Wie wir bereits im vorigen Paragraphen gesehen haben, hat ein Gitter jedoch im allgemeinen keine Orthogonalbasis; wir müssen wir uns daher mit weniger zufrieden geben. Trotzdem wollen wir eine Basis, die sich zumindest nicht allzu sehr von einer Orthogonalbasis unterscheidet. Letzteres können wir auch so formulieren, daß sich die Basis nicht zu sehr von ihrer GRAM-SCHMIDT-Orthogonalisierung unterscheiden soll, denn wenn wir dieses Verfahren auf eine Orthogonalbasis anwenden, ändert sich ja nichts.

Die nach GRAM-SCHMIDT konstruierten Vektoren der Orthogonalbasis sind

$$\vec{c}_i = \vec{b}_i - \sum_{j=1}^{i-1} \mu_{ij}\vec{c}_j \quad \text{mit} \quad \mu_{ij} = \frac{\vec{b}_i \cdot \vec{c}_j}{\vec{c}_j \cdot \vec{c}_j} ,$$

wobei die $\mu_{ij}$ aber im allgemeinen keine ganzen, sondern nur rationale Zahlen sind.

Wenn der Vektor $\vec{c}_i$ nicht im Gitter liegt, können wir wenigstens versuchen, ihn durch einen möglichst ähnlichen Gittervektor zu ersetzen, um so näher an das orthogonale Komplement zu kommen. Wir können beispielsweise alle Zahlen $\mu_{ij}$ ersetzen durch die jeweils nächstgelegene ganze Zahl (oder eine der beiden, falls $\mu_{ij}$ die Hälfte einer ungeraden Zahl sein sollte).

Wir können daher eine Basis findet, für die alle Koeffizienten $\mu_{ij}$ bei der GRAM-SCHMIDT-Orthogonalisierung höchstens den Betrag $\frac{1}{2}$ haben. LENSTRA, LENSTRA und LOVÁSZ stellen noch eine weitere Bedingung:

$$\left|\vec{c}_i + \mu_{i\,i-1}\vec{c}_{i-1}\right|^2 \geq \tfrac{3}{4}\left|\vec{c}_{i-1}\right|^2 \quad \text{für alle } i > 1 .$$

Um diese sogenannte LOVÁSZ-Bedingung zu verstehen, multiplizieren wir beiden Seiten aus. Wegen der Orthogonalität der $\vec{c}_j$ ist

$$\left|\vec{c}_i\right|^2 + \mu_{i\,i-1}^2\left|\vec{c}_{i-1}\right|^2 \geq \tfrac{3}{4}\left|\vec{c}_{i-1}\right|^2 \quad \text{oder} \quad \left|\vec{c}_i\right|^2 \geq \left(\tfrac{3}{4} - \mu_{i\,i-1}^2\right)\left|\vec{c}_{i-1}\right|^2 .$$

Da die Beträge der $\mu_{ij}$ höchstens $\frac{1}{2}$ sind, folgt insbesondere

$$\left|\vec{c}_i\right|^2 \geq \tfrac{1}{2}\left|\vec{c}_{i-1}\right|^2 \quad \text{oder} \quad \left|\vec{c}_{i-1}\right|^2 \leq 2\left|\vec{c}_i\right|^2 .$$

Diese Bedingung sorgt also dafür, daß sich die Längen der $\vec{c}_i$ nicht zu stark unterscheiden.



Man könnte sich fragen, warum hier ausgerechnet die Konstante $\frac{3}{4}$ verwendet wird. In der Tat funktioniert die folgende Konstruktion auch, wenn $\frac{3}{4}$ durch irgendeine Konstante $\alpha$ mit $\frac{1}{4} < \alpha < 1$ ersetzt wird. Die Zwei in der gerade bewiesenen Ungleichung wird dann zu $4/(4\alpha - 1)$, d.h. für $\alpha$ knapp unter eins können wir sie auf eine Zahl knapp über $4/3$ herunterdrücken, während $\alpha$-Werte nahe $\frac{1}{4}$ zu sehr schwachen Schranken führen. Starke Schranken sind zwar besser, allerdings wird dann auch der Aufwand für die Konstruktion einer entsprechenden Basis deutlich größer. Der Wert $\alpha = \frac{3}{4}$ ist ein Kompromiß, der sich bewährt hat und daher – soweit mir bekannt – praktisch überall verwendet wird.

Die formale Definition einer „geeigneten“ Basis ist somit

**Definition:** Eine Gitterbasis $\vec{b}_1, \ldots, \vec{b}_n$ mit GRAM-SCHMIDT-Orthogonalisierung

$$\vec{c}_i = \vec{b}_i - \sum_{j=1}^{i-1} \mu_{ij}\vec{c}_j$$

heißt LLL-reduziert, wenn

$$\left|\mu_{ij}\right| \leq \frac{1}{2} \quad \text{für alle } i, j \qquad \text{und}$$

$$\left|\vec{c}_i + \mu_{i\,i-1}\vec{c}_{i-1}\right|^2 \geq \tfrac{3}{4}\left|\vec{c}_{i-1}\right|^2 \quad \text{für alle } i > 1 .$$

Diese Basen können nur dann nützlich sein, wenn sie existieren. Wir wollen uns daher zunächst ansehen, wie LENSTRA, LENSTRA und LOVÁSZ eine solche Basis konstruieren. Der Algorithmus ist natürlich nahe an der GRAM-SCHMIDT-Orthogonalisierung; da es für Gitterbasen deutlich weniger Manipulationsmöglichkeiten gibt als für Vektorraumbasen und wir außerdem noch die LOVÁSZ-Bedingung erfüllen müssen, treten aber eine ganze Reihe zusätzlicher Komplikationen auf.

Wir gehen aus von irgendeiner Basis $(\vec{b}_1, \ldots, \vec{b}_n)$ eines Gitters $\Gamma \subset \mathbb{R}^n$ und wollen daraus eine LLL-reduzierte Basis konstruieren. Als erstes konstruieren wir dazu nach GRAM-SCHMIDT eine Orthogonalbasis bestehend aus den Vektoren

$$\vec{c}_i = \vec{b}_i - \sum_{j=1}^{i-1} \mu_{ij}\vec{c}_j \in \mathbb{R}^n \quad \text{mit} \quad \mu_{ij} = \frac{\vec{b}_i \cdot \vec{c}_j}{\vec{c}_j \cdot \vec{c}_j} . \qquad (*)$$

Im Laufe des Algorithmus werden die $\vec{b}_i, \vec{c}_j$ und die $\mu_{ij}$ in jedem Schritt verändert, allerdings stets so, daß die Gleichungen $(*)$ erfüllt bleiben.

Wie bei GRAM-SCHMIDT hangeln wir uns dimensionsweise hoch, d.h. wir realisieren die Bedingungen aus der Definition einer LLL-reduzierten Basis zunächst für Teilgitter. Dazu fordern wir für eine natürliche Zahl $k \geq 1$ die Bedingungen

$$|\mu_{ij}| \leq \tfrac{1}{2} \quad \text{für} \quad 1 \leq j < i \leq k \qquad (A_k)$$

und

$$|\vec{c}_i + \mu_{i\,i-1}\vec{c}_{i-1}|^2 \geq \tfrac{3}{4}\,|\vec{c}_{i-1}|^2 \quad \text{für} \quad 1 < i \leq k \qquad (B_k)$$

Für $k = 1$ gibt es keine Indizes $i, j$, für die die rechtsstehenden Ungleichungen erfüllt sind, die Bedinungen sind also leer und somit trivialerweise erfüllt. Für $k = n$ dagegen besagen diese beiden Bedingungen, daß $(\vec{b}_1, \ldots, \vec{b}_n)$ eine LLL-reduzierte Gitterbasis von $\Gamma$ ist. Wir müssen also $k$ schrittweise erhöhen. Im Gegensatz zur Verfahren von GRAM-SCHMIDT müssen wir hier allerdings $k$ gelegentlich auch *erniedrigen* statt erhöhen. Der Wert von $k$ wird aber stets zwischen Null und $n$ liegen und stets so gewählt sein, daß die Bedingungen $(A_k)$ und $(B_k)$ erfüllt sind.

Für jeden neuen Wert von $k$, egal ob er größer oder kleiner ist als sein Vorgänger, führen wir die folgenden Schritte durch:

Wir wollen die Gitterbasis und die davon abgeleitete Orthogonalbasis einschließlich der Koeffizienten $\mu_{ij}$ so verändern, daß auch $(A_{k+1})$ und $(B_{k+1})$ gelten.

Die Bedingung $|\mu_{k+1\,k}| \leq \frac{1}{2}$ ist kein großes Problem: Wir runden einfach $\mu_{k+1\,k}$ zur nächsten ganzen Zahl $q$ (oder einer der beiden nächsten) und ersetzen $\vec{b}_{k+1}$ durch $\vec{b}_{k+1} - q\vec{b}_k$. Damit $(*)$ weiterhin gilt, ersetzen wir $\mu_{k+1\,k}$ durch $\mu_{k+1\,k} - q$ und die $\mu_{k+1\,j}$ mit $j < k$ durch $\mu_{k+1\,j} - q\mu_{kj}$.

Für das weitere Vorgehen müssen wir zwei Fälle unterscheiden:

*Fall 1:* $k \geq 1$ und $|\vec{c}_{k+1} + \mu_{k+1\,k}\vec{c}_k|^2 < \frac{3}{4}\,|c_k|^2$

In diesem Fall vertauschen wir $\vec{b}_k$ und $\vec{b}_{k+1}$. Da die GRAM-SCHMIDT-Orthogonalisierung von der Reihenfolge der Basisvektoren abhängt, müssen wir dann eine neue Orthogonalbasis $(\vec{d}_1, \ldots, \vec{d}_n)$ berechnen.



An den $\vec{c}_j$ mit $j < k$ (so es welche gibt) ändert sich dabei nichts: Sie werden beim GRAM-SCHMIDTschen Orthogonalisierungsverfahren berechnet, bevor die Vektoren $\vec{b}_k$ und $\vec{b}_{k+1}$ ins Spiel kommen. Für $j < k$ ist somit $\vec{d}_j = \vec{c}_j$.

Auch für $j > k+1$ ist $\vec{d}_j = \vec{c}_j$, denn der $j$-te Vektor der Orthogonalbasis ist die Projektion des $j$-ten Vektors der Ausgangsbasis auf das orthogonale Komplement des von den ersten $j-1$ Basisvektoren aufgespannten Untervektorraums, und für $j > k+1$ ist

$$[\vec{c}_1, \ldots, \vec{c}_j] = [\vec{b}_1, \ldots, \vec{b}_j] = [\vec{d}_1, \ldots, \vec{d}_j]\,.$$

Bleiben noch die Vektoren $\vec{d}_k$ und $\vec{d}_{k+1}$. Die müssen verschieden sein von den Vektoren $\vec{c}_k$ und $\vec{d}_{k+1}$, denn $[\vec{c}_1, \ldots, \vec{c}_k] = [\vec{b}_1, \ldots, \vec{b}_k]$, aber $[\vec{d}_1, \ldots, \vec{d}_k] = [\vec{b}_1, \ldots, \vec{b}_{k-1}, \vec{b}_{k+1}]$.

Nach den Formeln zur GRAM-SCHMIDT-Orthogonalisierung ist

$$\vec{c}_k = \vec{b}_k - \sum_{j=1}^{k-1} \mu_{kj}\vec{c}_j \qquad \text{mit} \quad \mu_{kj} = \frac{\vec{b}_k \cdot \vec{c}_j}{\vec{c}_j \cdot \vec{c}_j}$$

und

$$\begin{aligned}
\vec{c}_{k+1} &= \vec{b}_{k+1} - \sum_{j=1}^{k} \mu_{k+1\,j}\vec{c}_j \qquad \text{mit} \quad \mu_{k+1\,j} = \frac{\vec{b}_{k+1} \cdot \vec{c}_j}{\vec{c}_j \cdot \vec{c}_j} \\
&= \vec{b}_{k+1} - \sum_{j=1}^{k-1} \mu_{k+1\,j}\vec{c}_j - \mu_{k+1\,k}\vec{b}_k + \sum_{j=1}^{k-1} \mu_{k+1\,k}\mu_{kj}\vec{c}_j \\
&= \vec{b}_{k+1} - \mu_{k+1\,k}\vec{b}_k - \sum_{j=1}^{k-1} (\mu_{k+1\,j} - \mu_{k+1\,k}\mu_{kj})\vec{c}_j\,;
\end{aligned}$$

entsprechend ist

$$\vec{d}_k = \vec{b}_{k+1} - \sum_{j=1}^{k-1} \nu_{kj}\vec{d}_j \qquad \text{mit} \quad \nu_{kj} = \frac{\vec{b}_{k+1} \cdot \vec{d}_j}{\vec{d}_j \cdot \vec{d}_j}$$

und

$$\vec{d}_{k+1} = \vec{b}_k - \sum_{j=1}^{k} \nu_{k+1\,j}\vec{d}_j \qquad \text{mit} \quad \nu_{k+1\,j} = \frac{\vec{b}_k \cdot \vec{d}_j}{\vec{d}_j \cdot \vec{d}_j}$$

$$= \vec{b}_k - \sum_{j=1}^{k-1} \nu_{k+1\,j}\vec{c}_j - \nu_{k+1\,k}\left(\vec{b}_{k+1} - \sum_{j=1}^{k-1}\nu_{kj}\vec{c}_j\right) .$$

Da $\vec{c}_j = \vec{d}_j$ für $j < k$, folgt insbesondere

$$\nu_{kj} = \mu_{k+1\,j} \quad \text{und} \quad \nu_{k+1\,j} = \mu_{kj} \qquad \text{für } j < k$$

$$\vec{d}_{k+1} = \vec{c}_k - \nu_{k+1\,k}\vec{d}_k \qquad \text{und}$$

$$\vec{d}_k = \vec{b}_{k+1} - \sum_{j=1}^{k-1}\nu_{kj}\vec{d}_j = \vec{b}_{k+1} - \sum_{j=1}^{k-1}\mu_{k+1\,j}\vec{c}_j = \vec{c}_{k+1} + \mu_{k+1\,k}\vec{c}_k .$$

Nach der Voraussetzung für das Eintreten von Fall 1 ist das Längenquadrat des letzten Vektors kleiner als $\frac{3}{4}$ des Längenquadrats von $\vec{c}_k$; zumindest ein Vektor der Orthogonalbasis wird also durch die Vertauschung von $\vec{b}_k$ und $\vec{b}_{k+1}$ kürzer. Das Quadrat seiner Länge ist

$$\vec{d}_k \cdot \vec{d}_k = \vec{c}_{k+1} \cdot \vec{c}_{k+1} + \mu_{k+1\,k}^2\, \vec{c}_k \cdot \vec{c}_k .$$

Damit können wir nun auch $\nu_{k+1\,k}$ durch Daten der „alten“ Basis ausdrücken: Der Zähler ist

$$\vec{b}_k \cdot \vec{d}_k = \left(\vec{c}_k + \sum_{j=1}^{k-1}\mu_{kj}\vec{c}_j\right) \cdot \vec{d}_k = \vec{c}_k \cdot \vec{d}_k ,$$

denn für $j \le k-1$ steht $\vec{d}_k$ senkrecht auf $\vec{d}_j = \vec{c}_j$. Deshalb ist auch

$$\vec{c}_k \cdot \vec{d}_k = \vec{c}_k \cdot \left(\vec{b}_{k+1} - \sum_{j=1}^{k-1}\nu_{kj}\vec{d}_j\right) = \vec{c}_k\vec{b}_{k+1} ,$$

also ist

$$\nu_{k+1\,k} = \frac{\vec{b}_k \cdot \vec{d}_k}{\vec{d}_k \cdot \vec{d}_k} = \frac{\vec{c}_k \cdot \vec{b}_{k+1}}{\vec{d}_k \cdot \vec{d}_k} = \frac{|\vec{c}_k|^2}{|\vec{d}_k|^2} \cdot \frac{\vec{c}_k \cdot \vec{b}_{k+1}}{\vec{d}_k \cdot \vec{d}_k} = \frac{|\vec{c}_k|^2}{|\vec{d}_k|^2} \cdot \mu_{k+1\,k}$$

$$= \frac{\mu_{k+1\,k}\,|\vec{c}_k|^2}{|c_{k+1}|^2 + \mu_{k+1\,k}^2\,|\vec{c}_k|^2} .$$



Dank dieser Formeln ist daher auch $\vec{d}_{k+1} = \vec{c}_k - \nu_{k+1\,k}\vec{d}_k$ vollständig durch Daten der „alten“ Basis ausdrückbar.

Die Vektoren $\vec{d}_j$ mit $j > k+1$ sind wieder gleich den entsprechenden $\vec{c}_j$, denn der $j$-te Vektor der Orthogonalbasis ist ja der Lotvektor von $\vec{b}_j$ auf den von $\vec{b}_1, \ldots, \vec{b}_{j-1}$ aufgespannten Untervektorraum, und für $j > k+1$ hat sich durch die Vertauschung von $\vec{b}_k$ und $\vec{b}_{k+1}$ weder an $\vec{b}_j$ noch an diesem Vektorraum etwas geändert. Aus diesem Grund sind auch die Koeffizienten $\nu_{ij}$ gleich den entsprechenden $\mu_{ij}$, sofern weder $i$ noch $j$ gleich $k$ oder $k+1$ sind.

Damit fehlen uns nur noch die Koeffizienten $\nu_{ik}$ und $\nu_{i\,k+1}$ für $i > k+1$. Um sie zu berechnen, drücken wir zunächst $\vec{c}_k$ und $\vec{c}_{k+1}$ aus durch $\vec{d}_k$ und $\vec{d}_{k+1}$: Nach den obigen Formeln ist

$$\vec{d}_k = \vec{c}_{k+1} + \mu_{k+1\,k}\vec{c}_k$$

$$\vec{d}_{k+1} = \vec{c}_k - \nu_{k+1\,k}\vec{d}_k = (1 - \nu_{k+1\,k}\mu_{k+1\,k})\vec{c}_k - \nu_{k+1\,k}\vec{c}_{k+1} .$$

Addition von $\nu_{k+1\,k}$-mal der ersten Gleichung zur zweiten eliminiert $\vec{c}_{k+1}$ und liefert uns die Gleichung

$$\vec{c}_k = \nu_{k+1\,k}\vec{d}_k + \vec{d}_{k+1} .$$

Setzen wir dies ein in die zweite Gleichung, erhalten wir

$$(1 - \nu_{k+1\,k}\mu_{k+1\,k})(\nu_{k+1\,k}\vec{d}_k + \vec{d}_{k+1}) - \nu\vec{c}_{k+1} = \vec{d}_{k+1}$$

und damit

$$\vec{c}_{k+1} = (1 - \nu_{k+1\,k}\mu_{k+1\,k})\vec{d}_k - \mu_{k+1\,k}\vec{d}_{k+1} .$$

Da in der Summe $\vec{d}_k = \vec{c}_{k+1} + \mu_{k+1\,k}\vec{c}_k$ die Vektoren $\vec{c}_k$ und $\vec{c}_{k+1}$ orthogonal sind, ist $|\vec{d}_k|^2 = |\vec{c}_{k+1}|^2 + \mu_{k+1\,k}^2\,|\vec{c}_k|^2$, also

$$\frac{|\vec{c}_{k+1}|^2}{|\vec{d}_k|^2} = \frac{|\vec{d}_k|^2 - \mu_{k+1\,k}^2\,|\vec{c}_k|^2}{|\vec{d}_k|^2} = 1 - \frac{|\vec{c}_k|^2}{|\vec{d}_k|^2}\mu_{k+1\,k}^2 = 1 - \nu_{k+1\,k}\mu_{k+1\,k} .$$

Somit ist

$$\vec{c}_{k+1} = \frac{|\vec{c}_{k+1}|^2}{|\vec{d}_{k+1}|^2}\vec{d}_k - \mu_{k+1\,k}\vec{d}_{k+1} .$$

Mit diesen beiden Formel gehen wir nun für $i > k+1$ in die Gleichungen

$$\vec{b}_i = \vec{c}_i + \sum_{j=1}^{i-1} \mu_{ij}\vec{c}_j \,,$$

Die Teilsumme $\mu_{ik}\vec{c}_k + \mu_{i\,k+1}\vec{c}_{k+1}$ aus den Termen für $j = k$ und $j = k+1$ wird dabei zu

$$\left(\mu_{ik}\nu_{k+1\,k} + \mu_{i\,k+1}\frac{|\vec{c}_{k+1}|^2}{|\vec{d}_k|^2}\right)\vec{d}_k + \left(\mu_{ik} - \mu_{i\,k+1}\mu_{k+1\,k}\right)\vec{d}_{k+1}\,.$$

Somit ist

$$\nu_{ik} = \mu_{ik}\nu_{k+1\,k} + \mu_{i\,k+1}\frac{|\vec{c}_{k+1}|^2}{|\vec{d}_k|^2} \quad \text{und} \quad \nu_{i\,k+1} = \mu_{ik} - \mu_{i\,k+1}\mu_{k+1\,k}\,.$$

Wir ersetzen nun alle $\vec{c}_i$ durch die entsprechenden $\vec{d}_i$ und alle $\mu_{ij}$ durch die entsprechenden $\nu_{ij}$; dann ist $(*)$ auch für die Gitterbasis mit vertauschten Positionen von $\vec{b}_k$ und $\vec{b}_{k+1}$ erfüllt. Die Bedingungen $(A_k)$ und $(B_k)$ sind nun allerdings nur noch für $k-1$ sicher erfüllt; wir müssen daher $k$ durch $k-1$ ersetzen und einen neuen Iterationsschritt starten.

*2. Fall:* $k = 0$ oder $|\vec{c}_{k+1} + \mu_{k+1\,k}\vec{c}_k|^2 \geq \frac{3}{4}\,|\vec{c}_k|^2$

In diesem Fall sorgen wir zunächst dafür, daß alle $\mu_{k+1\,j}$ einen Betrag von höchstens ein halb haben. (Im Fall $k = 0$ gibt es hier natürlich nichts zu tun.)

Für $j = k$ haben wir das bereits zu Beginn des Schritts für $k$ sichergestellt; wir wählen nun den größten Index $\ell < k$, für den $|\mu_{k+1\,\ell}|$ größer ist als $\frac{1}{2}$ und verfahren damit wie oben: Wir ersetzen $\vec{b}_{k+1}$ durch $\vec{b}_{k+1} - q\vec{b}_\ell$ und $\mu_{k+1\,\ell}$ durch $\mu_{k+1\,\ell} - q$, wobei $q$ die nächste ganze Zahl zu $\mu_{k+1\,\ell}$ ist, und wir ersetzen alle $\mu_{k+1\,j}$ mit $j < \ell$ durch $\mu_{k+1\,j} - q\mu_{\ell\,j}$. Sofern es danach immer noch ein $\mu_{k+1\,j}$ vom Betrag größer $\frac{1}{2}$ gibt, wählen wir wieder den größten Index $j$ mit dieser Eigenschaft und so weiter, bis alle $|\mu_{k+1\,j}| \leq \frac{1}{2}$ sind.

Falls $k = n$ ist, endet der Algorithmus an dieser Stelle, und wir haben eine LLL-reduzierte Basis gefunden. Andernfalls ersetzen wir $k$ durch $k+1$ und beginnen mit einem neuen Iterationsschritt.



Um zu sehen, daß das Verfahren nach endlich vielen Schritten abbricht, müssen wir uns überlegen, daß der obige Fall 1, in dem der Index $k$ erniedrigt wird, nicht unbegrenzt häufig auftreten kann. Ausgangspunkt dazu ist die Beobachtung, daß zumindest *ein* Vektor der Orthogonalbasis im ersten Fall verkürzt wird: Der $k$-te Vektor wird ersetzt durch einen neuen, dessen Längenquadrat höchstens gleich $\frac{3}{4}$ mal des alten ist.

Um dies auszunutzen, definieren wir für $k = 1, \ldots, n$ die reelle Zahl $d_k$ als Determinante der $k \times k$-Matrix $B_k$ mit $ij$-Eintrag $\vec{b}_i \cdot \vec{b}_j$. Für $k = n$ können wir sie leicht auf bekannte Größen zurückführen: Ist $B$ die $n \times n$-Matrix mit dem Basisvektor $\vec{b}_i$ als $i$-ter Spalte, so ist offensichtlich $B_n = B^T B$, also ist $d_n = (\det B)^2 = d(\Gamma)^2$. Insbesondere ist also $d_n$ unabhängig von der Gitterbasis und hängt nur ab vom Gitter.

Entsprechend können wir für $d_k$ das Gitter $\Gamma_k = \mathbb{Z}\vec{b}_1 \oplus \cdots \oplus \mathbb{Z}\vec{b}_n$ im Vektorraum $\mathbb{R}\vec{b}_1 \oplus \cdots \oplus \mathbb{R}\vec{b}_n$ betrachten. Auch dies ist ein EUKLIDischer Vektorraum; wenn wir dort eine Orthonormalbasis (d.h. eine Orthogonalbasis, deren sämtliche Vektoren Länge eins haben) auszeichnen, wird er isomorph zum $\mathbb{R}^k$ mit seinem üblichen Skalarprodukt. Daher hängt auch $d_k$ nur ab von $\Gamma_k$, nicht aber von den Vektoren $\vec{b}_1, \ldots, \vec{b}_k$.

Solange wir im LLL-Algorithmus nur die $\mu_{ij}$ auf Werte vom Betrag höchstens $\frac{1}{2}$ reduzieren, ändert sich nichts an den Gittern $\Gamma_k$, also bleiben auch die $d_k$ unverändert.

Wenn wir aber zwei Basisvektoren $\vec{b}_k$ und $\vec{b}_{k+1}$ miteinander vertauschen, ändert sich das Gitter $\Gamma_k$ *und nur dieses;* alle anderen $\Gamma_i$ bleiben unverändert. $d_k$ ist das Quadrat von $d(\Gamma_k)$, und $d(\Gamma_k)$ können wir auch als Produkt der Längen der Vektoren der zugeordneten Orthogonalbasis berechnen. Von diesen ändert sich nur der $k$-te, und dessen Längenquadrat wird kleiner als drei Viertel des Längenquadrats des entsprechenden Vektors der vorherigen Orthogonalbasis. Somit wird $d_k$ mit einem Faktor von höchstens $\frac{3}{4}$ multipliziert.

Dasselbe gilt dann auch für das Produkt $D$ aller $d_k$; wenn wir zeigen können, daß dieses eine nur vom Gitter abhängige untere Schranke hat, folgt also, daß wir nicht unbegrenzt oft im Fall 1 des Algorithmus sein

können und dieser daher nach endlich vielen Schritten enden muß. Diese untere Schranke liefert uns der

**Gitterpunktsatz von Minkowski:** $\Gamma \subset \mathbb{R}^n$ sei ein Gitter und $M \subset \mathbb{R}^n$ sei eine zum Nullpunkt symmetrische beschränkte konvexe Teilmenge von $\mathbb{R}^n$. Falls das Volumen von $M$ größer ist als $2^n d(\Gamma)$, enthält $M$ mindestens einen vom Nullpunkt verschiedenen Punkt des Gitters.

Bevor wir diesen Satz beweisen, überlegen wir uns zunächst, daß er uns wirklich untere Schranke für alle $d_i$ und damit auch für $D$ liefert. Dazu wenden wir ihn an auf das Gitter $\Gamma_i$, das wir – siehe oben – als Teilmenge eines Vektorraums $\mathbb{R}^i$ auffassen können, und den Würfel

$$M = \left\{(x_1, \ldots, x_i) \in \mathbb{R}^i \;\middle|\; |x_i| \le \varepsilon \text{ für alle } i\right\}.$$

Wenn dessen Volumen $(2\varepsilon)^i$ größer ist als $2^n d(\Gamma_i)$, gibt es in $\Gamma_i$ (und damit erst recht in $\Gamma$) einen vom Nullvektor verschiedenen Vektor aus $M$. Dessen Länge ist höchstens gleich der halben Diagonale von $M$, also $\varepsilon\sqrt{i}$. Somit gibt es in $\Gamma_i$ und damit erst recht in $\Gamma$ einen Vektor $\vec{v} \neq \vec{0}$, dessen Länge höchstens gleich $\varepsilon\sqrt{n}$ ist.

Da Gitter diskrete Mengen sind, gibt es in $\Gamma$ eine Untergrenze $\mu$ für die Länge eines vom Nullvektor verschiedenen Vektors: Andernfalls wäre der Nullvektor ein Häufungspunkt des Gitters.

Falls der gerade betrachtete Würfel die Voraussetzung des Satzes von MINKOWSKI erfüllt, muß daher $\mu \le \varepsilon\sqrt{n}$ sein, d.h. für $\varepsilon < \mu/\sqrt{n}$ kann die Voraussetzung nicht erfüllt sein. Somit ist

$$\left(\frac{\mu}{\sqrt{n}}\right)^i \le d(\Gamma_i).$$

Damit haben wir eine nur vom Gitter abhängige Untergrenze für $d(\Gamma_i)$ gefunden, also auch für $d_i$ und damit auch für das Produkt $D$ aller $d_i$. Dies zeigt, sofern wir den Gitterpunktsatz von MINKOWSKI voraussetzen, daß der LLL-Algorithmus zur Basisreduktion nach endlich vielen Schritten endet.

Bleibt also noch der Beweis des Satzes von MINKOWSKI:



$(\vec{b}_1, \ldots, \vec{b}_n)$ sei eine Gitterbasis und $B$ sei die Matrix mit den $\vec{b}_i$ als Spaltenvektoren. Dann ist

$$\varphi: \begin{cases} \mathbb{R}^n \to \mathbb{R}^n \\ \vec{v} \mapsto B\vec{v} \end{cases}$$

eine bijektive lineare Abbildung, die das Gitter $\mathbb{Z}^n \subset \mathbb{R}^n$ auf $\Gamma$ abbildet. Volumina werden bei dieser Abbildung mit $\det B = d(\Gamma)$ multipliziert; die Menge $\varphi^{-1}(M)$ hat also mindestens das Volumen $2^n$ und ist wegen der Linearität von $\varphi$ beschränkt, symmetrisch und konvex. Es reicht, wenn wir zeigen, daß diese Menge einen vom Nullpunkt verschiedenen Punkt aus $\mathbb{Z}^n$ enthält.

Dies ist auch die Version des Satzes, die MINKOWSKI selbst bewiesen hat; hier soll aber nicht sein Beweis wiedergegeben werden, sondern eine später gefundene Alternative von BLICHFELDT. Dieser bewies 1914 den folgenden

**Satz:** $B$ sei eine beschränkte Teilmenge von $\mathbb{R}^n$ mit einem Volumen größer eins. Dann enthält $B$ zwei verschiedene Punkte $P, Q$, deren Verbindungsvektor in $\mathbb{Z}^n$ liegt.

Wenden wir diesen Satz an auf die Menge $B = \frac{1}{2}\varphi^{-1}(M)$, so finden wir zwei Punkte $P, Q \in B$, deren Verbindungsvektor in $\mathbb{Z}^n$ liegt. Die Punkte $2P$ und $2Q$ liegen in $\varphi^{-1}(M)$, also –wegen der vorausgesetzten Symmetrie – auch $-2Q$. Wegen der Konvexität von $\varphi(M)$ liegt auch der Mittelpunkt der Verbindungsstrecke von $2P$ und $-2Q$ in $\varphi^{-1}(M)$; in Koordinaten ist dies der Punkt

$$\tfrac{1}{2}\big(2P + (-2Q)\big) = P - Q \in \mathbb{Z}^n.$$

Damit ist der Gitterpunktsatz vom MINKOWSKI modulo dem Satz von BLICHFELDT bewiesen.

Als letztes bleibt damit noch der Satz von BLICHFELDT zu zeigen.

Dazu sei $W = \left\{(x_1, \ldots, x_n) \in \mathbb{R}^n \;\middle|\; 0 \le x_i < 1 \text{ für alle } i\right\}$ und für $\vec{v} \in \mathbb{Z}^n$ sein $W_{\vec{v}} = W + \vec{v}$ der um $\vec{v}$ verschobene Würfel $W$. Dann sind offensichtlich alle $W_{\vec{v}}$ disjunkt und überdecken gemeinsam den $\mathbb{R}^n$. Da $B$ beschränkt ist, können nur endlich viele der $W_{\vec{v}}$ einen nichtleeren

Durchschnitt $B_{\vec{v}}$ mit $B$ haben. Für jeden dieser Durchschnitte betrachten wir seine Verschiebung $B_{\vec{v}} - \vec{v}$ um den Vektor $-\vec{v}$; das ist offensichtlich eine Teilmenge von $W$.

Die Summe der Volumina aller dieser Teilmengen ist gleich dem Volumen von $B$, denn $B$ ist die disjunkte Vereinigung aller $B_{\vec{v}}$. Damit ist diese Summe nach Voraussetzung größer als eins; da $W$ das Volumen eins hat, muß es also zwei Vektoren $\vec{v} \neq \vec{w}$ geben, so daß $B_{\vec{v}} \cap B_{\vec{w}}$ nicht leer ist. Für einen Punkt $R$ aus diesem Durchschnitt sei $P$ seine Translation um $\vec{v}$ und $Q$ die um $\vec{w}$. Dann liegen $P \in B_{\vec{v}}$ und $Q \in B_{\vec{w}}$ beide in $B$, und ihr Verbindungsvektor ist $\vec{w} - \vec{v} \in \mathbb{Z}^n$. ∎

HERMANN MINKOWSKI wurde 1864 als Sohn einer deutsch-jüdischen Kaufmannsfamilie im damals russischen Aleksotas (heute Kaunas in Litauen) geboren. Als er acht Jahre alt war, zog die Familie um nach Königsberg, wo er auch zur Schule und ab 1880 zur Universität ging. Einer seiner Komillitonen war HILBERT. Während seines Studiums ging er auch für drei Semester nach Berlin, promovierte aber 1885 in Königsberg über quadratische Formen. In seiner Habilitationsschrift von 1887, die ihm eine Stelle an der Universität Bonn verschaffte, beschäftigt er sich erstmalig mit seiner *Geometrie der Zahlen*, in der der Gitterpunktsatz ein wichtiges Hilfsmittel ist. 1892 ging er zurück nach Königsberg, 1894 dann an die ETH Zürich, wo EINSTEIN einer seiner Studenten war. 1902 folgte (auf Initiative von HILBERT) der Ruf auf einen Lehrstuhl in GÖTTINGEN, wo er sich vor allem mit mathematischer Physik beschäftigte. Die Geometrie des (in heutiger Terminologie) MINKOWSKI-Raums erwies sich als fundamental für die Entwicklung der Relativitätstheorie. Er starb 1909 im Alter von nur 44 Jahren an einem Blinddarmdurchbruch.

HANS FREDERIK BLICHFELDT wurde 1873 in Dänemark geboren, jedoch wanderte die Familie bereits 1888 aus nach USA. Er bestand zwar bereits in Dänemark die Aufnahmeprüfung zur Universität mit Auszeichnung, aber seine Eltern konnten die Studiengebüren nicht aufbringen. So konnte er erst nach vier Jahren Arbeit in Farms und Sägemühlen ab 1894 an der Stanford University in Palo Alto, Kalifornien studieren. Einer seiner dortigen Professoren lieh im das notwendige Geld zum Promotionsstudium bei SOPHUS LIE in Leipzig; seine Promotion beschäftigte sich mit Transformationsgruppen im $\mathbb{R}^3$. 1898 kehrte er zurück nach Stanford, wo er zunächst als *instructor* arbeitete. 1913 erhielt er einen Lehrstuhl, 1927 bis zu seiner Emeritierung 1938 war er Dekan der mathematischen Fakultät. Seine Arbeiten beschäftigen sich mit der Geometrie der Zahlen und der Gruppentheorie. Er starb 1945 in Palo Alto.



Als Beispiel für die LLL-Reduktion wollen wir eine LLL-reduzierte Basis des von

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} \quad \text{und} \quad \vec{b}_3 = \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix}$$

erzeugten Gitters $\Gamma \subset \mathbb{R}^3$ bestimmen.

Als erstes brauchen wir die zugehörige Orthogonalbasis. Nach GRAM-SCHMIDT setzen wir $\vec{c}_1 = \vec{b}_1$ und wählen dann $\mu_{21}$ so, daß

$$(\vec{b}_2 - \mu_{21}\vec{c}_1) \cdot \vec{c}_1 = 10 - 14\mu_{21} = 0$$

ist, d.h.

$$\mu_{21} = \frac{5}{7} \quad \text{und} \quad \vec{c}_2 = \frac{1}{7}\begin{pmatrix} 16 \\ 4 \\ -8 \end{pmatrix}.$$

Der dritte Vektor $\vec{c}_3 = \vec{b}_3 - \mu_{31}\vec{c}_1 - \mu_{32}\vec{c}_3$ wird so gewählt, daß

$$\vec{c}_3 \cdot \vec{c}_1 = 11 - 14\mu_{21} = 0 \quad \text{und} \quad \vec{c}_3 \cdot \vec{c}_2 = \frac{36}{7} - \frac{48}{7}\mu_{32} = 0\,,$$

wir haben also

$$\mu_{31} = \frac{11}{14}, \quad \mu_{32} = \frac{3}{4} \quad \text{und} \quad \vec{c}_3 = \frac{1}{2}\begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}.$$

Wir beginnen die LLL-Reduktion mit $k = 1$ und müssen als erstes testen, ob $\mu_{21} = \frac{5}{7}$ einen Betrag von höchstens $\frac{1}{2}$ hat. Das ist offensichtlich nicht der Fall; die nächste ganze Zahl ist $q = 1$, also ersetzen wir $\vec{b}_2$ durch

$$\vec{b}_2 - \vec{b}_1 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}$$

und $\mu_{21}$ durch $\mu_{21} - 1 = -\frac{2}{7}$.

$$\vec{c}_2 + \mu_{21}\vec{c}_1 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix}$$

hat Längenquadrat acht, was kleiner ist als $\frac{3}{4}\,|\vec{c}_1|^2 = \frac{21}{2}$. Daher sind wir in Fall 1 und müssen $\vec{b}_1$ und $\vec{b}_2$ vertauschen. Der neue erste Vektor der Orthogonalbasis ist der gerade berechnete Vektor $\vec{d}_1 = \vec{c}_2 + \mu_{21}\vec{c}_1$ und

$$\nu_{21} = \mu_{21}\frac{|\vec{c}_1|^2}{|\vec{c}_2|^2 + \mu_{21}\,|\vec{c}_1|^2} = -\frac{1}{2}\,.$$

Damit ist

$$\vec{d}_2 = \vec{c}_1 - \nu_{21}\vec{d}_1 = \begin{pmatrix}2\\2\\2\end{pmatrix};$$

die neuen Koeffizienten sind

$$\nu_{31} = \mu_{31}\nu_{21} + \mu_{32}\frac{|\vec{c}_2|^2}{|\vec{d}_1|^2} = \frac{1}{4} \quad \text{und} \quad \nu_{32} = \mu_{31} - \mu_{32}\mu_{21} = 1\,.$$

Wir ersetzen die $\vec{c}_i$ durch die $\vec{d}_i$ und die $\mu_{ij}$ durch die $\nu_{ij}$; außerdem müssen wir, da wir Basisvektoren vertauscht haben, $k$ um eins erniedrigen. Wir gehen also mit $k = 0$ in den nächsten Iterationsschritt.

Dort sind wir mit $k = 0$ automatisch im zweiten Fall, und es gibt nichts zu tun; also erhöhen wir $k$ wieder auf eins und starten mit einen neuen Iterationsschritt.

Als erstes muß sichergestellt werden, daß $\mu_{21}$ höchstens Betrag $\frac{1}{2}$ hat; da $\mu_{21} = -\frac{1}{2}$, ist dies der Fall.

$$\vec{c}_2 + \mu_{21}\vec{c}_1 = \begin{pmatrix}1\\2\\3\end{pmatrix}$$

hat Längenquadrat 14, was größer ist als $\frac{3}{4}\,|\vec{c}_1|^2$; wir sind also im zweiten Fall und müssen die $\mu_{2j}$ mit $j < 1$ auf Beträge von höchstens $\frac{1}{2}$ reduzieren. Da es keine $j < 1$ gibt, ist diese Bedingung leer; wir können also $k$ auf zwei erhöhen und zum nächsten Schritt gehen.

$\mu_{32} = 1$ hat zu großen Betrag, muß also auf Null reduziert werden; damit Bedingung $(*)$ erfüllt bleibt, müssen wir auch $\mu_{31}$ durch $\mu_{31} - \mu_{21} = \frac{3}{4}$



ersetzen und $\vec{b}_3$ durch

$$\vec{b}_3 - \vec{b}_2 = \begin{pmatrix}1\\1\\-2\end{pmatrix}\,.$$

Dann hat

$$\vec{c}_3 + \mu_{32}\vec{c}_2 = \frac{1}{2}\begin{pmatrix}-1\\2\\-1\end{pmatrix}$$

Längenquadrat $\frac{3}{2}$, während $\frac{3}{4}\,|\vec{c}_2|^2 = 9$ ist, wir sind also wieder im ersten Fall und müssen $\vec{b}_2$ mit $\vec{b}_3$ vertauschen. Neuer zweiter Vektor der Orthogonalbasis wird

$$\vec{d}_2 = \vec{c}_3 + \mu_{32}\vec{c}_2 = \frac{1}{2}\begin{pmatrix}-1\\2\\-1\end{pmatrix}$$

und

$$\nu_{32} = \mu_{32}\frac{|\vec{c}_2|^2}{|\vec{c}_3|^2 + \mu_{32}^2\,|\vec{c}_2|^2} = 0, \quad \nu_{21} = \mu_{31} = \frac{3}{4}, \quad \nu_{31} = \mu_{21} = -\frac{1}{2}\,.$$

Damit können wir nun auch den dritten Vektor der neuen Orthogonalbasis berechnen als

$$\vec{d}_3 = \vec{c}_2 - \nu_{32}\vec{d}_2 = \begin{pmatrix}2\\2\\2\end{pmatrix}\,.$$

Wir ersetzen $\vec{c}_2, \vec{c}_3$ durch $\vec{d}_2, \vec{d}_3$ und die $\mu_{ij}$ durch die entsprechenden $\nu_{ij}$, erniedrigen $k$ auf eins und beginnen mit einem neuen Iterationsschritt.

Dieser beginnt mit der Reduktion von $\mu_{21} = \frac{3}{4}$. Nächste ganze Zahl ist eins, also setzen wir

$$\vec{b}_2 \leftarrow \vec{b}_2 - \vec{b}_1 = \begin{pmatrix}-1\\1\\0\end{pmatrix} \quad \text{und} \quad \mu_{21} \leftarrow \mu_{21} - 1 = -\frac{1}{4}\,.$$

Um zu sehen, in welchem Fall wir sind, müssen wir das Längenquadrat zwei von

$$\vec{c}_2 + \mu_{21}\vec{c}_1 = \begin{pmatrix}-1\\1\\0\end{pmatrix}$$

mit $\frac{3}{4}\left|\vec{c}_1\right|^2 = 6$ vergleichen: Wir sind wieder in Fall 1 und müssen jetzt $\vec{b}_1$ und $\vec{b}_2$ miteinander vertauschen. Neuer erster Vektor der Orthogonalbasis wird $\vec{d}_1 = \vec{c}_2 + \mu_{21}\vec{c}_1$; nach GRAM-SCHMIDT muß das natürlich der neue Vektor $\vec{b}_1$ sein. Weiter ist

$$\nu_{21} = \mu_{21}\frac{\left|\vec{c}_1\right|^2}{\left|\vec{c}_2\right|^2 + \mu_{21}^2\left|\vec{c}_1\right|^2} = -1 \quad \text{und} \quad \vec{d}_2 = \vec{c}_1 - \nu_{21}\vec{d}_1 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Die restlichen $\nu_{ij}$ sind

$$\nu_{31} = \mu_{31}\nu_{21} + \mu_{23}\frac{\left|\vec{c}_2\right|^2}{\left|\vec{d}_1\right|^2} = \frac{1}{2} \quad \text{und} \quad \nu_{32} = \mu_{31} - \mu_{32}\mu_{21} = -\frac{1}{2}.$$

Wir ersetzen $\vec{c}_1, \vec{c}_2$ durch $\vec{d}_1, \vec{d}_2$ und die $\mu_{ij}$ durch $\nu_{ij}$, setzen $k = 0$ und beginnen einen neuen Iterationsschritt.

Für $k = 0$ gibt es nichts zu tun, also können wir gleich wieder auf $k = 1$ erhöhen und einen neuen Schritt starten. Hier muß als erstes $\mu_{21} = -1$ reduziert und die Basis entsprechend angepaßt werden:

$$\vec{b}_2 \leftarrow \vec{b}_2 + \vec{b}_1 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \quad \text{und} \quad \mu_{21} \leftarrow \mu_{21} + 1 = 0\,.$$

Da $\mu_{21}$ verschwindet, ist $\vec{c}_2 + \mu_{21}\vec{c}_1 = \vec{c}_2 = \vec{b}_2$, das alte $\vec{b}_1$; sein Längenquadrat ist sechs und damit größer als $\frac{3}{4}\left|\vec{c}_1\right|^2 = \frac{3}{2}$. Daher sind wir im Fall 2, wo es auch für $k = 1$ nichts zu tun gibt, wir können also gleich mit $k = 2$ weitermachen.

$\mu_{32} = \frac{1}{2}$ ist bereits klein genug, und

$$\vec{c}_3 + \mu_{32}\vec{c}_2 = \frac{1}{2}\begin{pmatrix} 3 \\ 3 \\ 6 \end{pmatrix}$$

hat Längenquadrat $\frac{27}{2}$, was größer ist als $\frac{3}{4}\left|\vec{c}_2\right|^2 = \frac{9}{2}$. Daher sind wir wieder im Fall 2 und müssen uns daher nur noch um die restlichen $\mu_{3j}$ kümmern, d.h. um $\mu_{31} = \frac{1}{2}$. Dessen Betrag ist nicht größer als $\frac{1}{2}$, somit



gibt es nichts mehr zu tun. Wir können also auf $k = 3$ erhöhen und der Algorithmus endet mit der LLL-reduzierten Basis aus

$$\vec{b}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \quad \text{und} \quad \vec{b}_3 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}.$$

Die zugehörige Orthogonalbasis des $\mathbb{R}^3$ besteht aus $\vec{c}_1 = \vec{b}_1$ und $\vec{c}_2 = \vec{b}_2$ sowie

$$\vec{c}_3 = \begin{pmatrix} 2 \\ 2 \\ 2 \end{pmatrix} = \vec{b}_3 - \frac{1}{2}\vec{c}_1 + \frac{1}{2}\vec{c}_2\,.$$

Nachdem wir nun gesehen haben, daß wir aus einer vorgegebenen Basis eine LLL-reduzierte Basis konstruieren können, stellt sich die Frage, was uns so eine Basis nützt bei der Suche nach kurzen Vektoren in einem Gitter. Unter den Vektoren einer LLL-reduzierten Basis muß kein kürzester Vektor des Gitters vorkommen, aber wir können immerhin obere Schranken finden für die Längen der Basisvektoren.

$(\vec{b}_1, \ldots, \vec{b}_n)$ sei eine LLL-reduzierte Basis eines Gitters $\Gamma \subset \mathbb{R}^n$ und $(\vec{c}_1, \ldots, \vec{c}_n)$ sei die dazu nach GRAM-SCHMIDT berechnete Orthogonalbasis. Dann ist

$$\vec{b}_i = \vec{c}_i + \sum_{j=1}^{i-1}\mu_{ij}\vec{c}_j \Longrightarrow \left|\vec{b}_i\right|^2 = \left|\vec{c}_i\right|^2 + \sum_{j=1}^{i-1}\mu_{ij}^2\left|\vec{c}_j\right|^2 \le \left|\vec{c}_i\right|^2 + \frac{1}{4}\sum_{j=1}^{i-1}\left|\vec{c}_j\right|^2,$$

denn für eine LLL-reduzierte Basis sind alle $\left|\mu_{ij}\right| \le \frac{1}{2}$. Außerdem ist wegen der LOVÁSZ-Bedingung $\left|\vec{c}_{j-1}\right|^2 \le 2\left|\vec{c}_j\right|^2$; mehrfache Anwenug dieser Ungleichung führt auf $\left|\vec{c}_j\right|^2 \le 2^{i-j}\left|\vec{c}_i\right|^2$ für alle $j < i$ und

$$\begin{aligned}\left|\vec{b}_i\right|^2 &\le \left|\vec{c}_i\right|^2 + \frac{1}{4}\sum_{j=1}^{i-1}2^{j-i}\left|\vec{c}_i\right|^2 = \left(1 + \frac{1}{4}\sum_{j=1}^{i-1}2^i\right)\left|\vec{c}_i\right|^2 = \frac{1+2^{i-1}}{2}\left|\vec{c}_i\right|^2 \\ &\le 2^{i-1}\left|\vec{c}_i\right|^2\,.\end{aligned}$$

Ebenfalls wegen der LOVÁSZ-Bedingung gilt für $j < i$ die Ungleichung $\left|\vec{c}_j\right|^2 \le 2^{i-j}\left|\vec{c}_i\right|^2$ und damit auch

$$\left|\vec{b}_j\right|^2 \le 2^{j-1}\left|\vec{c}_j\right|^2 \le 2^{j-1}2^{i-j}\left|\vec{c}_i\right|^2 = 2^{i-1}\left|\vec{c}_i\right|^2\,.$$

Nun seien $\vec{v}_1, \ldots, \vec{v}_m$ irgendwelche linear unabhängige Vektoren aus dem Gitter $\Gamma$ und $k$ sei die kleinste Zahl mit der Eigenschaft, daß alle $\vec{v}_i$ im von $\vec{b}_1$ bis $\vec{b}_k$ aufgespannten Teilgitter liegen. Dann gibt es Koeffizienten $\lambda_{ij}, \nu_{ij}$ derart, daß

$$\vec{v}_i = \sum_{j=1}^{k} \lambda_{ij} \vec{b}_j = \sum_{j=1}^{k} \nu_{ij} \vec{c}_j$$

ist. Die $\lambda_{ij}$ müssen dabei ganze Zahlen sein, die $\nu_{ij}$ natürlich nicht. Da wir die $\nu_{ij}$ aus den $\lambda_{ij}$ berechnen können, indem wir die Darstellung der $\vec{b}_j$ als Linearkombinationen der $\vec{c}_j$ einsetzen, muß aber $\nu_{ik} = \lambda_{ik}$ und somit ganzzahlig sein, denn außer $\vec{b}_k$ liefert kein anderes $\vec{b}_j$ einen Beitrag mit $\vec{c}_k$.

Wir wählen ein $i$, für das $\lambda_{ik} = \nu_{ik}$ nicht verschwindet; wegen der Minimalität von $k$ muß es das geben. Dann ist

$$\left|\vec{b}_k\right|^2 \le 2^{k-1} \left|\vec{c}_k\right|^2 \le 2^{k-1} \lambda_{ik}^2 \left|\vec{c}_k\right|^2 \le 2^{k-1} \left|\vec{v}_i\right|^2$$

und für $j < k$ ist

$$\left|\vec{b}_j\right|^2 \le 2^{k-1} \left|\vec{c}_k\right|^2 \le 2^{k-1} \left|\vec{v}_i\right|^2 .$$

Über die Zahlen $i$ und $k$ wissen wir nur, daß $k$ zwischen $m$ und $n$ liegen muß (sonst wären die $\vec{v}_j$ linear abhängig) und $i \le m$. Daher haben wir für alle $j \le m$ die Abschätzung

$$\left|\vec{b}_j\right|^2 \le 2^{n-1} \max\left\{ \left|\vec{v}_1\right|^2, \ldots, \left|\vec{v}_m\right|^2 \right\}$$

oder

$$\left|\vec{b}_j\right| \le 2^{(n-1)/2} \max\left\{ \left|\vec{v}_1\right|, \ldots, \left|\vec{v}_m\right| \right\} .$$

Speziell für $m = 1$ erhalten wir die Abschätzung

$$\left|\vec{b}_1\right| \le 2^{(n-1)/2} \left|\vec{v}_1\right|$$

für jeden vom Nullvektor verschiedenen Gittervektor $\vec{v}_1$. Dies gilt insbesondere für den kürzesten solchen Vektor; die Länge von $\vec{b}_1$ übersteigt dessen Länge also höchstens um den Faktor $2^{(n-1)/2}$.



## § 11: Anwendung auf Faktorisierungsprobleme

Wie in §9 betrachten wir wieder ein Polynom $f \in \mathbb{Z}[x]$ vom Grad $d$ sowie ein Polynom $h \in \mathbb{Z}[x]$ vom Grad $e$ mit führendem Koeffizienten eins, das modulo einer Primzahl $p$ irreduzibel ist und modulo einer gewissen $p$-Potenz $p^k$ Teiler von $f$. Wir nehmen außerdem an, daß $h^2$ modulo $p$ kein Teiler von $f$ mod $p$ ist; wenn wir von einem quadratfreien Polynom $f$ ausgehen und $p$ die Diskriminante nicht teilt, ist letzteres automatisch erfüllt.

Nach dem ersten Lemma aus §9 hat $f$ einen bis aufs Vorzeichen eindeutig bestimmten irreduziblem Faktor $h_0$, der modulo $p$ durch $h$ teilbar ist. Diesen Faktor wollen wir berechnen.

Dazu betrachten wir wieder das Gitter $\Lambda$ aller Polynome aus $\mathbb{Z}[x]$ vom Grad höchstens einer gewissen Schranke $m$, die modulo $p^k$ durch $h$ teilbar sind; nach dem letzten Lemma aus §9 ist ein Polynom $v \in \Lambda$ mit $\|f\|_2 \cdot \|v\|_2 < p^{ke}$ ein Vielfaches von $h_0$. Wenn wir genügend viele kurze Vektoren aus $\Lambda$ finden, können wir daher hoffen, daß deren ggT gleich $h_0$ ist.

Als erstes wollen wir eine Schranke für die $L^2$-Norm eines Teilers von $f$ finden. Aus den Überlegungen zur LANDAU-MIGNOTTE-Schranke in Kapitel 2, §8 folgt leicht

**Lemma:** Ist $g \in \mathbb{Z}[x]$ ein Teiler vom Grad $e$ des Polynoms $f \in \mathbb{Z}[x]$, so ist

$$\|g\|_2 \le \sqrt{\binom{2e}{e}} \, \|f\|_2 .$$

*Beweis:* Wenn $g$ Teiler von $f$ ist, muß auch der führende Koeffizient von $g$ Teiler des führenden Koeffizienten von $f$ sind; daher ist das Maß $\mu(g)$ kleiner oder gleich $\mu(f)$, und letzteres wiederum ist nach Lemma 4 aus Kapitel 2, §8 kleiner oder gleich $\|f\|_2$. Nach dem dortigen Lemma 2 ist außerdem der Betrag des $i$-ten Koeffizienten von $g$ kleiner oder gleich

$\binom{e}{i}\mu(g)$, also kleiner oder gleich $\binom{e}{i}\|f\|_2$. Somit ist

$$\|g\|_2^2 \le \sum_{i=0}^{e} \binom{e}{i}^2 \|f\|_2^2 \ .$$

Das Lemma ist bewiesen, wenn wir zeigen können, daß die Summe der $\binom{e}{i}^2$ gleich $\binom{2e}{e}$ ist. Dies läßt sich am einfachsten kombinatorisch einsehen: $\binom{2e}{e}$ ist die Anzahl von Möglichkeiten, aus einer Menge $\mathcal{M}$ von $2e$ Elementen $e$ auszuwählen. Wir zerlegen $\mathcal{M}$ in zwei disjunkte Teilmengen $\mathcal{M}_1$ und $\mathcal{M}_2$ mit je $e$ Elementen. Die Wahl von $e$ Elementen aus $\mathcal{M}$ ist gleichbedeutend damit, daß wir für irgendein $i$ zwischen 0 und $e$ zunächst $i$ Elemente von $\mathcal{M}_1$ auswählen und dann $e-i$ Elemente aus $\mathcal{M}_2$. Die Anzahl der Möglichkeiten dafür ist $\binom{e}{i}\binom{e}{e-i} = \binom{e}{i}^2$, die Summe aller dieser Quadrate ist also $\binom{2e}{e}$.

∎

Damit können wir nun zunächst eine Schranke für den Grad von $h_0$ finden:

**Lemma:** $(b_0, \ldots, b_m)$ sei eine LLL-reduzierte Basis des Gitters $\Lambda$ und $p^{ke} < 2^{md/2}\binom{2m}{m}^{d/2}\|f\|_2^{m+d}$. Genau dann hat $h_0$ höchstens den Grad $m$, wenn

$$\|b_0\|_2 < \sqrt[d]{\frac{p^{ke}}{\|f\|_2^m}} \ .$$

*Beweis:* Falls $b_0$ diese Ungleichung erfüllt, ist $\|f\|_2^m \|b_0\|_2^d < p^{ke}$, nach dem letzten Lemma aus §9 ist also $b_0$ ein Vielfaches von $h_0$, und damit kann $h_0$ höchstens den Grad $m$ haben.

Umgekehrt sei $\deg h_0 \le m$. Nach der oben bewiesenen LANDAU-MIGNOTTE-Schranke für die $L^2$-Norm eines Teilers von $f$ ist

$$\|h_0\|_2 \le \sqrt{\binom{2m}{m}}\,\|f\|_2 \ .$$



Kombinieren wir dies mit der Ungleichung am Ende des vorigen Paragraphen, erhalten wir die Abschätzung

$$\|b_0\|_2 \le 2^{m/2}\|h_0\|_2 \le 2^{m/2}\sqrt{\binom{2m}{m}}\,\|f\|_2 \ .$$

Nach Voraussetzung ist

$$p^{ke} < 2^{md/2}\binom{2m}{m}^{d/2}\|f\|_2^{m+d} \Longrightarrow \frac{p^{ke}}{\|f\|_2^m} < 2^{md/2}\binom{2m}{m}^{d/2}\|f\|_2^d \ .$$

Ziehen wir auf beiden Seiten der letzten Ungleichung die $d$-te Wurzel und lombinieren dies mit der obigen Abschätzung für $\|b_0\|_2$, folgt die Behauptung.

∎

**Lemma:** Angenommen, zusätzlich zu den Voraussetzungen des vorigen Lemmas existieren Indizes $j$, für die

$$\|b_j\|_2 < \sqrt[d]{\frac{p^{ke}}{\|f\|_2^m}}$$

ist, und $t$ ist der größte solche Index. Dann ist

$$\deg h_0 = m - t \quad \text{und} \quad h_0 = \mathrm{ggT}(b_0, \ldots, b_t) \ .$$

*Beweis:* Wir betrachten die Menge $J$ aller Indizes $j$ mit $\|b_j\| < \sqrt[d]{\frac{p^{ke}}{\|f\|_2^m}}$. Das Lemma am Ende von §9 sagt uns, daß $h_0$ jedes $b_j$ mit $j \in J$ teilt, also auch

$$h_1 = \mathrm{ggT}\big(\{b_j \mid j \in J\}\big) \ .$$

Da jedes dieser $b_j$ durch $h_1$ teilbar ist und höchstens den Grad $m$ hat, liegt es im Gitter

$$\mathbb{Z}\cdot h_1 \oplus \mathbb{Z}\cdot x h_1 \oplus \cdots \oplus \mathbb{Z}\cdot x^{m-\deg h_1} h_1 \ .$$

Da die $b_j$ als Elemente einer Basis linear unabhängig sind, ist die Elementanzahl von $J$ höchstens gleich $m+1-\deg h_1$. Nach der zu Beginn dieses Paragraphen bewiesenen LANDAU-MIGNOTTE-Schranke für die $L^2$-Norm eines Teilers ist außerdem

$$\|x^i h_0\|_2 = \|h_0\|_2 \le \sqrt{\binom{2m}{m}}\,\|f\|_2$$

für jedes $i$. Da die verschiedenen $x^i h_0$ linear unabhängig sind, ist daher nach der Abschätzung am Ende von §10

$$\|b_j\|_2 \le 2^{m/2}\sqrt{\binom{2m}{m}}\,\|f\|_2 \quad \text{für } j = 0,\ldots,m-\deg h_0\,.$$

Wegen der vom vorigen Lemma übernommenen Voraussetzung

$$p^{ke} > 2^{md/2}\binom{2m}{m}^{d/2}\|f\|_2^{m+d}$$

ist die rechte Seite kleiner als die $d$-te Wurzel aus $p^{ke}/\|f\|_2^m$, so daß alle $j \le m+1-\deg h_0$ in $J$ liegen. $J$ hat daher mindestest $m+1-\deg h_0$ Elemente und höchstens $m+1-\deg h_1$ Elemente. Da $h_0$ ein Teiler von $h_1$ ist, muß also $\deg h_0 = \deg h_1$ sein. Um zu sehen, daß sich die beiden Polynome höchstens durch das Vorzeichen unterscheiden, genügt es zu zeigen, daß auch $h_1$ primitiv ist.

$h_1$ ist der ggT von $b_0$ bis $b_t$; wäre $h_1$ nicht primitiv, könnten also auch diese $b_j$ nicht primitiv sein. Wenn aber eine ganze Zahl $d$ eines der Polynome $b_j$ teilt, ist wegen der Primitivität von $h_0$ auch $b_j/d$ ein Vielfaches von $h_0$, d.h. $b_j/d$ liegt im Gitter $\Lambda$. Da $(b_0,\ldots,b_m)$ eine Gitterbasis ist, geht das nur für $d = \pm 1$. Also ist $b_j$ und damit auch $h_1$ primitiv.

∎

Damit ist klar, wie wir den Algorithmus von ZASSENHAUS zur Faktorisierung eines primitiven Polynoms $f \in \mathbb{Z}[x]$ vom Grad $d$ so abändern können, daß nicht mehr im Extremfall alle Kombinationen der Faktorisierung modulo $p$ miteinander kombiniert werden müssen: Wir berechnen zunächst die Resultante von $f$ und $f'$ und wählen einen Primzahl $p$, die diese nicht teilt. Dann faktorisieren wir $f \bmod p$ nach dem Algorithmus von BERLEKAMP; die Faktoren positiven Grades seien so normiert, daß ihre höchsten Koeffizienten alle eins sind. Außerdem berechnen wir die LANDAU-MIGNOTTE-Schranke für die Faktoren von $f$ und wählen eine Zahl $M$, die größer ist, als deren Doppeltes.

Wir schreiben $f = f_1 f_2$ mit zwei Polynomen $f_1, f_2 \in \mathbb{Z}[x]$, wobei wir von $f_1$ die Faktorisierung in $\mathbb{Z}[x]$ kennen und von $f_2$ nur die modulo $p$. Zunächst ist natürlich $f_1 = 1$ und $f_2 = f$.



Solange $f_2$ positiven Grad hat, betrachten wir einen der irreduziblen Faktoren von $f_2 \bmod p$ aus $\mathbb{F}_p[x]$; sein Grad sei $e$. Mit Hilfe des HENSELschen Lemmas liften wir den Faktor zu einem Polynom $h \in \mathbb{Z}[x]$, der auch noch modulo einer $p$-Potenz $p^k \ge M$ Teiler von $f_2$ ist.

Wir wählen einen Grad $m \ge e$ und wollen entweder einen modulo $p$ durch $h$ teilbaren irreduziblen Faktor $h_0$ von $f_2$ mit Grad höchstens $m$ konstruieren oder aber beweisen, daß es keinen solchen Faktor gibt.

Dazu stellen wir zunächst sicher, daß

$$p^{ke} > 2^{md/2}\binom{2m}{m}^{d/2}\|f\|_2^{m+d}$$

und betrachten dann zum Gitter $\Lambda$ mit Basis

$$p^k X^i \quad \text{für } 0 \le i < e \quad \text{und} \quad x^j h \quad \text{für } 0 \le j \le m-e$$

die LLL-Reduktion $b_0,\ldots,b_m$ dieser Basis. Falls

$$\|b_0\|_2 \ge \sqrt[d]{\frac{p^{ke}}{\|f\|_2^m}}\,,$$

gibt es keinen Faktor $h_0$ vom Grad höchstens $m$. Wenn $m \ge [\frac{1}{2}]\deg f_2$ ist, wissen wir, daß $f_2$ irreduzibel, also $h_0 = f_2$ ist; anderfalls müssen wir entweder $m$ erhöhen oder einen anderen irreduziblen Faktor von $f_2 \bmod p$ betrachten.

Wenn obige Ungleichung nicht gilt, wählen wir für $t$ den größten Index $j$ mit

$$\|b_j\|_2 < \sqrt[d]{\frac{p^{ke}}{\|f\|_2^m}}$$

und können $h_0$ berechnen als ggT von $b_0,\ldots,b_t$.

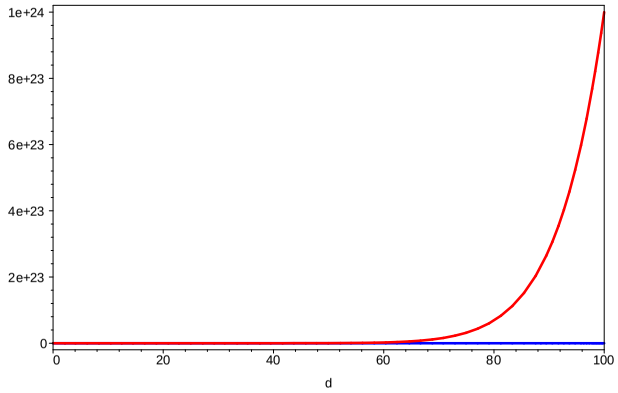
Wir ersetzen dann $f_1$ durch $f_1 h_0$ und $f_2$ durch $f_2/h_0$. Zur Faktorisierung des neuen $f_2$ modulo $p$ brauchen wir natürlich keinen BERLEKAMP-Algorithmus, sondern können einfach testen, welche Faktoren des alten $f_2 \bmod p$ in $h_0 \bmod p$ (oder im neuen $f_2 \bmod p$) stecken.

Was haben wir durch diese mathematisch deutlich kompliziertere Modifikation gewonnen? Theoretisch sehr viel: Wie das Beispiel der

SWINNERTON-DYER-Polynome zeigte, kann es uns im sechsten Schritt des Algorithmus von ZASSENHAUS passieren, daß wir bei der Faktorisierung eines Polynoms vom Grad $d$ bis zu $2^{[d/2]}$ Kombinationen ausprobieren müssen, bevor wir erkennen, daß das zu faktorisierende Polynom irreduzibel ist. Alle anderen Schritte erfordern einen Zeitaufwand, der als Funktion von $n$ sehr viel langsamer ansteigt als $2^{[d/2]}$, so daß asymptotisch betrachtet dieser Term dominiert.

Ein guter Teil der zitierten Arbeit von LENSTRA, LENSTRA und LOVASZ widmet sich der Frage, wie groß der entsprechende Aufwand mit LLL-Reduktion ist; sie zeigen, daß der domierende Term hier nur $d^{12}$ ist, was – wie aus der Analysis bekannt – deutlich schwächer ansteigt.

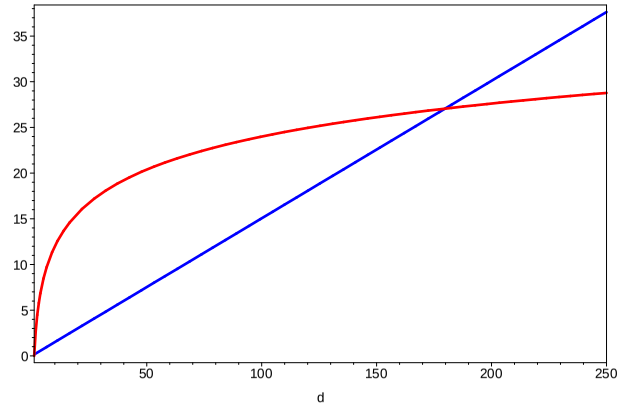
Läßt man zur Illustration die beiden Kurven im Bereich von $d = 0$ bis $d = 100$ von Maple zeichnen, so ist eine der beiden praktisch ununterscheidbar von der $d$-Achse, während die andere steil ansteigt:




Nachrechnen zeigt allerdings, daß die steil ansteigende rote Kurve der Graph von $d \mapsto d^{12}$ ist: Für $d = 100$ etwa ist $2^{50} \approx 1{,}125899907 \cdot 10^{15}$ und $50^{12} \approx 2{,}441406250 \cdot 10^{20}$ ist mehr als 200 000 mal so groß. Erst zwischen $d = 179$ und $d = 180$ schneiden sich die beiden Kurven, und ab dort dominiert dann die Exponentialkurve. Damit man etwas sehen kann, sind in der folgenden Abbildung die (dekadischen) Logarithmen der beiden Funktionen aufgezeichnet. Wie man sieht, liegt im unteren Bereich, mit dem wir es meist zu tun haben, die Kurve zu $d^{12}$ deutlich



über der für $2^{d/2}$; erst ab $d = 180$ wird $2^{d/2}$ größer.




Für Polynome in den Größenordnungen, mit denen wir es üblicherweise zu tun haben, sollte also der klassische ZASSENHAUS-Algorithmus schneller sein. Theoretisch könnte man den Exponenten 12 noch für jedes $\varepsilon > 0$ auf $9 + \varepsilon$ reduzieren, indem man asymptotisch schnellere Algorithmen zur Multiplikation ganzer Zahlen einsetzt, in der Praxis wird der Algorithmus dadurch allerdings deutlich langsamer: Die entsprechenden Methoden sind zwar nützlich für Zahlen mit Millionen von Dezimalstellen, nicht aber bei „nur“ ein paar hundert oder Tausend.

Man muß allerdings bedenken, daß sich alle hier angegebenen Schranken auf den schlechtestmöglichen Fall beziehen, der nur selten eintritt. In der Praxis dürften wohl beide Algorithmen deutlich schneller sein als es die asymptotischen Schranken vermuten lassen.

Trotzdem benutzt Maple anscheinend zur Faktorisierung keine Gittermethoden, obwohl die LLL-Reduktion für Gitterbasen als Funktion `lattice` zur Verfügung steht. Bislang scheint Faktorisierung mit LLL auf experimentelle zahlentheoretische Systeme beschränkt zu sein, in denen nicht über $\mathbb{Q}$, sondern über einem Erweiterungskörper faktorisiert wird. LLL-Basisreduktion hat heute Anwendungen in vielen Teilen der Mathematik, bei der ursprünglich vorgesehenen Anwendung der Faktorisierung von Polynomen aus $\mathbb{Z}[x]$ spielt er aber bislang in der Praxis nur eine ziemlich untergeordnete Rolle,