

**EXAMPLE SHEET 3 FOR DIOPHANTINE ANALYSIS,  
LENT 2026**

PÉTER VARJÚ

About this example sheet:

- Please send comments and corrections to pv270@dpmms.cam.ac.uk.
- Please submit your solutions of Problems 1 and 3, by Friday 6 March, 17:00.
- The purpose of this example sheet is to complement the material of the lectures. The level of difficulty of the problems varies considerably (in a non-monotone fashion), and they are not intended to be mock exam questions.

**1.** Let  $\alpha \neq 0$  be an algebraic number of degree  $d$  and  $\beta \neq 0$  an algebraic number of degree  $D$ . Suppose  $D > 2d$ .

Give a lower bound on  $|\alpha - \beta|$  by computing the height of  $\alpha - \beta$  and using the Liouville bound. The purpose of this question is to improve this bound. The actual values of the numerical constants are not important. If you obtain larger values, that is acceptable.

- (a) Let  $M \in \mathbf{Z}_{\geq 1}$  be such that  $D > 2Md$ . Prove that there is a polynomial  $P \in \mathbf{Z}[X]$  of degree  $D - 1$  that vanishes at  $\alpha$  to order  $M$  with

$$H(P) \leq D^{2M^2d/D} H(\alpha)^{2Md}.$$

- (b) Prove that

$$|P(\beta)| \geq H(P)^{-D} \cdot D^{-D} \cdot H(\beta)^{-D^2}.$$

- (c) Prove that

$$|P(\beta)| \leq D^{M+1} H(P) H(\alpha)^{dD} H(\beta)^{D^2} |\alpha - \beta|^M$$

- (d) Prove that there is an absolute constant  $C$  such that

$$|\alpha - \beta| > D^{-C(Dd)^{1/2}} M(\alpha)^{-CD} M(\beta)^{-C(Dd)^{1/2}}.$$

*Hint:* You should get  $|\alpha - \beta| > 2^{-dD} M(\alpha)^{-D} M(\beta)^{-d}$  from the Liouville bound. (a): Use Siegel's lemma and  $D - Md > D/2$ . (b): Estimate the height and use the Liouville bound. You also need to show  $P(\beta) \neq 0$ . (c): Use Taylor expansion around  $\alpha$  to order  $M$  with remainder term. Use  $\max(|\alpha|, |\beta|) < H(\alpha)^d H(\beta)^D$ .

## Solution

We have  $H(\alpha - \beta) \leq 2H(\alpha)H(\beta)$  and  $\deg(\alpha - \beta) \leq dD$ . Since  $\alpha$  and  $\beta$  have different degrees, they are not equal and  $\alpha - \beta \neq 0$ . Now Liouville's inequality gives

$$\begin{aligned} |\alpha - \beta| &\geq H(\alpha - \beta)^{-\deg(\alpha - \beta)} \geq (2H(\alpha)H(\beta))^{-dD} \\ &= 2^{-dD} M(\alpha)^{-D} M(\beta)^{-d}. \end{aligned}$$

(a) We use Siegel's lemma for the system of linear equations

$$\begin{aligned} 0 &= a_{D-1}\alpha^{D-1} + \dots + a_1\alpha + a_0 \\ 0 &= a_{D-1}\binom{D-1}{1}\alpha^{D-2} + \dots + a_1\binom{1}{1}\alpha^0 \\ &\vdots \\ 0 &= a_{D-1}\binom{D-1}{M-1}\alpha^{D-M} + \dots + a_{M-1}\binom{M-1}{M-1}\alpha^0, \end{aligned}$$

where  $a_0, \dots, a_{D-1}$  are the variables. We will then set  $P(X) = a_{d_1}X^{d-1} + \dots + a_1X + a_0$ , where the coefficients are integer solutions, not all zero, of the above system of equations.

We estimate the height of the equations. For a finite place  $v \in M_{\mathbf{Q}(\alpha)}$  and  $m \leq M - 1$ , we have

$$\max_{j=0, \dots, D-1} \left( \left| \binom{j}{m} \alpha^{j-m} \right|_v \right) = \max(1, |\alpha|_v)^{D-1-m} \leq \max(1, |\alpha|_v)^D.$$

For an infinite place, we have

$$\begin{aligned} \max_{j=0, \dots, D-1} \left( \left| \binom{j}{m} \alpha^{j-m} \right|_v \right) &\leq D^m \max(1, |\alpha|_v)^{D-1-m} \\ &\leq D^{M-1} \max(1, |\alpha|_v)^D, \end{aligned}$$

where we used the estimate  $\binom{j}{m} \leq D^m$ .

Using the definition, the height of the  $m$ th equation is bounded by

$$\begin{aligned} &H_p \left( \binom{D-1}{m} \alpha^{D-1-m}, \dots, \binom{m}{m} \alpha^{m-m} \right) \\ &\leq \prod_{v \in M_{\mathbf{Q}(\alpha), f}} \max(1, |\alpha|_v)^{d_v D/d} \prod_{v \in M_{\mathbf{Q}(\alpha), \infty}} D^{(M-1)d_v/d} \max(1, |\alpha|_v)^{d_v D/d} \\ &= D^{M-1} H(\alpha)^D. \end{aligned}$$

There are  $M$  equations with coefficients in the number field  $\mathbf{Q}(\alpha)$  that is of degree  $d$ . The number of variables is  $D$ . Siegel's lemma shows

that there is a non-zero integer solution  $a_0, \dots, a_{D-1}$  with

$$\begin{aligned} |a_j| &\leq (D \cdot D^{M-1} H(\alpha)^D)^{\frac{dM}{D-dM}} \leq (D \cdot D^{M-1} H(\alpha)^D)^{\frac{2dM}{D}} \\ &= D^{2M^2d/D} H(\alpha)^{2Md}, \end{aligned}$$

where we estimated the exponent as  $\frac{dM}{D-dM} \leq \frac{dM}{D/2}$ .

(b) We know that  $P(\beta) \neq 0$ , because  $\beta$  is of degree  $D > \deg P$ . We estimate the height of  $P(\beta)$  as

$$H(P(\beta)) \leq \mathcal{L}(P)H(\alpha)^{\deg(P)} \leq DH(P)H(\beta)^D.$$

Now the Liouville bound gives

$$|P(\beta)| \geq H(P(\beta))^{-D} \geq H(P)^{-D} D^{-D} H(\beta)^{-D^2},$$

where we used  $\deg P(\beta) \leq \deg \beta = D$ .

(c) We use that

$$P(\beta) = \int_{\alpha}^{\beta} \frac{1}{(M-1)!} (\beta - z)^{M-1} P^{(M)}(z) dz,$$

which can be proved using integration by parts and  $P^{(m)}(\alpha) = 0$  for  $m = 0, \dots, M-1$ . Here the integral is along the straight line connecting  $\alpha$  to  $\beta$ . (This is the integral form of the remainder term in Taylor's theorem.)

We have

$$|P(\beta)| \leq \frac{1}{(M-1)!} |\alpha - \beta|^M \max_z |P^{(M)}(z)|.$$

where  $z$  in the max runs over the line segment connecting  $\alpha$  and  $\beta$ . (When  $\alpha$  and  $\beta$  are real, you can also prove this using the Lagrange or the Cauchy form of the remainder.) We estimate

$$|P^{(M)}(z)| \leq DH(P^{(M)}) \max(|\alpha|, |\beta|)^D \leq D^{M+1} H(P) H(\alpha)^{dD} H(\beta)^{D^2},$$

and the claim follows if we plug this in to our previous bound on  $|P(\beta)|$ .

(d) Comparing the upper and lower bounds for  $|P(\beta)|$ , we get

$$\begin{aligned} |\alpha - \beta|^M &\geq D^{-M-1} H(P)^{-1} H(\alpha)^{-dD} H(\beta)^{-D^2} H(P)^{-D} D^{-D} H(\beta)^{-D^2} \\ &= D^{-M-1-D} H(P)^{-D-1} H(\alpha)^{-dD} H(\beta)^{-2D^2} \\ &\geq D^{-2D} H(P)^{-2D} H(\alpha)^{-dD} H(\beta)^{-2D^2}. \end{aligned}$$

Plugging in the bound for  $H(P)$  from part (a), we get

$$\begin{aligned} |\alpha - \beta|^M &\geq D^{-2D} D^{-4M^2d} H(\alpha)^{-4MdD} H(\alpha)^{-dD} H(\beta)^{-2D^2} \\ &= D^{-4M^2d-2D} H(\alpha)^{-5MdD} H(\beta)^{-2D^2}, \end{aligned}$$

which gives

$$|\alpha - \beta| \geq D^{-4Md-2D/M} H(\alpha)^{-5dD} H(\beta)^{-2D^2/M}.$$

The first factor is optimized when we choose  $M$  approximately  $(D/d)^{1/2}$ . It might be possible to improve this, when  $H(\beta)$  is large. To prove the bound asked for in the question, we need to choose an integer  $1 \leq M < D/2d$  such that  $c_1(D/d)^{1/2} \leq M \leq C_1(D/d)^{1/2}$  for some constants  $c_1$  and  $C_1$ . Taking  $M = \lfloor (D/2d)^{1/2} \rfloor$ , this holds with  $c_1 = 2^{-1/2}$  and  $C_1 = 2^{1/2}$ .

**2.** Let  $w_1, \dots, w_n$  be distinct real numbers, let  $d_1, \dots, d_n$  be non-negative rational integers, and let  $u_1, \dots, u_N$  be distinct real numbers for  $N = d_1 + \dots + d_n + n - 1$ . Show that there exist polynomials  $a_1, \dots, a_n \in \mathbf{R}[X]$  of degrees  $d_1, \dots, d_n$  respectively, such that the function

$$F(X) = \sum_{j=1}^n a_j(X) \exp(w_j X)$$

has a simple zero at each  $u_i$ , and no more zeros.

*Hint:* Use linear algebra to find a non-zero  $F$  that vanishes at the prescribed points. Use the zero estimate to show it has no more zeroes.

Solution

We consider  $\mathbf{R}[X]^{(d)}$  the vector space of polynomials of degree at most  $d$  and look for a solution of the system of equations

$$\sum_{j=1}^n a_j(u_i) \exp(w_j u_i)$$

for each  $i = 1, \dots, N$ , where we consider  $a_j \in \mathbf{R}[X]^{(d_j)}$  as variables. The dimension of the space  $\mathbf{R}[X]^{(d_1)} + \dots + \mathbf{R}[X]^{(d_n)}$  of variables is  $d_1 + \dots + d_n + n = N + 1$ . The number of equations is  $N$ , so there is a solution such that not all  $a_j$  is 0 by linear algebra.

By Proposition 42, the function  $\sum a_j(X) \exp(w_j X)$  cannot have more than  $N$  zeros counting multiplicities, so every  $u_i$  is a simple zero, and there are no other zeroes.

**3.** Let  $w_1, \dots, w_k$  be  $\mathbf{Q}$  linearly independent elements of  $\mathbf{C}^n$ . Show that the functions

$$\mathbf{C}^n \rightarrow \mathbf{C} : x = (x_1, \dots, x_n) \mapsto \exp(w_j \cdot x)$$

for  $j = 1, \dots, k$  are algebraically independent over the field  $\mathbf{Q}(x_1, \dots, x_n)$ .

*Hint:* You need to show that

$$\sum_{l_1, \dots, l_k} a_{l_1, \dots, l_k}(x) \exp((l_1 w_1 + \dots + l_k w_k) \cdot x) \neq 0$$

for any choice of non-zero  $a_{l_1, \dots, l_k} \in \mathbf{Q}[x_1, \dots, x_n]$ , where the indices  $l_1, \dots, l_k$  run through some finite range. Prove this by induction on the sum of the total degrees of the  $a$ 's by a similar argument to Proposition 42.

Solution

We need to show that if

$$P(X_1, \dots, X_k) = \sum_{l_1, \dots, l_k} a_{l_1, \dots, l_k}(x) X_1^{l_1} \cdots X_k^{l_k}$$

is a non-zero polynomial, that is, not all  $a_{l_1, \dots, l_k}$  is zero, then

$$P(\exp(w_1 \cdot x), \dots, \exp(w_k \cdot x)) \neq 0.$$

This is equivalent to

$$\sum_{l_1, \dots, l_k} a_{l_1, \dots, l_k}(x) \exp((l_1 w_1 + \dots + l_k w_k) \cdot x) \neq 0.$$

We note that by  $\mathbf{Q}$ -linear independence of  $w_1, \dots, w_k$ , the vectors  $l_1 w_1 + \dots + l_k w_k$  are all distinct. We prove the following formally stronger statement. Suppose that  $b_1, \dots, b_m \in \mathbf{Q}(x_1, \dots, x_n)$  are non-zero polynomials and  $v_1, \dots, v_m \in \mathbf{C}^n$  are distinct vectors. Then

$$(1) \quad \sum_{j=1}^m b_j(x) \exp(v_j \cdot x) \neq 0.$$

We write

$$N = \sum_{j=1}^m (\deg(b_j) + 1),$$

where the degree is the total degree, and prove the claim by induction on  $N$ . If  $N = 1$ , then there is only one term in the sum, and the claim is trivial.

Now suppose  $N > 1$ , and the claim holds for lower values on  $N$ . We note that

$$\sum_{j=1}^m b_j(x) \exp((v_j - v_1) \cdot x) = \exp(-v_1 \cdot x) \sum_{j=1}^m b_j(x) \exp(v_j \cdot x)$$

is non-zero if and only if (1) holds. We can, therefore, assume that  $v_1 = 0$ .

Now write

$$\frac{\partial}{\partial x_i} \sum_{j=1}^m b_j(x) \exp(v_j \cdot x) = \sum_{j=1}^m \tilde{b}_j(x) \exp(v_j \cdot x).$$

If this function is non-zero for some  $i$ , then (1) also holds. Note that  $\tilde{b}_1 = \partial b_1 / \partial x_i$  is either 0 or a non-zero polynomial of lower degree than  $b_1$ . In either case the value of  $N$  decreases, so we are done by the induction hypothesis if we can find a suitable  $i$  so that not all  $\tilde{b}_j$  is zero.

We consider two cases. The first is  $m = 1$ . In this case, the claim follows in the same way as in the  $N = 1$  case, because there is only

one term in the sum. Alternatively, we can finish the induction step, by noting that if  $N > 1$ , then  $b_1$  is not the constant polynomial and hence has a non-zero partial derivative.

The second case is  $m > 1$ . In this case,  $v_2 \neq 0$ , because the vectors  $v_j$  are distinct. In particular  $v_2$  has a non-zero component, say the  $i$  component. Then  $\tilde{b}_2$  has the same degree in  $x_i$  as  $b_2$ , so  $\tilde{b}_2 \neq 0$ , and we are done.

4. The goal of this question is to give an alternative proof of the real case of the Gelfond Schneider theorem based on Gelfond's proof. Now let  $\lambda_1, \lambda_2 \in \mathbf{R}_{\neq 0}$  and suppose  $\alpha_1 = e^{\lambda_1}, \alpha_2 = e^{\lambda_2}, \beta = \lambda_2/\lambda_1$  are all algebraic and  $\beta \notin \mathbf{Q}$ . We aim to derive a contradiction.

- (a) Let  $L \in \mathbf{Z}_{\geq 1}$  and let  $f_1, \dots, f_L : \mathbf{C} \rightarrow \mathbf{C}$  be entire functions. Let  $S_0, S_1 \in \mathbf{Z}_{\geq 0}$  with  $L = (S_0 + 1)S_1$ , and let  $\xi_1, \dots, \xi_{S_1} \in \mathbf{C}$ . Let  $r \in \mathbf{R}_{> 0}$  with  $|\xi_s| \leq r$  for  $s = 1, \dots, S_1$ . Let  $E \in \mathbf{R}_{\geq 1}$ . Prove that

$$\det[(d^\sigma/dz^\sigma)f_t(\xi_s)]_{\sigma,s}^t \leq E^{-L(L-1)/2+S_0(S_0+1)S_1/2} \cdot L! \cdot \prod_{t=1}^L \max_{\sigma=0,\dots,S_0} |(d^\sigma/dz^\sigma)f_t|_{Er}.$$

The indices in the determinant run through the ranges  $t = 1, \dots, L$ ,  $\sigma = 0, \dots, S_0$  and  $s = 1, \dots, S_1$ .

- (b) Now let  $T, S_0, S_1, L \in \mathbf{Z}_{\geq 0}$  be such that

$$L = (2T + 1)^2 = (S_0 + 1)(2S_1 + 1),$$

and consider the determinant

$$\Delta = \det[(d^\sigma/dz^\sigma) \exp((t_1 + \beta t_2)z)|_{z=\lambda_1 s}]_{\sigma,s}^{t_1, t_2},$$

where the indices run through the ranges  $t_1, t_2 = -T, \dots, T$ ,  $\sigma = 0, \dots, S_0$  and  $s = -S_1, \dots, S_1$ .

Use (a) to give an upper bound on  $|\Delta|$ .

- (c) Use Proposition 42 to show that  $\Delta \neq 0$ .  
 (d) Show that  $\Delta$  is algebraic, estimate its height and find a contradiction with an appropriate choice of the parameters.  
 (e) Compare  $\Delta$  with the determinant in Schneider's proof.

*Hint:* (a): Try the same proof as in Proposition 38, but now you cannot pull out factors of  $z^{\deg f_t}$  from the rows. Instead try to expand the determinant and estimate the degree of each term. (c): Same as the proof of Proposition 41, but now you need to count zeros with multiplicities. (e): Gelfond's  $\Delta$  is the transpose of Schneider's.

Solution

- (a) We use Schwarz's lemma for the function

$$F(u) := \det[(d^\sigma/dz^\sigma)f_t|_{z=\xi_s u}]_{\sigma,s}^t.$$

Estimating the determinant by expanding it into  $L!$  terms, we have

$$|F|_E \leq L! \cdot \prod_{t=1}^L \max_{\sigma=0, \dots, S_0} |(d^\sigma/dz^\sigma) f_t|_{Er},$$

so the claim will follow if we show that  $F$  has a zero of order at least  $L(L-1)/2 - S_0(S_0+1)S_1/2$  at  $z=0$ .

To do this, we assume similarly to the lectures that we have  $f_t = z^{k_t}$  for some  $k_t \in \mathbf{Z}_{\geq 0}$  for each  $t$ . (As we did in the lectures, it is possible to expand each  $f_t$  as a power series. Expanding the determinant, we get a series for  $F$ , which can be shown to converge and we show that each term in the series has the required order of vanishing at  $z=0$ .) If two  $k_t$  coincides, then the determinant is identically 0, so we can assume that this is not the case. Now the determinant is the sum of products

$$\prod_{t=1}^L \frac{d^{\sigma_t} z^{k_t}}{dz^{\sigma_t}} \Big|_{z=\xi_{s_t} u},$$

where  $(\sigma_t, s_t)$  runs through  $\{0, \dots, S_0\} \times \{1, \dots, S_1\}$  in some order. Such a product is divisible by  $u^{\sum_{t=1}^L (k_t - \sigma_t)}$  if  $k_t \geq \sigma_t$  for all  $t$ . Otherwise the product is identically 0.

Now

$$\sum_{t=1}^L k_t \geq \sum_{t=1}^L t = L(L-1)/2$$

and

$$\sum_{t=1}^L \sigma_t = S_1 \sum_{j=0}^{S_0} j = S_1 S_0 (S_0 + 1)/2$$

irrespectively of the order of  $\sigma_t$  running through  $\{0, \dots, S_0\}$   $S_1$  times. The claim follows.

(b) Using

$$f_{t_1, t_2} = \exp((t_1 + \beta t_2)z)$$

and  $\xi_s = \lambda_1 s$ , the determinant  $\Delta$  is of the form given in Part (a). We have  $|\xi_s| \leq CS_1 =: r$  for some constant  $C$  depending on  $\lambda_1$ , and

$$\left| \frac{d^\sigma}{dz^\sigma} f_{t_1, t_2} \right|_{Er} \leq \exp(C(\log T)S_0 + CETS_1)$$

for another constant  $C$  depending on  $\lambda_1$  and  $\lambda_2$ .

Now applying the bound in Part (a), we get

$$\log |\Delta| \leq -c(\log E)L^2 + C(\log E)S_0^2 S_1 + C(\log ET)S_0 L + CETS_1 L$$

for suitable  $c > 0$  and  $C$  depending on  $\lambda_1$  and  $\lambda_2$ .

(c) Suppose to the contrary that  $\Delta = 0$ . Then there are  $a_{t_1, t_2} \in \mathbf{R}$  for  $t_1$  and  $t_2$  running between  $-T, \dots, T$ , not all 0, such that the linear

combination of the rows of  $\Delta$  with these coefficients is the 0 vector. This means that the function

$$G(z) := \sum_{t_1, t_2} a_{t_1, t_2} \exp((t_1 + \beta t_2)z)$$

has the property that  $(d^\sigma/dz^\sigma)G|_{z=\lambda_1 s} = 0$  for  $\sigma = 0, \dots, S_0$  and  $s = -S_1, \dots, S_1$ , so  $G$  has  $(S_0 + 1)(2S + 1) = L$  zeroes counting multiplicities. This is more than  $L - 1$ , which contradicts Proposition 42.

(d) Using that  $\exp(\lambda_1) = \alpha_1$  and  $\exp(\beta\lambda_1) = \alpha_2$ , we can write

$$\Delta = [(t_1 + \beta t_2)^\sigma \alpha_1^{t_1 s} \alpha_2^{t_2 s}]_{\substack{t_1, t_2 \\ \sigma, s}}.$$

This is obtained from algebraic numbers with field operations, so it is algebraic. We estimate the height of  $\Delta$  using Proposition 28 thinking of  $\Delta$  as a polynomial in  $\beta$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_1^{-1}$  and  $\alpha_2^{-1}$ . We expand the determinant, we obtain  $L!$  terms, and in each term we expand the products of powers of  $(t_1 + \beta t_2)$ . This way, we obtain  $\leq L! \cdot 2^{S_0 L}$  terms. Each has an integer coefficient, which is the  $\leq S_0 L$ -fold product of numbers that arise as  $t_1$  or  $t_2$ . The length of the polynomial is at most  $L! \cdot (2T)^{S_0 L}$ . In this polynomial, the exponent of  $\beta$  is not more than  $LS_0$ , and the maximal exponents of each of  $\alpha_1$ ,  $\alpha_2$ ,  $\alpha_1^{-1}$  and  $\alpha_2^{-1}$  is not more than  $TS_1 L$ . By Proposition 28, we get

$$\begin{aligned} H(\Delta) &\leq L!(2T)^{LS_0} H(\beta)^{LS_0} H(\alpha_1)^{TS_1 L} H(\alpha_1^{-1})^{TS_1 L} \\ &\quad H(\alpha_2)^{TS_1 L} H(\alpha_2^{-1})^{TS_1 L} \\ &\leq \exp(L \log L + C(\log T)S_0 L + CTS_1 L) \end{aligned}$$

for a suitable constant  $C$  depending only on  $\lambda_1$  and  $\lambda_2$ .

By the Liouville bound and  $\Delta \neq 0$ , we have

$$|\Delta| \geq \exp(-CL \log L - C(\log T)S_0 L - CTS_1 L)$$

This contradicts the upper bound in Part (b) if  $c \log(E)L^2$  is larger than each of

$$CL \log L, \quad C(\log E)S_0^2 S_1, \quad C(\log ET)S_0 L, \quad CETS_1 L,$$

where  $c$  and  $C$  are some constants depending only on  $\lambda_1$  and  $\lambda_2$ .

Now we take  $E = 10$ ,  $T$  approximately  $L^{1/2}$ ,  $S_0$  approximately  $L^{1-\varepsilon}$  and  $S$  approximately  $L^\varepsilon$ . It is easy to check that this is suitable if  $L$  is sufficiently large.

(e) The determinant appearing in the question is simply the transpose of the determinant that appeared in the proof presented in the lectures. Therefore, the proofs of Gelfond and Schneider are duals of each other in a certain sense. This was not immediately obvious in their original formulation, which was based on constructions of auxiliary functions using Siegel's lemma.

5. The goal of this question is to give a proof of Dyson’s diophantine exponent without using Siegel’s lemma. Let  $\alpha \in \mathbf{R} \cap \overline{\mathbf{Q}} \setminus \mathbf{Q}$ , and let  $\varepsilon > 0$ . Suppose to the contrary that there are  $p_1/q_1, p_2/q_2 \in \mathbf{Q}$  with  $|\alpha - p_j/q_j| < q_j^{-\sqrt{2d}-\varepsilon}$  for  $j = 1, 2$  and such that  $q_1$  and  $\log q_2/\log q_1$  are both large in terms of  $\alpha$  and  $\varepsilon$ .

Consider the matrix

$$M = [\partial_{\sigma_1, \sigma_2} X^{t_1} Y^{t_2} |_{(X, Y) = \xi_s}]_{\substack{t_1, t_2 \\ \sigma_1, \sigma_2, s}},$$

where

$$\partial_{\sigma_1, \sigma_2} = \frac{\partial^{\sigma_1 + \sigma_2}}{\sigma_1! \sigma_2! \partial X^{\sigma_1} \partial Y^{\sigma_2}}.$$

The indices  $t_1$  and  $t_2$  run through the ranges  $0, \dots, n_1$  and  $0, \dots, n_2$ , where  $n_1, n_2 \in \mathbf{Z}_{\geq 0}$  are large and such that  $n_1 \log q_1$  and  $n_2 \log q_2$  are close to each other. We take  $\xi_1 = (\alpha, \alpha)$ , and let  $\xi_2, \dots, \xi_d$  be the Galois conjugates of  $(\alpha, \alpha)$ . We take  $\xi_{s+1} = (p_1/q_1, p_2/q_2)$ . The indices  $s, \sigma_1, \sigma_2$  run through

$$\{(s, \sigma_1, \sigma_2) : s = 1, \dots, d, \sigma_1/n_1 + \sigma_2/n_2 \leq \frac{2}{\sqrt{2d}} - \delta_1\} \cup \{(d+1, \sigma_1, \sigma_2) : \sigma_1/n_1 + \sigma_2/n_2 \leq \delta_2\}$$

for appropriate parameters  $\delta_1, \delta_2 > 0$ , which will be chosen in terms of  $\varepsilon$  and  $\alpha$ .

- (a) Use Dyson’s lemma to show that  $M$  has rank  $L := (n_1 + 1) \cdot (n_2 + 1)$ .
- (b) Show that you can find an  $L \times L$  submatrix with nonzero determinant  $\Delta$  of  $M$  that includes a maximal linearly independent subfamily of the columns that correspond to  $s \leq d$ .
- (c) Consider the Taylor expansion of all entries in the columns that correspond to  $s = d + 1$  around  $(X, Y) = (\alpha, \alpha)$ , and use this to give an upper bound on  $|\Delta|$ .
- (d) Estimate  $H(\Delta)$  and find a contradiction for an appropriate choice of the parameters.

*Hint:* (a): Consider a linear combination of the rows of  $M$  and show that the corresponding polynomial cannot have so much vanishing to make the linear combination all 0 if  $\delta_1, \delta_2$  are appropriately chosen and  $n_2/n_1$  is small enough, which you may force by requiring that  $\log q_2/\log q_1$  is large. (c): Use column operations on the determinant to remove low degree terms from the Taylor expansions. You need to use that you have enough columns in your determinant that span the space generated by all columns of  $M$  corresponding to  $s = 1$ . (d): The standard Liouville bound will not be enough, but you can prove that there is always a Galois embedding  $\pi : \mathbf{Q}(\alpha) \rightarrow \mathbf{C}$  such that  $|\pi(\Delta)| > H(\Delta)^{-1}$ . Observe that you may replace  $\Delta$  by  $\pi(\Delta)$  and the previous parts will not be affected.

Solution
(a) Suppose to the contrary that the rank of $M$ is less than $L$ . Then the rows of the matrix are linearly dependent, and there are some complex

numbers  $a_{t_1, t_2}$ , not all zero, such that the linear combination of the rows with these coefficients is the all 0 vector. That is, the polynomial

$$P(X, Y) = \sum_{t_1, t_2} a_{t_1, t_2} X^{t_1} Y^{t_2}$$

satisfies  $\partial_{\sigma_1, \sigma_2} P|_{(X, Y) = \xi_s} = 0$  for all  $(s, \sigma_1, \sigma_2)$  in the range specified in the question.

Now we estimate  $I_P(\xi_s; n_1^{-1}, n_2^{-1})$  for each  $s$ . We have

$$I_P(\xi_s; n_1^{-1}, n_2^{-1}) \geq \frac{2}{\sqrt{2d}} - \delta_1 =: t_s$$

for  $s = 1, \dots, d$  and

$$I_P(\xi_{d+1}; n_1^{-1}, n_2^{-1}) \geq \delta_2 =: t_{d+1}$$

using the vanishing of the derivatives stated above and the definition of the index. This contradicts Dyson's lemma if we have

$$\sum_{s=1}^{d+1} \frac{t_s^2}{2} > 1 + \frac{d-1}{2} \cdot \frac{n_2}{n_1}.$$

We have

$$\begin{aligned} \sum_{s=1}^{d+1} \frac{t_s^2}{2} &= d \left( \frac{2}{\sqrt{2d}} - \delta_1 \right)^2 / 2 + \delta_2^2 / 2 > d \left( \frac{2}{2d} - \frac{2\delta_1}{\sqrt{2d}} \right) + \delta_2^2 / 2 \\ &> 1 - 2\delta_1 \sqrt{d} + \delta_2^2 / 2. \end{aligned}$$

This satisfies our requirements if we choose  $\delta_1$  and  $\delta_2$  is such a way that  $3\sqrt{d}\delta_1 < \delta_2^2/2$  and then make sure that  $n_2/n_1 \approx \log q_1 / \log q_2 < \sqrt{d}\delta_1$ .

(b) Since  $M$  has rank  $L$ , its columns span  $\mathbf{C}^L$ . We begin by selecting linearly independent columns among the columns that correspond to  $s \leq d$ , and do this until there are no more columns among these that are linearly independent from those already selected. Finally we select columns among those that correspond to  $s = d+1$  until we get a basis of  $\mathbf{C}^L$ .

(c) Let  $f \in \mathbf{R}[X, Y]$  be a polynomial of degree at most  $n_1$  in  $X$  and at most  $n_2$  in  $Y$ . The Taylor expansion of  $f$  around  $(\alpha, \alpha)$  gives

$$f(p_1/q_1, p_2/q_2) = \sum_{\rho_1=1}^{n_1} \sum_{\rho_2=1}^{n_2} \partial_{\rho_1, \rho_2} f(\alpha, \alpha) (p_1/q_1 - \alpha)^{\rho_1} (p_2/q_2 - \alpha)^{\rho_2}.$$

Applying this for  $\partial_{\sigma_1, \sigma_2} f$  and taking into account the factorials in the definition of  $\partial_{\sigma_1, \sigma_2}$ , we have

$$\begin{aligned} \partial_{\sigma_1, \sigma_2} f(p_1/q_1, p_2/q_2) &= \sum_{\rho_1=1}^{n_1} \sum_{\rho_2=1}^{n_2} \binom{\rho_1 + \sigma_1}{\sigma_1} \binom{\rho_2 + \sigma_2}{\sigma_2} \\ &\quad \cdot \partial_{\rho_1 + \sigma_1, \rho_2 + \sigma_2} f(\alpha, \alpha) (p_1/q_1 - \alpha)^{\rho_1} (p_2/q_2 - \alpha)^{\rho_2}. \end{aligned}$$

We now consider a column of  $\Delta$  that corresponds to  $s = d + 1$  and some  $\sigma_1$  and  $\sigma_2$ . (Later we will see that  $\Delta$  contains quite a few such columns.) Writing  $f_{t_1, t_2} = X^{t_1} Y^{t_2}$ , this column has entries

$$(2) \quad \sum_{\rho_1=1}^{n_1} \sum_{\rho_2=1}^{n_2} \binom{\rho_1 + \sigma_1}{\sigma_1} \binom{\rho_2 + \sigma_2}{\sigma_2} \partial_{\rho_1 + \sigma_1, \rho_2 + \sigma_2} f_{t_1, t_2}(\alpha, \alpha) \left(\frac{p_1}{q_1} - \alpha\right)^{\rho_1} \left(\frac{p_2}{q_2} - \alpha\right)^{\rho_2}.$$

If  $\rho_1$  and  $\rho_2$  are such that

$$(\rho_1 + \sigma_1)/n_1 + (\rho_2 + \sigma_2)/n_2 \leq \frac{2}{\sqrt{2d}} - \delta_1,$$

then  $M$  contains the column whose entries are

$$\partial_{\rho_1 + \sigma_1, \rho_2 + \sigma_2} f_{t_1, t_2}(\alpha, \alpha).$$

This column is in the linear span of the columns of  $\Delta$  corresponding to  $s \leq d$ , so it can be written as a linear combination of columns of  $\Delta$  without using the column with entries (2). Performing an appropriate column operations, we can therefore replace the column with entries (2) with a column whose entries are

$$(3) \quad \sum_{\rho_1, \rho_2: \frac{\rho_1}{n_1} + \frac{\rho_2}{n_2} > \frac{2}{\sqrt{2d}} - \delta_1 - \frac{\sigma_1}{n_1} + \frac{\sigma_2}{n_2}} \binom{\rho_1 + \sigma_1}{\sigma_1} \binom{\rho_2 + \sigma_2}{\sigma_2} \cdot \partial_{\rho_1 + \sigma_1, \rho_2 + \sigma_2} f_{t_1, t_2}(\alpha, \alpha) (p_1/q_1 - \alpha)^{\rho_1} (p_2/q_2 - \alpha)^{\rho_2}$$

without changing the value of  $\Delta$ . This relies on the fact that the coefficient

$$\binom{\rho_1 + \sigma_1}{\sigma_1} \binom{\rho_2 + \sigma_2}{\sigma_2} (p_1/q_1 - \alpha)^{\rho_1} (p_2/q_2 - \alpha)^{\rho_2}$$

is independent of  $t_1$  and  $t_2$ .

We now estimate the entries (3). We have

$$\binom{\rho_1 + \sigma_1}{\sigma_1} \binom{\rho_2 + \sigma_2}{\sigma_2} |\partial_{\rho_1 + \sigma_1, \rho_2 + \sigma_2} f_{t_1, t_2}(\alpha, \alpha)| \leq 2^{n_1 + n_2} 2^{n_1 + n_2} |\alpha|^{n_1 + n_2},$$

where we estimated the product of the binomial coefficients, as well as the coefficient of the  $\partial_{\rho_1, \rho_2}$  derivative by  $2^{n_1 + n_2}$ . Next we are going to use

$$|p_i/q_i - \alpha| < q_i^{-\sqrt{2d} - \varepsilon}.$$

writing  $r = \min(q_1^{n_1}, q_2^{n_2})$  and  $R = \max(q_1^{n_1}, q_2^{n_2})$ , we have

$$|p_1/q_1 - \alpha|^{\rho_1} |p_2/q_2 - \alpha|^{\rho_2} \leq r^{-(\rho_1/n_1 + \rho_2/n_2)(\sqrt{2d} + \varepsilon)}.$$

Using that for all  $\rho_1, \rho_2$  that appear in the sum, we have

$$\frac{\rho_1}{n_1} + \frac{\rho_2}{n_2} > \frac{2}{\sqrt{2d}} - \delta_1 - \frac{\sigma_1}{n_1} - \frac{\sigma_2}{n_2}$$

and that  $\sigma_1/n_1 + \sigma_2/n_2 \leq \delta_2$  we have

$$(\rho_1/n_1 - \rho_2/n_1)(\sqrt{2d} + \varepsilon) > \left(\frac{2}{\sqrt{2d}} - \delta_1 - \delta_2\right)(\sqrt{2d} + \varepsilon) > 2 + \varepsilon/d$$

provided we choose  $\delta_1$  and  $\delta_2$  sufficiently small in terms of  $\varepsilon$ . Taking into account that there are at most  $(n_1 + 1)(n_2 + 1) \leq 2^{n_1+n_2}$  terms in (3), we get an upper bound of

$$|8\alpha|^{n_1+n_2} r^{-2-\varepsilon/d}$$

for these entries.

We write  $L_s$  for the number of columns corresponding to  $s$  in  $\Delta$ . Using a calculation similar to what was done in the lectures estimating the number of  $\sigma_1, \sigma_2$  with  $\sigma_1/n_1 + \sigma_2/n_2 \leq 2/\sqrt{2d} - \delta_1$  by considering the area of an appropriate triangle, we get that

$$L_s \leq \frac{(2/\sqrt{2d} - \delta_1/2)n_1(2/\sqrt{2d} - \delta_1/2)n_2}{2} \leq (1/d - \delta_1/d)L$$

for  $s \leq d$ , provided  $n_1$  and  $n_2$  are sufficiently large depending on  $\delta_1$ . Therefore,  $L_{d+1} \geq L - L_1 - \dots - L_d \geq \delta_1 L$ .

The entries of  $\Delta$  that are in columns corresponding to  $s \leq d$  are bounded by  $|2\alpha|^{n_1+n_2}$ . Expanding  $\Delta$  and using the bound we proved for the entries, we get

$$\begin{aligned} |\Delta| &\leq L! |2\alpha|^{(n_1+n_2)(L_1+\dots+L_d)} |8\alpha|^{(n_1+n_2)L_{d+1}} r^{-(2+\varepsilon/d)L_{d+1}} \\ &\leq L! |8\alpha|^{(n_1+n_2)L} r^{-(2+\varepsilon/d)L_{d+1}}. \end{aligned}$$

(d) We estimate the height of  $\Delta$  using its original form before we performed the column operations. Expanding that determinant, we have  $L!$  terms. Each term is a product of  $L$  entries of  $\Delta$ . Each term contribute a binomial coefficient that is an integer at most  $2^{n_1+n_2}$  and a power of  $\alpha$  or  $p_1/q_1$  and  $p_2/q_2$ . Therefore, it is possible to express  $\Delta$  as a polynomial expression of  $\alpha$ ,  $p_1/q_1$  and  $p_2/q_2$ . The polynomial has  $L!$  terms, each term has an integer coefficient at most  $2^{(n_1+n_2)L}$ , so the length of the polynomial is at most  $L!2^{(n_1+n_2)L}$ . The exponent of  $\alpha$  is at most  $(n_1 + n_2)L$ , and the exponent of  $p_j/q_j$  is at most  $n_j L_{d+1}$ . Therefore, we have

$$H(\Delta) \leq L! 2^{(n_1+n_2)L} H(\alpha)^{(n_1+n_2)L} H(p_1/q_1)^{n_1 L_{d+1}} H(p_2/q_2)^{n_2 L_{d+1}}$$

Using that  $H(p_j/q_j) = \max(|p_j|, |q_j|) \leq C|q_j|$  for some constant depending on  $\alpha$ , we have

$$H(\Delta) \leq L! C^{(n_1+n_2)L} R^{2L_{d+1}},$$

where  $C$  is another constant depending on  $\alpha$  and  $R$  was defined in Part (c).

Noting that  $\Delta \neq 0$  and hence

$$H(\Delta)^{[K:\mathbf{Q}]} = H(\Delta^{-1})^{[K:\mathbf{Q}]} \geq \prod_{v \in M_{K,\infty}} \max(|\Delta|_v^{-d_v}, 1),$$

it follows that there is an infinite place  $v$  such that

$$|\Delta|_v \geq H(\Delta)^{-1}.$$

So there is a Galois embedding  $\pi : K \rightarrow \mathbf{C}$  such that

$$|\pi(\Delta)| \geq H(\Delta)^{-1}.$$

We observe that  $\Delta$  is also a non-zero subdeterminant of  $M$  and the columns corresponding to  $s \leq d$  still span the space generated by the columns of  $M$  with  $s \leq d$ , so the upper bound we proved in Part (c) is also valid for  $|\pi(\Delta)|$ . There is no harm in assuming  $\pi(\Delta) = \Delta$ .

Comparing our upper and lower bounds we get

$$L!|\delta\alpha|^{(n_1+n_2)L} r^{-(2+\varepsilon/d)L_{d+1}} \geq L!^{-1} C^{-(n_1+n_2)L} R^{-2L_{d+1}}.$$

To get a contradiction, we observe the following. If  $n_1$  and  $n_2$  are sufficiently large depending on  $\varepsilon$ ,  $q_1$  and  $q_2$ , then  $n_2/n_1$  can be chosen sufficiently close to  $\log q_1 / \log q_2$  so that  $r = \min(q_1^{n_1}, q_2^{n_2})$  and  $R = \max(q_1^{n_1}, q_2^{n_2})$  are close enough so that

$$r^{(2+\varepsilon/d)L_{d+1}} R^{-2L_{d+1}} \geq R^{(\varepsilon/2d)L_{d+1}}.$$

We also have

$$L! \leq L^L = \exp((\log(n_1 + 1) + \log(n_2 + 1))L) \leq 2^{(n_1+n_2)L}.$$

Thus it is enough to show that

$$(32C|\alpha|)^{(n_1+n_2)L} < R^{(\varepsilon/2d)L_{d+1}}$$

and the contradiction will follow. To this end, we recall that  $L_{d+1} > \delta_1 L/d$ . If  $q_1$  and  $q_2$  are large enough in terms of  $C$ ,  $\alpha$ ,  $\varepsilon$  and  $d$ , then the required inequality holds.

**6.** The purpose of this question is to give a version of the argument in Question 6 for the Gelfond Schneider theorem using an auxiliary polynomial instead of an interpolation determinant. This time we do not need to assume that the logarithms are real.

Let  $\lambda_1, \lambda_2 \in \mathbf{C}_{\neq 0}$ . Suppose to the contrary that  $\alpha_1 = \exp(\lambda_1), \alpha_2 = \exp(\lambda_2), \beta = \lambda_2/\lambda_1 \in \overline{\mathbf{Q}}$  but  $\beta \notin \mathbf{Q}$ . We will derive a contradiction.

Let  $d = [\mathbf{Q}(\alpha_1, \alpha_2, \beta) : \mathbf{Q}]$  and we fix some positive integers  $T_0, T_1, S$ . In what follows,  $c$  and  $C$  are some constants that depend only on  $\alpha_1, \alpha_2$  and  $\beta$  and they may be a different one at each occurrence.

- (a) Under the assumption  $(2T + 1)^2 > 2d(S_0 + 1)(2S_1 + 1)$  find some  $a_{t_1, t_2} \in \mathbf{Z}$  not all 0 for  $t_1, t_2 = -T, \dots, T$  such that

$$\max_{t_1, t_2} |a_{t_1, t_2}| \leq \exp(CS_0 \log T + CS_1 T)$$

and the function

$$F(x) = \sum_{|t_1|, |t_2| \leq T} a_{t_1, t_2} \exp((t_1 + \beta t_2)x)$$

satisfies

$$(4) \quad \frac{d^\sigma}{dx^\sigma} F(\lambda_1 s) = 0$$

for  $\sigma = 0, \dots, S_0$  and  $s = -S_1, \dots, S_1$ .

(b) Fix a number  $E \geq 10$ . Let  $F$  be the function in part (a). Prove that

$$|F(x)| \leq \exp(CS_0 \log T + CTES_1)$$

for all complex  $|x| \leq ES_1|\lambda_1|$ .

(c) Suppose  $F$  satisfies (4) for  $\sigma = 0, \dots, S$  and  $s = -S_1, \dots, S_1$  with some  $S \geq S_0$ . Prove

$$|F(x)| \leq \exp(-cSS_1 \log E + CS_0 \log T + CTES_1)$$

for all complex  $|x| \leq 2S_1|\lambda_1|$ .

(d) Under the same assumptions prove that

$$\left| \frac{d^{S+1}}{dx^{S+1}} F(\lambda_1 s) \right| \leq \exp(-cSS_1 \log E + CS_0 \log T + CTES_1 + CS \log S)$$

for  $s = -S_1, \dots, S_1$ .

(e) Under the same assumptions prove that

$$\frac{d^{S+1}}{dx^{S+1}} F(\lambda_1 s) = 0$$

for  $s = -S_1, \dots, S_1$ .

(f) Conclude  $F = 0$ , a contradiction.

*Hint:* (a): Use Siegel's lemma. (c): Use the maximum modulus principle for the function

$$\frac{F(x)}{(x - S_1\lambda_1)^{S+1} \cdots (x + S_1\lambda_1)^{S+1}}.$$

(d): Use Cauchy's formula to express the derivative as a contour integral on the circle  $|x| = 2S_1|\lambda_1|$ . (e): Estimate the height of the number in question and compare the Liouville bound with the upper bound in the previous part. At this point you need to choose the parameters. Take  $S_1$  to be a large constant depending on  $\alpha_1, \alpha_2, \beta$ ,  $E = S^{1/10}$  and choose  $T$  and  $S_0$  to be sufficiently large satisfying the condition in part (a). (f) Run the arguments in parts (c)-(d) repeatedly to show that all derivatives of  $F$  vanish at 0 (and at all the other points we worked with).

Solution

(a) The equation  $d^\sigma/(dx^\sigma)F(\lambda_1 s) = 0$  can be written as

$$\begin{aligned} 0 &= \sum_{t_1, t_2} a_{t_1, t_2} (t_1 + \beta t_2)^\sigma \exp((t_1 + \beta t_2)\lambda_1 s) \\ &= \sum_{t_1, t_2} a_{t_1, t_2} (t_1 + \beta t_2)^\sigma \alpha_1^{t_1 s} \alpha_2^{t_2 s}. \end{aligned}$$

So we have  $d(S_0 + 1)(2S_1 + 1)$  linear constraints over  $\mathbf{Q}$  for  $(2T + 1)^2$  variables. To apply Siegel's lemma, we estimate the height of the equations. Let  $v$  be a place of  $\mathbf{Q}(\alpha_1, \alpha_2, \beta)$  which is either infinite or finite and at least one of  $|\alpha_1|_v$ ,  $|\alpha_2|_v$  and  $|\beta|_v$  is  $\geq 1$ . There are finitely many such places, and if  $v$  is another place than its contribution to the height of the equation is  $\leq 1$ . We estimate

$$\begin{aligned} &\max_{t_1, t_2, s, \sigma} |(t_1 + \beta t_2)^\sigma \alpha_1^{t_1 s} \alpha_2^{t_2 s}|_v \\ &\leq T^{S_0} (1 + |\beta|_v)^{S_0} \max(|\alpha_1|_v, |\alpha_1|_v^{-1})^{TS_1} \max(|\alpha_2|_v, |\alpha_2|_v^{-1})^{TS_1} \\ &\leq \exp(CS_0 \log T + CS_1 T), \end{aligned}$$

where  $C$  is a constant depending on  $\alpha_1, \alpha_2, \beta$ . Combining the contribution of all places, we get that the height of the equations is less than  $\exp(CS_0 \log T + CS_1 T)$  with another constant  $C$ .

Siegel's lemma implies that there is a solution for  $a_{t_1, t_2} \in \mathbf{Z}$ , not all 0, such that

$$|a_{t_1, t_2}| \leq ((2T + 1)^2 \exp(CS_0 \log T + CS_1 T))^{\frac{d(S_0+1)(2S_1+1)}{(2T+1)^2 - d(S_0+1)(2S_1+1)}}.$$

By the assumption  $(2T + 1)^2 > 2d(S_0 + 1)(2S_1 + 1)$ , the exponent is  $\leq 1$ . The factor  $(2T + 1)^2$  can be absorbed if we enlarge the constant  $C$ , so the claim follows.

(b) Suppose  $|x| \leq ES_1|\lambda_1|$ . Then

$$\exp((t_1 + \beta t_2)x) \leq \exp(|t_1 + \beta t_2|ES_1|\lambda_1|) \leq \exp(CTES_1)$$

for a suitable constant  $C$  depending on  $\beta$  and  $\lambda_1$ . Combining this with the bound in Part (a) for the coefficients, we get

$$\begin{aligned} |F(x)| &\leq \sum_{|t_1|, |t_2| \leq T} |a_{t_1, t_2} \exp((t_1 + \beta t_2)x)| \\ &\leq (2T + 1)^2 \exp(CS_0 \log T + CS_1 T + CTES_1) \\ &\leq \exp(C'S_0 \log T + C'TES_1), \end{aligned}$$

as required.

(c) Using the maximum principle, we can write

$$\left| \frac{F(x)}{\prod_{|s| \leq S_1} (x - \lambda_1 s)^{S+1}} \right|_{2S_1|\lambda_1|} \leq \left| \frac{F(x)}{\prod_{|s| \leq S_1} (x - \lambda_1 s)^{S+1}} \right|_{ES_1|\lambda_1|}.$$

For all  $s \in \{-S_1, \dots, S_1\}$ , we have

$$|x - \lambda_1 s| \leq 3S_1 |\lambda_1|$$

if  $|x| = 2S_1 |\lambda_1|$ , and

$$|x - \lambda_1 s| \geq (E - 1)S_1 |\lambda_1| \geq ES_1 |\lambda_1|/2$$

if  $|x| = ES_1 |\lambda_1|$ . Combining these estimates, we get

$$|F(x)|_{2S_1 |\lambda_1|} \leq \left(\frac{E}{6}\right)^{-(S+1)(2S_0+1)} |F(x)|_{ES_1 |\lambda_1|}.$$

We plug in the bound from Part (b), and the claim follows.

(d) We use the Cauchy integral formula

$$\frac{d^{S+1}}{dx^{S+1}} F(\lambda_1 s) = \frac{(S+1)!}{2\pi i} \int_{|x|=2S_1 |\lambda_1|} \frac{F(x)}{(x - \lambda_1 s)^{S+2}} dx.$$

Using that  $|x - \lambda_1 s| \geq S_1 |\lambda_1|$  on the contour of the integral, we get

$$\begin{aligned} \left| \frac{d^{S+1}}{dx^{S+1}} F(\lambda_1 s) \right| &\leq (S+1)! \cdot (2S_1 |\lambda_1|) \frac{|F(x)|_{2S_1 |\lambda_1|}}{(S_1 |\lambda_1|)^{S+2}} \\ &\leq \exp(-cSS_1 \log E + CS_0 \log T + CTES_1 + CS \log S) \end{aligned}$$

for a suitable constant  $C$  depending on  $\lambda_1, \lambda_2$ .

(e) Next we estimate the height of

$$\begin{aligned} \frac{d^{S+1}}{dx^{S+1}} F(\lambda_1 s) &= \sum_{t_1, t_2} a_{t_1, t_2} (t_1 + \beta t_2)^{S+1} \alpha_1^{t_1 s} \alpha_2^{t_2 s} \\ &= \sum_{t_1, t_2} \sum_{j=0}^{S+1} a_{t_1, t_2} \binom{S+1}{j} t_1^j t_2^{S+1-j} \beta^{S+1-j} \alpha_1^{t_1 s} \alpha_2^{t_2 s} \end{aligned}$$

This can be thought of as a polynomial of length

$$\begin{aligned} \sum_{t_1, t_2} \sum_j \binom{S+1}{j} |t_1|^j |t_2|^{S+1-j} |a_{t_1, t_2}| \\ \leq 2^{S+1} T^{S+1} (2T+1)^2 \exp(CS_0 \log T + CS_1 T) \end{aligned}$$

in the variables  $\beta, \alpha_1^{\pm 1}$  and  $\alpha_2^{\pm 1}$ . Therefore, the height in question is at most

$$\begin{aligned} 2^{S+1} T^{S+1} (2T+1)^2 \exp(CS_0 \log T + CS_1 T) H(\beta)^{S+1} H(\alpha_1)^{2TS_1} H(\alpha_2)^{2TS_1} \\ \leq \exp(CS_0 \log T + CS_1 T + CS \log T). \end{aligned}$$

If  $\frac{d^{S+1}}{dx^{S+1}} F(\lambda_1 s) \neq 0$ , then the Liouville bound will give

$$\left| \frac{d^{S+1}}{dx^{S+1}} F(\lambda_1 s) \right| > \exp(-CS_0 \log T - CS_1 T - CS \log T).$$

We contrast this with our previous upper bound, and we get contradiction, if we have

$$cSS_1 \log E > CS_0 \log T + CTES_1 + CS \log S + CS \log T.$$

We take  $E = S^{1/10}$ . Note that  $S \geq S_0$  and  $\log S \geq \log T$ . Among the terms,  $CS_0 \log T$ ,  $CS \log S$  and  $CS \log T$ ,  $CS \log S$  dominates. If  $S_1$  is a sufficiently large constant, then  $cSS_1 \log E = (c/10)S_1 S \log S$  will dominate all of these terms. In addition,  $S \geq S_0 > T^2/10S_1$ , so

$$S > S^{1/10} T^{18/10} / 10S_1 = ET \cdot T^{4/5} / 10S_1.$$

We can make  $T^{4/5}/10S_1$  as large as needed, and then  $cSS_1$  will dominate  $CTES_1$ . Therefore, we can choose the parameters to satisfy the requirements and reach the desired contradiction.

(f) Iterating (c)–(e), we can prove that (4) holds for  $\sigma = S_0 + 1, S_0 + 2, \dots$ . But  $F$  is an entire function, and all its derivatives vanish at 0, say, so  $F \equiv 0$ .