3.2 Sukzessive Minima und reduzierte Basen: Beweise

In diesem Abschnitt benutzen wir die Beschreibung eines Gitters (bis auf Isomorphie) durch eine (jede) seiner Gram-Matrizen sowie durch die zugehörige quadratische Form (bzw. ihre Äquivalenzklasse); wir erinnern an Abschnitt 1.1 für die verschiedenen Varianten des Begriffs „quadratische Form" sowie die zugehörigen Schreib- und Sprechweisen, z.B. $F[\mathbf{x}]$.

Proposition 3.2.1 (Dreieckszerlegung erster Art)

a) Es sei $F = (f_{ij})_{i,j=1,\ldots,n}$ eine positiv definite quadratische Form. Dann existiert eine eindeutige Zerlegung der Gestalt

$$\begin{aligned} F[\mathbf{x}] &= h_1(x_1,\ldots,x_n)^2 + h_2(x_2,\ldots,x_n)^2 + \ldots + h_{n-1}(x_{n-1},x_n)^2 + h_n(x_n)^2 \\ &= \textstyle\sum_{i=1}^n h_i(\mathbf{x})^2 \end{aligned}$$

wobei die h_i, $i = 1,\ldots,n$ Linearformen auf $\mathbb{R}^n$ sind, d. h. $h_i(\mathbf{x}) = \sum\limits_{j=1}^{n} h_{ij}x_j$, $(h_{ij} \in \mathbb{R})$, und ferner gilt $h_{ij} = 0$, falls $j < i$.

b) Unter den Voraussetzungen von a) gilt

i) $H^tH = F$, wobei $H = (h_{ij})$ eine obere Dreiecksatrix ist.

ii) $\det F = (h_{11} \cdot h_{22} \cdot \ldots \cdot h_{nn})^2$

iii) $h_{ii}^2 \leq f_{ii}$ für $i = 1,\ldots,n$.

Beweis: a) Wir zeigen die Existenz und Eindeutigkeit der Funktionen h_i bzw. ihrer Koeffizienten h_{ij} durch Induktion über n. Es gilt im Ansatz

$$\begin{aligned} F[\mathbf{x}] &= f_{11}x_1^2 + 2f_{12}x_1x_2 + \ldots + 2f_{1n}x_1x_n + (\ldots)\,, \\ h_1(\mathbf{x})^2 &= (h_{11}x_1 + h_{12}x_2 + \ldots + h_{1n}x_n)^2 \\ &= h_{11}^2x_1^2 + 2h_{11}h_{12}x_1x_2 + \ldots + 2h_{11}h_{1n}x_1x_n + (\ldots) \end{aligned}$$

wobei die Terme $(\cdots)$ kein x_1 mehr enthalten. Setze also

$$h_{11} = \sqrt{f_{11}},\ h_{1j} = f_{1j}/h_{11},\ i = 2,\ldots,n\,. \tag{3.2.1}$$

(Man beachte, dass $f_{11} > 0$ ist, denn die Matrix F ist positiv definit.) Unter der Voraussetzung $h_{i,1} = 0$, $i = 2,\ldots,r$ ist das die einzige Lösung für $h_{11},\ldots,h_{1n}$.

Nun ist die Funktion $F[\mathbf{x}] - h_1(\mathbf{x})^2$ unabhängig von x_1 und wieder eine quadratische Form. D. h. es gibt eine symmetrische Matrix G der Größe $n-1$ mit

$$F[\mathbf{x}] - h_1(\mathbf{x})^2 = G[\mathbf{x}'],\ \text{wobei } \mathbf{x}' = (x_2,\ldots,x_n)^t\,. \tag{3.2.2}$$

Wir behaupten, dass G wieder positiv definit ist. Sei dazu $\mathbf{x}' = (x_2, \ldots, x_n)^t$ vorgegeben. Setze $x_1 := -(h_{12}x_2 + \ldots + h_{1n}x_n)/h_{11}$, $\mathbf{x} = (x_1, \ldots, x_n)^t$. Dann ist $h_1(\mathbf{x}) = 0$, $G[\mathbf{x}'] = F[\mathbf{x}] \geq 0$, sogar > 0, falls $(x_2, \ldots, x_n) \neq (0, \ldots, 0)$.

Wir wiederholen nun diese Überlegung, bzw. wenden die Induktionsannahme auf G an, um die Koeffizienten h_{ij} für $2 \leq i \leq j \leq n$ zu erhalten.

b) Der Vektor $(h_1(\mathbf{x}), h_2(\mathbf{x}), \ldots, h_n(\mathbf{x}))^t$ kann als $H\mathbf{x}$ geschrieben werden, und folglich ist $\mathbf{x}^t F \mathbf{x} = F[\mathbf{x}] = (H\mathbf{x})^t H\mathbf{x}$ für alle $\mathbf{x} \in \mathbb{R}^n$, nach a), woraus sich sofort i) ergibt und damit auch ii). Die Abschätzung iii) folgt aus ii), denn $f_{ii} = \sum_{k=1}^n h_{ki}^2 \geq h_{ii}^2$. □

Es folgt nun eine Variante der Dreieckszerlegung, die für quadratische Formen über $\mathbb{Q}$ und Gitter oft etwas vorteilhafter ist.

Proposition 3.2.2 (Dreieckszerlegung zweiter Art)

a) Es sei $F = (f_{ij})_{i,j=1,\ldots,n}$ eine positiv definite quadratische Form. Dann existiert eine eindeutige Zerlegung der Gestalt

$$F[\mathbf{x}] = \sum_{i=1}^{n} d_i \, \bar{h}_i(x_i, \ldots, x_n)^2, \text{ mit } d_i, \bar{h}_{ij} \in \mathbb{R}, \ \bar{h}_{ij} = 0, \text{ falls } j < i \text{ und } \bar{h}_{ii} = 1\,.$$

Wenn alle $f_{ij} \in \mathbb{Q}$ sind, dann sind auch d_i, $\bar{h}_{ij} \in \mathbb{Q}$.

b) Unter den Voraussetzungen von a) gilt

i) $\bar{H}^t D \bar{H} = F$, wobei D die Diagonalmatrix aus $d_1, \ldots, d_n$ ist.

ii) $\det F = d_1 \cdot d_2 \cdot \ldots \cdot d_n$

iii) $d_i \leq f_{ii}$ für $i = 1, \ldots, n$.

c) Für h_{ii} wie in 3.2.1 gilt $d_i = h_{ii}^2$ für $i = 1, \ldots, n$.

Beweis: a) In leichter Modifikation des Beweises von 3.2.2 setze anstelle von (3.2.1)

$$d_1 = f_{11}, \ \bar{h}_{1j} = \frac{f_{1j}}{d_1} \text{ für } j > 1. \tag{3.2.3}$$

Die Induktion läuft wie oben (mit dem gleichen G). Wenn alle $f_{ij} \in \mathbb{Q}$ sind, gilt per Induktion auch $d_2, \ldots, d_n \in \mathbb{Q}$ und $\bar{h}_{ij} \in \mathbb{Q}$ für alle $i < j$.

Der Beweis von Teil b) läuft analog zum vorigen Satz 3.2.1.

Teil c) folgt aus der Eindeutigkeitsaussage in 3.2.1: offenbar muss sogar $h_{ij} = \sqrt{d_i} \cdot \bar{h}_{ij}$ für alle i, j gelten. □

Beispiel ($n = 2$; siehe auch Beispiel 1.3.6, Ungleichungen (1.3.3))

$$\begin{aligned} & ax^2 + 2bxy + cy^2 \\ = & \left(\sqrt{a}x + \tfrac{b}{\sqrt{a}}y\right)^2 + \left(\sqrt{c - \tfrac{b^2}{a}}y\right)^2 \\ = & \; a\left(x + \tfrac{b}{a}y\right)^2 + \left(c - \tfrac{b^2}{a}\right)y^2 \end{aligned}$$

Korollar 3.2.3 (Hadamard-Ungleichung) Für jede symmetrische, positiv definite Matrix $F = (f_{ij}) \in \mathbb{R}^{n\times n}$ gilt

$$\det F \leq \prod_{i=1}^{n} f_{ii}.$$

BEWEIS: Nach Proposition 3.2.2 b) ist $\det F = \det D = \prod d_i \leq \prod f_{ii}$. □

Folgendes Lemma wird für den Beweis von Satz 3.1.4 benutzt:

Lemma 3.2.4 Sei F eine positiv definite quadratische Form (symmetrische Matrix) und

$$F[\mathbf{x}] = \sum_{i=1}^{n} d_i \, \bar{h}_i(x_i, x_{i+1}, \ldots, x_n)^2$$

ihre Dreiecks-Zerlegung zweiter Art. Dann existiert zu jedem $\mathbf{x} \in \mathbb{R}^n$ ein $\mathbf{y} \in \mathbb{Z}^n$ mit

$$F[\mathbf{x} - \mathbf{y}] \leq \frac{1}{4} \sum_{i=1}^{n} d_i \,.$$

BEWEIS: Wir überlegen uns, dass wir $y_n, y_{n-1}, \ldots, y_k, \ldots, y_1$ in dieser Reihenfolge so wählen können, dass für alle $k = n, \ldots, 1$ gilt

$$|\bar{h}_k(\mathbf{x} - \mathbf{y})| = |\bar{h}_k(x_k - y_k, \ldots, x_n - y_n)| \leq \frac{1}{2}. \tag{3.2.4}$$

Es ist

$$\begin{aligned} \bar{h}_n(\mathbf{x} - \mathbf{y}) &= x_n - y_n \\ \bar{h}_k(\mathbf{x} - \mathbf{y}) &= (x_k - y_k) + \textstyle\sum_{j=k+1}^{n} \bar{h}_{kj}(x_j - y_j) \\ &= \left(x_k + \textstyle\sum_{j=k+1}^{n} \bar{h}_{kj}(x_j - y_j)\right) - y_k, \end{aligned}$$

woraus sich die Behauptung unmittelbar ergibt. □

Anmerkung: Das letzte Lemma bzw. sein Beweis hat den Charakter eines n-dimensionalen Rundungsverfahrens: zu einem beliebigen Vektor $\mathbf{x}$ wird ein ganzzahliger Vektor $\mathbf{y}$ bestimmt, der „nahe" (gemessen durch F) bei $\mathbf{x}$ liegt. Es wird allerdings nicht behauptet, dass dieses $\mathbf{y}$ bereits bestmöglich ist. Idelaerweise wäre

$$F[\mathbf{x}-\mathbf{y}] = \min_{\mathbf{z}\in\mathbb{Z}^n} F[\mathbf{x}-\mathbf{z}]. \tag{3.2.5}$$

Ein solches $\mathbf{y}$ existiert, da es wegen der Diskretheit von $\mathbb{Z}^n$ nur endlich viele Kandidaten gibt. Auf die Bestimmung von $\mathbf{y}$ kommen wir später im Kontext der *Dirichlet-Voronoi-Zellen* von Gittern noch kurz zurück. Siehe auch Übungsaufgabe 3.2.1.

3.2.5 Beweis von Satz 3.1.4:

Wähle linear unabhängige Vektoren $w_1, \ldots, w_n \in L$ mit

$$||w_i||^2 = \mu_i := \min L, \quad \text{für } i = 1, \ldots, n. \tag{3.2.6}$$

Sei $F = (b(w_i, w_j))$ die Gram-Matrix der w_i. (Wir bezeichnen also das zugrundeliegende Skalarprodukt mit b). Es gilt also

$$||v||^2 = b(v,v) = F[\mathbf{x}], \text{ wobei } v = \sum x_i w_i, \ x_i \in \mathbb{R}. \tag{3.2.7}$$

Es ist $\det F = m^2 \cdot \det L = m^2 \cdot \det(L, b)$, wobei m der Index des von $w_1, \ldots, w_n$ erzeugten Teilgitters von L ist. Wir benutzen nun die Dreieckszerlegung zweiter Art von

$$F[\mathbf{x}] = \sum_{i=1}^{n} d_i \, h_i(x_i, \ldots, x_n)^2. \tag{3.2.8}$$

Wir definieren hiermit eine neue quadratische Form $\widehat{F}$ durch

$$\widehat{F}[\mathbf{x}] = \sum_{i=1}^{n} \frac{d_i}{\mu_i} \, h_i(x_i, \ldots, x_n)^2. \tag{3.2.9}$$

Weiter definieren wir ein neues Skalarprodukt $\widehat{b}$ durch

$$\widehat{b}(v,v) = \widehat{F}[\mathbf{x}], \text{ wobei } v = \sum x_i w_i, \ x_i \in \mathbb{R}. \tag{3.2.10}$$

Dann gilt

$$\det \widehat{F} = \frac{1}{\mu_1 \cdot \cdots \cdot \mu_n} \det F, \tag{3.2.11}$$

und folglich auch

$$\det(L, \widehat{b}) = \frac{1}{\mu_1 \cdot \cdots \cdot \mu_n} \det(L, b). \tag{3.2.12}$$

Man überlegt sich nun als entscheidende Hilfsaussage:

$$\min(L, \widehat{b}) \geq 1, \text{ das heißt, } \widehat{b}(v,v) \geq 1 \text{ für alle } v \in L \smallsetminus 0. \tag{3.2.13}$$

Es sei hierzu $v = \sum x_i w_i \in L \smallsetminus 0$, und k sei der größte Index mit $h_k(\mathbf{x}) \neq 0$. Dann sind $w_1, \ldots, w_{k-1}, v$ linear unabhängig, denn sonst wäre v Linearkombination von $w_1, \ldots, w_{k-1}$, also $x_k = x_{k+1} = \cdots = x_n = 0$ und folglich auch $h_k(\mathbf{x}) = h_k(x_k, \ldots, x_n) = 0$. Nach Definition des k-ten sukzessiven Minimums ist $b(v,v) \geq \mu_k$, es folgt

$$\widehat{b}(v,v) = \sum_{i=1}^{k} \frac{d_i}{\mu_i} h_i(\mathbf{x})^2 \geq \frac{1}{\mu_k} \sum_{i=1}^{k} d_i h_i(\mathbf{x})^2 = \frac{1}{\mu_k} b(v,v) \geq 1.$$

Nun folgt aus (3.2.13) mit der Definition der Hermite-Konstanten

$$1 = 1^n \leq \min(L, \widehat{b})^n \leq \gamma_n^n \det(L, \widehat{b}) \leq \frac{\gamma_n^n \det(L,b)}{\mu_1 \cdot \cdots \cdot \mu_n},$$

also

$$\mu_1 \cdot \cdots \cdot \mu_n \leq \gamma_n^n \det(L,b),$$

wie behauptet. □

3.2.6 **Beweis** von Theorem 3.1.10

Setze wieder kurz $\mu_i := \min_i L,\ i = 1, \ldots, n$. Für $k = 1$ ist, wie bereits bemerkt, $\|v_1\|^2 = \mu_1$. Sei also $k > 1$. Es sei $L' := \mathbb{Z}v_1 \oplus \cdots \oplus \mathbb{Z}v_{k-1},\ V' := \mathbb{R}v_1 \oplus \cdots \oplus \mathbb{R}v_{k-1}$. Es gibt einen Vektor $w \in L$ mit $\|w\|^2 \leq \mu_k$ und $w \notin V'$. Die Vektoren $v_1, \ldots, v_{k-1}$ können zu einer Basis

$$v_1, \ldots, v_{k-1}, v \quad \text{von} \quad L \cap (V' \oplus \mathbb{R}w) \tag{3.2.14}$$

ergänzt werden. Da eine solche Basis auch zu einer Basis von L ergänzt werden kann (denn $L \cap (V' \oplus \mathbb{R}w)$ ist direkter Summand in L), gilt nach Definition der Minkowski-Reduktion

$$\|v_k\| \leq \|v\|. \tag{3.2.15}$$

All das bleibt richtig, wenn wir v um einen beliebigen Vektor aus L' abändern; hierauf kommen wir unten zurück. Nun stellen wir den Vektor w in der Basis dar, also

$$w = w' + mv,\ w' \in L',\ m \in \mathbb{Z}. \tag{3.2.16}$$

Wenn $|m| = 1$ ist, ist $v_1, \ldots, v_{k-1}, w$ ebenfalls eine Basis von $L \cap (V' \oplus \mathbb{R}w)$, nach (3.2.15) gilt also

$$\|v_k\|^2 \leq \|w\|^2 \leq \mu_k \leq \delta_k \mu_k. \tag{3.2.17}$$

Sei nun $|m| \geq 2$. Wir schreiben dann

$$v = v' + v^\perp,\ v' \in V',\ v^\perp \in V'^\perp. \tag{3.2.18}$$

Einsetzen in (3.2.16) liefert

$$w = (w' + mv') + mv^\perp,\ w' + mv' \in V',\ mv^\perp \in V'^\perp. \tag{3.2.19}$$

Es ist $\|w\|^2 \geq m^2\|v^\perp\|^2$, also

$$\|v^\perp\|^2 \leq \frac{\|w\|^2}{4} \leq \frac{\mu_k}{4}. \tag{3.2.20}$$

Wir möchten nun $\|v\|$ wenn möglich noch verkleinern, und zwar durch Anwendung von Lemma 3.2.4 auf L', genauer auf die Gram-Matrix F' von $v_1, \ldots, v_{k-1}$. Ersetze also v durch $v - u = (v' - u) + v^\perp$ mit $u \in L'$ so, dass

$$\|v' - u\|^2 \leq \tfrac{1}{4}(d_1 + \ldots d_{k-1}), \tag{3.2.21}$$

wobei die d_i aus der Dreieckszerlegung zweiter Art von F' kommen, also $d_i \leq f_{ii} = \|v_i\|^2$ gilt. Es gilt also nach (3.2.15), angewendet auf $v - u$, dass

$$\|v_k\|^2 \leq \|v - u\|^2 \leq \tfrac{1}{4}(\|v_1\|^2 + \cdots + \|v_{k-1}\|^2) + \tfrac{1}{4}\mu_k. \tag{3.2.22}$$

Wenn wir nun die Behauptung des Satzes für $v_1, v_2, \ldots, v_{k-1}$ als bewiesen annehmen und einsetzen, erhalten wir

$$\begin{aligned} \|v_k\|^2 \leq \|v - u\|^2 &\leq \tfrac{1}{4}(\delta_1\mu_1 + \cdots + \delta_{k-1}\mu_{k-1} + \mu_k) \\ &\leq \tfrac{1}{4}(\delta_1 + \cdots + \delta_{k-1} + 1)\mu_k \\ &\leq \delta_k\mu_k. \end{aligned}$$

Die letzte Abschätzung ergibt sich ohne Mühe aus der expliziten Definition $\delta_i = \left(\frac{5}{4}\right)^{i-4}$. (Hier wird gegenüber dem genauen Wert, der sich rekursiv aus dem vorgestellten Beweis ergibt, noch ein wenig verschenkt.) □