8 The Gelfond-Schneider Theorem and Some Related Results

In this section, we begin by stating some results without proofs.

In 1900, David Hilbert posed a general problem which included determining whether $2^{\sqrt{2}}$ is transcendental and whether e^{π} is transcendental. The problem was resolved independently by Gelfond and Schneider in 1934. Their result is the following

Theorem 19. If α and β are algebraic numbers with $\alpha \neq 0$, $\alpha \neq 1$, and $\beta \notin \mathbb{Q}$, then α^{β} is transcendental.

Observe that the theorem asserts that any value of α^{β} is transcendental under the conditions above. It is clear that $2^{\sqrt{2}}$ is transcendental follows from this result, and since e^{π} is a value of i^{-2i} , the transcendence of e^{π} also follows from this result. We note that the following are equivalent forms of this result:

- (i) If ℓ and β are complex numbers with $\ell \neq 0$ and $\beta \notin \mathbb{Q}$, then at least one of the three numbers e^{ℓ} , β , and $e^{\beta \ell}$ is transcendental.
- (ii) If α and β are non-zero algebraic numbers with $\log \alpha$ and $\log \beta$ linearly independent over the rationals, then $\log \alpha$ and $\log \beta$ are linearly independent over the algebraic numbers.

Observe that (ii) is clearly equivalent to the statement that if α and β are non-zero algebraic numbers with $\beta \neq 1$ and $\log \alpha / \log \beta \notin \mathbb{Q}$, then $\log \alpha / \log \beta$ is transcendental.

Proofs of Equivalences. To see that Theorem 19 implies (i), take $\alpha = e^{\ell}$. Then α is not 0 or 1. Theorem 19 implies that if α and β are algebraic, then $\alpha^{\beta} = e^{\beta \ell}$ is transcendental, which implies (i).

To see that (i) implies (ii), observe that the condition $\log \alpha$ and $\log \beta$ are linearly independent over the rationals implies that both α and β are not 1. Also, we get that $\log \alpha / \log \beta$ is not in \mathbb{Q} . Let $\ell = \log \beta$ and $\beta' = \log \alpha / \log \beta$. Then (i) implies that β' is transcendental, which implies (ii).

To see that (ii) implies Theorem 19, consider $\beta' = e^{\beta \log \alpha}$. Then $\log \alpha$ and $\log \beta'$ are linearly dependent over the algebraic numbers. Hence, by (ii), $\log \alpha$ and $\log \beta'$ are linearly dependent over the rationals. This contradicts that $\beta \notin \mathbb{Q}$.

There are results similar to (i). For example, Lang proved that

Theorem 20. Suppose ℓ_1 , ℓ_2 , and ℓ_3 are linearly independent over the rationals and that β_1 and β_2 are are linearly independent over the rationals. Then at least one of the numbers $e^{\ell_i\beta_j}$ is transcendental.

Let γ be a transcendental number. If α is an algebraic number different from 0 and 1, then we can set $\ell_1 = 1$, $\ell_2 = \gamma$, $\ell_3 = \gamma^2$, $\beta_1 = \log \alpha$, and $\beta_2 = \gamma \log \alpha$ to obtain that at least one of α^{γ} , α^{γ^2} , or α^{γ^3} is transcendental. Another similar result was independently obtained by Brownawell and Waldschmidt which implies that either e^e or e^{e^2} is transcendental.

In 1966, Baker established the following generalization of the Gelfond-Schneider Theorem (Theorem 19).

Theorem 21. If $\alpha_1, \ldots, \alpha_m$ are non-zero algebraic numbers with $\log \alpha_1, \ldots, \log \alpha_m$ linearly independent over the rationals, then $\log \alpha_1, \ldots, \log \alpha_m$ are linearly independent over the algebraic numbers. To further illustrate some directions that transcendental number theory has taken and how such results can be applied to other areas of number theory, we give such an application to a result which is of a similar flavor to Theorem 21. Basically, the result helps resolve the question as to how "far from zero" is a linear combination of logarithms of algebraic numbers.

We say that an algebraic number α has degree d and height A if α satisfies an irreducible polynomial $f(x) = \sum_{j=0}^{d} a_j x^j \in \mathbb{Z}[x]$ with $a_d \neq 0$, $gcd(a_d, \ldots, a_1, a_0) = 1$, and $\max_{0 \le j \le d} |a_j| = A$.

Theorem 22. Let $\alpha_1, \ldots, \alpha_r$ be non-zero algebraic numbers with degrees at most d and heights at most A. Let $\beta_0, \beta_1, \ldots, \beta_r$ be algebraic numbers with degrees at most d and heights at most B > 1. Suppose that

$$\Lambda = \beta_0 + \beta_1 \log \alpha_1 + \dots + \beta_r \log \alpha_r \neq 0.$$

Then there are numbers C = C(r, d) > 0 and $w = w(r) \ge 1$ such that

$$|\Lambda| > B^{-C(\log A)^u}$$

Possibly do an application here (not included in the notes).

We turn now to proving Theorem 19.* Suppose α is an algebraic number with $\alpha \neq 0$ and $\alpha \neq 1$. Further suppose β is an algebraic number and that α^{β} is algebraic. Then Theorem 19 will follow if we can show that $\beta \in \mathbb{Q}$. We consider now the special case that $\alpha > 0$ and β are real (it would suffice to have $\log \alpha$ real).

Observe that $\alpha^{s_1+s_2\beta}$ is an algebraic number for all integers s_1 and s_2 . To establish the theorem, it suffices to show that there are two distinct pairs of integers (s_1, s_2) and (s'_1, s'_2) for which

$$s_1 + s_2\beta = s_1' + s_2'\beta$$

We will choose S sufficiently large and show such pairs exist with $0 \le s_1, s_2, s'_1, s'_2 < S$.

Lemma 1. Let $a_1(t), \ldots, a_n(t)$ be non-zero polynomials in $\mathbb{R}[t]$ of degrees d_1, \ldots, d_n respectively. Let w_1, \ldots, w_n be pairwise distinct real numbers. Then

$$F(t) = \sum_{j=1}^{n} a_j(t) e^{w_j t}$$

has at most $d_1 + \cdots + d_n + n - 1$ real zeroes (counting multiplicities).

Proof. By multiplying through by $e^{-w_n t}$ if necessary, we may suppose that $w_n = 0$ and that otherwise $w_j \neq 0$. Let $k = d_1 + \cdots + d_n + n$. We do induction on k. If k = 1, then n = 1 and $d_1 = 0$, and the lemma easily follows. Let $\ell \geq 2$ be such that the lemma holds whenever $k < \ell$, and suppose $k = \ell$. Let N be the number of real roots of F(t). By Rolle's Theorem, the number of real roots of F'(t) is at least N - 1. On the other hand,

$$F'(t) = \sum_{j=1}^{n} b_j(t) e^{w_j t}$$

^{*}The arguments given here should be compared to those given in Chapter 2 of Michel Waldschmidt, Diophantine Approximation on Linear Algebraic Groups, Grundlehren der Mathematischen Wissenschaften 326, Springer-Verlag, Berlin-Heidelberg, 2000.

where

$$b_j(t) = a'_j(t) + w_j a_j(t).$$

Note that for $1 \le j \le n-1$, $b_j(t)$ has degree exactly d_j . Also, since $w_n = 0$, either there are only n-1 non-zero polynomials $b_j(t)$ in the expression for F'(t) above or there are n such polynomials and the degree of $b_n(t)$ is one less than the degree of $a_n(t)$. We get from the induction hypothesis that F'(t) has at most $d_1 + \cdots + d_n + n - 2$ real roots. Hence, $N - 1 \le d_1 + \cdots + d_n + n - 2$, and the result follows.

We will make use of the following result concerning analytic functions. We omit the proof. It is a version of the Maximum Modulus Principle (and follows fairly easily from the Open Mapping Theorem). We use the notation $|f|_r$ to denote the maximum value of |f(z)| for |z| = r.

Lemma 2. Suppose f(z) is an analytic function in the disk $D = \{z : |z| < R\}$ and that it is continuous on $\overline{D} = \{z : |z| \le R\}$. Then

$$|f(z)| \le |f|_R$$

for every $z \in \overline{D}$.

Lemma 3. Let r and R be 2 real numbers with $1 \le r \le R$. Let $f_1(z), f_2(z), \ldots, f_L(z)$ be analytic in $D = \{z : |z| < R\}$ and continuous on $\overline{D} = \{z : |z| \le R\}$. Let ζ_1, \ldots, ζ_L be such that $|\zeta_j| \le r$ for each $j \in \{1, 2, \ldots, L\}$. Then the determinant

$$\Delta = \det \begin{pmatrix} f_1(\zeta_1) & \cdots & f_L(\zeta_1) \\ \vdots & \ddots & \ddots \\ f_1(\zeta_L) & \cdots & f_L(\zeta_L) \end{pmatrix}$$

satisfies

$$|\Delta| \le \left(\frac{R}{r}\right)^{-L(L-1)/2} L! \prod_{\lambda=1}^{L} |f_{\lambda}|_{R}.$$

Proof. Let h(z) be the determinant of the $L \times L$ matrix $(f_j(\zeta_i z))$. Then h(z) is analytic in $D' = \{z : |z| < R/r\}$ and continuous on $\overline{D'} = \{z : |z| \le R/r\}$. Let M = L(L-1)/2, and write

$$f_j(\zeta_i z) = \sum_{k=0}^{M-1} b_k(j) \zeta_i^k z^k + z^M g_{i,j}(z)$$

where $b_k(j) \in \mathbb{C}$ for each k and $g_{i,j}(z)$ is analytic in D' and continuous on $\overline{D'}$. Then since the determinant is linear in its columns (to see this evaluate along the columns), we can view h(z) as z^M times an analytic function plus terms involving the factor

$$z^{n_1+\cdots+n_L} \det\left(\zeta_i^{n_j}\right),$$

where the n_j denote non-negative integers. Observe that the determinant in this last expression is zero if the n_j are not distinct. Therefore, the non-zero terms of this form satisfy

$$n_1 + n_2 + \dots + n_L \ge 0 + 1 + \dots + (L-1) = \frac{L(L-1)}{2}$$

Hence, we deduce that h(z) is divisible by z^M . More precisely, $h(z)/z^M$ is analytic in D' and continuous on $\overline{D'}$. Therefore, by Lemma 2, for any $w \in \overline{D'}$,

$$\left|\frac{h(w)}{w^M}\right| \le \left|\frac{h(z)}{z^M}\right|_{R/r} = \left(\frac{r}{R}\right)^M |h(z)|_{R/r}.$$

For |z| = R/r, we get that $|\zeta_i z| \le R$. We bound $|h(z)|_{R/r}$ by multiplying the number of terms in det $(f_j(\zeta_i z))$ by an obvious upper bound on the maximum such term. Thus,

$$|h(z)|_{R/r} \le L! \prod_{\lambda=1}^{L} |f_{\lambda}|_{R}$$

Observe that $|\Delta| = |h(1)|$ and $1 \le R/r \le R$. We deduce that

$$|\Delta| \le \left(\frac{r}{R}\right)^M |h(z)|_{R/r} \le \left(\frac{r}{R}\right)^M L! \prod_{\lambda=1}^L |f_\lambda|_R,$$

giving the desired conclusion.

To complete our proof, we will also want a lower bound on $|\Delta|$ when $\Delta \neq 0$. The specific Δ that we will use has not yet been specified. For now, we note that it will have the form given in our previous lemma as well as the form given in our next lemma.

Lemma 4. Let

$$\Delta = \det(\alpha_{i,j})_{L \times L}$$

where the $\alpha_{i,j}$ are algebraic numbers. Suppose that T is a positive rational integer for which $T\alpha_{i,j}$ is an algebraic integer for every $i, j \in \{1, 2, ..., L\}$. Finally, suppose that $\Delta \neq 0$. Then there is a conjugate of Δ with absolute value $\geq T^{-L}$.

Proof. Observe that $T^L \Delta$ is an algebraic integer so that one of its conjugates has absolute value ≥ 1 . The result follows.

Let c be a sufficiently large real number (to be specified momentarily). Consider integers L_0 , L_1 , and S each ≥ 2 . Let $L = (L_0 + 1)(L_1 + 1)$. Observe that we can find such L_0 , L_1 , and S (and we do so) with

 $cL_0 \log S \le L$, $cL_1S \le L$, and $L \le (2S-1)^2$;

for example, take S large and

$$L_0 = [S \log S]$$
 and $L_1 = [S/\log S]$.

(Observe that we could take $c = \log \log S$.) We consider a matrix \mathcal{M} described as follows. Consider some arrangement $(s_1(i), s_2(i))$ of the $(2S - 1)^2$ integral pairs (s_1, s_2) with $|s_1| < S$ and $|s_2| < S$. Also, consider some arrangement (u(j), v(j)), with $1 \le j \le L$, of the integral pairs (u, v) where $0 \le u \le L_0$ and $0 \le v \le L_1$. Then we define

$$\mathcal{M} = \left((s_1(i) + s_2(i)\beta)^{u(j)} \left(\alpha^{s_1(i) + s_2(i)\beta} \right)^{v(j)} \right)$$

so that \mathcal{M} is a $(2S-1)^2 \times L$ matrix. The idea is to:

- (i) Consider the determinant Δ of an arbitrary $L \times L$ submatrix of \mathcal{M} (any *one* would do).
- (ii) Use Lemma 3 to obtain an upper bound B_1 for the absolute value of Δ (or, more specifically, an upper bound for the $\log |\Delta|$).
- (iii) Use Lemma 4 to motivate that if $\Delta \neq 0$, then Δ has absolute value $\geq B_2$ for some $B_2 > B_1$ (and assume this to be the case).
- (iv) Conclude that Δ must be 0 and, hence, the rank of \mathcal{M} is < L.
- (v) Take a linear combination of the columns of \mathcal{M} to obtain an F(t) as in Lemma 1 with < L roots but with $F(s_1(i) + s_2(i)\beta) = 0$ for $1 \le i \le L$.
- (vi) Conclude that β is rational as described at the beginning of this section.

Since we have not specified the arrangements defining (u(j), v(j)) and $(s_1(i), s_2(i))$, it suffices to consider $\Delta = \det(f_j(\zeta_i))$ where

$$f_j(z) = z^{u(j)} \alpha^{v(j)z} \ (1 \le j \le L)$$
 and $\zeta_i = s_1(i) + s_2(i)\beta \ (1 \le i \le L).$

Observe that u(j) is a non-negative integer for each j. Also, $\alpha^{v(j)z} = \exp(v(j)z \log \alpha)$, and we fix $\log \alpha$ so that it is real. Hence, $f_j(z)$ is uniquely defined. Then $f_j(z)$ represents an entire function for each j. Observe that

$$|e^{z_1 z_2}| = e^{\operatorname{Re}(z_1 z_2)} \le e^{|z_1 z_2|} = e^{|z_1||z_2|}$$

for all complex numbers z_1 and z_2 . Hence, for any R > 0,

$$|f_j|_R \le R^{u(j)} e^{v(j)R|\log\alpha|}.$$

We use Lemma 3 with $r = S(1 + |\beta|)$ and $R = e^2 r$. Then for some constant $c_1 > 0$, we obtain that

$$\log |\Delta| \le -L(L-1) + \log L! + L \max_{1 \le j \le L} \{ \log |f_j|_R \}$$

$$\le -L(L-1) + L \log L + LL_0 \log R + LL_1 R |\log \alpha|$$

$$\le -L^2 + c_1 (LL_0 \log S + LL_1 S).$$

The constant c_1 above is independent of c. Therefore, if c is sufficiently large (namely, $c \ge 4c_1$), then

$$\log|\Delta| \le -L^2/2.$$

Suppose now that T' is a positive rational integer for which $T'\alpha$, $T'\beta$, and $T'\alpha^{\beta}$ are all algebraic integers. Then $T = (T')^{L_0+2SL_1}$ has the property that T times any element of \mathcal{M} (and, hence, T times any element of the matrix describing Δ) is an algebraic integer. Therefore, by Lemma 4, if $\Delta \neq 0$, then there is a conjugate of Δ with absolute value

$$\geq T^{-L} = (T')^{-LL_0 - 2SLL_1}$$

It is reasonable (maybe not) to expect a similar inequality might hold for $|\Delta|$ itself (rather than for the absolute value of a conjugate of Δ). In fact, it can be shown (and will be shown later) that if $\Delta \neq 0$, then there is a constant c_2 (independent of c) for which

$$\log |\Delta| \ge -c_2 \left(LL_0 \log S + SLL_1 \right). \tag{10}$$

By using our upper bound for $\log |\Delta|$ above, we see that for c sufficiently large ($c \ge 8c_2$ will do), we obtain that $\Delta = 0$. Since $\Delta = \det (f_j(\zeta_i))$ as defined above, we get that the columns of $(f_j(\zeta_i))$ must be linearly dependent (over the reals). In other words, there exist real numbers b_j , not all 0, such that

$$\sum_{j=1}^{L} b_j f_j(\zeta_i) = 0 \quad \text{ for } 1 \le i \le L.$$

By considering a particular ordering of the (u(j), v(j)), we deduce that

$$\sum_{v=0}^{L_1} \sum_{u=0}^{L_0} b_{(L_0+1)v+u+1} \zeta_i^u \alpha^{v\zeta_i} = 0 \quad \text{for } 1 \le i \le L.$$

But

$$\sum_{v=0}^{L_1} \sum_{u=0}^{L_0} b_{(L_0+1)v+u+1} \zeta_i^u \alpha^{v\zeta_i} = \sum_{v=0}^{L_1} a_v(t) e^{w_v t}$$

where

$$a_v(t) = \sum_{u=0}^{L_0} b_{(L_0+1)v+u+1} t^u, \quad w_v = v \log \alpha, \quad \text{and} \quad t = \zeta_i = s_1(i) + s_2(i)\beta.$$

Each of the L values of ζ_i is a root of $\sum_{v=0}^{L_1} a_v(t)e^{w_v t} = 0$. Since some $b_j \neq 0$, we deduce from Lemma 1 that there are at most $L_0(L_1 + 1) + (L_1 + 1) - 1 < L$ distinct real roots. Therefore, two roots ζ_i must be the same, and we can conclude that

$$s_1(i) + s_2(i)\beta = s_1(i') + s_2(i')\beta$$
 for some i, i' with $1 \le i < i' \le L$.

On the other hand, the pairs $(s_1(i), s_2(i))$ and $(s_1(i'), s_2(i'))$ are necessarily distinct, so we can conclude that β is rational, completing the proof of Lemma 1.

We should note what assumptions we have made to obtain our result. First, we have used that $\alpha > 0$ and β are real (this was used in applying Lemma 1 at the end of the argument). Also, we have assumed the inequality in (10) above. This is it. In other words, if we want to complete the proof for $\alpha > 0$ and β real or for some specific choice of $\alpha > 0$ and β real, then we merely need to establish (10) (where in (10) we assume that $\Delta \neq 0$).

Before continuing, we show how the material just presented leads to a proof that $2^{\sqrt{2}}$ is irrational (not transcendental). We take $\alpha = 2$ and $\beta = \sqrt{2}$, and assume $\alpha^{\beta} = a/b$ with a and b positive rational integers. We can deduce our result from the arguments in the previous section provided we can show that if $\Delta \neq 0$, then (10) holds. Note that we are dealing with an irrationality proof here rather than a transcendence proof because we will make use of $\alpha^{\beta} = a/b$ to deduce (10). Recall that

$$\Delta = \det\left(f_j(\zeta_i)\right)$$

where

$$f_j(z) = z^{u(j)} \alpha^{v(j)z} \ (1 \le j \le L)$$
 and $\zeta_i = s_1(i) + s_2(i)\beta \ (1 \le i \le L).$

Also,

$$|s_1(j)| < S$$
, $|s_2(j)| < S$, $0 \le u(j) \le L_0$, and $0 \le v(j) \le L_1$

for all relevant values of j. Suppose $\Delta \neq 0$. Observe that

$$b^{SL_1} f_j(\zeta_i) = b^{SL_1} \zeta_i^{u(j)} \alpha^{v(j)s_1(i)} (\alpha^\beta)^{v(j)s_2(i)}$$

= $\left(s_1(i) + s_2(i)\sqrt{2}\right)^{u(j)} 2^{v(j)s_1(i)} a^{v(j)s_2(i)} b^{SL_1 - v(j)s_2(i)}.$

Hence,

$$b^{SL_1L}\Delta = A + B\sqrt{2}$$

where A and B are rational integers satisfying

$$\max\{|A|, |B|\} \le L! (2ab)^{SL_1L} (3S)^{L_0L}.$$

On the other hand, the lemma to Liouville's Theorem (i.e., see the lemma to Theorem 11) implies that if $B \neq 0$, then

$$|A + B\sqrt{2}| \ge c_1/|B|$$

for some constant $c_1 > 0$. We deduce that

$$b^{SL_1L}|\Delta| = |A + B\sqrt{2}| \ge \frac{c_1}{|B|} \ge c_1(L!)^{-1}(2ab)^{-SL_1L}(3S)^{-L_0L}.$$

Hence,

$$\log |\Delta| \ge \log c_1 - L \log L - SL_1L \left(\log(2ab) + \log b\right) - L_0L \log(3S).$$

This easily implies (10) in the case that $B \neq 0$. On the other hand, if B = 0, the argument is even easier (and in fact follows from Lemma 4 as outlined before (10)). More specifically, if B = 0 (and $\Delta \neq 0$), then

$$|\Delta| = |A|b^{-SL_1L} \ge b^{-SL_1L}$$

so that

$$\log |\Delta| \ge -SL_1 L \log b$$

from which (10) follows. Therefore, $2^{\sqrt{2}}$ is irrational.

To complete the proof of Theorem 19 when $\alpha > 0$ and β are real, it suffices to show that if $\Delta \neq 0$, then (10) holds. We are working under the assumptions that $\alpha \neq 1$ and α , β , and α^{β} are algebraic. The proof that $2^{\sqrt{2}}$ is irrational made use of an inequality of Liouville. The idea is to modify the inequality now to obtain the more general result. We will make use of the basic idea given in Lemma 4. We let T' be a positive rational integer for which $T'\alpha$, $T'\beta$, and $T'\alpha^{\beta}$ are all algebraic integers. Then $T = (T')^{L_0+2SL_1}$ has the property that T times any element of \mathcal{M} (and, hence, T times any element of the matrix describing Δ) is an algebraic integer. It follows that $T^L\Delta$ is an algebraic integer.

We suppose that $\Delta \neq 0$. For an algebraic number w, we denote by ||w|| the maximum of the absolute value of a conjugate of w. Then we obtain that

$$||T^{L}\Delta|| = T^{L}||\Delta|| \le T^{L}L!S^{L_{0}L}(1+||\beta||)^{L_{0}L}||\alpha||^{SL_{1}L}||\alpha^{\beta}||^{SL_{1}L}$$

(where the last inequality should be done with care noting that possibly $||\alpha||$ and $||\alpha^{\beta}||$ are < 1; alternatively, one can replace these with $||\alpha|| + 1$ and $||\alpha^{\beta}|| + 1$, respectively, and continue as below). Since $T^{L}\Delta$ is an algebraic integer in $\mathbb{Q}(\alpha, \beta, \alpha^{\beta})$, we deduce that $T^{L}\Delta$ is a root of a monic polynomial g(x) of degree N where N is the product of the degrees of the minimal polynomials for α, β , and α^{β} . Note that each root of g(x) can be made to be a conjugate of $T^{L}\Delta$. Since the product of all the roots of g(x) has absolute value $|g(0)| \ge 1$ and since each root of g(x) has absolute value $\le ||T^{L}\Delta||$, we obtain that

$$|T^{L}\Delta| \geq |g(0)|/||T^{L}\Delta||^{N-1}$$

$$\geq T^{-(N-1)L}(L!)^{-N}S^{-NL_{0}L}(1+||\beta||)^{-NL_{0}L}(||\alpha||||\alpha^{\beta}||+1)^{-NSL_{1}L}.$$

Hence,

$$\log |\Delta| \ge -NL \log T - NL \log L - NL_0 L \log S$$
$$-NL_0 L \log(1 + ||\beta||) - NSL_1 L \log(||\alpha|| ||\alpha^\beta|| + 1).$$

Recall that $T = (T')^{L_0+2SL_1}$. Here, T' and N are constants only depending on α and β . We now get that (10) holds, and the proof of Theorem 19 (for $\alpha > 0$ and β real) follows.

Homework:

- 1. Using Theorem 19, explain why $\log 2 / \log 3$ is transcendental.
- 2. Let a, b, and k be fixed positive integers. Using Theorem 21, explain why there is a constant C(a, b, k) such that the number of pairs of positive integers (m, n) for which

$$0 < a^n - b^m \le k$$

is $\leq C(a, b, k)$.