Prof. Friedrich Eisenbrand
Alfonso Cevallos

Discussions from: May 19, 2015

Computer Algebra

Spring 2015

Assignment Sheet 7

Note: These are just notes and not necessarily full solutions to each exercise. Please report any mistakes you may find.

Exercise 1

See Exercise 1, Assignment 7, from the 2011 course (http://disopt.epfl.ch/CAL2011).

Exercise 2

Part 1. We need to present an element $q \in \mathbb{Z}_p$ with $q^2 = -1$. Recall that, for a prime p, the multiplicative group $\mathbb{Z}_p^*$ is cyclic, i.e. it contains an element g (a primitive root) of order $|\mathbb{Z}_p^*| = p-1$. Now, if we define $x := g^{\frac{p-1}{2}}$, we know that $x \neq 1$ (by definition of order). And we also saw in class that the equation $x^2 = 1$ admits only two solutions $x = \pm 1$ in $\mathbb{Z}_p^*$, when p is prime. So we conclude that $x = -1$. Finally, if $q = g^{\frac{p-1}{4}}$ (this is well defined because $\frac{p-1}{4}$ is an integer), we have $q^2 = x = -1$.

Part 2. Let Λ be the lattice generated by $B = \left(\begin{smallmatrix} 1 & 0 \\ q & p \end{smallmatrix}\right)$ (where q is the element defined above). Then any lattice element is of the form $v = x\binom{1}{q} + y\binom{0}{p} = \binom{x}{qx+py}$, for some integers x, y; so $\|v\|^2 \equiv x^2 + (qx+py)^2 \equiv x^2 + q^2x^2 \equiv x^2 - x^2 \equiv 0 \mod p$.

Part 3. The previous result implies that the norm of any non-zero lattice vector is either $\sqrt{p}$, or at least $\sqrt{2p}$. On the other hand, Minkowski's theorem says that the lattice Λ must have a non-zero vector of norm at most $2\sqrt{\frac{\det \Lambda}{\pi}} = 2\sqrt{\frac{p}{\pi}} < \sqrt{2p}$ (where $\det \Lambda = |\det B| = p$, and π is the area of the unit disc). Therefore there is a lattice vector of norm $\sqrt{p}$.

Part 4. If $v = \binom{a}{b}$ is a lattice vector of norm $\sqrt{p}$, for some integers a, b, then $a^2 + b^2 = p^2$.

Part 5. In contrast, we prove that for any prime $p \geq 3$ with $p \not\equiv 1 \mod 4$, p cannot be written as the sum of two squares. p is odd, so necessarily we have $p \equiv 3 \mod 4$. However, it is easy to prove that a perfect square is congruent to either 0 or 1 $\mod 4$. So the sum of two squares can never be congruent to 3 $\mod 4$.

Exercise 3

Using the same argument seen in class, we can assume without loss of generality that $\text{vol}(K) >$

$k \cdot 2^n$. We will extend a proof of Minkowski's theorem via Blichfeldt's theorem. Blichfeldt's theorem is the following: every bounded measurable set S with $\text{vol}(S) > k$ contains $k+1$ points $x_0, \cdots, x_k$ such that all differences $x_j - x_i$ are integral.[1]

Let us first see how this theorem can be used to prove the desired result. Define $S := \frac{1}{2} \cdot K$. Note that $\text{vol}(S) = 2^{-n}\text{vol}(K) > k$, so apply Blichfeldt's theorem to get $k+1$ points $x_0, \cdots, x_k \in S$ whose pairwise differences are integral. The difference between each pair of such points is a non-zero integer point in K due to the symmetry and convexity of K:

$$x_j - x_i = \frac{1}{2}\underbrace{(2x_j)}_{\in K} + \frac{1}{2}\underbrace{(-2x_i)}_{\in K} \in K$$

So we have $\binom{k+1}{2}$ potential non-zero integer points. Unfortunately, we cannot guarantee that they are distinct. However, it is easy to get k distinct non-zero integer points by taking $x_j - x_0$ for all $j = 1, \cdots, k$. And a simple trick lets us improve this to $2k$. Let $\succeq$ be a total order on $\mathbb{R}^n$ that is consistent with the vector space structure, e.g. the lexicographic order.[2] By reordering if necessary we can assume that, for each $j = 1, \cdots, j$, we have $x_j \succ x_0$, and this implies that $x_j - x_0 \succ 0$ and $x_0 - x_j \prec 0$. From this, it is easy to see that the $2n$ points of the form $x_j - x_0$ and $x_0 - x_j$ are all distinct.

It remains to prove Blichfeldt's theorem. The idea is fairly simple: slice S according to the fundamental parallelepiped of the lattice, translate all those slices on top of each other, and use a measure-theoretic pigeon-hole principle to deduce that at least one point must be covered by $k+1$ different slices: this gives us the desired $k+1$ points. To make this precise, denote by $P := [0,1)^n$ the half-open fundamental parallelepiped of the integer lattice. We will use the fact that the whole space can be partitioned into copies of P, shifted by points in the integer lattice, like this: $\mathbb{R}^n = \cup_{v \in \mathbb{Z}^n}(P + v)$. Also, for any set $A \subset \mathbb{R}^n$, we define its characteristic function $\chi_A(x) = 1$ if $x \in A$, 0 otherwise.

$$\begin{aligned}\text{vol}(S) &= \int_{\mathbb{R}^n} \chi_S(x)dx = \int_{\mathbb{R}^n} \sum_{v \in \mathbb{Z}^n} \chi_{S \cap (P+v)}(x)dx \\ &= \int_{\mathbb{R}^n} \sum_{v \in \mathbb{Z}^n} \chi_{S \cap P}(x+v)dx = \int_P \sum_{v \in \mathbb{Z}^n} \chi_S(x+v)dx = \int_P f(x)dx\end{aligned}$$

Where the function $f(x) := \sum_{v \in \mathbb{Z}^n} \chi_S(x+v)$ is defined only over P, and represents the "density" obtained at each point by slicing S along the lattice and translating all slices on top of each other inside P. Notice that f is well defined, because as S is bounded, there is only a finite number of non-zero terms in the sum over $\mathbb{Z}^n$. Also, $f(x)$ is an integer, and there must be a point $x^* \in P$ such that $f(x^*) > k$, because otherwise we get a contradiction:

$$k < \text{vol}(S) = \int_P f(x)dx \leq k \int_P dx = k$$

[1] As usual, this can be generalized for arbitrary lattices. If $\text{vol}(S) > k \cdot \det(\Lambda)$, then S contains $k+1$ points whose differences are lattice vectors.

[2] Let $x, y \in \mathbb{R}^n$, then in the lexicographic order one has $x \succ y$ if the first component where x and y differ is greater in x than in y.

Therefore $f(x^*) \geq k+1$. From the definition of f, this implies that there are vectors $v_0, \cdots, v_k \in \mathbb{Z}^n$ such that $\chi_S(x^* + v_0) = \cdots = \chi_S(x^* + v_k) = 1$, which is to say that all these last $k+1$ points are in S, and of course their differences are integral.

Exercise 4

Part 1. Given a real number $\alpha > 0$, and a denominator bound M, the best fractional approximation for α is defined as $\operatorname{argmin}\{|\alpha - p/q| \ : \ (p,q) \in \mathbb{N}^2, \ q \leq M\}$. In the last lecture, we proved this problem can be solved in polynomial time with the following algorithm:

Define $B = \left(\begin{smallmatrix} p_1 & p_2 \\ q_1 & q_2 \end{smallmatrix}\right) = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right)$;
while $q_1 + q_2 \leq M$ **do**
  if *(any of* $\frac{p_1}{q_1}$, $\frac{p_2}{q_2}$ *or* $\frac{p_1+p_2}{q_1+q_2}$ *is exactly* α*)* **then**
    return the corresponding solution (p_1, q_1), (p_2, q_2) or $(p_1 + p_2, q_1 + q_2)$;
  end
  if $\frac{p_1+p_2}{q_1+q_2} > \alpha$ **then**
    update $\binom{p_1}{q_1} \leftarrow \binom{p_1}{q_1} + c\binom{p_2}{q_2}$,
    where c is the largest integer s.t. $\frac{p_1+cp_2}{q_1+cq_2} \geq \alpha$ and $q_1 + cq_2 \leq M$;
  else
    update $\binom{p_2}{q_2} \leftarrow c\binom{p_1}{q_1} + \binom{p_2}{q_2}$,
    where c is the largest integer s.t. $\frac{cp_1+p_2}{cq_1+q_2} \leq \alpha$ and $cq_1 + q_2 \leq M$;
  end
end
return (p_1, q_1) or (p_2, q_2), whichever is a better approximation.

Throughout the execution of the algorithm, we have the properties that B is a unimodular matrix, and $\frac{p_1}{q_1} \geq \alpha \geq \frac{p_2}{q_2}$. We execute this algorithm for $\alpha = 365.2422$ and $M = 40$, and obtain the following sequence of updates for B: $\left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 365 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1461 & 365 \\ 4 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1461 & 10592 \\ 4 & 29 \end{smallmatrix}\right), \left(\begin{smallmatrix} 12053 & 10592 \\ 33 & 29 \end{smallmatrix}\right)$, and the corresponding sequence of parameters c is $365, 4, 7, 1$. The best approximation for α is therefore $\frac{12053}{33} = 365 + \frac{8}{33}$.

Part 2. We want a rule where 8 out of every 33 years are leap years. A possible rule is: a year is a leap year iff it is congruent to $0, 4, 8, 12, 16, 20, 24$ or $28 \mod 33$. Such a calendar will be off by one day in roughly 4460 years. Remark 1: This is actually more accurate than the Gregorian calendar we use in Switzerland (every year that is divisible by 4 is a leap year, except for those divisible by 100 but not by 400), which corresponds to an approximation of $365 + \frac{97}{400}$, and will be off by one day in roughly 3300 years. Remark 2: The Solar Hijri calendar, used in Iran, follows a 33-year cycle to determine leap years.

Part 3. The 4-th convergent of α is $(365, 4, 7, 1)$. This is precisely the sequence of parameters c computed by our algorithm. And the approximation $365 + \frac{8}{33} = 365 + \frac{1}{4 + \frac{1}{7 + \frac{1}{1}}}$ corresponds to the *continued fraction* of the 4-th convergent. Remark: The approximation $365 + \frac{1}{4}$ mentioned above is the continued fraction of the 2-convergent $(365, 4)$ of α.

Exercise 5

We perform the LLL algorithm for the lattice basis $B = (b1 \;\; b2 \;\; b3) = \begin{pmatrix} 1 & 4 & 0 \\ 0 & 2 & 0 \\ 0 & 15 & 3 \end{pmatrix}$. We compute the Gram-Schmidt orthogonalization $B = B^* R$, using the formula $r_{ij} = \frac{b_i^* \cdot b_j}{b_i^* \cdot b_i^*}$ for $i < j$, and in the normalization phase we modify the basis to make these r coefficients lie between $-1/2$ and $1/2$.

We start with $b_1^* = b_1$. Then, $r_{12} = \frac{b_1^* \cdot b_2}{b_1^* \cdot b_1^*} = \frac{4}{1}$, so to normalize we update $b_2 \leftarrow b_2 - \lfloor r_{12} \rceil b_1 = \begin{pmatrix} 4 \\ 2 \\ 15 \end{pmatrix} - 4 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 15 \end{pmatrix}$, and $r_{12} \leftarrow r_{12} - \lfloor r_{12} \rceil = 0$, and obtain $b_2^* = b_2 - r_{12} b_1^* = b_2$. We continue with $r_{13} = \frac{b_2^* \cdot b_3}{b_2^* \cdot b_2^*} = \frac{0}{1}$ and $r_{23} = \frac{b_2^* \cdot b_3}{b_2^* \cdot b_2^*} = \frac{45}{229}$ (no need to normalize since both values are <1/2), and obtain $b_3^* = b_3 - r_{13} b_1^* - r_{23} b_2^* = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} - \frac{45}{229} \begin{pmatrix} 0 \\ 2 \\ 15 \end{pmatrix} = \begin{pmatrix} 0 \\ -90/229 \\ 12/229 \end{pmatrix}$. We end the normalization phase with matrices $B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 15 & 3 \end{pmatrix}$ and $B^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & \frac{-90}{229} \\ 0 & 15 & \frac{12}{229} \end{pmatrix}$.

In the swapping phase, we will swap columns j and $j+1$ whenever $\|b_{j+1}^* + r_{j,j+1} b_j^*\|^2 < \frac{3}{4} \|b_j^*\|^2$, or equivalently whenever $\|b_{j+1}^*\|^2 < (\frac{3}{4} - r_{j,j+1}^2) \|b_j^*\|^2$. Columns 1 and 2 give $229 > (\frac{3}{4} - 0) \cdot 1$ (no need to swap); but columns 2 and 3 give $\frac{36}{229} < \left(\frac{3}{4} - (\frac{45}{229})^2\right) \cdot 229$, so we swap them, and normalize again. We have $b_2 = \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix}$, $r_{12} = 0$ and $b_2^* = b_2$ (no normalization); and $b_3 = \begin{pmatrix} 0 \\ 2 \\ 15 \end{pmatrix}$, $r_{13} = 0$, $r_{23} = \frac{45}{9} = 5$, so we update $b_3 \leftarrow b_3 - \lfloor r_{23} \rceil b_2 = \begin{pmatrix} 0 \\ 2 \\ 15 \end{pmatrix} - 5 \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}$, and $r_{23} \leftarrow r_{23} - \lfloor r_{23} \rceil = 0$, and obtain $b_3 = b_3^*$. At this point we have $B = B^* = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 2 \\ 0 & 3 & 0 \end{pmatrix}$, and start a new swapping phase.

Columns 1 and 2 give $9 > (\frac{3}{4} - 0) \cdot 1$, but columns 2 and 3 give $4 < (\frac{3}{4} - 0) \cdot 9$, so we swap them. We can check that no more normalization of swapping is required, and we stop with the basis $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix}$.