IMRN International Mathematics Research Notices
2003, No. 39

Almost Modular Functions and the Distribution of n^2x Modulo One

Jens Marklof

1 Introduction

It is well known that the sequence n^2x with $n = 1, 2, 3, 4, \ldots$ is equidistributed modulo one if x is irrational [22]. This means that, for every piecewise smooth function ψ of the circle $S^1 = \mathbb{R}/\mathbb{Z}$ to $\mathbb{C}$, we have

$$\frac{1}{N}\sum_{n=1}^{N}\psi(n^2x) \longrightarrow \int_0^1 \psi(t)dt \tag{1.1}$$

in the limit $N \to \infty$. Interesting choices for ψ are as follows:

(a) $\psi(t) = \chi_{[a,b]}(t)$, where $\chi_{[a,b]}$ is the indicator function of the interval $[a, b] + \mathbb{Z}$ on S^1, with $(b - a) \leq 1$;

(b) $\psi(t) = \{t\}$, where $\{t\}$ is the fractional part of t;

(c) $\psi(t) = e(t) := \exp(2\pi i t)$, leading to theta sums studied in [6, 7, 8, 14, 15];

(d) $\psi(t) = \log(1 - Ze(-t))$, for some $Z \in \mathbb{C}$, with $|Z| = 1$, and the sum in (1.1) becomes the logarithm of the polynomial

$$P_N(Z) := \prod_{n=1}^{N}\big(1 - Ze\big(-n^2x\big)\big). \tag{1.2}$$

The main objective of this work is to show that, for x uniformly distributed in $[0, 1]$, the

Received 31 January 2003.

fluctuations of the error term

$$E_{\psi}^{x}(N) := \sum_{n=1}^{N} \psi(n^2 x) - N \int_0^1 \psi(t)\,dt, \tag{1.3}$$

normalized by $1/\sqrt{N}$, have a limit distribution as $N \to \infty$, that is, there is a probability measure ν_ψ on $\mathbb{C}$ such that, for every bounded continuous function $g : \mathbb{C} \to \mathbb{C}$, we have

$$\lim_{N\to\infty} \int_0^1 g\left(\frac{E_{\psi}^{x}(N)}{\sqrt{N}}\right) dx = \int_{\mathbb{C}} g(w)\nu_\psi(dw). \tag{1.4}$$

The limit distribution can be expressed in terms of an almost modular function; in particular, it does not fall into the family of the classical stable limit laws. This is in contrast to the limit distribution of the error term for *lacunary* sequences, say $2^n x \bmod 1$, which is normal [9] (this result may in fact be viewed as a special case of the central limit theorem for dynamical systems [3]). Interestingly, the error term for the *linear* sequence $nx + y \bmod 1$, with $x, y \in [0, 1]$ random, has a limit distribution for the test function $\psi = \chi_{[a,b]}$ which is Cauchy and thus again stable [10, 11] (the normalization here is $1/\log N$, not $1/\sqrt{N}$).

It is very likely that the limit distribution of the error term of $n^2 x \bmod 1$ follows a stable limit law if the interval $[a, b]$ is no longer fixed but shrinks with $N \to \infty$. Of particular interest is the case when $(b - a)$ is of the order of the mean spacing $1/N$, where one expects a Poissonian limit distribution for the number of elements in $[a, b]$ (see [16, 17, 18] for details).

A nongeneric limit distribution has been observed as well for the error term in the classical circle problem [5] and more general lattice point counting problems in the plane [1, 2]. The limit distribution is, in these cases, given by almost periodic functions. Our proof of the limit theorem for almost modular functions in Section 8 is in fact modelled on that for almost periodic functions in [1].

2 Main results

It is natural to consider more general sums of the form

$$\frac{1}{\sqrt{N}} \sum_{n=1}^{\infty} f\left(\frac{n}{N}\right) \psi(n^2 x), \tag{2.1}$$

where f is a piecewise smooth cutoff function with compact support.

We think of the error term as a function $\Xi_{f,\psi} : \mathbb{C} \to \mathbb{C}$, where

$$\Xi_{f,\psi}(x+iy) = y^{1/4} \sum_{n=1}^{\infty} f(ny^{1/2})\psi(n^2x), \tag{2.2}$$

and $y = N^{-2}$.

Take $\psi \in L^2(S^1)$ real- or complex-valued with Fourier coefficients

$$\widehat{\psi}_k = \int_0^1 \psi(t)e(-kt)dt. \tag{2.3}$$

We assume in the following that (without loss of generality)

$$\widehat{\psi}_0 = 0, \tag{2.4}$$

and that there are constants $\beta > 1/2$ and $C(\psi) > 0$ such that

$$|\widehat{\psi}_k| \leq \frac{C(\psi)}{|k|^{\beta}}, \tag{2.5}$$

for all $k \neq 0$. These conditions are clearly satisfied for the examples (a), (b), (c), and (d) listed above.

We furthermore assume that $f \in PC_0^r(\mathbb{R}_+)$, the space of piecewise C^r functions $f : \mathbb{R}_+ \to \mathbb{R}$ with compact support supp_f ($\mathbb{R}_+$ includes the origin). *Piecewise* C^r means as usual that supp_f can be decomposed into finitely many intervals on each of which f is C^r and bounded.

For x uniformly distributed in $[0,1)$, $\Xi_{f,\psi}(x+iy)$ can be viewed as a family of random variables (parametrized by y) which are centered at expectation, that is,

$$\int_0^1 \Xi_{f,\psi}(x+iy)dx = 0. \tag{2.6}$$

We will see in Section 4 that the variance has a limit

$$\lim_{y\to 0} \int_0^1 \left|\Xi_{f,\psi}(x+iy)\right|^2 dx = \sigma^2(f,\psi), \tag{2.7}$$

where

$$\sigma^2(f,\psi) = \sum_{\substack{p,q=1 \\ \gcd(p,q)=1}}^{\infty} \int_0^{\infty} f(pr)f(qr)dr \int_0^1 \psi(p^2x)\overline{\psi}(q^2x)\,dx. \tag{2.8}$$

Our main result is the following.

Theorem 2.1. Let $f \in PC_0^\infty(\mathbb{R}_+)$ and $\psi \in L^2(S^1)$ satisfying (2.4) and (2.5). Then, for x uniformly distributed in $[0,1)$, $\Xi_{f,\psi}(x+iy)$ has a limit distribution as $y \to 0$. That is, there exists a probability measure $\nu_{f,\psi}$ on $\mathbb{C}$ such that, for any bounded continuous function $g : \mathbb{C} \to \mathbb{C}$,

$$\lim_{y\to 0} \int_0^1 g\big(\Xi_{f,\psi}(x+iy)\big)\,dx = \int_{\mathbb{C}} g(w)\nu_{f,\psi}(dw). \tag{2.9}$$

Furthermore, $\nu_{f,\psi}$ is symmetric with respect to $w \mapsto -w$. □

By establishing that $\Xi_{f,\psi}$ is almost modular (Section 11), this theorem follows directly from the limit theorem for almost modular functions (Section 8).

3 Decay of correlations

Lemma 3.1. For $\psi \in L^2(S^1)$ with (2.4) and (2.5),

$$\left| \int_0^1 \psi(ax)\overline{\psi}(bx)\,dx \right| \leq \sqrt{2\zeta(2\beta)}C(\psi)\|\psi\|_2 \frac{\gcd(a,b)^\beta}{b^\beta}, \tag{3.1}$$

for all $a, b \in \mathbb{N}$. (Here ζ denotes the Riemann zeta function.) □

Proof. Put $p = a/\gcd(a,b)$ and $q = b/\gcd(a,b)$. Then $\gcd(p,q) = 1$, and we have furthermore

$$\int_0^1 \psi(ax)\overline{\psi}(bx)\,dx = \int_0^1 \psi(px)\overline{\psi}(qx)\,dx. \tag{3.2}$$

Since $\psi \in L^2(S^1)$ and $\gcd(p,q) = 1$, we have

$$\int_0^1 \psi(px)\overline{\psi}(qx)\,dx = \sum_{\substack{k,l\neq 0\\ kp=lq}} \widehat{\psi}_k \overline{\widehat{\psi}_l} = \sum_{r\neq 0} \widehat{\psi}_{rq}\overline{\widehat{\psi}_{rp}}. \tag{3.3}$$

By the Cauchy-Schwartz inequality, the modulus of this last expression is less than or equal to

$$\left(\sum_{r\neq 0} |\widehat{\psi}_{rq}|^2\right)^{1/2} \left(\sum_{r\neq 0} |\widehat{\psi}_{rp}|^2\right)^{1/2} \leq \frac{C(\psi)}{q^\beta} \left(\sum_{r\neq 0} |r|^{-2\beta}\right)^{1/2} \|\psi\|_2, \tag{3.4}$$

which proves the claim. ■

Of course equation (3.3) also implies the bound

$$\left|\int_0^1 \psi(ax)\overline{\psi}(bx)\,dx\right| \leq 2\zeta(2\beta)C(\psi)^2\frac{\gcd(a,b)^{2\beta}}{(ab)^\beta} \tag{3.5}$$

which decays faster for a, b large. This, however, will be of no direct advantage, and the explicit dependence on $\|\psi\|_2$ in Lemma 3.1 will make the argument more transparent.

4 The variance

Lemma 4.1. There is a constant $K_\beta > 0$ such that

$$\limsup_{N\to\infty}\frac{1}{N}\int_0^1\left|\sum_{n=1}^{\infty} f\left(\frac{n}{N}\right)\psi(n^2x)\right|^2 dx \leq \sup\,(f^2)\,|\operatorname{supp}_f|\left(\|\psi\|_2^2 + K_\beta C(\psi)\|\psi\|_2\right) \tag{4.1}$$

holds uniformly for all $f \in PC_0(\mathbb{R}_+)$ and all $\psi \in L^2(S^1)$, satisfying (2.4) and (2.5). □

Proof. We have

$$\begin{aligned}\int_0^1\left|\sum_{n=1}^{\infty} f\left(\frac{n}{N}\right)\psi(n^2x)\right|^2 dx \\ = \sum_{n=1}^{\infty} f\left(\frac{n}{N}\right)^2\|\psi\|_2^2 + 2\,\mathrm{Re}\sum_{1\leq m<n} f\left(\frac{m}{N}\right)f\left(\frac{n}{N}\right)\int_0^1 \psi(m^2x)\overline{\psi}(n^2x)\,dx\end{aligned} \tag{4.2}$$

since

$$\int_0^1 \psi(n^2x)\overline{\psi}(n^2x)\,dx = \|\psi\|_2^2. \tag{4.3}$$

For $x \in \mathbb{R}$ and $S \subset \mathbb{R}$, denote by xS the set $\{xy : y \in S\}$. For the first term in (4.2), we then have

$$\left|\sum_{n=1}^{\infty} f\left(\frac{n}{N}\right)^2\right| \leq \sup\,(f^2)\#\{n \in \mathbb{N}\cap N\operatorname{supp}_f\} \leq \sup\,(f^2)N|\operatorname{supp}_f| + O_f(1). \tag{4.4}$$

We rewrite the second term in (4.2) as

$$\begin{aligned}&2\,\mathrm{Re}\sum_{\substack{1\leq p<q\\ \gcd(p,q)=1}}\sum_{r=1}^{\infty} f\left(\frac{pr}{N}\right)f\left(\frac{qr}{N}\right)\int_0^1 \psi(r^2p^2x)\overline{\psi}(r^2q^2x)\,dx \\ &= 2\,\mathrm{Re}\sum_{\substack{1\leq p<q\\ \gcd(p,q)=1}}\sum_{r=1}^{\infty} f\left(\frac{pr}{N}\right)f\left(\frac{qr}{N}\right)\int_0^1 \psi(p^2x)\overline{\psi}(q^2x)\,dx.\end{aligned} \tag{4.5}$$

Now

$$\left|\sum_{r=1}^{\infty} f\left(\frac{pr}{N}\right) f\left(\frac{qr}{N}\right)\right| \leq \sup\,(f^2)\#\left\{r \in \mathbb{N} \cap \left(\frac{N}{p}\operatorname{supp}_f\right) \cap \left(\frac{N}{q}\operatorname{supp}_f\right)\right\}$$
$$\leq \sup\,(f^2)\#\left\{r \in \mathbb{N} \cap \left(\frac{N}{q}\operatorname{supp}_f\right)\right\} \begin{cases} \leq \sup\,(f^2)\dfrac{N}{q}|\operatorname{supp}_f| + O_f(1) & \text{if } q \leq N\ell(f), \\ = 0 & \text{if } q > N\ell(f), \end{cases} \tag{4.6}$$

where $\ell(f)$ is the length of the shortest interval containing supp_f. The modulus of (4.5) is thus less than or equal to

$$\begin{aligned} &2N \sup\,(f^2)|\operatorname{supp}_f| \sum_{\substack{1\leq p<q\leq N\ell(f)\\ \gcd(p,q)=1}} \frac{1}{q}\left|\int_0^1 \psi(p^2x)\overline{\psi}(q^2x)\,dx\right| \\ &\qquad + 2O_f(1) \sum_{\substack{1\leq p<q\leq N\ell(f)\\ \gcd(p,q)=1}} \left|\int_0^1 \psi(p^2x)\overline{\psi}(q^2x)\,dx\right|. \end{aligned} \tag{4.7}$$

By Lemma 3.1, we have for the first sum

$$\sum_{\substack{1\leq p<q\leq N\ell(f)\\ \gcd(p,q)=1}} \frac{1}{q}\left|\int_0^1 \psi(p^2x)\overline{\psi}(q^2x)\,dx\right| \leq \sqrt{2\zeta(2\beta)}C(\psi)\|\psi\|_2 \sum_{\substack{1\leq p<q\\ \gcd(p,q)=1}} \frac{1}{q^{1+2\beta}}, \tag{4.8}$$

which converges for $\beta > 1/2$. Similarly, for the second sum in (4.7), assuming, without loss of generality, that $1/2 < \beta < 1$,

$$\begin{aligned} \sum_{\substack{1\leq p<q\leq N\ell(f)\\ \gcd(p,q)=1}} \left|\int_0^1 \psi(p^2x)\overline{\psi}(q^2x)\,dx\right| &\leq \sqrt{2\zeta(2\beta)}C(\psi)\|\psi\|_2 \sum_{\substack{1\leq p<q\leq N\ell(f)\\ \gcd(p,q)=1}} \frac{1}{q^{2\beta}} \\ &\leq C(\psi)\|\psi\|_2 O_\beta\left((N\ell(f))^{2-2\beta}\right). \end{aligned} \tag{4.9}$$

■

Lemma 4.2. Let $f \in PC_0(\mathbb{R}_+)$ and $\psi \in L^2(S^1)$, satisfying (2.4) and (2.5). Then

$$\lim_{N\to\infty} \frac{1}{N}\int_0^1 \left|\sum_{n=1}^{\infty} f\left(\frac{n}{N}\right)\psi(n^2x)\right|^2 dx = \sigma^2(f,\psi), \tag{4.10}$$

with

$$\sigma^2(f,\psi) = \sum_{\substack{p,q=1\\ \gcd(p,q)=1}}^{\infty} \int_0^{\infty} f(pr)f(qr)\,dr \int_0^1 \psi(p^2x)\overline{\psi}(q^2x)\,dx. \tag{4.11}$$

□

Proof. We have

$$\frac{1}{N}\int_0^1\left|\sum_{n=1}^{\infty} f\left(\frac{n}{N}\right)\psi(n^2x)\right|^2 dx = \sum_{\substack{p,q=1\\ \gcd(p,q)=1}}^{\infty} a_N(p,q), \tag{4.12}$$

with

$$a_N(p,q) = \frac{1}{N}\sum_{r=1}^{\infty} f\left(\frac{pr}{N}\right) f\left(\frac{qr}{N}\right)\int_0^1 \psi(p^2x)\overline{\psi}(q^2x)\,dx. \tag{4.13}$$

Next

$$\lim_{N\to\infty}\frac{1}{N}\sum_{r=1}^{\infty} f\left(\frac{pr}{N}\right) f\left(\frac{qr}{N}\right) = \int_0^{\infty} f(pr)f(qr)\,dr, \tag{4.14}$$

for p, q fixed, implies

$$\lim_{N\to\infty} a_N(p,q) = a(p,q) := \int_0^{\infty} f(pr)f(qr)\,dr \int_0^1 \psi(p^2x)\overline{\psi}(q^2x)\,dx. \tag{4.15}$$

It follows from the proof of Lemma 4.1 that there is a function $g(p,q)$ such that

$$\left|a_N(p,q)\right| \le g(p,q), \qquad \sum_{\substack{p,q=1\\ \gcd(p,q)=1}}^{\infty} g(p,q) < \infty. \tag{4.16}$$

Hence the dominated convergence theorem yields

$$\lim_{N\to\infty}\sum_{\substack{p,q=1\\ \gcd(p,q)=1}}^{\infty} a_N(p,q) = \sum_{\substack{p,q=1\\ \gcd(p,q)=1}}^{\infty} a(p,q). \tag{4.17}$$

■

5 Universal cover of $SL(2,\mathbb{R})$ and discrete subgroups

The action of $SL(2,\mathbb{R})$ on the upper half plane $\mathfrak{H} = \{z \in \mathbb{C} : \operatorname{Im} z > 0\}$ is given by fractional linear transformations, that is,

$$g : z \longmapsto gz = \frac{az+b}{cz+d}, \quad g = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in SL(2,\mathbb{R}). \tag{5.1}$$

We can define the continuous function $\varepsilon_g : \mathfrak{H} \to \mathbb{C}$ by $\varepsilon_g(z) = (cz+d)/|cz+d|$. One easily verifies that $\varepsilon_{gh}(z) = \varepsilon_g(hz)\,\varepsilon_h(z)$. In the following, we will denote by $C(\mathfrak{H})$ the space of continuous functions $\mathfrak{H} \to \mathbb{C}$. The universal covering group of $\mathrm{SL}(2,\mathbb{R})$ is defined as the set

$$\widetilde{\mathrm{SL}}(2,\mathbb{R}) = \left\{ [g, \beta_g] : g \in \mathrm{SL}(2,\mathbb{R}),\ \beta_g \in C(\mathfrak{H}) \text{ such that } e^{i\beta_g(z)} = \varepsilon_g(z) \right\}, \tag{5.2}$$

with multiplication law

$$[g, \beta_g^1][h, \beta_h^2] = [gh, \beta_{gh}^3], \qquad \beta_{gh}^3(z) = \beta_g^1(hz) + \beta_h^2(z). \tag{5.3}$$

We may identify $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ with $\mathfrak{H} \times \mathbb{R}$ via $[g, \beta_g] \mapsto (z, \phi) = (gi, \beta_g(i))$. The action of $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ on $\mathfrak{H} \times \mathbb{R}$ is then canonically defined by $[g, \beta_g](z, \phi) = (gz, \phi + \beta_g(z))$. The Haar measure of $\widetilde{\mathrm{SL}}(2,\mathbb{R})$ reads, in this parametrization,

$$d\mu(g) = \frac{dx\, dy\, d\phi}{y^2}. \tag{5.4}$$

For any integer $m > 0$, put

$$Z_m = \left\langle [-1, \beta_{-1}]^m \right\rangle, \quad \text{with } \beta_{-1}(z) = \pi, \tag{5.5}$$

that is, Z_m is the subgroup generated by the element $[-1, \beta_{-1}]^m$. The subgroup Z_m is contained in the center of $\widetilde{\mathrm{SL}}(2,\mathbb{R})$, and it is easily seen that $\mathrm{PSL}(2,\mathbb{R})$ is isomorphic to $\widetilde{\mathrm{SL}}(2,\mathbb{R})/Z_1$, and $\mathrm{SL}(2,\mathbb{R})$ is isomorphic to $\widetilde{\mathrm{SL}}(2,\mathbb{R})/Z_2$.

For any positive integer N, we define the congruence subgroups of $\mathrm{SL}(2,\mathbb{Z})$:

$$\Gamma_1(N) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{SL}(2,\mathbb{Z}) : a \equiv d \equiv 1,\ c \equiv 0 \bmod N \right\}, \tag{5.6}$$

and the following lift to the universal cover (assume now N is divisible by 4):

$$\Delta_1(N) = \left\{ [\gamma, \beta_\gamma] : \gamma \in \Gamma_1(N),\ \beta_\gamma \in C(\mathfrak{H}) \text{ such that } e^{i\beta_\gamma(z)/2} = j_\gamma(z) \right\}, \tag{5.7}$$

where

$$j_\gamma(z) = \left(\frac{c}{d}\right) \left(\frac{cz+d}{|cz+d|}\right)^{1/2}, \quad \gamma = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \Gamma_1(4). \tag{5.8}$$

Here $z^{1/2}$ denotes the principal branch of the square root of z, that is, the one for which $-\pi/2 < \arg z^{1/2} \leq \pi/2$; and $(\frac{c}{d})$ denotes the generalized quadratic residue symbol (see Appendix for details).

It is well known that j_γ forms a multiplier system for $\Gamma_1(4)$, that is, $j_{\gamma\eta}(z) = j_\gamma(\eta z)j_\eta(z)$ for all $\gamma, \eta \in \Gamma_1(4)$ (and hence for all $\gamma, \eta \in \Gamma_1(N) \subset \Gamma_1(4)$; recall that $4|N$). Therefore $\Delta_1(N)$ is indeed a subgroup of $\widetilde{SL}(2,\mathbb{R})$ if $4|N$.

We collect a few important properties which will be needed later on. For $y > 0$, we define

$$a_y = \begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix} \in SL(2,\mathbb{R}), \qquad A_y = \left[a_y, 0\right] \in \widetilde{SL}(2,\mathbb{R}). \tag{5.9}$$

Lemma 5.1. Assume N, N_1, and N_2 are positive integers divisible by 4, and k is any positive integer. Then

(a) $\Delta_1(N)$ is a finite index subgroup of $\Delta_1(4)$;
(b) $\Delta_1(4k) \subset A_k^{-1}\Delta_1(4)A_k$;
(c) $\Delta_1(\mathrm{lcm}(N_1,N_2)) \subset \Delta_1(N_1) \cap \Delta_1(N_2)$;
(d) $\mathcal{M}_N = \Delta_1(N)\backslash\widetilde{SL}(2,\mathbb{R})$ is a noncompact manifold of finite measure (with respect to Haar measure μ). □

Proof. For any integer N' divisible by 4, $\Delta_1(N')$ contains the subgroup $Z_4 = \{[1,\beta_1] : \beta_1(z) = 4\pi n, n \in \mathbb{Z}\}$, and $\Delta_1(N')/Z_4$ is isomorphic to $\Gamma_1(N')$. This proves (a).

A short calculation shows that

$$A_k \left[\begin{pmatrix} a & b \\ c & d \end{pmatrix}, \beta \right] A_k^{-1} = \left[\begin{pmatrix} a & kb \\ c/k & d \end{pmatrix}, \tilde{\beta} \right], \tag{5.10}$$

with $\tilde{\beta}(z) = \beta(z/k)$. Hence, if $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \Gamma_1(4k)$, then $a_k \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) a_k^{-1} = \left(\begin{smallmatrix} a & kb \\ c/k & d \end{smallmatrix}\right) \in \Gamma_1(4)$. Second, we need to show that

$$e^{i\tilde{\beta}(z)/2} = \left(\frac{(c/k)}{d}\right)\left(\frac{(c/k)z + d}{|(c/k)z + d|}\right)^{1/2} \tag{5.11}$$

holds. To this end, note that

$$e^{i\tilde{\beta}(z)/2} = e^{i\beta(z/k)/2} = \left(\frac{c}{d}\right)\left(\frac{cz/k + d}{|cz/k + d|}\right)^{1/2} \tag{5.12}$$

and that (using multiplicativity)

$$\left(\frac{c}{d}\right) = \left(\frac{(c/k)}{d}\right)\left(\frac{k}{d}\right). \tag{5.13}$$

Now $(\frac{k}{\cdot})$ is a character mod 4k and hence, for $d \equiv 1 \bmod 4k$, we have

$$\left(\frac{k}{d}\right) = \left(\frac{k}{1}\right) = 1. \tag{5.14}$$

This proves (b). Statement (c) is clear. Since $\mathrm{SL}(2,\mathbb{R})$ is isomorphic to $\widetilde{\mathrm{SL}}(2,\mathbb{R})/Z_2$, (d) follows from its analog for $\Gamma_1(N)\backslash \mathrm{SL}(2,\mathbb{R})$. ■

Because Z_4 is of index two in Z_2, $\Delta_1(N)\backslash\widetilde{\mathrm{SL}}(2,\mathbb{R})$ is in fact a double cover of $\Gamma_1(N)\backslash \mathrm{SL}(2,\mathbb{R})$. A fundamental domain for the action of $\Delta_1(N)$ on $\mathfrak{H}\times\mathbb{R}$ is $\mathcal{F}_{\Delta_1(N)} = \mathcal{F}_{\Gamma_1(N)}\times[0,4\pi)$ if $\mathcal{F}_{\Gamma_1(N)}$ is a fundamental region of $\Gamma_1(N)$ in $\mathfrak{H}$.

6 Equidistribution of closed horocycles

The manifold $\mathcal{M}_N$ has a finite number of cusps which are represented by the set $\eta_1,\ldots,\eta_\kappa \in \mathbb{Q}\cup\infty$ on the boundary of $\mathfrak{H}$. Let $\gamma_i \in \mathrm{PSL}(2,\mathbb{R})$ be a fractional linear transformation which maps the cusp at η_i to the standard cusp at ∞ of width one. Thus $(z_i,\phi_i) = \widetilde{\gamma}_i(z,\phi)$ yields a new set of coordinates, where the ith cusp appears as a cusp at ∞, which is invariant under $(z_i,\phi_i)\mapsto(z_i+1,\phi_i)$. The variable $y_i = \mathrm{Im}(\gamma_i z)$ measures the height into the ith cusp.

For any $\sigma \geq 0$, we denote by $B_\sigma(\mathcal{M}_N)$ the class of functions $F \in C(\mathcal{M}_N)$ such that, for all $i = 1,\ldots,\kappa$,

$$F(z,\phi) = O\big(y_i^\sigma\big), \quad y_i \longrightarrow \infty, \tag{6.1}$$

where the implied constant is independent of (z,ϕ). In view of the form of the invariant measure (5.4), we note that $B_\sigma(\mathcal{M}_N) \subset L^p(\mathcal{M}_N,\mu)$ if $\sigma < 1/p$.

Theorem 6.1. Let $0 \leq \sigma < 1$. Then, for every $F \in B_\sigma(\mathcal{M}_N)$,

$$\lim_{y\to 0}\int_0^1 F(x+iy,0)\,dx = \frac{1}{\mu(\mathcal{M}_N)}\int_{\mathcal{M}_N} F\,d\mu. \tag{6.2}$$

□

Proof. There are several ways to prove this theorem. One possibility is to use Eisenstein series of half-integral weight as in [15] which is based on Sarnak's approach [19]. The second variant is to use the mixing property of the flow

$$\Phi^t : \widetilde{\mathrm{SL}}(2,\mathbb{R}) \longrightarrow \widetilde{\mathrm{SL}}(2,\mathbb{R}), \qquad g \longmapsto g\begin{pmatrix} e^{-t/2} & 0 \\ 0 & e^{t/2}\end{pmatrix} \tag{6.3}$$

as in [4]. A further possibility is to quote Shah's theorem [20] on the distribution of translates of unipotent orbits. All three methods assume that F is bounded. The extension to $F \in B_\sigma(\mathcal{M}_N)$ is achieved by the argument given in [14, the proof of Proposition 4.3]. ■

7 Almost modular functions

In the following, we will consider functions $\Xi : \mathfrak{H} \to \mathbb{C}$ which are *periodic*, that is, for which $\Xi(z+1) = \Xi(z)$.

Definition 7.1. For any $p \geq 1$, let $\mathcal{B}^p$ be the class of periodic functions $\Xi : \mathfrak{H} \to \mathbb{C}$ with the property that for every $\varepsilon > 0$, there are an integer $N = N(\varepsilon) > 0$ and a function $F_\varepsilon \in B_\sigma(\mathcal{M}_N)$ with $0 \leq \sigma < 1/p$ so that

$$\limsup_{y \to 0} \int_0^1 \left|\Xi(x+iy) - F_\varepsilon(x+iy, 0)\right|^p dx < \varepsilon^p. \tag{7.1}$$

We will see below that the error term (2.2) falls into the class $\mathcal{B}^2$. A further example of an almost modular function of this type is

$$(\operatorname{Im} z)^{1/4} \log \prod_{n=1}^{\infty} \left(1 - e(n^2 z)\right), \tag{7.2}$$

which is discussed in more detail in [16].

Definition 7.2. Let $\mathcal{H}$ be the class of periodic functions $\Xi : \mathfrak{H} \to \mathbb{C}$ with the property that for every $\varepsilon > 0$, there are an integer $N = N(\varepsilon) > 0$ and a bounded continuous function $F_\varepsilon \in C(\mathcal{M}_N)$ such that

$$\limsup_{y \to 0} \int_0^1 \min\left\{1, \left|\Xi(x+iy) - F_\varepsilon(x+iy, 0)\right|\right\} dx < \varepsilon. \tag{7.3}$$

We will call functions in $\mathcal{B}^p$ or $\mathcal{H}$ *almost modular functions of class* $\mathcal{B}^p$ or $\mathcal{H}$, respectively.

Proposition 7.3. If $1 \leq q \leq p$, then

$$\mathcal{B}^p \subset \mathcal{B}^q \subset \mathcal{H}. \tag{7.4}$$

□

Proof. Hölder's inequality implies that if $f \in L^r(S^1)$, then $f \in L^1(S^1)$ and $\int_0^1 |f| dx \leq (\int_0^1 |f|^r dx)^{1/r}$. We put $f(x) = |\Xi(x+iy) - F_\varepsilon(x+iy, 0)|^q$ and $r = p/q$. Then

$$\int_0^1 \left|\Xi(x+iy) - F_\varepsilon(x+iy, 0)\right|^q dx \leq \left(\int_0^1 \left|\Xi(x+iy) - F_\varepsilon(x+iy, 0)\right|^p dx\right)^{q/p}. \tag{7.5}$$

Therefore, if (7.1) holds for p, it also holds for q, in fact with the same ε and F_ε.

To prove the second inclusion, it is enough to show that $\mathcal{B}^1 \subset \mathcal{H}$. Hence assume $\Xi \in \mathcal{B}^1$; we may then choose $F \in B_\sigma(\mathcal{M}_N)$ so that

$$\limsup_{y\to 0} \int_0^1 \left|\Xi(x+iy) - F(x+iy,0)\right| dx < \frac{\varepsilon}{2}. \tag{7.6}$$

We furthermore find a bounded continuous $F_\varepsilon \in C(\mathcal{M}_N)$ such that

$$\frac{1}{\mu(\mathcal{M}_N)} \int_{\mathcal{M}_N} \left|F - F_\varepsilon\right| d\mu < \frac{\varepsilon}{2}. \tag{7.7}$$

Then

$$\begin{aligned} \limsup_{y\to 0} \int_0^1 \min\left\{1, \left|\Xi(x+iy) - F_\varepsilon(x+iy,0)\right|\right\} dx \\ \leq \limsup_{y\to 0} \int_0^1 \left|\Xi(x+iy) - F(x+iy,0)\right| dx \\ + \limsup_{y\to 0} \int_0^1 \left|F(x+iy,0) - F_\varepsilon(x+iy,0)\right| dx. \end{aligned} \tag{7.8}$$

The first term is bounded by (7.6) and the second term converges to (7.7) by Theorem 6.1 since $|F - F_\varepsilon| \in B_\sigma(\mathcal{M}_N)$. ■

8 Limit theorems for almost modular functions

In this section, we follow Bleher's approach [1] for almost periodic functions. The main difference is that the equidistribution of irrational translations on tori is replaced by the equidistribution of closed horocycles on $\mathcal{M}_N$.

Proposition 8.1. If $\Xi \in \mathcal{B}^p$ and the approximants in Definition 7.1 satisfy

$$\frac{1}{\mu(\mathcal{M}_N)} \int_{\mathcal{M}_N} \left|F_\varepsilon\right|^p d\mu \leq R \tag{8.1}$$

for some constant $R > 0$, then

$$\|\Xi\|_{\mathcal{B}^p} := \left(\lim_{y\to 0} \int_0^1 \left|\Xi(x+iy)\right|^p dx \right)^{1/p} \tag{8.2}$$

exists. □

Proof. Minkowski's inequality and (7.1) yield, for all $0 < y < y_0(\varepsilon)$ small enough,

$$\left(\int_0^1 \left|\Xi(x+iy)\right|^p dx\right)^{1/p} < \left(\int_0^1 \left|F_\varepsilon(x+iy,0)\right|^p dx\right)^{1/p} + \varepsilon \tag{8.3}$$

and also

$$\left(\int_0^1 \left|F_\varepsilon(x+iy,0)\right|^p dx\right)^{1/p} < \left(\int_0^1 \left|\Xi(x+iy)\right|^p dx\right)^{1/p} + \varepsilon. \tag{8.4}$$

By Theorem 6.1, we then see that

$$\begin{aligned} \limsup_{y\to 0} \left(\int_0^1 \left|\Xi(x+iy)\right|^p dx\right)^{1/p} &< \left(\frac{1}{\mu(\mathcal{M}_N)}\int_{\mathcal{M}_N} \left|F_\varepsilon\right|^p d\mu\right)^{1/p} + \varepsilon, \\ \liminf_{y\to 0} \left(\int_0^1 \left|\Xi(x+iy)\right|^p dx\right)^{1/p} &> \left(\frac{1}{\mu(\mathcal{M}_N)}\int_{\mathcal{M}_N} \left|F_\varepsilon\right|^p d\mu\right)^{1/p} - \varepsilon. \end{aligned} \tag{8.5}$$

With condition (8.1), the upper and lower limit are arbitrarily close to the same constant $\leq R < \infty$. ■

Theorem 8.2. Let $\Xi \in \mathcal{H}$. Then, for x uniformly distributed in $[0,1)$, $\Xi(x+iy)$ has a limit distribution as $y \to 0$. That is, there exists a probability measure ν_Ξ on $\mathbb{C}$ such that, for every bounded continuous function $g : \mathbb{C} \to \mathbb{C}$,

$$\lim_{y\to 0} \int_0^1 g\big(\Xi(x+iy)\big)\, dx = \int_{\mathbb{C}} g(w)\nu_\Xi(dw). \tag{8.6}$$

□

We split the proof into two lemmas. We denote by ρ_y the distribution of the random variable $\Xi(x+iy)$, where y is fixed and x is uniformly distributed in $[0,1)$. We need to show that ρ_y converges weakly to some probability measure ν_Ξ.

Lemma 8.3. The family $\{\rho_y : 0 < y \leq 1\}$ is relatively compact. (I.e., every sequence of ρ_y has a weakly convergent subsequence.) □

Proof. We need to show that the family is tight, that is, for every $\varepsilon > 0$, there is a constant $K_\varepsilon > 0$ such that

$$\int_{|w|>K_\varepsilon} \rho_y(dw) = \left|\left\{x \in [0,1) : \left|\Xi(x+iy)\right| > K_\varepsilon\right\}\right| < \varepsilon \tag{8.7}$$

uniformly for $0 < y \leq 1$. To prove this, we start with the inequality

$$\begin{aligned} &\left|\left\{x \in [0,1) : \left|\Xi(x+iy)\right| > K_\varepsilon\right\}\right| \\ &\quad \leq \left|\left\{x \in [0,1) : \left|F_\varepsilon(x+iy,0)\right| > K_\varepsilon - 1\right\}\right| \\ &\quad\quad + \left|\left\{x \in [0,1) : \left|\Xi(x+iy) - F_\varepsilon(x+iy,0)\right| \geq 1\right\}\right|, \end{aligned} \tag{8.8}$$

where F_ε is an approximant as in Definition 7.2. So for the choice $K_\varepsilon = 1 + \sup_{\mathcal{M}_N} F_\varepsilon$, the first term is not present. From (7.3), we have, for all $0 < y < y_1(\varepsilon)$ small enough,

$$\int_0^1 \min\left\{1, \left|\Xi(x+iy) - F_\varepsilon(x+iy,0)\right|\right\} dx < \varepsilon, \tag{8.9}$$

which gives the desired upper bound for the second term in (8.8). In the range $y_1(\varepsilon) \le y \le 1$, relation (8.7) follows simply from the measurability of $\Xi(\cdot + iy)$. So (8.7) indeed holds uniformly for $0 \le y \le 1$.

The lemma now follows from the Helly-Prokhorov theorem [21] which asserts that every tight family is relatively compact. ∎

Lemma 8.4. For every $g \in C_0^\infty(\mathbb{C})$, the limit

$$I(g) := \lim_{y \to 0} \int_0^1 g\big(\Xi(x+iy)\big)\, dx \tag{8.10}$$

exists. □

Proof. Since $g \in C_0^\infty(\mathbb{C})$, we have

$$\left|g(w) - g(w')\right| \le C \min\left\{1, |w - w'|\right\} \tag{8.11}$$

for some $C > 0$. Hence

$$\begin{aligned} &\int_0^1 \left|g\big(\Xi(x+iy)\big) - g\big(F_\varepsilon(x+iy,0)\big)\right| dx \\ &\qquad \le C \int_0^1 \min\left\{1, \left|\Xi(x+iy) - F_\varepsilon(x+iy,0)\right|\right\} dx < C\varepsilon \end{aligned} \tag{8.12}$$

for $y < y_1(\varepsilon)$, as in (8.9).

Next we observe that, since $g \circ F_\varepsilon \in B_0(\mathcal{M}_N)$, Theorem 6.1 says that the sequence

$$\int_0^1 g\big(F_\varepsilon(x+iy,0)\big)\, dx \tag{8.13}$$

converges as $y \to 0$ and is therefore a Cauchy sequence. So for all $0 < y', y'' < y_2(\varepsilon, F_\varepsilon)$ small enough, we have

$$\left| \int_0^1 g\big(F_\varepsilon(x+iy',0)\big)\, dx - \int_0^1 g\big(F_\varepsilon(x+iy'',0)\big)\, dx \right| < \varepsilon. \tag{8.14}$$

Together with (8.12), this yields

$$\left| \int_0^1 g\big(\Xi(x+iy')\big)\,dx - \int_0^1 g\big(\Xi(x+iy'')\big)\,dx \right| < (2C+1)\varepsilon, \tag{8.15}$$

for $0 < y', y'' < \min\{y_1(\varepsilon), y_2(\varepsilon, F_\varepsilon)\}$, and thus $\int_0^1 g\big(\Xi(x+iy)\big)\,dx$ is a Cauchy sequence. ■

Proof of Theorem 8.2. For $g \in C_0^\infty(\mathbb{C})$, Lemma 8.3 shows that the limit in Lemma 8.4 is

$$I(g) = \int_{\mathbb{C}} g(w)\nu_\Xi(dw). \tag{8.16}$$

The theorem now follows for more general bounded continuous g from a standard approximation argument. ■

9 Shale-Weil representation and theta sums

For every $g \in SL(2,\mathbb{R})$, we have the unique Iwasawa decomposition

$$g = n_x a_y k_\phi = (z, \phi), \tag{9.1}$$

where $z = x + iy \in \mathfrak{H}$, $\phi \in [0, 2\pi)$,

$$n_x = \begin{pmatrix} 1 & x \\ 0 & 1 \end{pmatrix}, \qquad a_y = \begin{pmatrix} y^{1/2} & 0 \\ 0 & y^{-1/2} \end{pmatrix}, \qquad k_\phi = \begin{pmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{pmatrix}. \tag{9.2}$$

This can be extended to an Iwasawa decomposition of $\widetilde{SL}(2,\mathbb{R})$, which of course corresponds to the parametrization introduced after (5.3). We have, for any element $M = [g, \beta_g] \in \widetilde{SL}(2,\mathbb{R})$,

$$M = [g, \beta_g] = N_x A_y K_\phi = [n_x, 0][a_y, 0][k_\phi, \beta_{k_\phi}]. \tag{9.3}$$

The Shale-Weil representation is usually defined as a projective representation of $SL(2,\mathbb{R})$, which becomes a true representation on the metaplectic (i.e., double) cover of $SL(2,\mathbb{R})$. Therefore it is also a proper representation of the universal cover $\widetilde{SL}(2,\mathbb{R})$. In view of the decomposition (9.3), it is sufficient to define the representation on the three factors. For any Schwartz function $f \in \mathcal{S}(\mathbb{R})$, we set (cf. [12])

$$[R(N_x)f](t) = e(t^2 x)f(t), \qquad [R(A_y)f](t) = y^{1/4} f(y^{1/2}t), \tag{9.4}$$

and

$$[R(K_\phi)f](t) = \begin{cases} e\left(-\frac{\sigma_\phi}{8}\right) f(t) & (\phi = 0 \bmod 2\pi), \\ e\left(-\frac{\sigma_\phi}{8}\right) f(-t) & (\phi = \pi \bmod 2\pi), \\ e\left(-\frac{\sigma_\phi}{8}\right) 2^{1/2} |\sin\phi|^{-1/2} \int_{\mathbb{R}} e\left[\frac{(t^2 + t'^2)\cos\phi - tt'}{\sin\phi}\right] f(t')\,dt' & (\phi \neq 0 \bmod \pi), \end{cases} \tag{9.5}$$

where

$$\sigma_\phi = \begin{cases} 2\nu & \text{if } \phi = \nu\pi, \\ 2\nu + 1 & \text{if } \nu\pi < \phi < (\nu+1)\pi. \end{cases} \tag{9.6}$$

For $f \in \mathcal{S}(\mathbb{R})$ and $(z, \phi) \in \mathfrak{H} \times \mathbb{R} \simeq \widetilde{\mathrm{SL}}(2, \mathbb{R})$, we define the *theta sum* by

$$\Theta_f(z, \phi) := \Theta_f(M) := \sum_{n \in \mathbb{Z}} [R(M)f](n), \tag{9.7}$$

with $M = N_x A_y K_\phi$. More explicitly,

$$\Theta_f(z, \phi) = y^{1/4} \sum_{n \in \mathbb{Z}} f_\phi(ny^{1/2}) e(n^2 x), \tag{9.8}$$

where $f_\phi = R(K_\phi)f$.

Using integration by parts, one finds that for any $T > 1$, there is a constant c_T such that for all $t \in \mathbb{R}$, $\phi \in \mathbb{R}$, we have

$$|f_\phi(t)| \leq c_T (1 + |t|)^{-T}. \tag{9.9}$$

The series in (9.7) and (9.8) converges therefore rapidly and uniformly for (z, ϕ) with z in any compact set in $\mathfrak{H}$.

It is well known that Θ_f is invariant under the discrete subgroup $\Delta_1(4)$ (see, e.g., [14, Proposition 3.1]), that is,

$$\Theta_f(\gamma M) = \Theta_f(M), \tag{9.10}$$

for all $\gamma \in \Delta_1(4)$. We may therefore view Θ_f as a smooth function on the manifold $\mathcal{M}_4$.

Proposition 9.1. If $f \in \mathcal{S}(\mathbb{R})$, then $\Theta_f \in B_{1/4}(\mathcal{M}_4)$. □

Proof. The manifold $\mathcal{M}_4$ has three cusps at $z = 0, 1/2$ and ∞. We have the bounds (cf. [14, Proposition 3.2])

$$\Theta_f(z, \phi) = \begin{cases} e^{i\pi/4} f_{\phi_0}(0) y_0^{1/4} + O_T(y_0^{-T}) & (y_0 \geq 1), \\ O_T(y_{1/2}^{-T}) & (y_{1/2} \geq 1), \\ f_{\phi_\infty}(0) y_\infty^{1/4} + O_T(y_\infty^{-T}) & (y_\infty \geq 1), \end{cases} \tag{9.11}$$

for any $T > 1$, with the cuspidal coordinates

$$\begin{aligned} (z_0, \phi_0) &= \big(-(4z)^{-1}, \phi + \arg z \big), \\ (z_{1/2}, \phi_{1/2}) &= \left(-(4z-2)^{-1}, \phi + \arg\left(z - \frac{1}{2} \right) \right), \\ (z_\infty, \phi_\infty) &= (z, \phi). \end{aligned} \tag{9.12}$$

■

10 Smoothed error terms

We will now construct functions $E_{f,\psi}$ on $\mathcal{M}_N$ which represent smoothed error terms. For real-valued $f \in \mathcal{S}(\mathbb{R})$ and $\psi \in C^\infty(S^1)$ with $\widehat{\psi}_0 = 0$ and only finitely many Fourier coefficients nonzero, put

$$E_{f,\psi}(z, 0) = \frac{1}{2} y^{1/4} \sum_{n \in \mathbb{Z}} f(n y^{1/2}) \psi(n^2 x). \tag{10.1}$$

The building blocks of $E_{f,\psi}$ are theta sums. It is easily seen that we have the expansion

$$E_{f,\psi}(z, 0) = \frac{1}{2} \sum_{k \neq 0} \widehat{\psi}_k \Theta_f(kx + iy, 0). \tag{10.2}$$

The following theorem tells us that $E_{f,\psi}(z, 0)$ can be extended to values $\phi \neq 0$, yielding a smooth function on $\mathcal{M}_N$ of moderate growth in the cusps.

Theorem 10.1. Let $f \in \mathcal{S}(\mathbb{R})$ and $\psi \in C^\infty(S^1)$ with $\widehat{\psi}_k \neq 0$ only if $0 < |k| \leq K$, for some integer K. Then there is a function $E_{f,\psi} \in B_{1/4}(\mathcal{M}_N)$ with $N = 4\,\mathrm{lcm}(2, 3, \dots, K)$ such that

$$E_{f,\psi}(z, 0) = \frac{1}{2} y^{1/4} \sum_{n \in \mathbb{Z}} f(n y^{1/2}) \psi(n^2 x). \tag{10.3}$$

□

Proof. We can write $E_{f,\psi}(z,\phi) = E^+_{f,\psi}(z,\phi) + E^-_{f,\psi}(z,\phi)$, where

$$E^+_{f,\psi}(z,0) = \frac{1}{2}\sum_{k>0} \widehat{\psi}_k \Theta_f(kx+iy,0),$$
$$E^-_{f,\psi}(z,0) = \frac{1}{2}\sum_{k>0} \widehat{\psi}_{-k} \overline{\Theta_f(kx+iy,0)}. \tag{10.4}$$

Since $N_{kx} = A_k N_x A_k^{-1}$, we find

$$\Theta_f(kx+iy,0) = \sum_{n\in\mathbb{Z}} \big[R\big(N_{kx}A_y\big)f\big](n) = \sum_{n\in\mathbb{Z}} \big[R\big(A_k N_x A_y A_k^{-1}\big)f\big](n). \tag{10.5}$$

We extend (10.5) to $\phi \neq 0$ by setting

$$\Theta_f^{(k)}(z,\phi) := \Theta_f^{(k)}(M) := \Theta_f\big(A_k M A_k^{-1}\big), \tag{10.6}$$

where $M = N_x A_y K_\phi$. The invariance of Θ_f under $\Delta_1(4)$ implies that

$$\Theta_f^{(k)}(\gamma M) = \Theta_f^{(k)}(M)k, \tag{10.7}$$

for all $\gamma \in A_k^{-1}\Delta_1(4)A_k$, and hence for all $\gamma \in \Delta_1(4k)$, recall Lemma 5.1(b). The functions

$$E^+_{f,\psi}(M) = \frac{1}{2}\sum_{k>0} \widehat{\psi}_k \Theta_f^{(k)}(M), \qquad E^-_{f,\psi}(M) = \frac{1}{2}\sum_{k>0} \widehat{\psi}_{-k} \overline{\Theta_f^{(k)}(M)} \tag{10.8}$$

are therefore invariant under the group

$$\bigcap_{k=1}^{K} \Delta_1(4k), \tag{10.9}$$

which contains $\Delta_1(N)$ with $N = 4\,\mathrm{lcm}(2,3,\ldots,K)$ (see Lemma 5.1(c)).

The bound (6.1) on the growth of $E_{f,\psi}$ in the cusps follows from (9.11) and the fact that $E_{f,\psi}$ is a finite linear combination of theta sums. (The implied constant in (6.1) may depend on K.) ■

Lemma 10.2. With f, ψ as in Theorem 10.1,

$$E_{f,\psi}(z,\phi+\pi) = -i\big(E^+_{f,\psi}(z,\phi) - E^-_{f,\psi}(z,\phi)\big). \tag{10.10}$$

□

Note that this implies in particular $E_{f,\psi}(z,\phi+2\pi) = -E_{f,\psi}(z,\phi)$.

Proof. We have $f_{\phi+\pi}(t) = -if_\phi(-t)$ (compare (9.5)) and thus

$$\Theta_f^{(k)}(z, \phi + \pi) = -i\Theta_f^{(k)}(z, \phi). \tag{10.11}$$

The lemma follows from (10.8). ■

11 Error terms are almost modular

The central observation of our investigation is that the original error term $\Xi_{f,\psi}$ introduced in (2.2) is an almost modular function.

Theorem 11.1. If $f \in PC_0^\infty(\mathbb{R}_+)$ and if $\psi \in L^2(S^1)$ satisfies conditions (2.4) and (2.5), then $\Xi_{f,\psi} \in \mathcal{B}^2$. □

Proof. The aim is to apply Lemma 4.1. Suppose that the largest jump at a discontinuity of f is $D = \sup_{t\in\mathbb{R}_+} |f(t+0) - f(t-0)|$. We can now approximate f by an even function $f_\varepsilon \in C_0^\infty(\mathbb{R})$ so that $\sup_{t\in\mathbb{R}_+} |f(t) - f_\varepsilon(t)| \leq D$ and $\mathrm{supp}_{(f-f_\varepsilon^+)}$ is arbitrarily small; here f_ε^+ denotes the restriction of f_ε to $\mathbb{R}_+$. Similarly, the function

$$\psi_\varepsilon(x) = \sum_{0<|k|\leq K} \widehat{\psi}_k e(kx) \tag{11.1}$$

approximates ψ arbitrarily well in the L^2 norm, for K large enough. At the same time, $C(\psi - \psi_\varepsilon)$ in (2.5) is independent of K since

$$\left|\widehat{\psi}_k - \widehat{\psi}_{k,\varepsilon}\right| \leq \frac{C(\psi)}{|k|^\beta}. \tag{11.2}$$

This allows us to choose $C(\psi-\psi_\varepsilon) = C(\psi)$. Hence for any $\varepsilon > 0$, we can find approximants $f_\varepsilon, \psi_\varepsilon$ such that

$$\begin{aligned} &\sup\left((f - f_\varepsilon^+)^2\right)\left|\mathrm{supp}_{(f-f_\varepsilon^+)}\right|\left(\|\psi\|_2^2 + K_\beta C(\psi)\|\psi\|_2\right) < \left(\frac{\varepsilon}{2}\right)^2, \\ &\sup\left(f^2\right)\left|\mathrm{supp}_f\right|\left(\|\psi - \psi_\varepsilon\|_2^2 + K_\beta C(\psi - \psi_\varepsilon)\|\psi - \psi_\varepsilon\|_2\right) < \left(\frac{\varepsilon}{2}\right)^2. \end{aligned} \tag{11.3}$$

Now $f_\varepsilon, \psi_\varepsilon$ also satisfy the conditions of Theorem 10.1, so

$$E_{f_\varepsilon,\psi_\varepsilon}(z, 0) = \frac{1}{2} y^{1/4} \sum_{n\in\mathbb{Z}} f_\varepsilon(ny^{1/2})\psi_\varepsilon(n^2x) \tag{11.4}$$

can be extended to $\phi \neq 0$ to yield a function $E_{f_\varepsilon,\psi_\varepsilon} \in B_{1/4}(\mathcal{M}_N)$. If we set $y = N^{-2}$, the theorem follows from Lemma 4.1 (compare Definition 7.1). ■

Proof of Theorem 2.1. Since the error term is almost modular of class $\mathcal{B}^2$ (Theorem 11.1), Theorem 2.1 is a special case of Theorem 8.2. The symmetry of the limit distribution is a consequence of the observation after Lemma 10.2. ■

Appendix

Generalized quadratic residue symbol

For any integer x and any prime p, the standard quadratic residue symbol $(\frac{x}{p})$ is 1 if x is a square modulo p, and -1 otherwise. The *generalized quadratic residue symbol* $(\frac{a}{b})$ is, for any integer a and any odd integer b, characterized by the following properties (see [12, pages 160–161]):

(i) $(\frac{a}{b}) = 0$ if $\gcd(a, b) \neq 1$,

(ii) $(\frac{a}{-1}) = \operatorname{sgn} a$,

(iii) if $b > 0$, $b = \prod_i b_i$, b_j primes (not necessarily distinct), then $(\frac{a}{b}) = \prod_i (\frac{a}{b_i})$,

(iv) $(\frac{a}{-b}) = (\frac{a}{-1})(\frac{a}{b})$,

(v) $(\frac{0}{\pm 1}) = 1$.

It follows from these properties that the symbol is bimultiplicative

$$\left(\frac{a_1 a_2}{b}\right) = \left(\frac{a_1}{b}\right)\left(\frac{a_2}{b}\right), \qquad \left(\frac{a}{b_1 b_2}\right) = \left(\frac{a}{b_1}\right)\left(\frac{a}{b_2}\right). \tag{A.1}$$

Furthermore, if $b > 0$, then $(\frac{\cdot}{b})$ defines a character modulo b; if $a \neq 0$, then $(\frac{a}{\cdot})$ defines a character modulo $4a$.

Acknowledgments

I would like to thank Z. Rudnick and A. Strömbergsson for very helpful discussions. This research has been supported by an EPSRC Advanced Research Fellowship, the Nuffield Foundation (Grant NAL/00351/G), and the EC Research Training Network (Mathematical Aspects of Quantum Chaos) HPRN-CT-2000-00103.

References

[1] P. M. Bleher, *On the distribution of the number of lattice points inside a family of convex ovals*, Duke Math. J. **67** (1992), no. 3, 461–481.

[2] ———, *Trace formula for quantum integrable systems, lattice-point problem, and small divisors*, Emerging Applications Of Number Theory (Minneapolis, Minn, 1996) (D. A. Hejhal,

J. Friedman, M. C. Gutzwiller, and A. M. Odlyzko, eds.), IMA Vol. Math. Appl., vol. 109, Springer, New York, 1999, pp. 1–38.

[3] N. I. Chernov, *Limit theorems and Markov approximations for chaotic dynamical systems*, Probab. Theory Related Fields **101** (1995), no. 3, 321–362.

[4] A. Eskin and C. McMullen, *Mixing, counting, and equidistribution in Lie groups*, Duke Math. J. **71** (1993), no. 1, 181–209.

[5] D. R. Heath-Brown, *The distribution and moments of the error term in the Dirichlet divisor problem*, Acta Arith. **60** (1992), no. 4, 389–415.

[6] W. B. Jurkat and J. W. Van Horne, *The proof of the central limit theorem for theta sums*, Duke Math. J. **48** (1981), no. 4, 873–885.

[7] ———, *On the central limit theorem for theta series*, Michigan Math. J. **29** (1982), no. 1, 65–77.

[8] ———, *The uniform central limit theorem for theta sums*, Duke Math. J. **50** (1983), no. 3, 649–666.

[9] M. Kac, *On the distribution of values of sums of the type* $\sum f(2^k t)$, Ann. of Math. (2) **47** (1946), 33–49.

[10] H. Kesten, *Uniform distribution* mod 1, Ann. of Math. (2) **71** (1960), 445–471.

[11] ———, *Uniform distribution* mod 1. *II*, Acta Arith. **7** (1961/1962), 355–380.

[12] G. Lion and M. Vergne, *The Weil Representation, Maslov Index and Theta Series*, Progress in Mathematics, vol. 6, Birkhäuser Boston, Massachusetts, 1980.

[13] J. Marklof, *Holomorphic almost modular forms*, preprint, 2003, http://www.maths.bris.ac.uk/~majm/.

[14] ———, *Limit theorems for theta sums*, Duke Math. J. **97** (1999), no. 1, 127–153.

[15] ———, *Theta sums, Eisenstein series, and the semiclassical dynamics of a precessing spin*, Emerging Applications of Number Theory (Minneapolis, Minn, 1996) (D. A. Hejhal, J. Friedman, M. C. Gutzwiller, and A. M. Odlyzko, eds.), IMA Vol. Math. Appl., vol. 109, Springer, New York, 1999, pp. 405–450.

[16] J. Marklof and A. Strömbergsson, *Equidistribution of Kronecker sequences along closed horocycles*, to appear in Geom. Funct. Anal., http://www.maths.bris.ac.uk/~majm/.

[17] Z. Rudnick and P. Sarnak, *The pair correlation function of fractional parts of polynomials*, Comm. Math. Phys. **194** (1998), no. 1, 61–70.

[18] Z. Rudnick, P. Sarnak, and A. Zaharescu, *The distribution of spacings between the fractional parts of* $n^2\alpha$, Invent. Math. **145** (2001), no. 1, 37–57.

[19] P. Sarnak, *Asymptotic behavior of periodic orbits of the horocycle flow and Eisenstein series*, Comm. Pure Appl. Math. **34** (1981), no. 6, 719–739.

[20] N. A. Shah, *Limit distributions of expanding translates of certain orbits on homogeneous spaces*, Proc. Indian Acad. Sci. Math. Sci. **106** (1996), no. 2, 105–125.

[21] A. N. Shiryaev, *Probability*, Graduate Texts in Mathematics, vol. 95, Springer, New York, 1996.

[22] H. Weyl, *Über die Gleichverteilung von Zahlen mod. Eins*, Math. Ann. **77** (1916), 313–352 (German).

Jens Marklof: School of Mathematics, University of Bristol, Bristol BS8 1TW, UK

E-mail address: j.marklof@bristol.ac.uk