INDEFINITE THETA FUNCTIONS OF TYPE $(n, 1)$ I: DEFINITIONS AND EXAMPLES

LARRY ROLEN

1. Classical theta functions

Theta functions are classical examples of modular forms which play many roles in number theory (e.g., the counting of representations of integers as sums of squares), physics (e.g., as solutions to the heat equation), and geometry (e.g., they can be used to show that one cannot "hear the shape of a drum"). The most primitive example, due to Jacobi, is the following (where $q := e(\tau) := e^{2\pi i\tau}$ for $\tau \in \mathbb{H}$):

$$\theta(\tau) := \sum_{n\in\mathbb{Z}} q^{n^2}.$$

Theorem. *The function $\theta(\tau)$ is a modular form of weight $1/2$ (cf. Exercise 1 below).*

Sketch of Proof. The main idea is to apply *Poisson summation*, which states that if $\mathcal{F}$ denotes the Fourier transform, then for sufficiently nice real-valued functions f,

$$\sum_{n\in\mathbb{Z}} f(n) = \sum_{n\in\mathbb{Z}} (\mathcal{F}(f))(n).$$

This can be applied to θ by noting that showing a transformation equation of a holomorphic function on the upper half-place is equivalent, by the Identity Theorem, to showing it holds on the positive imaginary axis. The second critical observation in this case is the well-known fact that the Gaussian $e^{-\pi x^2}$ is its own Fourier transform. $\square$

We can extend this very simple example of a modular form in many ways. Firstly, in the theories of elliptic functions and q-series one encounters the important two-variable Jacobi theta function (here $\zeta := e(z)$ for $z \in \mathbb{C}$)

$$\vartheta(z;\tau) := \sum_{n\in\frac{1}{2}+\mathbb{Z}} (-1)^n q^{n^2/2}\zeta^n.$$

This is a Jacobi form of weight $1/2$ and index $1/2$, again transforming with a multiplier on $\mathrm{SL}_2(\mathbb{Z})$.

Analogous constructions can be given for quadratic forms in several variables. To describe these, let $Q\colon \mathbb{Z}^r \to \mathbb{Z}$ be a positive-definite, integral quadratic form for $r \in \mathbb{N}$, and let $B(x, y) := Q(x + y) - Q(x) - Q(y)$ be the associated *bilinear form.*

Theorem (Hecke, Schoenberg, Eichler, Zagier)**.** *For any $a \in \mathbb{Z}^r$, the function*

$$\Theta_a(z;\tau) := \sum_{n\in\mathbb{Z}^r} q^{Q(n)}\zeta^{B(n,a)}$$

is a Jacobi form of weight $r/2$ and index $Q(a)$ on a congruence subgroup of $\mathrm{SL}_2(\mathbb{Z})$.

To show this, one uses a generalized version of Poisson summation which applies to arbitrary lattices, such as $\mathbb{Z}^r$. Note that the condition of positive-definiteness is critical; indeed, $|q^{Q(n)}| = e^{-2\pi v Q(n)}$ (where $\tau = u + iv$), which will rapidly decay as $|n| \to \infty$ if and only if $Q(n)$ is positive in all directions.

Another way we can generalize our theta functions is to allow additional functions inside the summation. This anticipates generalizations necessary to discuss Zwegers' mock modular forms, and the claimed modularity properties will follow by our subsequent discussions of the proof of Zwegers' results on indefinite theta functions. For example, we can consider theta functions associated to *spherical polynomials.* To describe these, for any quadratic form Q, consider its associated symmetric matrix A, for which $Q(X) =: \frac{1}{2}\langle X, AX\rangle$ ($\langle\cdot,\cdot\rangle$ is the usual inner product on $\mathbb{R}^r$). Then the *Laplacian associated to* Q is the differential operator

$$\Delta_Q := \left\langle \frac{\partial}{\partial x}, A^{-1}\frac{\partial}{\partial x}\right\rangle.$$

After diagonalizing Q, this is the usual Laplacian on $\mathbb{R}^r$ with respect to the coordinate system with which $Q(x) = \sum_{j=1}^r x_r^2$. A homogenous polynomial P of degree d for which $\Delta_Q P = 0$ is called a *spherical polynomial.* Then, one has the associated theta function

$$\Theta_{Q,P}(\tau) := \sum_{n\in\mathbb{Z}^r} P(n)q^{Q(n)}.$$

It turns out, as we shall see, that this is a modular form of weight $r/2 + d$.

2. Restricted theta functions

Broadly speaking, the final type of generalization of theta functions we consider, which is the most complicated, is when we sum over only part of a lattice instead of over a full lattice like $\mathbb{Z}^r$. There are two main reasons one may want to do this: the sum over the whole lattice may be 0 (and hence uninteresting), or it may be divergent. The first case, which often arises from summing theta functions twisted by odd Dirichlet characters, is frequently of interest in combinatorics. In this case, one can try to recover an interesting object by summing over only *half* of a lattice. The resulting *partial theta series* were studied by Rogers, who showed, for example, that

$$\psi(q) := \sum_{n\geq 0}\left(\frac{-12}{n}\right) q^{\frac{n^2-1}{24}} = \sum_{n\geq 0}\frac{(-1)^n q^{\frac{n(n+1)}{2}}}{(-q)_n},$$

where $(a)_n = (a;q)_n := \prod_{j=0}^{n-1}(1-aq^j)$. One situation in which these series often arise is in the context of expanding Appell-type functions in terms of geometric series and then taking coefficients of powers of ζ. For example, consider the product

$$\frac{1}{(\zeta^{-1}q)_\infty(\zeta q)_\infty} =: \frac{C(\zeta;q)}{(q)_\infty}, \tag{1}$$

which by the Jacobi triple product is essentially $\eta(\tau)/\vartheta(z;\tau)$, where $\eta(\tau) := q^{\frac{1}{24}}(q)_\infty$. Here, the function $C(\zeta;q)$ turns out to be a two-variable generating function for *cranks* of integer partitions, which Andrews and Garvan devised (solving a problem of Dyson) to "combinatorially" prove the Ramanujan congruences for the integer partition function modulo $5, 7$, and 11. The ζ-Fourier coefficients of (1) have an important representation-theoretic interpretation. Namely, in 2001, Kac and Wakimoto showed that these are essentially the characters of irreducible highest weight $s\ell(0|1)^\wedge$-modules. This function is essentially an Appell-Lerch sum, as the following classical identity shows (cf. Exercise 4):

$$C(\zeta;q) = \frac{(1-\zeta)}{(q)_\infty}\sum_{n\in\mathbb{Z}}\frac{(-1)^n q^{\frac{n(n+1)}{2}}}{1-\zeta q^n}. \tag{2}$$

Setting

$$\Phi(\zeta;q) := \frac{i\zeta^{1/2}q^{-1/24}}{\eta^2(\tau)(1-\zeta)}C(\zeta;q),$$

a geometric series expansion in the range $|q| < |\zeta| < 1$ computes the Kac-Wakimoto characters (cf. Exercise 5):

$$\Phi(\zeta;q) =: \sum_{n\in\frac{1}{2}+\mathbb{Z}}\chi(r;\tau)\zeta^n$$

with

$$\chi(r;\tau) = \frac{iq^{-r^2/2}}{\eta^3(\tau)}\sum_{m\geq 0}(-1)^m q^{\frac{(m+|r-1/2|+1/2)^2}{2}}. \tag{3}$$

Such partial theta functions do have remnants of modularity properties, although of a somewhat more complicated nature. They are what are known as *quantum modular forms*, which we shall return to in the context of our applications to Maass waveforms and knot theory in the fourth lecture. Essentially, these functions have modularity properties "near" roots of unity on the boundary of the usual lower half-plane, which can be seen as a type of "leaking" behavior from ordinary mock modular forms on the upper half-plane. Many aspects of such connections are very mysterious. In the case of ψ above, as Ramanujan knew, his mock theta functions have q-hypergeometric series expressions which can be formally inverted by sending $q \mapsto q^{-1}$. Although this process yields non-unique answers (different representations of the same q-series inside the unit circle invert to different functions outside the unit circle), the function $\psi(q)$

can essentially be obtained (as shown by Bringmann-Folsom-Rhoades) by inverting a suitable representation for Ramanujan's third order mock theta function

$$f(q) := \sum_{n\geq 0} \frac{q^{n^2}}{(-q)_n^2}.$$

For a discussion and hint for showing these claims about $f(q)$ and $\psi(q)$, see Exercise 3.

The other type of generalized theta function which we shall be interested in throughout the remainder of my talks, is *indefinite theta functions.* These allow the quadratic form to take negative values. As per Section 3 below, these also can be seen to arise from geometric series expansions of Appell-Lerch series like μ.

Returning to the general situation, call the *type* of a general non-degenerate form Q the pair $(r-s, s)$ where $r-s$ is the number of positive eigenvalues of the associated matrix A and s is the number of negative eigenvalues. In other words, s is the maximal dimension of a subspace of $\mathbb{R}^r$ on which Q is negative-definite. In this case, some powers of the form $q^{Q(n)}$ will become arbitrarily large in absolute value as n ranges over a full lattice. Thus, we need to either insert appropriate prefactors dampening the growth of these exponential terms (like the spherical polynomials above, but replaced with a suitable decaying function), restrict the summation to part of a lattice, or some combination of the two. We will later discuss what prefactors are allowed to preserve modularity, but restricting to a piece of a lattice usually destroys modularity. However, as Zwegers showed in the case of type $(r-1, 1)$ forms, this lack of modularity can be corrected to produce a harmonic Maass form. His functions are defined as follows, where we now restrict Q to be of type $(r-1, 1)$. First, we consider the cone (where c_0 with $Q(c_0) < 0$ is fixed)

$$C_Q := \{c \in \mathbb{R}^r : Q(c) < 0, B(c, c_0) < 0\}$$

and consider the set of cusps (cf. Exercise 6 for a related discussion)

$$S_Q := \{c = (c_1, c_2, \ldots, c_r) \in \mathbb{Z}^r : \gcd(c_1, c_2, \ldots, c_r) = 1,\ Q(c) = 0,\ B(c, c_0) < 0\}.$$

We then set $\overline{C}_Q := C_Q \cup S_Q$, which serves as the compactification of C_Q.

Definition 2.1. Let Q have signature $(r-1, 1)$ and let $c_1, c_2 \in \overline{C}_Q$. Then the **theta function** associated to Q with **characteristic** a, b is the series

$$\Theta_{a,b}(\tau) = \Theta_{a,b,c_1,c_2}(\tau) = \Theta_{A,a,b,c_1,c_2}(\tau) := \sum_{n\in a+\mathbb{Z}^r} \rho(n;\tau) q^{Q(n)} e(B(n, b)).$$

Here,

$$\rho(n;\tau) = \rho_Q^{c_1,c_2}(n;\tau) := \rho^{c_1}(n;\tau) - \rho^{c_2}(n;\tau)$$

with

$$\rho^c(n;\tau) := \begin{cases} E\left(\frac{B(c,n)v^{\frac{1}{2}}}{\sqrt{-Q(c)}}\right) & \text{if } c \in C_Q, \\ \operatorname{sgn}(B(c,n)) & \text{if } c \in S_Q, \end{cases}$$

where $E(z) := 2\int_0^z e^{-\pi t^2}dt$.

We also have a Jacobi form version of $\Theta_{a,b}$, given by

$$\Theta(z;\tau) = \Theta_{A,c_1,c_2}(z;\tau) := \sum_{n\in\mathbb{Z}^r} \rho(n+a;\tau)e(B(n,z))q^{Q(n)}.$$

Theorem (Zwegers)**.** *Assuming a few small (mainly convergence-related) conditions, if Q has signature $((r-1),1)$, then the indefinite theta series $\Theta(z;\tau)$ is a vector-valued Jacobi form of weight $r/2$ on $\mathrm{SL}_2(\mathbb{Z})$, and the indefinite theta functions with characteristic are pieces of vector-valued mixed harmonic Maass forms.*

Remark. *A brief discussion of vector-valued forms is given in Exercise 2.*

Such *mixed* harmonic Maass forms are essentially sums of products of ordinary modular forms with harmonic Maass forms, and will be discussed in Professor Guerzhoy's lectures. Moreover, the shadows of the harmonic Maass forms in these decompositions will be postitive-definite theta functions on $(r-1)$-dimensional lattices (as we will discuss in the second lecture). It is only occasionally that all modular forms in these mixed harmonic Maass decompositions are all essentially the same, meaning that they can be divided out to yield ordinary harmonic Maass forms.

In the next lecture, we will discuss the proofs of these facts, and explain why special functions like E arise, as well as how to compute the shadows of these functions as discussed in the last paragraph.

3. Relation between Zwegers' μ function and indefinite theta functions

The indefinite theta functions above can in fact be thought of as analogous to the Appell-Lerch series from Professor Kane's lectures. For example, we can expand the denominator of Zwegers' μ-function and find, after a little rewriting, the following expression of μ in terms of a type $(1,1)$ indefinite theta series. Namely, for $0 < \frac{y_1}{v}, \frac{y_2-y_1}{v} + \frac{1}{2} < 1$ (with $z_j = x_j + iy_j$), we have that

$$\widehat{\mu}(z_1,z_2;\tau) = \frac{\zeta_1^{\frac{1}{2}}}{2\vartheta(z_2;\tau)}\Theta_{A,c_1,c_2}\left(z_1, z_2 - z_1 + \frac{\tau+1}{2};\tau\right), \tag{4}$$

where $A = \left(\begin{smallmatrix}1&1\\1&0\end{smallmatrix}\right)$, $c_1 = \left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$, $c_2 = \left(\begin{smallmatrix}-1\\1\end{smallmatrix}\right)$ (cf. Exercise 7).

4. Exercises

(1) Complete the proof of Theorem 1 above. Specifically, show that

$$\theta(\tau+1) = \theta(\tau),$$

$$\theta\left(-\frac{1}{4\tau}\right) = \sqrt{\frac{2\tau}{i}}\theta(\tau).$$

(2) In general, modular forms on congruence subgroups can be shown to be components of vector-valued modular forms on the full modular group $\mathrm{SL}_2(\mathbb{Z})$. This is frequently useful in computations. For example, it can be used to dramatically decrease the number of coefficients needed to check a "Sturm bound" to prove identities between modular or mock modular forms. For a stunning application of this principal to the famous "Mock theta conjectures", see Andersen's paper with arxiv identifier 1604.05294. The following example is meant to give you a flavor of what vector-valued modular forms look like, and is typical of examples which Zwegers gave for many of his examples of mock theta functions.

Consider the theta function

$$\theta_1(\tau) := \sum_{n\in\mathbb{Z}} (-1)^n q^{n^2},$$

which can be shown to equal

$$\theta_1(\tau) = \frac{\eta^2(\tau)}{\eta(2\tau)}.$$

Now consider the vector-valued function

$$\Theta(\tau) := \begin{pmatrix} \theta_1(\tau) \\ \frac{\eta^2(\tau)}{\eta(\tau/2)} \\ \frac{\eta^2(\tau)}{\eta(\tau/2+1/2)} \end{pmatrix}.$$

Show that this transforms as a vector-valued modular form of weight 1/2 on $\mathrm{SL}_2(\mathbb{Z})$. Specifically, show that

$$\Theta(\tau+1) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & \zeta_{12} \\ 0 & \zeta_{24} & 0 \end{pmatrix} \Theta(\tau),$$

and

$$\Theta(-1/\tau) = \left(\frac{\tau}{i}\right)^{\frac{1}{2}} \begin{pmatrix} 0 & \sqrt{2} & 0 \\ 1/\sqrt{2} & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$

where $\zeta_n := e(1/n)$. Hint: Show, or use, the following identity:

$$\eta(\tau+1/2) = \zeta_{48}\frac{\eta^3(2\tau)}{\eta(\tau)\eta(4\tau)}.$$

(3) In this problem, you will fill in the details of the partial theta/mock theta $\psi(q), f(q)$ example above, following Bringmann-Folsom-Rhoades (see the paper with arxiv identifier 1109.6560). It can be shown (for example, in Fine's famous

book on "Basic hypergeometric series and applications") that for $|q|<1$, the function $\widetilde{f}(q)$ is equal to $f(q)$, where

$$\widetilde{f}(q) := 1 + \sum_{n\geq 1} \frac{(-1)^{n+1}q^n}{(-q)_n}.$$

This function also converges for $|q|>1$; show that if we send $q\mapsto q^{-1}$, we obtain

$$\widetilde{f}(q^{-1}) = 2\psi(q).$$

The same trick can be played for the definition of $f(q)$ above, but instead one obtains

$$f(q^{-1}) = 2\psi(q) - \frac{1}{(-q)_\infty^2}\sum_{n\geq 0}(-1)^n q^{\frac{n(n+1)}{2}}.$$

(4) Prove (2). Hint: This can be done using partial fractions, and specifically, via an application of the Mittag-Leffler theorem.
(5) Perform geometric series expansions directly to show (3).
(6) The reason that the set S_Q is called the cusps can be seen in analogy with an isomorphism with the usual modular curve $X_0(1)$. In this exercise, convince yourself that the following details work. Take $r=3$ and set $Q(X,Y,Z) := \frac{1}{2}(Y^2-4XZ)$. We choose our cone C_Q to be the one containing $c_0 := (1,0,1)^T$, and we compute

$$B([a,b,c],c_0) = -2a-2c.$$

In particular,

$$C_Q = \left\{[a,b,c]\in\mathbb{R}^3 : b^2<4ac,\ a+c>0\right\}.$$

We next show that $C_Q/\mathbb{R}^+$ is isomorphic to $\mathbb{H}$, where $\mathbb{R}^+$ acts on C_Q by rescaling the coefficients by positive reals. The map from $C_Q/\mathbb{R}^+$ to $\mathbb{H}$ is constructed as follows: Given $[a,b,c]\in C_Q$ with $D:=b^2-4ac=-1$, the root $\frac{-b+\sqrt{D}}{2a}$ of $Q(1,X)$ gives the desired point in $\mathbb{H}$. Now given a point $\tau=u+iv\in\mathbb{H}$, there is a unique point in C_Q with τ as a root and with discriminant -1, given by

$$\left[\frac{1}{2v}, -\frac{u}{v}, \frac{u^2+v^2}{2v}\right]$$

(note that the condition $a+c>0$ is then satisfied automatically, since $v>0$ implies that $a,c>0$). This establishes that the map $C_Q/\mathbb{R}^+\to\mathbb{H}$ is a bijection. To realize this map as an isomorphism of metric spaces, we may naturally define a distance on $C_Q/\mathbb{R}^+$ by setting

$$d\left([a_1,b_1,c_1],[a_2,b_2,c_2]\right) := \operatorname{arccosh}\left(-B\left([a_1,b_1,c_1],[a_2,b_2,c_2]\right)\right).$$

Now if

$$[a_j,b_j,c_j] = \left[\frac{1}{2v_j}, -\frac{u_j}{v_j}, \frac{u_j^2+v_j^2}{2v_j}\right]$$

for $j=1,2$, then we compute that

$$\operatorname{arccosh}\left(-B\left([a_1,b_1,c_1],[a_2,b_2,c_2]\right)\right)=\operatorname{arccosh}\left(1+\frac{(v_2-v_1)^2+(u_2-u_1)^2}{2v_1v_2}\right),$$

which is just the usual distance in the hyperbolic metric on $\mathbb{H}$. We now discuss the notion of cusps on our hyperbolic space. The cusps of $C_Q/\mathbb{R}^+$ can then be represented by

$$\left\{[a,b,c]\in\mathbb{Q}^3 : b^2=4ac,\ a+c>0\right\}/\mathbb{Q}^+.$$

In fact, this set maps to $\mathbb{P}^1(\mathbb{Q})$ under the same map we used to give the isomorphism $C_Q/\mathbb{R}^+\to\mathbb{H}$. Namely, this quotient space is the same as the set of binary quadratic forms $[a,b,c](X,Y)=aX^2+bXY+cY^2$ with discriminant 0 and with trace $a+c=1$, so the corresponding root of such a quadratic form is equal to $\frac{-b}{2a}$. Then given any $\alpha\in\mathbb{Q}$, the set of $[a,b,c]$ which map to α is given by those which solve the equations

$$b^2=4ac,\qquad \frac{-b}{2a}=\alpha,\qquad a+c=1,$$

which is uniquely solved by setting

$$[a,b,c]=\left[\frac{1}{\alpha^2+1},\frac{-2\alpha}{\alpha^2+1},\frac{\alpha^2}{\alpha^2+1}\right].$$

The cusp at infinity corresponds to $[a,b,c]=[0,0,1]$.

(7) Use geometric series and manipulation of the resulting expressions to show (4).