Contemporary Mathematics
Volume **237**, 1999

On Bilu's Equidistribution Theorem

Robert Rumely

ABSTRACT. We place Bilu's equidistribution theorem for small points on algebraic tori in the context of potential theory, generalizing it to arbitrary compact sets of capacity 1.

Recently Szpiro, Ullmo, and Zhang ([**SUZ**]) proved that points of small height on an abelian variety A are equidistributed with respect to the Haar measure on $A(\mathbb{C})$. Bilu ([**Bi**]) established the analogue of this for algebraic tori. In particular he showed that if $h(\alpha)$ is the naive height on $\overline{\mathbb{Q}}$, the algebraic closure of $\mathbb{Q}$, and if $\{\alpha_n\}$ is a sequence of points in $\overline{\mathbb{Q}}$ such that

$$\deg(\alpha_n) \to \infty, \qquad h(\alpha_n) \to 0$$

as $n \to \infty$, then the discrete measures

$$\Delta_n = \frac{1}{\deg(\alpha_n)} \sum_{\sigma:\mathbb{Q}(\alpha_n)\to\mathbb{C}} \delta_{\sigma(\alpha_n)}(x)$$

converge weakly (relative to the space of continuous functions on $\mathbb{C}$ with compact support) to the uniform measure of mass 1 on the boundary of the unit disc $D(0,1)$.

The purpose of this note is to place Bilu's result in the context of potential theory. (Bombieri ([**Bo**]) has also touched on these ideas.) The set $E = D(0,1)$ has logarithmic capacity $\gamma(E) = 1$. Below, we will show that for an arbitrary compact set $E \subset \mathbb{C}$ with capacity $\gamma(E) = 1$, an analogue of Bilu's theorem is true.

1. The potential-theoretic setting

One way to define the naive height on $\overline{\mathbb{Q}}$ is as follows: for $\alpha \in \overline{\mathbb{Q}}$, if $K = \mathbb{Q}(\alpha)$ then

$$h(\alpha) \;=\; \sum_{p,\infty} h_p(\alpha)$$

where

$$h_p(\alpha) \;=\; \frac{1}{\deg(\alpha)} \sum_{\sigma:K\to\overline{\mathbb{Q}}_p} \max(0, \log(|\sigma(\alpha)|_p)) \;\; .$$

Here $|x|_p$ is the absolute value on $\overline{\mathbb{Q}}_p$ which extends the canonically normalized absolute value on $\mathbb{Q}_p$, and $\log(x) = \ln(x)$.

From the viewpoint of potential theory, at the archimedean place the function

$$\max(0, \log(|z|))$$

is simply the Green's function, relative to the point ∞, of $D(0,1)$. For an arbitrary compact $E \subset \mathbb{C}$ of positive capacity, define

$$h_{E,\infty}(\alpha) \;=\; \frac{1}{\deg(\alpha)} \sum_{\sigma: K \to \mathbb{C}} G(\sigma(\alpha), \infty; E)$$

where $G(z, \infty; E)$ is the the Green's function of the domain $\mathbb{C} \backslash E$, extended by 0 on E. Put

$$h_E(\alpha) \;=\; h_{E,\infty}(\alpha) + \sum_{p \text{ finite}} h_p(\alpha) \quad .$$

Thus $h_E(\alpha)$ is the "height of α with respect to the set E".

The uniform measure is the so-called "equilibrium distribution" of $D(0,1)$. In general, for a compact set E of positive capacity, the equilibrium distribution is the unique probability measure μ on E which minimizes the energy integral

$$\begin{aligned} I(\mu) \;&=\; \iint_{E \times E} -\log(|z-w|) d\mu(z) d\mu(w) \\ &:= \lim_{M \to \infty} \iint_{E \times E} \min\left(M, -\log(|z-w|)\right) d\mu(z) d\mu(w) \quad . \end{aligned}$$

The value of the integral is precisely

$$V(E) \;=\; -\log(\gamma(E)) \;;$$

if ν is any other probability measure on E, then

$$I(\nu) \;=\; \iint_{E \times E} -\log(|z-w|) d\nu(z) d\nu(w) \;>\; V(E) \quad .$$

The equilibrium distribution of E is always supported on the outer boundary of E. For proof of these facts, see ([**Ts**],pp.53–89).

THEOREM 1. *Suppose a compact set $E \subset \mathbb{C}$ has capacity $\gamma(E) = 1$ and is stable under complex conjugation. If $\{\alpha_n\} \subset \overline{\mathbb{Q}}$ is a sequence for which $\deg(\alpha_n) \to \infty$ and $h_E(\alpha_n) \to 0$, then the measures Δ_n converge weakly to the equilibrium distribution μ of E.*

The proof exploits the relationship between the energy integral and the discriminant, together with the fact that the equilibrium distribution is the *unique* probability measure minimizing the energy integral. The hypothesis that E is stable under complex conjugation is not needed for the proof; however, by the Fekete-Szegö theorem ([**F-S**]), it assures that the theorem is not vacuous.

For each place of $\mathbb{Q}$, fix an embedding of $\overline{\mathbb{Q}}$ into $\overline{\mathbb{Q}}_p$. Given $\alpha \in \overline{\mathbb{Q}}$, put

$$D(\alpha) \;=\; \prod_{i \neq j} (\sigma_i(\alpha) - \sigma_j(\alpha)) \;\in\; \mathbb{Q}$$

where σ_i, σ_j run over the embeddings of $K = \mathbb{Q}(\alpha)$ into $\overline{\mathbb{Q}}$. If $\alpha \neq 0$ then $D(\alpha) \neq 0$, and by the product formula

$$\prod_{p,\infty} |D(\alpha)|_p \;=\; 1 \quad .$$

For each nonarchimedean p, we can bound $|D(\alpha)|_p$ in terms of $h_p(\alpha)$. Indeed, for each $i \neq j$

$$\begin{aligned}\log(|\sigma_i(\alpha) - \sigma_j(\alpha)|_p) &\leq \log(\max(|\sigma_i(\alpha)|_p, |\sigma_j(\alpha)|_p)) \\ &\leq \max(0, \log(|\sigma_i(\alpha)|_p)) + \max(0, \log(|\sigma_j(\alpha)|_p))\end{aligned}$$

so that if $N = \deg(\alpha) = [\mathbb{Q}(\alpha) : \mathbb{Q}]$ then

$$\begin{aligned}\log(|D(\alpha)|_p) &= \sum_{i \neq j} \log(|\sigma_i(\alpha) - \sigma_j(\alpha)|_p) \\ &\leq 2(N-1) \sum_i \max(0, \log(|\sigma_i(\alpha)|_p)) \\ &= 2N(N-1) h_p(\alpha) \quad .\end{aligned}$$

At the archimedean place, if Δ is the discrete measure with mass $1/N$ at each conjugate of α, we have

$$\begin{aligned}-\frac{1}{N^2} \log(|D(\alpha)|_\infty) &= \sum_{i \neq j} -\log(|\sigma_i(\alpha) - \sigma_j(\alpha)|_\infty) \cdot \frac{1}{N^2} \\ &= \iint_{\mathbb{C} \times \mathbb{C} \setminus \{\text{diagonal}\}} -\log(|z - w|) d\Delta(z) d\Delta(w) \quad .\end{aligned}$$

Now let $\{\alpha_n\}$ be a sequence of points in $\overline{\mathbb{Q}}$ such that $\deg(\alpha_n) \to \infty$, $h_E(\alpha_n) \to 0$ as $n \to \infty$. Suppose for purpose of contradiction that the associated measures Δ_n do *not* converge weakly to the equilibrium distribution of μ of E. From the fact that $h_E(\alpha_n) \to 0$ it follows that the sequence of probability measures $\{\Delta_n\}$ is tight. Therefore, we can extract a subsequence converging to a probability measure $\nu \neq \mu$. Passing to this subsequence, we can assume that

$$\Delta_n \to \nu$$

weakly. Let E_0 be the "filled set" corresponding to E: $E_0 = \mathbb{C} \setminus U$, where U is the unbounded component of $\mathbb{C} \setminus E$. It is easy to see that ν is supported on E_0. On the other hand, it is well-known that E_0 has the same capacity and equilibrium distribution as E (see ([**Ts**],p.61)), so we can replace E by E_0. Since $\gamma(E) = 1$, we have

$$I(\nu) = \iint_{E \times E} -\log(|z - w|) d\nu(z) d\nu(w) > V(E) = 0 \quad .$$

Now we need a

LEMMA. *For each $M \geq 0$,*

$$\begin{aligned}\liminf_{n \to \infty} \iint_{\mathbb{C} \times \mathbb{C} \setminus \{\text{diagonal}\}} &\min\left(M, -\log(|z - w|)\right) d\Delta_n(z) d\Delta_n(w) \\ &\geq \iint_{E \times E} \min\left(M, -\log(|z - w|)\right) d\nu(z) d\nu(w) \quad .\end{aligned}$$

PROOF. Since M is fixed and the Δ_n-mass of the diagonal is $1/n$, without loss we can adjoin the diagonal to the integrals on the left. Fix bounded neighborhoods I, J of E such that the closure of I is contained in J, and choose a continuous function $\varphi(z)$ supported in J such that $0 \le \varphi(z) \le 1$ for all z and $\varphi(z) = 1$ on I. Put $\mathcal{O} = \mathbb{C}\backslash I$, and $\overline{\varphi}(z) = 1 - \varphi(z)$. Then

(*)

$$\begin{aligned}\iint_{\mathbb{C}\times\mathbb{C}} \min\big(M, -\log(|z-w|)\big) d\Delta_n(z) d\Delta_n(w) \ = \\ \iint_{J\times J} \min\big(M, -\log(|z-w|)\big)\varphi(z)\varphi(w) d\Delta_n(z) d\Delta_n(w) \\ +2 \iint_{\mathcal{O}\times J} \min\big(M, -\log(|z-w|)\big)\overline{\varphi}(z)\varphi(w) d\Delta_n(z) d\Delta_n(w) \\ + \iint_{\mathcal{O}\times\mathcal{O}} \min\big(M, -\log(|z-w|)\big)\overline{\varphi}(z)\overline{\varphi}(w) d\Delta_n(z) d\Delta_n(w) \ .\end{aligned}$$

It is not hard to see that there is a constant $L \ge 0$ such that

$$\begin{cases} -\log|z-w| \ge -G(z,\infty;E) - L & \text{for all } z \in \mathcal{O},\ w \in J \\ -\log|z-w| \ge -G(z,\infty;E) - G(w,\infty;E) - L & \text{for all } z \in \mathcal{O},\ w \in \mathcal{O}\ . \end{cases}$$

Since $M \ge 0$, the estimates above also apply to the kernels in (*). Using that $\sum_\sigma G(\sigma(\alpha_n),\infty;E)\cdot(1/N_n) \le h_E(\alpha_n)$, we see that

$$\begin{aligned}\iint_{\mathbb{C}\times\mathbb{C}} \min\big(M, -\log(|z-w|)\big) d\Delta_n(z) d\Delta_n(w) \ \ge \\ \iint_{J\times J} \min\big(M, -\log(|z-w|)\big)\varphi(z)\varphi(w) d\Delta_n(z) d\Delta_n(w) - 4h_E(\alpha_n) - 3\Delta_n(\mathcal{O})L\ .\end{aligned}$$

Since $\Delta_n \to \nu$ weakly as $n \to \infty$, the first term on the right approaches

$$\iint_{E\times E} \min\big(M, -\log(|z-w|)\big) d\nu(z) d\nu(w)\ .$$

Likewise, $\Delta_n(\mathcal{O}) \to 0$. But $h_E(\alpha_n) \to 0$ by hypothesis, so we obtain the assertion in the lemma.

Returning to the main proof, since $I(\nu)$ is a limit of integrals of the kernels $\min(M, -\log(|z-w|))$, it follows from the Lemma that

$$\begin{aligned}\liminf_{n\to\infty} \Big(\iint_{\mathbb{C}\times\mathbb{C}\backslash\{\text{diagonal}\}} -\log(|z-w|) d\Delta_n(z) d\Delta_n(w)\Big) \\ \ge \iint_{E\times E} -\log(|z-w|) d\nu(z) d\nu(w) \\ = I(\nu) \ > \ 0\ .\end{aligned}$$

Writing $N_n = \deg(\alpha_n)$, then since $N_n \to \infty$ it follows that

$$\limsup_{n\to\infty} \frac{1}{N_n(N_n-1)} \log(|D(\alpha_n)|_\infty) \ \le \ -I(\nu) \ < \ 0\ .$$

Fix $\varepsilon > 0$ small enough that $I(\nu) > 3\varepsilon$. Then for sufficiently large n, we have $h_E(\alpha_n) < \varepsilon$, so

$$\sum_{p \text{ finite}} h_p(\alpha_n) \leq h_E(\alpha_n) < \varepsilon .$$

Also, for sufficiently large n,

$$\frac{1}{N_n(N_n-1)} \log(|D(\alpha_n)|_\infty) < -I(\nu) + \varepsilon .$$

It follows that for such n,

$$\begin{aligned} 0 &= \sum_{p,\infty} \log(|D(\alpha_n)|_p) \\ &< 2N_n(N_n-1) \cdot \Big(\sum_{p \text{ finite}} h_p(\alpha_n)\Big) + N_n(N_n-1) \cdot (-I(\nu)+\varepsilon) \\ &\leq N_n(N_n-1) \cdot \big(-I(\nu) + 3\varepsilon\big) < 0 . \end{aligned}$$

This contradiction shows that $\Delta_n \to \mu$ weakly.

2. Applications and generalizations

Apart from the unit disc, arithmetically the most interesting set of capacity 1 is the line segment $E = [-2, 2]$. (In general, the capacity of a line segment $[a, b]$ is $(b-a)/4$).) The equilibrium distribution and Green's function of E are

$$\begin{aligned} d\mu &= \frac{1}{\pi} \cdot \frac{dx}{\sqrt{4-x^2}} , \\ G(z,\infty;E) &= \log(|\frac{1}{2}(z+\sqrt{z^2-4})|) . \end{aligned}$$

The algebraic integers whose conjugates all belong to E are just the numbers of the form $2\cos(a\pi/n) = \zeta_n^a + \zeta_n^{-a}$, where a and n are positive integers.

Put $H(\alpha) = \deg(\alpha) \cdot h(\alpha)$. Lehmer's conjecture asserts that there is a constant $C > 0$ such that unless α is 0 or a root of unity, then

$$H(\alpha) \geq C .$$

Correspondingly, if $H_E(\alpha) = \deg(\alpha) \cdot h_E(\alpha)$, the analogue of Lehmer's conjecture would assert there is a constant $C' > 0$ such that unless $\alpha = 2\cos(a\pi/n)$ for some a, n, then

$$H_E(\alpha) \geq C' .$$

By a theorem of Smythe ([**Sm**]), if θ is the real root of $\theta^3 - \theta - 1 = 0$, then for $C = \ln(\theta) \simeq .28119957$ the only possible exceptions to Lehmer's conjecture are units whose minimal polynomials are reciprocal polynomials, which we can assume have even degree, since the only irreducible reciprocal polynomial of odd degree is $x+1$. Given such a unit α, put $\beta = \alpha + 1/\alpha$. Then β is an algebraic integer, and we claim that

$$H_E(\beta) = H(\alpha) .$$

Indeed, the finite contributions to $H_E(\beta)$ and $H(\alpha)$ are both 0. At the archimedean place, the map $w = \varphi(z) := z+1/z$ takes $\partial D(0,1)$ onto E and takes the complement of $D(0,1)$ conformally onto the complement of E; the Green's functions satisfy

$$G(z,\infty;D(0,1)) + G(1/z,\infty;D(0,1)) = G(\varphi(z),\infty;E) .$$

For any given z, at least one of the terms on the left vanishes. The conjugates of α come in pairs $\sigma(\alpha)$, $1/\sigma(\alpha)$, and it is easy to see that $[\mathbb{Q}(\alpha):\mathbb{Q}(\beta)] = 2$. Thus, the archimedean components of $H_E(\beta)$ and $H(\alpha)$ also coincide.

Conversely, if β is an algebraic integer and $\alpha \in \mathbb{C}$ satisfies $\varphi(\alpha) = \beta$, then α is a unit with $[\mathbb{Q}(\alpha):\mathbb{Q}(\beta)] = 1$ or 2, and by the identity above

$$H(\alpha) \;\le\; H_E(\beta)\ .$$

Thus, Lehmer's conjecture is true if and only if its analogue holds for $E = [-2, 2]$. Moreover, the exceptions to the classical Lehmer's conjecture are precisely the pullbacks of exceptions to the Lehmer's conjecture for E.

For an infinite class of examples, consider the family of ellipses

$$E_{a,b} \;=\; \{x + yi \in \mathbb{C} : \frac{x^2}{a^2} + \frac{y^2}{b^2} \le 1\}\ .$$

The capacity of $E_{a,b}$ is $(a + b)/2$.

Suppose $a + b = 2$, so that $\gamma(E_{a,b}) = 1$. Then the Green's function of $E_{a,b}$ is

$$G(z, \infty; E_{a,b}) \;=\; \begin{cases} \log(|\frac{1}{2}(z + \sqrt{z^2 - 4(a-1)})|) & \text{if } z \notin E_{a,b} \\ 0 & \text{if } z \in E_{a,b} \end{cases} .$$

where $\sqrt{w}$ is chosen to be positive on the positive real axis. The level set of $G(z, \infty; E_{a,b}) = \varepsilon$ corresponding to $0 < \varepsilon = \ln(1 + \delta)$ is the ellipse $\partial E_{r,s}$ where

$$r \;=\; \frac{a + 2\delta + \delta^2}{1 + \delta}, \qquad s \;=\; \frac{b + 2\delta + \delta^2}{1 + \delta}\ .$$

If $\partial E_{a,b}$ is parametrized by $\varphi(\theta) = a\cos(\theta) + b\sin(\theta)i$, $0 \le \theta \le 2\pi$, then the equilibrium distribution of $E_{a,b}$ is given by

$$d\mu \;=\; \frac{1}{2\pi} \cdot \frac{d\theta}{\sqrt{b^2\cos^2(\theta) + a^2\sin^2(\theta)}}\ .$$

These facts can be deduced from the fact that $w = z + (a-1)/z$ takes the complement of $D(0, 1)$ conformally to the complement of $E_{a,b}$. The equilibrium distribution is given by the normal derivative of the Green's function, scaled to have total mass 1.

It is an interesting question whether the analogue of Lehmer's conjecture holds for some, or perhaps almost all, ellipses $E_{a,b}$ of capacity 1.

Theorem 1 can easily be generalized to an adelic setting, using the machinery developed in ([**R1**]). Let K be a global field, and let $\mathbb{E} = \prod_v E_v$ be an adelic set such that

1. for each v, $E_v \subset \overline{K}_v$, and E_v is stable under $\mathrm{Gal}(\overline{K}_v/K_v)$;
2. for all but finitely many v, $E_v = \{z \in \overline{K}_v : |z|_v \le 1\}$;
3. for the remaining v, either E_v is compact, or E_v is defined by an inequality $\{z : |P(z)|_v \le 1\}$ for some nonconstant $P(z) \in K_v[z]$.

In ([**R1**],§3,§4), the local capacity $\gamma(E_v)$ and Green's function $G(z, \infty; E_v)$ are defined, and for each v such that E_v is compact, the equilibrium distribution μ_v of E_v is defined.

For each $\alpha \in \overline{K}$ we can define the height relative to $\mathbb{E}$ by

$$h_{\mathbb{E}}(\alpha) = \sum_v h_{E_v}(\alpha)$$

where

$$h_{E_v}(\alpha) = \frac{1}{[K(\alpha):K]_{\text{sep}}} \sum_{\sigma:K(\alpha)\to\overline{K}_v} G(\sigma(\alpha),\infty;E_v) .$$

In this setting, the global capacity of $\mathbb{E}$, defined in ([**R1**],§5.1), is given by

$$\gamma(\mathbb{E},\{\infty\}) = \prod_v \gamma(E_v) .$$

Minor modifications to the proof of Theorem 1 yield

THEOREM 2. *Suppose that K and $\mathbb{E}$ are as above, and $\gamma(\mathbb{E},\{\infty\}) = 1$. Let $\{\alpha_n\} \subset \overline{K}$ be a sequence such that $[K(\alpha_n):K] \to \infty$, $h_{\mathbb{E}}(\alpha_n) \to 0$. Then for each v such that E_v is compact, the probability measures Δ_n on $\overline{K}_v$ attached to the α_n converge weakly to the equilibrium distribution μ_v of E_v.*

As an example, take $K = \mathbb{Q}$, and let S be a finite set of primes, including the archimedean one. For each nonarchimedean $p \in S$, let $E_p = \mathbb{Z}_p$. Then as shown in ([**R1**],p.353),

$$\gamma(E_p) = p^{-1/(p-1)} .$$

The equilibrium distribution of E_p is just additive Haar measure. At the archimedean prime, put $E_\infty = [-2r, 2r]$ where

$$r = \prod_{\text{finite } p\in S} p^{1/(p-1)} .$$

For each $p \notin S$, put $E_p = \{z \in \overline{\mathbb{Q}}_p : |z|_p \le 1\}$. For such p, $\gamma(E_p) = 1$. Thus, if $\mathbb{E} = \prod_{p,\infty} E_p$, then

$$\gamma(\mathbb{E},\{\infty\}) = 1 .$$

Using a more sophisticated formulation of the equilibrium distribution, such as that in ([**R2**]), Theorem 2 can undoubtedly be strengthened to allow more general sets E_v at nonarchimedean places. It can probably also be extended to adelic sets on curves, with capacities computed relative to several poles, as in ([**R1**]).

References

[Bi] Bilu, Y., *Limit distribution of small points on algebraic tori*, Duke Mathematical Journal 89 (1997), 465-476.

[Bo] Bombieri, E., *Subvarieties of linear tori and the unit equation*, in: Analytic Number Theory, ed. Y. Motohashi, LMS Lecture Notes 247 (1997), Cambridge University Press, Cambridge, U.K., 1-20.

[F-S] Fekete, M. and Szegö, G., *On algebraic equations with integral coefficients whose roots belong to a given point set*, Math. Zeit. 63 (1955), 158-172.

[R1] Rumely, R., *Capacity Theory on Algebraic Curves*, LNM 1378, Springer-Verlag, Berlin Heidelberg New York, 1989.

[R2] Rumely, R., *An intersection theory for curves, with analytic contributions from nonarchimedean places*, in: Canadian Math. Soc. Conference Proceedings, vol 15 (1995), AMS, Providence, R.I..

[Sm] Smythe, C., *On the product of the conjugates outside the unit circle of an algebraic integer*, Bull. London Math. Soc. 3 (1971), 169-175.

[SUZ] Szpiro, L., Ullmo, E., and Zhang, S., *Equidistribution des petits points*, Invent. Math. 127 (1997), 337-347.

[Ts] Tsuji, M., *Potential Theory in Modern Function Theory*, Maruzen, Tokyo 1959; reprinted: Chelsea, N.Y., 1975.

DEPARTMENT OF MATHEMATICS, UNIVERSITY OF GEORGIA, ATHENS GA 30602