Μάθημα 3

Απαλοίφουσα και Διακρίνουσα

Σε αυτήν την ενότητα θα συναντήσουμε δυο σημαντικές έννοιες που συνδέονται με τα πολυώνυμα μιας(ή και πολλών) μεταβλητής· την απαλοίφουσα και τη διακρίνουσα. Η απαλοίφουσα συνδέεται με τις ακολουθίες Sturm μέσω της ακολουθίας Stum-Habicht, η οποία είναι μια εξαιρετικά αποδοτική ακολουθία Sturm, με την έννοια ότι οι συντελεστές των πολυωνύμων που την αποτελούν έχουν αρκετά μικρό δυαδικό μήκος και, όπως έχουμε αναφέρει, ο υπολογισμός της είναι σχετικά ταχύς.

3.1 Παραγοντοποίηση και μέγιστος κοινός διαιρέτης

Στο *Μάθημα 1* είδαμε πως ένας δακτύλιος χωρίς μηδενοδιαιρέτες[1] ονομάζεται ακέραια περιοχή. Συνεχίζουμε ορίζοντας περιοχές μοναδικής παραγοντοποίησης και περιοχές με διαίρεση:

Ορισμός 3.1. *Έστω R ένας δακτύλιος και $a, b \in R$.*
(α) Θα λέμε ότι το a διαιρεί το b (συμβολισμός $a|b$) αν υπάρχει $c \in R$ τέτοιο ώστε $b = ac$.
(β) Τα a και b ονομάζονται συντροφικά(associate) στον R αν $a = ub$ για κάποιο αντιστρέψιμο $u \in R$.

Πχ τα μόνα συντροφικά στοιχεία του 5 στο $\mathbb{Z}$ είναι το 5 και το -5. Τα συντροφικά στοιχεία του $g(x) \in F[x]$ όπου F σώμα, είναι τα $ug(x)$ όπου $u \in F - \{0\}$.

Ορισμός 3.2. *Έστω R μια ακέραια περιοχή και $u \in R$. Το u ονομάζεται ανάγωγο(irreducible) στην R αν*
(α) το u δεν είναι μηδέν και δεν είναι αντιστρέψιμο, και
(β) Αν $u = ab$ με $a, b \in R$ τότε το a ή το b είναι αντιστρέψιμο.

Πχ τα ανάγωγα στοιχεία του $\mathbb{Z}$ είναι οι πρώτοι αριθμοί, ενώ τα ανάγωγα πολυώνυμα του $\mathbb{R}[x]$ είναι τα πρωτοβάθμια και τα δευτεροβάθμια με αρνητική διακρίνουσα.

Ορισμός 3.3. *Μια ακέραια περιοχή R καλείται περιοχή μοναδικής παραγοντοποίησης (unique factorization domain) ανν κάθε στοιχείο επιδέχεται «μοναδικής» παραγοντοποίησης. Πιο τυπικά,*
(α) κάθε μη μηδενικό και μη αντιστρέψιμο $a \in R$ παραγοντοποιείται ως $a = c_1 \cdots c_n$, όπου τα c_i ανάγωγα,
(β) αν υπάρχει κι άλλη παραγοντοποίηση $a = d_1 \cdots d_m$, τότε $m = n$ και υπάρχει μια 1-1 αντιστοίχιση των c_i, d_i τέτοια ώστε τα c_i, d_i να είναι συντροφικά(ή αλλιώς $c_i|d_i$ και $d_i|c_i$).

Ορισμός 3.4. Ευκλείδειος δακτύλιος *(Euclidean ring) λέγεται κάθε αντιμεταθετικός δακτύλιος D όπου ορίζεται μια συνάρτηση «διαβάθμισης» $\phi : D \to \mathbb{N}$, τέτοια ώστε $ab \neq 0 \Rightarrow \phi(ab) \leq \phi(a)$ και ορίζεται η διαίρεση $a = bq + r$ με υπόλοιπο r, όπου $r = 0$ ή $\phi(r) < \phi(b)$.*
Ένας ευκλείδειος δακτύλιος που είναι και ακέραια περιοχή λέγεται Ευκλείδεια περιοχή (Euclidean domain).

πχ το $\mathbb{Z}$ είναι Ευκλείδεια περιοχή με συνάρτηση διαβάθμισης την απόλυτο τιμή. Έστω σώμα K. Το K είναι Ευκλείδεια περιοχή με συνάρτηση διαβάθμισης $K \mapsto 1$. Επίσης το $K[x]$ είναι Ευκλείδεια περιοχή με συνάρτηση διαβάθμισης τον βαθμό πολυωνύμου.

Αποδεικνύεται πως κάθε Ευκλείδεια περιοχή είναι περιοχή μοναδικής παραγοντοποίησης. Συνεπώς, μια ιεραρχία δακτυλίων είναι: Αντιμεταθετικοί δακτύλιοι, Ακέραιες περιοχές, Περιοχές μοναδικής παραγοντοποίησης, Ευκλείδειες περιοχές, σώματα.

[1] μηδενοδιαιρέτες ενός δακτυλίου είναι δυο μη μηδενικά στοιχεία α, β αυτού, για τα οποία $\alpha\beta = 0$. Αν ο δακτύλιος δεν είναι μεταθετικός διακρίνουμε αριστερούς και δεξιούς μηδενοδιαιρέτες.

Ορισμός 3.5. *Ο μέγιστος κοινός διαιρέτης ΜΚΔ(a,b) σε μια Ευκλείδειο περιοχή είναι το μέγιστο(ως προς τη συνάρτηση διαβάθμισης) στοιχείο της που διαιρεί τα a και b. Στους ακέραιους πρόκειται για το μέγιστο θετικό, στα πολυώνυμα μίας μεταβλητής, αυτό με μέγιστο βαθμό. Το ελάχιστο κοινό πολλαπλάσιο ΕΚΠ(a,b) είναι το μικρότερο στοιχείο που διαιρεί τα a,b.*

Σημαντική είναι η εξής ιδιότητα: ΕΚΠ(a,b) ΜΚΔ$(a,b) = ab$.

Ο Αλγόριθμος του Ευκλείδη για την εύρεση του ΜΚΔ είναι ο αρχαιότερος αλγόριθμος σε χρήση σήμερα: Δεδομένων δυο στοιχείων $a,b \in D$ θέτουμε $c_0 = a, c_1 = b$. Για $i = 2,3,\dots$ εκτελούμε τις διαιρέσεις με υπόλοιπο: $c_{i-2} = c_{i-1}q_i + c_i$, δηλαδή το επόμενο στοιχείο στην ακολουθία c_i είναι το υπόλοιπο της προηγούμενης διαίρεσης. Η ακολουθία τερματίζεται όταν $c_k = 0$ οπότε και ο ΜΚΔ$(a,b) = c_{k-1}$.

3.2 Ο πίνακας Sylvester

Έστω $p_1 = a_{d_1}x^{d_1} + \cdots + a_0$ και $p_2 = b_{d_2}x^{d_2} + \cdots + b_0$ δυο πολυώνυμα, $p_1, p_2 \in D[x]$, όπου D μια Ευκλείδεια περιοχή. Ο πίνακας Sylvester των p_1, p_2 είναι ένας πίνακας $S \in D^{(d_1+d_2)\times(d_1+d_2)}$ του οποίου οι στήλες αντιστοιχούν σε δυνάμεις της μεταβλητής x και οι γραμμές σε πολλαπλάσια των πολυωνύμων ως εξής[2]:

$$S := \begin{array}{c} \\ p_1 \\ xp_1 \\ \vdots \\ x^{d_2-1}p_1 \\ p_2 \\ xp_2 \\ \vdots \\ x^{d_1-1}p_2 \end{array} \begin{array}{c} \begin{array}{ccccccc} 1 & x & x^2 & \dots & \dots & \dots & x^{d_1+d_2-1} \end{array} \\ \begin{bmatrix} a_0 & a_1 & \cdots & a_{d_1} & & & \\ & a_0 & a_1 & \cdots & a_{d_1} & & \mathbf{0} \\ \mathbf{0} & & \ddots & \ddots & \ddots & \ddots & \\ & & & a_0 & a_1 & \cdots & a_{d_1} \\ b_0 & b_1 & \cdots & b_{d_2} & & & \\ & b_0 & b_1 & \cdots & b_{d_2} & & \mathbf{0} \\ \mathbf{0} & & \ddots & \ddots & \ddots & \ddots & \\ & & & b_0 & b_1 & \cdots & b_{d_2} \end{bmatrix} \end{array}$$

Παρατηρήστε πως ο S αποτελείται από δυο υποπίνακες(blocks), οι οποίοι έχουν όλες τις διαγωνίους τους σταθερές. Τέτοιοι πίνακες ονομάζονται πίνακες Toeplitz:

Ορισμός 3.6. *Ενας πίνακας T, διάστασης $n \times m$ καλείται Toeplitz εάν τα έχει σταθερές διαγωνίους. Δηλαδή ο πίνακας ορίζεται από δυο διανύσματα $(d, t_1, \cdots, t_{m-1})$ και $(\tau_1, \tau_2, \dots, \tau_{n-1})$:*

$$T = \begin{bmatrix} d & t_1 & t_2 & \dots & \dots & t_{m-1} \\ \tau_1 & d & t_1 & \ddots & & \vdots \\ \tau_2 & \tau_1 & \ddots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & t_1 & t_2 \\ \vdots & & \ddots & \tau_1 & d & t_1 \\ \tau_{n-1} & \dots & \dots & \tau_2 & \tau_1 & d \end{bmatrix}$$

Οι πίνακες Toeplitz αλλά και οι block(κατά ομάδες) Toeplitz έχουν ενδιαφέρουσες ιδιότητες:

- Ο πολλαπλασιασμός $T\underline{w}$ ενός $n \times n$, κάτω τριγωνικού πίνακα Toeplitz από δεξιά με ένα διάνυσμα $\underline{w}$ εκφράζει τον πολλαπλασιασμό πολυωνύμων, όπου οι συντελεστές των πολυωνύμων βρίσκονται στην πρώτη στήλη του T και στα στοιχεία του $\underline{w}$ αντίστοιχα. Όμοια και για το γινόμενο $\underline{w}^tT$, όπου T άνω τριγωνικός.
- Όταν ο T είναι block Toeplitz, ο ο πολλαπλασιασμός με ένα πολυώνυμο με συντελεστές $\underline{w}$ εκφράζει το άθροισμα γινομένων πολυωνύμων.
- Η αντιμετάθεση στηλών(και γραμμών εφόσον δεν παραβιάζεται η block μορφή) δίνει ένα νέο πίνακα Toeplitz κατά block, ο οποίος έχει τις ίδιες ιδιότητες με τον αρχικό.

[2] Στη βιβλιογραφία μπορεί να συναντήσετε τον ορισμό του S ως τον ανάστροφο του πίνακα που δίνεται εδώ

3.3 Η Απαλοίφουσα δυο πολυωνύμων

Υπάρχουν αρκετοί τρόποι να ορίσεις την απαλοίφουσα. Για τις ανάγκες του μαθήματος δίνουμε τον παρακάτω

Ορισμός 3.7. *Αν $p_1, p_2 \in D[x]$ καλούμε απαλοίφουσα την ορίζουσα του πίνακα Sylvester: $R(p_1, p_2) := \det S$.*

Επειδή ο πίνακας Sylvester έχει στοιχεία από το D, η απαλοίφουσα ανήκει στο D. Επίσης από τη μορφή του S προκύπτει πως η απαλοίφουσα έχει βαθμό d_2 και d_1 ως προς τους συντελεστές του p_1 και του p_2 αντίστοιχα.

Με τις γνωστές ιδιότητες των οριζουσών μπορεί να δειχθεί η ιδιότητα $R[(x-r)p_1(x), p_2(x)] = p_2(r)R[p_1, p_2]$. Γενικότερα με επαγωγή αποδεικνύεται η λεγόμενη μορφή Poisson(Poisson formula):

$$R(p_1, p_2) = a_{d_1}^{d_2} \prod_{i=1}^{d_2} p(r_i) \tag{3.1}$$

όπου r_i οι d_2 ρίζες του p_2 στην αλγεβρική θήκη του D.

Θεώρημα 3.1. *Η απαλοίφουσα είναι μηδέν ανν τα p_1, p_2 έχουν κοινή ρίζα, με άλλα λόγια ισχύει η ισοδυναμία*

$$R(p_1, p_2) = 0 \iff \deg\left[\text{ΜΚΔ}(p_1, p_2)\right] \geq 1$$

Απόδειξη. ($\Leftarrow$) Έστω r κοινή ρίζα των p_1, p_2. Αρκεί ο πίνακας S να είναι ιδιάζων. Ισοδύναμα, αρκεί να υπάρχει μη μηδενικό διάνυσμα στον $\ker S$. Θέτουμε $\underset{\sim}{w} = \begin{bmatrix} 1 & r & \cdots & r^{d_1+d_2} \end{bmatrix}^t$. Λόγω της πρώτης συντεταγμένης, $\underset{\sim}{w} \neq 0$, όμως

$$S\underset{\sim}{w} = \begin{bmatrix} p_1(r) \\ \vdots \\ r^{d_2-1}p_1(r) \\ p_2(r) \\ \vdots \\ r^{d_1-1}p_2(r) \end{bmatrix} = \underset{\sim}{0} \quad \Rightarrow \quad \underset{\sim}{w} \in \ker S$$

($\Rightarrow$) Έστω ότι $\det S = 0$. Τότε στον αριστερό πυρήνα θα υπάρχει μη μηδενικό διάνυσμα: $\exists \underset{\sim}{v} \neq 0,\ \underset{\sim}{v} \in \ker S^T$. Έστω $\underset{\sim}{v} = [\kappa_0\ \kappa_1\ \cdots\ \kappa_{d_2-1}\ \lambda_0\ \lambda_1\ \cdots\ \lambda_{d_1-1}]^t$. Αν $q_1(x) = \sum_{i=0}^{d_2-1} \kappa_i x^i,\ q_2(x) = \sum_{i=0}^{d_1-1} \lambda_i x^i$, σύμφωνα με ιδιότητα των block Toelpitz πινάκων:

$$\underset{\sim}{v}^t S = \underset{\sim}{0} \ \Rightarrow\ q_1(x)p_1(x) + q_2(x)p_2(x) = 0 \ \Rightarrow\ q_1(x)p_1(x) = -q_2(x)p_2(x) \ \Rightarrow\ \deg \text{ΕΚΠ}[p_1, p_2] \leq d_1 + d_2 - 1$$

Γνωρίζουμε όμως ότι $\text{ΜΚΔ}[p_1, p_2]\text{ΕΚΠ}[p_1, p_2] = p_1 p_2 \Rightarrow \deg \text{ΕΚΠ}[p_1, p_2] + \deg \text{ΜΚΔ}[p_1, p_2] = d_1 + d_2$. Τελικά

$$\deg \text{ΜΚΔ}[p_1, p_2] = d_1 + d_2 - \deg \text{ΕΚΠ}[p_1, p_2] \geq d_1 + d_2 - d_1 - d_2 + 1 = 1$$

δηλαδή τα p_1, p_2 έχουν κοινή ρίζα. □

Με χρήση της απαλοίφουσας μπορούμε να κατασκευάσουμε άθροισμα, γινόμενο, πηλίκο κτλ αλγεβρικών αριθμών. Ένας αλγεβρικός αριθμός ορίζεται ως η ρίζα ενός πολυωνύμου σε ένα διάστημα απομόνωσης αυτής.

Παραδείγματος χάριν έστω οι: $a = \{p(x) = 0\ , x \in [t_1, t_2]\},\ b = \{q(x) = 0\ , x \in [r_1, r_2]\}$. Αν θέσουμε $b = a + y$, τότε τα $p(x),\ q(x+y)$ έχουν κοινή ρίζα (αν θεωρηθούν ως πολυώνυμα με μεταβλητή το x) το a, άρα η διαφορά είναι ρίζα της $R(y) \equiv R[p(x), q(x+y)] = 0$. Έτσι τελικά $b - a = \{R(y) = 0,\ y \in [r_1 - t_2, r_2 - t_1]\}$.

Μια άλλη εφαρμογή είναι στην απόδειξη του παρακάτω

Θεώρημα 3.2. *Έστω s η απόσταση δυο οποιονδήποτε ριζών ενός πολυωνύμου $p \in \mathbb{R}[x]$ βαθμού d με συντελεστές μήκους $O_B(\tau)$. Τότε είναι $-\log s = O_B(d\tau)$.*

Απόδειξη. Αν θέσουμε όπως πριν $b = a + y$, με $s = |y|$, θα είναι $R(y) \equiv R[p(x), p(x+y)] = 0$. Έστω $R(y) = c_r y^r + \cdots + c_1 y + c_0$, όπου $r \leq d^2$ από τον ορισμό της απαλοίφουσας ως ορίζουσας διάστασης $2d$. Έτσι $c_i = O_B(d^2 2^{d\tau})$. Από το (αντίστροφο) φράγμα του Cauchy τώρα

$$s \geq \frac{c_r}{1 + \max\limits_{0 \leq i \leq r} |c_i|} \succ \frac{1}{\max\limits_{0 \leq i \leq r} |c_i|}$$

Για $s \leq 1$, και αν c_{i_0} ο μέγιστος κατ' απόλυτο τιμή συντελεστής, παίρνουμε τελικά $-\lg s \geq \lg |c_{i_0}| - \lg 1 \cong 2 \lg d + d\tau = O_B(d\tau)$. □

Η ορίζουσα του S δεν αλλάζει αν προσθέσουμε την j-οστή στήλη της, πολλαπλασιασμένη με x^{j-1} για $j = 2, \dots, d_1 + d_2$, στην πρώτη στήλη. Όμως ο πίνακας έχει αλλάξει καθόσον η πρώτη στήλη περιέχει τα πολυώνυμα $p_1(x), \dots, x^{d_2-1}p_1(x), p_2(x), \dots, x^{d_1-1}p_2(x)$. Με βάση αυτήν την παρατήρηση, ορίζουμε:

Ορισμός 3.8. *Η μηδενική υπο-απαλοίφουσα R_0 είναι η απαλοίφουσα $R(p_1, p_2) = \det S$.*

Για $0 < i \leq \min\{d_1, d_2\}$, η i-οστή υπο-απαλοίφουσα είναι η ορίζουσα R_i.

Για $i = \min\{d_1, d_2\}$, η υπο-απαλοίφουσα R_i είναι η ορίζουσα του πίνακα που περιέχει μόνο συντελεστές του πολυωνύμου με το μεγαλύτερο βαθμό. Η πρώτη γραμμή της ξεκινά με το διάνυσμα των συντελεστών του μεγιστοβάθμιου πολυωνύμου και κατόπιν μηδενικά.

Για $\min\{d_1, d_2\} < i < \max\{d_1, d_2\}$ ορίζουμε $R_i = 0$. Σχηματικά:

$$R_i(p_1,p_2) = \left|\begin{array}{c:cccc:c} p_1(x) & a_{i+1} & \cdots & a_{d_1} & & \mathbf{0} \\ \vdots & \vdots & \ddots & & \ddots & \\ x^{d_2-i-1}p_1(x) & a_{2i-d_2+2} & \cdots & a_{i+1} & \cdots & a_{d_1} \\ p_2(x) & b_{i+1} & \cdots & b_{d_2} & & \\ \vdots & \vdots & \ddots & & \ddots & \mathbf{0} \\ x^{d_1-i-1}p_2(x) & b_{2i-d_1+2} & \cdots & a_{i+1} & \cdots & b_{d_2} \end{array}\right| \quad , \quad i = 0, 1, \dots, \min\{d_1, d_2\}$$

όπου ορίζουμε $a_k,\ b_k = 0$ για $k < 0$.

Η ορίζουσα R_i προκύπτει από εκείνη του μετασχηματισμένου πίνακα S, αν αφαιρέσουμε τις τελευταίες i γραμμές που περιέχουν συντελεστές του p_1, τις τελευταίες i γραμμές (που περιέχουν συντελεστές του p_2), τις τελευταίες i στήλες (οι οποίες πλέον περιέχουν μόνο μηδενικά) και, τέλος, τις i αριστερότερες στήλες που βρίσκονται δεξιά της πρώτης. Για $i = 0$ δεν αφαιρούμε γραμμές ή στήλες. Η διάσταση της ορίζουσας είναι προφανώς $(d_1 + d_2 - 2i) \times (d_1 + d_2 - 2i)$.

Χάριν ευκολίας στην παρουσίαση επιτρέψαμε αρνητικούς δείκτες, ερμηνεύοντας τα αντίστοιχα a_k ως μηδενικά. Εξαρτάται από το τον αριθμό των γραμμών/στηλών που θα αφαιρέσουμε αν θα υπάρχουν ή όχι μηδενικά στο κάτω αριστερά κομμάτι κάθε υποπίνακα του R_i. Παρατηρήστε επίσης ότι για $\kappa < \lambda$, η R_κ είναι υποορίζουσα της R_λ. Έτσι όλες οι υπο-απαλοίφουσες προκύπτουν από την αρχική R_0(σε κουτιά οι γραμμές της i-οστής υπο-απαλοίφουσας):

$$R_0(p_1,p_2) = \left|\begin{array}{c:ccccccc:c} p_1(x) & a_1 & \cdots & a_{i+1} & \cdots & a_{d_1} & & & \\ \vdots & a_0 & & \ddots & & \ddots & & & \mathbf{0} \\ \vdots & \mathbf{0} & \ddots & & \ddots & & \ddots & & \\ x^{d_2-i-1}p_1(x) & & & a_0 & \cdots & a_{i+1} & \cdots & a_{d_1} & \\ p_2(x) & b_1 & \cdots & b_{i+1} & \cdots & b_{d_1} & & & \\ \vdots & b_0 & & \ddots & & \ddots & & & \mathbf{0} \\ \vdots & & \ddots & & \ddots & & \ddots & & \\ x^{d_1-i-1}p_2(x) & \mathbf{0} & & b_0 & \cdots & b_{i+1} & \cdots & b_{d_2} & \end{array}\right|$$

Αν αναπτύξουμε την R_0 ως προς την πρώτη στήλη, συνάγεται η εξής ιδιότητα:

$$R_0(p_1,p_2) \equiv R(p_1,p_2) = p_1(x)t_0(x) + p_2(x)s_0(x), \text{ όπου } \deg t_0 = d_2 - 1,\ \deg s_0 = d_1 - 1$$

Όπως ήδη γνωρίζουμε $\deg R_0(x) = 0$. Αν $i = d_2 < d_1$, ο αντίστοιχος πίνακας είναι κάτω τριγωνικός και $R_{d_2}(x) = p_2(x)a_{d_2}^{d_1-d_2-1}$, άρα $\deg R_{d_2}(x) = d_2$. Γενικότερα ισχύει το παρακάτω

Θεώρημα 3.3. *Ο βαθμός της υπο-απαλοίφουσας είναι $\deg R_i(x) \leq i$ για $i = 0, \dots, \min\{d_1, d_2\}$.*

Ενδέχεται βέβαια κάποιες υπο-απαλοίφουσες να έχουν τον ίδιο βαθμό ή ο βαθμός δυο διαδοχικών υπο-απαλοιφουσών να διαφέρει περισσότερο από 1.

Η βασική ιδιότητα των υπο-απαλοιφουσών είναι η παρακάτω ισοδυναμία:

Θεώρημα 3.4. *Εφόσον τα $p_1, p_2 \in K[x]$ για K μια περιοχή μοναδικής παραγοντοποίησης με μονάδα:*

$$\text{τα } p_1, p_2 \text{ έχουν έναν κοινό διαιρέτη βαθμού } k \iff R_i = 0,\ \forall i < k \text{ και } R_k \neq 0$$

3.4 Διακρίνουσα πολυωνύμου

Ορισμός 3.9. *Έστω D ακέραια περιοχή και $p(x) = \sum_{i=0}^{d} c_i x^i$ με συντελεστές $c_i \in D, c_d \neq 0$ και $r_i \in \overline{D}$(αλγεβρική θήκη του D),$i = 1, \ldots, d$ οι ρίζες του p. Ορίζεται η διακρίνουσα:*

$$\Delta := c_d^{2d-2} \prod_{i<j} (r_i - r_j)^2 \tag{3.2}$$

Λήμμα 3.1. *$\Delta = 0$ τότε και μόνο τότε αν υπάρχει πολλαπλή ρίζα του p.*

Απόδειξη. $\Delta = 0 \iff \exists i \neq j : \ r_j = r_i \iff$ το r_i έχει πολλαπλότητα τουλάχιστον 2. □

Λήμμα 3.2. *Η διακρίνουσα συνδέεται με την απαλοίφουσα με τον τύπο $c_d\Delta = (-1)^{\binom{d}{2}} R(p, p')$.*

Απόδειξη. Έστω $p(x) = c_d \prod_{i=1}^{d} (x - r_i)$. Τότε παραγωγίζοντας με χρήση του κανόνα Leibnitz

$$p'(x) = c_d \left[\prod_{i=1}^{d} (x - r_i) \right]' = c_d \sum_{j=1}^{d} \prod_{i \neq j} (x - r_i) \quad \Rightarrow \quad p'(r_j) = c_d \prod_{i \neq j} (r_j - r_i)$$

Από τη μορφή Poisson (3.1) έχουμε:

$$R(p, p') = c_d^{d-1} \prod_{i=1}^{d} p'(r_i)$$

αντικαθιστώντας

$$R(p, p') = c_d^{d-1} \prod_{i=1}^{d} c_d \prod_{i \neq j} (r_j - r_i) = c_d^{2d-1} \prod_{i<j} \left[(-1)(r_i - r_j)^2 \right] = (-1)^{\binom{d}{2}} c_d^{2d-1} \prod_{i<j} (r_i - r_j)^2 = (-1)^{\binom{d}{2}} c_d \Delta$$

τελικά, $c_d\Delta = (-1)^{\binom{d}{2}} R(p, p')$. □

Λήμμα 3.3. *Η διακρίνουσα ανήκει στην ευκλείδεια περιοχή D.*

Απόδειξη. Με χρήση ιδιοτήτας της ορίζουσας, βγάζουμε κοινό παράγοντα το c_d από την τελευταία στήλη:

$$R(p, p') = \begin{vmatrix} c_0 & c_1 & \cdots & c_d & & & \\ & c_0 & c_1 & \cdots & c_d & & \mathbf{0} \\ & & \ddots & \ddots & \ddots & \ddots & \\ \mathbf{0} & & & c_0 & c_1 & \cdots & c_d \\ c_1 & 2c_2 & \cdots & dc_d & & & \\ & c_1 & 2c_2 & \cdots & dc_d & & \mathbf{0} \\ & & \ddots & \ddots & \ddots & \ddots & \\ \mathbf{0} & & & c_1 & 2c_2 & \cdots & dc_d \end{vmatrix} = \left| \begin{array}{ccccccc:c} c_0 & c_1 & \cdots & c_d & & & \mathbf{0} & \\ & c_0 & c_1 & \cdots & c_d & & & \mathbf{0} \\ & & \ddots & \ddots & \ddots & \ddots & & \\ \mathbf{0} & & & c_0 & c_1 & \cdots & & 1 \\ c_1 & 2c_2 & \cdots & dc_d & & & \mathbf{0} & \\ & c_1 & 2c_2 & \cdots & dc_d & & & \mathbf{0} \\ & & \ddots & \ddots & \ddots & \ddots & & \\ \mathbf{0} & & & c_1 & 2c_2 & \cdots & & d \end{array} \right| c_d$$

άρα $c_d \big| R(p, p') \Rightarrow \dfrac{R(p, p')}{c_d} \in D \Rightarrow (-1)^{\binom{d}{2}} \dfrac{R(p, p')}{c_d} \in D \Rightarrow \Delta \in D.$ □

Το επόμενο λήμμα συνδέει τη διακρίνουσα με μια ορίζουσα Vandermonde των ριζών του πολυωνύμου.

Λήμμα 3.4. *Το γινόμενο που εμφανίζεται στη διακρίνουσα μπορεί να εκφραστεί σαν το τετράγωνο μιας ορίζουσας Vandermonde, συγκεκριμένα $\prod_{i<j} (r_i - r_j) = \pm V(r_1, \ldots, r_d)$, όπου*

$$V(r_1, \ldots, r_d) := \begin{vmatrix} 1 & 1 & \cdots & 1 \\ r_1 & r_2 & \cdots & r_d \\ r_1^2 & r_2^2 & \cdots & r_d^2 \\ \vdots & \vdots & & \vdots \\ r_1^{d-1} & r_2^{d-1} & \cdots & r_d^{d-1} \end{vmatrix}$$

Απόδειξη. Θα δείξουμε πως τα δυο μέλη έχουν τον ίδιο βαθμό σαν πολυώνυμα των r_i και ότι διαιρούν το ένα το άλλο. Είναι

$$\deg\prod_{i<j}(r_i-r_j)=\binom{d}{2}=\frac{d!}{2!(d-2)!}=\frac{d(d-1)}{2}\quad,\quad \deg V(r_1,\dots,r_d)=(d-1)+(d-2)+\cdots+2+1=\frac{d(d-1)}{2}$$

Αν $\prod_{i<j}(r_i-r_j)=0$, τότε $r_i=r_j,\ i\neq j$. Η ορίζουσα Vandermonde έχει δυο στήλες ίδιες, δηλαδή γραμμικά εξαρτημένες άρα $V(r_1,\dots,r_d)=0\Rightarrow\prod_{i<j}(r_i-r_j)\big|V(r_1,\dots,r_d)$.

Αντίστροφα, αν $V(r_1,\dots,r_d)=0$ αναπτύσσοντας την ορίζουσα βλέπουμε ότι
$\prod_{i<j}(r_i-r_j)|V(r_1,\dots,r_d)\Rightarrow V(r_1,\dots,r_d)\big|\prod_{i<j}(r_i-r_j)$. □

Το αποτέλεσμα αυτό, σε συνδυασμό με το φράγμα Hadamard που ακολουθεί, χρησιμοποιούνται στην απόδειξη του Θεωρήματος Davenport-Mahler-Mignotte (2.3). Για την απόδειξη χρησιμοποιούμε το παρακάτω φράγμα στην ορίζουσα Vandermonde. Κατόπιν με γραμμοπράξεις σχηματίζουμε τις διαφορές και τα γινόμενα που θέλουμε:

Λήμμα 3.5. *(Φράγμα Hadamard) Αν* $A=[\underline{u}_1\ \cdots\ \underline{u}_n]=[\underline{w}_1\ \cdots\ \underline{w}_n]^T$ *ισχύει* $|\det A|\le\prod_{i=1}^{n}\|\underline{u}_i\|_2=\prod_{i=1}^{n}\|\underline{w}_i\|_2$.

Η ισότητα επιτυγχάνεται όταν οι στήλες(ή γραμμές) είναι ορθογώνιες, πχ $|\begin{vmatrix}1 & c\\ c & -1\end{vmatrix}|=1+c^2=\sqrt{1+c^2}\sqrt{c^2+1}$.

3.5 Ακολουθία Sturm-Habicht(Subresultant Sequence)

Ας δούμε τώρα πως συνδέεται η ακολουθία υπο-απαλοιφουσών με τις ακολουθίες Sturm. Υπενθυμίζουμε σύντομα την ψευδοδιαίρεση: Στον Ευκλείδειο αλγόριθμο η αύξηση του μεγέθους των (ενδιάμεσων) συντελεστών είναι εκθετική. Για το λόγο αυτό έχουν μελετηθεί μέθοδοι που γενικεύουν τη βασική σχέση, βασισμένες στην ψευδο-διαίρεση:

Ορισμός 3.10. *Στην ψευδο-διαίρεση* $\alpha p(x)=q(x)s(x)+r(x)$, *όπου* $p(x),s(x)\in\mathbb{Z}[x]$ *με* $\deg(p(x))>\deg(r(x))$, *έχουμε* $\alpha=c_d^{\delta}\in\mathbb{Z}$, *όπου* $\delta=\deg(p(x))-\deg(s(x))+1$ *και* c_d *ο μεγιστοβάθμιος συντελεστής του* p. *Έτσι το ψευδο-πηλίκο* $q(x)\in\mathbb{Z}[x]$, *συνεπώς και το ψευδο-υπόλοιπο* $r(x)$ *ανήκουν στο* $\mathbb{Z}[x]$. *Τα* $q(x),r(x)$ *είναι μοναδικά.*

Έστω $p_0(x)=a(x),p_1(x)=b(x)$ και για $i\ge 2: \alpha_i p_{i-2}(x)=p_{i-1}(x)q_i(x)+\beta_i p_i(x)$ όπου α_i και β_i σταθερές. Μερικές περιπτώσεις είναι:

- $\alpha_i=\beta_i=1$ στον *Ευκλείδειο* αλγόριθμο: δίνει το ελάχιστο β_i αλλά το μέγιστο μέγεθος συντελεστών, δηλ. εκθετικό στην χειρότερη περίπτωση [Zip93].
- $\beta_i p_i(x)=$ ψευδο-υπόλοιπο στη διαίρεση που έγινε στο προηγούμενο βήμα, όπου το β_i είναι ο ΜΚΔ των συντελεστών του ψευδο-υπολοίπου (άρα υπολογίζεται ως ένα ΜΚΔ ακεραίων), δηλ. το $p_i(x)$ είναι ένα *πρωτογενές (primitive)* πολυώνυμο: μέγιστο β_i, ελάχιστο μέγεθος πολυωνυμικών συντελεστών, αλλά υψηλό υπολογιστικό κόστος. Αυτός ο αλγόριθμος εφαρμόζεται επαγωγικά και με πολλές μεταβλητές $\mathbb{Z}[x_1,\dots,x_n]$.
- Το β_i δίνεται σε συνάρτηση των α_j,β_j για $j<i$ ενώ τα πολυώνυμα δίνονται από κάποια υπο-απαλοίφουσα(subresultant). Συγκεκριμένα, $\alpha_i=c^{d_{i-2}-d_{i-1}+1}$, όπου c ο μεγιστοβάθμιος συντελεστής του $p_{i-1}(x)$ και $\deg p_i=d_i$. Η θεωρία των Habicht, Collins, Brown αποδεικνύει πως το β_i διαιρεί το ψευδο-υπόλοιπο των $p_{i-2}(x),p_{i-1}(x)$. Επιτυγχάνονται ενδιάμεσες τιμές β_i και ενδιάμεσο μέγεθος συντελεστών των p_i σε σχέση με τις άλλες μεθόδους. Συγκεκριμένα, οι συντελεστές των p_i έχουν μήκος $O_B^*(d\tau)$, όπου τ το μέγεθος των συντελεστών στα δεδομένα πολυώνυμα. Η μέθοδος αυτή πετυχαίνει βέλτιστη δυαδική πολυπλοκότητα $O_B^*(d^3\tau)$ για τον υπολογισμό *ολόκληρης* της ακολουθίας [Lombardi-Roy-ElDin, Reischert].

Το παρακάτω θεώρημα συνδέει τις υπο-απαλοίουσες με την ακολουθία Sturm-Habicht:

Θεώρημα 3.5. *Έστω* $p+1,p_2\in\mathbb{Z}[x]$. *Το σύνολο των υπο-απαλοιφουσών προκύπτει σαν ένα σύνολο μιας ακολουθίας ψευδο-υπολοίπων, δηλαδή*

$$\{R_0(x),R_1(x),\dots\}=\{p_i(x)\ :\ b_ip_{i-1}(x)=q_i(x)p_i(x)+c_i^{\delta_i+1}p_{i+1}(x)\}$$

με $b_i,c_i\in D$, c_i *ο μεγιστοβάθμιος όρος του* p_i *και* $\delta_i=\deg q_i(x)$.