§6. Applications.

§6.1 The Classical Fekete and Fekete-Szegö Theorems

This section is purely expository. The proofs of the general Fekete and Fekete-Szegö Theorems in the next two sections are quite complicated due to technical difficulties from the generality of working adelically on algebraic curves and using capacity with respect to several centers. Here we present the Theorems in their classical form, with the hope that the ideas involved will be illustrated as clearly as possible.

Both Theorems deal with compact sets in the complex plane, and the capacity is the classical logarithmic capacity with respect to ∞, measured using the parameter $1/z$.

(6.1.1) Theorem (Fekete,[Fek]). *Let $E \subset \mathbb{C}$ be a compact set, stable under complex conjugation, having capacity $\gamma_\infty(E) < 1$. Then there exists a neighborhood U of E which contains only a finite number of complete conjugate sets of algebraic integers.*

Proof. The idea for the proof is to construct a monic polynomial $f(z) \in \mathbb{Z}[z]$ with sup norm $\|f(z)\|_E < 1$. The neighborhood can then be taken as

$$U = \{ z \in \mathbb{C} : |f(z)| < 1 \} .$$

If α is an algebraic integer which, together with all its conjugates $\sigma\alpha$, lies in U, then $\prod_\sigma f(\sigma\alpha)$ is a rational integer with absolute value less than one. Hence it is zero, so that some (and therefore every) conjugate of α is a root of $f(z)$. Since a polynomial has only finitely many roots, the result follows.

To construct $f(z)$, we first find a monic polynomial $h(z)$ with *real* coefficients which satisfies $\|h(z)\|_E < 1$, and then, by a process called "patching", make the transition from real to integer coefficients. The connection with capacity theory, and the idea behind the construction of $h(z)$, is that $(1/n)\log|h(z)|$ should closely approximate $G(z,\infty;E)-V_\infty(E)$ outside a small neighborhood of E.

As a preliminary reduction, we can replace E with a finite cover by suitably small, closed discs (stable under complex conjugation) and still maintain $\gamma_\infty(E) < 1$. Thus we can assume that ∂E is a union of arcs, which in turn implies that $G(z,\infty;E)$ is continuous for all $z \in \mathbb{C}$. Of course, $G(z,\infty;E) = 0$ for $z \in E$.

Because of our hypothesis that $\gamma_\infty(E) < 1$, we have

$$V_\infty(E) = -\log(\gamma_\infty(E)) > 0 \quad .$$

Fix r with $0 < r < V_\infty(E)$, and define two sets

$$(1) \qquad \begin{aligned} \mathcal{I}_r &= \{ z \in \mathbb{C} : G(z,\infty;E) \le r \} \\ \mathcal{E}_r &= \{ z \in \mathbb{C} : G(z,\infty;E) \ge r \} \quad . \end{aligned}$$

Then $G(z,\infty;E)-V_\infty(E) < 0$ throughout $\mathcal{I}_r$, and E is contained in the interior of $\mathcal{I}_r$.

Fix $\varepsilon > 0$ satisfying $r+\epsilon < V_\infty(E)$, and let μ be the equilibrium distribution of E. Since

$$G(z,\infty;E) - V_\infty(E) = -u_E(z,\infty) = \int_E \log|z-w| \, d\mu(w) \quad ,$$

we can construct a monic polynomial $h(z) \in \mathbb{R}[z]$ such that for all $z \in \mathcal{E}_r$

$$(2) \qquad | \, [G(z,\infty;E) - V_\infty(E)] - (1/n)\log|h(z)| \, | < \varepsilon$$

(here $n = \deg(h)$). The idea is that if we approximate μ sufficiently well by a discrete probability measure (in the weak topology), then the locations of the point masses determine the zeros of $h(z)$, and their weights determine the multiplicities. If the measure is stable under complex conjugation, $h(z)$ will have real coefficients. For details, see Hille ([Hil],vol.II,p.264) or the proof of Proposition 3.3.3.

Our interest in (2) is primarily that it holds on the boundary $\partial\mathcal{E}_r$, where it implies

$$(3) \qquad |h(z)| \le R := e^{n(r+\varepsilon-V(E))} < 1 \quad .$$

Since $\partial\mathcal{I}_r = \partial\mathcal{E}_r$, the Maximum Modulus Principle shows that (3) holds for all $z \in \mathcal{I}_r$, and in particular for all $z \in E$. Thus $\|h(z)\|_E < 1$, indeed, $\|h(z)\|_{\mathcal{I}_r} \le R < 1$. Consequently, for any power $h(z)^d$ of $h(z)$,

$$\|h(z)^d\|_E \le R^d < 1 \quad .$$

We now come to the patching argument, which has two steps. The first step takes the polynomials $h(z)^d$ and modifies their high-order coefficients to nearby integers in such a way that the norms $\|h(z)^d\|_{\mathcal{I}_r}$ remain less than 1. (Here, "high-order" means all terms with degree above a fixed bound, to be specified.) The second step subtracts two of the modified polynomials to construct a polynomial whose high-order coefficients are integers, and whose low-order coefficients are very close to integers (applying the pigeon-hole principle). Rounding off those coefficients, we obtain $f(z)$.

For the first step, patching the high order coefficients, we take $h(z)^d$ and successively modify its coefficients to be integers, working from highest to lowest order. The crude way to do this would be to round the coefficient of z^k to the nearest integer, or equivalently, to add a multiple of z^k. Instead, we write $k = mn+i$ where $0 \le i < n$, and add a multiple of $z^i h(z)^m$: this leaves the coefficients above z^{mn+i} unchanged, modifies the coefficient of z^{mn+i} to be an integer, and changes the coefficients below z^{mn+i} in unknown ways but leaves them in $\mathbb{R}$. The power of $h(z)$ insures that $z^i h(z)^m$ has small norm on E as long as m is sufficiently large. We patch the coefficients down through degree nT in this manner, where T is a parameter to be specified shortly. This gives a polynomial

$$h^{(d)}(z) = h(z)^d + \sum_{j=0}^{d-T} \sum_{k=1}^{n} \Delta_{j,k}^{(d)} z^{n-k} h(z)^{d-j}$$

whose high order coefficients are integers. Clearly the $\Delta_{j,k}^{(d)}$ have absolute value at most $1/2$.

We now specify the parameter T. Put $M = \sum_{i=0}^{n-1} \|z^i\|_E$; then

$$\|h^{(d)}(z)\|_E \le R^d + M \sum_{j=T}^{d} R^j \le R^d + M R^T/(1-R) \quad . \tag{4}$$

Since $R < 1$, we can choose T sufficiently large that whenever $d > T$,

$$\|f^{(d)}(z)\|_E < 1/3 \quad . \tag{5}$$

Observe that T depends only on $h(z)$ and E, not on d.

For the second step, patching the low-order coefficients, expand $h^{(d)}(z)$ in powers of z, writing

$$h^{(d)}(z) = \sum_{i=0}^{nd} c_i^{(d)} z^i \quad .$$

Let $[0,1]$ denote the interval $\{ x \in \mathbb{R} : 0 \le x \le 1 \}$, and map $h^{(d)}(z)$ to the vector

$$(\langle c_0^{(d)}\rangle, \langle c_1^{(d)}\rangle, \ldots, \langle c_{nT-1}^{(d)}\rangle) \in [0,1]^{nT}$$

where $0 \le \langle c\rangle < 1$ denotes the fractional part of $c \in \mathbb{R}$. By the Pigeon Hole Principle, or the compactness of $[0,1]$, we can find vectors corresponding to distinct values of d, which are arbitrarily near each other. In particular we can find $d_2 > d_1 > T$ such that

$$(6) \qquad \left\| \sum_{i=0}^{nT-1} (\langle c_i^{(d_2)} \rangle - \langle c_i^{(d_1)} \rangle) z^i \right\|_E < 1/3 \quad .$$

The desired polynomial $f(z)$ can then be taken as

$$(7) \qquad f(z) = h^{(d_2)}(z) - h^{(d_1)}(z) - \sum_{i=0}^{nT-1} (\langle c_i^{(d_2)} \rangle - \langle c_i^{(d_1)} \rangle) z^i \quad .$$

Clearly $f(z)$ is monic with integer coefficients, and by (5), (6) and the triangle inequality for $\|\cdot\|_E$,

$$\|f(z)\|_E < \tfrac{1}{3} + \tfrac{1}{3} + \tfrac{1}{3} = 1 \quad .$$

This completes the proof. //

The contrasting result is

(6.1.2) Theorem (Fekete-Szegö,[F-Sz]) *Let* $E \subset \mathbb{C}$ *be a compact set, stable under complex conjugation, having capacity* $\gamma_\infty(E) \geq 1$. *Then every neighborhood* U *of* E *contains infinitely many complete conjugate sets of algebraic integers.*

Remark. It is not necessarily true that E itself contains infinitely many conjugate sets of algebraic integers. For example, take E to be a circle with center at the origin and transcendental radius $r > 1$; then E contains no algebraic numbers at all.

Proof of Theorem 6.1.2. Again the idea for the proof is to construct a monic polynomial $f(z) \in \mathbb{Z}[z]$; however, this time the neighborhood U is specified in advance and we want $f(z)$ to satisfy

$$\{ z \in \mathbb{C} : |f(z)| \leq 1 \} \subset U \quad .$$

If that is so, the roots of

$$f(z)^m - 1 = 0 \; , \qquad m = 1,2,\ldots$$

will be the algebraic integers required by the Theorem. Clearly these numbers and their conjugates all belong to the set U, in fact to the level set $\{ z \in \mathbb{C} : |f(z)| = 1 \}$. There are infinitely many of them, since the fields they generate contain roots of unity of arbitrarily high order.

To construct $f(z)$, we start by approximating $G(z,\infty;E)-V_\infty(E)$ as before; however because $\gamma_\infty(E) \geq 1$, the patching process is somewhat different.

We begin with some preliminary reductions. Fix U. By enlarging E slightly within U, we can assume that the strict inequality $\gamma_\infty(E) > 1$ holds, not just $\gamma_\infty(E) \geq 1$ (see Proposition 3.3.1). Then, covering E with a finite number of closed discs contained in U, and replacing it by their union, we can assume that ∂E is composed of a finite union of arcs, and that $\mathbb{C}\backslash E$ has only finitely many connected components. Finally, by cutting slits through E, we can arrange that $\mathbb{C}\backslash E$ be connected; if the slits are sufficiently narrow, it will still be true that $\gamma_\infty(E) > 1$ (see the proof of Proposition 3.3.2). Observe that we can carry out all the modifications above in such a way that E remains stable under complex conjugation.

We are thus reduced to the case where $\gamma_\infty(E) > 1$, $\mathbb{C}\backslash E$ is connected and ∂E is a finite union of arcs. Because of the last two conditions, $G(z,\infty;E)$ satisfies

$$\begin{cases} G(z,\infty;E) = 0 & \text{for } z \in E \\ G(z,\infty;E) > 0 & \text{for } z \in \mathbb{C}\backslash E \end{cases} .$$

Hence there is a number $r > 0$ such that

$$\{ z \in \mathbb{C} : G(z,\infty;E) \leq r \} \subset U .$$

As in Theorem 6.1.1, define

$$\mathcal{F}_r = \{ z \in \mathbb{C} : G(z,\infty;E) \leq r \}$$

$$\mathcal{E}_r = \{ z \in \mathbb{C} : G(z,\infty;E) \geq r \} .$$

Fix ε in the range $0 < \varepsilon < r$; then as in Theorem 6.1.1 we can find a monic polynomial $h(z) \in \mathbb{R}[z]$ of degree n, say, such that

$$| [G(z,\infty;E) - V_\infty(E)] - (1/n)\log|h(z)| | < \varepsilon$$

for all $z \in \mathcal{E}_r$. In particular, on $\mathcal{E}_r$,

$$(8) \qquad (1/n)\log|h(z)| > -V_\infty(E) + r - \varepsilon .$$

Since $\gamma_\infty(E) > 1$, we have $-V_\infty(E) > 0$, so that

$$(9) \qquad R := e^{n(-V_\infty(E)+r-\varepsilon)} > 1 .$$

A reformulation of (8) is that $|h(z)| > R$ for all $z \in \mathcal{E}_r$. Since $\mathbb{C}\backslash\mathcal{E}_r \subset U$,

$$(10) \qquad \{ z \in \mathbb{C} : |h(z)| \leq R \} \subset U .$$

This inclusion, and the fact that $R > 1$, are the key points needed for the proof.

By perturbing the coefficients of $h(z)$ slightly we can arrange that $h(z) \in \mathbb{Q}[z]$, while still maintaining (10) for some $R > 1$. Put $\mathcal{X} = \{ z \in \mathbb{C} : |h(z)| \le R \}$, noting that $\mathcal{X} \subset U$ and that $\partial\mathcal{X}$ is the level curve $|h(z)| = R$.

The strategy is now to show that for a suitable integer $d > 0$ it is possible to patch $h(z)^d$ so that the resulting polynomial has integer coefficients, and still has absolute value > 1 for all $z \in \mathbb{C}\backslash\mathcal{X}$. This time a bounded number of high order coefficients give the trouble in the patching process, and we make them integral by a trick that goes back to Fekete and Szegö.

The trick is as follows: expand

$$h(z) = z^n + \sum_{i=1}^{n} c_i z^{n-i} \quad ,$$

where each c_i belongs to $\mathbb{Q}$ by our previous reduction. The Multinomial Theorem shows that for each $k \ge 1$ the coefficient of z^{nd-k} in $h(z)^d$ is

$$a_k(d) = \sum_{\substack{m_1+2m_2+\dots+km_n=k \\ m_0+m_1+\dots+m_n=d \\ \text{all } m_i\ge 0}} \binom{d}{m_0\ m_1 \dots m_n} c_1^{m_1} \dots c_n^{m_n} \quad . \tag{11}$$

The sum $a_k(d)$ contains a bounded number of terms, independent of d. Furthermore, each multinomial coefficient

$$\binom{d}{m_0\ m_1 \dots m_n} = \frac{d(d-1)\dots(d-m_1-m_2-\dots-m_k+1)}{m_1!\ m_2!\ \dots m_k!} \quad .$$

has a proper algebraic factor of d, since there is some $i \ge 1$ with $m_i > 0$. Hence (11) exhibits $a_k(d)$ as a polynomial in d, with rational coefficients, having no constant term. It follows that for all integers d which are multiples of some fixed number D_k, the coefficient $a_k(d)$ is an integer.

Let T be a parameter, to be specified shortly. By choosing d to be a multiple of $\mathrm{LCM}(D_1,\dots,D_{nT})$, we can force the nT highest order coefficients of $h(z)^d$ to be integral. Patch the remaining coefficients by the same method used in the first step of Theorem 6.1.1 (but continue the process until all the coefficients are integers, instead of stopping at a certain bound). We thus obtain a polynomial

$$h^{(d)}(z) = h(z)^d + \sum_{j=T+1}^{d} \sum_{k=1}^{n} \Delta_{j,k}^{(d)} z^{n-k} h(z)^{d-j} \in \mathbb{Z}[z] \tag{12}$$

where for all j, k, $\Delta_{j,k}^{(d)}$ satisfies $|\Delta_{j,k}^{(d)}| \le \frac{1}{2}$.

We claim that $|h^{(d)}(z)| > 1$ for all $z \notin \mathcal{X}$ if T and d are chosen appropriately. To see this, consider $|[h^{(d)}(z)-h(z)^d]/h(z)^d|$ on $\partial\mathcal{X}$. By the definition of $\mathcal{X}$, $|h(z)^d| = R^d$ for all $z \in \partial\mathcal{X}$. Let $M = \sum_{k=1}^{n} \|z^{n-k}\|_{\partial\mathcal{X}}$. The sum which forms the second term in (12) has magnitude on $\partial\mathcal{X}$ bounded by

$$\tfrac{1}{2}M \sum_{j=T+1}^{d} R^{d-j} \le \frac{M}{R^T(R-1)} R^d .$$

Since $R > 1$, by choosing T sufficiently large we can force $M/[R^T(R-1)] < \frac{1}{2}$; observe that the choice of T depends only on $h(z)$ and $\mathcal{I}_r$, not on d. Then on $\partial\mathcal{X}$

$$\left|\frac{h^{(d)}(z)-h(z)^d}{h(z)^d}\right| = \Big|\sum_{j=T+1}^{d} \sum_{k=1}^{n} \Delta_{j,k}^{(d)} z^{n-k} h(z)^{d-j}\Big| / |h(z)|^d$$
$$\le \frac{MR^d}{R^T(R-1)} \cdot \frac{1}{R^d} \le 1/2 .$$

Now $\partial\mathcal{X}$ is not only the boundary of $\mathcal{X}$, but of $\mathbb{C}\backslash\mathcal{X}$, and $[h^{(d)}(z)-h(z)^d]/h(z)^d$ is holomorphic in $\mathbb{C}\backslash\mathcal{X}$ and vanishes at ∞. Hence the Maximum Modulus Principle shows that $|[h^{(d)}(z)-h(z)^d]/h(z)^d| \le \frac{1}{2}$ holds for all $z \notin \mathcal{X}$. Thus, for such z,

$$|h^{(d)}(z)| > \tfrac{1}{2}|h(z)|^d > \tfrac{1}{2}R^d$$

which in turn implies

$$\{ z \in \mathbb{C} : |h^{(d)}(z)| < \tfrac{1}{2}R^d \} \subset \mathcal{X} \subset U . \tag{13}$$

In order to assure that $h^{(d)}(z)$ has integer coefficients, we have required that d be divisible by $\mathrm{LCM}(D_1,\dots,D_{nT})$. Clearly we can also choose d large enough that $\frac{1}{2}R^d > 1$.

If these conditions are met then $f(z) := h^{(d)}(z)$ satisfies the needs of the Theorem. For, by construction $f(z)$ is a monic polynomial in $\mathbb{Z}[z]$, and by (13),

$$\{ z \in \mathbb{C} : |f(z)| \le 1 \} \subset U .$$

This completes the proof. //