Το θεώρημα Thue–Siegel–Roth

Βασιλική Λαΐνα

Τμήμα Μαθηματικών
Πανεπιστήμιο Αθηνών
Αθήνα – 2010

Περιεχόμενα

1 Εισαγωγή **1**

1.1 Το θεώρημα του Liouville 1

1.2 Το θεώρημα Thue–Siegel–Roth 3

1.3 Λίγα λόγια για την ιδέα της απόδειξης 4

2 Το θεώρημα Thue–Siegel **7**

2.1 Περιγραφή της απόδειξης 7

2.2 Η απόδειξη 8

3 Δείκτης πολυωνύμου **15**

3.1 Ένα συνδυαστικό λήμμα 15

3.2 Το λήμμα του Siegel 17

3.3 Δείκτης πολυωνύμου 20

3.4 Ο δείκτης στο $(\alpha, \dots, \alpha)$ 23

3.5 Ο δείκτης σε ρητά σημεία κοντά στο $(\alpha, \dots, \alpha)$ 25

4 Γενικευμένες Wronskian και το Λήμμα του Roth **29**

4.1 Το λήμμα του Gauss 29

4.2 Γενικευμένες Wronskian 31

4.3 Το Λήμμα του Roth 33

5 Το θεώρημα του Roth **41**

5.1 Ανασκόπηση των προηγουμένων 41

5.2 Απόδειξη του θεωρήματος 43

Κεφάλαιο 1

Εισαγωγή

1.1 Το θεώρημα του Liouville

Το 1844, ο Liouville [5] απέδειξε το εξής θεώρημα για την προσέγγιση αλγεβρικών αριθμών από ρητούς.

Θεώρημα 1.1.1 (Liouville). *Έστω α πραγματικός αλγεβρικός αριθμός βαθμού d. Υπάρχει σταθερά $c(\alpha) > 0$ ώστε*

(1.1.1)
$$\left|\alpha - \frac{p}{q}\right| \geq \frac{c(\alpha)}{q^d}$$

για κάθε ρητό αριθμό $\frac{p}{q}$ διαφορετικό από τον α.

Απόδειξη. Αν $d = 1$ τότε ο $\alpha = \frac{m}{n}$ είναι ρητός και για κάθε $\frac{p}{q} \neq \frac{m}{n}$ έχουμε

(1.1.2)
$$\left|\frac{m}{n} - \frac{p}{q}\right| = \frac{|mq - np|}{nq} \geq \frac{c(\alpha)}{q},$$

όπου $c(\alpha) = 1/n$ (θεωρούμε την ανάγωγη μορφή m/n του α). Μπορούμε λοιπόν να υποθέσουμε ότι ο α έχει βαθμό $d > 1$, συνεπώς υπάρχει πολυώνυμο $T(x) = b_d x^d + \cdots + b_1 x + b_0$ βαθμού d, με σχετικά πρώτους ακέραιους συντελεστές και θετικό το συντελεστή του μεγιστοβάθμιου όρου, ώστε

(1.1.3)
$$T(\alpha) = 0.$$

Μπορούμε να αναπτύξουμε το T με κέντρο το α:

(1.1.4)
$$T(x) = \sum_{k=0}^{d} \frac{T^{(k)}(\alpha)}{k!}(x-\alpha)^k = \sum_{k=1}^{d} \frac{T^{(k)}(\alpha)}{k!}(x-\alpha)^k,$$

με την τελευταία ισότητα διότι $T^{(0)}(\alpha) = T(\alpha) = 0$. Συνεπώς, αν $\left|\frac{p}{q} - \alpha\right| \leq 1$ τότε

$$\begin{aligned} |T(p/q)| &= \left|\sum_{k=1}^{d} \frac{T^{(k)}(\alpha)}{k!}(p/q-\alpha)^k\right| \leq \sum_{k=1}^{d} \frac{|T^{(k)}(\alpha)|}{k!}|p/q-\alpha|^k \\ &\leq \sum_{k=1}^{d} \frac{|T^{(k)}(\alpha)|}{k!}|p/q-\alpha| = \frac{1}{c(\alpha)}\left|\frac{p}{q}-\alpha\right|, \end{aligned}$$

όπου

$$\frac{1}{c(\alpha)} = \sum_{k=1}^{d} \frac{|T^{(k)}(\alpha)|}{k!}. \tag{1.1.5}$$

Παρατηρούμε τώρα ότι $T(p/q) \neq 0$ αλλιώς ο βαθμός του α θα ήταν μικρότερος από d, άρα

$$|T(p/q)| = \frac{|b_d p^d + \cdots + b_1 p q^{d-1} + b_0 q^d|}{q^d} \geq \frac{1}{q^d}. \tag{1.1.6}$$

Έπεται ότι, αν $|p/q - \alpha| \leq 1$ τότε

$$\left|\frac{p}{q} - \alpha\right| \geq \frac{c(\alpha)}{q^d}. \tag{1.1.7}$$

Αντίστοιχη ανισότητα ισχύει προφανώς αν $|p/q-\alpha| > 1$, συνεπώς η απόδειξη είναι πλήρης. □

Ο Liouville χρησιμοποίησε αυτό το θεώρημα για να κατασκευάσει υπερβατικούς αριθμούς. Για παράδειγμα, ο

$$\alpha = \sum_{k=1}^{\infty} \frac{1}{2^{k!}} \tag{1.1.8}$$

είναι υπερβατικός. Πράγματι, αν θέσουμε $q_m = 2^{m!}$ και $p_m = 2^{m!}\sum_{k=1}^{m}\frac{1}{2^{k!}}$, τότε

$$\left|\alpha - \frac{p_m}{q_m}\right| = \sum_{k=m+1}^{\infty} \frac{1}{2^{k!}} < \frac{2}{2^{(m+1)!}} = \frac{2}{q_m^{m+1}}. \tag{1.1.9}$$

Αυτό σημαίνει ότι, για κάθε $c > 0$ και για κάθε $d > 1$, αν ο m είναι αρκετά μεγάλος έχουμε

$$\left|\alpha - \frac{p_m}{q_m}\right| < \frac{2}{q_m^{m+1}} < \frac{c}{q_m^d}. \tag{1.1.10}$$

Από το θεώρημα του Liouville συμπεραίνουμε ότι ο α δεν μπορεί να είναι αλγεβρικός (για κανένα βαθμό d). Συνεπώς, ο α είναι υπερβατικός.

1.2 Το θεώρημα Thue–Siegel–Roth

Άμεση συνέπεια του θεωρήματος του Liouville είναι ότι, αν α είναι αλγεβρικός αριθμός βαθμού $d \geq 2$, τότε, για κάθε $\delta > 0$, η ανισότητα

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{d+\delta}} \tag{1.2.1}$$

έχει πεπερασμένες το πλήθος λύσεις ως προς p, q. Το 1909, ο Thue [9] απέδειξε ότι, αν α είναι αλγεβρικός αριθμός βαθμού $d \geq 2$, τότε, για κάθε $\delta > 0$, η ανισότητα

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{\mu+\delta}} \tag{1.2.2}$$

έχει πεπερασμένες το πλήθος λύσεις, όπου $\mu = \frac{d}{2} + 1$. O Siegel ([7], 1921) απέδειξε το ίδιο με $\mu = 2\sqrt{d}$ και ο Dyson (Dyson, 1947) με $\mu = \sqrt{2d}$.

Το 1955, ο Roth [6] απέδειξε το εξής βέλτιστο αποτέλεσμα:

Θεώρημα 1.2.1 (θεώρημα του Roth). *Έστω α αλγεβρικός αριθμός βαθμού $d \geq 2$. Για κάθε $\delta > 0$, η ανισότητα*

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\delta}} \tag{1.2.3}$$

έχει πεπερασμένες το πλήθος λύσεις ως προς p, q.

Σύμφωνα με το θεώρημα του Dirichlet, για κάθε άρρητο αριθμό α, υπάρχουν άπειροι το πλήθος ρητοί p/q με την ιδιότητα

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^2}. \tag{1.2.4}$$

Συνεπώς, ο εκθέτης 2 στο θεώρημα του Roth είναι βέλτιστος. Για το αποτέλεσμα αυτό, ο Roth τιμήθηκε με το βραβείο Fields. Η απόδειξη που θα παρουσιάσουμε είναι σε γενικές γραμμές αυτή που δίνεται στα βιβλία των Cassels [2] και Schmidt [8].

Παρατήρηση 1.2.2. Παρατηρήστε ότι αρκεί να δείξουμε το θεώρημα για α ο οποίος είναι αλγεβρικός ακέραιος. Πράγματι, ας υποθέσουμε ότι για κάθε αλγεβρικό ακέραιο α και για κάθε $\delta > 0$, η $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\delta}}$ έχει πεπερασμένες το πλήθος λύσεις. Έστω ρ αλγεβρικός αριθμός και έστω $T(x) = b_d x^d + \cdots + b_1 x + b_0$ το ελάχιστο πολυώνυμο του ρ. Τότε, ο $\alpha = b_d \rho$ ικανοποιεί την

$$\alpha^d + b_{d-1}\alpha^{d-1} + \cdots + (b_1 b_d^{d-2})\alpha + (b_0 b_d^{d-1}) = 0 \tag{1.2.5}$$

και αυτό είναι το ελάχιστο πολυώνυμο του α, δηλαδή ο α είναι αλγεβρικός ακέραιος. Αν $\left|\rho - \frac{p}{q}\right| < \frac{1}{q^{2+\delta}}$, τότε

$$\left|\alpha - \frac{b_d p}{q}\right| < \frac{|b_d|}{q^{2+\delta}} < \frac{1}{q^{2+\frac{\delta}{2}}} \tag{1.2.6}$$

αν ο q είναι αρκετά μεγάλος. Συνεπώς, αν η $\left|\rho - \frac{p}{q}\right| < \frac{1}{q^{2+\delta}}$ έχει άπειρες λύσεις, τότε και η $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\frac{\delta}{2}}}$ έχει άπειρες λύσεις, το οποίο είναι άτοπο.

Υποθέτουμε λοιπόν, στη συνέχεια, ότι ο α είναι αλγεβρικός ακέραιος βαθμού $d \geq 2$ και γράφουμε

$$f(x) = x^d + b_{d-1}x^{d-1} + \cdots + b_1 x + b_0 \tag{1.2.7}$$

όπου $b_0, b_1, \ldots, b_{d-1} \in \mathbb{Z}$ για το ελάχιστο πολυώνυμο του α. Τέλος, θέτουμε

$$A := \max\{1, |b_{d-1}|, \ldots, |b_0|\}. \tag{1.2.8}$$

1.3 Λίγα λόγια για την ιδέα της απόδειξης

Όλοι όσοι συνεισέφεραν στο πρόβλημα (Thue, Siegel, Dyson και Roth) προσπάθησαν να βελτιώσουν το απλό επιχείρημα του Liouville. Ας υποθέσουμε ότι α είναι ένας αλγεβρικός αριθμός και $P(x)$ είναι ένα πολυώνυμο με ακέραιους συντελεστές, βαθμού r, το οποίο έχει ρίζα τάξης k τον α. Τότε, το ανάπτυγμα του $P(x)$ με κέντρο τον α παίρνει τη μορφή

$$P(x) = \sum_{j=k}^{r} \frac{P^{(j)}(\alpha)}{j!}(x-\alpha)^j. \tag{1.3.1}$$

Θεωρούμε $\mu > 0$ και διακρίνουμε δύο περιπτώσεις. Αν $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{\mu}}$ τότε

$$|P(p/q)| \leq \left|\alpha - \frac{p}{q}\right|^k \sum_{j=k}^{r} \frac{1}{j!}|P^{(j)}(\alpha)| = c(P)\left|\alpha - \frac{p}{q}\right|^k < \frac{c(P)}{q^{\mu k}}. \tag{1.3.2}$$

Από την άλλη πλευρά, αν εξαιρέσουμε πεπερασμένους το πλήθος ρητούς, έχουμε $P(p/q) \neq 0$, άρα

$$|P(p/q)| \geq \frac{1}{q^r}. \tag{1.3.3}$$

Έπεται ότι, αν η $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{\mu}}$ έχει άπειρες λύσεις, τότε πρέπει να ισχύει $\frac{1}{q^r} < \frac{c(P)}{q^{\mu k}}$ για οσοδήποτε μεγάλους $q \in \mathbb{N}$, δηλαδή $\mu \leq \frac{r}{k}$.

Το συμπέρασμα είναι ότι αν $\mu > \frac{r}{k}$ τότε η $\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^\mu}$ έχει πεπερασμένες το πλήθος λύσεις. Αυτό όμως δεν μπορεί να οδηγήσει σε βελτίωση του συμπεράσματος που προκύπτει από το θεώρημα του Liouville: αν d είναι ο βαθμός του α, τότε κάθε πολύωνυμο με ακέραιους συντελεστές που έχει τον α ως ρίζα τάξης k πρέπει να έχει βαθμό $r \geq kd$. Άρα, $\frac{r}{k} \geq d$.

Οι Thue και Siegel χρησιμοποίησαν πολυώνυμα δύο μεταβλητών. Στο Κεφάλαιο 2 θα παρουσιάσουμε συνοπτικά το επιχείρημα του Siegel. Όπως θα φανεί στα επόμενα Κεφάλαια, ο Roth «δανείστηκε» πολλές από τις ιδέες του Siegel.

Η νέα ιδέα του Roth ήταν να χρησιμοποιήσει πολυώνυμα πολλών μεταβλητών, με το πλήθος τους m να τείνει στο άπειρο. Ο ίδιος περιγράφει την στρατηγική του ως εξής: ας υποθέσουμε ότι οι ρητοί $\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}$ ικανοποιούν την ανισότητα

$$\left|\alpha - \frac{p_i}{q_i}\right| < \frac{1}{q_i^\mu}, \qquad i = 1, \ldots, m \tag{1.3.4}$$

για κάποιον $\mu > 0$. Θεωρούμε ένα πολυώνυμο $P(x_1, \ldots, x_m)$ με ακέραιους συντελεστές και βαθμό r_i ως προς τη μεταβλητή x_i. Αν

$$P\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right) \neq 0 \tag{1.3.5}$$

τότε

$$\left|P\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right)\right| \geq \frac{1}{q_1^{r_1} \cdots q_m^{r_m}}. \tag{1.3.6}$$

Θεωρούμε το ανάπτυγμα του $P(x_1, \ldots, x_m)$ με κέντρο το σημείο $(\alpha, \ldots, \alpha)$. Για κάποιους συντελεστές $C(j_1, \ldots, j_m) \in \mathbb{Q}(\alpha)$ έχουμε

$$P(x_1, \ldots, x_m) = \sum_{0 \leq j_i \leq r_i} C(j_1, \ldots, j_m)(x_1 - \alpha)^{j_1} \cdots (x_m - \alpha)^{j_m}. \tag{1.3.7}$$

Ακολουθώντας το επιχείρημα του Liouville, σκεφτόμαστε ότι η τιμή $|P(p_1/q_1, \ldots, p_m/q_m)|$ θα είναι μικρή αν (α) οι συντελεστές $C(j_1, \ldots, j_m)$ που αντιστοιχούν σε «μικρά» j_i μηδενίζονται, και (β) το άθροισμα των $|C(j_1, \ldots, j_m)|$ είναι «μικρό». Πράγματι, ας υποθέσουμε ότι υπάρχει πολυώνυμο $P(x_1, \ldots, x_m)$ με τις εξής δύο ιδιότητες:

(i) Για κάποιο «μικρό» $\delta > 0$ ισχύει

$$\sum_{0 \leq j_i \leq r_i} |C(j_1, \ldots, j_m)| \leq \left(q_1^{r_1} \cdots q_m^{r_m}\right)^\delta. \tag{1.3.8}$$

(ii) Για $\theta = 1/2$ ισχύει $C(j_1, \ldots, j_m) = 0$ αν

$$q_1^{j_1} \cdots q_m^{j_m} \leq \left(q_1^{r_1} \cdots q_m^{r_m}\right)^\theta. \tag{1.3.9}$$

Τότε, μπορούμε να γράψουμε

$$\begin{aligned}
|P(p_1/q_1,\dots,p_m/q_m)| &\le \sum_{0\le j_i\le r_i} |C(j_1,\dots,j_m)| \left|\alpha-\frac{p_1}{q_1}\right|^{j_1}\cdots\left|\alpha-\frac{p_m}{q_m}\right|^{j_m}\\
&\le \sum_{0\le j_i\le r_i} |C(j_1,\dots,j_m)| \frac{1}{\left(q_1^{j_1}\cdots q_m^{j_m}\right)^{\mu}}\\
&\le \left(\sum_{0\le j_i\le r_i} |C(j_1,\dots,j_m)|\right)\max_{0\le j_i\le r_i}\frac{1}{\left(q_1^{j_1}\cdots q_m^{j_m}\right)^{\mu}},
\end{aligned}$$

όπου το max παίρνεται πάνω από εκείνους τους $j_1,\dots,j_m$ για τους οποίους $C(j_1,\dots,j_m)\neq 0$. Παίρνοντας υπ' όψιν τις υποθέσεις μας για το P, έχουμε

$$\begin{aligned}
|P(p_1/q_1,\dots,p_m/q_m)| &\le \left(q_1^{j_1}\cdots q_m^{j_m}\right)^{\delta}\frac{1}{(q_1^{r_1}\cdots q_m^{r_m})^{\theta\mu}}\\
&= \left(q_1^{j_1}\cdots q_m^{j_m}\right)^{\delta-\theta\mu}.
\end{aligned}$$

Συνδυάζοντας με το κάτω φράγμα για το $|P(p_1/q_1,\dots,p_m/q_m)|$ συμπεραίνουμε ότι

$$\left(q_1^{j_1}\cdots q_m^{j_m}\right)^{1+\delta-\theta\mu}\ge 1, \tag{1.3.10}$$

δηλαδή

$$1+\delta-\theta\mu\ge 0. \tag{1.3.11}$$

Υποθέτοντας ότι $\mu>2$ και επιλέγοντας $\delta>0$ αρκετά μικρό θα καταλήξουμε σε άτοπο.

Το πρόβλημα είναι με ποιόν τρόπο μπορεί κανείς να κατασκευάσει ένα πολυώνυμο $P(x_1,\dots,x_m)$ με τις παραπάνω ιδιότητες. Εδώ θα χρησιμοποιηθεί πρώτα απ' όλα η υπόθεση ότι η $|\alpha-p/q|<\frac{1}{q^{\mu}}$ έχει άπειρες λύσεις. Θα χρειαστεί να θεωρήσουμε πολυώνυμο με πολλές μεταβλητές (m μεγάλο) ώστε το P να μην μηδενίζεται ταυτοτικά και να υπάρχει σημείο $(p_1/q_1,\dots,p_m/q_m)$ πολύ «κοντά» στο $(\alpha,\dots,\alpha)$ και με τους παρονομαστές $q_1,\dots,q_m$ πολύ «μεγάλους». Όπως θα φανεί κατά την απόδειξη, το πιο δύσκολο σημείο βρίσκεται στην απαίτηση να ισχύει $P(p_1/q_1,\dots,p_m/q_m)\neq 0$ που εξασφαλίζει το κάτω φράγμα για την $|P(p_1/q_1,\dots,p_m/q_m)|$. Οι συνθήκες (i) και (ii) για το P συμβιβάζονται αν και είναι αντιφατικές: στο Κεφάλαιο 3 θα εισάγουμε την έννοια του «δείκτη ενός πολυωνύμου $P(x_1,\dots,x_m)$ σε σημείο $(a_1,\dots,a_m)$». Η συνθήκη (ii) «αναγκάζει» τον δείκτη να είναι «μεγάλος» σε σημεία της μορφής $(p_1/q_1,\dots,p_m/q_m)$ ενώ η συνθήκη (i) τον «αναγκάζει» να είναι «μικρός».

Κεφάλαιο 2

Το θεώρημα Thue–Siegel

2.1 Περιγραφή της απόδειξης

Το θεώρημα Thue–Siegel ισχυρίζεται ότι αν α είναι ένας αλγεβρικός αριθμός βαθμού $d \geq 2$, τότε για κάθε $\varepsilon > 0$ υπάρχουν το πολύ πεπερασμένοι το πλήθος ρητοί αριθμοί $\frac{p_k}{q_k}$ που ικανοποιούν την

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_k^{2\sqrt{d}+\varepsilon}}. \tag{2.1.1}$$

Ο Dyson βελτίωσε αυτό τον εκθέτη σε $\sqrt{2d} + \varepsilon$ και, τελικά, ο Roth σε $2 + \varepsilon$.

Σε αυτή την παράγραφο δίνουμε μια σύντομη περιγραφή της μεθόδου του Siegel, ακολουθώντας τα άρθρα [3] του E. Croot και [1] των Bombieri, Hunt και van der Poorten. Υποθέτουμε ότι

$$\left|\alpha - \frac{p_1}{q_1}\right| < \frac{1}{q_1^{2\sqrt{d}+\varepsilon}} \quad \text{και} \quad \left|\alpha - \frac{p_2}{q_2}\right| < \frac{1}{q_2^{2\sqrt{d}+\varepsilon}} \tag{2.1.2}$$

για κάποιους ρητούς p_1/q_1 και p_2/q_2, όπου $q_1 < q_2$ και οι q_1, q_2 είναι «πολύ μεγάλοι». Ξεκινώντας με την υπόθεση ότι η (2.1.1) έχει άπειρες λύσεις, μπορούμε να πάρουμε τους q_1, q_2 οσοδήποτε μεγάλους χρειαστεί ώστε τελικά να καταλήξουμε σε άτοπο. Το άτοπο θα προκύψει από τα παρακάτω τρία βήματα:

Βήμα 1. Κατασκευάζουμε ένα πολυώνυμο $F(x, y)$ βαθμού N_1 ως προς x και N_2 ως προς y, με «μικρούς» ακέραιους συντελεστές, ώστε οι αρχικοί συντελεστές του αναπτύγματος Taylor του F με κέντρο το (α, α) να μηδενίζονται.

Βήμα 2. Η τιμή $F(p_1/q_1, p_2/q_2)$ είναι ρητός αριθμός με παρονομαστή το πολύ ίσο με $q_1^{N_1} q_2^{N_2}$. Αν λοιπόν, χρησιμοποιώντας το ανάπτυγμα Taylor του Βήματος 1 και χρησιμοποιώντας την (2.1.2), δείξουμε ότι αυτή είναι, κατ' απόλυτη τιμή, μικρότερη από τον $1/(q_1^{N_1} q_2^{N_2})$, τότε έχουμε $F(p_1/q_1, p_2/q_2) = 0$.

Βήμα 3. Δείχνουμε απευθείας (ή αντικαθιστώντας την F με κάποια μερική παράγωγό της) ότι $F(p_1/q_1, p_2, q_2) \neq 0$, οπότε συμπεραίνουμε ότι δεν υπάρχουν δύο λύσεις της (2.1.1) με παρονομαστές τόσο μεγάλους. Για το σκοπό αυτό, ο Thue και οι επόμενοι χρησιμοποιούσαν κατάλληλη γενίκευση (σε δύο μεταβλητές) του ισχυρισμού ότι αν ένα πολυώνυμο $g(x)$ με ακέραιους συντελεστές έχει ρητή ρίζα p/q τάξης m, τότε ο συντελεστής του μεγιστοβάθμιου όρου διαιρείται με q^m, άρα είναι «μεγάλος». Για την γενίκευση που χρειάζεται σε αυτό το σημείο, εισάγονται οι ορίζουσες Wronski $W(x,y)$ δύο μεταβλητών και σημαντικό ρόλο παίζει η ιδιότητά τους ότι παραγοντοποιούνται στη μορφή $W(x,y) = g(x)h(y)$.

2.2 Η απόδειξη

Επιλέγουμε N_1 και N_2 μεγαλύτερους από τον q_2, έτσι ώστε

$$\frac{1}{q_1^{N_1+1}} < \frac{1}{q_2^{N_2}} \leq \frac{1}{q_1^{N_1}}. \tag{2.2.1}$$

Αυτό γίνεται ως εξής: επιλέγουμε πρώτα κάποιον $N_2 > q_2$ και μετά παίρνουμε σαν N_1 τον μεγαλύτερο φυσικό για τον οποίον $q_1^{N_1} \leq q_2^{N_2}$.

<u>Βήμα 1</u>

Θεωρούμε ένα πολυώνυμο της μορφής

$$F(x,y) = \sum_{r=0}^{N_1}\sum_{s=0}^{N_2} c_{rs}x^r y^s, \tag{2.2.2}$$

όπου c_{rs} είναι ακέραιοι τους οποίους θα προσδιορίσουμε. Θεωρούμε τώρα το ανάπτυγμα Taylor

$$F(x,y) = \sum_{i=0}^{N_1}\sum_{j=0}^{N_1} d_{ij}(x-\alpha)^i(y-\alpha)^j \tag{2.2.3}$$

της F με κέντρο το (α,α), όπου

$$d_{ij} = \frac{1}{i!j!}\frac{\partial F}{\partial x^i \partial y^j}(\alpha,\alpha). \tag{2.2.4}$$

Κάθε d_{ij} είναι ακέραιος γραμμικός συνδυασμός δυνάμεων του α βαθμού το πολύ ίσου με $N_1 + N_2$. Οι συντελεστές σε αυτό τον γραμμικό συνδυασμό είναι μικρότεροι ή ίσοι από τον

$$B = 2^{N_1+N_2}\max_{r,s}|c_{rs}|. \tag{2.2.5}$$

Αυτό προκύπτει από την παρατήρηση ότι, για κάθε $0 \leq i \leq r \leq N_1$ και $0 \leq j \leq s \leq N_2$ έχουμε

$$\frac{1}{i!j!}\frac{\partial(x^r y^s)}{\partial x^i \partial y^j}(\alpha,\alpha) = \binom{r}{i}\binom{s}{j}\alpha^{r+s-(i+j)} \tag{2.2.6}$$

και

$$\binom{r}{i}\binom{s}{j} \leq 2^r 2^s \leq 2^{N_1+N_2}. \tag{2.2.7}$$

Θεωρούμε το ελάχιστο πολυώνυμο του α,

$$b_d x^d + b_{d-1}x^{d-1} + \cdots + b_1 x + b_0, \tag{2.2.8}$$

όπου $b_k \in \mathbb{Z}$ και ο μέγιστος κοινός διαιρέτης των $b_0, b_1, \ldots, b_d$ είναι 1. Με επαγωγή μπορούμε να δείξουμε ότι κάθε δύναμη α^i του α γράφεται στη μορφή

$$\alpha^i = b_{i,d-1}\alpha^{d-1} + \cdots + b_{i,1}\alpha + b_0, \tag{2.2.9}$$

όπου

$$b_{i,\ell} = \frac{n_{i\ell}}{b_d^i} \tag{2.2.10}$$

για κάποιον $n_{i\ell} \in \mathbb{Z}$, ο οποίος ικανοποιεί την

$$|n_{ij}| \leq C_1^{N_1+N_2} \tag{2.2.11}$$

για κάποια σταθερά $C_1 > 0$ που εξαρτάται μόνο από το ελάχιστο πολυώνυμο του α (δείτε το Λήμμα 3.2.3 στο επόμενο Κεφάλαιο).

Εισάγοντας αυτή την ανηγμένη μορφή των δυνάμεων του α στην αναπαράσταση των συντελεστών d_{ij}, βλέπουμε ότι

$$d_{ij} = \sum_{k=0}^{d-1} d_{ijk}\alpha^k, \tag{2.2.12}$$

όπου

$$d_{ijk} = \frac{n_{ijk}}{b_d^{N_1+N_2}}, \tag{2.2.13}$$

με τους n_{ijk} να είναι γραμμικοί συνδυασμοί των συντελεστών c_{rs} τους οποίους θέλουμε να προσδιορίσουμε, με συντελεστές που φράσσονται από $C_2^{N_1+N_2}$, για κάποια σταθερά $C_2 > 0$ που εξαρτάται μόνο από το ελάχιστο πολυώνυμο του α.

Θα θέλαμε τώρα να εξασφαλίσουμε ότι $d_{ij} = 0$ για κάθε $0 \leq i \leq tN_1$ και $0 \leq j \leq tN_2$, με το t όσο γίνεται μεγαλύτερο, πετυχαίνοντας ταυτόχρονα οι συντελεστές c_{rs} του F να είναι φραγμένοι απολύτως από $C_3^{N_1+N_2}$, όπου η σταθερά $C_3 = C_3(\alpha, \varepsilon) > 0$ να εξαρτάται μόνο από το ελάχιστο πολυώνυμο του α και τον ε.

Ας δούμε πρώτα πόσο μεγάλο μπορούμε να επιλέξουμε το t, αν δεν μας χρειάζεται το φράγμα για τους c_{rs}. Το πλήθος των συντελεστών c_{rs} που θέλουμε να προσδιορίσουμε είναι ίσο με $(N_1+1)(N_2+1)$. Για να έχουμε $d_{ij} = 0$ για κάποιους $0 \leq i \leq tN_1$ και $0 \leq j \leq tN_2$, πρέπει να ικανοποιείται η $d_{ijk} = 0$ για κάθε $k = 0, 1, \ldots, d-1$. Έτσι, έχουμε $dt^2N_1N_2$ ομογενείς εξισώσεις (από την (2.2.13)) ως προς c_{rs}. Αν

$$dt^2N_1N_2 < (N_1+1)(N_2+1), \tag{2.2.14}$$

τότε υπάρχει μη τετριμμένη λύση του συστήματος. Παίρνουμε λοιπόν την ικανή συνθήκη $t \ll 1/\sqrt{d}$.

Ελπίζουμε λοιπόν να εξασφαλίσουμε λύση $\{c_{rs}\}$ φραγμένη από $C_3^{N_1+N_2}$, κρατώντας το t να είναι της τάξης του $1/\sqrt{d}$. Το εργαλείο που μας δίνει αυτή τη δυνατότητα είναι το Λήμμα του Siegel (δείτε το Θεώρημα 3.2.2 στο επόμενο Κεφάλαιο). Χρησιμοποιώντας το βλέπουμε ότι:

Αν $0 < \delta \ll \varepsilon$ και $t = \frac{1}{\sqrt{d}} - \delta > 0$, τότε μπορούμε να βρούμε $C_4 = C_4(\alpha, \delta) > 0$ και $c_{rs} \in \mathbb{Z}$ με $|c_{rs}| \leq C_4^{N_1+N_2}$ ώστε, αν θεωρήσουμε το πολυώνυμο της (2.2.2) *και τους συντελεστές d_{ij} του αναπτύγματος Taylor* (2.2.3), *να ισχύει $d_{ij} = 0$ για κάθε $0 \leq i \leq tN_1$ και $0 \leq j \leq tN_2$.*

Βήμα 2

Από την (2.2.3) έχουμε

$$F(p_1/q_1, p_2/q_2) \leq \sum_{i=0}^{N_1}\sum_{j=0}^{N_1} |d_{ij}| \left|\frac{p_1}{q_1} - \alpha\right|^i \left|\frac{p_2}{q_2} - \alpha\right|^j. \tag{2.2.15}$$

Αν υποθέσουμε ότι οι $p_1/q_1, p_2/q_2$ ικανοποιούν τις (2.1.2) και (2.2.1), και ότι $d_{ij} = 0$ για κάθε $0 \leq i \leq tN_1$ και $0 \leq j \leq tN_2$ και $|c_{rs}| \leq C_4^{N_1+N_2}$, όπου $t = \frac{1}{\sqrt{d}} - \delta$ και $0 < \delta \ll c(\varepsilon, d)$, τότε

$$\begin{aligned}
F\left(\frac{p_1}{q_1}, \frac{p_2}{q_2}\right) &\leq N_1N_2C_5(\alpha,\delta)^{N_1+N_2} \max\left\{\frac{1}{q_1^{N_1(1/\sqrt{d}-\delta)(2\sqrt{d}+\varepsilon)}}, \frac{1}{q_2^{N_2(1/\sqrt{d}-\delta)(2\sqrt{d}+\varepsilon)}}\right\}\\
&< \max\left\{\frac{1}{q_1^{2N_1+1}}, \frac{1}{q_2^{2N_2}}\right\}\\
&\leq \frac{1}{q_1^{N_1}q_2^{N_2}},
\end{aligned}$$

όπου η τελευταία ανισότητα εξασφαλίζεται από την (2.2.1) και η προηγούμενη αν τα q_1, q_2 είναι αρκετά μεγάλα (θα εξαρτώνται από τα ε, δ και C_5). Αφού

$$F(p_1/q_1, p_2/q_2) = \frac{m}{q_1^{N_1} q_2^{N_2}} \tag{2.2.16}$$

για κάποιον $m \in \mathbb{Z}$, αναγκαστικά έχουμε $m = 0$ και

$$F(p_1/q_1, p_2/q_2) = 0. \tag{2.2.17}$$

Βήμα 3

Ας υποθέσουμε ότι $F(p_1/q_1, p_2/q_2) = 0$. Αναπτύσσουμε την F με κέντρο το $(p_1/q_1, p_2/q_2)$: είναι

$$F(x, y) = \sum_{r=0}^{N_1} \sum_{s=0}^{N_2} e_{rs} \left(x - \frac{p_1}{q_1}\right)^r \left(y - \frac{p_2}{q_2}\right)^s, \tag{2.2.18}$$

και $e_{00} = 0$. Ο ισχυρισμός είναι ότι πρέπει να υπάρχει κάποιος $e_{rs} \neq 0$ με τους $r \leq \varepsilon_1 N_1$ και $s \leq \varepsilon_1 N_2$, όπου το $\varepsilon_1 > 0$ μπορεί να επιλεγεί οσοδήποτε μικρό (σε σύγκριση με το ε) αν οι N_1, N_2 είναι αρκετά μεγάλοι. Αν αυτό ισχύει, τότε η

$$F_1(x, y) = \frac{1}{r!s!} \frac{\partial^{r+s} F(x, y)}{\partial x^r \partial y^s} \tag{2.2.19}$$

θα είναι διαφορετική από το μηδέν στο σημείο $(p_1/q_1, p_2/q_2)$ και ταυτόχρονα θα ικανοποιούνται οι προϋποθέσεις ώστε να εφαρμόσουμε το Βήμα 2 γι' αυτήν. Παρατηρήστε ότι οι συντελεστές της F_1 φράσσονται κι αυτοί από μια ποσότητα της μορφής $C^{N_1+N_2}$, όπου η σταθερά C εξαρτάται από τον α, το δ και το $\max|c_{rs}|$.

Η στρατηγική. Είναι χρήσιμο να δει κανείς πρώτα γιατί αυτός ο ισχυρισμός μοιάζει λογικός, ξεκινώντας από την περίπτωση της μιας μεταβλητής: έστω

$$f(x) = \sum_{r=0}^{N_1} e_r (x - p_1/q_1)^r \tag{2.2.20}$$

ένα πολυώνυμο με ακέραιους συντελεστές. Αν $e_r = 0$ για κάθε $r = 0, 1, \ldots, \varepsilon_1 N_1$, τότε το πολυώνυμο $(x - p_1/q_1)^{\lfloor \varepsilon_1 N_1 \rfloor + 1}$ διαιρεί το $f(x)$. Από το Λήμμα του Gauss (βλέπε Παράγραφο 4.1) έπεται ότι το πολυώνυμο $(q_1 x - p_1)^{\lfloor \varepsilon_1 N_1 \rfloor + 1}$ διαιρεί το $f(x)$. Ειδικότερα, ο $q_1^{\lfloor \varepsilon_1 N_1 \rfloor + 1}$ διαιρεί το μεγιστοβάθμιο συντελεστή του $f(x)$. Συνεπώς, οι συντελεστές του $f(x)$ πρέπει να είναι μεγάλοι. Αν όμως γνωρίζουμε ότι αυτοί οι συντελεστές φράσσονται από C^{N_1}, τότε ο ε_1 πρέπει να είναι μικρός αν οι q_1, N_1 είναι μεγάλοι.

Όταν το πολυώνυμο $F(x,y)$ είναι δύο μεταβλητών, το προηγούμενο επιχείρημα δεν δουλεύει. Υπάρχει όμως μία περίπτωση στην οποία δεν χρειάζεται καμία μετατροπή: ας υποθέσουμε ότι

$$F(x,y) = g(x)h(y) \tag{2.2.21}$$

για κάποια πολυώνυμα g, h μιας μεταβλητής, με ακέραιους συντελεστές. Αν το $F(x,y)$ δίνεται από την (2.2.18) και αν $e_{rs} = 0$ για όλους τους $r \leq \varepsilon_1 N_1$ και $s \leq \varepsilon_1 N_2$, τότε, γράφοντας

$$g(x) = \sum_{r=0}^{N_1} g_r(x - p_1/q_1)^r \;\text{ και }\; h(y) = \sum_{s=0}^{N_2} h_s(y - p_2/q_2)^s, \tag{2.2.22}$$

όπου $g_r, h_s \in \mathbb{Q}$, έχουμε: είτε $g_r = 0$ για κάθε $0 \leq r \leq \varepsilon_1 N_1$ ή $h_s = 0$ για κάθε $0 \leq s \leq \varepsilon_1 N_2$. Όπως και στην προηγούμενη παράγραφο, στην πρώτη περίπτωση καταλήγουμε στην $q_1^{\varepsilon_1 N_1} \leq C^{N_1+N_2}$ ενώ στην δεύτερη καταλήγουμε στην $q_2^{\varepsilon_1 N_2} \leq C^{N_1+N_2}$ (υποθέτοντας ότι οι συντελεστές του F φράσσονται από $C^{N_1+N_2}$). Αν υποθέσουμε ότι $N_1 \log q_1 \simeq N_2 \log q_2$, τότε το φράγμα που παίρνουμε για το ε_1 είναι $O(1/\log q_1)$.

Γενικά, ένα πολυώνυμο $F(x,y)$ δύο μεταβλητών δεν παραγοντοποιείται στη μορφή $g(x)h(y)$. Η ιδέα των Thue και Siegel ήταν να κατασκευάσουν ένα νέο πολυώνυμο $W(x,y)$ το οποίο έχει την ιδιότητα να παραγοντοποιείται στη μορφή $g(x)h(y)$, οι συντελεστές του φράσσονται από $C_1^{N_1+N_2}$ για μια σταθερά $C_1 > 0$ που εξαρτάται από τους α, ε, και ικανοποιεί το εξής: αν όλοι οι συντελεστές e_{rs} του F «μικρής τάξης» μηδενίζονται, τότε το ίδιο ισχύει για τους συντελεστές του αναπτύγματος του $W(x,y)$ με κέντρο το $(p_1/q_1, p_2/q_2)$. Μετά, χρησιμοποιεί κανείς το επιχείρημα της προηγούμενης παραγράφου για το πολυώνυμο W.

Ορισμός του $W(x,y)$. Γνωρίζουμε ότι υπάρχουν πολυώνυμα $a_i(x), b_i(y)$ με ρητούς συντελεστές, ώστε

$$F(x,y) = \sum_{i=1}^{k} a_i(x) b_i(y). \tag{2.2.23}$$

Για παράδειγμα, το ίδιο το ανάπτυγμα Taylor του $F(x,y)$ με κέντρο το $(0,0)$ μας δίνει μια τέτοια αναπαράσταση. Επιλέγουμε μια αναπαράσταση της μορφής (2.2.23) για την οποία το πλήθος k των προσθετέων είναι το μικρότερο δυνατό. Τότε, οι $a_1, \ldots, a_k$ είναι γραμμικά ανεξάρτητες πάνω από το $\mathbb{Q}$, και το ίδιο ισχύει για τις $b_1, \ldots, b_k$ (βλέπε (4.3.13)–(4.3.15)). Επίσης, $k \leq N_2$. Από την ανεξαρτησία των a_i και των b_i, $i = 1, \ldots, k$, προκύπτει ότι

$$g(x) := \det \left(\frac{a_i^{j-1}(x)}{(j-1)!} \right)_{i,j=1}^{k} \neq 0 \;\text{ και }\; h(y) := \det \left(\frac{b_i^{j-1}(x)}{(j-1)!} \right)_{i,j=1}^{k} \neq 0. \tag{2.2.24}$$

Δηλαδή, το πολυώνυμο

$$W(x,y) = \det \left(\frac{1}{(i-1)!(j-1)!} \frac{\partial^{i+j-2} F(x,y)}{\partial x^{i-1} \partial y^{j-1}} \right)_{i,j=1}^{k} \tag{2.2.25}$$

γράφεται στη μορφή

$$W(x,y)=g(x)h(y). \tag{2.2.26}$$

Μπορούμε επίσης να δώσουμε άνω φράγμα για τους συντελεστές του $W(x,y)$. Όλοι οι όροι που εμφανίζονται στην ορίζουσα (2.2.25) φράσσονται από $C(\alpha,\varepsilon,\delta)^{N_1+N_2}$. έπεται ότι οι συντελεστές του W φράσσονται από

$$k!c^{k(N_1+N_2)}\le N_2!c^{N_2(N_1+N_2)}\le C^{N_1N_2}, \tag{2.2.27}$$

αν οι N_1,N_2 υποτεθούν αρκετά μεγάλοι.

Αναπτύσσουμε το $W(x,y)$ με κέντρο το $(p_1/q_1,p_2/q_2)$ και θα θέλαμε να χρησιμοποιήσουμε το γεγονός ότι οι συντελεστές μικρής τάξης $(r,s)\le(\varepsilon_1N_1,\varepsilon_1N_2)$ σε αυτό το ανάπτυγμα

$$W(x,y)=\sum_{r=0}^{N_1}\sum_{s=0}^{N_2}e_{rs}(x-p_1/q_1)^r(y-p_2/q_2)^s \tag{2.2.28}$$

μηδενίζονται για να δείξουμε ότι κάθε μη μηδενικός όρος αυτού του αναπτύγματος είτε περιέχει πολλαπλάσιο μεγάλης δύναμης του $(x-p_1/q_1)$ ή περιέχει πολλαπλάσιο μεγάλης δύναμης του $(y-p_2/q_2)$. Μετά, θα μπορούσαμε να προχωρήσουμε όπως στην περίπτωση πολυωνύμου μιας μεταβλητής.

Σε αυτό το σημείο χρειάζεται να υποθέσουμε ότι ο q_2 είναι πολύ μεγαλύτερος από τον q_1 και ο N_1 πολύ μεγαλύτερος από τον N_2. Αυτό σημαίνει ότι $k\ll N_1$.

Ορίζουμε τώρα τον *δείκτη* $\mathrm{Ind}(P)$ ενός πολυωνύμου $P(x,y)$ (βαθμού (N_1,N_2) ως προς (x,y)) στο σημείο $(p_1/q_1,p_2/q_2)$ ως τη μέγιστη τιμή της ποσότητας $\frac{a}{N_1}+\frac{b}{N_2}$ πάνω από όλα τα ζευγάρια (a,b) για τα οποία εμφανίζεται όρος $(x-p_1/q_1)^a(y-p_2/q_2)^b$ στο ανάπτυγμα του P με κέντρο το $(p_1/q_1,p_2/q_2)$. Χρησιμοποιώντας τις βασικές ιδιότητες

$$\mathrm{Ind}(PQ)=\mathrm{Ind}(P)+\mathrm{Ind}(Q) \tag{2.2.29}$$

και

$$\mathrm{Ind}(P+Q)\ge\min\{\mathrm{Ind}(P),\mathrm{Ind}(Q)\} \tag{2.2.30}$$

του δείκτη, οι οποίες προκύπτουν από τον ορισμό, μπορούμε να δείξουμε ότι όλες οι συντεταγμένες στην j-στήλη του πίνακα (2.2.25) που ορίζει το $W(x,y)$ έχουν δείκτη μεγαλύτερο ή ίσο από

$$\varepsilon_1-\frac{j-1}{N_2}. \tag{2.2.31}$$

Τότε, πάλι χρησιμοποιώντας τις βασικές ιδιότητες του δείκτη, συμπεραίνουμε ότι

$$\mathrm{Ind}(W)\ge(\varepsilon_1-O(\varepsilon_1^2))N_2. \tag{2.2.32}$$

Παίρνοντας υπ' όψιν το φράγμα (2.2.27) για τους συντελεστές του $W(x,y)$ και το γεγονός ότι παραγοντοποιείται στη μορφή $W(x,y)=g(x)h(y)$, βλέπουμε ότι είτε $\mathrm{Ind}(g)\geq(\varepsilon_1/2-O(\varepsilon_1^2))N_2$ ή $\mathrm{Ind}(h)\geq(\varepsilon_1/2-O(\varepsilon_1^2))N_2$.

Στην πρώτη περίπτωση έχουμε

$$(q_1x-p_1)^{(\varepsilon_1/2+O(\varepsilon_1^2))N_1N_2}\mid g(x)\mid W(x,y), \tag{2.2.33}$$

το οποίο σημαίνει ότι

$$q_1^{(\varepsilon_1/2+O(\varepsilon_1^2))N_1N_2}\leq C^{N_1N_2}, \tag{2.2.34}$$

άρα

$$\varepsilon_1\leq\frac{C_1}{\log q_1} \tag{2.2.35}$$

αν ο q_1 είναι μεγάλος.

Στη δεύτερη περίπτωση έχουμε

$$(q_2y-p_2)^{(\varepsilon_1/2+O(\varepsilon_1^2))N_2^2}\mid W(x,y), \tag{2.2.36}$$

το οποίο σημαίνει ότι

$$(q_1^{N_1})^{(\varepsilon_1/2+O(\varepsilon_1^2))N_2}<q_2^{(\varepsilon_1/2+O(\varepsilon_1^2))N_2^2}\leq C^{N_1N_2}, \tag{2.2.37}$$

άρα, πάλι,

$$\varepsilon_1\leq\frac{C_1}{\log q_1}. \tag{2.2.38}$$

Με άλλα λόγια, αν οι q_1,q_2 είναι πολύ μεγάλοι και ο N_1 είναι πολύ μεγαλύτερος από τον N_2, τότε ο ε_1 πρέπει να είναι πολύ μικρός.

Αυτό με τη σειρά του σημαίνει ότι κάποια μερική παράγωγος μικρής τάξης $F_1(x,y)=\frac{\partial^{r+s}F(x,y)}{\partial x^r\partial y^s}$ του πολυωνύμου $F(x,y)$ που ορίστηκε στο Βήμα 1 δεν μηδενίζεται στο $(p_1/q_1,p_2/q_2)$. Την ίδια στιγμή, το Βήμα 2 μπορεί να επαναληφθεί για την $F_1(x,y)$ και αυτό οδηγεί σε άτοπο.

Κεφάλαιο 3

Δείκτης πολυωνύμου

3.1 Ένα συνδυαστικό λήμμα

Αποδεικνύουμε πρώτα ένα συνδυαστικό λήμμα που θα χρησιμοποιηθεί στη συνέχεια.

Λήμμα 3.1.1. *Έστω $r_1, \ldots, r_m \in \mathbb{N}$ και $0 < \varepsilon < 1$. Το πλήθος των m-άδων $(i_1, \ldots, i_m)$ μη αρνητικών ακεραίων που ικανοποιούν τις*

$$0 \leq i_k \leq r_k, \qquad k = 1, \ldots, m \tag{3.1.1}$$

και

$$\left| \sum_{k=1}^{m} \frac{i_k}{r_k} - \frac{m}{2} \right| \geq \varepsilon m \tag{3.1.2}$$

είναι μικρότερο ή ίσο από

$$2(r_1 + 1) \cdots (r_m + 1) e^{-\varepsilon^2 m/2}. \tag{3.1.3}$$

Απόδειξη. Η απόδειξη που θα δώσουμε είναι πιθανοθεωρητική. Γράφουμε M_1 για το πλήθος των m-άδων $(i_1, \ldots, i_m)$ για τις οποίες

$$\sum_{k=1}^{m} \frac{i_k}{r_k} - \frac{m}{2} \geq \varepsilon m \tag{3.1.4}$$

και M_2 για το πλήθος των m-άδων $(i_1, \ldots, i_m)$ για τις οποίες

$$\sum_{k=1}^{m} \frac{i_k}{r_k} - \frac{m}{2} \leq -\varepsilon m. \tag{3.1.5}$$

Θα δείξουμε ότι

$$\max\{M_1, M_2\} \leq (r_1+1)\cdots(r_m+1)e^{-\varepsilon^2 m/2}. \tag{3.1.6}$$

Θεωρούμε ανεξάρτητες τυχαίες μεταβλητές $X_1, \dots, X_m$ έτσι ώστε η X_k να παίρνει τις τιμές $0, 1, \dots, r_k$ με πιθανότητα $\frac{1}{r_k+1}$. Παρατηρήστε ότι

$$\mathbb{E}\,(X_k) = \sum_{j=0}^{r_k} \frac{j}{r_k+1} = \frac{r_k(r_k+1)}{2(r_k+1)} = \frac{r_k}{2}. \tag{3.1.7}$$

Συνεπώς,

$$\mathbb{E}\left(\sum_{k=1}^{m} \frac{X_k}{r_k}\right) = \frac{m}{2}. \tag{3.1.8}$$

Μπορούμε λοιπόν να γράψουμε

$$M_1 = (r_1+1)\cdots(r_m+1)\,\mathbb{P}\left(\sum_{k=1}^{m} \frac{X_k}{r_k} - \mathbb{E}\left(\sum_{k=1}^{m} \frac{X_k}{r_k}\right) \geq \varepsilon m\right). \tag{3.1.9}$$

Ορίζουμε

$$Y_k = \frac{X_k}{r_k} - \mathbb{E}\left(\frac{X_k}{r_k}\right). \tag{3.1.10}$$

Τότε, οι $Y_1, \dots, Y_m$ είναι ανεξάρτητες τυχαίες μεταβλητές και $\mathbb{E}\,(Y_k) = 0,\ k = 1, \dots, m$. Για κάθε $t > 0$ μπορούμε να γράψουμε

$$\begin{aligned} P &:= \mathbb{P}\left(\sum_{k=1}^{m} Y_k \geq \varepsilon m\right) = \mathbb{P}\left(e^{t\sum_{k=1}^{m} Y_k} \geq e^{t\varepsilon m}\right) \\ &\leq e^{-t\varepsilon m}\mathbb{E}\left(e^{t\sum_{k=1}^{m} Y_k}\right) \end{aligned}$$

χρησιμοποιώντας πρώτα το γεγονός ότι η εκθετική συνάρτηση είναι αύξουσα και μετά την ανισότητα του Markov. Από την ανεξαρτησία των Y_k έχουμε

$$\mathbb{E}\left(e^{t\sum_{k=1}^{m} Y_k}\right) = \mathbb{E}\left(\prod_{k=1}^{m} e^{tY_k}\right) = \prod_{k=1}^{m} \mathbb{E}\,(e^{tY_k}). \tag{3.1.11}$$

Για τον υπολογισμό της $\mathbb{E}\,(e^{tY_k})$ υπολογίζουμε πρώτα την

(3.1.12)

$$\mathbb{E}\,(Y_k^2) = \sum_{j=0}^{r_k} \frac{j^2}{r_k^2(r_k+1)} - [\mathbb{E}(X_k/r_k)]^2 = \frac{r_k(r_k+1)(2r_k+1)}{6r_k^2(r_k+1)} - \frac{1}{4} = \frac{2r_k+1}{6r_k} - \frac{1}{4} \leq \frac{1}{4},$$

και μετά, χρησιμοποιώντας την $|Y_k| \leq 1$ και το ανάπτυγμα Taylor της εκθετικής συνάρτηση-ς, γράφουμε

$$\begin{aligned} \mathbb{E}\left(e^{tY_k}\right) &= 1 + \sum_{s=2}^{\infty} \frac{t^s \mathbb{E}\left(Y_k^s\right)}{s!} \\ &\leq 1 + \frac{t^2}{2}\mathbb{E}\left(Y_k^2\right) + \sum_{s=3}^{\infty} \frac{t^s \mathbb{E}\left(Y_k^2\right)}{s!} \\ &\leq 1 + \frac{t^2}{2}\frac{1}{4}\left(1 + \sum_{s=3}^{\infty} \frac{2}{s!} t^{s-2}\right) \\ &\leq 1 + \frac{t^2}{2}\frac{1}{4}\left(1 + \sum_{s=3}^{\infty} \frac{1}{(s-2)!} t^{s-2}\right) \\ &\leq 1 + \frac{t^2}{8} e^t. \end{aligned}$$

Έπεται ότι

$$P \leq e^{-t\varepsilon m}\left(1 + \frac{t^2}{8}e^t\right)^m \leq e^{-t\varepsilon m} e^{\frac{t^2 m}{8} e^t} \tag{3.1.13}$$

για κάθε $t > 0$. Επιλέγουμε $t = \varepsilon$. Τότε,

$$-t\varepsilon m + \frac{t^2 m}{8} e^t = -\varepsilon^2 m + \frac{\varepsilon^2 m}{8} e^{\varepsilon} < -\frac{\varepsilon^2 m}{2}. \tag{3.1.14}$$

Τότε,

$$M_1 \leq (r_1 + 1) \cdots (r_m + 1) e^{-\varepsilon^2 m/2}, \tag{3.1.15}$$

και όμοια βλέπουμε ότι

$$M_2 \leq (r_1 + 1) \cdots (r_m + 1) e^{-\varepsilon^2 m/2}, \tag{3.1.16}$$

απ' όπου προκύπτει το συμπέρασμα. □

3.2 Το λήμμα του Siegel

Έστω T ένα πολυώνυμο m μεταβλητών, με ακέραιους συντελεστές. Μπορούμε να γράφουμε

$$T(x_1, \ldots, x_m) = \sum_{j_1, \ldots, j_m \geq 0} C(j_1, \ldots, j_m) x_1^{j_1} \cdots x_m^{j_m}, \tag{3.2.1}$$

όπου το άθροισμα είναι πάνω από όλες τις m-άδες $j_1, \ldots, j_m$ μη αρνητικών ακεραίων αλλά μόνο πεπερασμένοι το πλήθος από τους συντελεστές $C(j_1, \ldots, j_m)$ είναι μη μηδενικοί.

Το **ύψος** $h(T)$ του T ορίζεται από την

$$h(T) = \max\{|C(j_1, \ldots, j_m)| : j_1, \ldots, j_m \in \mathbb{Z}^+\}. \tag{3.2.2}$$

Αν $\vec{i} = (i_1, \ldots, i_m)$ είναι μια m-άδα μη αρνητικών ακεραίων, γράφουμε $T_{\vec{i}}$ για το πολυώνυμο

$$T_{\vec{i}} = T_{i_1,\ldots,i_m} = \frac{1}{i_1! \cdots i_m!} \frac{\partial^{i_1+\cdots+i_m}}{\partial x_1^{i_1} \cdots \partial x_m^{i_m}} T. \tag{3.2.3}$$

Λήμμα 3.2.1. *Έστω T ένα πολυώνυμο m μεταβλητών, με ακέραιους συντελεστές. Αν $\vec{i} = (i_1, \ldots, i_m)$ είναι μια m-άδα μη αρνητικών ακεραίων, τότε το $T_{\vec{i}}$ έχει κι αυτό ακέραιους συντελεστές. Επιπλέον, αν ο βαθμός του T ως προς x_k είναι ίσος με r_k $(k = 1, \ldots, m)$ τότε*

$$h(T_{\vec{i}}) \leq 2^{r_1+\cdots+r_m} h(T). \tag{3.2.4}$$

Απόδειξη. Γράφουμε το T στη μορφή

$$T(x_1, \ldots, x_m) = \sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} C(j_1, \ldots, j_m) x_1^{j_1} \cdots x_m^{j_m}, \tag{3.2.5}$$

όπου $C(j_1, \ldots, j_m) \in \mathbb{Z}$. Τότε,

$$T_{\vec{i}}(x_1, \ldots, x_m) = \sum_{j_1=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} \binom{j_1}{i_1} \cdots \binom{j_m}{i_m} x_1^{j_1-i_1} \cdots x_m^{j_m-i_m}. \tag{3.2.6}$$

Σε αυτή την αναπαράσταση συμφωνούμε ότι $\binom{j_k}{i_k} = 0$ αν $j_k < i_k$. Οι συντελεστές του $T_{\vec{j}}$ είναι ακέραιοι, διότι $\binom{j_k}{i_k} \in \mathbb{Z}$ για κάθε $k = 1, \ldots, m$.

Τέλος, από την ανισότητα

$$\binom{j_k}{i_k} \leq 2^{j_k} \leq 2^{r_k} \tag{3.2.7}$$

είναι φανερό ότι οι συντελεστές του $T_{\vec{j}}$ φράσσονται απολύτως από

$$2^{r_1} \cdots 2^{r_m} \max |C(j_1, \ldots, j_m)|,$$

δηλαδή $h(T_{\vec{i}}) \leq 2^{r_1+\cdots+r_m} h(T)$. □

Θεώρημα 3.2.2 (λήμμα του Siegel). *Θεωρούμε m γραμμικές μορφές*

$$L_j(z_1, \ldots, z_n) = \sum_{k=1}^{n} a_{jk} z_k \tag{3.2.8}$$

με ακέραιους συντελεστές. Αν $n > m$ και $|a_{jk}| \leq A$ για κάθε $1 \leq j \leq m$ και $1 \leq k \leq n$, όπου A φυσικός αριθμός, τότε υπάρχουν ακέραιοι $y_1, \ldots, y_n$, όχι όλοι μηδέν, ώστε

$$L_j(y_1, \ldots, y_n) = 0, \qquad j = 1, \ldots, m \tag{3.2.9}$$

και

$$\max_{1 \leq k \leq n} |y_k| \leq (nA)^{\frac{m}{n-m}}. \tag{3.2.10}$$

Απόδειξη. Αφού $n > m$, υπάρχουν μη μηδενικές ρητές λύσεις του συστήματος (3.2.9). Παρατηρώντας ότι αν $(z_1, \ldots, z_n)$ είναι μια λύση του (3.2.9) τότε και η $(tz_1, \ldots, tz_n)$ είναι λύση του, συμπεραίνουμε ότι υπάρχουν μη τετριμμένες ακέραιες λύσεις.

Θέτουμε

$$B := \lfloor (nA)^{\frac{m}{n-m}} \rfloor. \tag{3.2.11}$$

Τότε, $(nA)^{\frac{m}{n-m}} < B + 1$, άρα

$$nA < (B+1)^{\frac{n-m}{m}}. \tag{3.2.12}$$

Έπεται ότι

$$nAB + 1 \leq nA(B+1) < (B+1)^{n/m}. \tag{3.2.13}$$

Τώρα, παρατηρούμε ότι αν $z = (z_1, \ldots, z_n)$ και $0 \leq z_k \leq B$, τότε

$$-r_j B \leq L_j(z) \leq s_j B \tag{3.2.14}$$

για κάθε $j = 1, \ldots, m$, όπου $-r_j$ και s_j είναι τα αθροίσματα των αρνητικών και θετικών συντελεστών a_{jk} αντίστοιχα. Αφού $r_j + s_j \leq nA$, κάθε $L_j(z)$ ανήκει σε ένα διάστημα μήκους το πολύ ίσου με nAB. Δηλαδή, κάθε $L_j(z)$ μπορεί να πάρει το πολύ $nAB + 1$ διαφορετικές τιμές. Έπεται ότι το διάνυσμα $(L_1(z), \ldots, L_m(z))$ μπορεί να πάρει το πολύ

$$(nAB+1)^m < (B+1)^n \tag{3.2.15}$$

διαφορετικές τιμές.

Από την άλλη πλευρά, το πλήθος των ακεραίων n-άδων $z = (z_1, \ldots, z_n)$ που ικανοποιούν την $0 \leq z_k \leq B$ είναι ίσο με $(B+1)^n$. Συνεπώς, υπάρχουν διακεκριμένες τέτοιες n-άδες $y^{(1)}$ και $y^{(2)}$ με την ιδιότητα

$$L_j(y^{(1)}) = L_j(y^{(2)}) \tag{3.2.16}$$

για κάθε $j = 1, \ldots, m$. Τότε, η n-άδα $y = y^{(1)} - y^{(2)}$ είναι ακέραια, μη τετριμμένη λύση του συστήματος και $\max_{1 \leq k \leq n} |y_k| \leq (nA)^{\frac{m}{n-m}}$. □

Λήμμα 3.2.3. *Έστω α αλγεβρικός ακέραιος, με ελάχιστο πολυώνυμο το*

$$Q(x) = x^d + t_1 x^{d-1} + \cdots + t_{d-1}x + t_d. \tag{3.2.17}$$

Για κάθε μη αρνητικό ακέραιο s, υπάρχουν ακέραιοι $t_1^{(s)}, \ldots, t_d^{(s)}$ που ικανοποιούν τις

$$\alpha^s = t_1^{(s)}\alpha^{d-1} + \cdots + t_{d-1}^{(s)}\alpha + t_d^{(s)} \tag{3.2.18}$$

και

$$|t_i^{(s)}| \leq (h(Q)+1)^s, \qquad 1 \leq i \leq d. \tag{3.2.19}$$

Απόδειξη. Το συμπέρασμα ισχύει προφανώς αν $s < d$. Συνεχίζουμε με επαγωγή ως προς s. Αν το συμπέρασμα ισχύει για τον a^{s-1}, γράφουμε

$$\begin{aligned}
\alpha^s = \alpha^{s-1}\alpha &= \left(t_1^{(s-1)}\alpha^{d-1} + \cdots + t_d^{(s-1)}\right)\alpha \\
&= t_1^{(s-1)}\alpha^d + t_2^{(s-1)}\alpha^{d-1} + \cdots + t_d^{(s-1)}\alpha \\
&= \left(t_2^{(s-1)} - t_1 t_1^{(s-1)}\right)\alpha^{d-1} + \cdots + \left(t_d^{(s-1)} - t_{d-1}t_1^{(s-1)}\right)\alpha - t_d t_1^{(s-1)}.
\end{aligned}$$

Δηλαδή,

$$\alpha^s = t_1^{(s)}\alpha^{d-1} + \cdots + t_{d-1}^{(s)}\alpha + t_d^{(s)}, \tag{3.2.20}$$

όπου $t_i^{(s)} = t_{i+1}^{(s-1)} - t_i t_1^{(s-1)}$ αν $1 \leq i < d$ και $t_d^{(s)} = -t_d t_1^{(s-1)}$. Τώρα, για κάθε $i = 1, \ldots, d$ έχουμε την εκτίμηση

$$|t_i^{(s)}| \leq (h(Q)+1)^{s-1} + h(Q)(h(Q)+1)^{s-1} = (h(Q)+1)^s. \tag{3.2.21}$$

Έτσι, ολοκληρώνεται το επαγωγικό βήμα. □

3.3 Δείκτης πολυωνύμου

Ορισμός 3.3.1 (δείκτης πολυωνύμου). Έστω $P(x_1, \ldots, x_m)$ πολυώνυμο m μεταβλητών με ακέραιους συντελεστές. Σταθεροποιούμε φυσικούς $r_1, \ldots, r_m$. Αν $P \neq 0$, για κάθε $(a_1, \ldots, a_m) \in \mathbb{R}^m$ ορίζουμε ως **δείκτη του P ως προς** $(a_1, \ldots, a_m; r_1, \ldots, r_m)$ την ελάχιστη τιμή της ποσότητας

$$\frac{i_1}{r_1} + \cdots + \frac{i_m}{r_m} \tag{3.3.1}$$

πάνω από όλους τους $i_1, \ldots, i_m \geq 0$ για τους οποίους $P_{i_1,\ldots,i_m}(a_1, \ldots, a_m) \neq 0$. Ειδικότερα, ο δείκτης $\mathrm{Ind}P(a_1, \ldots, a_m)$ του P στο $(a_1, \ldots, a_m)$ είναι ίσος με 0 αν $P(a_1, \ldots, a_m) \neq 0$.

Στην περίπτωση που $P \equiv 0$ ορίζουμε $\mathrm{Ind}P(a_1, \ldots, a_m) = +\infty$ για κάθε $(a_1, \ldots, a_m) \in \mathbb{R}^m$.

Στο επόμενα Λήμματα αποδεικνύουμε κάποιες βασικές ιδιότητες του δείκτη:

Λήμμα 3.3.2. *Έστω $P(x_1, \ldots, x_m)$, $i = 1, 2$, πολυώνυμο m μεταβλητών με ακέραιους συντελεστές, έστω $r_1, \ldots, r_m \in \mathbb{N}$ και $(a_1, \ldots, a_m) \in \mathbb{R}^m$. Τότε, για κάθε $i_1, \ldots, i_m \geq 0$,*

$$\mathrm{Ind}P_{i_1,\ldots,i_m}(a_1, \ldots, a_m) \geq \mathrm{Ind}P(a_1, \ldots, a_m) - \sum_{k=1}^{m} \frac{i_k}{r_k}. \tag{3.3.2}$$

Απόδειξη. Θέτουμε $T = P_{\vec{i}} = P_{i_1,\ldots,i_m}$. Παρατηρούμε ότι, αν $\vec{j} = (j_1, \ldots, j_m)$ είναι μια m-άδα μη αρνητικών ακεραίων, τότε $T_{\vec{j}} = P_{\vec{i}+\vec{j}}$. Αν λοιπόν $T_{\vec{j}}(a_1, \ldots, a_m) \neq 0$, τότε $P_{\vec{i}+\vec{j}}(a_1, \ldots, a_m) \neq 0$, το οποίο σημαίνει ότι

$$\sum_{k=1}^{m} \frac{i_k + j_k}{r_k} \geq \mathrm{Ind}P(a_1, \ldots, a_m). \tag{3.3.3}$$

Ισοδύναμα,

$$\sum_{k=1}^{m} \frac{j_k}{r_k} \geq \mathrm{Ind}P(a_1, \ldots, a_m) - \sum_{k=1}^{m} \frac{i_k}{r_k}. \tag{3.3.4}$$

Από τον ορισμό του δείκτη (για το T) προκύπτει το συμπέρασμα. □

Λήμμα 3.3.3. *Έστω $P^{(i)}(x_1, \ldots, x_m)$, $i = 1, 2$, πολυώνυμα m μεταβλητών με ακέραιους συντελεστές και έστω $r_1, \ldots, r_m \in \mathbb{N}$. Τότε, για κάθε $(a_1, \ldots, a_m) \in \mathbb{R}^m$,*

$$\mathrm{Ind}(P^{(1)} + P^{(2)})(a_1, \ldots, a_m) \geq \min\{\mathrm{Ind}P^{(1)}(a_1, \ldots, a_m), \mathrm{Ind}P^{(2)}(a_1, \ldots, a_m)\} \tag{3.3.5}$$

και

$$\mathrm{Ind}(P^{(1)} \cdot P^{(2)})(a_1, \ldots, a_m) = \mathrm{Ind}P^{(1)}(a_1, \ldots, a_m) + \mathrm{Ind}P^{(2)}(a_1, \ldots, a_m). \tag{3.3.6}$$

Απόδειξη. Για τον πρώτο ισχυρισμό, υποθέτουμε ότι $(P^{(1)} + P^{(2)})_{\vec{j}}(a_1, \ldots, a_m) \neq 0$. Τότε, είτε $P^{(1)}_{\vec{j}}(a_1, \ldots, a_m) \neq 0$ ή $P^{(2)}_{\vec{j}}(a_1, \ldots, a_m) \neq 0$. Άρα, είτε

$$\sum_{k=1}^{m} \frac{j_k}{r_k} \geq \mathrm{Ind}\, P^{(1)}(a_1, \ldots, a_m) \tag{3.3.7}$$

ή

$$\sum_{k=1}^{m} \frac{j_k}{r_k} \geq \mathrm{Ind}\, P^{(2)}(a_1, \ldots, a_m). \tag{3.3.8}$$

Σε κάθε περίπτωση,

$$\text{(3.3.9)}\quad \mathrm{Ind}(P^{(1)}+P^{(2)})(a_1,\dots,a_m)\geq \min\{\mathrm{Ind}P^{(1)}(a_1,\dots,a_m),\mathrm{Ind}P^{(2)}(a_1,\dots,a_m)\}.$$

Για τον δεύτερο ισχυρισμό, παρατηρούμε πρώτα ότι για κάθε $\vec{j}=(j_1,\dots,j_m)$ μπορούμε να γράφουμε

$$(P^{(1)}P^{(2)})_{\vec{j}}=\sum_{\vec{i}+\vec{i'}=\vec{j}} C(\vec{i},\vec{i'})P^{(1)}_{\vec{i}}P^{(2)}_{\vec{i'}}. \tag{3.3.10}$$

Από τον τρόπο ορισμού του $T_{\vec{i}}$ (στην (3.2.3)) μπορούμε μάλιστα να δούμε ότι $C(\vec{i},\vec{i'})=1$.

Υπάρχει m-άδα $\vec{j}$ για την οποία

$$\mathrm{Ind}\,(P^{(1)}P^{(2)})=\sum_{k=1}^{m}\frac{j_k}{r_k} \tag{3.3.11}$$

και

$$(P^{(1)}P^{(2)})(a_1,\dots,a_m)\neq 0. \tag{3.3.12}$$

Τότε, από την (3.3.10) συμπεραίνουμε ότι υπάρχουν $\vec{i}$ και $\vec{i'}$ με $\vec{i}+\vec{i'}=\vec{j}$ ώστε $P^{(1)}_{\vec{i}}(a_1,\dots,a_m)\neq 0$ και $P^{(2)}_{\vec{i'}}(a_1,\dots,a_m)\neq 0$. Αυτό σημαίνει ότι

$$\sum_{k=1}^{m}\frac{i_k}{r_k}\geq \mathrm{Ind}\,P^{(1)} \tag{3.3.13}$$

και

$$\sum_{k=1}^{m}\frac{i'_k}{r_k}\geq \mathrm{Ind}\,P^{(2)}. \tag{3.3.14}$$

Έπεται ότι

$$\begin{aligned}\mathrm{Ind}\,(P^{(1)}P^{(2)}) &= \sum_{k=1}^{m}\frac{j_k}{r_k}=\sum_{k=1}^{m}\frac{i_k}{r_k}+\sum_{k=1}^{m}\frac{i'_k}{r_k}\\ &\geq \mathrm{Ind}\,P^{(1)}+\mathrm{Ind}\,P^{(2)}.\end{aligned}$$

Για την αντίστροφη ανισότητα, θεωρούμε πρώτα όλες τις m-άδες $\vec{i}$ για τις οποίες

$$\mathrm{Ind}\,P^{(1)}=\sum_{k=1}^{m}\frac{i_k}{r_k} \tag{3.3.15}$$

και

$$P^{(1)}_{\vec{i}}(a_1, \ldots, a_m) \neq 0. \tag{3.3.16}$$

Από αυτές, επιλέγουμε την μικρότερη ως προς τη λεξικογραφική διάταξη, ας την πούμε $\vec{\bar{i}} = (\bar{i}_1, \ldots, \bar{i}_m)$. Τελείως ανάλογα, θεωρούμε τη μικρότερη λεξικογραφικά m-άδα $\vec{\bar{i'}} = (\bar{i'}_1, \ldots, \bar{i'}_m)$ για την οποία

$$\mathrm{Ind}\, P^{(2)} = \sum_{k=1}^{m} \frac{i'_k}{r_k} \tag{3.3.17}$$

και

$$P^{(2)}_{\vec{i'}}(a_1, \ldots, a_m) \neq 0. \tag{3.3.18}$$

Θέτουμε $\vec{j} = \vec{\bar{i}} + \vec{\bar{i'}}$. Τότε, από την (3.3.10) βλέπουμε ότι

$$(P^{(1)}P^{(2)})_{\vec{j}}(a_1, \ldots, a_m) = C(\vec{\bar{i}}, \vec{\bar{i'}}) P^{(1)}_{\vec{\bar{i}}}(a_1, \ldots, a_m) P^{(2)}_{\vec{\bar{i'}}}(a_1, \ldots, a_m) \neq 0. \tag{3.3.19}$$

Έπεται ότι

$$\begin{aligned} \mathrm{Ind}\,(P^{(1)}P^{(2)}) &\leq \sum_{k=1}^{m} \frac{j_k}{r_k} = \sum_{k=1}^{m} \frac{i_k}{r_k} + \sum_{k=1}^{m} \frac{i'_k}{r_k} \\ &= \mathrm{Ind}\, P^{(1)} + \mathrm{Ind}\, P^{(2)} \end{aligned}$$

και η απόδειξη έχει ολοκληρωθεί. □

3.4 Ο δείκτης στο $(\alpha, \ldots, \alpha)$

Το επόμενο θεώρημα εξασφαλίζει την ύπαρξη πολυωνύμων με «μικρούς συντελεστές» και «μεγάλο δείκτη» στο σημείο $(\alpha, \ldots, \alpha) \in \mathbb{R}^m$, όπου α αλγεβρικός αριθμός βαθμού $d \geq 2$:

Θεώρημα 3.4.1 (το θεώρημα του δείκτη). *Έστω α αλγεβρικός αριθμός βαθμού $d \geq 2$. Θεωρούμε τυχόν $\varepsilon \in (0,1)$ και έναν φυσικό αριθμό m που ικανοποιεί την*

$$m > 16\varepsilon^{-2} \log(4d). \tag{3.4.1}$$

Τότε, για κάθε επιλογή φυσικών αριθμών $r_1, \ldots, r_m$, υπάρχει πολυώνυμο $P(x_1, \ldots, x_m) \neq 0$ με ακέραιους συντελεστές, το οποίο έχει τις παρακάτω ιδιότητες:

(i) *Για κάθε $k = 1, \ldots, m$, ο βαθμός του P ως προς x_k είναι μικρότερος ή ίσος του r_k.*

(ii) *Ο δείκτης του P ως προς $r_1, \ldots, r_m$ στο σημείο $(\alpha, \ldots, \alpha)$ είναι μεγαλύτερος ή ίσος του $\frac{m}{2}(1 - \varepsilon)$.*

(iii) *Ισχύει* $h(P) \leq B^{r_1+\cdots+r_m}$ *για κάποια σταθερά* $B = B(\alpha) > 0$ *που εξαρτάται μόνο από τον* α.

Απόδειξη. Θεωρούμε πολυώνυμα της μορφής

$$P(x_1, \ldots, x_m) = \sum_{j=0}^{r_1} \cdots \sum_{j_m=0}^{r_m} C(j_1, \ldots, j_m) x_1^{j_1} \cdots x_m^{j_m} \tag{3.4.2}$$

με ακέραιους συντελεστές, τα οποία να ικανοποιούν τις συνθήκες (ii) και (iii). Το πλήθος των συντελεστών $C(j_1, \ldots, j_m)$ τους οποίους θέλουμε να προσδιορίσουμε είναι ίσο με

$$N = (r_1 + 1) \cdots (r_m + 1). \tag{3.4.3}$$

Για να πετύχουμε να είναι μεγαλύτερος ή ίσος του $\frac{m}{2}(1-\varepsilon)$ ο δείκτης του P στο $(\alpha, \ldots, \alpha)$, θα πρέπει να ικανοποιούνται οι

$$P_{i_1,\ldots,i_m}(\alpha, \ldots, \alpha) = 0 \tag{3.4.4}$$

για όλες τις m-άδες $(i_1, \ldots, i_m)$ που ικανοποιούν την

$$\sum_{k=1}^{m} \frac{i_k}{r_k} < (1 - \varepsilon)\frac{m}{2}. \tag{3.4.5}$$

Από το Λήμμα 3.1.1, το πλήθος αυτών των m-άδων είναι μικρότερο ή ίσο από

$$2(r_1 + 1) \cdots (r_m + 1)e^{-\varepsilon^2 m/16} \leq \frac{N}{2d} \tag{3.4.6}$$

αν πάρουμε υπ' όψιν μας την υπόθεση ότι $m > 16\varepsilon^{-2} \log(4d)$.

Αν σταθεροποιήσουμε μια m-άδα $(i_1, \ldots, i_m)$, τότε η εξίσωση $P_{i_1,\ldots,i_m}(\alpha, \ldots, \alpha) = 0$ είναι γραμμική ως προς τους αγνώστους $C(j_1, \ldots, j_m)$. Οι συντελεστές των αγνώστων είναι ακέραια πολλαπλάσια δυνάμεων του α, άρα είναι αλγεβρικοί αριθμοί. Αφού κάθε δύναμη του α είναι γραμμικός συνδυασμός των $1, \alpha, \alpha^2, \ldots, \alpha^{d-1}$ με ακέραιους συντελεστές, ισοδύναμα έχουμε ένα σύστημα d γραμμικών εξισώσεων ως προς $C(j_1, \ldots, j_m)$, με ακέραιους συντελεστές. Αν A είναι η μέγιστη απόλυτη τιμή αυτών των ακέραιων συντελεστών, το Λήμμα 3.2.3 μας δίνει το φράγμα

$$\begin{aligned} A &\leq \binom{j_1}{i_1} \cdots \binom{j_m}{i_m} (h(Q) + 1)^{(j_1-i_1)+\cdots+(j_m-i_m)} \\ &\leq 2^{j_1+\cdots+j_m}(h(Q) + 1)^{(j_1-i_1)+\cdots+(j_m-i_m)} \\ &\leq (2(h(Q) + 1))^{r_1+\cdots+r_m}, \end{aligned}$$

όπου $Q(x)$ είναι το ελάχιστο πολυώνυμο του α. Το συνολικό πλήθος των εξισώσεων είναι το πολύ ίσο με $M \leq d \cdot \frac{N}{2d}$, άρα το Λήμμα του Siegel εξασφαλίζει την ύπαρξη λύσης με

$$\begin{aligned} \max |C(j_1, \ldots, j_m)| &\leq (NA)^{\frac{M}{N-M}} \\ &\leq NA \leq 2^{r_1+\cdots+r_m}(2(h(Q) + 1))^{r_1+\cdots+r_m} \\ &= B(\alpha)^{r_1+\cdots+r_m}, \end{aligned}$$

με τη σταθερά $B(\alpha) = 4(h(Q) + 1)$ να εξαρτάται μόνο από τον α. □

3.5 Ο δείκτης σε ρητά σημεία κοντά στο $(\alpha, \ldots, \alpha)$

Σε αυτή την παράγραφο υποθέτουμε ότι ο α είναι αλγεβρικός ακέραιος βαθμού $d \geq 2$, θεωρούμε $0 < \varepsilon < 1$ και $m = m(\alpha, \varepsilon) > 16\varepsilon^{-2}\log(4d)$, σταθεροποιούμε φυσικούς $r_1, \ldots, r_m$ και θεωρούμε ένα πολυώνυμο $P(x_1, \ldots, x_m)$ με τις ιδιότητες που εξασφαλίζει το Θεώρημα 3.4.1.

Θεώρημα 3.5.1. *Έστω $0 < \delta < 1$ και*

$$0 < \varepsilon < \frac{\delta}{36}. \tag{3.5.1}$$

Υποθέτουμε ότι οι ρητοί $\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}$ ικανοποιούν τις

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_k^{2+\delta}}, \qquad j = 1, \ldots, m \tag{3.5.2}$$

και ότι

$$q_k^\delta > D, \qquad k = 1, \ldots, m \tag{3.5.3}$$

για κάποια σταθερά $D = D(\alpha) > 0$ που εξαρτάται μόνο από τον α. Αν, επιπλέον,

$$r_1 \log q_1 \leq r_k \log q_k \leq (1 + \varepsilon) r_1 \log q_1, \qquad k = 1, \ldots, m \tag{3.5.4}$$

τότε

$$\mathrm{Ind} P\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right) \geq \varepsilon m. \tag{3.5.5}$$

Απόδειξη. Έστω $j_1, \ldots, j_m$ μη αρνητικοί ακέραιοι για τους οποίους

$$\sum_{k=1}^m \frac{j_k}{r_k} < \varepsilon m. \tag{3.5.6}$$

Θεωρούμε το πολυώνυμο

$$T(x_1, \ldots, x_m) = P_{\vec{j}}(x_1, \ldots, x_m), \tag{3.5.7}$$

όπου $\vec{j} = (j_1, \ldots, j_m)$. Θα δείξουμε ότι

$$T\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right) = 0, \tag{3.5.8}$$

οπότε

$$\mathrm{Ind}P\left(\frac{p_1}{q_1},\dots,\frac{p_m}{q_m}\right)\geq \varepsilon m. \tag{3.5.9}$$

Από το Θεώρημα 3.4.1(iii) έχουμε

$$h(P)\leq B^{r_1+\cdots+r_m}. \tag{3.5.10}$$

Από το Λήμμα 3.2.1 βλέπουμε ότι

$$h(T)\leq (2B)^{r_1+\cdots+r_m}, \tag{3.5.11}$$

και εφαρμόζοντας ξανά το Λήμμα 3.2.1 παίρνουμε

$$h(T_{\vec{i}})\leq (4B)^{r_1+\cdots+r_m} \tag{3.5.12}$$

για κάθε m-άδα $\vec{i}=(i_1,\dots,i_m)$ μη αρνητικών ακεραίων. Θέτοντας $x_1=\cdots=x_m=\alpha$ έχουμε ότι η τιμή του $T_{\vec{i}}(\alpha,\dots,\alpha)$ προκύπτει από μονώνυμα που φράσσονται απολύτως από

$$(4B)^{r_1+\cdots+r_m}[\max\{1,|\alpha|\}]^{r_1+\cdots+r_m}. \tag{3.5.13}$$

Το πλήθος αυτών των μονωνύμων είναι μικρότερο ή ίσο από

$$(r_1+1)\cdots(r_m+1)\leq 2^{r_1+\cdots+r_m}. \tag{3.5.14}$$

Συνεπώς,

$$|T_{\vec{i}}(\alpha,\dots,\alpha)|\leq C(\alpha)^{r_1+\cdots+r_m}, \tag{3.5.15}$$

όπου $C(\alpha)=8B\max\{1,|\alpha|\}$.

Από το Θεώρημα 3.4.1(ii) έχουμε

$$\mathrm{Ind}\,P(\alpha,\dots,\alpha;r_1,\dots,r_m)\geq\frac{(1-\varepsilon)m}{2}. \tag{3.5.16}$$

Τότε, το Λήμμα 3.3.2 δείχνει ότι

$$\mathrm{Ind}\,T(\alpha,\dots,\alpha;r_1,\dots,r_m)\geq\frac{(1-\varepsilon)m}{2}-\sum_{k=1}^{m}\frac{j_k}{r_k}>\frac{(1-3\varepsilon)m}{2}. \tag{3.5.17}$$

Χρησιμοποιώντας το θεώρημα Taylor γράφουμε

$$T(p_1/q_1,\dots,p_m/q_m)=\sum_{i_1=0}^{r_1}\cdots\sum_{i_m=0}^{r_m}T_{i_1,\dots,i_m}(\alpha,\dots,\alpha)\left(\frac{p_1}{q_1}-\alpha\right)^{i_1}\cdots\left(\frac{p_m}{q_m}-\alpha\right)^{i_m}. \tag{3.5.18}$$

Σε αυτό το άθροισμα έχουμε $T_{i_1,\ldots,i_m}(\alpha,\ldots,\alpha)=0$ αν $\frac{i_1}{r_1}+\cdots+\frac{i_m}{r_m}\leq\frac{(1-3\varepsilon)m}{2}$. Χρησιμοποιώντας επίσης το άνω φράγμα για την $|T_{i_1,\ldots,i_m}(\alpha,\ldots,\alpha)|$ και το γεγονός ότι οι p_k/q_k είναι προσεγγίσεις του α, βλέπουμε ότι

$$|T(p_1/q_1,\ldots,p_m/q_m)|\leq\sum{}' C(\alpha)^{r_1+\cdots+r_m}\left(q_1^{i_1}\cdots q_m^{i_m}\right)^{-2-\delta}, \tag{3.5.19}$$

όπου $\sum'$ είναι το άθροισμα πάνω από τις m-άδες $\vec{i}=(i_1,\ldots,i_m)$ για τις οποίες $\frac{i_1}{r_1}+\cdots+\frac{i_m}{r_m}>\frac{(1-3\varepsilon)m}{2}$. Όμως, για κάθε τέτοια m-άδα έχουμε

$$\begin{aligned} q_1^{i_1}q_2^{i_2}\cdots q_m^{i_m} &= q_1^{r_1\frac{i_1}{r_1}}q_2^{r_2\frac{i_2}{r_2}}\cdots q_m^{r_m\frac{i_m}{r_m}} \\ &\geq q_1^{r_1\left(\frac{i_1}{r_1}+\cdots\frac{i_m}{r_m}\right)} \\ &> q_1^{r_1\frac{(1-3\varepsilon)m}{2}} \\ &\geq \left(q_1^{r_1}q_2^{r_2}\cdots q_m^{r_m}\right)^{\frac{1-3\varepsilon}{1+\varepsilon}} \\ &> \left(q_1^{r_1}q_2^{r_2}\cdots q_m^{r_m}\right)^{\frac{1-6\varepsilon}{2}}. \end{aligned}$$

Το πλήθος των προσθετέων στην (3.5.19) είναι μικρότερο ή ίσο από $2^{r_1+\cdots+r_m}$, άρα

$$|T(p_1/q_1,\ldots,p_m/q_m)|\leq\prod_{k=1}^{m}\left(2C(\alpha)q_k^{-\frac{1}{2}(1-6\varepsilon)(2+\delta)}\right)^{r_k}. \tag{3.5.20}$$

Όμως,

$$\frac{1}{2}(1-6\varepsilon)(2+\delta)>1+\frac{\delta}{2}-9\varepsilon>1+\frac{\delta}{4}, \tag{3.5.21}$$

άρα,

$$2C(\alpha)q_k^{-\frac{1}{2}(1-6\varepsilon)(2+\delta)}<2C(\alpha)q_k^{-1-\frac{\delta}{4}}<\frac{1}{q_k} \tag{3.5.22}$$

αν $q_k^{\delta}>(2C(\alpha))^4$. Αυτό ισχύει αν επιλέξουμε $D=(2C(\alpha))^4$. Τότε,

$$|T(p_1/q_1,\ldots,p_m/q_m)|<\frac{1}{q_1^{r_1}q_2^{r_2}\cdots q_m^{r_m}}.$$

Από την άλλη πλευρά, το P έχει βαθμό μικρότερο ή ίσο από r_k ως προς x_k, άρα το ίδιο ισχύει για το T. Συνεπώς,

$$T(p_1/q_1,\ldots,p_m/q_m)=\frac{s}{q_1^{r_1}q_2^{r_2}\cdots q_m^{r_m}} \tag{3.5.23}$$

για κάποιον ακέραιο s. Αναγκαστικά, $s=0$. Αυτό αποδεικνύει ότι $T(p_1/q_1,\ldots,p_m/q_m)=0$. □

Κεφάλαιο 4

Γενικευμένες Wronskian και το Λήμμα του Roth

4.1 Το λήμμα του Gauss

Αποδεικνύουμε πρώτα ένα Λήμμα του Gauss το οποίο θα χρησιμοποιήσουμε στη συνέχεια.

Λήμμα 4.1.1. *Εστω $f(x_1, \ldots, x_m)$ και $g(x_1, \ldots, x_m)$ πολυώνυμα m μεταβλητών. Υποθέτουμε ότι καθένα από τα f και g έχει σχετικά πρώτους ακέραιους συντελεστές. Τότε, οι συντελεστές του πολυωνύμου fg είναι σχετικά πρώτοι.*

Απόδειξη. Γράφουμε

$$f(x_1, \ldots, x_m) = \sum a_{i_1,\ldots,i_m} x_1^{i_1} \cdots x_m^{i_m} = \sum_I a_I I \tag{4.1.1}$$

και

$$g(x_1, \ldots, x_m) = \sum b_{j_1,\ldots,j_m} x_1^{j_1} \cdots x_m^{j_m} = \sum_J b_J J, \tag{4.1.2}$$

όπου το άθροισμα είναι πάνω από όλα τα μονώνυμα $I = x_1^{i_1} \cdots x_m^{i_m}$, $i_s \geq 0$, και μόνο πεπερασμένοι το πλήθος συντελεστές a_I, b_J δεν μηδενίζονται. Τότε,

$$(fg)(x_1, \ldots, x_m) = \sum_I \left(\sum_{JU=I} a_J b_U \right) I. \tag{4.1.3}$$

Συμφωνούμε να λέμε ότι το μονώνυμο $I := x_1^{i_1} \cdots x_m^{i_m}$ είναι *μικρότερο* από το $J := x_1^{j_1} \cdots x_m^{j_m}$ αν, από τις διαφορές $j_1 - i_1, \ldots, j_m - i_m$, η πρώτη η οποία δεν μηδενίζεται είναι θετική. Παρατηρήστε ότι αν $IJ = I_0 J_0$ τότε ισχύει ένα από τα παρακάτω:

(i) $I = I_0$ και $J = J_0$.

(ii) Το I είναι μικρότερο από το I_0.

(iii) Το J είναι μικρότερο από το J_0.

Έστω p ένας πρώτος αριθμός. Οι συντελεστές του f είναι σχετικά πρώτοι, άρα υπάρχει κάποιο ελάχιστο μονώνυμο I_0 ώστε ο p να μην διαιρεί τον a_{I_0}. Όμοια, υπάρχει ελάχιστο μονώνυμο J_0 ώστε ο p να μην διαιρεί τον b_{J_0}. Θεωρούμε τον

$$c_{I_0J_0} = \sum_{IJ=I_0J_0} a_I b_J. \tag{4.1.4}$$

Αν το I είναι μικρότερο από το I_0 τότε $p \mid a_I$ και αν το J είναι μικρότερο από το J_0 τότε $p \mid b_J$. Συνεπώς, ο p διαιρεί όλους τους προσθετέους του $c_{I_0J_0}$ εκτός από τον $a_{I_0}b_{J_0}$ τον οποίο δεν διαιρεί. Έπεται ότι ο p δεν διαιρεί τον $c_{I_0J_0}$, άρα δεν είναι κοινός διαιρέτης των συντελεστών του fg.

Αφού ο p ήταν τυχών πρώτος αριθμός, οι συντελεστές του πολυωνύμου fg είναι σχετικά πρώτοι. □

Πόρισμα 4.1.2. *Έστω $f(x_1, \ldots, x_m)$ και $g(x_1, \ldots, x_m)$ πολυώνυμα m μεταβλητών με ακέραιους συντελεστές. Αν a είναι ο μέγιστος κοινός διαιρέτης των συντελεστών του f και b είναι ο μέγιστος κοινός διαιρέτης των συντελεστών του g, τότε ο μέγιστος κοινός διαιρέτης των συντελεστών του fg είναι ο $c = ab$.*

Απόδειξη. Εφαρμόζουμε το Λήμμα 4.1.1 για τα πολυώνυμα $f_1 = \frac{1}{a}f$ και $g_1 = \frac{1}{b}g$. □

Πόρισμα 4.1.3. *Έστω $f(x_1, \ldots, x_m)$ και $g(x_1, \ldots, x_m)$ πολυώνυμα m μεταβλητών. Υποθέτουμε ότι το f έχει σχετικά πρώτους ακέραιους συντελεστές και το g έχει ρητούς συντελεστές. Αν το fg έχει ακέραιους συντελεστές, τότε οι συντελεστές του g είναι ακέραιοι.*

Απόδειξη. Υπάρχει $t \in \mathbb{Z}$ ώστε οι συντελεστές του tg να είναι ακέραιοι. Τότε, το πολυώνυμο $t(fg) = f(tg)$ έχει ακέραιους συντελεστές οι οποίοι έχουν, από την υπόθεση, κοινό διαιρέτη τον t. Από το Πόρισμα 4.1.2 συμπεραίνουμε ότι $t \mid ab$, όπου a είναι ο μέγιστος κοινός διαιρέτης των συντελεστών του f και b είναι ο μέγιστος κοινός διαιρέτης των συντελεστών του tg. Όμως, $a = 1$ από την υπόθεση. Έπεται ότι $t \mid b$. Αυτό αποδεικνύει ότι όλοι οι συντελεστές του g είναι ακέραιοι. □

Πόρισμα 4.1.4. *Έστω $f(x_1, \ldots, x_m)$ και $g(x_1, \ldots, x_m)$ πολυώνυμα m μεταβλητών με ρητούς συντελεστές. Αν το fg έχει ακέραιους συντελεστές, τότε υπάρχει ρητός $q \neq 0$ ώστε οι συντελεστές των qf και $\frac{1}{q}g$ να είναι ακέραιοι.*

Απόδειξη. Βρίσκουμε πρώτα ρητό q ώστε οι συντελεστές του qf να είναι σχετικά πρώτοι ακέραιοι. Τότε, αν γράψουμε $fg = (qf)\left(\frac{1}{q}g\right)$, μπορούμε να εφαρμόσουμε το Πόρισμα 4.1.3 για τα πολυώνυμα qf και $\frac{1}{q}g$. □

4.2 Γενικευμένες Wronskian

Έστω $\phi_1, \dots, \phi_s$ ρητές συναρτήσεις m μεταβλητών $x_1, \dots, x_m$. Θεωρούμε διαφορικούς τελεστές της μορφής

$$\Delta = \Delta_{i_1+\cdots+i_m} = \frac{\partial^{i_1+\cdots+i_m}}{\partial x_1^{i_1}\cdots \partial x_m^{i_m}}, \tag{4.2.1}$$

όπου $i_k \in \mathbb{Z}^+$. Ορίζουμε *τάξη* του τελεστή $\Delta_{i_1,\dots,i_m}$ τον μη αρνητικό ακέραιο $i_1+\cdots+i_m$.

Ορισμός 4.2.1 (γενικευμένη Wronskian). Με τον όρο *γενικευμένη Wronskian των* $\phi_1, \dots, \phi_s$ θα εννοούμε κάθε ορίζουσα της μορφής

$$\det\left(\Delta_i \phi_k\right)_{i,k=1}^{s}$$

όπου Δ_i τελεστής της μορφής (2.4.1) τάξης μικρότερης ή ίσης από $i-1$.

Παρατήρηση 4.2.2. Στην περίπτωση $m=1$ οι $\phi_1, \dots, \phi_s$ είναι συναρτήσεις μιας μεταβλητής. Για μια συνάρτηση $\phi(x)$ έχουμε

$$\Delta_1\phi = \phi, \quad \Delta_2\phi = \phi \text{ ή } \phi', \quad \Delta_3\phi = \phi \text{ ή } \phi' \text{ ή } \phi'' \text{ κλπ.} \tag{4.2.2}$$

Μπορεί κανείς να ελέγξει ότι, σε αυτή την περίπτωση, μια γενικευμένη Wronskian των $\phi_1, \dots, \phi_s$ μπορεί να μην είναι ταυτοτικά μηδενική μόνο αν είναι η συνηθισμένη Wronskian των ϕ_k, δηλαδή αν

$$\Delta_i \phi_k = \phi_k^{(i-1)}, \qquad i,k = 1, \dots, s. \tag{4.2.3}$$

Ο λόγος για την εισαγωγή των γενικευμένων Wronskian από τον Roth είναι ο εξής: αν $P(x_1, \dots, x_m)$ είναι ένα μη μηδενικό πολυώνυμο m μεταβλητών, τότε δεν είναι γενικά σωστό ότι το P παραγοντοποιείται στη μορφή

$$P(x_1, \dots, x_{m-1}, x_m) = P_1(x_1, \dots, x_{m-1}) P_2(x_m). \tag{4.2.4}$$

Κάθε όμως μη μηδενική γενικευμένη Wronskian έχει αυτή την ιδιότητα. Αντιστοιχίζοντας σε ένα πολυώνυμο P μια γενικευμένη Wronskian W, μπορούμε να χρησιμοποιήσουμε μια επαγωγική διαδικασία για να πάρουμε πληροφορίες για την W και, ενδεχομένως, για το ίδιο το P.

Λήμμα 4.2.3. *Έστω $\phi_1, \dots, \phi_k$ ρητές συναρτήσεις m μεταβλητών $x_1, \dots, x_m$ με πραγματικούς συντελεστές, οι οποίες είναι γραμμικά ανεξάρτητες πάνω από το $\mathbb{R}$. Τότε, υπάρχει γενικευμένη Wronskian των $\phi_1, \dots, \phi_k$ η οποία δεν μηδενίζεται ταυτοτικά.*

Απόδειξη. Με επαγωγή ως προς k. Αν $k=1$ τότε ο Δ_1 είναι αναγκαστικά ο ταυτοτικός τελεστής και η γενικευμένη Wronskian είναι η συνάρτηση ϕ_1. Αφού το $\{\phi_1\}$ είναι γραμμικά ανεξάρτητο πάνω από το $\mathbb{R}$, έχουμε $\phi_1 \neq 0$.

Υποθέτουμε τώρα ότι $k \geq 2$ και ότι $\phi_1, \ldots, \phi_k$ είναι ρητές συναρτήσεις που ικανοποιούν τις υποθέσεις του Λήμματος. Έστω T μια μη μηδενική ρητή συνάρτηση των $x_1, \ldots, x_m$ με πραγματικούς συντελεστές. Θεωρούμε τις συναρτήσεις

$$\psi_i = T\phi_i, \qquad i = 1, \ldots, k. \tag{4.2.5}$$

Οι $\psi_1, \ldots, \psi_k$ είναι κι αυτές γραμμικά ανεξάρτητες πάνω από το $\mathbb{R}$. Επίσης, κάθε γενικευμένη Wronskian των $\psi_1, \ldots, \psi_k$ είναι γραμμικός συνδυασμός γενικευμένων Wronskians των $\phi_1, \ldots, \phi_k$, με συντελεστές κάποιες ρητές συναρτήσεις στις οποίες υπεισέρχονται οι μερικές παράγωγοι της T. Για να δείξουμε το Λήμμα, αρκεί να δείξουμε ότι κάποια γενικευμένη Wronskian των $\psi_1, \ldots, \psi_k$ δεν μηδενίζεται ταυτοτικά.

Επιλέγουμε $T = \frac{1}{\phi_1}$. Τότε,

$$\psi_1 = 1, \psi_2 = \frac{\phi_2}{\phi_1}, \ldots, \psi_k = \frac{\phi_k}{\phi_1}. \tag{4.2.6}$$

Με βάση αυτό το συλλογισμό, μπορούμε να υποθέσουμε ότι για τις αρχικές ρητές συναρτήσεις $\phi_1, \ldots, \phi_k$ ισχύει $\phi_1 \equiv 1$.

Παρατηρούμε ότι το σύνολο όλων των γραμμικών συνδυασμών

$$t_1\phi_1 + \cdots + t_k\phi_k \tag{4.2.7}$$

με πραγματικούς συντελεστές $t_1, \ldots, t_k$, είναι γραμμικός χώρος V διάστασης k. Αφού $k > 1$ και οι $\phi_1 \equiv 1, \phi_2$ είναι γραμμικά ανεξάρτητες, η ϕ_2 δεν είναι σταθερή συνάρτηση. Άρα, υπάρχει j ώστε

$$\frac{\partial \phi_2}{\partial x_j} \neq 0. \tag{4.2.8}$$

Χωρίς περιορισμό της γενικότητας υποθέτουμε ότι $j = 1$. Θεωρούμε τον υπόχωρο W του V που αποτελείται από όλες τις συναρτήσεις $t_1\phi_1 + \cdots + t_k\phi_k$ για τις οποίες

$$\frac{\partial}{\partial x_1}\left(t_1\phi_1 + \cdots + t_k\phi_k\right) = 0. \tag{4.2.9}$$

Αφού $\phi_1 \in W$, ο W δεν είναι ο τετριμμένος υπόχωρος του V. Από την άλλη πλευρά, έχουμε $W \neq V$ διότι $\phi_2 \notin W$. Αν λοιπόν θέσουμε $s = \dim(W)$, τότε $1 \leq s \leq k-1$.

Επιλέγουμε μια βάση $\{\psi_1, \ldots, \psi_k\}$ του V ώστε το $\{\psi_1, \ldots, \psi_s\}$ να είναι βάση του W. Από την επαγωγική υπόθεση, υπάρχουν τελεστές $\Delta_1^*, \ldots, \Delta_s^*$ με τάξη $0, 1, \ldots, s-1$ αντίστοιχα, ώστε

$$W_1 = \det(\Delta_i^*\psi_j) \neq 0, \qquad 1 \leq i, j \leq s. \tag{4.2.10}$$

Παρατηρήστε ότι, αν $t_{s+1}, \ldots, t_k$ είναι πραγματικοί αριθμοί, όχι όλοι μηδέν, τότε

$$\frac{\partial}{\partial x_1}(t_{s+1}\psi_{s+1} + \cdots + t_k\psi_k) \neq 0. \tag{4.2.11}$$

Αυτό προκύπτει άμεσα από το γεγονός ότι ο υπόχωρος που παράγουν οι $\psi_{s+1}, \ldots, \psi_k$ έχει τετριμμένη τομή με τον W.

Η παρατήρηση αυτή δείχνει ότι οι ρητές συναρτήσεις

$$\frac{\partial}{\partial x_1}\psi_{s+1}, \ldots, \frac{\partial}{\partial x_1}\psi_k \tag{4.2.12}$$

είναι γραμμικά ανεξάρτητες πάνω από το $\mathbb{R}$. Χρησιμοποιώντας την επαγωγική υπόθεση βρίσκουμε τελεστές $\Delta^*_{s+1}, \ldots, \Delta^*_k$ με τάξη $0, 1, \ldots, k-s-1$ αντίστοιχα, ώστε

$$W_2 = \det\left(\Delta^*_i \frac{\partial}{\partial x_1}\psi_j\right) \neq 0, \qquad s+1 \leq i, j \leq k. \tag{4.2.13}$$

Ορίζουμε τελεστές Δ_i, $i = 1, \ldots, k$, θέτοντας

$$\Delta_i = \Delta^*_i, \qquad 1 \leq i \leq s \tag{4.2.14}$$

και

$$\Delta_i = \Delta^*_i \frac{\partial}{\partial x_1}, \qquad s+1 \leq i \leq k. \tag{4.2.15}$$

Κάθε τελεστής Δ_i έχει τάξη μικρότερη ή ίση από $i-1$, και

$$\det(\Delta_i \psi_j) = W_1 W_2 \neq 0. \tag{4.2.16}$$

Αφού το $\{\psi_1, \ldots, \psi_k\}$ είναι βάση του V, συμπεραίνουμε ότι

$$\det(\Delta_i \phi_j) \neq 0 \tag{4.2.17}$$

και η απόδειξη είναι πλήρης. □

Σημείωση. Το αντίστροφο του Λήμματος 3.1.3 ισχύει κι αυτό. Αν οι $\phi_1, \ldots, \phi_k$ είναι γραμμικά εξαρτημένες πάνω από το $\mathbb{R}$, τότε κάθε γενικευμένη Wronskian των $\phi_1, \ldots, \phi_k$ μηδενίζεται ταυτοτικά.

4.3 Το Λήμμα του Roth

Θεώρημα 4.3.1. *Έστω $0 < \varepsilon < \frac{1}{12}$. Σταθεροποιούμε $m \in \mathbb{N}$ και ορίζουμε*

$$\omega = \omega(m, \varepsilon) = 24 \cdot 2^{-m} \left(\frac{\varepsilon}{12}\right)^{2^{m-1}}. \tag{4.3.1}$$

Έστω $r_1, \ldots, r_m$ φυσικοί αριθμοί που ικανοποιούν τις

$$\omega r_k \geq r_{k+1}, \qquad k = 1, \ldots, m-1. \tag{4.3.2}$$

Έστω $(p_1, q_1), \ldots, (p_m, q_m)$ *ζευγάρια σχετικά πρώτων ακεραίων με* $q_k > 0$ *και*

$$q_k^{r_k} \geq q_1^{r_1}, \quad q_k^{\omega} \geq 2^{3m}, \qquad k = 1, \ldots, m. \tag{4.3.3}$$

Έστω $P(x_1, \ldots, x_m)$ *μη μηδενικό πολυώνυμο βαθμού μικρότερου ή ίσου από* r_k *ως προς* x_k $(1 \leq k \leq m)$ *με ακέραιους συντελεστές και*

$$h(P) \leq q_1^{\omega r_1}. \tag{4.3.4}$$

Τότε,

$$\operatorname{Ind} P\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right) \leq \varepsilon. \tag{4.3.5}$$

Απόδειξη. Με επαγωγή ως προς m.

$m = 1$: Μπορούμε να γράψουμε το P στη μορφή

$$P(x) = \left(x - \frac{p_1}{q_1}\right)^s P_1(x), \tag{4.3.6}$$

όπου $P_1(x)$ είναι πολυώνυμο με ρητούς συντελεστές, για το οποίο $P_1(p_1/q_1) \neq 0$. Ισοδύναμα, γράφουμε

$$P(x) = (q_1 x - p_1)^s R(x), \tag{4.3.7}$$

όπου $R(x) = q_1^{-s} P_1(x)$. Αφού οι p_1, q_1 είναι σχετικά πρώτοι, το Πόρισμα 4.1.3 δείχνει ότι το $R(x)$ έχει ακέραιους συντελεστές.

Αυτό σημαίνει ότι ο συντελεστής του μεγιστοβάθμιου όρου του $P(x)$ διαιρείται με q_1^s. Συνεπώς,

$$q_1^s \leq h(P) \leq q_1^{\omega r_1} = q^{\varepsilon r_1}, \tag{4.3.8}$$

διότι $\omega(1, \varepsilon) = \varepsilon$. Αφού $q_1 > 1$, έπεται ότι

$$s \leq \varepsilon r_1. \tag{4.3.9}$$

Όμως,

$$\operatorname{Ind} P\left(\frac{p_1}{q_1}, r_1\right) = \frac{s}{r_1}. \tag{4.3.10}$$

Αυτό αποδεικνύει το ζητούμενο στην περίπτωση $m = 1$.

Επαγωγικό βήμα $(m-1) \to m$: Θεωρούμε όλες τις αναπαραστάσεις της μορφής

$$P(x_1, \ldots, x_m) = \sum_{j=1}^{k} \phi_j(x_1, \ldots, x_{m-1}) \psi_j(x_m), \tag{4.3.11}$$

όπου $\phi_1,\dots,\phi_k$ και $\psi_1,\dots,\psi_k$ είναι πολυώνυμα με ρητούς συντελεστές. Το P έχει τουλάχιστον μία τέτοια αναπαράσταση με $k = r_m+1$: παίρνουμε $\psi_j(x_m) = x_m^j$, $j = 0,1,\dots,r_m$ και κατάλληλες ϕ_j. Επιλέγουμε αναπαράσταση με το ελάχιστο δυνατό μήκος k. Τότε,

$$k \leq r_m + 1. \tag{4.3.12}$$

Επίσης, αφού το k είναι το ελάχιστο δυνατό, μπορούμε να ελέγξουμε ότι οι $\phi_1,\dots,\phi_k$ είναι γραμμικά ανεξάρτητες πάνω από το $\mathbb{R}$. Πράγματι, αν αυτό δεν συμβαίνει, τότε υπάρχουν $t_1,\dots,t_k\in\mathbb{R}$ – όχι όλοι ίσοι με μηδέν – ώστε

$$t_1\phi_1+\cdots+t_k\phi_k = 0. \tag{4.3.13}$$

Μπορούμε μάλιστα να υποθέσουμε ότι όλοι οι t_j είναι ρητοί, διότι τα πολυώνυμα ϕ_j έχουν ρητούς συντελεστές. Τότε, αν για παράδειγμα $t_k\neq 0$, μπορούμε να γράψουμε

$$\phi_k = -\sum_{j=1}^{k-1}\frac{t_j}{t_k}\phi_j, \tag{4.3.14}$$

άρα

$$P = \sum_{j=1}^{k-1}\phi_j\psi_j - \sum_{j=1}^{k-1}\frac{t_j}{t_k}\phi_j\psi_k = \sum_{j=1}^{k-1}\phi_j\left(\psi_j - \frac{t_j}{t_k}\psi_k\right), \tag{4.3.15}$$

το οποίο είναι άτοπο από την ελαχιστική ιδιότητα του k. Με τον ίδιο τρόπο ελέγχουμε ότι οι $\psi_1,\dots,\psi_k$ είναι γραμμικά ανεξάρτητες πάνω από το $\mathbb{R}$.

Ορίζουμε

$$U(x_m) = \det\left(\frac{1}{(i-1)!}\frac{\partial^{i-1}}{\partial x_m^{i-1}}\psi_j(x_m)\right)_{1\leq i,j\leq k}. \tag{4.3.16}$$

Από το Λήμμα 4.2.3 συμπεραίνουμε ότι

$$U(x_m)\neq 0. \tag{4.3.17}$$

Επίσης, υπάρχουν τελεστές

$$\Delta_s = \frac{1}{i_1!\cdots i_{m-1}!}\frac{\partial^{i_1+\cdots+i_{m-1}}}{\partial x_1^{i_1}\cdots\partial x_{m-1}^{i_{m-1}}},\qquad 1\leq s\leq k \tag{4.3.18}$$

τάξης

$$i_1+\cdots+i_{m-1}\leq s-1\leq k-1\leq r_m \tag{4.3.19}$$

ώστε

$$V(x_1,\dots,x_{m-1}) := \det(\Delta_s\phi_j)_{1\le s,j\le k} \neq 0. \tag{4.3.20}$$

Ορίζουμε

$$W(x_1,\dots,x_m) = \det\left(\frac{1}{(j-1)!}\frac{\partial^{j-1}}{\partial x_m^{j-1}}\Delta_s P\right)_{1\le s,j\le k}. \tag{4.3.21}$$

Τότε,

$$W(x_1,\dots,x_m) = \det\left(\sum_{r=1}^{k}(\Delta_s\phi_r)\left(\frac{1}{(j-1)!}\frac{\partial^{j-1}}{\partial x_m^{j-1}}\psi_r\right)\right) = V(x_1,\dots,x_{m-1})U(x_m) \neq 0. \tag{4.3.22}$$

Παρατηρήστε ότι οι συντεταγμένες της ορίζουσας – μέσω της οποίας ορίζεται το W – έχουν ακέραιους συντελεστές. Συνεπώς, το W είναι πολυώνυμο με ακέραιους συντελεστές.

Για τη συνέχεια της απόδειξης θα χρειαστούμε το ακόλουθο Λήμμα.

Λήμμα 4.3.2. *Έστω Θ ο δείκτης της W ως προς $\left(\frac{p_1}{q_1},\dots,\frac{p_m}{q_m};r_1,\dots,r_m\right)$. Τότε,*

$$\Theta \le \frac{\varepsilon^2 k}{6}. \tag{4.3.23}$$

Απόδειξη του Λήμματος. Μπορούμε να υποθέσουμε, πολλαπλασιάζοντας τα V και U με κατάλληλο ρητό και τον αντίστροφό του (βλέπε Πόρισμα 4.1.4) ότι

$$W(x_1,\dots,x_m) = V(x_1,\dots,x_{m-1})U(x_m) \tag{4.3.24}$$

με τα V και U να έχουν ακέραιους συντελεστές.

Πρώτα φράσσουμε τα ύψη $h(U)$ και $h(V)$. Παρατηρήστε ότι

$$h(P_{i_1,\dots,i_{m-1},j-1}) \le 2^{r_1+\cdots+r_m}h(P) \le 2^{r_1+\cdots+r_m}q_1^{\omega r_1}. \tag{4.3.25}$$

Επίσης, το πλήθος των όρων στο πολυώνυμο $P_{i_1,\dots,i_{m-1},j-1}$ είναι το πολύ ίσο με $2^{r_1+\cdots+r_m}$ και το πλήθος των προσθετέων στο ανάπτυγμα της ορίζουσας που μας δίνει το W είναι

$$k! \le k^{k-1} \le k^{r_m} \le 2^{kr_m}. \tag{4.3.26}$$

Συνεπώς,

$$h(W) \le 2^{kr_m}\left(2^{r_1+\cdots+r_m}2^{r_1+\cdots+r_m}q_1^{\omega r_1}\right)^k \le \left(2^{3mr_1}q_1^{\omega r_1}\right)^k, \tag{4.3.27}$$

όπου χρησιμοποιήσαμε το γεγονός ότι

$$r_1 \ge r_2 \ge \cdots \ge r_m. \tag{4.3.28}$$

Από την υπόθεση,

$$h(W) \le \left(q_1^{2\omega r_1}\right)^k = q_1^{2\omega r_1 k}. \tag{4.3.29}$$

Έπεται ότι

$$h(U) \le q_1^{2\omega r_1 k} \tag{4.3.30}$$

και

$$h(V) \le q_1^{2\omega r_1 k}. \tag{4.3.31}$$

Εφαρμόζουμε τώρα την επαγωγική υπόθεση με το $m-1$ στη θέση του m, τους $kr_1, \dots, kr_m$ στη θέση των $r_1, \dots, r_m$, τον $\varepsilon^2/12$ στη θέση του ε, το $V(x_1, \dots, x_{m-1})$ στη θέση του $P(x_1, \dots, x_m)$. Οι υποθέσεις ικανοποιούνται για τον $\omega(m-1, \varepsilon^2/12) = 2\omega(m, \varepsilon)$ και

$$h(V) \le q_1^{\omega(m-1,\varepsilon^2/12)(kr_1)}. \tag{4.3.32}$$

Συμπεραίνουμε έτσι ότι

$$\operatorname{Ind} V\left(\frac{p_1}{q_1}, \dots, \frac{p_{m-1}}{q_{m-1}}; kr_1, \dots, kr_{m-1}\right) \le \frac{\varepsilon^2}{12}. \tag{4.3.33}$$

Έπεται ότι

$$\operatorname{Ind} V\left(\frac{p_1}{q_1}, \dots, \frac{p_{m-1}}{q_{m-1}}; r_1, \dots, r_{m-1}\right) \le \frac{k\varepsilon^2}{12}. \tag{4.3.34}$$

Θεωρώντας το V σαν πολυώνυμο των m μεταβλητών $x_1, \dots, x_m$, βλέπουμε ότι

$$\operatorname{Ind} V\left(\frac{p_1}{q_1}, \dots, \frac{p_{m-1}}{q_{m-1}}, \frac{p_m}{q_m}; r_1, \dots, r_{m-1}, r_m\right) \le \frac{\varepsilon^2}{12}. \tag{4.3.35}$$

Όμοια, οι υποθέσεις ικανοποιούνται για το $U = U(x_m)$ με $m = 1$, τον kr_m στη θέση του r_1 και τον $\varepsilon^2/12$ στη θέση του ε (παρατηρήστε ότι $\omega(1, \varepsilon^2/12) \ge 2\omega(m, \varepsilon)$). Χρησιμοποιώντας την επαγωγική υπόθεση βλέπουμε ότι

$$\operatorname{Ind} U\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m\right) \le \frac{k\varepsilon^2}{12}. \tag{4.3.36}$$

Αφού $P = UV$, από τις βασικές ιδιότητες του δείκτη παίρνουμε

$$\Theta = \operatorname{Ind} W\left(\frac{p_1}{q_1}, \dots, \frac{p_m}{q_m}; r_1, \dots, r_m\right) \le \frac{k\varepsilon^2}{12} + \frac{k\varepsilon^2}{12} = \frac{k\varepsilon^2}{6}. \tag{4.3.37}$$

Αυτό αποδεικνύει το Λήμμα. □

Συνέχεια της απόδειξης του Θεωρήματος. Έστω

$$\gamma = \operatorname{Ind} P\left(\frac{p_1}{q_1},\dots,\frac{p_m}{q_m};r_1,\dots,r_m\right). \tag{4.3.38}$$

Παρατηρούμε ότι

$$\begin{aligned}\operatorname{Ind} P_{i_1,\dots,i_{m-1},j-1} &\geq \gamma - \frac{i_1}{r_1} - \cdots - \frac{i_{m-1}}{r_{m-1}} - \frac{j-1}{r_m}\\ &\geq \gamma - \frac{i_1+\cdots+i_{m-1}}{r_{m-1}} - \frac{j-1}{r_m}\\ &\geq \gamma - \frac{r_m}{r_{m-1}} - \frac{j-1}{r_m}\\ &\geq \gamma - \omega - \frac{j-1}{r_m}\\ &\geq \gamma - \frac{\varepsilon^2}{24} - \frac{j-1}{r_m}.\end{aligned}$$

Κάθε συντεταγμένη της j-στής στήλης στην ορίζουσα που ορίζει το W είναι της μορφής $P_{i_1,\dots,i_{m-1},j-1}$. Από την ταυτότητα

$$\operatorname{Ind}(P^{(1)}P^{(2)}) = \operatorname{Ind}(P^{(1)}) + \operatorname{Ind}(P^{(2)}) \tag{4.3.39}$$

και την ανισότητα

$$\operatorname{Ind}(P^{(1)}+P^{(2)}) \geq \min\{\operatorname{Ind}(P^{(1)}),\operatorname{Ind}(P^{(2)})\}, \tag{4.3.40}$$

και αφού το W είναι ένα άθροισμα γινομένων καθένα από τα οποία έχει k όρους, έναν από κάθε στήλη, συμπεραίνουμε ότι

$$\begin{aligned}\Theta &\geq \sum_{j=1}^{k}\max\left\{\gamma - \frac{\varepsilon^2}{24} - \frac{j-1}{r_m},0\right\}\\ &\geq -\frac{k\varepsilon^2}{24} + \sum_{i=0}^{m-1}\max\left\{\gamma - \frac{i}{r_m},0\right\}.\end{aligned}$$

Άρα,

$$\sum_{i=0}^{k-1}\max\left\{\gamma - \frac{i}{r_m},0\right\} \leq \Theta + \frac{k\varepsilon^2}{24} \leq \frac{k\varepsilon^2}{6} + \frac{k\varepsilon^2}{24} < \frac{k\varepsilon^2}{4}. \tag{4.3.41}$$

Διακρίνουμε δύο περιπτώσεις:

Πρώτη περίπτωση: $\gamma > \frac{k-1}{r_m}$. Τότε, η τελευταία ανισότητα παίρνει τη μορφή

$$\frac{k}{2}\left(\gamma + \gamma - \frac{k-1}{r_m}\right) < \frac{k\varepsilon^2}{4}, \tag{4.3.42}$$

ή, ισοδύναμα,

$$\gamma + \left(\gamma - \frac{k-1}{r_m}\right) < \frac{\varepsilon^2}{2}. \tag{4.3.43}$$

Αφού $\gamma - \frac{k-1}{r_m} > 0$, έπεται ότι

$$\gamma < \frac{\varepsilon^2}{2} < \varepsilon. \tag{4.3.44}$$

Δεύτερη περίπτωση: $\gamma \leq \frac{k-1}{r_m}$. Τότε, η τελευταία ανισότητα παίρνει τη μορφή

$$\sum_{i=0}^{\lfloor \gamma r_m \rfloor} \left(\gamma - \frac{i}{r_m}\right) < \frac{k\varepsilon^2}{4}, \tag{4.3.45}$$

απ' όπου έπεται ότι

$$\frac{\gamma}{2}\big(\lfloor \gamma r_m \rfloor + 1\big) < \frac{k\varepsilon^2}{4}. \tag{4.3.46}$$

Τότε,

$$\frac{\gamma^2 r_m}{2} < \frac{k\varepsilon^2}{4}, \tag{4.3.47}$$

και αφού $k \leq r_m + 1 \leq 2r_m$, βλέπουμε ότι $\gamma^2 r_m < \varepsilon^2 r_m$, δηλαδή $\gamma < \varepsilon$.

Τόσο στην πρώτη όσο και στη δεύτερη περίπτωση, είδαμε ότι

$$\mathrm{Ind}\, P\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}; r_1, \ldots, r_m\right) < \varepsilon. \tag{4.3.48}$$

Δηλαδή, ισχύει το συμπέρασμα του Θεωρήματος. □

Κεφάλαιο 5

Το θεώρημα του Roth

5.1 Ανασκόπηση των προηγουμένων

Είμαστε τώρα σε θέση να ολοκληρώσουμε την απόδειξη του θεωρήματος του Roth. Θεωρούμε έναν αλγεβρικό αριθμό α βαθμού $d \geq 2$. Όπως είδαμε στο Κεφάλαιο 1, μπορούμε να υποθέσουμε ότι ο α είναι αλγεβρικός ακέραιος. Υποθέτουμε ότι, για κάποιον $\delta > 0$ η

$$\left|\alpha - \frac{p}{q}\right| < \frac{1}{q^{2+\delta}} \tag{5.1.1}$$

έχει άπειρες ρητές λύσεις, και θα οδηγηθούμε σε άτοπο.

Θα χρησιμοποιήσουμε τα παρακάτω αποτελέσματα που αποδείχτηκαν στα Κεφάλαια 3 και 4:

Θεώρημα 5.1.1. *Θεωρούμε τυχόν $\varepsilon \in (0,1)$ και έναν φυσικό αριθμό m που ικανοποιεί την*

$$m > 16\varepsilon^{-2}\log(4d). \tag{5.1.2}$$

Τότε, για κάθε επιλογή φυσικών αριθμών $r_1, \dots, r_m$, υπάρχει πολυώνυμο $P(x_1, \dots, x_m) \neq 0$ με ακέραιους συντελεστές, το οποίο έχει τις παρακάτω ιδιότητες:

(i) *Για κάθε $k = 1, \dots, m$, ο βαθμός του P ως προς x_k είναι μικρότερος ή ίσος του r_k.*

(ii) *Ο δείκτης του P ως προς $r_1, \dots, r_m$ στο σημείο $(\alpha, \dots, \alpha)$ είναι μεγαλύτερος ή ίσος του $\frac{m}{2}(1-\varepsilon)$.*

(iii) *Ισχύει $h(P) \leq B^{r_1+\cdots+r_m}$ για κάποια σταθερά $B = B(\alpha) > 0$ που εξαρτάται μόνο από τον α.*

Για το πολυώνυμο $P(x_1, \dots, x_m)$ με τις ιδιότητες που εξασφαλίζει το Θεώρημα 5.1.1 αποδείξαμε το εξής:

Θεώρημα 5.1.2. *Έστω* $0 < \delta < 1$ *και*

$$0 < \varepsilon < \frac{\delta}{36}. \tag{5.1.3}$$

Υποθέτουμε ότι οι ρητοί $\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}$ *ικανοποιούν τις*

$$\left|\alpha - \frac{p_k}{q_k}\right| < \frac{1}{q_k^{2+\delta}}, \qquad j = 1, \ldots, m \tag{5.1.4}$$

και ότι

$$q_k^{\delta} > D, \qquad k = 1, \ldots, m \tag{5.1.5}$$

για κάποια σταθερά $D = D(\alpha) > 0$ *που εξαρτάται μόνο από τον* α. *Αν, επιπλέον,*

$$r_1 \log q_1 \leq r_k \log q_k \leq (1+\varepsilon) r_1 \log q_1, \qquad k = 1, \ldots, m \tag{5.1.6}$$

τότε

$$\mathrm{Ind} P\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right) \geq \varepsilon m. \tag{5.1.7}$$

Τέλος, θα χρησιμοποιήσουμε το *Λήμμα του Roth*:

Θεώρημα 5.1.3. *Έστω* $0 < \varepsilon < \frac{1}{12}$. *Σταθεροποιούμε* $m \in \mathbb{N}$ *και ορίζουμε*

$$\omega = \omega(m, \varepsilon) = 24 \cdot 2^{-m} \left(\frac{\varepsilon}{12}\right)^{2^{m-1}}. \tag{5.1.8}$$

Έστω $r_1, \ldots, r_m$ *φυσικοί αριθμοί που ικανοποιούν τις*

$$\omega r_k \geq r_{k+1}, \qquad k = 1, \ldots, m-1. \tag{5.1.9}$$

Έστω $(p_1, q_1), \ldots, (p_m, q_m)$ *ζευγάρια σχετικά πρώτων ακεραίων με* $q_k > 0$ *και*

$$q_k^{r_k} \geq q_1^{r_1}, \quad q_k^{\omega} \geq 2^{3m}, \qquad k = 1, \ldots, m. \tag{5.1.10}$$

Έστω $P(x_1, \ldots, x_m)$ *μη μηδενικό πολυώνυμο βαθμού μικρότερου ή ίσου από* r_k *ως προς* x_k $(1 \leq k \leq m)$ *με ακέραιους συντελεστές και*

$$h(P) \leq q_1^{\omega r_1}. \tag{5.1.11}$$

Τότε,

$$\mathrm{Ind}\, P\left(\frac{p_1}{q_1}, \ldots, \frac{p_m}{q_m}\right) \leq \varepsilon. \tag{5.1.12}$$

5.2 Απόδειξη του θεωρήματος

1. Θεωρούμε αλγεβρικό ακέραιο α βαθμού $d \geq 2$ για τον οποίο η (5.1.1) έχει άπειρες λύσεις. Χωρίς περιορισμό της γενικότητας, υποθέτουμε ότι $0 < \delta < 1$.

2. Επιλέγουμε $\varepsilon > 0$ με $0 < \varepsilon < \delta/36$. Δηλαδή, ικανοποιείται η (5.1.3). Αφού $\delta < 1$, ικανοποιείται έτσι και η υπόθεση $0 < \varepsilon < \frac{1}{12}$ του Θεωρήματος 5.1.3.

3. Επιλέγουμε φυσικό $m > 16\varepsilon^{-2}\log(4d)$. Τότε, ικανοποιείται η (5.1.2). Επίσης, ορίζουμε $\omega(m,\varepsilon)$ όπως στην (5.1.8): $\omega = 24 \cdot 2^{-m}\left(\frac{\varepsilon}{12}\right)^{2^{m-1}}$.

4. Θεωρούμε μια λύση $\frac{p_1}{q_1}$ της (5.1.1) με τους p_1, q_1 σχετικά πρώτους, τον q_1 θετικό, και

$$q_1^{\omega} > B^m, \tag{5.2.1}$$

όπου B η σταθερά στο Θεώρημα 5.1.1(iii). Ζητάμε επιπλέον να ισχύουν οι

$$q_1^{\delta} > D \ \text{ και } \ q_1^{\omega} \geq 2^{3m} \tag{5.2.2}$$

(βλέπε (5.1.5) και (5.1.10)). Οι τρεις αυτοί περιορισμοί ικανοποιούνται αν ο q_1 είναι αρκετά μεγάλος, κάτι που μπορούμε να εξασφαλίσουμε διότι η (5.1.1) έχει άπειρες λύσεις.

5. Στη συνέχεια επιλέγουμε διαδοχικά λύσεις $\frac{p_2}{q_2}, \ldots, \frac{p_m}{q_m}$ της (5.1.1) με τους p_k, q_k σχετικά πρώτους, κάθε q_k θετικό, και

$$\omega \log q_{k+1} \geq 2\log q_k, \qquad k = 1, \ldots, m-1. \tag{5.2.3}$$

Εδώ χρησιμοποιούμε πάλι την υπόθεση ότι η (5.1.1) έχει άπειρες λύσεις. Από την επιλογή των q_k έχουμε

$$q_1 < q_2 < \cdots < q_k < q_{k+1} < \cdots < q_m. \tag{5.2.4}$$

Ειδικότερα,

$$q_k^{\delta} > D \ \text{ και } \ q_k^{\omega} \geq 2^{3m}. \tag{5.2.5}$$

Δηλαδή, οι $q_1, q_2, \ldots, q_m$ ικανοποιούν τις (5.1.5) και (5.1.10).

6. Θεωρούμε φυσικό r_1 αρκετά μεγάλο ώστε να ισχύει

$$\varepsilon r_1 \log q_1 \geq \log q_m. \tag{5.2.6}$$

7. Για κάθε $k = 2, \ldots, m$ ορίζουμε

$$r_k = \left\lfloor \frac{r_1 \log q_1}{\log q_k} \right\rfloor + 1. \tag{5.2.7}$$

Τότε, για κάθε $k = 2, \ldots, m$ έχουμε

$$r_1 \log q_1 < r_k \log q_k \leq r_1 \log q_1 + \log q_k \leq (1+\varepsilon) r_1 \log q_1. \tag{5.2.8}$$

Δηλάδή, ικανοποιούνται οι (5.1.6) και (5.1.10). Επίσης, βλέπουμε ότι

$$r_{k+1}\log q_{k+1} \leq (1+\varepsilon) r_k \log q_k, \qquad k=1,\ldots,m-1. \tag{5.2.9}$$

Άρα,

$$\omega r_k \geq \omega \frac{r_{k+1}\log q_{k+1}}{(1+\varepsilon)\log q_k} \geq \frac{2r_{k+1}}{1+\varepsilon}, \tag{5.2.10}$$

δηλαδή

$$\omega r_k \geq r_{k+1}. \tag{5.2.11}$$

Ικανοποιείται έτσι και η (5.1.9).
8. Μπορούμε τώρα, εφαρμόζοντας το Θεώρημα 5.1.1, να βρούμε πολυώνυμο $P(x_1,\ldots,x_m) \neq 0$ με ακέραιους συντελεστές, με βαθμό ως προς x_k μικρότερο ή ίσο του r_k, δείκτη ως προς $r_1,\ldots,r_m$ στο σημείο $(\alpha,\ldots,\alpha)$ μεγαλύτερο ή ίσο του $\frac{m}{2}(1-\varepsilon)$, και $h(P) \leq B^{r_1+\cdots+r_m}$.
9. Όλες οι υποθέσεις του Θεωρήματος 5.1.2 ικανοποιούνται από το P. Συνεπώς,

$$\mathrm{Ind} P\left(\frac{p_1}{q_1},\ldots,\frac{p_m}{q_m}\right) \geq \varepsilon m. \tag{5.2.12}$$

10. Από την άλλη πλευρά, ικανοποιούνται και όλες οι υποθέσεις του Θεωρήματος 5.1.3. Η μόνη που μένει να ελεγχθεί είναι η (5.1.11). Όμως,

$$h(P) \leq B^{r_1+\cdots+r_m} \leq B^{mr_1} \leq q_1^{\omega r_1} \tag{5.2.13}$$

λόγω των (5.2.11) και (5.2.1). Από το Θεώρημα 5.1.3 έπεται ότι

$$\mathrm{Ind}\, P\left(\frac{p_1}{q_1},\ldots,\frac{p_m}{q_m}\right) \leq \varepsilon. \tag{5.2.14}$$

Οι (5.2.12) και (5.2.14) οδηγούν σε άτοπο. □

Βιβλιογραφία

[1] E. Bombieri, D. C. Hunt and A. J. van der Poorten, *Determinants in the study of Thue's method and curves with prescribed singularities*, Experimental Mathematics **4** (1995), 87–96.

[2] J. W. S. Cassels, *An introduction to Diophantine Approximation*, Cambridge Tracts in Mathematics and Mathematical Physics **35** (1957), New York, Cambridge University Press.

[3] E. Croot, *An outline of the Thue–Siegel Theorem*, Lecture Notes (2007).

[4] F. J. Dyson, *The approximation to algebraic numbers by rationals*, Acta Mathematica **79** (1947), 225–240.

[5] J. Liouville, *Sur les classes très étendues de quantités dont la valeur n'est ni algébrique, ni même réductible a des irrationelles algébriques*, C. R. Acad. Sci. Paris **18** (1844), 883–885, 910–911.

[6] K. F. Roth, *Rational approximations to algebraic numbers*, Mathematika **2** (1955), 1–20.

[7] C. L. Siegel, *Approximation algebraischer Zahlen*, Mathematische Zeitschrift **10** (1921), 173–213.

[8] W. M. Schmidt, *Diophantine Approximation*, Lecture Notes in Mathematics **785** (1980), Springer-Verlag, Berlin.

[9] A. Thue, *Über Annäherungswerte algebraischer Zahlen*, J. Reine Ang. Math. **135** (1909), 284–305.