded in terms of the component fields gives

$$\mathcal{L} = \frac{1}{g^2} \operatorname{Tr} \left[-\frac{1}{4} F_{\mu\nu} F^{\mu\nu} - i\lambda \sigma^{\mu} \nabla_{\mu} \bar{\lambda} + \frac{1}{2} D^2 \right] + \frac{\theta}{32\pi^2} F_{\mu\nu} \widetilde{F}^{\mu\nu}$$

$$+ \operatorname{Tr} \left[\left| \nabla_{\mu} \phi \right|^2 - i \bar{\psi} \bar{\sigma}^{\mu} \nabla_{\mu} \psi + F^{\dagger} F - g \left(\phi^{\dagger} [D, \phi] - i \sqrt{2} \phi^{\dagger} \{\lambda, \psi\} + i \sqrt{2} \bar{\psi} [\bar{\lambda}, \phi] \right) \right]$$

$$(1.61)$$

We recognize among the terms in the above equation various familiar expressions: kinetic energies for the gauge fields, gauginos, quarks, scalars, and the second Chern class $F_{\mu\nu}\tilde{F}^{\mu\nu}$ which (after integration over spacetime) counts the instanton number of the gauge field configuration, in addition to various interaction terms.

$\mathcal{N} = 2$ Vector Multiplet

An $\mathcal{N} = 2$ vector multiplet is built up out of an $\mathcal{N} = 1$ chiral multiplet and an $\mathcal{N} = 1$ vector multiplet, and has a gauge field, two Weyl fermions, and a scalar. Now, the most general $\mathcal{N} = 1$ action we have written down in (1.59) cannot be $\mathcal{N} = 2$ supersymmetric as the superpotential term only couples to the ψ field, and not the λ field. In an $\mathcal{N} = 2$ vector multiplet, the fermions ψ and λ form an R-symmetry doublet; consequently, the requirement of $\mathcal{N} = 2$ supersymmetry forces the superpotential to be zero. For the same reasons, the kinetic terms of the gauginos and quarks should have the same coefficients. Thus, the full $\mathcal{N} = 2$ supersymmetric Lagrangian in terms of the $\mathcal{N} = 1$ superfields is

$$\mathcal{L} = \frac{1}{8\pi} \operatorname{Im} \operatorname{Tr} \left[\tau \left(\int \mathrm{d}^2 \theta \ W^{\alpha} W_{\alpha} + 2 \int \mathrm{d}^2 \theta \ \mathrm{d}^2 \bar{\theta} \ \Phi^{\dagger} \mathrm{e}^{-2V} \Phi \right) \right] .$$
(1.62)

In (1.61), we see that the auxiliary fields are still present. The equations of motion for auxiliary fields is algebraic; consequently, they are not dynamical and may be replaced by solutions to their equations of motion. Once this is done, we find that this yields a bosonic potential

$$\mathcal{L}_{\text{aux.}} = -\frac{1}{2g^2} \operatorname{Tr} \left[\phi^{\dagger}, \phi\right]^2 , \qquad (1.63)$$

which tells us that in order for supersymmetry to remain unbroken at low energies, it is both necessary and sufficient that ϕ and ϕ^{\dagger} commute, i.e. that ϕ sits in the Cartan subalgebra of the gauge Lie algebra.

$\mathcal{N} = 2$ Superspace

The 2-extended supersymmetry enjoyed by the action (1.62) is manifest when working in the $\mathcal{N} = 2$ superspace. A discussion of the same will allow us to introduce the notion of a *prepotential* corresponding to an $\mathcal{N} = 2$ gauge theory.

The $\mathcal{N} = 2$ superspace requires a pair of coordinates $(\tilde{\theta}, \tilde{\tilde{\theta}})$ over and above the $\mathcal{N} = 1$ superspace, thus coordinatized by $(x^{\mu}, \theta, \bar{\theta}, \tilde{\theta}, \tilde{\tilde{\theta}})$. A chiral superfield in this superspace admits the expansion

$$\Psi = \Phi(\tilde{y}, \theta) + \sqrt{2}\,\tilde{\theta}W(\tilde{y}, \theta) + \tilde{\theta}^2 G(\tilde{y}, \theta) , \qquad (1.64)$$

where $\tilde{y}^{\mu} = y^{\mu} + i\tilde{\theta}\sigma^{\mu}\bar{\tilde{\theta}}$, Φ is an $\mathcal{N} = 1$ chiral superfield as given in (1.50), W is an $\mathcal{N} = 1$ vector multiplet field strength as given in (1.58), and G is

$$G(\tilde{y},\theta) = -\frac{1}{2} \int d^2\bar{\theta} \left[\Phi(\tilde{y} - i\theta\sigma\bar{\theta},\theta,\bar{\theta}) \right]^{\dagger} \exp^{-2gV(\tilde{y} - i\theta\sigma\bar{\theta},\theta,\bar{\theta})} , \qquad (1.65)$$

required to eliminate unphysical degrees of freedom [35]. In terms of this $\mathcal{N} = 2$ superfield, the action (1.62) is

$$\mathcal{L} = \operatorname{Im} \left[\frac{\tau}{4\pi} \int d^2 \theta \, d^2 \tilde{\theta} \, \frac{1}{2} \operatorname{Tr} \Psi^2 \right] \,. \tag{1.66}$$

It is of fundamental importance to note that the above action is holomorphic in Ψ . In fact, one can show that any action of the form

$$\mathcal{L} = \frac{1}{4\pi} \operatorname{Im} \left[\int d^2 \theta \, d^2 \tilde{\theta} \, \mathcal{F}(\Psi) \right] \,, \qquad (1.67)$$

is $\mathcal{N} = 2$ supersymmetric, and the function \mathcal{F} is referred to as the *prepotential* of the theory. Considerations of renormalizability fix the prepotential of the classical $\mathcal{N} = 2$ super Yang-Mills as

$$\mathcal{F}_{\text{class.}} = \operatorname{Tr} \frac{\tau}{2} \Psi^2 . \qquad (1.68)$$

When studying the low-energy effective action, the above prepotential receives corrections at 1-loop, and there exists an infinite tower of instanton corrections as well. In the following section, we will derive the form of 1-loop corrections in pure $\mathcal{N} = 2$ super Yang-Mills theory.

1.2.3 1-Loop Corrections

Nonrenormalization theorems [5] that use the holomorphy of the prepotential allow us to conclude that it is perturbatively exact at 1-loop. In this section we will review the computation of this 1-loop contribution to the pure $\mathcal{N} = 2$ gauge theory with gauge group SU(N) [36]. The low-energy effective theory consists of SU(N) broken down to its maximal torus U(1)^{N-1}, and has an effective action given in terms of a prepotential \mathcal{F} as

$$\frac{1}{4\pi} \operatorname{Im} \left[\int \mathrm{d}^4\theta \, \frac{\partial \mathcal{F}(A)}{\partial A^i} \bar{A}^i + \frac{1}{2} \int \mathrm{d}^2\theta \, \frac{\partial^2 \mathcal{F}(A)}{\partial A^i \, \partial A^j} \, \left(W^i \right)_{\alpha} \left(W^j \right)^{\alpha} \right] \,, \tag{1.69}$$

where as we have emphasized, \mathcal{F} is a holomorphic function of the variables A_i , and the index *i* runs over the Cartan directions. Here, A^i is the $\mathcal{N} = 1$ chiral multiplet in the $\mathcal{N} = 2$ vector multiplet whose adjoint scalar takes the value a_i .

Supersymmetric field theories have an R-symmetry that rotates the spinor supercharges into each other; in (1.47)–(1.49) the R-symmetry rotates the supercharge indices I, J, \dots . For $\mathcal{N} = 2$ supersymmetry, this R-symmetry is $SU(2)_R \times U(1)_R$. The fields in an $\mathcal{N} = 2$ vector multiplet organize themselves into representations of the $SU(2)_R$ symmetry: the scalar and the gauge field each form a singlet, and the fermions form a doublet. The $U(1)_R$ symmetry is inherited from the $\mathcal{N} = 1$ superspace, and is anomalous; the chiral current is no longer conserved in the quantum theory, and instead we find

$$\partial_{\mu} j_{5}^{\mu} = -\frac{N}{8\pi^{2}} F_{\mu\nu} \widetilde{F}^{\mu\nu} . \qquad (1.70)$$

This in turn tells us that under an infinitesimal $U(1)_R$ transformation, the La-

grangian changes as

$$\delta \mathcal{L} = -\frac{\alpha N}{8\pi^2} F \widetilde{F} . \qquad (1.71)$$

This change is proportional to the second Chern class, and is quantized, i.e.

$$\frac{1}{32\pi^2} \int F\widetilde{F} \in \mathbb{Z} , \qquad (1.72)$$

and consequently the U(1)_R is broken down to \mathbb{Z}_{4N} . Given that the change in the Lagrangian is as above, we obtain

$$\frac{1}{16\pi} \operatorname{Im} \left[\mathcal{F}''(\mathrm{e}^{2\mathrm{i}\alpha}A)(-F^2 + \mathrm{i}F\widetilde{F}) \right] = \frac{1}{16\pi} \operatorname{Im} \left[\mathcal{F}''(\Phi)(-F^2 + \mathrm{i}F\widetilde{F}) \right] - \frac{\alpha N}{8\pi^2} F\widetilde{F} ,$$
(1.73)

which effectively constrains the prepotential to satisfy

$$\mathcal{F}''(e^{2i\alpha}A) = \mathcal{F}''(A) - \frac{2\alpha N}{\pi} , \qquad (1.74)$$

and for infinitesimal α , we find

$$\frac{\partial^3 \mathcal{F}}{\partial A^3} = \frac{\mathrm{i}N}{\pi} \frac{1}{A} \,. \tag{1.75}$$

As an example, we may set N = 2 and solve the above equation for $\mathcal{F}_{1-\text{loop}}$

$$\mathcal{F}_{1\text{-loop}} = \frac{\mathrm{i}}{2\pi} A^2 \log\left(\frac{A^2}{\Lambda^2}\right) , \qquad (1.76)$$

where Λ is the dynamically generated scale.

The nonrenormalization theorem does not preclude the possibility of the prepotential receiving contributions from instantons. Indeed, as anticipated in [5], the k-instanton contribution to the prepotential takes the form

$$\mathcal{F}_{\text{inst.}} = \sum_{k=1}^{\infty} \mathcal{F}_k \, \left(\frac{\Lambda}{A}\right)^{4k} A^2 \,. \tag{1.77}$$

To "count instantons" is to compute the functions \mathcal{F}_k . The multi-instanton calculus [7] is able to do this for small k, but it is difficult to extend these computations to higher instanton number or more general gauge group. We discuss two more fruitful methods to count instantons in the following chapter: Seiberg-Witten theory, and equivariant localization.

Seiberg-Witten, Calogero-Moser, and Nekrasov

2

This chapter will develop various techniques that serve as a substrate to later chapters of this thesis: Seiberg-Witten theory, the elliptic Calogero-Moser system, and the method of equivariant localization as applied to the ADHM moduli space.

2.1 Seiberg-Witten Theory

As we have emphasised in the previous chapter, to *solve* an $\mathcal{N} = 2$ gauge theory is to compute all its non-perturbative (instanton) corrections. In 1994, Seiberg and Witten [1, 2] presented an essentially geometrical solution to this problem, that was soon generalized to higher-rank gauge groups [37, 38, 39]. In this section, we will review the Seiberg-Witten solution for pure $\mathcal{N} = 2$ gauge theory (i.e. without matter) with gauge group SU(2).

2.1.1 Moduli Space

As we have seen in (1.63), the Lagrangian of this theory has a scalar potential

$$V = -\frac{1}{2g^2} \operatorname{Tr} \left[\phi, \phi^{\dagger}\right]^2 , \qquad (2.1)$$

On general grounds [40], we expect that supersymmetry-preserving vacua have V = 0, which requires (at least) that ϕ is valued in the Cartan subalgebra of SU(2); we choose

$$\phi = \frac{1}{2} a \tau_3 , \qquad (2.2)$$

where $a \in \mathbb{C}$ and τ_3 is the familiar Pauli matrix. The space spanned by a serves as a 1-complex dimensional parameter space of the low-energy effective theory: the *moduli space* of vacua, where the gauge group is broken down to U(1). Thus, at low energies one has a theory of electromagnetism. For this reason, this space of vacua is called the Coulomb moduli space. Notice, however, that Weyl reflections send $a \to -a$, and consequently this parametrization of the Coulomb moduli space is not gauge invariant. It is helpful to use as coordinates the gauge invariant moduli

$$u = \operatorname{Tr} \phi^2 , \qquad (2.3)$$

$$\simeq \frac{1}{2} a^2 , \qquad (2.4)$$

to leading order in the semiclassical expansion. Classically, it might seem reasonable to expect that at u = 0, the full gauge symmetry is restored. We will soon see that this does not happen, but for the moment we note that a fully quantum mechanical characterization of the Coulomb moduli space requires that we consider vacuum expectation values, i.e.

$$u = \left\langle \operatorname{Tr} \phi^2 \right\rangle \,, \tag{2.5}$$

Might this moduli space be characterized by a metric as well? The answer is yes. From (1.67) we see that that the Lagrangian of this theory in terms of the $\mathcal{N} = 1$ superspace may be arrived at by doing the $\tilde{\theta}$ integrals, which yields

$$\mathcal{L} = \frac{1}{4\pi} \left[\int d^4\theta \, \frac{\partial \mathcal{F}}{\partial \Phi} \, \Phi^\dagger + \int d^2\theta \, \frac{1}{2} \frac{\partial^2 \mathcal{F}}{\partial \Phi^2} \, W^\alpha W_\alpha \right] \,, \tag{2.6}$$

where Φ is the $\mathcal{N} = 1$ chiral superfield, related to *a* as the vacuum expectation value of its scalar component. When viewed as a σ -model on the space of fields, we can read off the metric on the Coulomb moduli space:

$$ds^{2} = \operatorname{Im} \frac{\partial^{2} \mathcal{F}}{\partial a^{2}} da \, d\bar{a} = \operatorname{Im} \tau(a) \, da \, d\bar{a} \, .$$
(2.7)

Physically, $\text{Im }\tau(a)$ must be positive because it is the coefficient of the gauge kinetic term, and without positivity the action would be ill-defined. Further, the metric $\text{Im }\tau(a)$ is a harmonic function as it is the second derivative of the holomorphic prepotential, and consequently cannot have a local minimum. This in turn means it is not bounded from below! We conclude from this that if the metric is globally defined, it cannot be positive-definite, indicating in turn that such a description of the metric can at best be locally valid.

2.1.2 Electric-Magnetic Duality

We will require a complimentary description of the theory in regions of the moduli space where $\text{Im } \tau(a) < 0$. Consider terms involving the gauge fields in (2.6), which may be written as

$$\frac{1}{32\pi} \operatorname{Im} \int \tau(a) \left(F + \mathrm{i}\widetilde{F}\right)^2 = \frac{1}{16\pi} \operatorname{Im} \int \tau(a) \left(F^2 + \mathrm{i}\widetilde{F}F\right) , \qquad (2.8)$$

We want to treat F as an independent field, and in order to do so we need to implement the Bianchi identity dF = 0 as a constraint. We do this using a Lagrange multiplier as follows: add to the action a term of the form

$$\frac{1}{8\pi} \int (V_{\rm D})_{\mu} \epsilon^{\mu\nu\rho\sigma} \partial_{\nu} F_{\rho\sigma} = \frac{1}{8\pi} \int \widetilde{F}_D F , \qquad (2.9)$$

where $F_{\rm D} = \mathrm{d}V_{\rm D}$. On adding the above term to the action (2.8) and integrating over F, we get

$$\frac{1}{32\pi} \operatorname{Im} \int \left(-\frac{1}{\tau}\right) \left(F_{\rm D} + \mathrm{i}\widetilde{F}_{\rm D}\right)^2 = \frac{1}{16\pi} \operatorname{Im} \int \left(-\frac{1}{\tau}\right) \left(F_{\rm D}^2 + \mathrm{i}\widetilde{F}_{\rm D}F_{\rm D}\right) , \qquad (2.10)$$

and we see that the effect of this duality transformation is to replace the gauge field A_{μ} that couples to electric charges with a dual gauge field V_D that couples to magnetic monopoles. Simultaneously, the complex gauge coupling is transformed as

$$\tau \to \tau_{\rm D} = -\frac{1}{\tau} \tag{2.11}$$

Our action has, in addition to this duality symmetry, another invariance. The Yang-Mills action is invariant under the transformation $\tau \to \tau + b$ where $b \in \mathbb{R}$, which in turn sends $\theta \to \theta + 2\pi b$. In order for the path integral over gauge fields to be well-defined, it is necessary that $b \in \mathbb{Z}$. Together, the invariances

$$\tau \to \tau + 1 \quad \text{and} \quad \tau \to -\frac{1}{\tau} ,$$
 (2.12)

generate the modular group $SL(2, \mathbb{Z})$. A general element of the modular group acts on τ as a Möbius transformation. It is natural, in light of this electric-magnetic duality, to ask if the "electric" vevs *a* have "magnetic" duals.

Introduce

$$a_{\rm D} = \frac{\partial \mathcal{F}}{\partial a} , \qquad (2.13)$$

and note that this relation implies

$$\tau = \frac{\partial a_{\rm D}}{\partial a} \ . \tag{2.14}$$

We can use this definition to write the metric on the moduli space (2.7) in a sym-

metric form

$$\mathrm{d}s^2 = \mathrm{Im}\,\mathrm{d}a_\mathrm{D}\,\mathrm{d}a\ .\tag{2.15}$$

This metric may be written in an $SL(2, \mathbb{R})$ invariant manner, by introducing the vector $a^{\alpha} = (a_{\rm D}, a)$ and the antisymmetric ϵ -tensor as

$$ds^{2} = -\frac{i}{2}\epsilon_{\alpha\beta}\frac{da^{\alpha}}{du}\frac{da^{\beta}}{d\bar{u}}du\,d\bar{u}\;.$$
(2.16)

 $SL(2,\mathbb{R})$ is generated by the matrices

$$\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix} , \qquad (2.17)$$

Once again, for the path integral to be well-defined we require that $b \in \mathbb{Z}$, and we see that the theory has a duality group that is the modular group.

2.1.3 Monodromies, Monopoles, and Dyons

At large |a|, the theory is asymptotically free, and the prepotential is well-approximated by its 1-loop evaluation, which from (1.76) is

$$\mathcal{F}(a) = \frac{\mathrm{i}}{2\pi} a^2 \log \frac{a^2}{\Lambda^2} , \qquad (2.18)$$

and in terms of which we may compute the magnetic vev:

$$a_{\rm D} = \frac{2\mathrm{i}a}{\pi} \log \frac{a}{\Lambda} + \frac{\mathrm{i}a}{\pi} \ . \tag{2.19}$$

A closed loop in the *u*-plane around the point at infinity—where the above expressions are valid—mixes the electric and magnetic vevs in the vector $(a_{\rm D}, a)^{\rm T}$ via the

monodromy matrix

$$M_{\infty} = \begin{pmatrix} -1 & 2\\ 0 & -1 \end{pmatrix} , \qquad (2.20)$$

thereby signaling that there must be a non-trivial monodromy about some finite-upoint in the moduli space. The simplest guess we alluded to earlier—that there is a singularity in the moduli space at the origin, which is consistent with the $u \rightarrow -u$ symmetry—cannot be true, because if it was then the two monodromies would commute, rendering the coordinate a globally valid. As we have argued, this cannot be true, so we require that there be at least two points on the Coulomb moduli space with non-trivial monodromies about them. Let these points be at $\pm u_0$.

Singularities in the moduli space are points at which ordinarily massive particles go massless, since at these points, the notion of a Wilsonian effective action is no longer applicable. These singularities cannot be because of gauge bosons becoming massless, as this would imply a conformally invariant theory in the infrared. However, conformal invariance implies either u = 0 or that $\text{Tr } \phi^2$ is a dimension zero operator i.e. the unit operator, neither of which is true. Since there are no other elementary multiplets in the theory, we are forced to conclude that the singularities in the moduli space are due to collective excitations — monopoles and dyons going massless.

The mass of particles in a centrally extended $\mathcal{N} = 2$ supersymmetric theory is bounded by its central charge — this is a straightforward consequence of the supersymmetry algebra for massive irreducible representations — and we find

$$M \ge \sqrt{2}|Z| \ . \tag{2.21}$$

There also exists a BPS bound on the mass, which in turn is determined in terms of electric and magnetic charges [41]. For a low-energy effective theory determined by a prepotential, the central charge is given by

$$Z = n_e a + n_m a_\mathrm{D} , \qquad (2.22)$$

where n_e and n_m count the units of electric and magnetic charge. Let us suppose that at the point $+u_0$, a magnetic monopole goes massless. We may then conclude that

$$a_{\rm D}(u_0) = 0$$
 . (2.23)

Monopoles are described by $\mathcal{N} = 2$ hypermultiplets and do not couple to "electric" fields in a local manner, although they do so locally with the dual "magnetic" fields, in much the same way that electrons (also described by $\mathcal{N} = 2$ hypermultiplets) couple locally to "electric" fields. Schematically,

electrons
$$\longleftrightarrow (\Phi, W)$$
, (2.24)

monopoles
$$\longleftrightarrow (\Phi_{\rm D}, W_{\rm D})$$
, (2.25)

where (Φ, W) are the $\mathcal{N} = 1$ multiplets that constitute an $\mathcal{N} = 2$ chiral multiplet and $(\Phi_{\rm D}, W_{\rm D})$ are their magnetic duals.

To summarize, near the u_0 singularity we have a low-energy theory consisting of a "magnetic" $\mathcal{N} = 2$ vector multiplet coupled to light monopoles, which is precisely $\mathcal{N} = 2$ SQED. From the 1-loop β -function, the magnetic coupling is

$$\tau_{\rm D} \sim -\frac{\mathrm{i}}{\pi} \log a_{\rm D} \ . \tag{2.26}$$

Since $a_{\rm D}$ is a good coordinate near u_0 , let us suppose

$$a_{\rm D}(u) \sim c(u - u_0)$$
, (2.27)

from which we can conclude that

$$a(u) \sim a_0 + \frac{\mathrm{i}}{\pi} c(u - u_0) \log(u - u_0) ,$$
 (2.28)

and a monodromy about u_0 is governed by the matrix

$$M_{u_0} = \begin{pmatrix} 1 & 0 \\ -2 & 1 \end{pmatrix} . (2.29)$$

The third monodromy matrix may be obtained from M_{u_0} above and M_{∞} from (2.20) via the condition $M_{u_0}M_{-u_0} = M_{\infty}$, and is

$$M_{-u_0} = \begin{pmatrix} -1 & 2\\ -2 & 3 \end{pmatrix} . (2.30)$$

The particle that goes massless at this singularity has electric and magnetic charges governed by the relation $(n_m, n_e)M_{-u_0} = (n_m, n_e)$, which imposes the condition $n_m = -n_e$.

Let us take a moment to summarize: we determined that our theory has a moduli space — the *u*-plane — of vacua equipped with a metric, whose expression in terms of the electric coordinate *a* is not globally valid. We concluded on the basis of this that the moduli space has singularities — three of them to be precise, at ∞ and $\pm u_0$, consistent with Weyl reflection symmetry — and determined the monodromies of the electric and magnetic coordinates about each of them. These singularities were interpreted as due to electrically and magnetically charged particles going massless. With these three monodromies, we have elucidated the structure of the Coulomb moduli space, and now move on to discuss the manner in which these insights may be used to solve the theory. For what follows, it will be important to keep in mind that for our theory to make sense physically, we require Im $\tau > 0$.

2.1.4 Solution

Here we briefly discuss the Seiberg-Witten solution. The monodromy matrices that we have determined in the previous section all generate a subgroup $\Gamma(2) \in SL(2, \mathbb{Z})$. Further, the *u*-plane with the singularities at ∞ and $\pm u_0$ is the same as the quotient of the upper-half plane $\mathbb{H}/\Gamma(2)$, and without loss of generality, we set $u_0 = \Lambda^2$. At this point, we draw on intuition from the theory of elliptic curves, where the space $\mathbb{H}/\Gamma(2)$ is the moduli space of a family of the following family of elliptic curves

$$y^{2} = (x - \Lambda^{2})(x + \Lambda^{2})(x - u) . \qquad (2.31)$$

Let us arrange our branch cuts such that we have one between $\pm \Lambda^2$ and the other running between u and ∞ . An elliptic curve has the topology of a torus, to which one can associate two independent homology cycles: let us call the cycle running once around the cut $(-\Lambda^2, \Lambda^2)$ the A cycle, and the cut going from $+\Lambda^2$ to u on the first sheet, and then cycling back from u to $+\Lambda^2$ on the second sheet the B cycle. Then the electric and magnetic vevs are identified with the cycles in the following manner:

$$a = \oint_{A} \lambda_{\rm SW} , \qquad (2.32)$$

$$a_{\rm D} = \oint_B \lambda_{\rm SW} , \qquad (2.33)$$

where λ_{SW} is a suitably chosen differential 1-form that is called the Seiberg-Witten differential.

What do the monodromies do in this picture, and what properties must λ_{SW} satisfy? When u executes a monodromy around one of the singular points, the cycles on the elliptic curve change into linear combinations of themselves, which in turn implies that the periods and dual periods are transformed into linear combinations of themselves, just as our u-plane analysis indicated. Further, as

$$\tau = \frac{\partial a_{\rm D}}{\partial a} = \frac{\mathrm{d}a_{\rm D}/\mathrm{d}u}{\mathrm{d}a/\mathrm{d}u} , \qquad (2.34)$$

we find that the complex structure parameter of the torus is identified with the complex gauge coupling. This quantity naturally satisfies $\text{Im } \tau > 0$, and thus we have self-consistently satisfied the requirement of positivity of the metric on the moduli space. For our elliptic curve (2.31) the Seiberg-Witten differential [1, 2] is

$$\lambda_{\rm SW} = \frac{1}{\sqrt{2\pi}} \frac{\sqrt{x-u}}{\sqrt{x^2 - \Lambda^4}} \,\mathrm{d}x \;, \qquad (2.35)$$

and it is easy to check that this choice is consistent with — indeed, fixed by! — the monodromy data we found for the electric and magnetic variables in the previous section.

As the integrals in question are elliptic, the periods may be evaluated in terms of hypergeometric functions. We can now invert a(u) and insert the result in the expression for a_D order-by-order in Λ . On integrating this, we recover the instanton expansion as in (1.77). In conclusion, for the pure $\mathcal{N} = 2$ super Yang-Mills theory with gauge group SU(2), the Seiberg-Witten solution at the first few orders in the instanton expansion [39] is

$$\mathcal{F}_1 = \frac{1}{2} ,$$
 (2.36)

$$\mathcal{F}_2 = \frac{5}{64} , \qquad (2.37)$$

$$\mathcal{F}_3 = \frac{3}{64} \ . \tag{2.38}$$

Some of these results were first checked against explicit multi-instanton calculus results [7], in addition to later being verified at higher orders via Nekrasov's equivariant localization [3, 4].
2.2 Calogero-Moser System

In this section we recall some concepts from the theory of integrable systems. After this, we introduce the Calogero-Moser integrable system and discuss its spectral curve.

2.2.1 A Brief Interlude: Integrable Systems

A mechanical system is said to be *integrable* if there exist as many integrals of motion (i.e. conserved quantities) as there are degrees of freedom. More precisely, if the mechanical system under investigation has n degrees of freedom, we require that there exist n functionally independent conserved quantities $I_k(x, p)$; this is embodied in the conditions

$$\{I_k, H\} = 0 \text{ and } \{I_k, I_\ell\} = 0.$$
 (2.39)

Once these conditions are satisfied we can canonically transform to action (I_k) and angle (ψ_k) variables that evolve linearly in time:

$$I_k(t) = I_k(0)$$
 and $\psi_k(t) = \psi_k(0) + c_k I_k t$. (2.40)

There exists a more fruitful (and fully equivalent) formulation of the criterion for integrability: a system is said to be integrable if one can find a *Lax pair* of $N \times N$ matrix-valued functions on phase space (L, M) such that the equation

$$\dot{L} = [L, M] , \qquad (2.41)$$

is equivalent to Hamilton's equations of motion. There are two points to note here.

The first is that the size N of the Lax matrices—for reasons that will become clear shortly—is not related to the dimension 2n of phase space in a straightforward manner, and the best we can do is bound $n \leq N$. The second is that once found, Lax pairs are not unique: for any matrix S of the same size, it is easy to verify that the pair (L', M') defined by

$$L' = S^{-1}LS , (2.42)$$

$$M' = S^{-1}MS - S^{-1}\dot{S} . (2.43)$$

is also a Lax pair. Modulo this ambiguity, it may be verified that once a Lax pair is found, the integrals of motion are given by

$$I_k = \operatorname{Tr} L^k . (2.44)$$

From Cayley's theorem it is clear that the I_k with $k \ge N$ are not independent and can be expressed in terms of lower I_k .

It is often the case that one can find a family of Lax pairs labelled by a complex parameter called a *spectral parameter*. That is, the equation (2.41) is lifted to a one-parameter family of equations

$$\dot{L}(z) = [L(z), M(z)],$$
 (2.45)

which is equivalent to Hamilton's equations of motion for all values of z. To such a Lax pair with a spectral parameter, we may naturally associate a spectral curve

$$\Gamma = \{ (k, z) \in \mathbb{C} \times \mathbb{C} : \det [kI - L(z)] = 0 \} , \qquad (2.46)$$

and a differential 1-form

$$\mathrm{d}\lambda = k\,\mathrm{d}z\;.\tag{2.47}$$

2.2.2 Calogero-Moser Systems

The Calogero-Moser system is a family of integrable classical mechanical systems that have (n + 1) particles interacting pairwise

$$H = \frac{1}{2} \sum_{i=1}^{n+1} p_i^2 - \frac{m^2}{2} \sum_{i \neq j}^{n+1} V(x_i - x_j) , \qquad (2.48)$$

the potential function is either

rational :
$$V(x) = \frac{1}{x^2}$$
, (2.49)

trigonometric:
$$V(x) = \frac{1}{\sin^2 x}$$
, (2.50)

elliptic:
$$V(x) = \wp(x; \omega_1, \omega_2)$$
. (2.51)

The function $\wp(x;\omega_1,\omega_2)$ is called the Weierstraß elliptic function, and is defined as

$$\wp(x;\omega_1,\omega_2) = \frac{1}{x^2} + \sum_{(m,n)\neq(0,0)} \left(\frac{1}{(x+2m\omega_1+2n\omega_2)^2} - \frac{1}{(2m\omega_1+2n\omega_2)^2} \right) , \quad (2.52)$$

and is doubly periodic with periods $2\omega_1$ and $2\omega_2$. These three models are integrable in the sense we have discussed in the previous section—the equations of motion admit a representation in terms of a Lax pair with spectral parameter—and they are related to each other: taking the limit $\omega_2 \to \infty$ while holding ω_1 fixed interpolates between the elliptic and trigonometric Calogero-Moser systems; in addition to this, taking the subsequent limit $\omega_1 \to \infty$ yields the rational Calogero-Moser system. Thus,

elliptic CM $\xrightarrow{\omega_2 \to \infty}$ trigonometric CM $\xrightarrow{\omega_1 \to \infty}$ rational CM. (2.53)

It is easy to anticipate which of these will be of interest to supersymmetric gauge theories. We know that $\mathcal{N} = 2$ supersymmetric gauge theories have a complexified Yang-Mills coupling constant τ , which courtesy of the Seiberg-Witten solution may be interpreted as a modular parameter of an algebraic curve. If this object is protected from renormalization effects, it is reasonable to anticipate its importance at low energies, i.e. in the Seiberg-Witten curve. An example of such a theory is $\mathcal{N} = 2$ supersymmetric Yang-Mills theory coupled to a single adjoint hypermultiplet, since its β -function vanishes; this is referred to as the $\mathcal{N} = 2^*$ gauge theory. Further, it is natural to expect the corresponding integrable system to make reference to a modular parameter as well, and this leads us to conclude that the integrable system of interest is the elliptic Calogero-Moser system, whose two periods may be combined to give a modular parameter.

2.2.3 Lax Pairs for Elliptic Calogero-Moser Systems

The elliptic Calogero-Moser system has the following Lax pair with spectral parameter [42]:

$$L_{ij}(z) = p_i \delta_{ij} - m(1 - \delta_{ij}) \Phi(x_i - x_j; z) , \qquad (2.54)$$

$$M_{ij}(z) = d_i(x)\delta_{ij} + m(1 - \delta_{ij})\Phi'(x_i - x_j; z) , \qquad (2.55)$$

where

$$d_i(x) = m \sum_{k \neq i} \wp(x_i - x_k) ,$$
 (2.56)

and the function $\Phi(x; z)$ is the Lamé function, defined as

$$\Phi(x;z) = \frac{\sigma(z-x)}{\sigma(x)\sigma(z)} e^{x\zeta(z)} , \qquad (2.57)$$

is a solution to the Lamé equation

$$\left(\frac{\mathrm{d}^2}{\mathrm{d}z^2} - \wp(x)\right)\Phi(x;z) = 2\wp(z)\Phi(x;z) , \qquad (2.58)$$

and the functions ζ and σ are auxiliary Weierstrass functions, related to \wp as

$$\wp(z) = -\zeta'(z) \quad \text{and} \quad \zeta(z) = \frac{\sigma'(z)}{\sigma(z)} .$$
(2.59)

The spectral curve corresponding to the elliptic Calogero-Moser system is the same as the Seiberg-Witten curve of the $\mathcal{N} = 2^*$ gauge theory with gauge group U(N). This correspondence will be expanded upon and developed further in Chapter 3.

2.3 Equivariant Localization

In this section we outline the manner in which k-instanton contributions to the prepotential of a pure $\mathcal{N} = 2$ gauge theory may be computed.

2.3.1 Topological Twisting

We start with the pure $\mathcal{N} = 2$ super Yang-Mills action (1.61). The procedure we are going to introduce is referred to as *topological twisting* [43], and begins with the identification of the global symmetry group of the theory, which is

$$H = \underbrace{\mathrm{SU}(2)_l \times \mathrm{SU}(2)_r}_{K} \times \mathrm{SU}(2)_R , \qquad (2.60)$$

where the underbraced part K = SO(4) is the group of Euclidean rotations in \mathbb{R}^4 , i.e. the analogue of the Lorentz group. The question of how to embed K in H has many possible answers, of which the above resolution is one; indeed, one can use non-standard embeddings as well. Let $SU(2)_d = \text{diag } SU(2)_r \times SU(2)_R$, the diagonal subgroup, and let

$$K' = \operatorname{SU}(2)_l \times \operatorname{SU}(2)_d . \tag{2.61}$$

Topological twisting has the effect of turning the supercharge indices I, J, \cdots into spacetime indices. Since we have changed what we call the Lorentz group, our fields have to be reconstituted into irreducible representations of this new Lorentz group. This regrouping of degrees of freedom breaks up our spinor supercharges into tensor, vector, and scalar components defined as

$$\bar{Q}^+_{\mu\nu} = \bar{\sigma}^{\dot{\alpha}I}_{\mu\nu} \bar{Q}_{\dot{\alpha}I} , \qquad (2.62)$$

$$Q_{\mu} = \bar{\sigma}^{\alpha I}_{\mu} Q_{\alpha I} , \qquad (2.63)$$

$$\bar{Q} = \epsilon^{\dot{\alpha}I} \bar{Q}_{\dot{\alpha}I} , \qquad (2.64)$$

Topological twisting has the happy consequence that all fields will now transform as integer spin fields under K'; for example, the fermions become

$$\bar{\psi}_{\mu\nu} = \bar{\sigma}^{\dot{\alpha}I}_{\mu\nu} \bar{\psi}_{\dot{\alpha}I} , \qquad (2.65)$$

$$\psi_{\mu} = \bar{\sigma}_{\mu}^{\alpha I} \psi_{\alpha I} , \qquad (2.66)$$

$$\bar{\psi} = \epsilon^{\dot{\alpha}I} \bar{\psi}_{\dot{\alpha}I} . \tag{2.67}$$

The Lagrangian (1.61) can now be written in \overline{Q} -exact form [44]. That is, in terms of the corresponding actions in component form, we find

$$\mathcal{S}_{\rm YM} = \mathcal{S}_{\rm top.} + \bar{Q} \underbrace{\left[\frac{\tau}{16\pi} \int d^4 x' \operatorname{Tr} \left(F^-_{\mu\nu} \bar{\psi}^{\mu\nu} - i\sqrt{2} \psi^{\mu} \nabla_{\mu} \phi^{\dagger} - i \bar{\psi} [\phi^{\dagger}, \phi] \right) \right]}_{\mathcal{S}_{\rm YM}}, \quad (2.68)$$

where $S_{\text{top.}}$ is the topological term (1.9) which is \bar{Q} -closed, and

$$F_{\mu\nu}^{-} = \frac{1}{2} \left(F_{\mu\nu} - \tilde{F}_{\mu\nu} \right) .$$
 (2.69)

In (2.68), let us call the underbraced expression $V_{\rm YM}$. It is possible to show that any correlation function of the form

$$\langle O \rangle = \int \mathrm{D}X \, O \,\mathrm{e}^{-S_{\mathrm{YM}}} \,, \qquad (2.70)$$

where DX is the path integral measure over all fields, is insensitive to the addition of \bar{Q} -exact terms in the action. To see this, deform the action like so

$$\langle O \rangle' = \int \mathrm{D}X \, O \,\mathrm{e}^{-S_{\mathrm{YM}} + \bar{Q}\delta V} , \qquad (2.71)$$

$$= \langle O \rangle + \int \mathbf{D}X \ O \,\mathrm{e}^{-S_{\mathrm{YM}}} \,\bar{Q} \delta V , \qquad (2.72)$$

$$= \langle O \rangle + \int \mathrm{D}X \,\bar{Q} \left(O \,\mathrm{e}^{-S_{\mathrm{YM}}} \delta V \right) \,, \qquad (2.73)$$

$$=\langle O \rangle$$
 . (2.74)

Here, we have used the fact that physical observables sit in the \bar{Q} -cohomology, so the correlator of \bar{Q} -exact terms identically vanishes, as in the second last step above.

It is possible to use this freedom to add \bar{Q} -exact terms to modify the action in such a way [44] that we get

$$S_{\rm YM} = S_{\rm top.} + \int d^4 x' \,{\rm Tr} \,\left(-t^2 F^-_{\mu\nu} F^-_{\mu\nu} + \cdots\right) \,,$$
 (2.75)

where the \cdots are terms of $\mathcal{O}(t)$. Since the path integral is insensitive to the value of t, for convenience we may evaluate it in the limit $t \to \infty$, which gives us the equations of motion

$$F_{\mu\nu}^{-} = 0 . (2.76)$$

From (2.69), this is nothing but the self-dual Yang-Mills equations

$$F_{\mu\nu} = \widetilde{F}_{\mu\nu} \ . \tag{2.77}$$

Now, in Section 1.1.1 we studied the ADHM moduli space: the set of solutions to the ADHM equations modulo gauge transformations. This method constructs all self-dual solutions to the Yang-Mills equations. Notice that in the above discussion we started with an infinite-dimensional path integral over the space of fields, and *localized* it onto the finite-dimensional space of solutions to the self-dual equations, i.e. precisely the ADHM moduli space!

Since the partition function of a gauge theory may be thought of as the expectation value of the unit operator, we can conclude that the object of interest is the volume of the k-instanton moduli space \mathcal{M}_k for all k, and thus the partition function we wish to compute may be written formally as

$$Z = \sum_{k=0}^{\infty} q^k \oint_{\mathcal{M}_k} 1 .$$
 (2.78)

where the parameter q formally keeps track of the contributions from each kinstanton sector. For an asymptotically free theory,

$$q = \Lambda^b , \qquad (2.79)$$

where Λ is the dynamically generated scale, and b is the coefficient of the β -function for the Yang-Mills coupling constant. Of course, the conformally invariant theories we are considering have b = 0, and in this case

$$q = e^{2\pi i \tau} , \qquad (2.80)$$

where τ is the bare coupling constant.

2.3.2 The Ω Background

The computation of volumes of k-instanton moduli spaces is difficult because the spaces \mathcal{M}_k are non-compact, i.e. naively, they have a divergent volume. Thus, a physically sensible result will require that we adopt a regularization scheme: the Ω background. We will do these integrals using localization theorems, and will introduce them in a series of examples, closely following [45].

Supersymmetric Quantum Mechanics and Localization

Consider a supersymmetric particle on \mathbb{C}^2 coordinatized by (z, w), such that the coordinates z, w are supersymmetry invariant, and (\bar{z}, ψ_z) and (\bar{w}, ψ_w) are supersymmetric partners. The angular momenta orthogonal to the two orthogonal complex planes is such that $J_1 = 1$ for z and $J_2 = 1$ for w. The partition function of the theory is given by

$$Z(\beta;\epsilon_1,\epsilon_2) = \operatorname{Tr}_{\mathcal{H}_{susv}} e^{i\beta\epsilon_1 J_1} e^{i\beta\epsilon_2 J_2} , \qquad (2.81)$$

where \mathcal{H}_{susy} is the space of supersymmetric states, which (apart from a Gaussian that isn't holomorphic) can only be polynomials in z and w i.e.

$$\mathcal{H}_{\text{susy}} = \bigoplus_{m,n \ge 0} \mathbb{C} \, z^m w^n \, . \tag{2.82}$$

The partition function is then given by

$$Z(\beta;\epsilon_1,\epsilon_2) = \frac{1}{1 - e^{i\beta\epsilon_1}} \frac{1}{1 - e^{i\beta\epsilon_2}} .$$
(2.83)

Next, let us consider a charged particle moving in a magnetic flux j that pierces a \mathbb{CP}^1 , in turn coordinatized by z. The supersymmetry-invariant Hilbert space this

time is

$$\mathcal{H}_{\text{susy}} = \bigoplus_{k=0}^{2j} \mathbb{C} \, z^k (\partial_z)^{\otimes j} \,, \qquad (2.84)$$

If z has charge 1 under the generator J, then the partition function

$$Z(\beta;\epsilon) = \operatorname{Tr}_{\mathcal{H}_{susy}} e^{i\beta\epsilon J} , \qquad (2.85)$$

$$=\sum_{k=-j}^{+j} e^{\mathbf{i}\beta\epsilon k} , \qquad (2.86)$$

which can be rewritten in a more enlightening form

$$Z(\beta;\epsilon) = \frac{\mathrm{e}^{\mathrm{i}\beta\epsilon\,j}}{1 - \mathrm{e}^{-\mathrm{i}\beta\epsilon}} + \frac{\mathrm{e}^{-\mathrm{i}\beta\epsilon\,j}}{1 - \mathrm{e}^{\mathrm{i}\beta\epsilon}} \ . \tag{2.87}$$

These examples hint at a localization theorem, which we will now state without proof: for a supersymmetric particle moving on a complex manifold M of dimension d with an isometry $U(1)^n$ generated by the set $\{J_k\}_{k=1}^n$ that has isolated fixed points, and in the presence of a magnetic flux—corresponding to a line bundle L on M, and where the supersymmetric states are holomorphic sections of this bundle—the partition function is

$$Z(\beta; \epsilon_1, \cdots, \epsilon_n) = \operatorname{Tr}_{\mathcal{H}}(-1)^F e^{i\beta \sum_k \epsilon_k J_k} , \qquad (2.88)$$

$$=\sum_{p} \frac{\mathrm{e}^{\mathrm{i}\beta\sum_{k}j(p)_{k}\epsilon_{k}}}{\prod_{a=1}^{d} \left(1 - \mathrm{e}^{\mathrm{i}\beta\sum_{k}\ell(p)_{k,a}\epsilon_{k}}\right)} \,. \tag{2.89}$$

The theorem gives a result that is "localized" onto fixed points p of the U(1)ⁿ isometry on M. The quantities j(p) and $\ell(p)$ are defined in terms of characters—traces over the fiber and tangent space at p, as

$$\operatorname{Tr}_{L|_{p}} \mathrm{e}^{\mathrm{i}\beta\sum_{k}\epsilon_{k}J_{k}} = \mathrm{e}^{\mathrm{i}\beta\sum_{k}j(p)_{k}\epsilon_{k}} , \qquad (2.90)$$

$$\operatorname{Tr}_{\operatorname{T}M|_{p}} e^{\mathrm{i}\beta \sum_{k} \epsilon_{k} J_{k}} = \sum_{a=1}^{d} e^{\mathrm{i}\beta \sum_{k} \ell(p)_{k,a} \epsilon_{k}} .$$
(2.91)

It is convenient to adopt the following convention: vector spaces, like people, shall henceforth be recognized by their characters. Thus, we write

$$L|_{p} = e^{i\beta\sum_{k}j(p)_{k}\epsilon_{k}} , \qquad (2.92)$$

$$TM|_{p} = \sum_{a=1}^{d} e^{i\beta \sum_{k} \ell(p)_{k,a}\epsilon_{k}} .$$
(2.93)

It is easy to check that the general expression (2.89) reproduces the partition functions (2.83) and (2.87).

Lifts and Twists

We return in this section to the study of $\mathcal{N} = 2$ gauge theories, characterized by a gauge group $G = \mathrm{SU}(N)$ with rank r = N - 1, a flavour summetry group Fwith hypermultiplets in the representation $R \oplus \overline{R}$. This data can also be used to define a five-dimensional $\mathcal{N} = 1$ gauge theory, which we will put on $\mathbb{C}^2 \times [0, \beta]$ and coordinatized by (z, w) for the two complex planes, and ξ^5 for the circle on identifying

$$(z, w, 0) \sim (\mathrm{e}^{\mathrm{i}\beta\epsilon_1} z, \mathrm{e}^{\mathrm{i}\beta\epsilon_2} w, \beta) .$$
 (2.94)

This is the Ω background. Observe in particular that the non-zero Ω background parameters (ϵ_1, ϵ_2) introduce an off-set in the identifications of the coordinates as we move once around the circle. The limit $\beta \to 0$ will be the four-dimensional limit, and the vacuum expectation values of the gauge field and the masses of the hypermultiplets when integrated along the circle are

$$\operatorname{diag}(\mathrm{e}^{\mathrm{i}\beta a_1},\cdots,\mathrm{e}^{\mathrm{i}\beta a_r}) \in \mathrm{U}(1)^r , \qquad (2.95)$$

$$\operatorname{diag}(\mathrm{e}^{\mathrm{i}\beta m_1}, \cdots, \mathrm{e}^{\mathrm{i}\beta m_f}) \in \mathrm{U}(1)^f .$$

$$(2.96)$$

We are interested in the supersymmetric partition function of this $\mathcal{N} = 1$ gauge theory on this background

$$Z(\beta,\Omega;a_1,\cdots,a_r;m_1,\cdots,m_f) = \operatorname{Tr}_{\mathcal{H}_{5d}}(-1)^F e^{\mathrm{i}\beta\left(\epsilon_1 J_1 + \epsilon_2 J_2 + \sum_{s=1}^r a_s Q_s + \sum_{s=1}^f m_s F_s\right)},$$
(2.97)

where $J_{1,2}$ generate spatial rotations in the two complex planes, $Q_{1,\dots,r}$ generate global gauge rotations, and $F_{1,\dots,f}$ generate flavour rotations. The trace itself is over the Hilbert space of the five-dimensional gauge theory.

Let us address the question of what gauge field configurations dominate the above trace. At each time slice, from considerations of energetics it is easy to see that gauge field configurations that are "close" to k-instanton configurations will contribute. In the five-dimensional picture, one can think of the ξ^5 direction as a time direction, in which case the collective coordinates describing the k-instanton configuration can be thought of as varying slowly along the circle. This can be approximated by a quantum mechanical particle moving on the k-instanton moduli space, and is called the moduli space approximation [46]. In supersymmetric theories this approximation is exact, and we have

$$Z = \sum_{k=0}^{\infty} e^{-\beta \frac{8\pi^2}{g^2}k} \operatorname{Tr}_{\mathcal{H}_k}(-1)^F e^{\mathrm{i}\beta \left(\epsilon_1 J_1 + \epsilon_2 J_2 + \sum_{s=1}^r a_s Q_s + \sum_{s=1}^f m_s F_s\right)}, \qquad (2.98)$$

where g is the five-dimensional gauge coupling constant, and \mathcal{H}_k is the Hilbert space of the supersymmetric particle moving on the k-instanton moduli space, whose bosonic part is the same as the ADHM moduli space \mathcal{M}_k , which has real dimension 4Nk. There are fermionic directions $\mathcal{V}(R)$ which come from the zero modes of hypermultiplets in the representation R, which have real dimension 2C(R)k. Here, C(R) is the quadratic Casimir in the representation R that is normalized such that it is equal to 2N for the adjoint representation.

Accounting for the spatial, gauge, and flavour rotations, if the fixed points p of the

 $U(1)^2 \times U(1)^r$ action were isolated, we can use the localization theorem to evaluate this integral.

$$Z = \sum_{k=0}^{\infty} e^{-\beta \frac{8\pi^2}{g^2}k} \sum_{p} \frac{\prod_{t=1}^{Ck} (1 - e^{i\beta w_t(p)})}{\prod_{t=1}^{2Nk} (1 - e^{i\beta v_t(p)})} , \qquad (2.99)$$

in analogy with (2.89). As before, the functions $v_t(p)$ and $w_t(p)$ are the characters

$$\mathcal{V}(R)|_{p} = \sum_{t=1}^{C(R)k} e^{i\beta w_{t}(p)} ,$$
 (2.100)

$$\mathcal{T}\mathcal{M}_k|_p = \sum_{t=1}^{2Nk} e^{\mathrm{i}\beta v_t(p)} , \qquad (2.101)$$

that are functions of the Ω deformation parameters, the vacuum expectation values of the adjoint scalar, and the masses of the hypermultiplets.

We wish to take the four-dimensional limit $\beta \to 0$ while keeping the Ω -deformation parameters and the a_s finite — this will yield a partition function for a *deformed* $\mathcal{N} = 2$ gauge theory, which will soon be related to the prepotential. Notice that in the sum over instantons, at fixed k there are more factors in the denominator than the numerator, so naively the limit $\beta \to 0$ will diverge. In order to extract sensible answers, we rescale our classical instanton contribution as

$$e^{-\frac{8\pi^2 k}{g^2}\beta} = \left[(-i\beta)^{2N-C(R)} q \right]^k .$$
 (2.102)

The limit $\beta \to 0$ is taken while keeping q (soon to be related to the instanton counting parameter) fixed. Finally, the four-dimensional limit is determined to be

$$Z = \sum_{k=0}^{\infty} q^k \sum_{p} \frac{\prod_{t=1}^{Ck} w_t(p)}{\prod_{t=1}^{2Nk} v_t(p)} , \qquad (2.103)$$

and the four-dimensional coupling is related to the five-dimensional coupling by

$$\frac{1}{g_{\rm YM}^2} = \frac{\beta}{g^2} \ . \tag{2.104}$$

This identification is nice because we can now interpret β^{-1} as an ultraviolet scale, and (2.102) may be understood as incorporating the logarithmic running of the fourdimensional Yang-Mills coupling. Indeed, the expression 2N - C(R) is the coefficient of the β -function, and the dynamical scale is related to q as

$$q = \Lambda^{2N - C(R)} {.} (2.105)$$

In this thesis, we will consider conformal theories, i.e. theories where

$$C(R) = 2N$$
 . (2.106)

which can be achieved when one has (i) a massive hypermultiplet in the adjoint representation, or (ii) 2N hypermultiplets in the fundamental representation. In these cases, the instanton counting parameter is dimensionless, and simply given by

$$q = e^{2\pi i \tau} , \qquad (2.107)$$

where τ is, as before, the bare complexified Yang-Mills coupling.

The expression (2.103) is quite formal, and it would be desirable to have a more explicit or constructive expression. We turn to this in the following section.

Deformations and Tableaux

Recall from Section 1.1.1 the ADHM construction of the k-instanton moduli space: the space of operators (I, J, B_1, B_2) that satisfy the ADHM equations

$$\vec{\mu} = (\mu_{\mathbb{R}}, \operatorname{Re} \mu_{\mathbb{C}}, \operatorname{Im} \mu_{\mathbb{C}}) = 0 , \qquad (2.108)$$

modulo gauge transformations, and consider the following modification [47, 48]:

$$\mathcal{M}_{k,t} = \{\mu_{\mathbb{C}}(x) = t \,|\, x \in X_k\} \,/ \mathrm{GL}(k) \,, \tag{2.109}$$

where the vector space X_k is defined by the character

$$X_k = \left(e^{-i\beta\epsilon_1} + e^{-i\beta\epsilon_2} \right) VV^* + W^*V + e^{-i\beta(\epsilon_1 + \epsilon_2)}V^*W , \qquad (2.110)$$

where as before $V = \mathbb{C}^k$ and $W = \mathbb{C}^N$ and it is straightforward to check that when t = 0, we recover the usual ADHM construction. The proposed modification was introduced [47] in order to resolve small-instanton singularities, and our strategy will be to apply the localization theorem to $\mathcal{M}_{k,t}$ and then take the limit $t \to 0$. From [49], we learn that the fixed points of the U(1)^{2+r} action on can be classified in the following manner, with the use of Young tableaux.

Fixed points on $\mathcal{M}_{k,t}$ are labelled by N Young diagrams $\mathbf{Y} = (Y_1, \cdots, Y_N)$ subject to the constraint that the total number of boxes in the collection \mathbf{Y} is k. In each of the Y_s , let us use the pair of positive integers (i, j) to specify the position of a particular box. Define

$$W_p = \sum_{s=1}^{N} e^{i\beta a_s}$$
, (2.111)

$$V_p = \sum_{s=1}^{N} \sum_{(i,j)\in Y_s} e^{i\beta\phi(a,s)} , \qquad (2.112)$$

where

$$\phi(a,s) = a + (1-i)\epsilon_1 + (1-j)\epsilon_2 . \qquad (2.113)$$

We then determine the functions $v(p)_t$ from

$$T\mathcal{M}_{k,t}|_{p} = W_{p}^{*}V_{p} + e^{i\beta(\epsilon_{1}+\epsilon_{2})}V_{p}^{*}W_{p} - (1 - e^{i\beta\epsilon_{1}})(1 - e^{i\beta\epsilon_{2}})V_{p}V_{p}^{*}$$
(2.114)

and $w(t)_p$ for the case of hypermultiplets in the fundamental and adjoint representations

$$\mathcal{V}(\mathbf{fund})|_p = \mathrm{e}^{-\mathrm{i}\beta m} V_p , \qquad (2.115)$$

$$\mathcal{V}(\mathbf{adj})|_p = e^{-i\beta m} T \mathcal{M}_{k,t}|_p . \qquad (2.116)$$

For each tableau $Y = (\lambda_1 \ge \lambda_2 \ge \cdots)$ where λ_i is the height of the *i*-th column, and $Y^T = (\lambda'_1 \ge \lambda'_2 \ge \cdots)$ is the transposed tableau. A box s = (i, j) has arm- and leg-lengths

$$A_Y(s) = \lambda_i - j , \qquad (2.117)$$

$$L_Y(s) = \lambda'_j - i . (2.118)$$

In terms of the arm- and leg-lengths defined above, we define a function E by

$$E(a, Y_a, Y_b, s) = a - \epsilon_1 L_{Y_b}(s) + \epsilon_2 \left(A_{Y_1}(s) + 1 \right) .$$
(2.119)

We now have all the ingredients to construct the Nekrasov partition function corresponding to the field content of the $\mathcal{N} = 2$ gauge theory under consideration. For example, the contribution of the SU(N) vector multiplet to the Nekrasov integrand is derived by plugging (2.111) and (2.112) into (2.114). The answer, in terms of the E functions we have defined above, is

$$z_{\text{vect.}}(\mathbf{a}, \mathbf{Y}) = \left[\prod_{i,j=1}^{N} \prod_{s \in Y_i} E(a_i - a_j, Y_i, Y_j, s) \prod_{t \in Y_j} (\epsilon_1 + \epsilon_2 - E(a_j - a_i, Y_j, Y_i, t))\right]^{-1}.$$
(2.120)

The contribution of the adjoint hypermultiplet is

$$z_{\text{adj.}}(\mathbf{a}, \mathbf{Y}, m) = \prod_{i,j=1}^{N} \prod_{s \in Y_i} \left(E(a_i - a_j, Y_i, Y_j, s) + m \right) \times \cdots$$
$$\cdots \times \prod_{t \in Y_j} \left(\epsilon_1 + \epsilon_2 - E(a_j - a_i, Y_j, Y_i, t) - m \right) . \tag{2.121}$$

The contribution from each fundamental matter multiplet is

$$z_{\text{fund.}}(\mathbf{a}, \mathbf{Y}, m) = \prod_{i=1}^{N} \prod_{s \in Y_i} (\phi(a_i, s) - m + \epsilon_1 + \epsilon_2) . \qquad (2.122)$$

These pieces may be put together in various combinations depending on the theory under consideration. For example, when interested in the partition function of an $\mathcal{N} = 2^*$ theory—a theory with a vector multiplet and an adjoint hypermultiplet—the partition function is constructed as

$$Z(\mathbf{a}, m; \epsilon_1, \epsilon_2) = \sum_{k=0}^{\infty} q^k \sum_{\mathbf{Y}} z_{\text{vec.}}(\mathbf{a}, \mathbf{Y}; \epsilon_1, \epsilon_2) z_{\text{adj.}}(\mathbf{a}, \mathbf{Y}, m; \epsilon_1, \epsilon_2) .$$
(2.123)

This is the Nekrasov partition function, denoted Z_{Nek} .

2.3.3 The Prepotential

The above procedure yields a partition function for a *deformed* $\mathcal{N} = 2$ gauge theory, i.e. one that depends on the Ω deformation parameters. Nekrasov and Okounkov [3, 4] showed that the prepotential of the undeformed $\mathcal{N} = 2$ gauge theory may be recovered from the Nekrasov partition function as

$$F = -\lim_{\epsilon_1, \epsilon_2 \to 0} \epsilon_1 \epsilon_2 \log Z_{\text{Nek}} . \qquad (2.124)$$

Our strategy throughout this thesis will be to first compute quantities in the deformed gauge theory, and then take the limit in which the Ω deformation parameters vanish. More details relevant to the specific theories we will turn our attention to will be contained in the chapters that follow.

Chiral Rings

In this chapter, we study $\mathcal{N} = 2^*$ theories with gauge group U(N) and use equivariant localization to calculate the quantum expectation values of the simplest chiral ring elements. These are expressed as an expansion in the mass of the adjoint hypermultiplet, with coefficients given by quasi-modular forms of the S-duality group. Under the action of this group, we construct combinations of chiral ring elements that transform as modular forms of definite weight. As an independent check, we confirm these results by comparing the spectral curves of the associated Hitchin system and the elliptic Calogero-Moser system. We also propose an exact and compact expression for the 1-instanton contribution to the expectation value of the chiral ring elements.

3.1 Introduction

 $\mathcal{N} = 2$ super Yang-Mills (SYM) theories in four dimensions are an extraordinarily fertile ground to search for exact results. Indeed, their non-perturbative behaviour can be tackled both via the Seiberg-Witten (SW) description of their low-energy effective theory [1, 2], and via the microscopic computation of instanton effects by means of localization techniques [3, 50, 4, 51, 52, 20]. Understanding the farreaching consequences of strong/weak coupling dualities in the effective theory has always been a crucial ingredient in the SW approach. On the other hand, the same dualities can also be exploited in the microscopic description through the associated modular structure. The comparison of how these dualities may be used to constrain physical observables in the two approaches is one of the main themes of this paper.

Among the $\mathcal{N} = 2$ models, much effort has been devoted to gaining a deeper understanding of superconformal theories and their massive deformations (see for example the collection of reviews [53] and references therein), where many different approaches have been investigated. Among these we can mention the relation to integrable models [54], the 2d/4d AGT correspondence [55, 56], the use of matrix model techniques [57, 58] and the link to topological string amplitudes through geometric engineering [59, 30, 60]. Furthermore, the pioneering work of Gaiotto [22] has taught us that the duality properties are of the utmost relevance.

In this chapter we focus on $\mathcal{N} = 2^*$ theories, which we briefly review in Section 3.2. Besides the gauge vector multiplet, they contain an adjoint hypermultiplet of mass m that interpolates between the $\mathcal{N} = 4$ SYM theories (when $m \to 0$) and the pure $\mathcal{N} = 2$ SYM theories (when $m \to \infty$). The $\mathcal{N} = 2^*$ theories inherit from the $\mathcal{N} = 4$ models an interesting action of the S-duality group; in particular, their prepotential satisfies a modular anomaly equation which greatly constrains its form. Modular anomaly relations in gauge theories were first noticed in [10] and are related to the holomorphic anomaly equations that occur in topological string theories on local Calabi-Yau manifolds [61, 62, 63, 64]. These equations have been studied in a variety of settings, for example in an Ω background [65, 66, 67, 68, 30, 69, 24, 23, 70, 25, 71, 72, 73], from the point of view of the AGT correspondence [74, 75, 76, 77], in the large-N limit [25], and in SQCD models with fundamental matter [24, 23, 26, 78].

Recently, the modular anomaly equation for $\mathcal{N} = 2^*$ theories with arbitrary gauge groups has been linked in a direct way to S-duality [11, 12, 13]. This approach has led to a very efficient way of determining the mass expansion of the prepotential in terms of: *i*) quasi-modular functions of the gauge coupling and *ii*) the vacuum expectation values a_u of the scalar field Φ of the gauge multiplet such that only particular combinations, defined purely in terms of sums over the root lattice of the corresponding Lie algebra, appear. These results have been checked against explicit computations using equivariant localization.

In this work, we take the first steps towards showing that similar modular structures also exist for other observables of $\mathcal{N} = 2^*$ gauge theories. We choose to work with U(N) gauge groups, and consider the quantum expectation values

$$\langle \operatorname{Tr} \Phi^n \rangle$$
 . (3.1)

The supersymmetry algebra implies that correlators of chiral operators factorize and can therefore be expressed in terms of the expectation values in (3.1).¹

A priori, it is not obvious that these chiral observables exhibit modular behaviour. However, we show that it is always possible to find combinations that transform as modular forms of definite weight under the non-perturbative duality group $SL(2, \mathbb{Z})$. These combinations have a natural interpretation as modular-covariant coordinates on the Coulomb moduli space, and can be analysed using two different techniques: *i*) the SW approach via curves and differentials, and *ii*) equivariant localization combined with the constraints arising from S-duality.

For $\mathcal{N} = 2^*$ theories there are many distinct forms of the SW curve that capture different properties of the chiral observables. In one approach, due to Donagi and Witten [14, 15], the SW curve has coefficients A_n that have a natural interpretation as modular-covariant coordinates on the Coulomb moduli space. Thus, this approach provides us with a natural setting to study the elliptic and modular properties

¹More general correlators involving also one anti-chiral operator have recently been considered in [79].

of the observables (3.1). Another form of the SW curve was found by using the relation with integrable systems [16]. For the $\mathcal{N} = 2^*$ theory, the relevant curve was proposed by D'Hoker and Phong [17, 6], who used the close relation between the gauge theory and the elliptic Calogero-Moser system [18]. In this second formulation, the coefficients of the spectral curve of the integrable system are interpreted as symmetric polynomials built out of the quantum chiral ring elements (3.1). A third form of the SW curve for the $\mathcal{N} = 2^*$ theories was proposed by Nekrasov and Pestun [21] together with an extension to general quiver models. In Section 3.3 we review and relate the first two descriptions of the SW curve which are suitable for our purposes. This comparison will lead to interesting relationships between the coefficients of the respective curves. Along the way, we will find it necessary to modify the analysis of [14] in a subtle but important way.

It is clearly desirable to work with chiral observables that in the classical limit coincide with the symmetric polynomials built out of the vacuum expectation values a_u . As we discuss in Section 3.4, this can be done in two ways. The first is to compute the period integrals in the Donagi-Witten form of the curve as a series expansion in the mass m of the adjoint hypermultiplet. Inverting this expansion order by order in m gives us an expression for the A_n in terms of the a_u . The second way is to postulate that the A_n have a definite modular weight under the S-duality group, and use the well-understood action of S-duality to derive a modular anomaly equation that recursively determines them up to modular pieces. In this derivation, it is crucial that the prepotential and hence the dual periods of the $\mathcal{N} = 2^*$ theory are known in terms of quasi-modular forms. In both ways it turns out that the chiral observables can be expressed in terms of quasi-modular forms and of particular functions of the a_u involving only sums over the weight and root lattices of the Lie algebra $\mathfrak{u}(N)$, generalizing those appearing in the prepotential.

In Section 3.5 we test our findings against explicit microscopic computations of the

observables (3.1) using equivariant localization techniques [3, 50, 4, 51, 52, 20] (for further technical details see also [80]). We find that the chiral observables computed using localization can be matched with those obtained from the SW curves by a redefinition of the chiral ring elements. Such a redefinition contains only a finite number of terms and is exact both in the mass of the hypermultiplet and in the gauge coupling. It is well known that the localization results for the chiral observables do not, in general, satisfy the classical chiral ring relations [19, 20, 21]. Strikingly, we show that the redefinition of the chiral ring elements which allow the matching of the two sets of results can be interpreted as a judicious choice of coordinates on the Coulomb moduli space in which the *classical* chiral ring relations are naturally satisfied.

In Section 3.6, we focus on the 1-instanton contributions and, just as it was done for the prepotential in [11, 12], we manage to resum the mass expansion to obtain an exact expression involving only sums over roots and weights of the corresponding Lie algebra.

3.2 Review: $\mathcal{N} = 2^* U(N)$ SYM Theories

The $\mathcal{N} = 2^*$ SYM theories are massive deformations of the $\mathcal{N} = 4$ SYM theories arising when the adjoint hypermultiplet is given a mass m. The classical vacua of these theories on the Coulomb branch are parametrized by the expectation values of the scalar field Φ in the vector multiplet, which in the U(N) case is

$$\langle \Phi \rangle \equiv a = \operatorname{diag}\left(a_1, a_2, \dots, a_N\right).$$
 (3.2)

When the complex numbers a_u are all different, the gauge group is broken to its maximal torus $U(1)^N$. The low-energy effective action of this abelian theory is com-

pletely determined by a single holomorphic function F(a), called the prepotential. It consists of a classical term

$$F_{\text{class}} = i\pi\tau a^2 \equiv i\pi\tau \sum_{u=1}^N a_u^2 , \qquad (3.3)$$

where τ is the complexified gauge coupling

$$\tau = \frac{\theta}{2\pi} + i \frac{4\pi}{g^2} , \qquad (3.4)$$

and a quantum part

$$f = F_{1-\text{loop}} + F_{\text{inst}} \tag{3.5}$$

accounting for the 1-loop and instanton corrections.

The 1-loop term $F_{1-\text{loop}}$ is τ -independent and takes the simple form (see for instance [6])

$$F_{1-\text{loop}} = -\frac{1}{4} \sum_{\alpha \in \Psi} \left[(\alpha \cdot a)^2 \log \left(\frac{\alpha \cdot a}{\Lambda} \right)^2 - (\alpha \cdot a + m)^2 \log \left(\frac{\alpha \cdot a + m}{\Lambda} \right)^2 \right] , \quad (3.6)$$

where Λ is an arbitrary scale and α is an element of the root system Ψ of the gauge algebra. The first and seconds terms in (4.8) are, respectively, contributions from the vector multiplet and the massive hypermultiplet.

The instanton corrections to the prepotential are proportional to q^k , where

$$q = e^{2\pi i\tau} \tag{3.7}$$

is the instanton counting parameter and k is the instanton number. These nonperturbative terms can be calculated either using the SW curve and corresponding holomorphic differential λ_{SW} [1, 2], or by a microscopic evaluation of the prepotential using localization [3, 50, 4, 51, 52, 20]. In the SW approach, besides the "electric" variables a_u , one introduces dual or "magnetic" variables defined by

$$a_u^{\rm D} = \frac{1}{2\pi i} \frac{\partial F}{\partial a_u} . \tag{3.8}$$

The pairs $(a_u, a_u^{\rm D})$ describe the period integrals of the holomorphic differential $\lambda_{\rm SW}$ over cycles of the Riemann surface defined by the SW curve. More precisely, one has

$$a_u = \oint_{A_u} \lambda_{SW}$$
 and $a_u^D = \oint_{B_u} \lambda_{SW}$. (3.9)

Here, the A- and B-cycles form a canonically conjugate symplectic basis of cycles with intersection matrix $A_u \cap B_v = \delta_{uv}$.

For the $\mathcal{N} = 2^* \operatorname{U}(N)$ theory, the non-perturbative S-duality group has a simple embedding into the symplectic duality group $\operatorname{Sp}(4N,\mathbb{Z})$ of the Riemann surface. In particular, the S-transformation acts by exchanging electric and magnetic variables, while inverting the coupling constant, namely

$$S(a_u) = a_u^{\rm D}, \qquad S(a_v^{\rm D}) = -a_v, \qquad S(\tau) = -\frac{1}{\tau}.$$
 (3.10)

Along with the T-transformation, given by

$$T(a_u) = a_u, \qquad T(a_v^{\rm D}) = a_v^{\rm D} + a_v, \qquad T(\tau) = \tau + 1, \qquad (3.11)$$

one generates the modular group $SL(2,\mathbb{Z})$.

To discuss the $\mathcal{N} = 2^*$ prepotential and the action of the duality group on it, it is convenient to organize its quantum part (3.5) as an expansion in powers of the hypermultiplet mass, as

$$f = \sum_{n=1}^{\infty} f_n \, m^{2n} \, . \tag{3.12}$$

Notice that only even powers of m occur in this expansion as a consequence of the \mathbb{Z}_2 symmetry that sends $m \to -m$. In order to write the coefficients f_n in a compact form, it is useful to introduce the following lattice sums

$$C^{p}_{n;m_{1}\cdots m_{\ell}} = \sum_{\lambda \in \mathcal{W}} \sum_{\alpha \in \Psi_{\lambda}} \sum_{\beta_{1} \neq \cdots \neq \beta_{\ell} \in \Psi_{\alpha}} \frac{(\lambda \cdot a)^{p}}{(\alpha \cdot a)^{n} (\beta_{1} \cdot a)^{m_{1}} \cdots (\beta_{\ell} \cdot a)^{m_{\ell}}}$$
(3.13)

where \mathcal{W} is the set of weights λ of the fundamental representation of U(N), while Ψ_{λ} and Ψ_{α} are the subsets of the root system Ψ defined, respectively, by

$$\Psi_{\lambda} = \left\{ \alpha \in \Psi \, \middle| \, \lambda \cdot \alpha = 1 \right\} \,, \tag{3.14}$$

for any $\lambda \in \mathcal{W}$, and by

$$\Psi_{\alpha} = \left\{ \beta \in \Psi \, \middle| \, \alpha \cdot \beta = 1 \right\} \,, \tag{3.15}$$

for any $\alpha \in \Psi$. Notice that

$$C^{0}_{n;m_{1}\cdots m_{\ell}} = C_{n;m_{1}\cdots m_{\ell}} \tag{3.16}$$

where $C_{n;m_1\cdots m_\ell}$ are the lattice sums introduced in [11, 12, 13]. Furthermore, we have

$$C_{0;0\cdots 0}^{\ell} = \sum_{u=1}^{N} a_{u}^{\ell} \equiv C^{\ell} . \qquad (3.17)$$

Using this notation, the first few coefficients in the mass expansion of the U(N) prepotential were shown to be given by [11]²

$$f_{1} = \frac{1}{4} \sum_{\alpha \in \Psi} \log \left(\frac{\alpha \cdot a}{\Lambda}\right)^{2} ,$$

$$f_{2} = -\frac{1}{24} E_{2} C_{2}^{0} ,$$

$$f_{3} = -\frac{1}{720} \left(5E_{2}^{2} + E_{4}\right) C_{4}^{0} - \frac{1}{576} \left(E_{2}^{2} - E_{4}\right) C_{2;11}^{0} ,$$
(3.18)

²We warn the reader that, for later convenience, we have changed notation with respect to [11] and have explicitly factored out the mass-dependence. So, $f_n^{\text{there}} = f_n^{\text{here}} m^{2n}$.

where E_{2k} are the Eisenstein series (see Appendix A). These formulas encode the *exact* dependence on the coupling constant τ . Indeed, by expanding the Eisenstein series in powers of q, one can recover the perturbative contributions, corresponding to the terms proportional to q^0 , and the k-instanton contributions proportional to q^k . Analogous expressions can be obtained for the higher order mass terms in the U(N) theory and for other gauge algebras as well [11, 12, 13].

As discussed in great detail in [24, 23, 25] the prepotential coefficients f_n satisfy the recursion relation

$$\frac{\partial f_n}{\partial E_2} = -\frac{1}{24} \sum_{m=1}^{n-1} \frac{\partial f_m}{\partial a} \cdot \frac{\partial f_{n-m}}{\partial a} , \qquad (3.19)$$

which in turn implies that the quantum prepotential f obeys the non-linear differential equation

$$\frac{\partial f}{\partial E_2} + \frac{1}{24} \left(\frac{\partial f}{\partial a}\right)^2 = 0 . \qquad (3.20)$$

This equation, which is a direct consequence of the S-duality action (3.10) on the prepotential, is referred to as the modular anomaly equation since E_2 has an anomalous modular behavior

$$E_2\left(-\frac{1}{\tau}\right) = \tau^2 \left(E_2(\tau) + \frac{6}{\mathrm{i}\pi\tau}\right). \tag{3.21}$$

3.3 Seiberg-Witten Curves

In this section we review and compare two distinct algebraic approaches to describe the low-energy effective quantum dynamics of the $\mathcal{N} = 2^* U(N)$ SYM theory. The first approach is due to Donagi and Witten [14] (see also [15]), while the second approach is due to D'Hoker and Phong [17]. Even though some of the following considerations already appeared in the literature [81, 82], we are going to revisit the comparison between the two curves with the purpose of introducing the essential ingredients for the non-perturbative analysis presented in later sections.

3.3.1 The Donagi-Witten Curve

In this first approach, the algebraic curve of the $\mathcal{N} = 2^* \, \mathrm{U}(N)$ theory is given as an N-fold cover of an elliptic genus-one curve. The latter takes the standard Weierstraß form

$$y^{2} = (x - e_{1})(x - e_{2})(x - e_{3})$$
, (3.22)

where the e_i sum to zero and their differences are given in terms of the Jacobi θ -constants [2] as ³

$$e_2 - e_3 = \frac{1}{4} \theta_2(\tau)^4$$
, $e_2 - e_1 = \frac{1}{4} \theta_3(\tau)^4$, $e_3 - e_1 = \frac{1}{4} \theta_4(\tau)^4$. (3.23)

Here τ is the complex structure parameter of the elliptic curve which is identified with the gauge coupling (3.4) and the θ -constants have the following Fourier expansions

$$\theta_2(\tau) = \sum_{n \in \mathbb{Z}} q^{\frac{1}{2}\left(n - \frac{1}{2}\right)^2} , \qquad \theta_3(\tau) = \sum_{n \in \mathbb{Z}} q^{\frac{1}{2}n^2} , \qquad \theta_4(\tau) = \sum_{n \in \mathbb{Z}} (-1)^n q^{\frac{1}{2}n^2} , \quad (3.24)$$

where q is as in (3.7). Using the relations between the θ -constants and the Eisenstein series (see A.15), the elliptic curve (3.22) can be rewritten as

$$y^2 = x^3 - \frac{E_4}{48}x + \frac{E_6}{864} . ag{3.25}$$

Since E_4 and E_6 are modular forms of weight 4 and 6, for consistency x and y must have modular weight 2 and 3 respectively. If we recall the uniformizing solution in

³We use a different notation and normalization as compared to [2]. In particular our normalizations are such that the α -period of the uniformizing coordinate of the torus is $\omega_1 = 2\pi i$.

terms of the Weierstraß function $\wp(z)$, which obeys

$$\wp'(z)^2 = 4\,\wp^3(z) - \frac{4\pi^4 \,E_4}{3}\,\wp(z) - \frac{8\pi^6 \,E_6}{27} \tag{3.26}$$

when $z \sim z + 1$ and $z \sim z + \tau$, then by comparing with (3.25) we straightforwardly obtain the following identifications:

$$2y = \frac{\wp'(z)}{(2\pi i)^3} , \qquad x = \frac{\wp(z)}{(2\pi i)^2} . \qquad (3.27)$$

In this framework, the curve of the $\mathcal{N} = 2^* U(N)$ theory is described by the equation

$$F(t, x, y) = 0$$
 (3.28)

where F(t, x, y) is a polynomial of degree N. Modular covariance is extended to this equation by assigning modular weight 1 to the variable t. Certain technical conditions described in detail in [14, 15] allow one to fix the form of F to be

$$F(t, x, y) = \sum_{n=0}^{N} (-1)^n A_n P_{N-n}(t, x, y), \qquad (3.29)$$

where $A_0 = 1$ and the remaining N quantities A_n parametrize the Coulomb branch of the moduli space. The polynomials $P_n(t, x, y)$ are of degree n and are *almost* completely determined by the recursion relations [14]

$$\frac{\mathrm{d}P_n}{\mathrm{d}t} = n \, P_{n-1} \,, \tag{3.30}$$

combined with physical requirements related to the behaviour of F in the limits $x, y \to \infty$.

At the first two levels, n = 0 and n = 1, in view of the weights assigned to x and y,

the polynomials are uniquely fixed to be

$$P_0 = 1$$
, $P_1 = t$. (3.31)

At the next order, n = 2, the solution to the recursion equation (3.30) is

$$P_2 = t^2 + c \, m^2 \tag{3.32}$$

where the second term is an integration constant depending on the hypermultiplet mass that is allowed since P_2 has mass dimension 2. In addition, since P_2 has modular weight 2, the coefficient c must be an elliptic or modular function of weight 2. There is a unique such function, namely x, and thus P_2 must be of the form

$$P_2 = t^2 + \alpha \, x \, m^2 \tag{3.33}$$

where α is a numerical coefficient which is fixed by requiring a specific behavior at infinity [14].

If we choose coordinates such that u = 0 parametrizes the point at infinity, then taking into account that x is an elliptic function of weight 2, we can write

$$x = \frac{1}{u^2} . (3.34)$$

In terms of this variable, the required behavior at infinity is that under the shift

$$t \to t + \frac{m}{u} , \qquad (3.35)$$

the function F, and therefore all polynomials P_n , must have at most a *simple* pole

in u, namely for $u \to 0$ they must behave as

$$P_n\left(t+\frac{m}{u}\right) \sim \frac{\alpha_n}{u} + \text{regular}$$
 (3.36)

This follows from the requirement in [14] that the adjoint scalar field Φ has the following behaviour near the point u = 0 on the torus:

$$\Phi = \frac{m}{u} \operatorname{diag}(1, 1, \dots, -(N-1)) + \operatorname{regular terms}.$$

The residue *m* is identified with the mass of the adjoint hypermultiplet. The function F(t, x, y), which defines the *N*-fold spectral cover of the torus, is identified with the equation $\det(t\mathbf{1} - \Phi) = 0$. The shift in *t* above ensures that N - 1 of the eigenvalues of Φ have no pole as $u \to 0$ and this is what constrains the growth of the polynomials P_n near infinity (see [14] for more details).

The requirement that all higher order poles in u cancel constrains the integration constants that are allowed to appear. For example, imposing this behavior, one can easily fix the constant α in (3.33) and find that final form of P_2 is

$$P_2 = t^2 - x \, m^2 \,. \tag{3.37}$$

To fix the higher order polynomials, it is necessary to know the behaviour of y near u = 0. Using the algebraic equation (3.25), we easily find

$$y = \frac{1}{u^3}\sqrt{1 - \frac{E_4}{48}u^4 + \frac{E_6}{84}u^6} = \frac{1}{u^3} - \frac{E_4}{96}u - \frac{E_6}{1728}u^3 + \cdots$$
(3.38)

Using this and (3.34), we can completely determine the polynomial P_3 and get

$$P_3 = t^3 - 3t x m^2 + 2y m^3 . (3.39)$$

However, at the next level, we find that

$$P_4 = t^4 - 6t^2 x m^2 + 8t y m^3 - (3x^2 - \alpha E_4) m^4$$
(3.40)

satisfies all requirements for any value of α . In [14, 15] the simplest choice $\alpha = 0$ was made, but we will find that it is actually essential to keep the α -dependence and fix it to a different value.

This procedure can be iterated without any difficulty and in Appendix B we list a few of the higher degree polynomials P_n that we find in this way. They differ from the ones listed in [14, 15] by elliptic and modular functions. At first glance, these might seem trivial modifications since, for example in (3.40), the difference is proportional to E_4 , which is a modular form of weight 4. However, for $\alpha \neq 0$, this new term feeds into the iterative procedure to calculate the higher P_n , which in turn depend on these coefficients. These modified higher degree polynomials will play a crucial role in the following.

Using the explicit form of the polynomials P_n given in Appendix B and collecting the powers of t, we find that the curve equation (3.3.1) is

$$F(t, x, y) = t^{N} - A_{1} t^{N-1} + t^{N-2} \left[A_{2} - \binom{N}{2} m^{2} x \right] - t^{N-3} \left[A_{3} - \binom{N-1}{2} m^{2} A_{1} x - \binom{N}{3} 2 m^{3} y \right] + t^{N-4} \left[A_{4} - \binom{N-2}{2} m^{2} A_{2} x - \binom{N-1}{3} 2 m^{3} A_{1} y - \binom{N}{4} m^{4} (3 x^{2} - \alpha E_{4}) \right] + O(t^{N-5}) = 0.$$

$$(3.41)$$

Since F is a linear combination of the P_n , which are modular with weight n, it will transform homogeneously (with weight N) if the coefficients A_n are modular with weight n. To verify this fact and provide a precise identification between the A_n and the the gauge invariant quantum observables $\langle \operatorname{Tr} \Phi^n \rangle$ which naturally parametrize the moduli space, we find that the modifications that we have made to the P_n as compared to those of [14, 15] are essential.

3.3.2 The D'Hoker-Phong Curve

The second form of the curve for the $\mathcal{N} = 2^* \operatorname{U}(N)$ theory is due to D'Hoker and Phong and was originally derived by using the relation between the SW curve and the spectral curve of the elliptic Calogero-Moser system [17]. This spectral curve is abstractly defined as

$$R(t,z) \equiv \det\left[t\,1\!\!1 - L(z)\right] = 0\,, \qquad (3.42)$$

where L(z) is the Lax matrix of the integrable system. We refer the reader to [17] for details and here we merely present the curve in the form that is most convenient for our purposes.

First, we define the degree N polynomial H(t):

$$H(t) = \prod_{u=1}^{N} (t - e_u) = \sum_{n=0}^{N} (-1)^n W_n t^{N-n}$$
(3.43)

where

$$W_n = \sum_{u_1 < \dots < u_n} e_{u_1} \cdots e_{u_n} . \tag{3.44}$$

The e_u are interpreted as the quantum-corrected vacuum expectation values of the scalar field Φ and, at weak coupling, they have the following form

$$e_u = a_u + O(q) \tag{3.45}$$

in terms of the classical vacuum expectation values a_u (see (3.2)). Thus, the gauge invariant quantum expectation values, which parametrize the quantum moduli space, can be written as

$$\langle \operatorname{Tr} \Phi^n \rangle = \sum_{u=1}^N e_u^n .$$
 (3.46)

Next, we define the function

$$f(t,z) = \sum_{n=0}^{N} (-1)^n \frac{m^n}{n!} h_n(z) H^{(n)}(t)$$
(3.47)

where

$$H^{(n)}(t) \equiv \frac{d^n H(t)}{dt^n} = \sum_{\ell=0}^{N-n} (-1)^\ell \frac{(N-\ell)!}{(N-\ell-n)!} W_\ell t^{N-n-\ell} , \qquad (3.48)$$

and

$$h_n(z) \equiv \frac{1}{\theta_1(z|\tau)} \left(\frac{1}{2\pi i} \frac{d}{dz}\right)^n \theta_1(z|\tau)$$
(3.49)

with $\theta_1(z|\tau)$ being the first Jacobi θ -function

$$\theta_1(z|\tau) = \sum_{n \in \mathbb{Z}} e^{i\pi\tau(n-\frac{1}{2})^2 + 2\pi i(z-\frac{1}{2})(n-\frac{1}{2})} .$$
(3.50)

Notice we have chosen normalizations so that the uniformizing coordinate z on the torus obeys $z \sim z + 1$ and $z \sim z + \tau$, and that, as before, the complex structure parameter τ is identified with the gauge coupling (3.4).

Using this notation, the spectral curve of the Calogero-Moser system (3.42), and hence the SW curve for the U(N) theory, takes the form [17]

$$R(t,z) = f(t+mh_1,z) = 0.$$
(3.51)

To make the modular properties of the curve more manifest, we rewrite the function f(t, z) in (3.47) in a slightly different way. We first observe that

$$h_n(z) = \left(\frac{1}{2\pi i}\frac{d}{dz} + h_1(z)\right)^n 1 , \qquad (3.52)$$

as one can easily check recursively. Plugging this into the definition (3.47) of f and

using (3.48) (and after a simple rearrangement of the sums), we get

$$f(t,z) = \sum_{n=0}^{N} \sum_{\ell=0}^{N-n} (-1)^{\ell+n} {\binom{N-\ell}{n}} W_{\ell} t^{N-\ell-n} m^n \left(\frac{1}{2\pi i} \frac{d}{dz} + h_1(z)\right)^n 1$$

$$= \sum_{\ell=0}^{N} (-1)^{\ell} W_{\ell} \left[t - m \left(\frac{1}{2\pi i} \frac{d}{dz} + h_1(z)\right) \right]^{N-\ell} 1.$$
 (3.53)

From this we see that the shift in t in (3.51) simply amounts to setting $h_1=0$ after taking the derivatives. Thus, the curve equation for the $\mathcal{N} = 2^* U(N)$ theory in this formulation becomes

$$R(t,z) = \sum_{\ell=0}^{N} (-1)^{\ell} W_{\ell} \left[t - m \left(\frac{1}{2\pi i} \frac{d}{dz} + h_{1}(z) \right) \right]^{N-\ell} 1 \Big|_{h_{1}=0}$$

$$= t^{N} - t^{N-1} W_{1} + t^{N-2} \left[W_{2} + \binom{N}{2} m^{2} h_{1}' \right]$$

$$- t^{N-3} \left[W_{3} + \binom{N-1}{2} m^{2} h_{1}' W_{1} + \binom{N}{3} m^{3} h_{1}'' \right]$$

$$+ t^{N-4} \left[W_{4} + \binom{N-2}{2} m^{2} h_{1}' W_{2} + \binom{N-1}{3} m^{3} h_{1}'' W_{1} + \binom{N}{4} m^{4} \left(h_{1}''' + 3(h_{1}')^{2} \right) \right] + O(t^{N-5}) = 0$$
(3.54)

where the ' stands for the derivative with respect to $2\pi i z$.

3.3.3 Comparing Curves

By comparing the two forms of the SW curve presented in the previous subsections, one can establish a relation between the W_n , which are related to the quantum expectation values $\langle \operatorname{Tr} \Phi^n \rangle$, and the modular covariant combinations A_n on which Sduality acts in a simple way. A different method to relate the A_n and the W_n , which only involves the D'Hoker-Phong form of the curve, is presented in Appendix C. Equating the coefficients of the same power of t in (3.41) and (3.54), we easily get

$$A_{1} = W_{1} ,$$

$$A_{2} = W_{2} + {\binom{N}{2}} m^{2} (h_{1}' + x) ,$$

$$A_{3} = W_{3} + {\binom{N-1}{2}} m^{2} (h_{1}' + x) W_{1} + {\binom{N}{3}} m^{3} (h_{1}'' + 2y) ,$$

$$A_{4} = W_{4} + {\binom{N-2}{2}} m^{2} (h_{1}' + x) W_{2} + {\binom{N-1}{3}} m^{3} (h_{1}'' + 2y) + {\binom{N}{4}} m^{4} (h_{1}''' + 3(h_{1}')^{2} + 6h_{1}' x + 9x^{2} - \alpha E_{4})$$
(3.55)

and so on. Recalling that x and y are related to the Weierstraß function as shown in (3.27), and using the properties of $\theta_1(z|\tau)$ and its derivatives, one can show that all z-dependence cancels in the right hand side of (3.55) as it should, since

$$h'_{1} + x = \frac{E_{2}}{12} ,$$

$$h''_{1} + 2y = 0 , \qquad (3.56)$$

$$h'''_{1} + 3(h'_{1})^{2} + 6h'_{1}x + 9x^{2} = \frac{E_{2}^{2}}{48} + \frac{E_{4}}{24} .$$

We have included proofs of these identities in Appendix A. Using these results, the relations (3.55) simplify and reduce to

$$A_{1} = W_{1} ,$$

$$A_{2} = W_{2} + {\binom{N}{2}} \frac{m^{2} E_{2}}{12} ,$$

$$A_{3} = W_{3} + {\binom{N-1}{2}} \frac{m^{2} E_{2}}{12} W_{1} ,$$

$$A_{4} = W_{4} + {\binom{N-2}{2}} \frac{m^{2} E_{2}}{12} W_{2} + {\binom{N}{4}} \left(\frac{m^{4} E_{2}^{2}}{48} + \frac{m^{4} E_{4}(1-24\alpha)}{24}\right) .$$
(3.57)

Notice that all terms proportional to m^3 cancel and that the formula for A_4 can be further simplified by setting the free parameter to $\alpha = \frac{1}{24}$. With this choice we eliminate the modular form E_4 , leaving only the quasi-modular form E_2 .
The same procedure may be carried out for the higher coefficients A_n without any difficulty. Exploiting the freedom of fixing the parameters in front of the modular forms to systematically eliminate them, we obtain the following rather compact result:

$$A_n = \sum_{\ell=0}^{\lfloor n/2 \rfloor} \binom{N-n+2\ell}{2\ell} (2\ell-1)!! \left(\frac{m^2 E_2}{12}\right)^\ell W_{n-2\ell} .$$
(3.58)

This formula can be easily inverted and one gets

$$W_n = \sum_{\ell=0}^{[n/2]} (-1)^{\ell} \binom{N-n+2\ell}{2\ell} (2\ell-1)!! \left(\frac{m^2 E_2}{12}\right)^{\ell} A_{n-2\ell} .$$
(3.59)

We have verified these relations by working to higher orders in both n and N. It is interesting to observe that, although both the Donagi-Witten curve and the D'Hoker-Phong curve separately have coefficients that are elliptic functions, the maps between the two sets of coefficients can be written entirely in terms of quasimodular forms. For this to happen and, more importantly, in order that all dependence on the uniformizing coordinate z disappears in the relations between the A_n and the W_n , it is essential to use a set of polynomials P_n that are differ from those originally defined in [14, 15].

Both W_n and A_n are good sets of coordinates for the Coulomb moduli space of the $\mathcal{N} = 2^* \operatorname{U}(N)$ SYM theory. The former naturally incorporate the quantum corrections that are calculable using either the curve analysis or by localization calculations while the latter are distinguished by their simple behavior under Sduality. In the following sections, we will independently calculate the A_n and the W_n in a weak-coupling expansion and show that they satisfy the general relations (3.58) and (3.59) provided some important caveats are taken into account.

3.4 Period Integrals and Modular Anomalies

In this section, we present two methods to compute the modular covariant quantities A_n and express them in terms of the classical vacuum expectation values a_u of the adjoint scalar field Φ given in (3.2). The first method is based on a direct use of the curve and the associated differential, while the second exploits an extension of the modular anomaly equation (3.20).

3.4.1 Period Integrals

By solving the Donagi-Witten curve equation (3.28) one can express the variable t as a function of x and y, and hence of the uniformizing coordinate of the torus z through the identifications (3.27). Once this is done, the SW differential is given by [17]:

$$\lambda_{\rm SW} = t(z) \, dz \,\,, \tag{3.60}$$

and its periods are identified with the pairs of dual variables a_u and a_u^D according to (3.9). Of course, in order to obtain explicit expressions, a canonical basis of 1-cycles is needed. Since the curve is an N-fold cover of a torus, there is a natural choice for such a basis, as we now demonstrate. In fact, F being a polynomial of degree N, we can factorize it as

$$F = \prod_{u=1}^{N} \left(t - t_u(x(z), y(z)) \right) = 0 , \qquad (3.61)$$

and then define

$$a_{u} = \oint_{A_{u}} \lambda_{SW} := \oint_{\alpha} t_{u}(x(z), y(z)) dz ,$$

$$a_{u}^{D} = \oint_{B_{u}} \lambda_{SW} := \oint_{\beta} t_{u}(x(z), y(z)) dz ,$$
(3.62)

where α and β are, respectively, the A and B cycles of the torus. To see that this identification is correct, let us (for a moment) consider switching off the mass of the adjoint hypermultiplet. If we do so, the supersymmetry is enhanced to $\mathcal{N} = 4$ and Donagi-Witten polynomials simply become $P_n = t^n$, so that the curve takes the form

$$F = \sum_{n=0}^{N} (-1)^n t^{N-n} A_n = 0 . \qquad (3.63)$$

Since in the $\mathcal{N} = 4$ SYM theory the classical moduli space does not receive quantum corrections, it makes sense to identify the modular covariant coordinates A_n with the symmetric polynomials constructed from the classical vacuum expectation values, namely

$$A_n = \sum_{u_1 < \dots < u_n} a_{u_1} \cdots a_{u_n} \ . \tag{3.64}$$

Substituting this into (3.63), we see that F factorizes as

$$F = \prod_{u=1}^{N} (t - a_u) = 0, \qquad (3.65)$$

so we may conclude that in the massless limit we have $t_u = a_u$. This is clearly consistent with our ansatz (3.62), since the integral over the α -cycle gives unity. The integral over the β -cycle, instead, gives

$$a_u^{\rm D} = \oint_\beta a_u = \tau \, a_u \,\,, \tag{3.66}$$

which is the expected answer in the $\mathcal{N} = 4$ gauge theory.

Let us now revert to our original problem, and consider the scenario where the adjoint hypermultiplet has a mass m. In general, it is not possible to compute the period integrals (3.62) explicitly, as each of the $t_u(x, y)$ is a solution of a generic polynomial equation of degree N. However, progress can be made by assuming that each of these solutions has a expansion in powers of the hypermultiplet mass, of the

form

$$t_u(x,y) = a_u + \sum_{\ell \in \mathbb{N}/2} t_u^{(\ell)}(x,y) \, m^{2\ell} \,, \qquad (3.67)$$

and by working perturbatively order by order in m. Notice that in (3.67) the sum is over both integers and half-integers in order to have in principle both even and odd powers of m, even though in the end only the even ones will survive. Of course, this assumption implies that the modular covariant coordinates on moduli space have a mass expansion of the form

$$A_n = \sum_{u_1 < \dots < u_n} a_{u_1} \cdots a_{u_n} + \sum_{\ell \in \mathbb{N}/2} A_n^{(\ell)} m^{2\ell} .$$
 (3.68)

Using this ansatz in the curve equation (3.41) leads to constraints on the $t_u^{(\ell)}$, which we solve in terms of the $A_n^{(\ell)}$. Finally, we substitute these into the expressions for the A-periods in (3.62) and demand that all higher order terms in m vanish for self-consistency as that equation is already solved by $t_u^{(0)}$. The integrals for these higher order terms typically involve integrals of powers of the Weierstraß function and its derivative, which are known in terms of quasi-modular forms. In this way we can construct the various mass corrections $A_n^{(\ell)}$ in terms of the classical a_u and of quasi-modular forms.

Let us first illustrate this procedure in the simple case of the U(2) gauge theory. For N = 2 the Donagi-Witten curve is

$$t^{2} - tA_{1} + (A_{2} - m^{2}x) = 0. ag{3.69}$$

Inserting the mass expansions (3.67) and (3.68) and collecting the powers of m, we

obtain

$$a_u^2 - a_u(a_1 + a_2) + a_1 a_2 + m \left(A_2^{(1/2)} + (2 a_u - a_1 - a_2) t_u^{(1/2)} - a_u A_1^{(1/2)} \right)$$

$$+ m^2 \left(A_2^{(1)} + \left(t_u^{(1/2)} \right)^2 + (2 a_u - a_1 - a_2) t_u^{(1)} - t_u^{(1/2)} A_1^{(1/2)} - a_u A_1^{(1)} - x \right) + O(m^3) = 0$$
(3.70)

for u = 1, 2. It is easy to check that the zeroth order term in the mass vanishes, as it should. Requiring the cancellation of the term at linear order in m amounts to setting

$$t_u^{(1/2)} = \frac{A_2^{(1/2)} - a_u A_1^{(1/2)}}{a_1 + a_2 - 2a_u}$$
(3.71)

for u = 1, 2. Now, in order to maintain the relation (3.62), the integral of $t_u^{(\ell)}$ over the A-cycles has to vanish for all ℓ . In particular, for $\ell = 1/2$ and taking into account that $t_u^{(1/2)}$ in (3.71) is constant with respect to z, one has

$$\oint_{\alpha} t_u^{(1/2)} dz = \frac{A_2^{(1/2)} - a_u A_1^{(1/2)}}{a_1 + a_2 - 2a_u} = 0$$
(3.72)

for both u = 1 and u = 2. In turn this leads to

$$A_1^{(1/2)} = A_2^{(1/2)} = 0 . (3.73)$$

Substituting this into (3.70) and demanding the cancellation of the m^2 terms, we get

$$t_u^{(1)} = \frac{A_2^{(1)} - a_u A_1^{(1)} - x}{a_1 + a_2 - 2a_u} .$$
(3.74)

Imposing that

$$\oint_{\alpha} t_u^{(1)} dz = 0 \tag{3.75}$$

for u = 1, 2, and using the fact that, in view of the identification (3.27),

$$\oint_{\alpha} x \, dz = \frac{1}{(2\pi i)^2} \oint_{\alpha} \wp(z) \, dz = \frac{E_2}{12} \,, \qquad (3.76)$$

we get

$$A_1^{(1)} = 0 , \qquad A_2^{(1)} = \frac{E_2}{12} .$$
 (3.77)

Recapitulating, we have obtained

$$A_1 = a_1 + a_2 ,$$

$$A_2 = a_1 a_2 + \frac{m^2}{12} E_2 + O(m^3) .$$
(3.78)

This process can be repeated in similar fashion to obtain all mass corrections in a systematic way. This procedure requires that we compute period integrals of polynomials in the Weirstraß function and its derivative which can be done using standard techniques (see for example [75] and references therein). We stress that although this approach is perturbative in m, it is *exact* in the gauge coupling constant, since the coefficients are fully resummed quasi-modular forms in τ .

The same procedure can of course be carried out for $\mathcal{N} = 2^*$ theories with higher rank gauge groups, even if the calculations quickly become more involved as Nincreases. The results, however, can be organized in a rather compact way by using the lattice sums $C_{n;m_1\cdots}^p$ defined in (3.13). In fact, the expressions we find for the first few A_n at the first few non-trivial orders in m in the $\mathrm{U}(N)$ theory are

$$A_1 = \sum_u a_u aga{3.79}$$

$$A_{2} = \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} + {\binom{N}{2}} \frac{m^{2}}{12} E_{2} + \frac{m^{4}}{288} \left(E_{2}^{2} - E_{4}\right) C_{2}^{0} + \frac{m^{6}}{4320} \left(5E_{2}^{3} - 3E_{2}E_{4} - 2E_{6}\right) C_{4}^{0} + \frac{m^{6}}{3456} \left(E_{2}^{3} - 3E_{2}E_{4} + 2E_{6}\right) C_{2;11}^{0} + O(m^{8}) , \qquad (3.80)$$

$$\begin{split} A_{3} &= \sum_{u_{1} < u_{2} < u_{3}} a_{u_{1}} a_{u_{2}} a_{u_{3}} + \binom{N-1}{2} \frac{m^{2}}{12} E_{2} \sum_{u} a_{u} + \frac{m^{4}}{288} \left(E_{2}^{2} - E_{4}\right) \left(C_{2}^{0} \sum_{u} a_{u} - 2C_{2}^{1}\right) \\ &+ \frac{m^{6}}{4320} \left(5E_{2}^{3} - 3E_{2}E_{4} - 2E_{6}\right) \left(C_{4}^{0} \sum_{u} a_{u} - 2C_{4}^{1}\right) \\ &+ \frac{m^{6}}{3456} \left(E_{2}^{3} - 3E_{2}E_{4} + 2E_{6}\right) \left(C_{2;11}^{0} \sum_{u} a_{u} - 2C_{2;11}^{1}\right) + O(m^{8}) , \quad (3.81) \end{split}$$

$$\begin{aligned} A_{4} &= \sum_{u_{1} < \cdots < u_{4}} a_{u_{1}} a_{u_{2}} a_{u_{3}} a_{u_{4}} + \binom{N-2}{2} \frac{m^{2}}{12} E_{2} \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} + \binom{N}{4} \frac{m^{4}}{48} E_{2}^{2} \\ &+ \frac{m^{4}}{288} \left(E_{2}^{2} - E_{4}\right) \left(C_{2}^{0} \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} - 2C_{2}^{1} \sum_{u} a_{u} + 3C_{2}^{2} - \binom{N}{2}\right) \\ &+ \frac{m^{6}}{4320} \left(5E_{2}^{3} - 3E_{2}E_{4} - 2E_{6}\right) \left(C_{4}^{0} \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} - 2C_{4}^{1} \sum_{u} a_{u} + 3C_{4}^{2} - \frac{1}{2}C_{2}^{0}\right) \\ &+ \frac{m^{6}}{3456} \left(E_{2}^{3} - 3E_{2}E_{4} + 2E_{6}\right) \left(C_{2;11}^{0} \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} - 2C_{1}^{1} \sum_{u} a_{u} + 3C_{4}^{2} - \frac{1}{2}C_{2}^{0}\right) \\ &+ \left(\frac{N-2}{2}\right) \frac{m^{6}}{3456} E_{2} \left(E_{2}^{2} - E_{4}\right) C_{2}^{0} + O(m^{8}) . \end{aligned}$$

Of course, only the A_n with $n \leq N$ are the independent coordinates that can be used to parametrize the moduli space of the theory. Despite their appearance, it is not difficult to recognize a regular pattern in these expressions, which contain the same combinations of Eisenstein series appearing in the prepotential coefficients. Notice also that only even powers of m are present, this being in full agreement with the \mathbb{Z}_2 symmetry of the theory that sends $m \to -m$.

We have explicitly verified that under S-duality the above A_n transform with weight

n, namely

$$S(A_n) = \tau^n A_n . aga{3.83}$$

To do so we used the properties of the Eisenstein series under inversion, and replaced each a_u with the corresponding dual variable a_u^D , which can be computed either by evaluating the periods of the SW differential along the *B*-cycles according to (3.62) or, more efficiently, by taking the derivative of the prepotential with respect to a_u according to (4.16). The fact that (3.83) holds true despite the explicit presence of the quasi-modular Eisenstein series E_2 in the A_n is a highly non-trivial consistency check. Finally, we observe that by inserting (3.79)–(3.82) in the map (3.59), one can obtain the quantum expectation values W_n in terms of the classical variables a_u . The result is

$$W_{1} = \sum_{u} a_{u} , \qquad (3.84)$$

$$W_{2} = \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} + \frac{m^{4}}{288} \left(E_{2}^{2} - E_{4} \right) C_{2}^{0} + \frac{m^{6}}{4320} \left(5E_{2}^{3} - 3E_{2}E_{4} - 2E_{6} \right) C_{4}^{0} + \frac{m^{6}}{3456} \left(E_{2}^{3} - 3E_{2}E_{4} + 2E_{6} \right) C_{2;11}^{0} + O(m^{8}) , \qquad (3.85)$$

$$W_{3} = \sum_{u_{1} < u_{2} < u_{3}} a_{u_{1}} a_{u_{2}} a_{u_{3}} + \frac{m^{4}}{288} \left(E_{2}^{2} - E_{4} \right) \left(C_{2}^{0} \sum_{u} a_{u} - 2C_{2}^{1} \right) + \frac{m^{6}}{4320} \left(5E_{2}^{3} - 3E_{2}E_{4} - 2E_{6} \right) \left(C_{4}^{0} \sum_{u} a_{u} - 2C_{4}^{1} \right) + \frac{m^{6}}{3456} \left(E_{2}^{3} - 3E_{2}E_{4} + 2E_{6} \right) \left(C_{2;11}^{0} \sum_{u} a_{u} - 2C_{4}^{1} \right) + O(m^{8}) , \qquad (3.86)$$

$$W_{4} = \sum_{u_{1} < \dots < u_{4}} a_{u_{1}} a_{u_{2}} a_{u_{3}} a_{u_{4}} + \frac{m^{4}}{288} (E_{2}^{2} - E_{4}) \left(C_{2}^{0} \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} - 2 C_{2}^{1} \sum_{u} a_{u} + 3 C_{2}^{2} - \binom{N}{2} \right) + \frac{m^{6}}{4320} (5E_{2}^{3} - 3E_{2}E_{4} - 2E_{6}) \left(C_{4}^{0} \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} - 2 C_{4}^{1} \sum_{u} a_{u} + 3 C_{4}^{2} - \frac{1}{2} C_{2}^{0} \right) + \frac{m^{6}}{3456} (E_{2}^{3} - 3E_{2}E_{4} + 2E_{6}) \left(C_{2;11}^{0} \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} - 2 C_{2;11}^{1} \sum_{u} a_{u} + 3 C_{2;11}^{2} \right) + O(m^{8}) .$$

$$(3.87)$$

It is interesting to notice that these expressions are a bit simpler than the ones for the A_n ; in particular, all m^2 terms disappear and, up to a constant term in W_4 , all other explicit dependence on N drops out. These formulas will be useful in later sections, where we compare them with results from explicit localization calculations. An important consistency check on our results is the fact that both W_3 and W_4 vanish for U(2), and that W_4 vanishes for U(3). This has to happen since the W_n are symmetric polynomials in the quantum variables e_u , see (3.44).

3.4.2 A Modular Anomaly Equation

We now explore an alternative route to express the A_n in terms of the classical parameters a_u , which is based on the S-duality transformation properties. The main idea is simple: if we assume the mass expansion (3.68), then the requirement that A_n transforms with weight n under S-duality constrains the form of $A_n^{(\ell)}$ once the previous mass terms are known. So, starting from the classical part it is possible to systematically reconstruct in this way all subleading terms.

Let us recall from Section 3.2 that ⁴

$$S(a) = a^{\mathrm{D}} = \frac{1}{2\pi \mathrm{i}} \frac{\partial F}{\partial a} = \tau \left(a + \frac{\delta}{12} \frac{\partial f}{\partial a} \right)$$
(3.88)

where f is the quantum part of the prepotential and $\delta = \frac{6}{i\pi\tau}$. Furthermore, in order for the A_n to have the correct mass dimension, the subleading terms $A_n^{(\ell)}$ must be homogeneous functions of a with weight $n - 2\ell$:

$$A_n^{(\ell)}(\tau, \lambda a) = \lambda^{n-2\ell} A_n^{(\ell)}(\tau, a) .$$
(3.89)

The other basic requirement is that they are quasi modular forms of weight 2ℓ . This

⁴For simplicity we suppress the subscripts and denote the pair $(a_u, a_u^{\rm D})$ as $(a, a^{\rm D})$.

implies that the $A_n^{(\ell)}$ depend on the coupling constant τ only through the Eisenstein series E_2 , E_4 and E_6 , namely

$$A_n^{(\ell)}(\tau, a) = A_n^{(\ell)} \left(E_2(\tau), E_4(\tau), E_6(\tau), a \right) , \qquad (3.90)$$

so that

$$A_n^{(\ell)}\left(-\frac{1}{\tau},a\right) = A_n^{(\ell)}\left(E_2\left(-\frac{1}{\tau}\right), E_4\left(-\frac{1}{\tau}\right), E_6\left(-\frac{1}{\tau}\right),a\right)$$

= $\tau^{2\ell} A_n^{(\ell)}\left(E_2 + \delta, E_4, E_6, a\right),$ (3.91)

where in the last step we have used the anomalous modular transformation (3.21) of the second Eisenstein series E_2 . From now on, for ease of notation, we only exhibit the dependence on E_2 . Putting everything together, we find

$$S(A_n^{(\ell)}) = A_n^{(\ell)} \left(E_2\left(-\frac{1}{\tau}\right), a^{\mathrm{D}} \right) = \tau^n A_n^{(\ell)} \left(E_2 + \delta, a + \frac{\delta}{12} \frac{\partial f}{\partial a} \right)$$
$$= \tau^n \left[A_n^{(\ell)} + \left(\frac{\partial A_n^{(\ell)}}{\partial E_2} + \frac{1}{12} \frac{\partial A_n^{(\ell)}}{\partial a} \cdot \frac{\partial f}{\partial a} \right) \delta + O(\delta^2) \right] .$$
(3.92)

The requirement that under S-duality A_n be a modular form of weight n leads to a modular anomaly equation:

$$\frac{\partial A_n}{\partial E_2} + \frac{1}{12} \frac{\partial A_n}{\partial a} \cdot \frac{\partial f}{\partial a} = 0 .$$
(3.93)

Notice that if (3.93) is satisfied, then all terms in (3.92) which are of higher order in δ , vanish. Expanding both the A_n and the quantum prepotential f in powers of m, we can rewrite the above modular anomaly equation in the form of a recursion relation for the $A_n^{(\ell)}$, namely

$$\frac{\partial A_n^{(\ell)}}{\partial E_2} + \frac{1}{12} \sum_{k=0}^{\ell} \frac{\partial A_n^{(k)}}{\partial a} \cdot \frac{\partial f_{\ell-k}}{\partial a} = 0 .$$
(3.94)

This shows that starting from the classical symmetric polynomials

$$A_n^{(0)} = \sum_{u_1 < \dots < u_n} a_{u_1} \cdots a_{u_n}$$
(3.95)

and the prepotential coefficients (some of which have been listed in (3.18)), one can systematically calculate the higher order terms and obtain the modular completion iteratively by integrating the modular anomaly equation (3.94). For example, at the first step ($\ell = 1$) we have

$$\frac{\partial A_n^{(1)}}{\partial E_2} = -\frac{1}{12} \frac{\partial A_n^{(0)}}{\partial a} \cdot \frac{\partial f_1}{\partial a} = -\frac{1}{12} \sum_{u \neq v} \frac{\partial A_n^{(0)}}{\partial a_u} \frac{1}{a_u - a_v} , \qquad (3.96)$$

which is solved by

$$A_n^{(1)} = \binom{N-n+2}{2} \frac{E_2}{12} A_{n-2}^{(0)} .$$
(3.97)

The higher order corrections $A_n^{(\ell)}$ can be similarly derived up to terms that are purely composed of modular forms of weight 2ℓ . These cannot be determined from the recursion relation alone, which is a symmetry requirement, and some extra dynamical input is needed. To illustrate this point let us consider the explicit expressions of A_1 and A_2 for the U(N) theory that can be derived using the above procedure. Up to order m^8 we find

$$A_{1} = \sum_{u} a_{u} , \qquad (3.98)$$

$$A_{2} = \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} + {\binom{N}{2}} \frac{m^{2}}{12} E_{2} + \frac{m^{4}}{288} (E_{2}^{2} - \alpha E_{4}) C_{2}^{0} + \frac{m^{6}}{4320} (5E_{2}^{3} + (2 - 5\alpha)E_{2}E_{4} - \beta E_{6}) C_{4}^{0} + \frac{m^{6}}{3456} (E_{2}^{3} - (2 + \alpha)E_{2}E_{4} + \gamma E_{6}) C_{2;11}^{0} + O(m^{8}) , \qquad (3.99)$$

where α, β, γ are free parameters. As anticipated, the terms that only depend on E_2 are completely fixed by the modular anomaly equation, while those involving

also the modular forms E_4 and E_6 depend on integration constants. One can fix them by requiring that the perturbative limit of the above expressions, in which all Eisenstein series effectively are set to 1, matches with the known perturbative behavior that can be deduced from the relations between the modular A_n and the quantum W_n discussed in Section 3.3.3. In particular, from (3.58) with n = 2 we see that

$$A_2\big|_{\rm cl} = W_2\big|_{\rm cl} + \binom{N}{2}\frac{m^2}{12} = \sum_{u_1 < u_2} a_{u_1}a_{u_2} + \binom{N}{2}\frac{m^2}{12} .$$
(3.100)

This perturbative behavior is matched by (3.99) only if

$$\alpha = 1 \quad \text{and} \quad \beta = \gamma = 2 \;. \tag{3.101}$$

It is reassuring to see that with this choice of parameters one precisely recovers the expression for A_2 in (3.80) that was obtained from the calculation of the period integrals. By extending this procedure to higher order we can also derive A_3 and A_4 and verify that they exactly agree with (3.81) and (3.82). This match is a very strong indication of the correctness of our calculations and the validity of the approach based on the modular anomaly equation (3.93).

Finally, we would like to remark that up to order m^{10} the matching with the perturbative results is enough to completely fix all integration constants, since there is a unique modular form of weight 2n up to n = 5. At n = 6, i.e. at order m^{12} there are two independent modular forms of weight 12, namely E_4^3 and E_6^2 . So the knowledge of the perturbative behavior is not enough to fix all parameters and more information, for example from the 1-instanton sector, is needed. At n = 7, again the perturbative information is sufficient since only one modular form of weight 14 exists. However from that point on, some extra data from the non-perturbative sectors is necessary. This is exactly the same situation occurring also for the prepotential coefficients, as pointed out for instance in [11, 12, 13]. The discussion of the previous section clearly shows that in order to confirm the general relations among the chiral observables and their modular properties, and also to have data to fix the coefficients left undetermined by the modular anomaly equation, it is necessary to explicitly compute some instanton contributions. This is possible using the equivariant localization techniques.

Following the discussion in [11], we first deform the $\mathcal{N} = 2^*$ theory by introducing the Ω -background [3, 4] and then calculate the partition function in a multi-instanton sector. The Ω -deformation parameters will be denoted ϵ_1 and ϵ_2 . The partition function Z_k for the U(N) theory in the presence of k-instantons is obtained by doing the following multi-dimensional contour integral:

$$Z_k = \oint \prod_{i=1}^k \frac{d\chi_i}{2\pi i} z_k^{\text{gauge}} z_k^{\text{matter}} , \qquad (3.102)$$

where the integrand is given by

$$z_k^{\text{gauge}} = \frac{(-1)^k}{k!} \left(\frac{\epsilon_1 + \epsilon_2}{\epsilon_1 \epsilon_2}\right)^k \frac{\Delta(0)\Delta(\epsilon_1 + \epsilon_2)}{\Delta(\epsilon_1)\Delta(\epsilon_2)} \prod_{i=1}^k \frac{1}{P(\chi_i + \frac{\epsilon_1 + \epsilon_2}{2})P(\chi_i - \frac{\epsilon_1 + \epsilon_2}{2})}$$
(3.103a)

$$z_k^{\text{matter}} = \left(\frac{(\epsilon_1 + \epsilon_3)(\epsilon_1 + \epsilon_4)}{\epsilon_3 \epsilon_4}\right)^k \frac{\Delta(\epsilon_1 + \epsilon_3)\Delta(\epsilon_1 + \epsilon_4)}{\Delta(\epsilon_3)\Delta(\epsilon_4)} \prod_{i=1}^k P(\chi_i + \frac{\epsilon_3 - \epsilon_4}{2})P(\chi_i - \frac{\epsilon_3 - \epsilon_4}{2})$$
(3.103b)

with

$$P(x) = \prod_{u=1}^{N} (x - a_u) \qquad \Delta(x) = \prod_{i < j}^{k} (x^2 - \chi_{ij}^2) , \qquad (3.104)$$

and $\chi_{ij} = \chi_i - \chi_j$. The parameters ϵ_3 and ϵ_4 are related the hypermultiplet mass m

according to

$$\epsilon_3 = m - \frac{\epsilon_1 + \epsilon_2}{2} , \qquad \epsilon_4 = -m - \frac{\epsilon_1 + \epsilon_2}{2} . \qquad (3.105)$$

The contour integrals are computed by closing the contours in the upper half planes of the χ_i variables, assigning imaginary parts to the ϵ 's, with the prescription [11]:

$$\operatorname{Im}(\epsilon_4) \gg \operatorname{Im}(\epsilon_3) \gg \operatorname{Im}(\epsilon_2) \gg \operatorname{Im}(\epsilon_1) > 0$$
. (3.106)

This prescription allows one to calculate the residues without ambiguity and obtain the partition function

$$Z_{\rm inst} = 1 + \sum_{k} q^k Z_k \,, \tag{3.107}$$

from which one can derive the instanton part of prepotential

$$F_{\text{inst}} = \lim_{\epsilon_1, \epsilon_2 \to 0} \left(-\epsilon_1 \epsilon_2 \log Z_{\text{inst}} \right) = \sum_{k=1}^{\infty} q^k F_k .$$
(3.108)

In this way one can compute the non-perturbative contributions to the coefficients f_n and verify the agreement with the resummed expressions like those given in (3.18) (for details we refer to [11, 80] and references therein).

The same localization methods can be used to compute the chiral correlators, which are known to receive quantum corrections from all instanton sectors. In this framework the expectation value for the generating function of such chiral observables is given by [51, 52, 20, 80]

$$\langle \operatorname{Tr} e^{z \Phi} \rangle \Big|_{\operatorname{loc}} = \sum_{n=0}^{\infty} \frac{z^n}{n!} \langle \operatorname{Tr} \Phi^n \rangle \Big|_{\operatorname{loc}}$$

$$= \sum_{u=1}^{N} e^{za_u} - \frac{1}{Z_{\operatorname{inst}}} \sum_{k=1}^{\infty} \frac{q^k}{k!} \oint \prod_{i=1}^{k} \frac{d\chi_i}{2\pi \mathrm{i}} \mathcal{O}(z,\chi_i) \, z_k^{\operatorname{gauge}} \, z_k^{\operatorname{matter}} \,,$$

$$(3.109)$$

where the operator insertion in the instanton partition function is explicitly given

by

$$\mathcal{O}(z,\chi_i) = \sum_{i=1}^k e^{z\chi_i} (1 - e^{z\epsilon_1})(1 - e^{z\epsilon_2}) , \qquad (3.110)$$

and the prescription to perform the contour integrals in (3.106) is the same as the one used for the instanton partition function. By explicitly computing these integrals order by order in k and then taking multiple derivatives with respect to z, one obtains the various instanton contributions to the chiral observables $\langle \operatorname{Tr} \Phi^n \rangle \big|_{\operatorname{loc}}$. Up to three instantons and for $n \leq 5$, we have explicitly verified that these instanton corrections can be compactly written using the lattice sums (3.13) as follows

$$\langle \operatorname{Tr} \Phi^{n} \rangle \big|_{\operatorname{loc}} = C^{n} - \binom{n}{2} 2 m^{2} (q + 3q^{2} + 4q^{3} + \cdots) C^{n-2} + \binom{n}{2} 2 m^{4} (q + 6q^{2} + 12q^{3} + \cdots) C_{2}^{n-2} + \binom{n}{4} 2 m^{4} (3q^{2} + 20q^{3} + \cdots) C^{n-4} - \binom{n}{2} 24 m^{6} (q^{2} + 8q^{3} + \cdots) C_{4}^{n-2} + \binom{n}{2} m^{6} (q + 12q^{2} + 36q^{3} + \cdots) C_{2;11}^{n-2} - \binom{n}{4} 24 m^{6} (q^{3} + \cdots) C_{2}^{n-4} + O(m^{8}) .$$

$$(3.111)$$

Recall that $C^n = \sum_u a_u^n$ and that one should set the Cs to zero when the superscript of the C's is negative. Based on our previous experience we expect that the coefficients of the various structures in (3.111) are just the first terms of the instanton expansion of (quasi)-modular forms built out of Eisenstein series. This is indeed what happens. In fact, we find

$$\langle \operatorname{Tr} \Phi^{n} \rangle \Big|_{\operatorname{loc}} = C^{n} + {\binom{n}{2}} \frac{m^{2}}{12} (E_{2} - 1) C^{n-2} - {\binom{n}{2}} \frac{m^{4}}{144} (E_{2}^{2} - E_{4}) C_{2}^{n-2} + {\binom{n}{4}} \frac{m^{4}}{720} (21 - 30E_{2} + 10E_{2}^{2} - E_{4}) C^{n-4} - {\binom{n}{2}} \frac{m^{6}}{2160} (5E_{2}^{3} - 3E_{2}E_{4} - 2E_{6}) C_{4}^{n-2} - {\binom{n}{2}} \frac{m^{6}}{1728} (E_{2}^{3} - 3E_{2}E_{4} + 2E_{6}) C_{2;11}^{n-2} + {\binom{n}{4}} \frac{m^{6}}{4320} (15E_{2}^{2} - 5E_{2}^{3} - 15E_{4} + 9E_{2}E_{4} - 4E_{6}) C_{2}^{n-4} + O(m^{8}) .$$

$$(3.112)$$

By expanding the Eisenstein series in powers of q we can obtain the contributions at any instanton number. We have verified the correctness of our extrapolation by computing the 4 and 5 instanton terms in the U(4) theory and the 4 instanton terms in the U(5) theory, finding perfect match with the "predictions" coming from the Fourier expansion of (3.112). We also note that using the Matone relation [83], the result for n = 2 matches perfectly with the mass expansion of the prepotential obtained in [11, 12]. Another noteworthy feature of the formula (3.112) is that the same quasi-modular functions appear for all values of n. Our results can therefore be thought of as a natural generalization of the result for the prepotential to other observables of the gauge theory.

To compare with our findings of the previous sections, it is convenient to change basis and make combinations of the above operators that describe the quantum version of the symmetric polynomials in the classical vacuum expectation values. At the first few levels the explicit map is

$$\begin{split} W_{1}^{\text{loc}} &= \langle \operatorname{Tr} \Phi \rangle \big|_{\text{loc}} , \\ W_{2}^{\text{loc}} &= \frac{1}{2} \Big(\langle \operatorname{Tr} \Phi \rangle \big|_{\text{loc}}^{2} - \langle \operatorname{Tr} \Phi^{2} \rangle \big|_{\text{loc}} \Big) , \\ W_{3}^{\text{loc}} &= \frac{1}{6} \Big(\langle \operatorname{Tr} \Phi \rangle \big|_{\text{loc}}^{3} - 3 \langle \operatorname{Tr} \Phi \rangle \big|_{\text{loc}} \langle \operatorname{Tr} \Phi^{2} \rangle \big|_{\text{loc}} + 2 \langle \operatorname{Tr} \Phi^{3} \rangle \big|_{\text{loc}} \Big) , \end{split}$$
(3.113)
$$\begin{split} W_{4}^{\text{loc}} &= \frac{1}{24} \Big(\langle \operatorname{Tr} \Phi \rangle \big|_{\text{loc}}^{4} - 6 \langle \operatorname{Tr} \Phi \rangle \big|_{\text{loc}}^{2} \langle \operatorname{Tr} \Phi^{2} \rangle \big|_{\text{loc}} + 3 \langle \operatorname{Tr} \Phi^{2} \rangle \big|_{\text{loc}} \\ &+ 8 \langle \operatorname{Tr} \Phi \rangle \big|_{\text{loc}} \langle \operatorname{Tr} \Phi^{3} \rangle \big|_{\text{loc}} - 6 \langle \operatorname{Tr} \Phi^{4} \rangle \big|_{\text{loc}} \Big) , \end{split}$$

and so on. Plugging the localization results (3.112), after some long but straight-

forward algebra, we find

$$W_1^{\text{loc}} = \sum_u a_u , \qquad (3.114)$$

$$W_2^{\text{loc}} = \sum_{u_1 < u_2} a_{u_1} a_{u_2} - \frac{N m^2}{24} (E_2 - 1) + \frac{m^4}{288} (E_2^2 - E_4) C_2^0$$
(3.115)

$$+ \frac{m^{6}}{4320} \left(5E_{2}^{3} - 3E_{2}E_{4} - 2E_{6} \right) C_{4}^{0} + \frac{m^{6}}{3456} \left(E_{2}^{3} - 3E_{2}E_{4} + 2E_{6} \right) C_{2;11}^{0} + O(m^{8}) ,$$

$$W_{3}^{\text{loc}} = \sum_{u_{1} < u_{2} < u_{3}} a_{u_{1}} a_{u_{2}} a_{u_{3}} - \frac{(N-2)m^{2}}{24} \left(E_{2} - 1 \right) \sum_{u} a_{u} + \frac{m^{4}}{288} \left(E_{2}^{2} - E_{4} \right) \left(C_{2}^{0} \sum_{u} a_{u} - 2C_{2}^{1} \right)$$

$$+ \frac{m^{6}}{4320} \left(5E_{2}^{3} - 3E_{2}E_{4} - 2E_{6} \right) \left(C_{4}^{0} \sum_{u} a_{u} - 2C_{4}^{1} \right)$$

$$+ \frac{m^{6}}{3456} \left(E_{2}^{3} - 3E_{2}E_{4} + 2E_{6} \right) \left(C_{2;11}^{0} \sum_{u} a_{u} - 2C_{2;11}^{1} \right) + O(m^{8}) , \qquad (3.116)$$

$$W_{4}^{\text{loc}} = \sum_{u_{1} < \dots < u_{4}} a_{u_{1}} a_{u_{2}} a_{u_{3}} a_{u_{4}} - \frac{m^{2}}{24} \left(E_{2} - 1 \right) \left(\left(\sum_{u} a_{u} \right)^{2} + (N-6) \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} \right)$$

$$+ \frac{m^{4}}{288} \left(E_{2}^{2} - E_{4} \right) \left(C_{2}^{0} \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} - 2C_{2}^{1} \sum_{u} a_{u} + 3C_{2}^{2} - \binom{N}{2} \right)$$

$$+ \frac{Nm^{4}}{5760} \left(5N \left(3E_{2}^{2} - 2E_{4} - 2E_{2} + 1 \right) - 30E_{2}^{2} + 12E_{4} + 60E_{2} - 42 \right)$$

$$+ \frac{m^{6}}{6} \left(5E^{3} - 3E_{2}E_{4} - 2E_{2} \right) \left(C_{4}^{0} \sum_{u} a_{u} a_{u} - 2C_{1}^{1} \sum_{u} a_{u} + 3C_{2}^{2} - \binom{N}{2} \right)$$

$$+ \frac{m}{4320} \left(5E_2^3 - 3E_2E_4 - 2E_6\right) \left(C_4^0 \sum_{u_1 < u_2} a_{u_1}a_{u_2} - 2C_4^1 \sum_u a_u + 3C_4^2 - \frac{1}{2}C_2^0\right) + \frac{m^6}{3456} \left(E_2^3 - 3E_2E_4 + 2E_6\right) \left(C_{2;11}^0 \sum_{u_1 < u_2} a_{u_1}a_{u_2} - 2C_{2;11}^1 \sum_u a_u + 3C_{2;11}^2\right) - \frac{(N-6)m^6}{6912} \left(E_2 - 1\right) \left(E_2^2 - E_4\right) C_2^0 + O(m^8) .$$
(3.117)

It is remarkable to see in these expressions the same combinations of Eisenstein series and of lattice sums appearing in the W_n presented in (3.84)–(3.87). However, there are also some important differences which we are going to discuss.

The first observation is that, even though the classical part of the W_n^{loc} is the degree n symmetric polynomial in the vacuum expectation values, the full W_n^{loc} do not satisfy the corresponding chiral ring relations.⁵ Indeed, it is not difficult to verify

⁵This was already noted in [19, 20, 21] for pure $\mathcal{N} = 2$ SYM theories.

that 6

$$W_{2,3,4}^{\text{loc}}\Big|_{\mathrm{U}(1)} \neq 0 , \qquad W_4^{\text{loc}}\Big|_{\mathrm{U}(2)} \neq 0 , \qquad W_4^{\text{loc}}\Big|_{\mathrm{U}(3)} \neq 0 , \qquad (3.118)$$

whereas in all these cases one should expect a vanishing result if the W_n^{loc} were the quantum version of the classical symmetric polynomials. We find that enforcing the chiral ring relations allows us to make contact with the results for the W_n coming from the Seiberg-Witten curves. This amounts a redefinition of W_n^{loc} , and thereby a different choice of the generators for the chiral ring.

The second observation is that our explicit localization results allow us to perform this redefinition in a systematic way. Indeed, from

$$W_2^{\text{loc}}\Big|_{\mathrm{U}(1)} = -\frac{m^2}{24} (E_2 - 1) , \qquad (3.119)$$

we immediately realize that the "good" operator at level 2 can be obtained from W_2^{loc} by removing the constant m^2 term proportional to $(E_2 - 1)$. We are thus led to define⁷

$$\widehat{W}_2 = W_2^{\text{loc}} + \frac{N m^2}{24} \left(E_2 - 1 \right) \,. \tag{3.120}$$

Similarly, at level 3 we find that the term responsible for the inequalities in (3.118) is again the m^2 part proportional to $(E_2 - 1)$, so that the desired operator is

$$\widehat{W}_3 = W_3^{\text{loc}} + \frac{(N-2)m^2}{24} \left(E_2 - 1\right) \sum_u a_u . \qquad (3.121)$$

At level 4 we see that the non-vanishing results in (3.118) are due again to the m^2 terms proportional to $(E_2 - 1)$ but also to the *a*-independent terms at order m^4 and

⁶Recall that the localization formulas formally hold true also for N = 1.

⁷It is interesting to note that also the prepotential of $\mathcal{N} = 2^*$ theories satisfies the duality properties discussed in [11, 12] only if an *a*-independent term proportional to m^2 , which is not quasi-modular, is discarded. Such a constant term in the prepotential does not, however, influence the effective action.

to the m^6 terms in the last line of (3.117). This motivates us to introduce

$$\widehat{W}_{4} = W_{4}^{\text{loc}} + \frac{m^{2}}{24} \left(E_{2} - 1 \right) \left(\left(\sum_{u} a_{u} \right)^{2} + (N - 6) \sum_{u_{1} < u_{2}} a_{u_{1}} a_{u_{2}} \right) - \frac{N m^{4}}{5760} \left(5N \left(3E_{2}^{2} - 2E_{4} - 2E_{2} + 1 \right) - 30E_{2}^{2} + 12E_{4} + 60E_{2} - 42 \right)$$
$$+ \frac{(N - 6) m^{6}}{6912} \left(E_{2} - 1 \right) \left(E_{2}^{2} - E_{4} \right) C_{2}^{0} .$$
(3.122)

It is interesting to observe that the difference between \widehat{W}_n and W_n^{loc} only consists of terms whose coefficients are polynomials in the Eisenstein series that do not have a definite modular weight, whereas the common terms at order $m^{2\ell}$ are quasi-modular forms of weight 2ℓ . Removing all such inhomogeneous terms from the W_n^{loc} yields the one-point functions that satisfy the classical chiral ring relations. Furthermore, it is worth noticing that (3.122) can be rewritten as

$$\widehat{W}_{4} = W_{4}^{\text{loc}} + \frac{(N-6)m^{2}}{24} (E_{2}-1)W_{2}^{\text{loc}} + \frac{m^{2}}{24} (E_{2}-1)W_{1}^{2} - \frac{Nm^{4}}{5760} \left(5N(E_{2}^{2}-2E_{4}+2E_{2}-1)+30E_{2}^{2}+12E_{4}-60E_{2}+18\right).$$
(3.123)

The fact that the m^6 terms are exactly reabsorbed is a very strong indication that the above formula is exact in m. Notice also that this redefinition, like the previous ones (3.120) and (3.121), is exact in the gauge coupling.

The most important point, however, is that the resulting expressions for the \widehat{W}_n derived from the localization formulas precisely match those for the W_n obtained from the SW curves in the previous section. Indeed, comparing (3.120)–(3.122) with (3.85)–(3.87), we have

$$\widehat{W}_n = W_n \ . \tag{3.124}$$

Our calculations provide an explicit proof of this equivalence for $n \leq 4$, but of course they can be generalized to higher levels. Summarizing, we have found that the quantum coordinates of the moduli space computed using the SW curves for the $\mathcal{N} = 2^* U(N)$ theory agree with those obtained from the localization formulas provided on the latter we enforce the classical chiral ring relations obeyed by the symmetric polynomials. Enforcing these relations is clearly a choice that amounts to selecting a particular basis for the generators of the chiral ring. It would be interesting to explore the possibility of modifying the localization prescription in order to obtain chiral observables that automatically satisfy such relations without the need for subtracting the non-quasi-modular terms.

3.6 1-Instanton Results

In the previous sections we have presented a set of results that are exact in the gauge coupling constant for quantities that have been evaluated order by order in the hypermultiplet mass. Here instead, we exhibit a result that is exact in m but is valid only at the 1-instanton level. To do so let us consider the localization results (3.111) for the one-point functions $\langle \operatorname{Tr} \Phi^n \rangle \big|_{\operatorname{loc}}$, and focus on the terms proportional to q corresponding to k = 1. Actually, the calculations at k = 1 can be easily performed also for higher rank groups and pushed to higher order in the mass without any problems. Collecting these results, it is does not take long to realize that they have a very regular pattern and can be written compactly as

$$\left\langle \operatorname{Tr} \Phi^{n} \right\rangle \Big|_{k=1} = -n(n-1) q \, m^{2} \left(C^{n-2} - m^{2} C_{2}^{n-2} - \frac{m^{4}}{2} C_{2;11}^{n-2} - \frac{m^{6}}{24} C_{2;1111}^{n-2} + \cdots \right)$$

$$= -n(n-1) q \, m^{2} \left(C^{n-2} - \sum_{\ell=0} \frac{m^{2+\ell}}{\ell!} C_{2;\underline{1}\ldots\underline{1}}^{n-2} \right).$$
(3.125)

Notice that $C_{2;1\cdots 1}^{p}$ with an odd number of 1's is zero, and that for a U(N) theory only N-1 terms are present in the sum over ℓ . Using the explicit form of the lattice sums (3.13), one can resum the above expression and find

$$\left\langle \operatorname{Tr} \Phi^{n} \right\rangle \Big|_{k=1} = -n(n-1) q m^{2} \sum_{\lambda \in \mathcal{W}} (\lambda \cdot \phi)^{n-2} \left[1 - \sum_{\alpha \in \Psi_{\lambda}} \frac{m^{2}}{(\alpha \cdot \phi)^{2}} \prod_{\beta \in \Psi_{\alpha}} \left(1 + \frac{m}{\beta \cdot \phi} \right) \right].$$
(3.126)

This is a generalization of an analogous formula for the prepotential found in [11, 12], to the case of the chiral observables of the $\mathcal{N} = 2^*$ theory. Being exact in m, we can use (3.126) to decouple the hypermultiplet by sending its mass to infinity and thus obtain the 1-instanton contribution to the one-point function of the single trace operators in the pure $\mathcal{N} = 2$ U(N) gauge theory. More precisely, this decoupling limit is

$$m \to \infty$$
 and $q \to 0$ with $q m^{2N} \equiv \Lambda^{2N}$ fixed. (3.127)

Recalling that the number of roots β in Ψ_{α} is 2N - 4, we see that the highest mass power in (3.126) is precisely m^{2N} , so that in the decoupling limit we get

$$\langle \operatorname{Tr} \Phi^n \rangle \Big|_{k=1} = n(n-1) \Lambda^{2N} \sum_{\lambda \in \mathcal{W}} \sum_{\alpha \in \Psi_\lambda} \frac{(\lambda \cdot \phi)^{n-2}}{(\alpha \cdot \phi)^2} \prod_{\beta \in \Psi_\alpha} \frac{1}{\beta \cdot \phi}$$
 (3.128)

We remark that for n = 2 this formula agrees with the 1-instanton prepotential of the pure $\mathcal{N} = 2$ theory, which was derived in [84, 85] using completely different methods. Indeed, through the Matone relation [83] $\langle \operatorname{Tr} \Phi^2 \rangle$ and the prepotential at 1 instanton are proportional to each other.

Moreover, if we restrict to SU(N), it is possible to verify that (3.128) is in full agreement with the chiral ring relations of the pure $\mathcal{N} = 2$ SYM theory that follow by expanding in inverse powers of z the identity [19, 20, 21]

$$\left\langle \operatorname{Tr} \frac{1}{z - \Phi} \right\rangle = \frac{P_N'(z)}{\sqrt{P_N^2(z) - 4\Lambda^{2N}}}$$
(3.129)

where

$$P_N(z) = z^N + \sum_{\ell=2}^N u_\ell \, z^{N-\ell} \,, \qquad (3.130)$$

is a degree N polynomial that encodes the Coulomb moduli u_{ℓ} appearing in the SW curve of the pure SU(N) SYM theory.

It would be nice to see whether the formulas (3.126) and (3.128) for generic n are valid also for other groups, as is the case for the n = 2 case [85, 11, 12].

Fundamental Matters

4

In this chapter, we study $\mathcal{N} = 2$ superconformal theories with gauge group SU(N) and 2N fundamental flavours in a locus of the Coulomb branch with a \mathbb{Z}_N symmetry. In this special vacuum, we calculate the prepotential, the dual periods and the period matrix using equivariant localization. In the conformal limit, we find that the period matrix is completely specified by $\left[\frac{N}{2}\right]$ effective couplings. On each of these, we show that the S-duality group acts as a generalized triangle group and that its hauptmodul can be used to write a non-perturbatively exact relation between each effective coupling and the bare one. For N = 2, 3, 4 and 6, the generalized triangle group is an arithmetic Hecke group which contains a subgroup that is also a congruence subgroup of the modular group PSL(2, \mathbb{Z}). For these cases, we introduce mass deformations that respect the symmetries of the special vacuum and show that the constraints arising from S-duality make it possible to resum the instanton contributions to the period matrix in terms of meromorphic modular forms which solve modular anomaly equations.

4.1 Introduction

One of the most interesting properties of supersymmetric gauge theories is the existence of non-perturbative S-dualities that relate their weak- and strong-coupling behaviour.¹ Recently, there has been much progress in understanding these dualities in conformally invariant $\mathcal{N} = 2$ supersymmetric gauge theories in four dimensions, especially following the seminal work of Gaiotto [22]. In that work, the fourdimensional $\mathcal{N} = 2$ theories were realized as compactifications of the six-dimensional (2,0) theory on a punctured Riemann surface Σ . One of the important results of this approach was to identify the complex structure moduli space of Σ with the space of gauge couplings modulo the action of the S-duality group. For linear quiver gauge theories in the weak coupling limit, the Riemann surface degenerates into a collection of three-punctured spheres connected by long thin tubes, and the sewing parameters are identified with the bare coupling constants of the superconformal gauge theory.

This approach is fruitfully contrasted with the original solution of $\mathcal{N} = 2$ gauge theories due to Seiberg and Witten [1, 2], where the quantum effective action on the Coulomb branch is obtained from an algebraic curve describing a Riemann surface, and an associated holomorphic differential. For generic vacuum expectation values of the scalar fields in the adjoint gauge multiplet, the quantum effective action describes a $\mathcal{N} = 2$ supersymmetric theory with gauge group $U(1)^r$, where r is the rank of the original non-abelian gauge group. The matrix of effective couplings τ_{ij} between the various U(1)'s is identified with the period matrix of the Seiberg-Witten curve. For gauge groups with large r, it becomes difficult to use the Seiberg-Witten curve and the corresponding differential to do explicit calculations. In such cases, however, it is possible to make progress using equivariant localization methods [3, 4, 51, 87]

¹For a recent review we refer the reader to [86].

which allow one to compute the prepotential, the dual periods and the period matrix of the effective action order by order in an instanton expansion. Interestingly, the instanton counting parameters in this expansion have a natural interpretation as the bare coupling constants of the superconformal gauge theory [22, 88].

Whichever approach one uses to study the low-energy theory, a natural question to ask is whether the non-perturbative S-duality group can be used to solve for the effective action. For $\mathcal{N} = 2^{\star}$ theories (i.e. mass deformed $\mathcal{N} = 4$ theories) with unitary gauge groups it has been shown [10, 23, 24, 25] that the constraints coming from S-duality take the form of a modular anomaly equation whose solution allows one to reconstruct the prepotential on the Coulomb branch order by order in the mass of the adjoint hypermultiplet to all orders in the gauge coupling. To achieve this result one has to organize the low-energy effective prepotential as a semi-classical expansion in inverse powers of the vacuum expectation values of the scalar fields in the gauge vector multiplet and realize that the coefficients of this expansion satisfy a recursion relation whose solution can be written in terms of quasi-modular forms of $PSL(2,\mathbb{Z})$ acting on the bare gauge coupling. These modular forms resum the instanton series and therefore provide an exact result. It is of particular importance that $\mathcal{N} = 2^{\star}$ theories are characterized by the absence of any renormalization of the coupling constant, even non-perturbatively; thus, the bare coupling is the only coupling that is present in the effective theory. This procedure has been applied also to $\mathcal{N} = 2^*$ theories with arbitrary gauge groups in [12, 11], where it has been observed that for non-simply laced algebras the effective prepotential is expressed in terms of quasi-modular forms of congruence subgroups of $PSL(2,\mathbb{Z})$.

In this work we study $\mathcal{N} = 2$ gauge theories with gauge group SU(N) and 2N fundamental flavours, generalizing the analysis of the SU(3) gauge theory with six flavours recently presented in [26]. When all flavours are massless, these SQCD theories are superconformal. However, unlike the case of $\mathcal{N} = 2^*$ theories, the bare

gauge coupling in $\mathcal{N} = 2$ SQCD is renormalized by quantum corrections which arise from a finite 1-loop contribution as well as from an infinite series of non-perturbative contributions due to instantons. In general these corrections are different for the various U(1) factors and thus one expects to find several effective couplings in the low-energy theory.

This chapter is divided into two parts. In the first part, we work in the conformal limit with all flavour masses set to zero, and calculate various observables of the effective theory such as the prepotential, the period integrals and the period matrix, using equivariant localization. In particular, we work in a special locus of the Coulomb branch which possesses a \mathbb{Z}_N symmetry and which we call the "special vacuum" [27]. In this special vacuum, the period matrix has fewer independent components than it does at a generic point of the moduli space. More precisely, when all quantum corrections are taken into account there are $\left[\frac{N}{2}\right]$ distinct matrix structures which correspond to $\left[\frac{N}{2}\right]$ renormalized coupling constants in the effective theory.² Of course, at leading order such renormalized couplings are all equal to the bare coupling, but when 1-loop and instanton corrections are taken into account, they begin to differ from one another. Given that the S-duality group naturally acts on the bare coupling, an obvious question to ask is how S-duality is realized on the various parameters of the quantum theory. The answer we provide in this paper is that on each individual effective coupling S-duality acts as a generalized triangle group (see for example [28]). Moreover, using this insight, we propose a non-perturbatively *exact* relation between the bare coupling and the renormalized ones that takes a universal form in terms of the *j*-invariants of the triangle groups. We perform several successful checks of this proposal by comparing the instanton contributions predicted by the exact relation with the explicit results obtained from multi-instanton localization. As a further evidence in favour of our proposal, we show that the action of S-duality on the renormalized couplings is fully consistent

²Here $[\cdot]$ denotes the floor function.

with the action on the bare coupling as obtained from Gaiotto's analysis [22]. We believe that our results, and in particular the exact relation we propose, can play an important role in the study of these SQCD theories at strong coupling [29]. This is because the *j*-invariants have a well-understood behaviour near those cusp points where the coupling constants become large and the usual weak-coupling expansion cannot be used.

In the second part of the chapter we move away from the conformal limit by giving a mass to the fundamental flavour hypermultiplets. For generic masses the \mathbb{Z}_N symmetry of the special vacuum is broken; to avoid this, we restrict our analysis to \mathbb{Z}_N -symmetric mass configurations so that the modular structure uncovered in the massless limit gets deformed in a natural and smooth manner. In particular, with these \mathbb{Z}_N -symmetric mass configurations we find that the $\left[\frac{N}{2}\right]$ matrix structures of the massless theories are preserved, while the $\left[\frac{N}{2}\right]$ effective couplings simply receive further contributions proportional to the hypermultiplet masses. Building on earlier literature [30, 31], this analysis was already carried out for the SU(2) theory in [23, 24], where it was shown that the prepotential can be written in terms of quasimodular forms of the modular group $PSL(2,\mathbb{Z})$. Moreover, after expanding the prepotential in powers of the flavour masses, it was realized that the coefficients of this expansion satisfy a modular anomaly equation that takes the form of a recursion relation, similar to that of the $\mathcal{N} = 2^*$ case. These results have been recently extended to the SU(3) theory with six massive flavours in [26], where it has been shown that the prepotential, the dual periods and the period matrix are constrained by S-duality to obey again a recursion relation that can be written as a modular anomaly equation. In this case, the solutions of this equation are quasi-modular forms of $\Gamma_1(3)$, which is a subgroup of the S-duality group that is also a congruence subgroup of $PSL(2,\mathbb{Z})$.³ Here we further extend these results to the general SU(N)

³The relevance of $\Gamma_1(3)$ and of its modular forms for the effective SU(3) theory with six flavours was already observed long ago in [89, 90, 91].

theory with 2N massive flavours and show that the constraints arising from S-duality can always be written as a recursion relation for any N. However, beyond this step, the analysis crucially depends on the arithmetic properties of the S-duality group. It turns out that for N = 2, 3, 4 and 6, the S-duality group acting on each quantum coupling always has a subgroup which is a congruence subgroup of PSL $(2, \mathbb{Z})$. For these theories, which we call arithmetic, the discussion proceeds along the same lines described in [26] for the SU(3) theory, with one important modification: in the higher rank cases, the S-duality constraints are written as *coupled* modular anomaly equations. These coupled equations are nevertheless integrable and their solutions turn out to be polynomials in meromorphic quasi-modular forms of congruence subgroups of PSL $(2, \mathbb{Z})$. For all non-arithmetic theories, instead, S-duality acts as generalized triangle groups and one would need to use their automorphic forms to solve for various observables. Here, we restrict our analysis only to the massive arithmetic cases, leaving the study of the non-arithmetic cases for the future.

Massless $\mathcal{N} = 2$ SQCD and Duality Groups

In this part, we discuss $\mathcal{N} = 2$ SQCD theories with massless fundamental hypermultiplets.

Massless SQCD and the Special Vacuum

We begin by reviewing the main features of $\mathcal{N} = 2$ SQCD theories with unitary gauge groups U(N). These theories are superconformal invariant if the number of flavours is 2N.

As usual, we can combine the bare Yang-Mills coupling g and the θ -angle into the complex variable

$$\tau_0 = \frac{\theta}{2\pi} + \mathrm{i} \frac{4\pi}{g^2} , \qquad (4.1)$$

so that the instanton counting parameter q_0 is defined as

4.2

$$q_0 = e^{2\pi i \tau_0}$$
 (4.2)

The low-energy effective dynamics of these $\mathcal{N} = 2$ theories is completely determined by the prepotential, which we now describe. The prepotential F admits a decomposition into classical (tree-level), perturbative (1-loop), and non- perturbative (instanton) contributions:

$$F = F_{\text{class}} + F_{1\text{-loop}} + F_{\text{inst}} .$$

$$(4.3)$$

Classical Contribution

For the U(N) gauge theory the classical prepotential is given by

$$F_{\text{class}} = \mathrm{i}\pi\tau_0 \operatorname{tr} \langle A \rangle^2 = \mathrm{i}\pi\tau_0 \sum_{u=1}^N A_u^2, \qquad (4.4)$$

where the vacuum expectation value of the adjoint scalar A is

$$\langle A \rangle = \operatorname{diag}\left(A_1, \cdots, A_N\right)$$
 (4.5)

For unitary gauge groups the A_u 's are unrestricted, while for special unitary groups we have to impose the tracelessness condition

$$\sum_{u=1}^{N} A_u = 0 . (4.6)$$

Throughout this paper we satisfy this constraint by taking

$$A_{u} = \begin{cases} a_{u} & \text{for } u = 1, \cdots, N - 1, \\ -(a_{1} + \dots + a_{N-1}) & \text{for } u = N. \end{cases}$$
(4.7)

When referring to the SU(N) theory we will use the indices $i, j, \dots \in \{1, \dots, N-1\}$ to label the Cartan directions.

Perturbative Contribution

The perturbative (1-loop) contribution to the prepotential is independent of the bare coupling τ_0 and is given by

$$F_{1-\text{loop}} = \sum_{u \neq v=1}^{N} \gamma(A_u - A_v) - 2N \sum_{u=1}^{N} \gamma(A_u) , \qquad (4.8)$$

where (see for example [6])

$$\gamma(x) = -\frac{x^2}{4} \log\left(\frac{x^2}{\Lambda^2}\right) . \tag{4.9}$$

Here Λ is an arbitrary mass scale, which actually drops out from $F_{1-\text{loop}}$ due to conformal invariance.

Instanton Contribution

The non-perturbative contributions to the prepotential can be explicitly calculated using the methods of equivariant localization [3, 4, 51, 87] (see also [80] for technical details) and are of the form

$$F_{\text{inst.}} = \sum_{k=1}^{\infty} F_k(u_r) \, q_0^k \tag{4.10}$$

where

$$u_r = \sum_{u=1}^N A_u^r \tag{4.11}$$

for $r = 1, \dots, N$ are the Casimir invariants of the gauge group. The function F_k represents the k-instanton contribution to the prepotential and, on dimensional grounds, must have mass dimension 2.

4.2.2 The Special Vacuum

In the following we will study the massless SQCD theories in the so-called *special* vacuum [27] which is defined as the locus of points on the moduli space where

$$u_r = 0 \quad \text{for} \quad r = 1, \cdots, N - 1 \;.$$
 (4.12)

For SU(N) theories the condition $u_1 = 0$ is nothing but (4.7), while the other conditions select vacuum configurations with special properties.⁴

The special vacuum restriction (4.12) can be implemented by choosing the vacuum expectation values of the adjoint SU(N) scalar as

$$a_i = \omega^{i-1}a \tag{4.13}$$

for $i = 1, \dots, N-1$, where

$$\omega = \mathrm{e}^{\frac{2\pi\mathrm{i}}{N}} \,. \tag{4.14}$$

We thus see that the special vacuum can be parametrized by a single scale a and that it possesses a \mathbb{Z}_N symmetry.

4.2.3 Observables

We now discuss the properties of some observables in the special vacuum, starting with the prepotential.

⁴In the SU(2) theory there is clearly only one condition, namely $u_1 = 0$ and the notion of special vacuum does not apply in this case. Despite this fact, most of the subsequent formulas formally hold also for SU(2).

The Prepotential

In the special vacuum several simplifications occur when one evaluates the prepotential. For instance, the classical prepotential (4.4) vanishes and the \mathbb{Z}_N -invariance of the special vacuum implies that for large *a* the prepotential has a semi-classical expansion of the form

$$F = \sum_{n=1}^{\infty} \frac{f_n(q_0)}{a^{nN}} .$$
 (4.15)

The coefficients f_n 's must have mass dimension equal to (nN + 2); however, since the flavours are massless, the only available scale is a and it is not possible to give f_n the required mass dimensions. Thus the prepotential identically vanishes in the special vacuum.⁵

Dual Periods

In the SU(N) theory the dual periods $a_i^{\rm D}$ are defined by

$$a_i^{\rm D} = \frac{1}{2\pi i} \frac{\partial F}{\partial a_i} \,. \tag{4.16}$$

As in the special vacuum all a_i 's are proportional to each other, this is also true of the dual periods. For example one can verify that

$$a_i^{\mathrm{D}} = -\left(\omega + \omega^2 + \dots + \omega^i\right)a_{N-1}^{\mathrm{D}}$$

$$(4.17)$$

for any $i = 1, \dots, N-1$. Therefore, in the special vacuum without any loss of generality we can choose the following conjugate pair of variables: $(a_{N-1}^{\rm D}, a_1)$. To simplify notation, we will omit the subscripts and denote these just by $(a^{\rm D}, a)$.

⁵The case N = 2 is clearly an exception. Indeed, the prepotential of the massless SU(2) theory is proportional to a^2 , which has the right mass dimension and is \mathbb{Z}_2 -symmetric (see for instance [24, 23]).

The classical contribution to $a^{\rm D}$ is given by

$$a_{\text{class}}^{\text{D}} = \tau_0 \left(a_1 + a_2 + \dots + 2a_{N-1} \right)$$

= $c_N \tau_0 a$, (4.18)

where the second line follows upon using the special vacuum values (4.13) which lead to

$$c_N = \frac{(1-\omega)}{\omega^2} \ . \tag{4.19}$$

The classical dual period receives both 1-loop and instanton corrections, even in the massless theory. Physically, this corresponds to a non-perturbative redefinition of the bare coupling constant τ_0 into a new renormalized coupling constant that we denote τ . This renormalized coupling constant is defined in such a way that the quantum corrected dual period takes the simple form

$$a^{\mathrm{D}} = c_N \,\tau \,a \,, \tag{4.20}$$

namely the same classical expression (4.18) with τ_0 replaced by τ . The latter admits the following non-perturbative expansion

$$2\pi i \tau = 2\pi i \tau_0 + i\pi + \log b_0 + \sum_{k=1}^{\infty} b_k q_0^k .$$
(4.21)

In this expression, the logarithmic term represents a finite contribution at 1-loop, while the term proportional to q_0^k is the k-instanton contribution.

The Period Matrix

In the SU(N) theory the period matrix Ω is the $(N-1) \times (N-1)$ matrix defined as

$$\Omega_{ij} = \frac{1}{2\pi i} \frac{\partial^2 F}{\partial a_i \partial a_j} . \qquad (4.22)$$

The classical part of the period matrix is simply given by

$$\Omega_{\text{class}} = \tau_0 \,\mathcal{C} \tag{4.23}$$

where

$$C = \begin{pmatrix} 2 & 1 & \cdots & 1 \\ 1 & 2 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 1 & 1 & \cdots & 2 \end{pmatrix}$$
(4.24)

is the Cartan matrix corresponding to our parametrization of SU(N).

For N > 3 this simple structure is lost [90, 92] when perturbative and instanton contributions are taken into account, even in the special vacuum. For example, at 1-loop from (4.8) one finds

$$\Omega_{1-\text{loop}} = \frac{\mathrm{i}}{\pi} \left(\log(2N) \,\mathcal{C} + \mathcal{G} \right) \tag{4.25}$$

where the matrix elements of \mathcal{G} are given by [90]

$$\mathcal{G}_{ii} = 2\log\sin\left(\frac{i\pi}{N}\right) ,$$

$$\mathcal{G}_{ij} = \log\sin\left(\frac{i\pi}{N}\right) + \log\sin\left(\frac{j\pi}{N}\right) - \log\sin\left(\frac{|i-j|\pi}{N}\right) \quad \text{for } i \neq j .$$

$$(4.26)$$

For N = 3 it is easy to see that \mathcal{G} is proportional to the Cartan matrix \mathcal{C} , but this relation does not hold for N > 3. Indeed, a closer inspection of (4.26) reveals that

it is possible to identify $\left[\frac{N}{2}\right]$ different matrix structures. A similar result is found even after the instanton contributions are taken into account. Thus, in general the complete period matrix Ω can be written as

$$\Omega = \underbrace{\tau_1 \ \mathcal{M}_1 + \tau_2 \ \mathcal{M}_2 + \cdots}_{\left[\frac{N}{2}\right] \text{ terms}}$$
(4.27)

where the \mathcal{M}_k 's are independent matrix structures and the τ_k are distinct complex couplings that characterize the effective theory. Of course, one could in principle use any basis of matrices \mathcal{M}_k to write Ω , but a particularly insightful choice is the one that "diagonalizes" the action of the S-duality group. In such a basis, under S-duality each \mathcal{M}_k stays invariant and each τ_k transforms *individually* as

$$\tau_k \to -\frac{1}{\lambda_k \, \tau_k} \tag{4.28}$$

for some positive λ_k . We will explicitly show in a series of examples that the spectrum of λ_k is given by

$$\lambda_k = 4\sin^2\left(\frac{k\,\pi}{N}\right) \ . \tag{4.29}$$

Note that for $N \in \{2, 3, 4, 6\}$ all the λ_k 's take integer values. We call these cases *arithmetic*. If instead $N \notin \{2, 3, 4, 6\}$, then the λ_k 's are not necessarily integer. We refer to the latter as the *non-arithmetic* cases. Moreover, we will find that for any N the coupling τ_1 in (4.27) coincides with the coupling τ that appears in the expression (4.20) for the dual period $a^{\rm D}$.

In order to show these facts, we now turn to a detailed discussion of the S-duality group.
The S-duality group of $\mathcal{N} = 2$ SQCD has been derived in [93]. Here, we focus on the massless case in the special vacuum, for which the Seiberg-Witten curve takes the following hyperelliptic form

$$y^{2} = (x^{N} - u_{N})^{2} - h x^{2N} . (4.30)$$

Here u_N is the only non-zero Coulomb modulus labeling the special vacuum and h is a function of the gauge coupling given by (see for example [80, 26])

$$h = \frac{4q_0}{(1+q_0)^2} \ . \tag{4.31}$$

The Seiberg-Witten curve degenerates when its discriminant vanishes and from (4.30) it is easy to see that this happens at $h = 0, 1, \infty$. The monodromies around these points generate the S-duality group [93]. We will take this to be our working definition of the S-duality group in what follows. In Section 4.5 we will rederive this result by a completely different method.

We begin by choosing a canonical homology basis of cycles for the U(N) theory described by (4.30), which we denote by hatted variables. Specifically, we introduce $\hat{\alpha}$ and $\hat{\beta}$ cycles with the following intersections

$$\hat{\alpha}_{u} \cap \hat{\alpha}_{v} = \hat{\beta}_{u} \cap \hat{\beta}_{v} = 0 ,$$

$$\begin{pmatrix} \hat{\alpha}_{u} \cap \hat{\beta}_{v} \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 \\ 0 & 1 & -1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & 0 & \cdots & 1 \end{pmatrix} ,$$
(4.32)

for $u, v = 1, \dots, N$. These cycles are not linearly independent since

$$\sum_{u=1}^{N} \hat{\alpha}_{u} = 0 \quad \text{and} \quad \sum_{v=1}^{N} \hat{\beta}_{v} = 0 .$$
 (4.33)

In the special vacuum there is a natural \mathbb{Z}_N symmetry that rotates this basis clockwise and is generated by

$$\Phi : \begin{cases} \hat{\alpha}_u \longrightarrow \hat{\alpha}_{u-1} ,\\ \hat{\beta}_v \longrightarrow \hat{\beta}_{v-1} . \end{cases}$$

$$(4.34)$$

Physical observables are insensitive to this \mathbb{Z}_N rotation.

The action of S- and T-transformations on this basis of cycles has been determined in [93] from the monodromy around the points $h = \infty$ and h = 0, respectively, and is given by

$$S : \begin{cases} \hat{\alpha}_u \to \hat{\beta}_u , \\ \hat{\beta}_v \to \hat{\alpha}_{v-1} , \end{cases} \quad \text{and} \quad T : \begin{cases} \hat{\alpha}_u \to \hat{\alpha}_u , \\ \hat{\beta}_v \to \hat{\beta}_v + \hat{\alpha}_v - \hat{\alpha}_{v-1} . \end{cases}$$
(4.35)

In the SU(N) theory we can choose the independent cycles as follows:

$$\alpha_i = \hat{\alpha}_i \quad \text{and} \quad \beta_j = \sum_{i=1}^j \hat{\beta}_i$$
(4.36)

for $i, j = 1, \dots, N-1$. Using (4.32) one can easily check that this basis is symplectic, in the sense that $\alpha_i \cap \alpha_j = \beta_i \cap \beta_j = 0$ and $\alpha_i \cap \beta_j = \delta_{ij}$.

The restriction of the S and T transformations to the SU(N) basis (4.36) follows directly from (4.35). If we represent them as $(2N - 2) \times (2N - 2)$ matrices acting

on the
$$(2N-2)$$
 vector $\begin{pmatrix} \beta \\ \alpha \end{pmatrix}$, we find

$$S = \begin{pmatrix} 0 & \mathcal{B} \\ -(\mathcal{B}^{t})^{-1} & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} \mathbb{1} & \mathcal{C} \\ 0 & \mathbb{1} \end{pmatrix}$$

where

$$\mathcal{B} = \begin{pmatrix} -1 & -1 & -1 & \cdots & -1 \\ 0 & -1 & -1 & \cdots & -1 \\ 0 & 0 & -1 & \cdots & -1 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & -1 \end{pmatrix} , \qquad (4.38)$$

and C is the Cartan matrix (4.24). It is interesting to observe that

$$S^2 = V av{4.39}$$

(4.37)

where V is an Sp $(2N - 2, \mathbb{Z})$ matrix that implements the \mathbb{Z}_N transformation (4.34) on the SU(N) basis of cycles (α_i, β_j) , given by

$$V = \begin{pmatrix} (\mathcal{V}^{t})^{-1} & 0\\ 0 & \mathcal{V} \end{pmatrix}$$
(4.40)

where

$$\mathcal{V} = \begin{pmatrix} -1 & -1 & -1 & \cdots & -1 \\ 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{pmatrix} .$$
(4.41)

Notice that $\mathcal{V}^N = 1$. Such a transformation leaves the period matrix invariant and is a symmetry of the theory. This means that S effectively squares to the identity. From the transformations (4.37) on the homology cycles we can straightforwardly deduce how S and T act on the periods a_i and their duals $a_j^{\rm D}$, which are the integrals of the Seiberg-Witten differential associated to the curve (4.30) over the cycles α_i and β_j respectively. Focusing in particular on the S-transformation, we find

$$S(a_j^{\mathrm{D}}) = \left(\mathcal{B} \cdot a\right)_j,$$

$$S(a_i) = -\left(\left(\mathcal{B}^{\mathrm{t}}\right)^{-1} \cdot a^{\mathrm{D}}\right)_i$$
(4.42)

where \mathcal{B} is the matrix in (4.38). Thus, for our conjugate pair of variables $(a_{N-1}^{D}, a_{1}) \equiv (a^{D}, a)$ we have

$$S(a^{\rm D}) = -a_{N-1} = -\frac{1}{\omega^2} a , \qquad (4.43)$$
$$S(a) = a_1^{\rm D} = -\omega a^{\rm D} ,$$

where in each line the second equality follows upon using the special vacuum relations (4.13) and (4.17). As expected, the period and dual period integrals are exchanged under S-duality.

This result is quite useful since it allows us to deduce how S-duality acts on the gauge coupling τ . Indeed, if we take into account the link (4.20) between $a^{\rm D}$ and a, and apply to it the S-duality transformations, we find

$$S(a^{\rm D}) = -c_N S(\tau) \,\omega \, a^{\rm D} = -c_N^2 S(\tau) \,\tau \,\omega \,a \,.$$
(4.44)

Consistency with (4.43) implies that

$$\tau \to -\frac{1}{\lambda \, \tau}$$
 (4.45)

with

$$\lambda = -c_N^2 \,\omega^3 = -\frac{(1-\omega)^2}{\omega} = 4\sin^2\frac{\pi}{N} \,. \tag{4.46}$$

Here we have used (4.19) and (4.14). This is precisely the case k = 1 of the general formula (4.29), and thus we conclude that the coupling τ that appears in the relation between a and a^{D} is actually τ_{1} , according to our definition in (4.28). As we have already noticed, in the arithmetic cases $N \in \{2, 3, 4, 6\}$, the constant λ in (4.46) takes integer values.

4.4 The Arithmetic Theories

In this section, we collect the results obtained from localization calculations for the lower rank SQCD models and accumulate evidence for our conjecture regarding the form of the period matrix and the S-duality transformations of the quantum couplings that we anticipated at the end of Section 4.2.3. While the SU(2) and SU(3) theories have already been studied in the literature, for completeness we start by briefly recalling the main results for these cases.

4.4.1 *N* = 2

In this case, the period matrix is just a complex constant given by

$$\Omega = 2\tau_1 \tag{4.47}$$

where τ_1 is the only effective coupling of this theory. Using multi-instanton calculations (see for example [80, 88]), one can show that

$$2\pi i \tau_1 = \log q_0 + i\pi - \log 16 + \frac{1}{2} q_0 + \frac{13}{64} q_0^2 + \frac{23}{192} q_0^3 \cdots, \qquad (4.48)$$

which can be inverted order by order to give

$$q_0 = -16 q_1 \left(1 + 8 q_1 + 44 q_1^2 + \dots\right) = -16 \left(\frac{\eta(4\tau_1)}{\eta(\tau_1)}\right)^8 , \qquad (4.49)$$

where $q_1 = e^{2\pi i \tau_1}$ and η is the Dedekind η -function.

The analysis of the previous section shows that under the S-transformation, the period matrix transforms under a symplectic $\text{Sp}(2,\mathbb{Z})$ transformation⁶:

$$S : \Omega \to -\frac{1}{\Omega}$$
 (4.50)

From this it follows that the effective coupling τ_1 transforms as

$$S : \tau_1 \to -\frac{1}{4\tau_1},$$
 (4.51)

in agreement with (4.28) since $\lambda_1 = 4$ for N = 2. Using this in (4.49), we get

$$S \quad : \quad q_0 \to \frac{1}{q_0} \ . \tag{4.52}$$

Furthermore, by computing the dual period we find

$$a^{\mathrm{D}} = 2\tau_1 a \tag{4.53}$$

in agreement with the general formula (4.20) for $\omega = -1$.

 $^{^{6}}$ See section 4.5.1 for a more detailed discussion of S-duality for the SU(2) theory.

4.4.2 *N* = 3

In this case the period matrix turns out to be proportional to the SU(3) Cartan matrix

$$\Omega = \tau_1 \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \tag{4.54}$$

where τ_1 has the following instanton expansion (see for example [80, 26])

$$2\pi i \tau_1 = \log q_0 + i\pi - \log 27 + \frac{4}{9} q_0 + \frac{14}{81} q_0^2 + \frac{1948}{19683} q_0^3 \cdots$$
 (4.55)

As for the SU(2) case, the SU(3) theory in the special vacuum has a single τ_1 parameter even after the quantum corrections are taken into account. On inverting the above expansion, we get

$$q_0 = -27 q_1 \left(1 + 12 q_1 + 90 q_1^2 + \dots\right) = -27 \left(\frac{\eta(3\tau_1)}{\eta(\tau_1)}\right)^{12}$$
(4.56)

where, as before, $q_1 = e^{2\pi i \tau_1}$. Again, we have provided a non-perturbatively exact expression in terms of η -quotients. Using the Sp(4, Z) matrices derived in Section 4.3, one can check that S-duality leaves the SU(3) Cartan matrix invariant and acts on τ_1 as [26]:

$$S \quad : \quad \tau_1 \to -\frac{1}{3\tau_1} \ . \tag{4.57}$$

in agreement with (4.29) since $\lambda_1 = 3$ for N = 3; using this in (4.56), we easily see again that

$$S \quad : \quad q_0 \to \frac{1}{q_0} \ . \tag{4.58}$$

Finally, on computing the dual period in this case we find

$$a^{\mathrm{D}} = \mathrm{i}\sqrt{3}\tau_1 a \tag{4.59}$$

which confirms the general formula (4.20) since for SU(3) $\omega = e^{\frac{2\pi i}{3}}$.

4.4.3 *N* = 4

We now turn to the SU(4) theory. As always, the classical period matrix is proportional to the Cartan matrix of the gauge Lie algebra but this time another independent matrix structure appears when one takes into account the 1-loop and the instanton corrections. We have explicitly checked up to three instantons that it is possible to write the quantum period matrix as

$$\Omega = \tau_1 \,\mathcal{M}_1 + \tau_2 \,\mathcal{M}_2 \;, \tag{4.60}$$

where \mathcal{M}_1 and \mathcal{M}_2 are two 3×3 matrices given by

$$\mathcal{M}_{1} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \text{and} \quad \mathcal{M}_{2} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix} , \quad (4.61)$$

and the two couplings τ_1 and τ_2 have the following instanton expansions

$$2\pi i \tau_1 = \log q_0 + i\pi - \log 64 + \frac{3}{8}q_0 + \frac{141}{1024}q_0^2 + \frac{311}{4096}q_0^3 + \cdots , \qquad (4.62a)$$

$$2\pi i \tau_2 = \log q_0 + i\pi - \log 16 + \frac{1}{2} q_0 + \frac{13}{64} q_0^2 + \frac{23}{192} q_0^3 + \cdots$$
 (4.62b)

On inverting these expansions we find

$$q_0 = -64 q_1 \left(1 + 24 q_1 + 300 q_1^2 + \dots\right) = -64 \left(\frac{\eta(2\tau_1)}{\eta(\tau_1)}\right)^{24} , \qquad (4.63a)$$

$$q_0 = -16 q_2 \left(1 + 8 q_2 + 44 q_2^2 + \dots\right) = -16 \left(\frac{\eta(4\tau_2)}{\eta(\tau_2)}\right)^8 , \qquad (4.63b)$$

where we have introduced the notation

$$q_k = e^{2\pi i \tau_k} \tag{4.64}$$

for k = 1, 2. Once again, as for N = 2 and 3, the bare coupling can be expressed as a quotient of η - functions of the renormalized couplings. Notice that the q_0 -expansion of τ_2 is the same as that of the effective coupling of the SU(2) theory (see (4.48)). Although this coincidence may appear surprising at first glance, it is actually a consequence of the fact that this pair of couplings transform in the same way under S-duality. We will explicitly show this below, but this result can be anticipated by noticing that the general formula (4.29) implies that λ_2 for N = 4 and λ_1 for N = 2are both equal to 4.⁷

Let us now consider the action of S-duality on the period matrix (4.60). Using the $Sp(6,\mathbb{Z})$ transformations (4.37), we find that \mathcal{M}_1 and \mathcal{M}_2 are left invariant while

$$S \quad : \quad \tau_k \to -\frac{1}{\lambda_k \, \tau_k} \tag{4.65}$$

with $\lambda_1 = 2$ and $\lambda_2 = 4$, exactly as predicted by (4.28) and (4.29). Using these transformations in (4.63), we can check also in this case that

$$S \quad : \quad q_0 \to \frac{1}{q_0} \ . \tag{4.66}$$

By computing the dual period $a^{\rm D} \equiv a_3^{\rm D}$ in terms of $a \equiv a_1$, we find

$$a^{\rm D} = (i-1)\tau_1 a$$
 (4.67)

in agreement with (4.20) for $\omega = i$.

⁷Indeed, for all even N, $\lambda_{\frac{N}{2}} = 4$.

4.4.4 N = 6

We now turn to the last arithmetic case, namely the SU(6) theory. We have verified using localization techniques up to two instantons that in the special vacuum the period matrix can be written as a sum of three independent structures as follows

$$\Omega = \tau_1 \mathcal{M}_1 + \tau_2 \mathcal{M}_2 + \tau_3 \mathcal{M}_3 , \qquad (4.68)$$

where

$$\mathcal{M}_{1} = \begin{pmatrix} +\frac{1}{3} & +\frac{1}{2} & +\frac{1}{3} & 0 & -\frac{1}{6} \\ +\frac{1}{2} & +1 & +1 & +\frac{1}{2} & 0 \\ +\frac{1}{3} & +1 & +\frac{4}{3} & +1 & +\frac{1}{3} \\ 0 & +\frac{1}{2} & +1 & +1 & +\frac{1}{2} \\ -\frac{1}{6} & 0 & +\frac{1}{3} & +\frac{1}{2} & +\frac{1}{3} \end{pmatrix},$$

$$\mathcal{M}_{2} = \begin{pmatrix} +1 & +\frac{1}{2} & 0 & +1 & +\frac{1}{2} \\ +\frac{1}{2} & +1 & 0 & +\frac{1}{2} & +1 \\ 0 & 0 & 0 & 0 & 0 \\ +1 & +\frac{1}{2} & 0 & +1 & +\frac{1}{2} \\ +\frac{1}{2} & +1 & 0 & +\frac{1}{2} & +1 \end{pmatrix},$$

$$\mathcal{M}_{3} = \begin{pmatrix} +\frac{2}{3} & 0 & +\frac{2}{3} & 0 & +\frac{2}{3} \\ 0 & 0 & 0 & 0 & 0 \\ +\frac{2}{3} & 0 & +\frac{2}{3} & 0 & +\frac{2}{3} \\ 0 & 0 & 0 & 0 & 0 \\ +\frac{2}{3} & 0 & +\frac{2}{3} & 0 & +\frac{2}{3} \end{pmatrix}$$

$$(4.69)$$

and

$$2\pi i \tau_1 = \log q_0 + i\pi - \log 432 + \frac{5}{18} q_0 + \frac{485}{5184} q_0^2 + \cdots , \qquad (4.70a)$$

$$2\pi i \tau_2 = \log q_0 + i\pi - \log 27 + \frac{4}{9} q_0 + \frac{14}{81} q_0^2 + \cdots , \qquad (4.70b)$$

$$2\pi i \tau_3 = \log q_0 + i\pi - \log 16 + \frac{1}{2} q_0 + \frac{13}{64} q_0^2 + \cdots$$
 (4.70c)

We easily recognize that the q_0 -expansion of τ_2 is the same as that of the effective coupling of the SU(3) theory (see (4.55)), and that the q_0 -expansion of τ_3 is the same as that of the coupling τ_2 appearing in the SU(4) theory which, as we already remarked, is also the same as the coupling τ_1 of the SU(2) theory. Again these facts are a consequence of the symmetries of the formula (4.29) which imply that these pairs of couplings have the same transformations under S-duality.

Inverting the expansions (4.70), we obtain

$$q_0 = -432 q_1 \left(1 + 120 q_1 + 4140 q_1^2 + \cdots\right) , \qquad (4.71a)$$

$$q_0 = -27 q_2 \left(1 + 12 q_2 + 90 q_2^2 + \dots\right) = -27 \left(\frac{\eta(3\tau_2)}{\eta(\tau_2)}\right)^{12} , \qquad (4.71b)$$

$$q_0 = -16 q_3 \left(1 + 8 q_3^2 + 44 q_3^2 + \dots\right) = -16 \left(\frac{\eta(4\tau_3)}{\eta(\tau_3)}\right)^8 , \qquad (4.71c)$$

where we have used the notation (4.64). We observe that there appears to be no simple way to express q_0 in terms of η -quotients of τ_1 . However, we will revisit this issue in Section 4.5 where we will provide for all SU(N) models a universal formula for q_0 in terms of modular functions of *any* renormalized couplings τ_k , thus including also the τ_1 of the SU(6) theory.

Let us now consider the S-duality action on the period matrix (4.68). Under the $Sp(10,\mathbb{Z})$ transformation given in (4.37), we find that the three matrices (4.69) re-

main invariant while the couplings transform as

$$S \quad : \quad \tau_k \to -\frac{1}{\lambda_k \, \tau_k} \tag{4.72}$$

with $\lambda_1 = 1$, $\lambda_2 = 3$ and $\lambda_3 = 4$ in full agreement with (4.29). Exploiting the η -quotient expressions in (4.71), one can easily prove that the S-transformations of τ_2 and τ_3 lead again to

$$S \quad : \quad q_0 \to \frac{1}{q_0} \ . \tag{4.73}$$

Finally, by computing the dual period $a^{D} \equiv a_{5}^{D}$ in terms of $a \equiv a_{1}$ in the special vacuum, we obtain

$$a^{\mathrm{D}} = -a\tau_1 \tag{4.74}$$

which confirms once more (4.20), since for SU(6) we have $\omega = e^{\pi i/3}$.

Besides the S-duality action we should also consider the T-duality transformation of the effective couplings which is simply 8

$$T \quad : \quad \tau_k \to \tau_k + 1 \; . \tag{4.75}$$

Thus the previous results can be summarized by saying that in the arithmetic cases the duality transformations act as fractional linear transformations on each of the τ_k and form a subgroup of PSL(2, \mathbb{R}) generated by

$$S = \begin{pmatrix} 0 & 1/\sqrt{\lambda_k} \\ -\sqrt{\lambda_k} & 0 \end{pmatrix} \quad \text{and} \quad T = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$
(4.76)

with $\lambda_k \in \{1, 2, 3, 4\}$ as given by (4.29). We call this subgroup $\Gamma^*(\lambda_k)$. For $\lambda_k = 1$ this is the usual modular group $PSL(2,\mathbb{Z})$.

⁸See Section 4.5.1 for the SU(2) case, which has a distinct T- transformation.

ories

In this section we collect the results obtained so far and explain how our definition of S-duality fits in the general discussion presented in [22]. If we describe the SU(N)theory in the special vacuum by the Seiberg-Witten curve in the Gaiotto form

$$x^{N} = \frac{u_{N}}{t^{N-1}(t-1)(t-q_{0})} , \qquad (4.77)$$

then S-duality can be described as an action on the (x, t) variables given by [22]

$$S : (x, t) \to \left(-t^2 x, \frac{1}{t}\right) , \qquad (4.78)$$

which effectively amounts to an inversion of the bare coupling 9 :

$$S \quad : \quad q_0 \to \frac{1}{q_0} \ . \tag{4.79}$$

We have already seen in various explicit examples that the rule (4.79) is implied by the S-duality transformations of the renormalized couplings τ_k of the arithmetic theories, namely

$$S \quad : \quad \tau_k \to -\frac{1}{\lambda_k \, \tau_k} \tag{4.80}$$

with $\lambda_k \in \{2, 3, 4\}$. Actually, all these cases can be combined together by observing that the η -quotients in (4.49), (4.56), (4.63), (4.71) can be written as

$$q_0 = -(\lambda_k)^{\frac{6}{\lambda_k - 1}} \left(\frac{\eta(\lambda_k \tau_k)}{\eta(\tau_k)}\right)^{\frac{24}{\lambda_k - 1}} , \qquad (4.81)$$

⁹The curve (4.77) retains its form if (4.79) is accompanied by $u_N \to (-1)^N u_N/q_0$.

from which the inversion rule (4.79) immediately follows upon using the transformation properties of the Dedekind η -function under (4.80). The only case that is not covered by this formula is the relation between q_0 and τ_1 in the SU(6) theory, given by the first line of (4.71), for which there seems to be no simple expression in terms of η -quotients.¹⁰ However, the argument based on the transformation properties of the curve (4.77) is completely general; thus, also in this case the S-duality transformation $\tau_1 \rightarrow -1/\tau_1$ should imply, for consistency, an inversion of q_0 . We will solve this problem in the following subsections, and in doing so we will actually find a new way of writing a non-perturbatively exact relation between the bare coupling and the effective ones. This will turn out to be valid not only in all arithmetic cases, including the SU(6) theory mentioned above, but also in the non-arithmetic theories, thus opening the way to make further progress. Before doing this, however, we briefly revisit the SU(2) theory in order to clarify some issues that are specific to the N = 2 case.

4.5.1 The S-duality Group of Conformal SU(2) Gauge Theory

In the SU(2) gauge theory with four fundamental flavours there is only one renormalized coupling constant τ_1 , which is related to the bare coupling constant by the non-perturbative relation (4.81) with $\lambda_1 = 4$. This might seem unfamiliar, given that it was already proven in [2] that the S-duality group for this theory is the full modular group PSL(2, Z). We now explain how this enhancement takes place within the formalism of our paper.

Let us rewrite the non-perturbative relation between the bare coupling and the renormalized coupling using the standard Jacobi θ -functions as follows:

$$q_0 = -\left(\frac{\theta_2(2\tau_1)}{\theta_4(2\tau_1)}\right)^4 .$$
 (4.82)

¹⁰Note that in this case we have $\lambda_1 = 1$ which cannot be used in (4.81).

One can check that this coincides with the η -quotient expression in (4.49). We have already seen that the S-transformation acts as follows on the renormalized coupling τ_1 :

$$S : \tau_1 \to -\frac{1}{4\tau_1}$$
 (4.83)

The key point is that only for the conformal SU(2) theory, there is a shift symmetry of the form

$$T : \tau_1 \to \tau_1 + \frac{1}{2}$$
 (4.84)

This is because, in the presence of massless hypermultiplets in the doublet pseudoreal representation, the SU(2) theory enjoys a shift symmetry of the effective θ -angle:

$$\theta \to \theta + \pi$$
, (4.85)

which implies (4.84) (see for example appendix B.3 of [55]). Defining $\tilde{\tau} = 2\tau_1$, we see that (4.83) and the above shift become, respectively, $\tilde{\tau} \to -1/\tilde{\tau}$ and $\tilde{\tau} \to \tilde{\tau} + 1$, which generate the modular group $PSL(2,\mathbb{Z})$ in full agreement with the original analysis of [2].

Using the non-perturbative relation (4.82), the *T*-transformation (4.84) leads to the following action on the bare coupling constant:

$$T : q_0 \to \frac{q_0}{q_0 - 1}$$
 (4.86)

Note that this symmetry transformation exists only for the conformal SU(2) gauge theory because in all other cases the *T*-action leaves the bare coupling invariant, since it shifts τ by an integer. Combined with the *S*-transformation, which inverts q_0 , one can check that

$$TST : q_0 \to 1 - q_0 .$$
 (4.87)

We now show that this is completely consistent with the Gaiotto formulation of the S-duality group on the bare coupling constant. The Gaiotto curve for the SU(2) case is [22]:

$$x^{2} = \frac{u_{2}}{t(t-1)(t-q_{0})} .$$
(4.88)

In this expression there is a symmetry between the poles at t = 0 and t = 1.¹¹ Thus, besides (4.78), there is another transformation which leaves the curve invariant, namely [22]

$$\widetilde{T}$$
 : $(x, t) \to (x, 1-t)$. (4.89)

It is easy to check that \widetilde{T} precisely generates the transformation (4.87). Therefore, in the conformal SU(2) theory the S-duality group is enhanced to the full modular group PSL(2,Z), on account of the half-integer shift of the τ -parameter.

4.5.2 Renormalized Couplings and *j*-Invariants

Let us now return to the issue of finding a non-perturbative relation between the renormalized coupling τ_1 of the conformal SU(6) theory and the bare coupling constant. The new and key ingredient is the Klein *j*- invariant function $j(\tau_1)$ for the modular group PSL(2,Z) which is the S-duality group for the τ_1 coupling of the SU(6) theory. The *j*-invariant has the following weak-coupling expansion

$$j(\tau_1) = \frac{1}{q_1} + 744 + 196844 \, q_1 + 21493760 \, q_1^2 + \cdots , \qquad (4.90)$$

with $q_1 = e^{2\pi i \tau_1}$, and is such that

$$j(i) = 1728$$
, $j(e^{\frac{2\pi i}{3}}) = 0$ and $j(i\infty) = \infty$. (4.91)

¹¹For generic N, there is a higher order pole at t = 0.

The *j*-invariant is also called *hauptmodul* (see for example [94]), and is such that all rational functions of j are modular.

Using (4.90), it is possible to verify that

$$\frac{\sqrt{j(\tau_1) - 1728} - \sqrt{j(\tau_1)}}{\sqrt{j(\tau_1) - 1728} + \sqrt{j(\tau_1)}} = -432 q_1 \left(1 + 120q_1 + 4140q_1^2 + \cdots\right)$$
(4.92)

which is precisely the same expansion appearing in the first line of (4.71) that was obtained by inverting the instanton series. Based on this evidence, we propose that the exact relation between the bare coupling q_0 and the renormalized coupling τ_1 is

$$q_0 = \frac{\sqrt{j(\tau_1) - 1728} - \sqrt{j(\tau_1)}}{\sqrt{j(\tau_1) - 1728} + \sqrt{j(\tau_1)}} .$$
(4.93)

Further evidence in support of this proposal is its behaviour under $\tau_1 \rightarrow -1/\tau_1$. This is derived from the monodromy of j around the fixed point of this action, *i.e.* $\tau_1 = i$, namely

$$(j(\tau_1) - 1728) \rightarrow e^{2\pi i} (j(\tau_1) - 1728)$$
, (4.94)

which implies the inversion of q_0 as it should be.

This approach is easily generalized, since hauptmoduln have been studied for the duality groups $\Gamma^*(\lambda_k)$ of the arithmetic theories.¹² Indeed, following [95] for $\lambda_k \in$

¹²Recall that the duality group is generated by S and T as defined in (4.76).

 $\{1,2,3,4\}$ we introduce the functions j_{λ_k} given by

$$j_1(\tau) = \left(\frac{E_4(\tau)}{\eta^8(\tau)}\right)^3 , \qquad (4.95a)$$

$$j_2(\tau) = \left[\left(\frac{\eta(\tau)}{\eta(2\tau)} \right)^{12} + 64 \left(\frac{\eta(2\tau)}{\eta(\tau)} \right)^{12} \right]^2 , \qquad (4.95b)$$

$$j_3(\tau) = \left[\left(\frac{\eta(\tau)}{\eta(3\tau)} \right)^6 + 27 \left(\frac{\eta(3\tau)}{\eta(\tau)} \right)^6 \right]^2 , \qquad (4.95c)$$

$$j_4(\tau) = \left[\left(\frac{\eta(\tau)}{\eta(4\tau)} \right)^4 + 16 \left(\frac{\eta(4\tau)}{\eta(\tau)} \right)^4 \right]^2 .$$
(4.95d)

where in the first line E_4 is the Eisenstein series of weight 4. It is possible to check that j_1 coincides with the *j*-invariant introduced above, while j_2 , j_3 and j_4 are generalizations thereof.¹³ Notice that in (4.95b)–(4.95d) we find precisely the η -quotients appearing in the relations between the bare coupling q_0 and the renormalized couplings τ_k . Solving for these quotients in terms of the j_{λ_k} 's and inserting the result in (4.81), we obtain

$$q_{0} = \frac{\sqrt{j_{\lambda_{k}}(\tau_{k}) - d_{\lambda_{k}}^{-1}} - \sqrt{j_{\lambda_{k}}(\tau_{k})}}{\sqrt{j_{\lambda_{k}}(\tau_{k}) - d_{\lambda_{k}}^{-1}} + \sqrt{j_{\lambda_{k}}(\tau_{k})}}$$
(4.96)

where

$$d_2^{-1} = 256$$
, $d_3^{-1} = 108$, $d_4^{-1} = 64$. (4.97)

Eq. (4.96) has the same structure as (4.93); however, this is more than a formal analogy. On consulting Fig. 1, one sees that the location of the corners of the fundamental domain — which are the fixed points of the S, ST^{-1} , and T transformations — are given by

$$\tau_k^A = \frac{\mathrm{i}}{\sqrt{\lambda_k}} , \quad \tau_k^B = \frac{1}{2} + \frac{\mathrm{i}}{2} \sqrt{\frac{4 - \lambda_k}{\lambda_k}} , \quad \tau_k^C = \mathrm{i}\infty , \qquad (4.98)$$

¹³Our definition of the *j*-invariants differ from those in [95] by a constant term, which does not affect its invariance under the duality group.



Figure 4.1: The fundamental domain F' of $\Gamma^*(\lambda_k)$. The point τ_k^A is the fixed point of the S, τ_k^B is the fixed point of ST^{-1} , while τ_k^C is the fixed point of T.

respectively. Furthermore, one can show that [95]

$$j_{\lambda_k}(\tau_k^A) = d_{\lambda_k}^{-1} , \quad j_{\lambda_k}(\tau_k^B) = 0 , \quad \text{and} \quad j_{\lambda_k}(\tau_k^C) = \infty , \qquad (4.99)$$

which is a direct generalization of (4.91), while from the monodromy of j_{λ_k} around the fixed points of S, namely

$$\left(j(\tau_{\lambda_k}) - d_{\lambda_k}^{-1}\right) \to e^{2\pi i} \left(j(\tau_{\lambda_k}) - d_{\lambda_k}^{-1}\right), \qquad (4.100)$$

one easily deduces from (4.96) that q_0 gets inverted under S-duality, as expected.

In Tab. 1 we collect the relevant properties of these *j*-invariants together with their expansions around the cusp point at infinity. In particular we observe in the last column that the weak-coupling expansions of the bare coupling q_0 are in perfect agreement with the results presented in Section 4.4.

λ_k	$d_{\lambda_k}^{-1}$	q-expansion of j_{λ_k}	$4d_{\lambda_k} q_0$
1	1728	$q^{-1} + 744 + 196884 q + 21493760 q^2 + \cdots$	$q(1+120 q+4140 q^2+\cdots)$
2	256	$q^{-1} + 104 + 4372 q + 96256 q^2 + \cdots$	$q(1+24 q+300 q^2+\cdots)$
3	108	$q^{-1} + 42 + 783 q + 8672 q^2 + \cdots$	$q(1+12q+90q^2+\cdots)$
4	64	$q^{-1} + 24 + 276 q + 2048 q^2 + \cdots$	$q(1+8q+44q^2+\cdots)$

Table 4.1: Relevant parameters for the j_{λ_k} functions, their *q*-expansions, and the weak-coupling expansion of the bare coupling q_0 defined in (4.96).

4.6

SU(N) Theories and Triangle Groups

We now proceed to generalize the discussion of the previous sections to SU(N)SQCD theories with arbitrary N. To this end, we note that for the arithmetic cases $-\lambda_k \in \{1, 2, 3, 4\}$ — the S-duality groups $\Gamma^*(\lambda_k)$ are particular instances of Hecke groups. A Hecke group H(p) is a discrete subgroup of $PSL(2,\mathbb{R})$ whose generators T and S satisfy

$$S^2 = 1$$
, $(ST)^p = 1$ (4.101)

where p is an integer $\geq 3.^{14}$ When p = 3 the Hecke group is the modular group $PSL(2,\mathbb{Z})$.

Using the results of Section 4.4, it is not difficult to realize that $\Gamma^*(\lambda_k) = H(p_k)$ where

$$\lambda_k = 4\cos^2\left(\frac{\pi}{p_k}\right) \ . \tag{4.102}$$

For the four arithmetic cases the correspondence between λ_k and p_k is summarized

$$S : \quad \widetilde{\tau} \rightarrow -\frac{1}{\widetilde{\tau}} \quad \text{and} \quad T : \widetilde{\tau} \rightarrow \widetilde{\tau} + 2\cos\left(\frac{\pi}{p}\right)$$

By setting $\tilde{\tau} = 2\cos\left(\frac{\pi}{p}\right)\tau$, we see that on τ the group H(p) coincides with $\Gamma^*(\lambda)$ with $\lambda = 4\cos^2(\frac{\pi}{p})$.

¹⁴The constraints (4.101) are usually implemented by

in Tab. 2.

λ_k	1	2	3	4
p_k	3	4	6	∞

Table 4.2: The correspondence between λ_k and p_k according to (4.102) in the arithmetic cases.

Notice that these are the only cases in which both λ_k and p_k are integers. By combining (4.29) and (4.102), we find

$$\frac{1}{p_k} = \frac{1}{2} - \frac{k}{N} \tag{4.103}$$

for $k = 1, ..., \left[\frac{N}{2}\right]$. This formula can be formally extended beyond the arithmetic cases where, in general, p_k becomes a rational number.

The Hecke groups H(p) also exist when $p \notin \{3, 4, 6, \infty\}$; moreover they admit a generalization into the so-called triangle groups [28] which we conjecture can be further extended for rational p. In the following we show that the action of the S-duality group on the renormalized couplings of the SU(N) SQCD theories for arbitrary N is precisely that of a generalized triangle group. Furthermore, we show that the *j*-invariant or hauptmodul associated to these triangle groups appears in the non-perturbative relation between the bare coupling and the renormalized ones, exactly as in the arithmetic cases.

4.6.1 A Short Digression on Triangle Groups

We follow closely the presentation of [28], often considering special cases of their formulas for our purposes.

Triangle groups are defined by a triple of integer numbers m_i that form the so-called

type $\mathbf{t} = (m_1, m_2, m_3)$ and correspond to the orders of the stabilizers. These groups are Fuchsian, *i.e.* they are discrete subgroups of $PSL(2, \mathbb{R})$. The type \mathbf{t} defines a set of angular parameters $v_i = 1/m_i$, which are related to deficit angles πv_i at the cusps of the corresponding fundamental domain. In what follows, we will analyze in particular types of the form $\mathbf{t} = (2, p, \infty)$ corresponding to the Hecke groups H(p)if p is an integer, and to their generalizations if p is a rational number. In the latter case the associated triangle groups are not discrete.

Let us first consider a type $\mathbf{t} = (m_1, m_2, \infty)$. Using Theorem 1 of [28], we define the parameter $d_{\mathbf{t}}$ according to

$$d_{\mathbf{t}}^{-1} = b'd' \prod_{k=1}^{b'-1} \left(2 - 2\cos\left(2\pi\frac{k}{b'}\right)\right)^{-\frac{1}{2}\cos\left(\frac{2\pi ka'}{b'}\right)} \prod_{\ell=1}^{d'-1} \left(2 - 2\cos\left(\frac{2\pi\ell}{d'}\right)\right)^{-\frac{1}{2}\cos\left(2\pi\frac{kc'}{d'}\right)}$$
(4.104)

where the primed variables are given by

$$\frac{a'}{b'} = \frac{1+v_1-v_2}{2} \quad \text{and} \quad \frac{c'}{d'} = \frac{1+v_1+v_2}{2} \tag{4.105}$$

with $v_i = 1/m_i$. Introducing the rescaled variable

$$\widetilde{q} = \frac{q}{d_{\mathbf{t}}} \tag{4.106}$$

with $q = e^{2\pi i \tau}$, the hauptmodul J_t for this triangle group has a weak-coupling expansion in \tilde{q} of the form

$$J_{\mathbf{t}}(\tau) = \frac{1}{\widetilde{q}} + \sum_{k=0}^{\infty} c_k \, \widetilde{q}^k \,. \tag{4.107}$$

The coefficients c_k are uniquely determined by the following Schwarzian equation

$$-2\ddot{J}_{\mathbf{t}}\dot{J}_{\mathbf{t}} + 3\ddot{J}_{\mathbf{t}}^{2} = \dot{J}_{\mathbf{t}}^{4} \left(\frac{1 - v_{2}^{2}}{J_{\mathbf{t}}^{2}} + \frac{1 - v_{1}^{2}}{(J_{\mathbf{t}} - 1)^{2}} + \frac{v_{1}^{2} + v_{2}^{2} - 1}{J_{\mathbf{t}}(J_{\mathbf{t}} - 1)} \right) .$$
(4.108)

Here the dots denote the logarithmic τ -derivatives. The hauptmodul that will be relevant for us is the one whose weak-coupling expansion begins with q^{-1} . This is simply obtained by rescaling J_t according to

$$j_{\mathbf{t}}(\tau) = \frac{J_{\mathbf{t}}(\tau)}{d_{\mathbf{t}}} . \tag{4.109}$$

Let us check these formulas for $\mathbf{t} = (2, 3, \infty)$ which corresponds to $H(3) = PSL(2, \mathbb{Z})$. When p = 3 the corresponding λ is 1 as we see from Tab. 2, and thus instead of the subscript $_{\mathbf{t}}$ we can use the subscript $_{1}$ in all relevant quantities. In this case we have

$$v_1 = \frac{1}{2}$$
 and $v_2 = \frac{1}{3}$, (4.110)

and

$$\frac{a'}{b'} = \frac{7}{12} \qquad \frac{c'}{d'} = \frac{11}{12}$$
 (4.111)

Substituting this into (4.104), we find

$$d_1^{-1} = 1728 , \qquad (4.112)$$

while the Schwarzian equation (4.108) becomes

$$-2\ddot{J}_{1}\dot{J}_{1} + 3\ddot{J}_{1}^{2} = \dot{J}_{1}^{4} \left(\frac{32 - 41J_{1} + 36J_{1}^{2}}{36J_{1}^{2}(J_{1}^{2} - 1)}\right) .$$
(4.113)

Solving for J_1 and rescaling the solution with d_1 according to (4.109), one gets

$$j_1(\tau) = 1728 J_1(\tau) = \frac{1}{q} + 744 + 196884 q + \cdots$$
 (4.114)

which exactly matches the expansion of the absolute j-invariant of the modular group (see (4.90)).

In a similar way one can check that for $\mathbf{t} = (2, p, \infty)$ with $p \in \{4, 6, \infty\}$, the above

formulas correctly lead to the expressions of the *j*-invariants and the *d* parameters of the other arithmetic cases that are summarized in Tab. 1. However, as we have already mentioned, these same formulas can be used also for other integer values of p and formally extended to the case in which p is a rational number. As a first example of this extension we consider the conformal SU(5) SQCD theory.

4.6.2 *N* = 5

Using localization techniques we have computed the prepotential and the period matrix of the SU(5) theory with 10 massless flavours up to 2 instantons. In the special vacuum we find that the period matrix Ω can be conveniently written as a sum of two independent structures, in agreement with the general formula (4.27). Defining

$$\lambda_{1} = 4 \sin^{2} \frac{\pi}{5} = 4 \cos^{2} \frac{3\pi}{10} = \frac{\sqrt{5}}{2} \left(\sqrt{5} - 1\right) ,$$

$$\lambda_{2} = 4 \sin^{2} \frac{2\pi}{5} = 4 \cos^{2} \frac{\pi}{10} = \frac{\sqrt{5}}{2} \left(\sqrt{5} + 1\right) ,$$
(4.115)

the quantum corrected period matrix can be written as

$$\Omega = \tau_1 \,\mathcal{M}_1 + \tau_2 \,\mathcal{M}_2 \tag{4.116}$$

where

$$\mathcal{M}_{1} = \begin{pmatrix} \frac{2\lambda_{1}}{5} & \frac{1}{\lambda_{1}} & \frac{\lambda_{1}}{5} & -\frac{\lambda_{1}^{2}}{5\sqrt{5}} \\ \frac{1}{\lambda_{1}} & \frac{2}{\lambda_{1}} & \frac{\sqrt{5}}{\lambda_{1}^{2}} & \frac{\lambda_{1}}{5} \\ \frac{\lambda_{1}}{5} & \frac{\sqrt{5}}{\lambda_{1}^{2}} & \frac{2}{\lambda_{1}} & \frac{1}{\lambda_{1}} \\ -\frac{\lambda_{1}^{2}}{5\sqrt{5}} & \frac{\lambda_{1}}{5} & \frac{1}{\lambda_{1}} & \frac{2\lambda_{1}}{5} \end{pmatrix} , \qquad \mathcal{M}_{2} = \begin{pmatrix} \frac{2}{\lambda_{1}} & \frac{\lambda_{1}}{5} & \frac{1}{\lambda_{1}} & \frac{\sqrt{5}}{\lambda_{1}^{2}} \\ \frac{\lambda_{1}}{5} & \frac{2\lambda_{1}}{5} & -\frac{\lambda_{1}^{2}}{5\sqrt{5}} & \frac{1}{\lambda_{1}} \\ \frac{1}{\lambda_{1}} & -\frac{\lambda_{1}^{2}}{5\sqrt{5}} & \frac{2\lambda_{1}}{5} & \frac{\lambda_{1}}{5} \\ \frac{\sqrt{5}}{\lambda_{1}^{2}} & \frac{1}{\lambda_{1}} & \frac{\lambda_{1}}{5} & \frac{2}{\lambda_{1}} \end{pmatrix} , \quad (4.117)$$

and

$$2\pi i \tau_1 = \log q_0 + i\pi - \log \left[25\sqrt{5} \left(\frac{2}{\sqrt{5} - 1} \right)^{\sqrt{5}} \right] + \frac{8q_0}{25} + \frac{14q_0^2}{125} + \cdots , \qquad (4.118)$$

$$2\pi i \tau_2 = \log q_0 + i\pi - \log \left[25\sqrt{5} \left(\frac{2}{\sqrt{5}+1}\right)^{\sqrt{5}} \right] + \frac{12q_0}{25} + \frac{24q_0^2}{125} + \cdots$$
 (4.119)

This structure is less cumbersome than it appears at first sight. Indeed, one can check that the parameters λ_1 and λ_2 in (4.115) are another instance of the general formula (4.29) and that

$$\mathcal{M}_1 + \mathcal{M}_2 = \mathcal{C} \tag{4.120}$$

where C is the Cartan matrix of SU(5). Moreover, the classical and the logarithmic terms of τ_1 and τ_2 exactly coincide with the results already reported in [90]. But, most importantly, using the S-duality transformations described in Section 4.3, one finds that the matrices \mathcal{M}_1 and \mathcal{M}_2 remain invariant while the effective couplings transform simply as

$$\tau_1 \to -\frac{1}{\lambda_1 \tau_1} \quad \text{and} \quad \tau_2 \to -\frac{1}{\lambda_2 \tau_2}$$
 (4.121)

These observations show that the SU(5) theory has the same general features we encountered in the arithmetic cases. Therefore it is natural to expect that also this theory can be understood along the same lines, and in particular that it is possible to write non-perturbatively exact expressions for the relations between the bare coupling and the renormalized ones in terms of hauptmoduln. We now confirm that these expectations are correct.

Let us first put k = 2. The form of λ_2 in (4.115) indicates to us that the relevant Hecke group is H(10). Indeed, for k = 2 and N = 5, equation (4.103) yields $p_2 = 10$ so that the type of the triangle group is $\mathbf{t}_2 = (2, 10, \infty)$. Using this in (4.104), with a little bit of algebra we obtain

$$d_{\lambda_2}^{-1} = 4 \left[25\sqrt{5} \left(\frac{2}{\sqrt{5}+1} \right)^{\sqrt{5}} \right] = 76.2385 \cdots , \qquad (4.122)$$

while from (4.108) we get the following rescaled hauptmodul

$$j_{\lambda_2}(\tau_2) = \frac{1}{q_2} + \frac{19}{50} \frac{1}{d_{\lambda_2}} + \frac{673}{10000} \frac{q_2}{d_{\lambda_2}^2} + \frac{701}{93750} \frac{q_2^2}{d_{\lambda_2}^3} + \cdots$$
(4.123)

with $q_2 = e^{2\pi i \tau_2}$. This function is such that

$$j_{\lambda_2}(\tau_2^A) = d_{\lambda_2}^{-1}, \quad j_{\lambda_2}(\tau_2^B) = 0 \quad \text{and} \quad j_{\lambda_2}(\tau_2^C) = \infty$$
 (4.124)

where $\tau_2^{A,B,C}$ are the cusp locations in the τ_2 -plane given by (4.98) with the current value of λ_2 . Notice also that the quantity in square brackets in (4.122) also appears in the 1-loop logarithmic term of (4.119).

These facts and our experience with the arithmetic theories indicate that it is in fact not too bold to propose that the relation between the bare coupling q_0 and the renormalized coupling τ_2 be of the general form (4.96), namely

$$q_0 = \frac{\sqrt{j_{\lambda_2}(\tau_2) - d_{\lambda_2}^{-1} - \sqrt{j_{\lambda_2}(\tau_2)}}}{\sqrt{j_{\lambda_2}(\tau_2) - d_{\lambda_2}^{-1} + \sqrt{j_{\lambda_2}(\tau_2)}}} = -\frac{q_2}{4d_{\lambda_2}} \left(1 + \frac{3}{25}\frac{q_2}{d_{\lambda_2}} + \frac{6}{625}\frac{q_2^2}{d_{\lambda_2}^2} + \cdots\right) \quad (4.125)$$

Inverting this series and taking the logarithm, we obtain

$$2\pi i \tau_2 = \log q_0 + i\pi + \log \left(4d_{\lambda_2}\right) + \frac{12}{25}q_0 + \frac{25}{125}q_0^2 + \cdots$$
 (4.126)

which precisely matches the instanton expansion for τ_2 in (4.119) obtained using equivariant localization! Furthermore, from the monodromy around τ_2^A which is the fixed point under S, namely

$$\left(j(\tau_{\lambda_2}) - d_{\lambda_2}^{-1}\right) \to e^{2\pi i} \left(j(\tau_{\lambda_2}) - d_{\lambda_2}^{-1}\right), \qquad (4.127)$$

we see that q_0 gets inverted, in agreement with the general expectations. This analysis shows that the action of the S-duality group on the effective coupling τ_2 of the SU(5) theory is that of the Hecke group H(10).

We now turn to k = 1 and the quantum coupling τ_1 . The form of λ_1 in (4.115) indicates that we are dealing with a non-Hecke group. Indeed, setting k = 1 and N = 5 in (4.103), we get $p_1 = \frac{10}{3}$ which leads to the type $\mathbf{t}_1 = (2, \frac{10}{3}, \infty)$. Despite the non-integer entry of \mathbf{t}_1 , we still proceed and apply the formulas we have described in the previous subsection to obtain d_{λ_1} and the hauptmodul $j_{\lambda_1}(\tau_1)$. Specifically, from (4.104) after some algebraic manipulations we get

$$d_{\lambda_1}^{-1} = 4 \left[25\sqrt{5} \left(\frac{2}{\sqrt{5} - 1} \right)^{\sqrt{5}} \right] = 655.8364 \cdots , \qquad (4.128)$$

while from the Schwarzian equation (4.108) we find the following rescaled hauptmodul

$$j_{\lambda_1}(\tau_1) = \frac{1}{q_1} + \frac{21}{50} \frac{1}{d_{\lambda_1}} + \frac{663}{10000} \frac{q_1}{d_{\lambda_1}^2} + \frac{227}{46875} \frac{q_1^2}{d_{\lambda_1}^3} + \dots$$
(4.129)

with $q_1 = e^{2\pi i \tau_1}$. This function is such that

$$j_{\lambda_1}(\tau_1^A) = d_{\lambda_1}^{-1}, \quad j_{\lambda_1}(\tau_1^B) = 0 \quad \text{and} \quad j_{\lambda_1}(\tau_1^C) = \infty$$
 (4.130)

where $\tau_1^{A,B,C}$ are the three cusps in the τ_1 -plane (see (4.98)). Plugging these results into our universal formula (4.96), we get

$$q_{0} = \frac{\sqrt{j_{\lambda_{1}}(\tau_{1}) - d_{\lambda_{1}}^{-1}} - \sqrt{j_{\lambda_{1}}(\tau_{1})}}{\sqrt{j_{\lambda_{1}}(\tau_{1}) - d_{\lambda_{1}}^{-1}} + \sqrt{j_{\lambda_{1}}(\tau_{1})}} = -\frac{q_{1}}{4d_{\lambda_{1}}} \left(1 + \frac{2}{25}\frac{q_{1}}{d_{\lambda_{1}}} + \frac{13}{5000}\frac{q_{1}^{2}}{d_{\lambda_{1}}^{2}} + \cdots\right).$$

$$(4.131)$$

Inverting this and taking the logarithm of q_1 , we obtain

$$2\pi i \tau_1 = \log q_0 + i\pi + \log \left(4d_{\lambda_1}\right) + \frac{8}{25}q_0 + \frac{14}{125}q_0^2 + \cdots$$
(4.132)

which is in perfect agreement with the explicit result (4.118) derived from localization! Again, from the monodromy around τ_1^A , which is the fixed point of S, we easily see that under S-duality q_0 is correctly mapped into its inverse.

In conclusion, the SU(5) theory has two non-arithmetic couplings which are related to the bare coupling by the same universal formula that holds in the arithmetic theories.

4.6.3 Generalization to Higher *N*

The analysis of the previous subsection can be extended to arbitrary values of N. Even if the algebraic manipulations become more and more involved as N increases, it is possible to prove that the quantum period matrix can always be written as

$$\Omega = \sum_{k=1}^{\left\lfloor \frac{N}{2} \right\rfloor} \tau_k \,\mathcal{M}_k \tag{4.133}$$

where each individual coefficient τ_k transforms under the duality group according to

$$S : \tau_k \to -\frac{1}{\lambda_k \tau_k}$$
 and $T : \tau_k \to \tau_k + 1$ (4.134)

for some positive λ_k .

To show this, let us first consider N to be an odd number. In this case a careful analysis [91] of the duality transformations on the homology cycles of the Seiberg-

Witten curve shows that S and T have to satisfy the constraint

$$(STS^{-1}T)^{N} = \mathbf{1} . (4.135)$$

Given (4.134), it is not difficult to show that

$$STS^{-1}T = \begin{pmatrix} 1 & 1\\ -\lambda_k & 1 - \lambda_k \end{pmatrix} .$$
(4.136)

The N^{th} power of this matrix projectively equals the identity, as required by (4.135), if

$$\lambda_k = 4\sin^2\left(\frac{k\pi}{N}\right) = 4\cos^2\left(\frac{(N-2k)\pi}{2N}\right)$$
(4.137)

or

$$\lambda_k = 4\cos^2\left(\frac{k\,\pi}{N}\right) \ . \tag{4.138}$$

The latter solution, however, leads to an additional constraint of the form $(ST)^N = -1$, which is not found in the explicit realization of the S and T transformations as $\text{Sp}(2N - 2, \mathbb{Z})$ matrices [91]. This leaves us with the solution (4.137) which is precisely the spectrum we conjectured and found to be true in all cases we have considered so far.

The matrices \mathcal{M}_k can be given an explicit expression too. The key ingredient for this is the matrix \mathcal{G} appearing at 1-loop (see (4.26)). Decomposing it into its $\left[\frac{N}{2}\right]$ independent components according to

- -

$$\mathcal{G} = \sum_{k=1}^{\left\lfloor \frac{N}{2} \right\rfloor} \log \sin \left(\frac{k \pi}{N} \right) \mathcal{G}_k , \qquad (4.139)$$

it turns out that the matrices

$$\mathcal{M}_{k} = \sum_{\ell=1}^{\left\lfloor\frac{N}{2}\right\rfloor} \lambda_{k\ell} \,\mathcal{G}_{\ell} = 4 \sum_{\ell=1}^{\left\lfloor\frac{N}{2}\right\rfloor} \sin^{2}\left(\frac{k\,\ell\,\pi}{N}\right) \,\mathcal{G}_{\ell} \tag{4.140}$$

satisfy the required properties. Notice also that these matrices add up to the Cartan matrix:

$$\sum_{k=1}^{\left\lfloor\frac{N}{2}\right\rfloor} \mathcal{M}_k = \mathcal{C} \ . \tag{4.141}$$

We have explicitly checked and verified these statements up to N = 15.

Having the spectrum of the allowed λ_k 's, from the cosine expression in (4.137) we see that the type of the generalized triangle group that we should consider is

$$\mathbf{t}_k = \left(2, \frac{2N}{N-2k}, \infty\right) \tag{4.142}$$

whose second entry is in general a rational number. As we have seen in the SU(5) theory, there are no obstructions in extending the formulas (4.104), (4.107) and (4.109) to types with a rational entry. Thus, proceeding as we described in the previous subsections, we can determine d_{t_k} and the hauptmoduln j_{t_k} corresponding to (4.142) and use the resulting expressions into the universal formula (4.96) to find the exact relation between the bare coupling q_0 and the renormalized one τ_k . If this procedure is correct, inverting this map order by order in q_0 we should retrieve the multi-instant expansion produced by the localization method, exactly as we showed for N = 5. In Appendix E we give some details for the case N = 7, where again we finding perfect agreement. At this point, it should be clear that our procedure works for arbitrary values of N. We regard the complete agreement between these two approaches as a highly non-trivial and quite remarkable check on the consistency of the procedure.

The above results are valid also when N is even. In this case, the spectrum of λ_k

is still given by (4.137) while the matrices \mathcal{G}_k and \mathcal{M}_k are defined by (4.139) and (4.140) with the caveat that for $k = \frac{N}{2}$ one should use

$$\mathcal{G}_{\frac{N}{2}} = \mathcal{C} - \sum_{k=1}^{\frac{N}{2}-1} \mathcal{G}_k \quad \text{and} \quad \mathcal{M}_{\frac{N}{2}} = \mathcal{C} - \sum_{k=1}^{\frac{N}{2}-1} \mathcal{M}_k . \quad (4.143)$$

We have checked this is indeed the case up to N = 14.

4.6.4 Relation to Earlier Work

We now show that our analysis is consistent with earlier discussions of S-duality in conformal SQCD theories and that it extends them in several aspects. Consider the Seiberg-Witten curve (4.30) for the massless case and in the special vacuum:

$$y^{2} = (x^{N} - u_{N})^{2} - h x^{2N} . (4.144)$$

Using our results, we can write the function h in terms of the renormalized couplings as follows:

$$h = \frac{4q_0}{(1+q_0)^2} = \frac{1}{1-d_{\lambda_k} j_{\lambda_k}(\tau_k)} .$$
(4.145)

This shows that for any N the Seiberg-Witten curve can be expressed in terms of j-invariants. Of course, any of the renormalized couplings can be chosen as long as the appropriate j-invariant is used.

Let us now consider the behaviour of h near the cusp points (4.98). Using (4.99), it is easy to find that

$$h(\tau_k) \to \infty \quad \text{near } \tau_k^A ,$$

$$h(\tau_k) \to 1 \quad \text{near } \tau_k^B , \qquad (4.146)$$

$$h(\tau_k) \to 0 \quad \text{near } \tau_k^C .$$

Given the meaning of the fixed points, we conclude that the monodromy of h around

 ∞ , 1 and 0 yields, respectively, the behaviour under *S*, ST^{-1} and *T*. This is precisely what we began with in Section 4.3. There, we obtained the *T* and *S* matrices by associating them with monodromies around the points h = 0 and $h = \infty$, respectively, and by following their effects on the $\hat{\alpha}$ - and $\hat{\beta}$ -cycles of the Seiberg-Witten curve. It is reassuring to rederive this very same result by studying the action of the duality group on the quantum couplings τ_k . This provides additional confirmation for our proposal (4.96).

There are a number of novel elements in our discussion compared with earlier works [90, 93, 91]. To begin with, we note that (4.96) represents a non-perturbatively exact relation between the bare and renormalized coupling constants. As we have shown in a case-by-case study, this completely specifies the manner in which all $\left[\frac{N}{2}\right]$ coupling constants are renormalized for all SU(N) theories in the special vacuum.

Furthermore, we observe that previous investigations have focused on a specific renormalized coupling, which in our notation, is $\tau_{\left[\frac{N}{2}\right]}$. For odd N, the type (4.142) corresponding to $k = \left[\frac{N}{2}\right]$ is $(2, 2N, \infty)$, which identifies the Hecke group H(2N), while for even N, the type becomes $(2, \infty, \infty)$ corresponding to the Hecke group H(∞) which is isomorphic to $\tilde{\Gamma}^0(2)$. Thus, we have successfully reproduced the observations of [96, 97, 91] that these Hecke groups are relevant when considering the duality properties of SU(N) theories. However, as we have tried to emphasize, one does not need to single out any specific quantum coupling τ_k in order to understand the S-duality group. Indeed, one could choose to express q_0 in terms of any of the τ_k 's since the behaviour of the curve near the cusps is universal and independent of this choice.

While this remains true away from the conformal limit (provided the mass deformations are turned on in a controlled manner), we find that expressing the observables in terms of specific effective coupling constants τ_k instead of the bare coupling q_0 expedites the identification of modular structures. In particular, the choice of which effective couplings to consider follows solely from S-duality constraints. This, in turn, makes it possible to resum the non-perturbative data of the gauge theory into modular forms associated to congruence subgroups of the full modular group. This analysis is the subject of Part II.

Massive $\mathcal{N} = 2$ SQCD and Modular Anomaly Equations

In this part, we discuss $\mathcal{N} = 2$ SQCD theories with 2N massive fundamental hypermultiplets in the special vacuum. In order to retain the \mathbb{Z}_N symmetry of the special vacuum, we will consider only mass configurations that preserve this symmetry. Furthermore, we will restrict our attention to the arithmetic theories. The reason for this is just a matter of simplicity. Indeed, as we will see, in the arithmetic theories the S-duality groups $\Gamma^*(\lambda_k)$ contain subgroups that are also congruence subgroups of the modular group $PSL(2,\mathbb{Z})$, so that the analysis of the modular properties of the various observables can be done using standard modular forms, without the need of introducing the more involved theory of automorphic forms. Since the SU(2) and SU(3) SQCD theories have already been considered from this point of view in [24, 23] and in [26] respectively, we will discuss in detail the other two arithmetic cases, namely N = 4 and N = 6, even if many of the subsequent formulas are valid for arbitrary N. While the classical prepotential (4.4) is unaffected by mass deformations, the 1-loop prepotential (4.8) becomes

$$F_{1-\text{loop}} = \sum_{u \neq v=1}^{N} \gamma(A_u - A_v) - \sum_{u=1}^{N} \sum_{f=1}^{2N} \gamma(A_u + m_f) . \qquad (4.147)$$

Expanding for small masses, one obtains an expression in which the 2N fundamental masses appear through the Casimir invariants of the flavour group, namely

$$T_{\ell} = \sum_{f=1}^{2N} (m_f)^{\ell}$$
 (4.148)

for $\ell = 1, ..., 2N$. As we mentioned above, in order not to spoil the \mathbb{Z}_N symmetry of the special vacuum, we turn on only those flavour Casimirs that are \mathbb{Z}_N -symmetric. This can be done by choosing the following mass configuration

$$m_f = \begin{cases} \omega^{f-1} \ m & , \quad f \in \{1, \cdots, N\} \ , \\ \omega^{f-1} \ \widetilde{m} & , \quad f \in \{N+1, \cdots, 2N\} \ , \end{cases}$$
(4.149)

where $\omega = e^{\frac{2\pi i}{N}}$, which in turn implies

$$T_N = N\left(m^N + \widetilde{m}^N\right) \quad \text{and} \quad T_{2N} = N\left(m^{2N} + \widetilde{m}^{2N}\right), \qquad (4.150)$$

with all other T_{ℓ} vanishing. In what follows, by special vacuum we will mean both the restriction (4.12) on the scalar vacuum expectation values and the above choice of masses.

As discussed in Section 4.2.3, the \mathbb{Z}_N -invariance of the special vacuum implies that the prepotential has a semi-classical expansion of the form (4.15), but now the coefficients f_n depend also on the mass invariants (4.150), namely

$$F = \sum_{n} \frac{f_n(q_0; T_N, T_{2N})}{a^{Nn}} .$$
(4.151)

The f_n 's must have mass-dimension equal to (nN+2), but since q_0 is dimensionless and T_N and T_{2N} have dimensions N and 2N respectively, it is not possible to satisfy this requirement. As a result, in the massive case as well, the special vacuum prepotential vanishes identically.

Let us now turn to the dual period a^{D} . When the 1-loop and instanton corrections are taken into account, we find

$$a^{\rm D} = c_N \, a \, \tau_1 + \frac{c_N}{2\pi \mathrm{i}} \, \sum_{n=0}^{\infty} \frac{g_n^{(1)}(\tau_1; T_N, T_{2N})}{a^{Nn+N-1}}$$
(4.152)

where c_N is defined in (4.19). This form, which will be confirmed by the explicit examples worked out in the later sections, can be argued simply using dimensional analysis because $g_n^{(1)}$ has mass dimension (Nn + N) and can be constructed out of the \mathbb{Z}_N -invariant Casimirs T_N and T_{2N} .

Finally, we consider the period matrix Ω . Its decomposition in terms of the matrices \mathcal{M}_k that diagonalize the S-action remains valid

$$\Omega = \widetilde{\tau}_1 \,\mathcal{M}_1 + \widetilde{\tau}_2 \,\mathcal{M}_2 + \cdots , \qquad (4.153)$$

but now the coefficients acquire terms proportional to the flavour Casimirs. In particular one finds

$$\widetilde{\tau}_k = \tau_k - \frac{1}{2\pi i} \sum_{n=0}^{\infty} (Nn + N - 1) \frac{g_n^{(k)}(\tau_k; T_N, T_{2N})}{a^{Nn+N}}$$
(4.154)

for $k = 1, \dots, \left\lfloor \frac{N}{2} \right\rfloor$. Detailed examples will be given in the following sections.
4.8 S-Duality in Massive SQCD

To see the implications of S-duality in massive SQCD theories, we use the same approach described in [26] for the SU(3) theory and introduce the following combination

$$X := a^{\mathrm{D}} - c_{N} a \tau_{1}$$

$$= \frac{c_{N}}{2\pi \mathrm{i}} \sum_{n} \frac{g_{n}}{a^{Nn+N-1}}$$
(4.155)

where $g_n \equiv g_n^{(1)}(\tau_1; T_N, T_{2N})$. We now perform an S-duality transformation on the first line of (4.155) and use (4.43), (4.45) and (4.46); after some simple algebra we get

$$S(X) = \frac{1}{c_N \,\omega^2 \,\tau_1} X \,. \tag{4.156}$$

On the other hand, applying S-duality to the second line of (4.155) we get

$$S(X) = \frac{c_N}{2\pi i} \sum_n \frac{S(g_n)}{(-\omega \, a^{\mathrm{D}})^{Nn+N-1}} \,.$$
(4.157)

If we now substitute the expression (4.152) for a^{D} and equate the two different expressions for S(X) order by order in the large-*a* expansion, we can deduce how the coefficients g_n transform under *S*. From the leading term, we simply find

$$S(g_0) = \left(i\sqrt{\lambda_1}\tau_1\right)^{N-2}g_0 , \qquad (4.158)$$

where λ_1 is as in (4.46). For the higher order terms, however, we find non-linear contributions that lead to a recursion relation

$$S(g_n) = (-1)^n \left(i\sqrt{\lambda_1}\tau_1 \right)^{Nn+N-2} \left(g_n + \frac{1}{2\pi i\tau_1} \sum_m (Nm+N-1)g_m g_{n-m-1} + \cdots \right).$$
(4.159)

The summand on the right hand side is symmetric under $m \to (n - m - 1)$, and thus $S(g_n)$ can be more conveniently written as

$$S(g_n) = (-1)^n \left(i\sqrt{\lambda_1}\tau_1 \right)^{Nn+N-2} \left(g_n + \frac{(Nn+N-2)}{4\pi i\tau_1} \sum_m g_m g_{n-m-1} + \cdots \right).$$
(4.160)

The presence of the $(-1)^n$ factor suggest to us that the notion of S-parity or charge under S-duality will be a useful one. We define it to be (+1) when n is even and (-1) when n is odd.

So far N has been generic, but to make further progress from now on we will restrict our attention to the arithmetic cases for which λ_1 is an integer. In fact, in these cases the S- duality group $\Gamma^*(\lambda_1)$ contains a subgroup, denoted as $\Gamma_1(\lambda_1)$, which is also a congruence subgroup of PSL(2,Z). The modular forms of such a subgroup, which are well-known and classified (see for instance [98, 99]), will play a crucial role in our analysis and will appear in the exact expressions of the coefficients g_n . To see this, let us first recall that $\Gamma_1(\lambda_1)$ is generated by T and $S' = STS^{-1}$, the latter acting on the effective coupling as

$$S' : \tau_1 \to \frac{\tau_1}{1 - \lambda_1 \tau_1}$$
 (4.161)

When λ_1 is an integer, this is indeed an element of $PSL(2,\mathbb{Z})$. Combining the actions of S and T, we can easily deduce how the conjugate periods a and a^{D} transform under S'. The result is

$$S'(a^{\rm D}) = a^{\rm D}$$
 and $S'(a) = a + \omega(1 - \omega)a^{\rm D}$. (4.162)

Using these rules on X, from the first line of (4.155) we get

$$S'(X) = \frac{1}{1 - \lambda_1 \tau_1} X , \qquad (4.163)$$

while from the second line of (4.155) we find

$$S'(X) = \frac{c_N}{2\pi i} \sum_n \frac{S'(g_n)}{((1-\lambda_1\tau_1)a)^{Nn+N-1}} \left(1 + \frac{c_N}{2\pi i(1-\lambda_1\tau_1)} \sum_m \frac{g_m}{a^{Nm+N-2}}\right)^{Nn+N-1}.$$
(4.164)

Equating these two expressions, to leading order we obtain

$$S'(g_0) = (1 - \lambda_1 \tau_1)^{N-2} g_0 , \qquad (4.165)$$

while at higher orders we get a recursion relation very similar to the one obtained before for S, namely

$$S'(g_n) = (1 - \lambda_1 \tau_1)^{Nn+N-2} \left(g_n + \frac{(Nn+N-2)}{4\pi i \tau_1} \sum_m g_m g_{n-m-1} + \cdots \right) . \quad (4.166)$$

Eq. (4.165) shows that g_0 is a modular form of $\Gamma_1(\lambda_1)$ with weight (N-2). As we will see in the specific examples in the next section, such a modular form behaves under S exactly as required by (4.158), thus proving the consistency of our analysis. On the other hand, the presence of non-linear terms in the right hand side of (4.166) implies that the coefficients g_n for n > 0 are quasi-modular forms of $\Gamma_1(\lambda_1)$ with weight (Nn + N - 2) that satisfy a modular anomaly equation to which we now turn.

4.8.1 The Modular Anomaly Equation

In [24, 23, 26] it has been shown that in the massive SU(2) and SU(3) theories the quasi-modularity is due to the presence of the anomalous Eisenstein series E_2 . The same conclusion has been reached for the $\mathcal{N} = 2^*$ theories with arbitrary gauge groups in [25, 11, 12]. Therefore it is very natural to expect that for the massive higher rank SQCD theories too, the Eisenstein series E_2 plays a fundamental role.

Let us recall that E_2 is a quasi-modular form of weight 2 such that

$$E_2\left(-\frac{1}{\tau_1}\right) = -\left(i\tau_1\right)^2 \left(E_2(\tau_1) + \frac{6}{i\pi\tau_1}\right) \,. \tag{4.167}$$

In the arithmetic cases under consideration, it is always possible to form a linear combination of E_2 and a modular form of $\Gamma_1(\lambda_1)$, which under the *S* transformation $\tau_1 \rightarrow -\frac{1}{\lambda_1 \tau_1}$ transforms in a way similar to (4.167). More precisely, if we denote such a combination by $\widetilde{E}_2^{(\lambda_1)}$, we will have

$$\widetilde{E}_{2}^{(\lambda_{1})}\left(-\frac{1}{\lambda_{1}\tau_{1}}\right) = -\left(\mathrm{i}\sqrt{\lambda_{1}}\tau_{1}\right)^{2}\left(\widetilde{E}_{2}^{(\lambda_{1})}(\tau_{1}) + \frac{6}{\mathrm{i}\pi\tau_{1}}\right) .$$

$$(4.168)$$

Notice that the existence of such a combination is a priori not obvious since the S-transformation lies outside both the modular group and its congruence subgroup $\Gamma_1(\lambda_1)$. Nevertheless this combination exists and the explicit examples for the relevant cases are given in Appendix D (see in particular (D.15), (D.24) and (D.30)).

Following [24, 23, 26] we propose that the coefficients g_n depend on τ_1 only through $\widetilde{E}_2^{(\lambda_1)}$ and the modular forms of $\Gamma_1(\lambda_1)$, in such a way that they are globally quasimodular forms of $\Gamma_1(\lambda_1)$ with total weight (Nn + N - 2). For simplicity, in the following we will only exhibit the dependence on $\widetilde{E}_2^{(\lambda_1)}$ and just write $g_n[\widetilde{E}_2^{(\lambda_1)}]$. Then, applying S-duality, we have

$$S\left(g_{n}\left[\widetilde{E}_{2}^{(\lambda_{1})}\right]\right) = (-1)^{n}(\mathrm{i}\sqrt{\lambda_{1}}\tau_{1})^{Nn+N-2} g_{n}\left[\widetilde{E}_{2}^{(\lambda_{1})} + \frac{6}{\mathrm{i}\pi\tau_{1}}\right]$$
$$= (-1)^{n}(\mathrm{i}\sqrt{\lambda_{1}}\tau_{1})^{Nn+N-2} \left(g_{n}\left[\widetilde{E}_{2}^{(\lambda_{1})}\right] + \frac{6}{\mathrm{i}\pi\tau_{1}}\frac{\partial g_{n}}{\partial\widetilde{E}_{2}^{(\lambda_{1})}} + \cdots\right)$$
(4.169)

where the second line follows upon expanding for large τ_1 . Comparing with (4.160) we obtain the modular anomaly equation

$$\frac{\partial g_n}{\partial \widetilde{E}_2^{(\lambda_1)}} = \frac{Nn + N - 2}{24} \sum_{m=0}^{n-1} g_m g_{n-m-1}$$
(4.170)

which has the form of a recursion relation. Indeed, given the initial condition that specifies g_0 as a modular form, the \tilde{E}_2 -dependent part of g_1 can be unambiguously obtained by integrating the modular anomaly equation. This leaves room for a truly modular piece, which can be fixed by comparing with the explicit instanton expansion obtained using localization. Once g_1 is fully fixed, we can use it in (4.170) to find g_2 , and recursively proceed in this way for the higher g_n 's. This approach has been successfully applied to the SU(3) theory in [26]. In the next sections we complete the analysis for the SU(4) and SU(6) theories.

4.8.2 Coupled Modular Anomaly Equations

We now consider the period matrix Ω . As we mentioned in Section 4.7, after including the quantum corrections it can be decomposed as in (4.153) where, under S-duality, the flavour deformed couplings $\tilde{\tau}_k$ transform as

$$S : \tilde{\tau}_k \to -\frac{1}{\lambda_k \tilde{\tau}_k}$$
 (4.171)

The fact that $\tilde{\tau}_k$ behave like τ_k is a simple consequence of the algebraic properties of the matrices \mathcal{M}_k . Applying S-duality to both sides of (4.154), we get

$$-\frac{1}{\lambda_k \tau_k} \left(1 - \frac{1}{2\pi i \tau_k} \sum_{m=1}^{\infty} \frac{Nm + N - 1}{a^{Nm}} g_{m-1}^{(k)} \right)^{-1} = -\frac{1}{\lambda_k \tau_k} - \frac{1}{2\pi i} \sum_{n=1}^{\infty} \frac{Nn + N - 1}{(-\omega a^{\mathrm{D}})^{Nn}} S(g_{n-1}^{(k)})$$
(4.172)

which, after inserting the semi-classical expansion (4.152) for the dual period, yields the S-duality transformation rules for the coefficients $g_n^{(k)}$. In particular, at leading order we find

$$S(g_0^{(k)}) = \frac{\left(i\sqrt{\lambda_1}\tau_1\right)^N}{\left(i\sqrt{\lambda_k}\tau_k\right)^2} g_0^{(k)} , \qquad (4.173)$$

while at higher orders we get

$$S(g_{n}^{(k)}) = \frac{(-1)^{n} (i\sqrt{\lambda_{1}}\tau_{1})^{Nn+N}}{(i\sqrt{\lambda_{k}}\tau_{k})^{2}} \left[g_{n}^{(k)} + \frac{1}{2\pi i\tau_{k}} \sum_{m} \left(\frac{(Nm+N-1)(N(n-m)-1)}{Nn+N-1} g_{m}^{(k)} g_{n-m-1}^{(k)} \right) + \frac{1}{2\pi i\tau_{1}} \sum_{m} \left(\frac{(Nm+N)(Nm+N-1)}{Nn+N-1} g_{m}^{(k)} g_{n-m-1} \right) + \cdots \right].$$

$$(4.174)$$

When k = 1, both (4.173) and (4.174) reduce to (4.158) and (4.160), respectively. This is a simple but important consistency check of our analysis.

We now perform a similar analysis for the S' transformation under which each effective coupling $\tilde{\tau}_k$ changes as

$$\widetilde{\tau}_k \to \frac{\widetilde{\tau}_k}{1 - \lambda_k \, \widetilde{\tau}_k} \,.$$
(4.175)

Since in the arithmetic theories the λ_k 's are integers, this is a $PSL(2,\mathbb{Z})$ transformation. Using the general technique of comparing coefficients in the semi-classical expansions, we obtain the following constraint for $g_0^{(k)}$:

$$S'(g_0^{(k)}) = \frac{(1 - \lambda_1 \tau_1)^N}{(1 - \lambda_k \tau_k)^2} g_0^{(k)} , \qquad (4.176)$$

while for the higher coefficients $g_n^{(k)}$ we get

$$S'(g_n^{(k)}) = \frac{(1 - \lambda_1 \tau_1)^{Nn+N}}{(1 - \lambda_k \tau_k)^2} \left[g_n^{(k)} + \frac{1}{2\pi i \tau_k} \sum_m \left(\frac{(Nm + N - 1)(N(n - m) - 1)}{Nn + N - 1} g_m^{(k)} g_{n-m-1}^{(k)} \right) + \frac{1}{2\pi i \tau_1} \sum_m \left(\frac{(Nm + N)(Nm + N - 1)}{Nn + N - 1} g_m^{(k)} g_{n-m-1} \right) + \cdots \right].$$
(4.177)

Again it is not difficult to check that for k = 1 these two equations reduce respec-

tively to (4.165) and (4.166), as it should be.

From (4.176) combined with (4.173), we can infer that $g_0^{(k)}$ is a ratio of a modular form of $\Gamma_1(\lambda_1)$ with weight N and a modular form of $\Gamma_1(\lambda_k)$ with weight 2. Likewise, by combining (4.177) with (4.174) we deduce that for n > 0 the coefficients $g_n^{(k)}$ are quasi-modular meromorphic forms of $\Gamma_1(\lambda_1)$ and $\Gamma_1(\lambda_k)$ which receive contributions from both $\widetilde{E}_2^{(\lambda_1)}$ and $\widetilde{E}_2^{(\lambda_k)}$. Taking into account the factors multiplying the square brackets in (4.174) and (4.177), we are led to the following ansatz:

$$g_n^{(k)} = \sum_{\ell=0}^n G_n^{nN+N-2\ell;2+2n-2\ell}(\tau_1,\tau_k) \left(\widetilde{E}_2^{(\lambda_1)}(\tau_1)\right)^\ell \left(\widetilde{E}_2^{(\lambda_k)}(\tau_k)\right)^{n-\ell}$$
(4.178)

where the coefficients $G_n^{r_1;r_k}(\tau_1,\tau_k)$ are made of modular forms of $\Gamma_1(\lambda_1)$ and $\Gamma_1(\lambda_k)$ with weights r_1 and r_k respectively. Using the anomalous transformation properties of the second Eisenstein series, from (4.178) we get

$$S(g_n^{(k)}) = \frac{\left(-1\right)^n \left(\mathrm{i}\sqrt{\lambda_1}\tau_1\right)^{Nn+N}}{\left(\mathrm{i}\sqrt{\lambda_k}\tau_k\right)^2} \left(g_n^{(k)} + \frac{6}{\pi\mathrm{i}\tau_1}\frac{\partial g_n^{(k)}}{\partial \widetilde{E}_2^{(\lambda_1)}} + \frac{6}{\pi\mathrm{i}\tau_k}\frac{\partial g_n^{(k)}}{\partial \widetilde{E}_2^{(\lambda_k)}} + \cdots\right),\tag{4.179}$$

and, after comparison with (4.174), we arrive at the following coupled equations

$$\frac{\partial g_n^{(k)}}{\partial \widetilde{E}_2^{(\lambda_k)}} = \frac{1}{12} \sum_{m=0}^{n-1} \frac{(Nm+N-1)(N(n-m)-1)}{Nn+N-1} g_m^{(k)} g_{n-m-1}^{(k)} ,$$

$$\frac{\partial g_n^{(k)}}{\partial \widetilde{E}_2^{(\lambda_1)}} = \frac{1}{12} \sum_{m=0}^{n-1} \frac{(Nm+N)(Nm+N-1)}{Nn+N-1} g_m^{(k)} g_{n-m-1} .$$
(4.180)

In order for these equations to be consistent and integrable, it is necessary that the mixed second derivatives computed from either line of (4.180) match. We find that this is indeed the case, since we have

$$\frac{\partial}{\partial \widetilde{E}_{2}^{(\lambda_{1})}} \left(\frac{\partial g_{n}^{(k)}}{\partial \widetilde{E}_{2}^{(\lambda_{k})}} \right) - \frac{\partial}{\partial \widetilde{E}_{2}^{(\lambda_{k})}} \left(\frac{\partial g_{n}^{(k)}}{\partial \widetilde{E}_{2}^{(\lambda_{1})}} \right) = 0 .$$
(4.181)

Given the structure of the modular anomaly equations (4.180), this is a non-trivial

check which makes it possible to "integrate-in" the quasi-modular terms in a consistent manner.

4.9 Resummation:
$$N = 4$$
 and $N = 6$

In this section we study in detail the SU(4) and SU(6) gauge theories along the lines discussed before. Throughout this section, we use special cases of the formulas derived in the previous section, i.e. setting N = 4 or N = 6 as the case may be.

For the SU(4) theory the relevant parameters are:

$$\omega = i$$
, $c_4 = i - 1$, $k = 1, 2$, $\lambda_1 = 2$, $\lambda_2 = 4$. (4.182)

The Dual Period

We have computed the SU(4) prepotential, the dual periods, and the period matrix up to three instantons using localization methods. From these results, after using the relation (4.63a) to rewrite the instanton counting parameter q_0 in terms of the renormalized coupling q_1 , we find that the dual period can be written as

$$a^{\rm D} = (i-1) a \tau_1 + \frac{(i-1)}{2\pi i} \sum_{n=0}^{\infty} \frac{g_n(q_1; T_4, T_8)}{a^{4n+3}}$$
(4.183)

in agreement with the general form (4.152). The first coefficients g_n are

$$g_{0} = \frac{T_{4}}{12} \left(1 + 24 q_{1} + 24 q_{1}^{2} + 96 q_{1}^{3} + \cdots \right) , \qquad (4.184a)$$

$$g_{1} = \frac{T_{4}^{2}}{4} \left(q_{1} + 26 q_{1}^{2} + 84 q_{1}^{3} + \cdots \right) + \frac{T_{8}}{56} \left(1 - 56 q_{1} - 2296 q_{1}^{2} - 13664 q_{1}^{3} + \cdots \right) , \qquad (4.184b)$$

where, as usual, we have set $q_1 = e^{2\pi i \tau_1}$.

Our goal is to show that these expressions arise from a weak-coupling expansion of quasi-modular forms of $\Gamma_1(2)$. Indeed, according to the discussion of the previous section, we should have

$$S(g_n) = (-1)^n (i\sqrt{2}\tau_1)^{4n+2} [g_n + \cdots] ,$$

$$S'(g_n) = (1 - 2\tau_1)^{4n+2} [g_n + \cdots] .$$
(4.185)

In particular for n = 0 when there are no extra terms beyond leading order, these equations tell us that g_0 should be a modular form of $\Gamma_1(2)$ with weight 2 and Sparity (+1). As shown in Appendix D there is only one such form, namely $f_{2,+}^{(2)}$ whose weak-coupling expansion is

$$f_{2,+}^{(2)} = 1 + 24q_1 + 24q_1^2 + 96q_1^3 + 24q_1^4 + 144q_1^5 \cdots$$
(4.186)

Comparing with (4.184a), we are led to conclude

$$g_0 = \frac{T_4}{12} f_{2,+}^{(2)} , \qquad (4.187)$$

which, to be consistent with (4.185), implies also that T_4 is invariant under both S and S' transformations, namely $S(T_4) = S'(T_4) = T_4$. We would like to stress that once we assume that g_0 is a modular form $\Gamma_1(2)$ of weight 2, the only freedom we have is the overall coefficient which is fixed by matching with the perturbative

contribution. After this is done, *all* non-perturbative terms are fixed by the Fourier expansion of the modular form. The fact that these terms perfectly match the explicit multi-instanton results coming from localization up to three instantons is a very strong and highly non-trivial test of our general strategy.

To obtain the coefficients g_n for n > 0 we can use the recursion relation (4.170), which in the present case is

$$\frac{\partial g_n}{\partial \widetilde{E}_2^{(2)}} = \frac{2n+1}{12} \sum_{m=0}^{n-1} g_m g_{n-m-1}$$
(4.188)

where $\widetilde{E}_{2}^{(2)}$ is the quasi-modular form introduced in Appendix D (see in particular (D.15)). Let us now determine g_1 which according to our general analysis should be a quasi-modular form of of $\Gamma_1(2)$ with weight 6 and with S-parity (-1) that solves the above modular anomaly equation for n = 1, namely

$$\frac{\partial g_1}{\partial \tilde{E}_2^{(2)}} = \frac{1}{4} g_0^2 . aga{4.189}$$

Integrating with respect to $\widetilde{E}_2^{(2)}$ and using the exact expression for g_0 obtained above, we find

$$g_1 = \frac{T_4^2}{576} (f_{2,+}^{(2)})^2 \widetilde{E}_2^{(2)} + \text{modular piece}, \qquad (4.190)$$

where by 'modular piece' we mean a modular form of $\Gamma_1(2)$ with weight 6 and with S-parity (-1). As shown in Appendix D there is only one such form, namely

$$f_{2,+}^{(2)} f_{4,-}^{(2)} = 1 - 56q_1 - 2296q_1^2 - 13664q_1^3 + \cdots$$
 (4.191)

Comparing with the localization result (4.184b), obtain the following exact expression

$$g_1 = \frac{T_4^2}{576} \left[\left(f_{2,+}^{(2)} \right)^2 \widetilde{E}_2^{(2)} - \frac{3}{2} f_{2,+}^{(2)} f_{4,-}^{(2)} \right] + \frac{T_8}{56} f_{2,+}^{(2)} f_{4,-}^{(2)} .$$
(4.192)

As before, all coefficients are fixed by matching with the perturbative terms and following that, all non- perturbative contributions follow from the Fourier expansions of the modular forms. The agreement with the explicit multi-instanton results in (4.184b) is rather remarkable.

The above procedure can be iteratively used to determine the higher coefficients g_n . In this way we have determined up to g_3 , always finding perfect agreement with the localization results.

The Period Matrix

In the special vacuum the period matrix Ω of the massive SU(4) theory can be compactly written as

$$\Omega = \tilde{\tau}_1 \,\mathcal{M}_1 + \tilde{\tau}_2 \,\mathcal{M}_2 \tag{4.193}$$

where the two matrices \mathcal{M}_k are given in (4.61) and

$$\widetilde{\tau}_k = \tau_k - \frac{1}{2\pi i} \sum_{n=0}^{\infty} \frac{4n+3}{a^{4n+4}} \, \widehat{g}_n^{(k)} \left(q_0; T_4, T_8 \right) \,. \tag{4.194}$$

This has the same form as (4.154), except that the coefficients are expressed in terms of the bare coupling q_0 instead of the renormalized ones; this is the meaning of the $\hat{g}_k^{(n)}$ notation. From our explicit calculations, using the non-perturbative relation (4.63a) we find

$$\widehat{g}_n^{(1)}(q_0; T_4, T_8) = g_n(q_1; T_4, T_8) \tag{4.195}$$

where the g_n 's are the same coefficients appearing in the dual period, for which we have already given exact expressions. On the other hand, we find that the first $\hat{g}_n^{(2)}$

coefficients are

$$\widehat{g}_{0}^{(2)} = \frac{T_{4}}{12} \left(1 - \frac{1}{4} q_{0} - \frac{25}{256} q_{0}^{2} - \frac{29}{512} q_{0}^{3} + \cdots \right) , \qquad (4.196a)$$

$$\widehat{g}_{1}^{(2)} = -\frac{T_{4}^{2}}{224} \left(q_{0} + \frac{7}{64} q_{0}^{2} + \frac{7}{512} q_{0}^{3} + \cdots \right) + \frac{T_{8}}{56} \left(1 - q_{0} - \frac{5}{128} q_{0}^{2} - \frac{39}{512} q_{0}^{3} + \cdots \right) . \qquad (4.196b)$$

The challenge is now to show that, once the bare coupling is mapped into the renormalized ones, the resulting expressions $g_n^{(2)}$ have good modular properties. In particular for $g_0^{(2)}$, according to the general analysis of the previous section (see (4.173) and (4.176) for N = 4 and k = 2), we should have

$$S(g_0^{(2)}) = \frac{\left(\sqrt{2}\,\mathrm{i}\,\tau_1\right)^4}{\left(2\,\mathrm{i}\,\tau_2\right)^2} g_0^{(2)} \qquad \text{and} \qquad S'(g_0^{(2)}) = \frac{\left(1-2\tau_1\right)^4}{\left(1-4\tau_2\right)^2} g_0^{(2)} \ . \tag{4.197}$$

These equations tell us that $g_0^{(2)}$ is the ratio of a modular form of $\Gamma_1(2)$ in τ_1 with weight 4 and a modular form of $\Gamma_1(4)$ in τ_2 with weight 2, with total S-parity (+1). From the list of the modular forms presented in Appendices D for $\Gamma_1(2)$ and $\Gamma_1(4)$, we see that the most general ansatz which satisfies these properties is

$$g_0^{(2)} = \frac{T_4}{12} \left[x \, \frac{\left(f_{2,+}^{(2)}\right)^2}{f_{2,+}^{(4)}} + \left(1-x\right) \frac{f_{4,-}^{(2)}}{f_{2,-}^{(4)}} \right] \,, \tag{4.198}$$

where the overall coefficient has been fixed to match with the perturbative result in (4.196a) and x is a free parameter. By Fourier expanding the modular forms and expressing the result in terms of the bare coupling q_0 , one sees that both meromorphic forms within square brackets are identical and both match the q_0 expansion in (4.196a). In the following we choose for simplicity x = 1, so that ¹⁵

$$g_0^{(2)} = \frac{T_4}{12} \frac{\left(f_{2,+}^{(2)}\right)^2}{f_{2,+}^{(4)}} . \tag{4.199}$$

¹⁵We could just as well have picked x = 0; the Fourier expansions do not distinguish between these choices, and it is clear that the modular anomaly equations are not affected by this choice.

For the higher coefficients $g_n^{(2)}$, we have to use the coupled modular anomaly equations (4.180). For n = 1 they become

$$\frac{\partial g_1^{(2)}}{\partial \tilde{E}_2^{(4)}} = \frac{3}{28} \left(g_0^{(2)} \right)^2 \quad \text{and} \quad \frac{\partial g_1^{(2)}}{\partial \tilde{E}_2^{(2)}} = \frac{1}{7} g_0^{(2)} g_0 , \quad (4.200)$$

where $\widetilde{E}_2^{(2)}$ and $\widetilde{E}_2^{(4)}$ are the quasi-modular forms of $\Gamma_1(2)$ and $\Gamma_1(4)$ defined in (D.15) and (D.30) respectively. Integrating (4.200) we find

$$g_1^{(2)} = \frac{3}{28} \left(g_0^{(2)} \right)^2 \widetilde{E}_2^{(4)} + \frac{1}{7} g_0^{(2)} g_0 \, \widetilde{E}_2^{(2)} + \text{modular piece} \,. \tag{4.201}$$

As before, the 'modular piece' is determined by considerations of weight and S-parity and by demanding agreement with the perturbative terms in (4.196b). Explicitly, we find

$$g_{1}^{(2)} = \frac{T_{4}^{2}}{1344} \left(\frac{\left(f_{2,+}^{(2)}\right)^{4} \widetilde{E}_{2}^{(4)}}{\left(f_{2,+}^{(4)}\right)^{2}} + \frac{4}{3} \frac{\left(f_{2,+}^{(2)}\right)^{3} \widetilde{E}_{2}^{(2)}}{f_{2,+}^{(4)}} - \frac{9}{2} \frac{\left(f_{2,+}^{(2)}\right)^{2} f_{4,-}^{(2)}}{f_{2,+}^{(4)}} \right) + \frac{T_{8}}{56} \frac{\left(f_{2,+}^{(2)}\right)^{2} f_{4,-}^{(2)}}{f_{2,+}^{(4)}} .$$

$$(4.202)$$

Once again, the perturbative terms are enough to fix all coefficients that are not determined by the modular anomaly equations; then, the instanton contributions follow by Fourier expanding the modular forms. The perfect agreement with the explicit result (4.196b) obtained from localization confirms in a very non-trivial way the validity of our procedure.

Using this approach iteratively, we have computed higher $g_k^{(2)}$ coefficients, finding complete agreement with the multi-instanton results.

4.9.2 *N* = 6

We now repeat the above analysis for the massive SU(6) theory. In this case the relevant parameters are:

$$\omega = e^{\frac{\pi 1}{3}}, \quad c_6 = -1, \quad k = 1, 2, 3, \quad \lambda_1 = 1, \quad \lambda_2 = 3, \quad \lambda_3 = 4.$$

(4.203)

The Dual Period

The large-a expansion of dual period of the massive SU(6) theory takes the form

$$a^{\rm D} = -a\,\tau_1 - \frac{1}{2\pi {\rm i}} \sum_n \frac{g_n(q_1; T_6, T_{12})}{a^{6n+5}} \ . \tag{4.204}$$

Using localization methods we have computed the coefficients g_n up to two instantons and rewritten them in terms of the effective parameter q_1 by means the relation (4.93). The explicit expressions of the first coefficients are

$$g_0 = \frac{T_6}{5} \left(1 + 240q_1 + 2160q_1^2 + \cdots \right) , \qquad (4.205a)$$

$$g_1 = 12 T_6^2 \left(q_1 + 258q_1^2 + \dots \right) + \frac{T_{12}}{22} \left(1 - 264q_1 - 135432q_1^2 + \dots \right) .$$
 (4.205b)

Since $\lambda_1 = 1$ we expect to resum these expansions into standard modular forms of $PSL(2,\mathbb{Z})$. In particular, from (4.158) and (4.165) we have

$$S(g_0) = (i \tau_1)^4 g_0$$
 and $S'(g_0) = (1 - \tau_1)^4 g_0$, (4.206)

which tell us that g_0 is a modular form of weight 4 with positive S-parity. The unique form of this kind is the Eisenstein series E_4 ; thus, matching the perturbative contribution we find

$$g_0 = \frac{T_6}{5} E_4 \ . \tag{4.207}$$

Again, all instanton terms are dictated by the Fourier expansion of E_4 and are in perfect agreement with the localization result (4.205a). We also verify that g_0 satisfies (4.206) provided that T_6 is invariant under S-duality, namely $S(T_6) = S'(T_6) = T_6$.

To obtain the coefficients g_n with n > 0, we use the modular anomaly equation (4.170) which in this case becomes

$$\frac{\partial g_n}{\partial E_2} = \frac{3n+2}{12} \sum_{m=0}^{n-1} g_m g_{n-m-1} . \qquad (4.208)$$

For example, integrating this equation for n = 1 and fixing the E_2 -independent part by comparing with the perturbative contributions, we get

$$g_1 = \frac{T_6^2}{60} \left(E_4^2 E_2 - E_4 E_6 \right) + \frac{T_{12}}{22} E_4 E_6 . \qquad (4.209)$$

Using the Fourier expansion of the Eisenstein series it is easy to check that the instanton terms precisely match those in (4.205b). Proceeding iteratively in this manner one can derive the exact expressions of the higher coefficients g_n . In particular we have explicitly computed a few higher g_n , always finding perfect agreement with the localization results.

The Period Matrix

In the special vacuum the period matrix Ω of the massive SU(6) theory can be written compactly as

$$\Omega = \widetilde{\tau}_1 \,\mathcal{M}_1 + \widetilde{\tau}_2 \,\mathcal{M}_2 + \widetilde{\tau}_3 \,\mathcal{M}_3 \tag{4.210}$$

where the three matrices \mathcal{M}_k are given in (4.69) while, using a notation similar to that of the SU(4) theory, the three effective couplings turn out to have the following semi-classical expansion

$$\widetilde{\tau}_k = \tau_k - \frac{1}{2\pi i} \sum_{n=0}^{\infty} \frac{6n+5}{a^{6n+6}} \, \widehat{g}_n^{(k)}(q_0; T_6, T_{12}) \,. \tag{4.211}$$

The coefficients $\widehat{g}_n^{(1)}$ coincide with the g_n 's already discussed, while the first coefficients for k = 2 are

$$\widehat{g}_{0}^{(2)} = \frac{T_{6}}{5} \left(1 - \frac{7}{18} q_{0} - \frac{319}{2592} q_{0}^{2} + \cdots \right) , \qquad (4.212a)$$

$$\widehat{g}_{1}^{(2)} = -\frac{7 T_{6}^{2}}{198} \left(q_{0} - \frac{65}{504} q_{0}^{2} + \dots \right) + \frac{T_{12}}{22} \left(1 + \frac{7}{9} q_{0} - \frac{443}{1296} q_{0}^{2} + \dots \right) , \quad (4.212b)$$

and for k = 3 are

$$\widehat{g}_0^{(3)} = \frac{T_6}{5} \left(1 - \frac{1}{3} q_0 - \frac{47}{432} q_0^2 + \cdots \right) , \qquad (4.213a)$$

$$\widehat{g}_{1}^{(3)} = -\frac{5T_{6}^{2}}{132} \left(q_{0} - \frac{23}{360} q_{0}^{2} + \dots \right) + \frac{T_{12}}{22} \left(1 + \frac{5}{6} q_{0} - \frac{227}{864} q_{0}^{2} + \dots \right) .$$
(4.213b)

We now show that these are the first few terms in the semi-classical expansion of rational functions of quasi-modular forms. Since the procedure is similar to that of the SU(4) theory, we will be brief in our discussion.

Let us first consider $g_0^{(2)}$ and $g_0^{(3)}$, whose S and S' transformations are

$$S(g_0^{(2)}) = \frac{(\mathrm{i}\,\tau_1)^6}{(\sqrt{3}\,\mathrm{i}\,\tau_2)^2} g_0^{(2)} , \qquad S'(g_0^{(2)}) = \frac{(1-\tau_1)^6}{(1-3\tau_2)^2} g_0^{(2)} ,$$

$$S(g_0^{(3)}) = \frac{(\mathrm{i}\,\tau_1)^6}{(2\,\mathrm{i}\,\tau_3)^2} g_0^{(3)} , \qquad S'(g_0^{(3)}) = \frac{(1-\tau_1)^6}{(1-4\tau_3)^2} g_0^{(3)} .$$
(4.214)

These formulas suggest that $g_0^{(2)}$ should be expressed as a ratio of a modular form in τ_1 with weight 6 and a modular form of $\Gamma_1(3)$ in τ_2 with weight 2, with an overall S-parity equal to (+1). Likewise, $g_0^{(3)}$ should be expressed as a ratio of a modular form in τ_1 with weight 6 and a modular form of $\Gamma_1(4)$ in τ_3 with weight 2, with an overall S-parity equal to (+1). Using the results collected in Appendix D, matching the weights and S-parities and fixing the overall normalization in agreement with the perturbative contributions, we find that a solution is

$$g_0^{(2)} = \frac{T_6}{5} \frac{f_{2,+}^{(1)} E_4}{\left(f_{1,-}^{(3)}\right)^2} \quad \text{and} \quad g_0^{(3)} = \frac{T_6}{5} \frac{E_6}{f_{2,-}^{(4)}} .$$
 (4.215)

By Fourier expanding the modular forms and expressing the result with bare coupling q_0 , we do not only recover the multi-instanton terms in (4.212a) and (4.213a) but also predict all other higher instanton contributions.

As before, the coefficients $g_n^{(k)}$ with n > 0 are obtained from the coupled modular anomaly equations (4.180), which in this case become

$$\frac{\partial g_1^{(k)}}{\partial E_2} = \frac{5}{22} g_0^{(k)} g_0 \quad \text{and} \quad \frac{\partial g_1^{(k)}}{\partial \widetilde{E}_2^{(\lambda_k)}} = \frac{25}{132} \left(g_0^{(k)}\right)^2 \quad (4.216)$$

where the quasi-modular forms $\widetilde{E}_2^{(3)}$ for k = 2 and $\widetilde{E}_2^{(4)}$ for k = 3 are given in (D.24) and (D.30), respectively. These equations can be solved in a straightforward manner and the undetermined modular terms can be fixed by comparing with the perturbative contributions in (4.212b) and (4.213b). In this way one obtains

$$g_1^{(2)} = \frac{T_6^2}{110} \left(\frac{f_{2,+}^{(1)} E_4^2 E_2}{\left(f_{1,-}^{(3)}\right)^2} + \frac{5}{6} \frac{E_4^3 \widetilde{E}_2^{(3)}}{\left(f_{1,-}^{(3)}\right)^4} - \frac{8}{3} \frac{f_{2,+}^{(1)} E_4 E_6}{\left(f_{1,-}^{(3)}\right)^2} \right) + \frac{T_{12}}{22} \frac{f_{2,+}^{(1)} E_4 E_6}{\left(f_{1,-}^{(3)}\right)^2} , \quad (4.217a)$$

$$g_1^{(3)} = \frac{T_6^2}{110} \left(\frac{E_4 E_6 E_2}{f_{2,-}^{(4)}} + \frac{5}{6} \frac{E_6^2 \widetilde{E}_2^{(4)}}{\left(f_{2,-}^{(4)}\right)^2} - \frac{37}{12} \frac{E_6^2}{f_{2,-}^{(4)}} \right) + \frac{T_{12}}{22} \frac{E_6^2}{f_{2,-}^{(4)}} .$$
(4.217b)

Again, by Fourier expanding the right hand sides and expressing everything in terms of the bare coupling q_0 , we retrieve the first instanton corrections in perfect agreement with the localization results (4.212) and (4.213), and predict all successive non-perturbative contributions. Similar analyses can be performed at higher orders; we have checked that the coefficients $g_n^{(k)}$ are successfully determined in this manner.

Outlook

5.1 Chiral Rings

In this work we have performed a detailed analysis of the simplest chiral observables constructed from the adjoint scalar Φ of the $\mathcal{N} = 2^* \operatorname{U}(N)$ SYM theory. The expressions for $\langle \operatorname{Tr} \Phi^n \rangle$ that we obtained using localization methods are written as mass expansions, with the dependence on the gauge coupling constant being completely resummed into quasi-modular forms, and the dependence on the classical vacuum expectation values expressed through lattice sums involving the roots and weights of the gauge algebra. Therefore, these findings can be thought of as a natural generalization of the results obtained in [11, 12, 13] for the prepotential to other observables of the $\mathcal{N} = 2^*$ theory.

We also found that the symmetric polynomials W_n constructed out of $\langle \operatorname{Tr} \Phi^n \rangle$ do not satisfy the classical chiral ring relations [21], while some simple redefinitions allow one to enforce them. The redefined chiral observables obtained in this way perfectly match those we derived by completely independent means, namely from the SW curves and the associated period integrals, or from modular anomaly equations. We then identified particular combinations A_n of chiral observables that transform as modular forms of weight n under the non-perturbative S-duality group, and derived a relation between the W_n and the A_n which is exact both in the hypermultiplet mass and in the gauge coupling constant.

Given that our results are a generalization of what was found in [11, 12, 13], it is natural to ask ourselves about the possibility of extending the above analysis to $\mathcal{N} = 2^*$ theories with other classical groups. In this respect we recall that the integrable system that governs the quantum gauge theory for these cases and the associated Lax pair have been obtained in [100, 101]. However, for the D_n series, the explicit form of the spectral curves in terms of elliptic and modular forms is only known for cases with low rank [17]. Thus, it would be very interesting to revisit this problem in the present context, especially given the significant progress that has been made relating gauge theories and integrable systems over the past decade [102, 103, 54]. The localization results available for a generic group G would provide additional checks on the correctness of the proposed solution. Another important class of theories to consider would be the superconformal ADE quiver-type models studied in [21, 104].

It would also be worthwhile to calculate these chiral observables for other theories, such as SQCD-like theories. In these cases, the prepotential has been resummed in terms of quasi-modular forms of generalized triangle groups in a special locus on the moduli space [26, 78] and thus it would be interesting to see if one can obtain similar results for the one point functions of chiral observables as well.

It would also be very interesting to investigate the modular properties of the chiral observables in presence of the Ω -deformation. We expect that in the Nekrasov-Shatashvili limit [103] the chiral observables A_n still satisfy the modular anomaly equation (3.93) and hence have the same behaviour under S-duality. For a generic Ω -background, instead, we expect a modification of the modular anomaly equation with the addition of a term proportional to $\epsilon_1 \epsilon_2$. This is in analogy with what happens for the modular anomaly equation satisfied by the prepotential whose S-

dual in the Ω -deformed theories is obtained via a Fourier transform which generalizes the Legendre transform occurring at the classical level [23, 70].

Finally, we remark that the calculation of the one point functions $\langle \operatorname{Tr} \Phi^n \rangle$ has an important role in the physics of surface operators [105, 106] (for a review see for instance [107]). The infrared physics of surface operators in $\mathcal{N} = 2$ gauge theories is in fact captured by a twisted effective superpotential in a two dimensional theory. As shown in [108], one of the ways in which this twisted superpotential can be determined is from the generating function of the expectation values of chiral ring elements in the bulk four dimensional theory. Our results can be interpreted as a first step in this direction. Furthermore, it would be interesting to explore if the existence of combinations of chiral ring elements that have simple modular behaviour under S-duality can be useful to improve our understanding of the two dimensional theory that captures the infrared physics of surface operators.

5.2 Fundamental Matters

In this work we have obtained two sets of largely independent, but complementary results. In the first part, we calculated the period matrix for massless $\mathcal{N} = 2$ SQCD theories with gauge group SU(N) in the conformal limit in a locus of vacua possessing a \mathbb{Z}_N symmetry. We uncovered an interesting modular structure that becomes manifest only when the observables are written in terms of the $\left[\frac{N}{2}\right]$ renormalized couplings τ_k . In particular, we have shown that on each of these couplings, the S-duality group acts as a (generalized) triangle group. We also proposed a nonperturbatively *exact* relation between the bare coupling and the renormalized ones in terms of the hauptmodul of the corresponding triangle group, namely

$$q_{0} = \frac{\sqrt{j_{\lambda_{k}}(\tau_{k}) - d_{\lambda_{k}}^{-1}} - \sqrt{j_{\lambda_{k}}(\tau_{k})}}{\sqrt{j_{\lambda_{k}}(\tau_{k}) - d_{\lambda_{k}}^{-1}} + \sqrt{j_{\lambda_{k}}(\tau_{k})}} .$$
(5.1)

This relation correctly reproduces the instanton expansion and we showed that it is consistent with expectations from S-duality. While previous investigations [91, 93] concentrated essentially on only one of these effective couplings, which in our notation is $\tau_{\left[\frac{N}{2}\right]}$, our analysis shows that S-duality is more transparent if we consider all individual couplings τ_k . Of course, we could select one of them and express all the others in terms of it using the exact relation (5.1) via the bare coupling q_0 , but then the modular structure we have described is hidden.

There are many questions that remain to be explored. For example, it would be interesting to understand from "first principles" the spectrum of λ_k , for which our case-by-case analysis provides the simple answer

$$\lambda_k = 4\sin^2 \frac{k\pi}{N} \,. \tag{5.2}$$

Using the universal formula (5.1), questions about the strong coupling properties of the gauge theory could be addressed in an explicit way because the behaviour of the hauptmodul j_{λ_k} around the strong coupling cusps in the τ_k plane is well understood [28].

As an interesting curiosity, we observe that if we define

$$j_{\lambda_k}^* = -\frac{1}{4d_{\lambda_k} q_0} \tag{5.3}$$

then, in the arithmetic cases (see Tab. 1) the pairs $(j_{\lambda_k}, j^*_{\lambda_k})$ satisfy remarkable

identities, called the Ramanujan-Sato identities, that take the form [109]:

$$\sum_{k=0}^{\infty} s_{\lambda_k}^A(k) \, \frac{1}{(j_{\lambda_k})^{k+1/2}} = \pm \sum_{k=0}^{\infty} s_{\lambda_k}^B(k) \, \frac{1}{(j_{\lambda_k}^*)^{k+1/2}} \,, \tag{5.4}$$

where the $s_{\lambda_k}^{A,B}(k)$ are integers. It would be interesting to understand if these mathematical identities hold also in the non-arithmetic cases and if they have any interpretation within the gauge theory.

In the second part, we considered massive SQCD theories with SU(N) gauge groups, and restricted our analysis to mass configurations that respect the \mathbb{Z}_N symmetry of the special vacuum. We then showed that in this case the modular structure of the massless theory is deformed in an interesting manner. In particular we have proved that the period matrix maintains the same structure as in the massless case, while the renormalized couplings have a semiclassical expansion with mass dependent coefficients. In the arithmetic theories, these coefficients are constrained by S-duality to satisfy coupled modular anomaly equations whose solutions are meromorphic functions of quasi-modular forms of the congruence subgroups of the modular group.

A natural question to pose is whether these results can be extended to all SU(N) theories. Since in the non-arithmetic cases the S-duality group has no subset in common with the modular group, we expect that the automorphic forms and Eisenstein series of the (generalized) triangle groups should play an important role. This subject seems to be of recent interest in the mathematical literature [28] and it might be worthwhile to explore this possibility.

Another extension of our work would be to study the modular structure in the special vacuum with generic masses or in the Ω -deformed theory [3, 4]. An incentive to study this problem comes from the AGT correspondence [55, 110]. Indeed, in the SU(2) theory with four flavours the non-perturbative relation between the bare and renormalized coupling plays an important role in writing the prepotential as quasi-

modular functions. The quantum-corrected coupling constant is used to rewrite the null- vector decoupling equation as an elliptic equation [75, 76, 77]. This in turn can be used to obtain the Ω -dependent corrections to the prepotential in terms of modular functions in the Nekrasov-Shatashvili limit [54]. It would be nice to extend this approach to higher rank gauge theories using the non-perturbative relation (5.1).

A more difficult but very interesting problem is to release the special vacuum constraints and analyse the theory at a generic point on the Coulomb moduli space, to see how the modular structures we have obtained are generalized. We hope to return to some of these issues in the near future.

Eisenstein Series and Elliptic Functions

Eisenstein Series

The Eisenstein series E_{2n} are holomorphic functions of $\tau \in \mathbb{H}_+$ defined as

$$E_{2n} = \frac{1}{2\zeta(2n)} \sum_{m,n \in \mathbb{Z}^2 \setminus \{0,0\}} \frac{1}{(m+n\tau)^{2n}} .$$
 (A.1)

For n > 1, they are modular forms of weight 2n, namely under an $SL(2, \mathbb{Z})$ transformation

$$\tau \to \tau' = \frac{a\tau + b}{c\tau + d}$$
 with $a, b, c, d \in \mathbb{Z}$ and $ad - bc = 1$, (A.2)

they transform as

$$E_{2n}(\tau') = (c\tau + d)^{2n} E_{2n}(\tau) .$$
(A.3)

For n = 1, the E_2 series is instead quasi-modular. Its modular transformation has in fact an anomalous term:

$$E_2(\tau') = (c\tau + d)^2 E_2(\tau) + \frac{6}{i\pi}c(c\tau + d) .$$
 (A.4)

All modular forms of weight 2n > 6 can be expressed as polynomials of E_4 and E_6 ; the quasi-modular forms instead can be expressed as polynomials in E_2 , E_4 and E_6 .

The Eisenstein series admit a Fourier expansion in terms of $q = e^{2\pi i \tau}$ of the form

$$E_{2n} = 1 + \frac{2}{\zeta(1-2n)} \sum_{k=1}^{\infty} \sigma_{2n-1}(k) q^k , \qquad (A.5)$$

where $\sigma_p(k)$ is the sum of the *p*-th powers of the divisors of *k*. In particular, this amounts to

$$E_{2} = 1 - 24 \sum_{k=1}^{\infty} \sigma_{1}(k)q^{k} = 1 - 24q - 72q^{2} - 96q^{3} + \cdots ,$$

$$E_{4} = 1 + 240 \sum_{k=1}^{\infty} \sigma_{3}(k)q^{k} = 1 + 240q + 2160q^{2} + 6720q^{3} + \cdots ,$$

$$E_{6} = 1 - 504 \sum_{k=1}^{\infty} \sigma_{5}(k)q^{k} = 1 - 504q - 16632q^{2} - 122976q^{3} + \cdots .$$
(A.6)

The quasi-modular and modular forms are connected to each other by logarithmic q-derivatives as

$$q\frac{dE_2}{dq} = \frac{1}{12} \left(E_2^2 - E_4 \right) , \quad q\frac{dE_4}{dq} = \frac{1}{3} \left(E_2 E_4 - E_6 \right) , \quad q\frac{dE_6}{dq} = \frac{1}{2} \left(E_2 E_6 - E_4^2 \right) .$$
(A.7)

As an interesting aside, we mention in passing that these logarithmic q-derivative relations between Eisenstein series of $SL(2,\mathbb{Z})$ has recently been generalized to all Hecke groups (introduced in Chapter 4) in [111, 112].

The Eisenstein series E_2 is related to the derivative of the Dedekind η -function

$$\eta(q) = q^{1/24} \prod_{k=1}^{\infty} (1 - q^k) .$$
 (A.8)

In fact, we have

$$q\frac{d}{dq}\log\left(\frac{\eta}{q^{1/24}}\right) = -\sum_{k=1}^{\infty}\sigma_1(k)q^k = \frac{E_2 - 1}{24} .$$
 (A.9)

Jacobi *θ*-Functions

The Jacobi θ -functions are defined as

$$\theta \begin{bmatrix} a \\ b \end{bmatrix} (z|\tau) = \sum_{n} e^{\pi i \tau \left(n - \frac{a}{2}\right)^2 + 2\pi i \left(z - \frac{b}{2}\right) \left(n - \frac{a}{2}\right)} , \qquad (A.10)$$

for a, b = 0, 1. These functions are quasi-periodic, in a multiplicative fashion, for shifts of the variable z by a lattice element $\lambda = p\tau + q$, with $p, q \in \mathbb{R}$; in fact one has

$$\theta \begin{bmatrix} a \\ b \end{bmatrix} (z + \lambda | \tau) = e(\lambda, z) \theta \begin{bmatrix} a \\ b \end{bmatrix} (z | \tau) , \qquad (A.11)$$

where

$$e(\lambda, z) = e^{-\pi i \tau p^2 - 2\pi i p \left(z - \frac{b}{2}\right) - \pi i a q}$$
 (A.12)

As customary, we use the notation

$$\theta_1(z|\tau) = \theta \begin{bmatrix} 1\\1 \end{bmatrix} (z|\tau) , \qquad \theta_2(z|\tau) = \theta \begin{bmatrix} 1\\0 \end{bmatrix} (z|\tau) ,$$

$$\theta_3(z|\tau) = \theta \begin{bmatrix} 0\\0 \end{bmatrix} (z|\tau) , \qquad \theta_4(z|\tau) = \theta \begin{bmatrix} 0\\1 \end{bmatrix} (z|\tau) .$$
(A.13)

By evaluating these functions at z = 0, one obtains the so-called θ -constants $\theta_a(\tau)$, which satisfy the abstruse identity:

$$\theta_3(\tau)^4 - \theta_2(\tau)^4 - \theta_4(\tau)^4 = 0 , \qquad (A.14)$$

while $\theta_1(\tau) = 0$.

The Eisenstein series E_4 and E_6 can be written as polynomials in the θ -constants

according to

$$E_{4} = \frac{1}{2} \left(\theta_{2}(\tau)^{8} + \theta_{3}(\tau)^{8} + \theta_{4}(\tau)^{8} \right) ,$$

$$E_{6} = \frac{1}{2} \left(\theta_{3}(\tau)^{4} + \theta_{4}(\tau)^{4} \right) \left(\theta_{2}(\tau)^{4} + \theta_{3}(\tau)^{4} \right) \left(\theta_{4}(\tau)^{4} - \theta_{4}(\tau)^{4} \right) .$$
(A.15)

Weierstraß Function

The Weierstraß function $\wp(z|\tau)$ defined by

$$\wp(z|\tau) = \frac{1}{z^2} + \sum_{m,n \in \mathbb{Z}^2 \setminus \{0,0\}} \left(\frac{1}{(z+m\tau+n)^2} - \frac{1}{(m\tau+n)^2} \right) , \qquad (A.16)$$

is a meromorphic function in the complex z-plane with a double pole in z = 0, which is doubly periodic with periods 1 and τ . We often leave the τ -dependence implicit, and write simply $\wp(z)$.

It is a Jacobi form of weight 2 and index 0, namely under a modular transformation (A.2) combined with $z \to z' = z/(c\tau + d)$, it transforms as

$$\wp(z'|\tau') = (c\tau + d)^2 \wp(z|\tau).$$
 (A.17)

It also satisfies the following differential equation

$$\wp'(z|\tau)^2 = 4\,\wp^3(z|\tau) - \frac{4\pi^4 \,E_4}{3}\,\wp(z|\tau) - \frac{8\pi^6 \,E_6}{27} \,. \tag{A.18}$$

Using the quasi-periodicity properties of the θ -functions given in (A.11), it is easy to show that second derivative of θ_1 is a proper periodic function; indeed

$$\frac{d^2}{dz^2}\log\theta_1(z+m+n\tau|\tau) = \frac{d^2}{dz^2}\log\theta_1(z|\tau) .$$
 (A.19)

Furthermore, by studying its pole structure, it is possible to show that it coincides with the Weierstraß function, up to a z-independent term:

$$\wp(z|\tau) = -\frac{d^2}{dz^2} \log \theta_1(z|\tau) + c .$$
 (A.20)

The explicit evaluation of the constant shows that

$$c = -\frac{\pi^2}{3} \left(1 - 24 \sum_{k=1}^{\infty} \frac{q^k}{(1-q^k)^2} \right) = -\frac{\pi^2}{3} \left(1 - 24 \sum_{k=1}^{\infty} \sigma_1(k) q^k \right) = -\frac{\pi^2}{3} E_2 ,$$
(A.21)

so that we have

$$\wp(z|\tau) = -\frac{d^2}{dz^2} \log \theta_1(z|\tau) - \frac{\pi^2}{3} E_2 . \qquad (A.22)$$

Using the notation of Section 3.3.2 (see in particular (3.49)), from (A.22) one can easily show that

$$h_1' = \frac{1}{2\pi i} \frac{d}{dz} h_1(z) = \frac{1}{(2\pi i)^2} \frac{d^2}{dz^2} \log \theta_1(z|\tau) = -\frac{\wp(z|\tau)}{(2\pi i)^2} + \frac{E_2}{12}$$
(A.23)

which proves the first identity in (3.56). By taking further derivatives of this equation with respect to $2\pi i z$ and using the differential equation (A.18), one can straightforwardly prove the other identities in (3.56).

Using the periodicity property (A.19), it is possible to exploit the relation (A.22) to deduce the values of the integral of the \wp function along the α and β cycles of the torus, that are parametrized respectively by $z = \gamma$ and $z = \gamma \tau$, with $\gamma \in [0, 1]$; for instance we have

$$\oint_{\alpha} \wp(z|\tau) = -\frac{\pi^2}{3} E_2 . \qquad (A.24)$$

This result has been used in Section 3.4, see in particular (3.76).

By differentiating the differential equation (A.18) and using the previous result, one can compute also the integral of higher powers of \wp . For instance, the first derivative

of (A.18) yields the relation

$$\wp(z|\tau)'' = 6\,\wp(z|\tau)^2 - \frac{2\pi^4}{3}E_4 \tag{A.25}$$

from which we find

$$\oint_{\alpha} \wp^2(z|\tau) = \frac{\pi^4}{9} E_4 .$$
 (A.26)

Proceeding in this way, one can easily compute the period integrals for higher powers of \wp , (see for example [75] and references therein).

B

Generalized Donagi-Witten Polynomials

In Section 3.3.1 we obtained the expression of the first polynomials P_n that appear in the Donagi-Witten curve, by imposing the requirements that they satisfy the recursion relation

$$\frac{dP_n}{dt} = nP_{n-1} , \qquad (B.1)$$

and that their behaviour at infinity is

$$P_n\left(t+\frac{m}{u}\right) \sim \frac{\alpha_n}{u} + \text{regular}$$
 (B.2)

This procedure can be iteratively carried out order by order in n. The general form of the P_n required from (B.1) is

$$P_n = t^n - \sum_{p=2}^n (-1)^p (p-1) x_p m^p \binom{n}{p} t^{n-p} , \qquad (B.3)$$

where the coefficients x_p are elliptic and modular forms of weight p that can be fixed recursively. As discussed in the main text, up to n = 3 the solution to the constraints is unique, namely

$$P_0 = 1 , \qquad P_2 = t^2 - m^2 x ,$$

$$P_1 = t , \qquad P_3 = t^3 - 3t m^2 x + 2m^3 y .$$
(B.4)

From n = 4 on, several combinations of elliptic and modular forms start to appear and their relative coefficients are not uniquely fixed by the requirement of the behaviour at infinity. For instance, for n = 4 and n = 5 one finds a one-parameter family of solutions, and for n = 6 a two-parameter family of solutions, given by

$$P_{4} = t^{4} - 6t^{2} x m^{2} + 8t y m^{3} - (3x^{2} - \alpha E_{4}) m^{4} ,$$

$$P_{5} = t^{5} - 10t^{3} m^{2} x + 20t^{2} m^{3} y - 5t (3x^{2} - \alpha E_{4}) m^{4} + 4m^{5} x y ,$$

$$P_{6} = t^{6} - 15t^{4} m^{2} x + 40t^{3} m^{3} y - 15t^{2} (3x^{2} - \alpha E_{4}) m^{4} + 24t m^{5} x y$$

$$- m^{6} \left((5 + \beta) x^{3} - \beta y^{2} - \frac{E_{4}}{48} (32 - 720 \alpha + \beta) x \right) .$$
(B.5)

These polynomials correspond to the expression in (B.3) where the first few x_p are

$$x_{2} = x , x_{3} = y , x_{4} = x^{2} - \frac{\alpha}{3}E_{4} ,$$

$$x_{5} = xy , x_{6} = \frac{1}{5}\left(\left(5 + \beta\right)x^{3} - \beta y^{2} - \frac{E_{4}}{48}\left(32 - 720 \alpha + \beta\right)x\right) .$$
(B.6)

Modular Covariance and the D'Hoker-Phong Curve

In this Appendix we explain how to obtain the relation (3.58) between the modular covariant A_n and the W_n , directly from the D'Hoker-Phong form of the SW curve instead of comparing it with the Donagi-Witten curve as we did in Section 3.3.3.

Recall that in the D'Hoker-Phong approach the SW curve is given by

$$R(t,z) = \sum_{\ell=0}^{N} (-1)^{\ell} W_{\ell} \left[t - m \left(\frac{1}{2\pi i} \frac{d}{dz} + h_1(z) \right) \right]^{N-\ell} 1 \bigg|_{h_1=0} = 0$$
(C.1)

As discussed in the main text, the coefficients W_{ℓ} do not transform homogeneously under S-duality. One can see this clearly by analyzing how the other objects appearing in (C.1) transform. In fact, using (A.23), the modular property (A.17) of the Weierstraß function implies that h'_1 transforms as a quasi-modular form of weight 2. Acting with additional derivatives on both sides of (A.23) kills the term proportional to E_2 so that the *n*-th derivative of h_1 for n > 1 transforms homogeneously with weight n + 1. On the other hand, from the analysis in section 3.3.1, we know that one can rewrite the equation for the curve such that it becomes modular of weight N. Hence there must exist some inhomogeneous transformation law of the W_n , compensating the inhomogeneous transformation of h'_1 , such that the whole polynomial is modular covariant. Indeed, if not for this inhomogeneous transformation of h'_1 , the curve would be manifestly modular covariant. These observations suggest to introduce a new function $R_{\text{mod}}(t, z)$ with coefficients A_{ℓ} , by substituting the quasi-modular h'_1 for the modular expression $h'_1 - E_2/12$, namely

$$R_{\text{mod}}(t,z) = \sum_{\ell=0}^{N} (-1)^{\ell} A_{\ell} \left[t - m \left(\frac{1}{2\pi i} \frac{d}{dz} + h_1(z) \right) \right]^{N-\ell} 1 \Big|_{h_1=0, h_1' \to h_1' - E_2/12} .$$
(C.2)

By construction, this polynomial is modular of weight N if the coefficients A_{ℓ} are modular of weight ℓ . Equating $R_{\text{mod}} = R$ then yields a relation between the modular covariant A_{ℓ} and the expectation values of symmetric polynomials W_{ℓ} , which agrees exactly with (3.58). In fact, the asymptotic expansion at large t of R_{mod} reads

$$R_{\text{mod}}(t,z) = t^{N} - t^{N-1} A_{1} + t^{N-2} \left[A_{2} + \binom{N}{2} m^{2} \left(h_{1}' - \frac{E_{2}}{12} \right) \right] - t^{N-3} \left[A_{3} + \binom{N-1}{2} m^{2} \left(h_{1}' - \frac{E_{2}}{12} \right) A_{1} + m^{3} \binom{N}{3} h_{1}'' \right] + t^{N-4} \left[A_{4} + \binom{N-2}{2} m^{2} \left(h_{1}' - \frac{E_{2}}{12} \right) A_{2} + m^{3} \binom{N-1}{3} h_{1}'' A_{1} + \binom{N}{4} m^{4} \left(h_{1}^{(3)} + 3 \left(h_{1}' - \frac{E_{2}}{12} \right)^{2} \right) \right] + O(t^{N-5}) .$$
(C.3)

By comparing this with (3.54) and equating the coefficients of the various t powers we can easily find the relation (3.58).

Congruence Subgroups

In this appendix we collect a few results on the (quasi-)modular forms of the modular group PSL(2, \mathbb{Z}) and its congruence subgroups $\Gamma_1(2)$, $\Gamma_1(3)$ and $\Gamma_1(4)$ which occur in the arithmetic theories. We refer to the literature for the proofs of the various statements (see for example [98, 99]) and only quote the main results that are relevant for the calculations described in the main text.

The Modular Group, Eisenstein Series, and S-Parity

The Eisenstein series E_{2n} are holomorphic functions of τ (with $\text{Im}(\tau) \ge 0$), defined as

$$E_{2n} = \frac{1}{2\zeta(2n)} \sum_{m,n \in \mathbb{Z}^2 \setminus \{0,0\}} \frac{1}{(m+n\tau)^{2n}}$$
(D.1)

where ζ denotes the Riemann ζ -function. For n > 1, the E_{2n} 's are modular forms of degree 2n. In particular, under $\tau \to -1/\tau$ they transform as

$$E_{2n}\left(-\frac{1}{\tau}\right) = \tau^{2n} E_{2n}(\tau) = (-1)^n \left(i\tau\right)^{2n} E_{2n}(\tau) .$$
 (D.2)

This shows that the S-parity of E_{2n} is $(-1)^n$. The E_2 series is instead quasi-modular:

$$E_2\left(-\frac{1}{\tau}\right) = -\left(\mathrm{i}\tau\right)^2 \left(E_2(\tau) + \frac{6}{\mathrm{i}\pi\tau}\right),\tag{D.3}$$

and has odd S-parity.

All modular forms of degree 2n > 6 can be expressed in terms of E_4 and E_6 ; the quasi-modular forms instead can be expressed as polynomials in E_2 , E_4 and E_6 . The Fourier expansions of the first Eisenstein series are

$$E_{2} = 1 - 24q - 72q^{2} - 96q^{3} + \cdots ,$$

$$E_{4} = 1 + 240q + 2160q^{2} + 6720q^{3} + \cdots ,$$

$$E_{6} = 1 - 504q - 16632q^{2} - 122976q^{3} + \cdots$$
(D.4)

where $q = e^{2\pi i \tau}$.

Let us now consider the subgroup Γ' generated by T and $S' = STS^{-1}$. As the results on the SU(6) theory reported in Section 4.9 explicitly indicate, the following expression

$$f_{2,+}^{(1)} = 1 + 120q - 6120q^2 + 737760q^3 + \cdots$$
 (D.5)

plays a crucial role in matching the modular structure of the period matrix with the multi-instanton calculations. We notice that this expansion is accounted for if we write

$$f_{2,+}^{(1)} = \left(E_4\right)^{\frac{1}{2}}$$
. (D.6)

The presence of the square root seems to suggest that the modular group can be viewed as a two-sheeted cover of Γ' . Moreover we observe that everything is consistent by requiring that $f_{2,+}^{(1)}$ be a modular form of weight 2 under Γ' and with positive *S*-parity. This also explains the notation we have used. To construct the modular forms of $\Gamma_1(2)$ we first define the following functions

$$f_{4,\pm}^{(2)}(\tau) = \left(\frac{\eta^2(\tau)}{\eta(2\tau)}\right)^8 \pm 64 \left(\frac{\eta^2(2\tau)}{\eta(\tau)}\right)^8$$
(D.7)

where $\eta(\tau)$ is the Dedekind η -function. Their Fourier expansions are

$$f_{4,+}^{(2)} = 1 + 48q + 624q^2 + 1344q^3 + \cdots ,$$

$$f_{4,-}^{(2)} = 1 - 80q - 400q^2 - 2240q^3 + \cdots ,$$
(D.8)

where as usual $q = e^{2\pi i \tau}$. These functions are modular forms $\Gamma_1(2)$ of weight 4 [98, 99], as evinced by their behavior under the S'-transformation:

$$f_{4,\pm}^{(2)}\left(\frac{\tau}{1-2\tau}\right) = (1-2\tau)^4 f_{4,\pm}^{(2)}(\tau) \ . \tag{D.9}$$

In addition, using the modular transformation properties of the Dedekind η -function and in particular

$$\eta\left(-\frac{1}{\tau}\right) = \sqrt{-\mathrm{i}\tau}\,\eta(\tau) \;, \tag{D.10}$$

one can easily check that

$$f_{4,\pm}^{(2)}\left(-\frac{1}{2\tau}\right) = \pm \left(i\sqrt{2\tau}\right)^4 f_{4,\pm}^{(2)}(\tau) .$$
 (D.11)

Thus $f_{4,\pm}^{(2)}$ have weight 4 and S-parity (+1) and (-1) respectively, as the notation itself suggests.

Now consider the square-root of $f_{4,+}^{(2)}$, namely

$$f_{2,+}^{(2)} := \left(f_{4,+}^{(2)}\right)^{\frac{1}{2}} = 1 + 24q + 24q^2 + 96q^3 + \cdots$$
 (D.12)
This is a modular form of $\Gamma_1(2)$ with weight 2 and positive S-parity. Indeed,

$$f_{2,+}^{(2)}\left(-\frac{1}{2\tau}\right) = \left(i\sqrt{2\tau}\right)^2 f_{2,+}^{(2)}(\tau) .$$
 (D.13)

The modular forms of $\Gamma_1(2)$ form a ring generated by $f_{2,+}^{(2)}$ and $f_{4,-}^{(2)}$.

In order to study quasi-modular forms of $\Gamma_1(2)$ let us consider the second Eisenstein series E_2 which satisfies

$$E_2\left(-\frac{1}{2\tau}\right) = \left(4\tau^2\right)E_2(2\tau) + \frac{12\tau}{i\pi} ,$$

$$E_2(2\tau) = \frac{1}{2}E_2(\tau) + \frac{1}{2}f_{2,+}^{(2)}(\tau) .$$
(D.14)

These equations naturally lead us to introduce the following combination

$$\widetilde{E}_{2}^{(2)} = E_{2} + \frac{1}{2}f_{2,+}^{(2)} = \frac{3}{2} - 12q - 60q^{2} - 48q^{3} + \cdots$$
 (D.15)

Using (D.13), it is easy to check that

$$\widetilde{E}_{2}^{(2)}\left(-\frac{1}{2\tau}\right) = -\left(\sqrt{2}\mathrm{i}\tau\right)^{2}\left(\widetilde{E}_{2}^{(2)}(\tau) + \frac{6}{\mathrm{i}\pi\tau}\right),\qquad(\mathrm{D.16})$$

which shows that $\widetilde{E}_2^{(2)}$ transforms under S-duality similarly to E_2 and has negative S-parity.

The Congruence Subgroup $\Gamma_1(3)$

To construct the modular forms of $\Gamma_1(3)$ we first define the following functions

$$f_{3,\pm}^{(3)}(\tau) = \left(\frac{\eta^3(\tau)}{\eta(3\tau)}\right)^3 \mp 27 \left(\frac{\eta^3(3\tau)}{\eta(\tau)}\right)^3$$
(D.17)

whose Fourier expansions are

$$f_{3,+}^{(3)} = 1 - 36q - 54q^2 - 252q^3 + \cdots ,$$

$$f_{3,-}^{(3)} = 1 + 18q + 108q^2 + 234q^3 + \cdots ,$$
(D.18)

where as usual $q = e^{2\pi i \tau}$. These functions are modular forms $\Gamma_1(3)$ of weight 3 [98, 99]. Under the S'- transformation, they behave as

$$f_{3,\pm}^{(3)}\left(\frac{\tau}{1-3\tau}\right) = (1-3\tau)^3 f_{3,\pm}^{(3)}(\tau) , \qquad (D.19)$$

while under the S-transformation they change as

$$f_{3,\pm}^{(3)}\left(-\frac{1}{3\tau}\right) = \pm \left(i\sqrt{3}\tau\right)^3 f_{3,\pm}^{(3)}(\tau) , \qquad (D.20)$$

as one can easily check using the modular properties of the Dedekind function. The last equation shows that $f_{3,\pm}^{(3)}$ have S-parity (+1) and (-1), respectively, as also the notation suggests.

Now consider the cube-root of $f_{3,-}^{(3)}$, namely

$$f_{1,-}^{(3)} := \left(f_{3,-}^{(3)}\right)^{\frac{1}{3}} = 1 + 6q + 6q^3 + \cdots$$
 (D.21)

This is a modular form of $\Gamma_1(3)$ with weight 1 and negative S-parity. Indeed,

$$f_{1,-}^{(3)}\left(-\frac{1}{3\tau}\right) = -\left(i\sqrt{3}\tau\right)f_{1,-}^{(3)}(\tau) . \qquad (D.22)$$

The modular forms of $\Gamma_1(3)$ form a ring generated by $f_{1,-}^{(3)}$ and $f_{3,+}^{(3)}$.

In order to study quasi-modular forms of $\Gamma_1(3)$ we have to consider the second

Eisenstein series E_2 which satisfies

$$E_2\left(-\frac{1}{3\tau}\right) = \left(9\tau^2\right)E_2(3\tau) + \frac{18\tau}{i\pi} ,$$

$$E_2(3\tau) = \frac{1}{3}E_2(\tau) + \frac{2}{3}\left(f_{1,-}^{(3)}(\tau)\right)^2 .$$
(D.23)

These equations naturally lead us to introduce the following combination

$$\widetilde{E}_{2}^{(3)} = E_{2} + \left(f_{1,-}^{(3)}\right)^{2} = 2 - 12q - 36q^{2} - 84q^{3} + \cdots$$
(D.24)

Using (D.22), it is easy to check that

$$\widetilde{E}_2^{(3)}\left(-\frac{1}{3\tau}\right) = -\left(\sqrt{3}\mathrm{i}\tau\right)^2 \left(\widetilde{E}_2^{(3)}(\tau) + \frac{6}{\mathrm{i}\pi\tau}\right),\qquad(\mathrm{D.25})$$

which shows that $\widetilde{E}_2^{(3)}$ transforms under S-duality similarly to E_2 and has negative S-parity.

The Congruence Subgroup $\Gamma_1(4)$

The ring of modular forms of $\Gamma_1(4)$ is generated by the weight-2 modular forms which we denote $f_{2,\pm}^{(4)}$. They are defined as

$$f_{2,+}^{(4)}(\tau) := \theta_3^4(2\tau) = 1 + 8q + 24q^2 + 32q^3 + \cdots ,$$

$$f_{2,-}^{(4)}(\tau) := \theta_4^4(2\tau) - \theta_2^4(2\tau) = 1 - 24q + 24q^2 - 96q^3 + \cdots ,$$
(D.26)

where the θ_a 's are the standard Jacobi θ -functions and as usual $q = e^{2\pi i \tau}$. Using the modular properties of the θ -functions and in particular

$$\theta_2\left(-\frac{1}{\tau}\right) = \sqrt{-i\tau}\,\theta_4(\tau)\,,\quad \theta_3\left(-\frac{1}{\tau}\right) = \sqrt{-i\tau}\,\theta_3(\tau)\,,\quad \theta_4\left(-\frac{1}{\tau}\right) = \sqrt{-i\tau}\,\theta_2(\tau)\,,\quad (D.27)$$

it is easy to show that

$$f_{2,\pm}^{(4)}\left(-\frac{1}{4\tau}\right) = \pm (2i\tau)^2 f_{2,\pm}^{(4)}(\tau) . \qquad (D.28)$$

Thus $f_{2,\pm}^{(4)}$ have S-parity (+1) and (-1) respectively.

In order to study quasi-modular forms of $\Gamma_1(4)$ we have to consider the second Eisenstein series E_2 which satisfies

$$E_2\left(-\frac{1}{4\tau}\right) = \left(4\tau\right)^2 E_2(4\tau) + \frac{24\tau}{i\pi} ,$$

$$E_2(4\tau) = \frac{1}{4} E_2(\tau) + \frac{3}{4} f_{2,+}^{(4)}(\tau) .$$
(D.29)

These equations suggest to introduce the following combination

$$\widetilde{E}_{2}^{(4)} = E_{2} + \frac{3}{2} f_{2,+}^{(4)} = \frac{5}{2} - 12q - 36q^{2} - 48q^{3} + \cdots , \qquad (D.30)$$

which under S-duality transforms in a way similar to E_2 , namely

$$\widetilde{E}_{2}^{(4)}\left(-\frac{1}{4\tau}\right) = -\left(2i\tau\right)^{2}\left(\widetilde{E}_{2}^{(4)} + \frac{6}{i\pi\tau}\right).$$
(D.31)

This equation shows that $\widetilde{E}_2^{(4)}$ is a quasi-modular form with weight 2 and negative S-parity.

$\mathbf{\mathcal{F}}$ Massless SU(7) SQCD

In this appendix we briefly report the results for the massless SU(7) SQCD theory in the special vacuum.

The quantum corrected period matrix takes the form

$$\Omega = \tau_1 \mathcal{M}_1 + \tau_2 \mathcal{M}_2 + \tau_3 \mathcal{M}_4 \tag{E.1}$$

where

$$\mathcal{M}_k = \sum_{\ell=1}^3 \lambda_{k\ell} \,\mathcal{G}_\ell \tag{E.2}$$

with

$$\lambda_k = 4\sin^2 \frac{k\pi}{7} = 4\cos^2 \frac{(7-2k)\pi}{14}$$
(E.3)

and

$$\mathcal{G}_{1} = \begin{pmatrix} 2 & 0 & 1 & 1 & 1 & 2 \\ 0 & 0 & -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & -1 & 0 & 1 \\ 1 & 0 & -1 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 & 0 & 0 \\ 2 & 1 & 1 & 1 & 0 & 2 \end{pmatrix}, , \\
\mathcal{G}_{2} = \begin{pmatrix} 0 & 1 & -1 & 0 & 1 & -1 \\ 1 & 2 & 1 & 0 & 2 & 1 \\ -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 \\ 1 & 2 & 0 & 1 & 2 & 1 \\ -1 & 1 & 0 & -1 & 1 & 0 \end{pmatrix}, , \quad (E.4)$$

$$\mathcal{G}_{3} = \begin{pmatrix} 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 \\ 0 & 0 & 1 & 1 & -1 & -1 \\ 1 & 1 & 2 & 2 & 1 & 0 \\ 0 & 1 & 2 & 2 & 1 & 1 \\ -1 & -1 & 1 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

These are specific examples of the matrices defined through Eq. (4.139) of the main text. Finally, up to two instantons we find that the three renormalized couplings are given by

$$2\pi i \tau_{1} = \log q_{0} + i\pi + \log \left(4d_{\lambda_{1}}\right) + \frac{12}{49}q_{0} + \frac{192}{2401}q_{0}^{2} + \cdots ,$$

$$2\pi i \tau_{2} = \log q_{0} + i\pi + \log \left(4d_{\lambda_{2}}\right) + \frac{20}{49}q_{0} + \frac{370}{2401}q_{0}^{2} + \cdots ,$$

$$2\pi i \tau_{3} = \log q_{0} + i\pi + \log \left(4d_{\lambda_{3}}\right) + \frac{24}{49}q_{0} + \frac{474}{2401}q_{0}^{2} + \cdots ,$$

(E.5)

with

$$d_{\lambda_1} = 4611.1803\cdots, \quad d_{\lambda_2} = 163.6225\cdots, \quad d_{\lambda_3} = 69.8572\cdots.$$
 (E.6)

It is worth noticing that all coefficients in the instanton expansion of the renormalized couplings are rational. According to the general discussion of Section 4.6, these formulas should follow upon using the hauptmoduln of certain (generalized) triangle groups in the universal formula (4.96). We now show that this is indeed the case for the SU(7) theory.

Let us start from k = 3. Here we have $\lambda_3 = 4 \cos^2 \frac{\pi}{14}$ and thus the S-duality group is simply the Hecke group H(14) whose type is $\mathbf{t} = (2, 14, \infty)$. Applying the formulas of Section 4.6.1, it is not difficult to find that the corresponding hauptmodul is

$$j_{\lambda_3} = \frac{1}{q_3} + \frac{37}{98} \frac{1}{d_{\lambda_3}} + \frac{2587}{38416} \frac{q_3}{d_{\lambda_3}^2} + \frac{899}{117649} \frac{q_3^2}{d_{\lambda_3}^3} + \dots$$
(E.7)

where d_{λ_3} is precisely the same number given in (E.6).

Now let us put k = 2. In this case we have $\lambda_2 = 4 \cos^2 \frac{3\pi}{14}$ which implies that the S-duality group is a generalized triangle group with type $t = (2, \frac{14}{3}, \infty)$. As we observed in the main text, the formulas for the hauptmoduln can be formally extended also when the type has a rational entry. In this case we find

$$j_{\lambda_2} = \frac{1}{q_2} + \frac{39}{98} \frac{1}{d_{\lambda_2}} + \frac{2571}{38416} \frac{q_2}{d_{\lambda_2}^2} + \frac{4435}{705894} \frac{q_2^2}{d_{\lambda_2}^3} + \dots$$
(E.8)

where d_{λ_2} is exactly as in (E.6).

Finally for k = 1, we have $\lambda_1 = 4\cos^2\frac{5\pi}{14}$ leading to a generalized triangle group with type $t = (2, \frac{14}{5}, \infty)$. In this case the corresponding hauptmodul is

$$j_{\lambda_1} = \frac{1}{q_1} + \frac{43}{98} \frac{1}{d_{\lambda_1}} + \frac{2521}{38416} \frac{q_1}{d_{\lambda_1}^2} + \frac{2573}{705894} \frac{q_1^2}{d_{\lambda_1}^3} + \dots$$
(E.9)

with d_{λ_1} given in (E.6).

If we now plug these expansions in the universal formula (4.96) and invert the resulting series, we perfectly match the instanton results (E.5) obtained from localization, thus confirming also in this case the consistency of our proposal.

Bibliography

- [1] N. Seiberg and E. Witten, Electric magnetic duality, monopole condensation, and confinement in N=2 supersymmetric Yang-Mills theory, Nucl. Phys. B426 (1994) 19-52, arXiv:hep-th/9407087 [hep-th].
 [Erratum: Nucl. Phys.B430,485(1994)].
- [2] N. Seiberg and E. Witten, Monopoles, duality and chiral symmetry breaking in N=2 supersymmetric QCD, Nucl. Phys. B431 (1994) 484-550, arXiv:hep-th/9408099 [hep-th].
- [3] N. A. Nekrasov, Seiberg-Witten prepotential from instanton counting, Adv. Theor. Math. Phys. 7 no. 5, (2003) 831-864, arXiv:hep-th/0206161
 [hep-th].
- [4] N. Nekrasov and A. Okounkov, Seiberg-Witten theory and random partitions, Prog. Math. 244 (2006) 525-596, arXiv:hep-th/0306238 [hep-th].
- [5] N. Seiberg, Supersymmetry and Nonperturbative beta Functions, Phys. Lett.
 B206 (1988) 75–80.
- [6] E. D'Hoker and D. H. Phong, Lectures on supersymmetric Yang-Mills theory and integrable systems, in Theoretical physics at the end of the twentieth

century. Proceedings, Summer School, Banff, Canada, June 27-July 10, 1999, pp. 1-125. 1999. arXiv:hep-th/9912271 [hep-th].

- [7] N. Dorey, T. J. Hollowood, V. V. Khoze, and M. P. Mattis, *The Calculus of many instantons*, *Phys. Rept.* **371** (2002) 231-459, arXiv:hep-th/0206063
 [hep-th].
- [8] E. Witten, Topological Quantum Field Theory, Commun. Math. Phys. 117 (1988) 353.
- [9] M. F. Atiyah, N. J. Hitchin, V. G. Drinfeld, and Yu. I. Manin, Construction of Instantons, Phys. Lett. A65 (1978) 185–187. [,133(1978)].
- [10] J. A. Minahan, D. Nemeschansky, and N. P. Warner, Instanton expansions for mass deformed N=4 superYang-Mills theories, Nucl. Phys. B528 (1998) 109-132, arXiv:hep-th/9710146 [hep-th].
- [11] M. Billò, M. Frau, F. Fucito, A. Lerda, and J. F. Morales, S-duality and the prepotential in N = 2* theories (I): the ADE algebras, JHEP 11 (2015) 024, arXiv:1507.07709 [hep-th].
- [12] M. Billò, M. Frau, F. Fucito, A. Lerda, and J. F. Morales, S-duality and the prepotential of N = 2^{*} theories (II): the non-simply laced algebras, JHEP 11 (2015) 026, arXiv:1507.08027 [hep-th].
- [13] M. Billò, M. Frau, F. Fucito, J. F. Morales, and A. Lerda, *Resumming instantons in N = 2* theories with arbitrary gauge groups*, in *Proceedings*, 14th Marcel Grossmann Meeting on Recent Developments in Theoretical and Experimental General Relativity, Astrophysics, and Relativistic Field Theories (MG14) (In 4 Volumes): Rome, Italy, July 12-18, 2015, vol. 4, pp. 4139-4150. 2017. arXiv:1602.00273 [hep-th]. http://inspirehep.net/record/1418886/files/arXiv:1602.00273.pdf.

- [14] R. Donagi and E. Witten, Supersymmetric Yang-Mills theory and integrable systems, Nucl. Phys. B460 (1996) 299-334, arXiv:hep-th/9510101
 [hep-th].
- [15] R. Y. Donagi, Seiberg-Witten integrable systems, arXiv:alg-geom/9705010 [alg-geom].
- [16] A. Gorsky, I. Krichever, A. Marshakov, A. Mironov, and A. Morozov, *Integrability and Seiberg-Witten exact solution*, *Phys. Lett.* B355 (1995) 466-474, arXiv:hep-th/9505035 [hep-th].
- [17] E. D'Hoker and D. H. Phong, Calogero-Moser systems in SU(N) Seiberg-Witten theory, Nucl. Phys. B513 (1998) 405-444, arXiv:hep-th/9709053 [hep-th].
- [18] E. J. Martinec, Integrable structures in supersymmetric gauge and string theory, Phys. Lett. B367 (1996) 91-96, arXiv:hep-th/9510204 [hep-th].
- [19] F. Cachazo, M. R. Douglas, N. Seiberg, and E. Witten, Chiral rings and anomalies in supersymmetric gauge theory, JHEP 12 (2002) 071, arXiv:hep-th/0211170 [hep-th].
- [20] R. Flume, F. Fucito, J. F. Morales, and R. Poghossian, Matone's relation in the presence of gravitational couplings, JHEP 04 (2004) 008, arXiv:hep-th/0403057 [hep-th].
- [21] N. Nekrasov and V. Pestun, Seiberg-Witten geometry of four dimensional N=2 quiver gauge theories, arXiv:1211.2240 [hep-th].
- [22] D. Gaiotto, N=2 dualities, JHEP 08 (2012) 034, arXiv:0904.2715 [hep-th].

- [23] M. Billò, M. Frau, L. Gallot, A. Lerda, and I. Pesando, Modular anomaly equation, heat kernel and S-duality in N = 2 theories, JHEP 11 (2013) 123, arXiv:1307.6648 [hep-th].
- [24] M. Billò, M. Frau, L. Gallot, A. Lerda, and I. Pesando, Deformed N=2 theories, generalized recursion relations and S-duality, JHEP 04 (2013) 039, arXiv:1302.0686 [hep-th].
- [25] M. Billò, M. Frau, F. Fucito, A. Lerda, J. F. Morales, R. Poghossian, and
 D. Ricci Pacifici, Modular anomaly equations in N = 2* theories and their large-N limit, JHEP 10 (2014) 131, arXiv:1406.7255 [hep-th].
- [26] S. K. Ashok, M. Billò, E. Dell'Aquila, M. Frau, A. Lerda, and M. Raman, Modular anomaly equations and S-duality in N = 2 conformal SQCD, JHEP 10 (2015) 091, arXiv:1507.07476 [hep-th].
- [27] P. C. Argyres and S. Pelland, Comparing instanton contributions with exact results in N=2 supersymmetric scale invariant theories, JHEP 03 (2000) 014, arXiv:hep-th/9911255 [hep-th].
- [28] C. F. Doran, T. Gannon, H. Movasati, and K. M. Shokri, Automorphic forms for triangle groups, Commun. Num. Theor Phys. 07 (2013) 689-737, arXiv:1307.4372 [math.NT].
- [29] P. C. Argyres and N. Seiberg, S-duality in N=2 supersymmetric gauge theories, JHEP 12 (2007) 088, arXiv:0711.0054 [hep-th].
- [30] M.-x. Huang, A.-K. Kashani-Poor, and A. Klemm, The Ω deformed B-model for rigid N = 2 theories, Annales Henri Poincare 14 (2013) 425-497, arXiv:1109.5728 [hep-th].
- [31] M. Billò, M. Frau, L. Gallot, and A. Lerda, The exact 8d chiral ring from 4d recursion relations, JHEP 11 (2011) 077, arXiv:1107.3691 [hep-th].

- [32] S. Vandoren and P. van Nieuwenhuizen, Lectures on instantons, arXiv:0802.1862 [hep-th].
- [33] E. Corrigan, D. B. Fairlie, S. Templeton, and P. Goddard, A Green's Function for the General Selfdual Gauge Field, Nucl. Phys. B140 (1978) 31-44.
- [34] E. Corrigan and P. Goddard, Construction of Instanton and Monopole Solutions and Reciprocity, Annals Phys. 154 (1984) 253.
- [35] M. F. Sohnius, Introducing Supersymmetry, Phys. Rept. 128 (1985) 39–204.
- [36] L. Alvarez-Gaume and S. F. Hassan, Introduction to S duality in N=2 supersymmetric gauge theories: A Pedagogical review of the work of Seiberg and Witten, Fortsch. Phys. 45 (1997) 159-236, arXiv:hep-th/9701069 [hep-th].
- [37] A. Klemm, W. Lerche, S. Yankielowicz, and S. Theisen, Simple singularities and N=2 supersymmetric Yang-Mills theory, Phys. Lett. B344 (1995) 169-175, arXiv:hep-th/9411048 [hep-th].
- [38] P. C. Argyres and A. E. Faraggi, The vacuum structure and spectrum of N=2 supersymmetric SU(n) gauge theory, Phys. Rev. Lett. 74 (1995) 3931-3934,
 arXiv:hep-th/9411057 [hep-th].
- [39] A. Klemm, W. Lerche, and S. Theisen, Nonperturbative effective actions of N=2 supersymmetric gauge theories, Int. J. Mod. Phys. A11 (1996)
 1929–1974, arXiv:hep-th/9505150 [hep-th].
- [40] E. Witten, Dynamical breaking of supersymmetry, Nuclear Physics B 188 no. 3, (1981) 513 - 554. http: //www.sciencedirect.com/science/article/pii/0550321381900067.

- [41] E. Witten and D. I. Olive, Supersymmetry Algebras That Include Topological Charges, Phys. Lett. 78B (1978) 97–101.
- [42] I. M. Krichever, Elliptic solutions of the Kadomtsev-Petviashvili equation and integrable systems of particles, Functional Analysis and Its Applications 14 no. 4, (1980) 282–290.
- [43] E. Witten, Topological quantum field theory, Comm. Math. Phys. 117 no. 3, (1988) 353-386.
 https://projecteuclid.org:443/euclid.cmp/1104161738.
- [44] S. Shadchin, On certain aspects of string theory/gauge theory correspondence.
 PhD thesis, Orsay, LPT, 2005. arXiv:hep-th/0502180 [hep-th].
- [45] Y. Tachikawa, A review on instanton counting and W-algebras, in New Dualities of Supersymmetric Gauge Theories, Teschner, Jörg, ed., pp. 79-120. 2016. arXiv:1412.7121 [hep-th]. https://inspirehep.net/record/1335337/files/arXiv:1412.7121.pdf.
- [46] N. Manton, A remark on the scattering of bps monopoles, Physics Letters B 110 no. 1, (1982) 54 - 56. http: //www.sciencedirect.com/science/article/pii/0370269382909509.
- [47] N. Nekrasov and A. S. Schwarz, Instantons on noncommutative R**4 and (2,0) superconformal six-dimensional theory, Commun. Math. Phys. 198 (1998) 689-703, arXiv:hep-th/9802068 [hep-th].
- [48] A. Losev, N. Nekrasov, and S. L. Shatashvili, The Freckled instantons, arXiv:hep-th/9908204 [hep-th].
- [49] H. Nakajima and K. Yoshioka, Lectures on instanton counting, in CRM Workshop on Algebraic Structures and Moduli Spaces Montreal, Canada, July 14-20, 2003. 2003. arXiv:math/0311058 [math-ag].

- [50] R. Flume and R. Poghossian, An Algorithm for the microscopic evaluation of the coefficients of the Seiberg-Witten prepotential, Int. J. Mod. Phys. A18 (2003) 2541, arXiv:hep-th/0208176 [hep-th].
- [51] U. Bruzzo, F. Fucito, J. F. Morales, and A. Tanzini, Multiinstanton calculus and equivariant cohomology, JHEP 05 (2003) 054, arXiv:hep-th/0211108
 [hep-th].
- [52] A. S. Losev, A. Marshakov, and N. A. Nekrasov, Small instantons, little strings and free fermions, arXiv:hep-th/0302191 [hep-th].
- [53] Teschner, Jörg, Exact Results on N = 2 Supersymmetric Gauge Theories, in New Dualities of Supersymmetric Gauge Theories, Teschner, Jörg, ed., pp. 1-30. 2016. arXiv:1412.7145 [hep-th]. http://inspirehep.net/record/1335348/files/arXiv:1412.7145.pdf.
- [54] N. A. Nekrasov and S. L. Shatashvili, Quantization of Integrable Systems and Four Dimensional Gauge Theories, in Proceedings, 16th International Congress on Mathematical Physics (ICMP09): Prague, Czech Republic, August 3-8, 2009, pp. 265-289. 2009. arXiv:0908.4052 [hep-th]. http://inspirehep.net/record/829640/files/arXiv:0908.4052.pdf.
- [55] L. F. Alday, D. Gaiotto, and Y. Tachikawa, Liouville Correlation Functions from Four-dimensional Gauge Theories, Lett. Math. Phys. 91 (2010) 167-197, arXiv:0906.3219 [hep-th].
- [56] L. F. Alday, D. Gaiotto, S. Gukov, Y. Tachikawa, and H. Verlinde, Loop and surface operators in N=2 gauge theory and Liouville modular geometry, JHEP 01 (2010) 113, arXiv:0909.0945 [hep-th].
- [57] R. Dijkgraaf and C. Vafa, Toda Theories, Matrix Models, Topological Strings, and N=2 Gauge Systems, arXiv:0909.2453 [hep-th].

- [58] M. C. N. Cheng, R. Dijkgraaf, and C. Vafa, Non-Perturbative Topological Strings And Conformal Blocks, JHEP 09 (2011) 022, arXiv:1010.4573
 [hep-th].
- [59] I. Antoniadis, S. Hohenegger, K. S. Narain, and T. R. Taylor, Deformed Topological Partition Function and Nekrasov Backgrounds, Nucl. Phys. B838 (2010) 253-265, arXiv:1003.2832 [hep-th].
- [60] I. Florakis and A. Zein Assi, N = 2* from Topological Amplitudes in String Theory, Nucl. Phys. B909 (2016) 480-506, arXiv:1511.02887 [hep-th].
- [61] M. Bershadsky, S. Cecotti, H. Ooguri, and C. Vafa, Holomorphic anomalies in topological field theories, Nucl. Phys. B405 (1993) 279-304, arXiv:hep-th/9302103 [hep-th]. [AMS/IP Stud. Adv. Math.1,655(1996)].
- [62] E. Witten, Quantum background independence in string theory, in Conference on Highlights of Particle and Condensed Matter Physics (SALAMFEST) Trieste, Italy, March 8-12, 1993, pp. 0257-275. 1993.
 arXiv:hep-th/9306122 [hep-th].
- [63] M. Aganagic, V. Bouchard, and A. Klemm, Topological Strings and (Almost) Modular Forms, Commun. Math. Phys. 277 (2008) 771-819, arXiv:hep-th/0607100 [hep-th].
- [64] M. Gunaydin, A. Neitzke, and B. Pioline, Topological wave functions and heat equations, JHEP 12 (2006) 070, arXiv:hep-th/0607200 [hep-th].
- [65] M.-x. Huang and A. Klemm, Holomorphic Anomaly in Gauge Theories and Matrix Models, JHEP 09 (2007) 054, arXiv:hep-th/0605195 [hep-th].
- [66] T. W. Grimm, A. Klemm, M. Marino, and M. Weiss, Direct Integration of the Topological String, JHEP 08 (2007) 058, arXiv:hep-th/0702187 [HEP-TH].

- [67] M.-x. Huang and A. Klemm, Holomorphicity and Modularity in Seiberg-Witten Theories with Matter, JHEP 07 (2010) 083, arXiv:0902.1325 [hep-th].
- [68] M.-x. Huang and A. Klemm, Direct integration for general Ω backgrounds, Adv. Theor. Math. Phys. 16 no. 3, (2012) 805-849, arXiv:1009.1126
 [hep-th].
- [69] D. Galakhov, A. Mironov, and A. Morozov, S-duality as a beta-deformed Fourier transform, JHEP 08 (2012) 067, arXiv:1205.4998 [hep-th].
- [70] N. Nemkov, S-duality as Fourier transform for arbitrary ϵ_1, ϵ_2 , J. Phys. A47 no. 10, (2014) 105401, arXiv:1307.0773 [hep-th].
- [71] N. Lambert, D. Orlando, and S. Reffert, Alpha- and Omega-Deformations from fluxes in M-Theory, JHEP 11 (2014) 162, arXiv:1409.1219 [hep-th].
- [72] M. Beccaria, On the large Ω -deformations in the Nekrasov-Shatashvili limit of $\mathcal{N} = 2^*$ SYM, JHEP 07 (2016) 055, arXiv:1605.00077 [hep-th].
- [73] M. Beccaria and G. Macorini, Exact partition functions for the Ω -deformed $\mathcal{N} = 2^* SU(2)$ gauge theory, JHEP **07** (2016) 066, arXiv:1606.00179 [hep-th].
- [74] A. Marshakov, A. Mironov, and A. Morozov, Zamolodchikov asymptotic formula and instanton expansion in N=2 SUSY N(f) = 2N(c) QCD, JHEP 11 (2009) 048, arXiv:0909.3338 [hep-th].
- [75] A.-K. Kashani-Poor and J. Troost, The toroidal block and the genus expansion, JHEP 03 (2013) 133, arXiv:1212.0722 [hep-th].
- [76] A.-K. Kashani-Poor and J. Troost, Transformations of Spherical Blocks, JHEP 10 (2013) 009, arXiv:1305.7408 [hep-th].

- [77] A.-K. Kashani-Poor and J. Troost, Quantum geometry from the toroidal block, JHEP 08 (2014) 117, arXiv:1404.7378 [hep-th].
- [78] S. K. Ashok, E. Dell'Aquila, A. Lerda, and M. Raman, S-duality, triangle groups and modular anomalies in N = 2 SQCD, JHEP 04 (2016) 118, arXiv:1601.01827 [hep-th].
- [79] E. Gerchkovitz, J. Gomis, N. Ishtiaque, A. Karasik, Z. Komargodski, and
 S. S. Pufu, Correlation Functions of Coulomb Branch Operators, JHEP 01 (2017) 103, arXiv:1602.05971 [hep-th].
- [80] M. Billo, M. Frau, F. Fucito, L. Giacone, A. Lerda, J. F. Morales, and D. R. Pacifici, Non-perturbative gauge/gravity correspondence in N=2 theories, JHEP 08 (2012) 166, arXiv:1206.3914 [hep-th].
- [81] H. Itoyama and A. Morozov, Integrability and Seiberg-Witten theory: Curves and periods, Nucl. Phys. B477 (1996) 855-877, arXiv:hep-th/9511126
 [hep-th].
- [82] I. P. Ennes, C. Lozano, S. G. Naculich, and H. J. Schnitzer, *Elliptic models and M theory*, *Nucl. Phys.* B576 (2000) 313-346, arXiv:hep-th/9912133
 [hep-th].
- [83] M. Matone, Instantons and recursion relations in N=2 SUSY gauge theory, Phys. Lett. B357 (1995) 342-348, arXiv:hep-th/9506102 [hep-th].
- [84] S. Benvenuti, A. Hanany, and N. Mekareeya, The Hilbert Series of the One Instanton Moduli Space, JHEP 06 (2010) 100, arXiv:1005.3026 [hep-th].
- [85] C. A. Keller, N. Mekareeya, J. Song, and Y. Tachikawa, The ABCDEFG of Instantons and W-algebras, JHEP 03 (2012) 045, arXiv:1111.5624
 [hep-th].

- [86] Teschner, Jörg, ed., New Dualities of Supersymmetric Gauge Theories.
 Mathematical Physics Studies. Springer, Cham, Switzerland, 2016.
- [87] F. Fucito, J. F. Morales, and R. Poghossian, Multi instanton calculus on ALE spaces, Nucl. Phys. B703 (2004) 518-536, arXiv:hep-th/0406243
 [hep-th].
- [88] S. K. Ashok, M. Billò, E. Dell'Aquila, M. Frau, R. R. John, and A. Lerda, Non-perturbative studies of N=2 conformal quiver gauge theories, Fortsch. Phys. 63 (2015) 259-293, arXiv:1502.05581 [hep-th].
- [89] J. A. Minahan and D. Nemeschansky, Hyperelliptic curves for supersymmetric Yang-Mills, Nucl. Phys. B464 (1996) 3-17, arXiv:hep-th/9507032 [hep-th].
- [90] J. A. Minahan and D. Nemeschansky, N=2 superYang-Mills and subgroups of SL(2,Z), Nucl. Phys. B468 (1996) 72-84, arXiv:hep-th/9601059
 [hep-th].
- [91] J. A. Minahan, Duality symmetries for N=2 supersymmetric QCD with vanishing Beta functions, Nucl. Phys. B537 (1999) 243-259, arXiv:hep-th/9806246 [hep-th].
- [92] O. Aharony and S. Yankielowicz, Exact electric magnetic duality in N=2 supersymmetric QCD theories, Nucl. Phys. B473 (1996) 93-120, arXiv:hep-th/9601011 [hep-th].
- [93] P. C. Argyres and A. Buchel, The Nonperturbative gauge coupling of N=2 supersymmetric theories, Phys. Lett. B442 (1998) 180-184, arXiv:hep-th/9806234 [hep-th].

- [94] T. Gannon, The Algebraic Meaning of Being a Hauptmodul, p. 204218.
 London Mathematical Society Lecture Note Series. Cambridge University Press, 2010.
- [95] D. Zagier, Traces of Singular Moduli, in Motives, Polylogarithms, and Hodge theory, Part I, L. K. F. Bogomolov, ed. International Press, Somerville, 2002.
- [96] P. C. Argyres, M. R. Plesser, and A. D. Shapere, The Coulomb phase of N=2 supersymmetric QCD, Phys. Rev. Lett. 75 (1995) 1699-1702,
 arXiv:hep-th/9505100 [hep-th].
- [97] P. C. Argyres, S duality and global symmetries in N=2 supersymmetric field theory, Adv. Theor. Math. Phys. 2 (1998) 293-302, arXiv:hep-th/9706095
 [hep-th].
- [98] N. Koblitz, Introduction to Elliptic Curves and Modular Forms. Graduate Texts in Mathematics. Springer-Verlag, 1993. https://books.google.co.in/books?id=F9CPQgAACAAJ.
- [99] T. Apostol, Modular Functions and Dirichlet Series in Number Theory. Graduate Texts in Mathematics. Springer New York, 1976.
 https://books.google.co.in/books?id=rAJcQgAACAAJ.
- [100] E. D'Hoker and D. H. Phong, Calogero-Moser Lax pairs with spectral parameter for general Lie algebras, Nucl. Phys. B530 (1998) 537-610, arXiv:hep-th/9804124 [hep-th].
- [101] E. D'Hoker and D. H. Phong, Spectral curves for superYang-Mills with adjoint hypermultiplet for general Lie algebras, Nucl. Phys. B534 (1998)
 697-719, arXiv:hep-th/9804126 [hep-th].

- [102] N. A. Nekrasov and S. L. Shatashvili, Supersymmetric vacua and Bethe ansatz, Nucl. Phys. Proc. Suppl. 192-193 (2009) 91-112, arXiv:0901.4744
 [hep-th].
- [103] N. A. Nekrasov and S. L. Shatashvili, Quantum integrability and supersymmetric vacua, Prog. Theor. Phys. Suppl. 177 (2009) 105-119, arXiv:0901.4748 [hep-th].
- [104] N. Nekrasov, V. Pestun, and S. Shatashvili, Quantum geometry and quiver gauge theories, Commun. Math. Phys. 357 no. 2, (2018) 519-567, arXiv:1312.6689 [hep-th].
- [105] S. Gukov and E. Witten, Gauge Theory, Ramification, And The Geometric Langlands Program, arXiv:hep-th/0612073 [hep-th].
- [106] S. Gukov and E. Witten, *Rigid Surface Operators, Adv. Theor. Math. Phys.*14 no. 1, (2010) 87–178, arXiv:0804.1561 [hep-th].
- [107] S. Gukov, Surface Operators, in New Dualities of Supersymmetric Gauge Theories, Teschner, Jörg, ed., pp. 223-259. 2016. arXiv:1412.7127 [hep-th]. http://inspirehep.net/record/1335340/files/arXiv:1412.7127.pdf.
- [108] D. Gaiotto, S. Gukov, and N. Seiberg, Surface Defects and Resolvents, JHEP
 09 (2013) 070, arXiv:1307.2578 [hep-th].
- [109] H. H. CHAN and S. COOPER, Rational analogues of ramanujan's series for 1/π, Mathematical Proceedings of the Cambridge Philosophical Society 153 no. 2, (2012) 361383.
- [110] N. Wyllard, A_{N-1} conformal Toda field theory correlation functions from conformal N = 2 SU(N) quiver gauge theories, JHEP 11 (2009) 002, arXiv:0907.2189 [hep-th].

- [111] M. Raman, Aspects of Hecke Symmetry I: Ramanujan Identities and Inversion Formulas, arXiv:1803.10224 [math.NT].
- [112] S. K. Ashok, D. P. Jatkar, and M. Raman, Aspects of Hecke Symmetry II: Anomalies, Curves, and Chazy Equations, arXiv:1810.07919 [hep-th].