Electronic Colloquium on Computational Complexity - Reports Series 1996 - available via:
ECCC TR96-007
FTP: ftp.eccc.uni-trier.de:/pub/eccc/
WWW: http://www.eccc.uni-trier.de/eccc/
Email: ftpmail@ftp.eccc.uni-trier.de with subject 'help eccc'

Generating Hard Instances of Lattice Problems

Extended abstract

M. Ajtai
IBM Almaden Research Center
650 Harry Road, San Jose, CA, 95120
e-mail: ajtai@almaden.ibm.com

ABSTRACT. We give a random class of lattices in $\mathbf{Z}^n$ so that, if there is a probabilistic polynomial time algorithm which finds a short vector in a random lattice with a probability of at least $\frac{1}{2}$ then there is also a probabilistic polynomial time algorithm which solves the following three lattice problems in *every* lattice in $\mathbf{Z}^n$ with a probability exponentially close to one. (1) Find the length of a shortest nonzero vector in an n-dimensional lattice, approximately, up to a polynomial factor. (2) Find the shortest nonzero vector in an n-dimensional lattice L where the shortest vector v is unique in the sense that any other vector whose length is at most $n^c\|v\|$ is parallel to v, where c is a sufficiently large absolute constant. (3) Find a basis $b_1, ..., b_n$ in the n-dimensional lattice L whose length, defined as $\max_{i=1}^n \|b_i\|$, is the smallest possible up to a polynomial factor.

A large number of the existing techniques of cryptography include the generation of a specific instance of a problem in NP (together with a solution) which for some reason is thought to be difficult to solve. As an example we may think about factorization. Here a party of a cryptographic protocol is supposed to provide a composite number m so that the factorization of m is known to her but she has some serious reason to believe that nobody else will be able to factor m. The most compelling reason for such a belief would be a mathematical proof of the fact that the prime factors of m cannot be found in less then k step in some realistic model of computation, where k is a very large number. For the moment we do not have any proof of this type, neither for specific numerical values of m and k, nor in some assymptotic sense.

In spite of the lack of mathematical proofs, in two cases at least, we may expect that a problem will be difficult to solve. One is the class of NP-complete problems. Here we may say that if there is a problem at all which is difficult to solve, then an NP-complete problem will provide such an example. The other case is, if the problem is a very famous question (e.g. prime factorization), which for a long time were unsuccesfully attacked by the most able scientists. In both cases it is reasonable to expect that the problem is difficult to solve. Unfortunately the expression "difficult to solve" means difficult to solve in the worst case. If our task is to provide a specific instance of the problem, these general principles do not provide any guidence about how to create one.

It has been realized a long time ago that a possible solution would be to find a set of randomly generated problems and show that if there is an algorithm which finds a solution of a random instance with a positive probability, then there is also an algorithm which solves one of the famous unsolved problems in the worst case. (It does not really matter whether this "positive probability" is $\frac{1}{2}$, ϵ or $\frac{1}{n^c}$, because taking many instances of the problem and asking for a solution for each of them, we may create a new problem so that even if it can be solved with an exponentially small positive probability then the "famous" worst case problem can be solved with a probability exponentially close to one.)

In this paper we give such a class of random problems. In fact we give a random problem: find a short vector in a certain class of random lattices (whose elements can be generated together with a short vector in them), whose solution in the mentioned sense would imply the solution of a group of related "famous" problems in the worst case. We mention here three of these worst-case problems:

(P1) *Find the length of a shortest nonzero vector in a n dimensional lattice, approximately, up to a polynomial factor.*

(P2) *Find the shortest nonzero vector in an n dimensional lattice L where the shortest vector v is unique in the sense that any other vector whose length is at most $n^c\|v\|$ is parallel to v, where c is a sufficiently large absolute constant.*

(P3) *Find a basis $b_1, ..., b_n$ in the n-dimensional lattice L whose length, defined as $\max_{i=1}^n \|b_i\|$, is the smallest possible up to a polynomial factor.*

Remarks. 1. (P2) can be given in a more general form. *If a lattice $L \subseteq \mathbf{Z}^n$ is given, then find all sublattices $L' = V \cap L$ (by giving a basis in them), where V is a d-dimensional subspace of $\mathbf{Z}^n$ so that $\min\{d, n-d\}$ is smaller than a constant and $V \cap L$ has a basis $v_1, ..., v_d$ so that for all $w \in L \backslash V$, $n^{c^d} \max_{i=1}^d \|v_i\| \leq \|w\|$, where $c_d > 0$ is sufficiently large with respect to d, but does not depend on anything else.*

2. The random problem can be also formulated as a linear simultaneous Diophantine approximation problem.

3. Although (P1) is not in NP (we are not able to check whether our estimate is good), still, our algorithm will give a one-sided certificate. Namely we may get a certificate which shows that there is no shorter vector than the lower bound in our estimate. (This certificate will be a basis with small length in the dual lattice.) In problem (P3) we get an estimate on the minimal basis length of the lattice. Since

we get it together with a basis, we have a cerificate for the upper bound. We get no certificate on the lower bound.

4. There are problems, e.g. find the discrete logarithm of a number modulo p or decide wether a number is quadratic residue modulo $m = pq$, where it is known that for any fixed choice of p resp. m the worst case problem can be easily reduced to the average case problem. For the choice of p resp. m however, there is no known method wich would guarantee that we get a problem as hard as the worst case.

Notation. $\mathbf{R}$ is the field of real numbers, $\mathbf{Z}$ is the ring of integers, $\mathbf{R}^n$ is the Euclidean space of n-dimensional real vectors with the usual Euclidean norm $\|a\|$. $\mathbf{Z}^n$ is the set of vectors in $\mathbf{R}^n$ with integer coordinates.

Definitions. 1. If $a_1, ..., a_n$ are linearly independent vectors in an $\mathbf{R}^n$, then we say that the set $\{\sum_{i=1}^n k_i a_i | k_1, ..., k_n$ are integers $\}$ is a lattice in $\mathbf{R}^n$. We will denote this lattice by $L(a_1, ..., a_n)$. The set $a_1, ..., a_n$ is called a basis of the lattice. The determinant of a lattice L will be the absolute value of the determinant whose rows are the vectors $a_1, ..., a_n$. $\mathrm{sh}(L)$ will be the length of a shortest nonzero vector in L, and $\mathrm{bl}(L)$ the length of the shortest basis as defined in (P3)

Historical remarks. We give here only a few facts to show that the mentioned lattice problems are sufficiently "famous" for our purposes. The question of finding a short vector in a lattice was already formulated by Dirichlet in 1842 in the form of simultaneous Diophantine approximation problems. Although the lattices where these Diophatine problems can be formulated in terms of finding a short vector or estimating the length of a short vector, form only a special class of lattices in $\mathbf{R}^n$ the random class that we will define later is an element of this special class. Moreover Dirichlet's theorem about the existence of a good approximation, as we will see is very relevant to our topic. His theorem is actually an upper bound on $\mathrm{sh}(L)$. His proof is non-constructive.

Minkowski's theorem about convex, central symmetric bodies (published in 1896) is also an esitmate about the length of the shortest non-zero vector (with repsect to a norm defined by the convex body). In the case of Euclidean norm, when the convex body is a sphere, it gives the upper bound $\mathrm{sh}(L) \le cn^{\frac{1}{2}}(\det L)^{\frac{1}{n}}$ where $\det L$ is the determinant of the lattice. This inequality and its consequences play an important role in our proof. Minkowski's proof is also nonconstructive. Minkowski's theory of successive minima formulates (as the two extreme cases) the problem of finding the length of a shortest vector and the length of the shortest basis (in the sense given in our problems).

A.K. Lenstra, H.W. Lenstra and L. Lovasz gave a deterministic polynomial time algorithm (the basis reduction or L^3 algorithm) which finds a vector in each lattice

$L \subseteq \mathbf{R}^n$ whose length is at most $2^{\frac{n-1}{2}}\mathrm{sh}(L)$. C.P. Schnorr proved that the factor $2^{\frac{n-1}{2}}$ can be replaced by $(1+\epsilon)^n$ for any fixed $\epsilon > 0$. These algorithms naturally give an esitmate on $\mathrm{sh}(L)$ up to a factor of $2^{\frac{n-1}{2}}$ resp. $(1+\epsilon)^n$. The L^3 algorithm was used in successfull attacks on different knapsack cryptosystems. (Cf. Adleman [Ad], Lagarias and Odlyzko [LaOd], Brickell [Br]). Lattices, where the shortest vector is unique in a sense similar to that of (P2), play an important role (see [LaOd]). (The polynomial factor of (P2) is substituted by an exponential one.)

The definition of the random class. Since a lattice is an infinite set we have to fix a finite representation of the lattices in the random class, that can serve as an input for our algorithm. The lattices of the random class will consist of vectors with integer coordinates. Moreover these lattices will be defined modulo q (where q will be an integer depending only on n), in the sense that if two vectors are congruent modulo q then either both of them or neither of them belong to the lattice. Finally the lattices of the random class will be defined as the set of all sequences of integers of length m, (m will depend only on n) which are orthogonal to a given sequence of vectors $u_1, ..., u_m \in \mathbf{Z}^m$ modulo q. More precisely if $\nu = \langle u_1, ..., u_m \rangle$ where $u_i \in \mathbf{Z}^n$ then let $\Lambda(\nu, q)$ be the lattice of all sequences of integers $h_1, ..., h_m$ so that $\sum_{i=1}^m h_i u_i \equiv 0 \pmod q$ where the mod q congruence of two vectors means that all of their coordinates are congruent. Every lattice in our random class will be of the form $\Lambda(\nu, q)$ for some ν and for a single fixed q (depending only on n).

Our definition of the random class will depend on the choice of two absolute constant c_1 and c_2. If n is given let $m = [c_1 n \log n]$ and $q = [n^{c_2}]$. For each n we will give a single random variable λ so that $\Lambda = \Lambda(\lambda, q)$ is a lattice with dimension m. (The existence of a polynomial time algorithm which finds a short vector in Λ will imply the existence of such an algorithm which solves the mentioned problems in every lattice $L \subseteq \mathbf{R}^n$.)

First we define an "idealized" version λ' of λ, whom we can define in a simpler way. The disadvantage of λ' is that we do not know how to generate λ' together with short vector in $\Lambda(\lambda', q)$. Then we define λ (in a somewhat more complicated way) so that we can generate it together with a short vector in $\Lambda(\lambda, q)$ and we will also have that $P(\lambda \neq \lambda')$ is exponentially small. This last inequality implies that if we prove our theorem for $\Lambda(\lambda', q)$ then it will automatically hold for $\Lambda(\lambda, q)$ too.

Let $\lambda' = \langle v_1, ..., v_m \rangle$ where $v_1, ..., v_m$ are chosen independently and with uniform distribution from the set of all vectors $\langle x_1, ..., x_n \rangle$ where $x_1, ..., x_n$ are integers and $0 \leq x_i < q$. To find a short vector in the lattice $\Lambda(\lambda', q)$ is equivalent of finding a solution for a linear simultaneous Diophantine approximation problem. Dirichlet's theorem implies that if c_1 is sufficiently large with respect to c_2 then there is always a vector shorter than n. (The proof of Dirichlet's theorem is not constructive, it is based on the Pigeonhole Principle applied to a set of exponential size.)

Definition of λ. We randomize the vectors $v_1,...,v_{m-1}$ independently and with uniform distribution on the set of all vectors $\langle x_1,...,x_n\rangle \in \mathbf{Z}^n$, with $0 \le x_i < q$. Independently of this randomization we also randomize a $0,1$-sequence $\delta_1,...,\delta_{m-1}$ where the numbers δ_i are chosen independently and with uniform distribution from $\{0,1\}$. We define v_m by $v_m \equiv -\sum_{i=1}^{m-1}\delta_i v_i \pmod q$ with the additional constraint that every component of v_m is an integer in the interval $[0,q-1]$. Let $\lambda = \langle v_1,...,v_m\rangle$. (If we want to emphasize the dependence of λ on n, c_1, c_2 then we will write λ_{n,c_1,c_2}.) We prove that the distribution of λ is exponentially close to the uniform distribution in the sense that $\sum_{a\in A}|P(\lambda = a) - |A|^{-1}| \le 2^{-cn}$, where A is the set of possible values of λ. This will imply that the random variable λ' with the given distribution can be chosen in a way that $P(\lambda' \ne \lambda)$ is exponentially small.

With this definition our theorem will be formulated in the following way: "if there is an algorithm which finds a short vector in $\Lambda(\lambda,q)$ given λ as an input, then etc." That is, we allow the algorithm whose existense is assumed in the theorem to use λ.

The representation of the lattice vectors. To give an exact formulation of our results we have to fix some representation of the lattice vectors in problems (P1),(P2),(P3). As we have seen already, the vectors in the random lattice Λ have integer coordinates, that is, they are in $\mathbf{Z}^m$. We will formulate problems (P1), (P2), (P3) in terms of vectors in $\mathbf{Z}^n$ as well. (Another possible approach would be to have lattice vectors in $\mathbf{R}^n$ given by oracles. In that case it is natural (and possible) to give the random class in terms of vectors whose components are random real numbers. The modulo q arithmetic can be substituted by arithmetic modulo 1.) The simplest approach is to assume that the lattices in $\mathbf{Z}^n$ are presented with a basis where each coordinate of each vector is an integer given by a polynomial (in n) number of bits. However our results remain valid even if the numbers are longer. Naturally in this case the input size is not n (the dimension of the lattice) but the total number of bits in the presentation of the lattice, so our algorithm will be polynomial in this number.

Definitions. 1. If v is a shortest nonzero vector in the lattice $L \subseteq \mathbf{R}^n$, and $\alpha > 1$, we say that v is α-unique if for any $w \in L$, $\|w\| \le \alpha\|v\|$ implies that v and w are parallel.

2. If k is an integer then size(k) will denote the number of bits in the binary representation of k, (size$(0) = 1$). If $v = \langle x_1,...,x_n\rangle \in \mathbf{Z}^n$ then size$(v) = \sum_{i=1}^{n}$ size(x_i). Our definition implies that for all $v \in \mathbf{Z}^n$, size$(v) \ge n$.

Theorem 1 . *There are absolute constants c_1, c_2, c_3 so that the following holds. Suppose that there is a probabilistic polynomial time algorithm $\mathcal{A}$ which given a value of the random variable λ_{n,c_1,c_2} as an input, with a probability of at least $1/2$ outputs a vector of $\Lambda(\lambda_{n,c_1,c_2}, [n^{c_2}])$ of length at most n. Then, there is a probabilistic algorithm $\mathcal{B}$ with the following properties. If the linearly independent vectors $a_1, ..., a_n \in \mathbf{Z}^n$ are given as an input, then $\mathcal{B}$, in time polynomial in $\sigma = \sum_{i=1}^{n} \mathrm{size}(a_i)$, gives the outputs $z, u, \langle d_1, ..., d_n \rangle$ so that, with a probability of greater than $1-2^{-\sigma}$, the following three requirements are met:*

(1.1) *if v is a shortest non-zero vector in $L(a_1, ..., a_n)$ then $z \le \|v\| \le n^{c_3} z$*

(1.2) *if v is an n^{c_3}-unique shortest nonzero vector in $L(a_1, ..., a_n)$ then $u = v$ or $u = -v$*

(1.3) *$d_1, ..., d_n$ is a basis with $\max_{i=1}^{n} \|d_i\| \le n^{c_3} \mathrm{bl}(L)$.*

Remarks. 1. The probability $1/2$ in the assumption about $\mathcal{A}$ can be replaced by n^{-c}. This will increase the running time of $\mathcal{B}$ by a factor of at most n^c but does not affect the constants c_1, c_2 and c_3.

2. If we assume that $\mathcal{A}$ produces a vector of length at most $n^{c'}$ for some $c' > 1$ then the theorem remains true but c_1, c_2 and c_3 will depend on c'.

Sketch of the proof. (We give a detailed the proof in the attached appendix.) We show first that there is an algorithm $\mathcal{B}$ so that (1.3) holds. By (1.3) we have an estimate H on the minimal basis length up to a polynomial factor. It is a consequence of Minkowski's upper bound on $\mathrm{sh}(L)$ that H^{-1} is an estimate (up to a polynomial factor) on $\mathrm{sh}(L^*)$, where L^* is the dual lattice of $L \subseteq \mathbf{R}^n$. (The dual lattice is the lattice of all linear functionals on $\mathbf{R}^n$ that take integer values on every vectors of L. Each element of L^* is identified, in the natural way, with an element of the Euclidean space $\mathbf{R}^n$.) Therefore by estimating the minimal basis length of L^* we get also an estimate on $\mathrm{sh}((L^*)^*) = \mathrm{sh}(L)$.

We will construct an algorithm which produces the output with property (1.2) by using an algorithm which satisfies (1.3). In this step we will not use the assumption about our random class directly. Therefore, the critical part of the proof is the proof of (1.3).

First we note that it is easy to see that from a set of n linearly independent vectors $r_1, ..., r_n \in L$ we can construct in polynomial time a basis of $s_1, ..., s_n$ of L so that $\max_{i=1}^{n} \|s_i\| \le n \max_{i=1}^{n} \|r_i\|$. Therefore it is enough to construct a set of linearly independent elements of L so that each of them is shorter than $n^{c_3-1}\mathrm{bl}(L)$.

Assume now that we have a lattice $L \subseteq \mathbf{Z}^n$ and assume that we have a set of linearly independent elements $a_1, ..., a_n \in L$ so that $\max_{i=1}^{n} \|a_i\| = M$. If $M \le$

$n^{c_3-1}\mathrm{bl}(L)$ then we have already found a basis with the required properties. Assume that $M > n^{c_3-1}\mathrm{bl}(L)$. We will construct another set of linearly independent elements, $b_1, ..., b_n \in L$ so that $\max_{i=1}^n \|b_i\| \le \frac{M}{2}$. Iterating this procedure we find a linearly independent set of elements $d'_1, ..., d'_n$ with $\max_{i=1}^n \|d'_i\| \le n^{c_3-1}\mathrm{bl}(L)$ in less than $\log_2 M \le 2\sigma$ steps.

Starting from the set $a_1, ..., a_n$, we construct a set of linearly independent elements in L, $f_1, ..., f_n$ so that $\max_{i=1}^n \|f_i\| \le n^3 M$ and the parallelepiped $W = \mathcal{P}(f_1, ..., f_n)$ defined by the vectors $f_1, ..., f_n$ is very close to a cube. Closeness will mean that the distance of each vertex of $\mathcal{P}(f_1, ..., f_n)$ from the vertices of a fixed cube will be at most nM and as a consequence the volume, the width and the surface area of W will be about the same as that of a cube of similar size. This will imply that if we cover the space with the cells of the lattice determined by a short basis, then most of the cells intersecting W will be compltetely in its interior. (The number of exceptional cells is polynomially small compared to the total.) As a consequence we get that all of the parallelepipeds $u + W$ where u is an arbitrary element of $\mathbf{R}^n$ have about the same number of lattice points. The error again will be polynomially small fraction of the total. These remain true even if we consider all of the parallelepipeds $u + \frac{1}{q}W$ where $q = [n^{c_2}]$ and c_3 is sufficiently large with respect to c_2. This fact will ensure that if we pick a lattice point at random from a set D of almost disjoint parallelepipeds of type $u + \frac{1}{q}W$, then the distribution induced on D is very close to the uniform distribution. (We will consider to parallelepiped almost disjoint if their interiors are disjoint.)

Now we cut W into q^n small parallelepipeds each of the form $(\sum_{i=1}^n \frac{t_i}{q} f_i) + \frac{1}{q}W$, where $0 \le t_i < q$, $i = 1, ..., n$ is a sequence of integers. We take a random sequence of lattice points $\xi_1, ..., \xi_m$, $m = [c_1 n \log n]$ from the parallelepiped $W = \mathcal{P}(f_1, ..., f_n)$ independently and with (almost) uniform distribution. (Such a random sequence can be generated in the following way. Let $b_1, ..., b_n$ be a basis. We take random sums $s = \sum_{i=1}^n \alpha_i b_i$ with random integer coefficients $\alpha_i \in [0, T]$, where T is a very large integer, and then we reduce s into a point in W modulo $(f_1, ..., f_n)$.)

Assume that $\xi_j \in (\sum_{i=1}^n \frac{t_i^{(j)}}{q} f_i) + \frac{1}{q}W$. Let $v_j = \langle t_1^{(j)}, ..., t_n^{(j)} \rangle$. We will consider the sequence $v_1, ..., v_m$ as a value of the random variable λ. (The distribution of $v_1, ..., v_m$ is not identical to that of λ, still we will prove that it is so close to it that this identification does not change our conclusions.) Applying algorithm $\mathcal{A}$ to the input $v_1, ..., v_n$ we get a vector $\langle h_1, ..., h_m \rangle \in \mathbf{Z}^n$ so that with a probability of at least $1/2$ its length is at most n and $\sum_{j=1}^n h_j v_j = 0$. We claim that with a positive probability $u = \sum_{j=1}^n h_j \xi_j \ne 0$ and $\|u\| \le \frac{M}{2}$. Indeed if $\eta_j = \sum_{i=1}^n \frac{t_i^{(j)}}{q} f_i$ then

$u = \sum h_j\xi_j = (\sum_{j=1}^n h_j(\xi_j - \eta_j)) + (\sum_{j=1}^n h_j\eta_j)$. $\sum_{j=1}^n h_j v_j = 0$ implies that the second term is 0. We may get an estimate on the first term using that $|\sum_{j=1}^n h_j^2| \le n$ and since ξ_j and η_j are in the same parallelepiped $\eta_j + \frac{1}{q}W$ we have that $\|\xi_j - \eta_j\| < nn^3M\frac{1}{q} \le n^4 n^{-c_3}M$. Therefore we get $\|u\| \le n^4 n^{-c_3}Mn^2 = n^{6-c_3}M$ if $c_3 \ge 7$ this implies that $\|u\| \le \frac{M}{2}$.

We prove that $u \ne 0$ with a positive probability by performing the randomization of the vectors ξ_j in a different way. First we randomize the sequence of vectors $v_1, ..., v_m$. This will uniquely determine both the numbers $h_1, ..., h_m$ and the vectors η_j. Now we have to randomize the vectors $\xi_j - \eta_j$. Assume that we have randomized them for $j = 1, ..., m-1$, and assume that $h_m \ne 0$. The distribution of $\xi_j - \eta_j$ is almost uniform in $\frac{1}{q}W$. Since $u - (\xi_m - \eta_m) = \eta_j + \sum_{j=1}^{m-1} h_j\xi_j$ is already fixed, we get that with high probability u is not 0. By the same argument we also get that with high probability u is not in any fixed hyperplane. Therefore if we are getting many (say n^2) independent values of u then with high probability there will be n linearly independent among them and so we have constructed n linearly independent elements in L each of length at most $M/2$.

(1.3)$\to$(1.2). Let $L_0 = L^*$ be the dual lattice of L. We show that if L has an n^{c_3}-unique shortest vector then L_0 has an $n-1$-dimensional sublattice $L' = L_0 \cap F$ where F is an $n-1$ dimensional subspace, so that the distances between the cosets of F intersecting L_0 are at least $n^c \mathrm{bl}(L')$. We prove that it is possible to compute a basis of L', and using that, a shortest vector v in L. (v will be orthogonal to L'.)

Although we give a deterministic algorithm for finding L' (using the algorithm of (1.3) as a black box), it is easier to sketch the idea of a probabilistic one. Assume that we take points of L_0 at random from a parallelepiped whose center is 0 and whose diameter is at most $n^{c'}\mathrm{bl}(L')$, where c' is large with respect to c. (An inductive argument shows that we are able to construct such a parallelepiped.) If we take enough, but still a polynomial number, of random points then at least two of them will be in the same coset of L'. With high probability they will be distinct. Therefore taking all of the differences of the random lattice points we get, among them, a non-zero lattice vector u_1 in $L' = L_0 \cap F$. The most important part of this proof is to show that we are able to decide whether a vector is an L', that is, we are able to select the vector u_1 from the set of differences. If this can be done, then by repeating this procedure many times we will get a sequence $u_1, ..., u_{2n}$. The independence of the vectors u_i implies that there will be n linearly independent among them.

To decide whether u is in L' we consider the lattice L_1 generated by the vectors of L_0 and the vector $\frac{1}{t}u$, where $t \ge n^c$ is a prime number. (It is easy to see that this is indeed a lattice.) Using (1.3) we estimate $\mathrm{bl}(L_0)$ and $\mathrm{bl}(L_1)$. If the estimates do not differ more than allowed by the error, then u is in L'. If the estimate decreases more than that, then $u \notin L'$. This follows from the fact that in the case of $u \in L'$, L_1 will be covered by the cosets of F intersecting L_0, and so $\mathrm{bl}(L_1)$ will be at least the distance of these cosets. In the case $u \notin L'$ there will be new cosets of F which intersect L_1 but not L_0. Between two consecutive cosets intersecting L_0 there will be $t-1$ intersecting only L_1. We get a short basis of L_1 from a short basis of L' and a lattice vector of minimal length connecting two consecutive cosets of F intersecting L_1.

REFERENCES

[Ad] L. Adleman, "On breaking the iterated Merkle-Hellman public key cryptosystem", in: Advances in Cryptology, Proceedings of CRYPTO 82, Plenum Press, New York, 1983, 303-308.

[Br] E.F. Brickell, "Breaking iterated knapsacks", in: Advances in Cryptology, Proceedings of CRYPTO 84, Springer, Berlin, 1985

[LaOd] J.C. Lagarias, A.M. Odlyzko (1983), "Solving low-density subset sum problems", Journal of the Association for Computing Machinery 32 (1985) 229-246.

[LGS] M. Grötschel, L. Lovász, A. Schrijver, "Geometric Algorithms and Combinatorial Optimization", Springer, Algorithms and Combinatorics, 1988

[LG] P.M. Gruber, C.G.Lekkerkerker, "Geometry of Numbers", North-Holland, 1987

[LLL] A.K. Lenstra, H.W. Lenstra, L. Lovász "Factoring polynomials with rational coefficients", Math. Ann. 261, 515-534 (1982)

APPENDIX

Generating Hard Instances of Lattice Problems

M. Ajtai
IBM Almaden Research Center
650 Harry Road, San Jose, CA, 95120
e-mail: ajtai@almaden.ibm.com

We give here the proof of the theorem formulated in the abstract in a detailed but preliminary form. We will prove the theorem for the random variable λ' instead of λ. As we will show the corollary of lemma 1 implies that $P(\lambda = \lambda') \geq 1 - 2^{-cn}$ for some absolute constant c, therefore if we have an algortihm which solves the random problem defined by λ' with probability p then we also have a algorithm which solved the problem defined by λ' with probability $p - 2^{-cn}$. (Although we formulate our theorem for $p = 1/2$ the proof is actually the same for any $p \in [n^{-c'}, 1 - n^{-c'}]$.) We formulate the statement of the theorem again in lemmata 11,12,14 with a slight notational difference.

Notation. $\mathbf{R}$ is the field of real numbers, $\mathbf{Z}$ is the ring of integers, $\mathbf{R}^n$ is the Euclidean space of n dimensional real vectors with the inner porduct $a \cdot b$ and the Euclidean norm $\|a\| = (a \cdot a)^{1/2}$. $\mathbf{Z}^n$ is the set of vectors in $\mathbf{R}^n$ with integer coordinates, we frequently will consider it as a $\mathbf{Z}$-module.

Definitions. 1. If $a_1, ..., a_n$ are linearly independent vectors in an n dimensional Euclidean space E, then we say that the set $\{\sum_{i=1}^n k_i a_i | k_1, ..., k_n$ are integers $\}$ is a lattice in E. We will denote this lattice by $L(a_1, ..., a_n)$. The set $a_1, ..., a_n$ is called a basis of the lattice. The determinant of a lattice L, $\det L$ will be the absolute value of the determinant whose rows are the coordinates of the vectors $a_1, ..., a_n$ in some orthonormal basis of E.

2. If k is an integer then size(k) will denote the number of bits in the binary representation of k, (size$(0) = 1$). If $v = \langle x_1, ..., x_n \rangle \in \mathbf{Z}^n$ then size$(v) = \sum_{i=1}^n$ size(x_i). Our definition implies that for all $v \in \mathbf{Z}^n$, size$(v) \geq n$.

Some of the technical lemmata of the proof are probably known, but we have not yet located an appropriate reference. We give the complete proof for these statements. (Lemmata 1, 2, 4, 8, 9, 10 belong to this category.)

The following lemma and its corollary implies that if λ' is the random variables defined in the abstract, then for a suitable choice of λ (with the distribution defined there) we have $P(\lambda_{n,c_1,c_2} = \lambda_{n,c_1,c_2}) \geq 1 - 2^{-cn}$ where $c > 0$ depends only on c_1 and c_2.

Lemma 1. *There exists a $c > 0$ so that if A is a finite Abelian group with n elements and k is a positive integer and $b = \langle b_1, ..., b_k \rangle$ is a sequence of length k whose elements are chosen independently and with uniform distribution from A, then with a probability of at least $1 - 2^{-ck}$ the following holds:*

Assume that b is fixed and we randomize a $0,1$-sequence $\delta_1, ...\delta_k$, where the numbers δ_i are chosen independetly and with uniform distribution from $\{0,1\}$. For each $a \in A$ let $p_a = P(a = \sum_{i=1}^{k} \delta_i b_i))$. Then

(a) $\sum_{a\in A}(p_a - |A|^{-1})^2 \leq 2^{-2ck}$ *and*

(b) $\sum_{a\in A} |p_a - |A|^{-1}| \leq |A|^{\frac{1}{2}} 2^{-ck}$.

Corollary. *There is a $c' > 0$ so that the following holds. Suppose that $b_1, ..., b_k$,$\delta_1, ..., \delta_k$ are mutally independent random variables with the distributions given in the lemma. Then there is a random variable η with uniform distribution on A so that the random variables $b_1, ..., b_k, \eta$ are mutually independent and*

$P(\eta = \sum_{i=1}^{k} \delta_i b_i) \geq 1 - |A|^{1/2} 2^{-c'k}$.

First we show how can we use the corllary to prove our claim about λ and λ'. We apply the corollary with $A \to \mathbf{Z}^n/(q)$, $k \to m = [c_1 n \log n]$, $b_i \to v_i$, $\delta_i \to \delta_i$, where q, n, v_i, δ_i were given in the definition of λ. Let $\lambda' = \langle v_1, ..., v_{m-1}, -\eta \rangle$. By the assumption of the corollary λ' has the required uniform distribution on A. (We may always take the smallest nonnegative residues mod q.) By the corollary $P(\lambda_{n,c_1,c_2} = \lambda'_{n,c_1,c_2}) \geq 1 - |A|^{1/2} 2^{-c'k}$. $|A| = q^n \leq n^{c_2 n} \leq 2^{c_2 n \log n}$. Therefore if c_1 is sufficiently large with respect to c_2 and c then we have $P(\lambda_{n,c_1,c_2} = \lambda'_{n,c_1,c_2}) \geq 2^{-cn}$.

As a first step in the proof of the lemma we prove the following.

Lemma 2. *There are absolute constants $0 < c_1 < 1$, $0 < c_2 < 1$ so that if A is a finite Abelian group, f is a real-valued function on A and $\sum_{a\in A} f(a) = 0$, then the following holds. For any $b \in A$ we have $\sum_{a\in A} f(a)^2 \geq \sum_{a\in A}(\frac{f(a)+f(a+b)}{2})^2$. Moreover if we pick a random element b of A with uniform distribution then*

$P(c_1 \sum_{a\in A} f(a)^2 \geq \sum_{a\in A}(\frac{f(a)+f(a+b)}{2})^2) > c_2$.

Proof. Our first inequality can be written in the form of

(1) $\frac{1}{2}(\sum_{a\in A} f(a)^2 + \sum_{a\in A} f(a+b)^2) \ge \sum_{a\in A}(\frac{f(a)+f(a+b)}{2})^2$.

This holds since for any fixed a and b the difference of the two sides is $\frac{1}{4}(f(a) - f(a+b))^2 \ge 0$. This also implies that if e.g. $f(a) \ge 0$ and $f(a) \ge 2f(a+b)$ then the differnce of the two sides is at least $\frac{1}{4}(\frac{f(a)}{2})^2 = \frac{f(a)^2}{16}$. We will use this in the proof of the second inequality which can be written in the form. $P(\frac{c_1}{2}(\sum_{a\in A} f(a)^2 + \sum_{a\in A} f(a+b)^2) \ge \sum_{a\in A}(\frac{f(a)+f(a+b)}{2})^2) > c_2$

We will show that there are constants $0 < c_3 < 1$, $0 < c_4 < 1$, $0 < c_5 < 1$ and sets $D \subseteq A$, $F \subseteq A$ so that for each $a \in D$, $b \in F$ we have

(2) $\frac{c_3}{2}(f(a)^2 + f(a+b)^2) \ge (\frac{f(a)+f(a+b)}{2})^2$,

morevover $|F| \ge c_4|A|$ and $\sum_{A\in D} f(a)^2 \ge c_5 \sum_{a\in A} f(a)^2$. This together with the inequality $\frac{1}{2}(f(a)^2 + f(a+b)^2) \ge (\frac{f(a)+f(a+b)}{2})^2$, which holds for all $a, b \in A$ will imply our statement. Indeed $|F| \ge c_4|A|$ implies that it is enough to prove the inequality with the condition $b \in F$. We claim that for each fixed $b \in F$ we have $\frac{c_1}{2}(\sum_{a\in A} f(a)^2 + \sum_{a\in A} f(a+b)^2) \ge \sum_{a\in A}(\frac{f(a)+f(a+b)}{2})^2$. This is a consequence of the fact that for each $a \notin D$ we have (1) and for the ramaining ones we have (ei).

Let $P = \{a \in A | f(a) > 0\}$. We may assume that e.g. $\sum_{a\in P} f(a)^2 \ge \sum_{a\in A\backslash P} f(a)^2$. As we have seen inequality (2) holds if $f(a) \ge 0$ and $f(a) \ge 2f(a+b)$. (This includes the $f(b) < 0$ case too.) Let $\epsilon > 0$ be a small constant. If $|A\backslash P| > \epsilon|A|$ then with $D = P$, $F = A\backslash P$ our conditions are satisfied. Assume now that $|A\backslash P| < \epsilon|A|$. This implies that if $M = |A|^{-1}\sum_{a\in P} f(a) = |A|^{-1}\sum_{a\in A\backslash P} -f(a)$, then $\sum_{a\in A\backslash P} f(a)^2 \ge \epsilon^{-2}|A||M^2|$. We claim that the sets $D = \{a | f(a) > 4M\}$, $F = \{a \in A | f(a) < 2M\}$ meet our requirements. Indeed $\sum_D f(a)^2 \ge \sum_P f(a)^2 - \sum_{P\backslash D} f(a)^2$ and $\sum_{P\backslash D} f(a)^2 \le 16|A|M^2$. If $\epsilon > 0$ is sufficiently small, this implies that $\sum_D f(a)^2 \ge c_5 \sum_A f(a)^2$. $\sum_{(A\backslash F)\cap P} f(a) \le |A||M|$ implies that $|(A\backslash F) \cap P| \le \frac{1}{2}|A|$ and so $|F| \ge |\frac{1}{2} - \epsilon||A|$. *Q.E.D.*(Lemma 2)

Proof of lemma 1. By the Cauchy-Scwarz inequality, (b) follows from (a) so we will prove only (a). Suppose that we randomize both b_i and δ_i independently. Let $f_i(a) = |A|^{-1} - P(a = \sum_{r=1}^{i} \delta_i b_i)$ and let $H(i) = \sum_{a\in A} f_i(a)^2$. We prove that $H(i) \le H(i+1)$ for all $i = 1, ..., k$, $H(1) \le 2$ and with a probability greater than $2^{-c'k}$

(3) there are at least $c''k$ numbers $i \in [1, k]$ so that $H(i+1) < c_1 H(i)$,

where $c' > 0$, $c'' > 0$ and $0 < c_1 < 1$ are absolute constants. This would imply that if we randomize $b_1, ..., b_k$ only, then with a probability of at least $1 - 2^{-\frac{c'}{2}k}$ we get a sequence $b_1, ..., b_k$ so that for the randomization of $\delta_1, ..., \delta_k$ with a probability of at least $1-2^{-\frac{c'}{2}k}$ we get a sequence so that with (3). This clearly will imply the assertion of the lemma.

Assume now that $b_1, ..., b_i$, $\delta_1, ..., \delta_i$ has been randomized. Lemma 2 implies that for the randomization of b_{i+1}, δ_{i+1} the probability of $H(i+1) < c'''H(i)$ is at least c_2 where $c_2 > 0$ is an absolute constant. Therefore Chernoff's inequality implies our statement. *Q.E.D.*(Lemma 1)

Proof of the corollary. We define η separately on each subset of the probability space where $b_1, ..., b_k$ take some fixed value. Assume that these fixed $b_1, ..., b_k$ do not staisfy condition (b) of lemma 1. In this case let η be an arbitrary random variable with uniform distribution. Suppose now that (b) holds Let $\zeta = \sum_{i=1}^{k} \delta_i b_i$. For each $a \in A$ let B_a be the event $\zeta = a$. We choose a $B'_a \subseteq B_a$ for all $a \in A$ so that either $P(B'_a) = |A|^{-1}$ or $P(B'_a) = P(B_a) < |A|^{-1}$. Let η be a random variable with uniform distribution on A so that B'_a implies $\eta = a$. $b_1, ..., b_k$ are mutually independent since with any condition on $b_1, ..., b_k$, η is uniform on A. If we randomize $b_1, ..., b_k$ first then the probability that (b) does not hold is smaller the 2^{-ck} therefore we may assume that η is defined in the second way. In this case however (b) implies that $P(\zeta \neq \eta) < |A|\frac{1}{2}2^{-ck}$. *Q.E.D.*(Corollary)

Definition. 1. If $b_1, ..., b_n \in \mathbf{R}^n$ then $\mathcal{P}(b_1, ..., b_n)$ will denote the parallelepiped $\{\sum_{i=1}^{n} \gamma_i b_i | 0 \leq \gamma_j \leq 1\}$.

2. The minimal height (or width) of $\mathcal{P}(b_1, ..., b_n)$ will be the minimum of the heights belonging to the various faces of $\mathcal{P}(b_1, ..., b_n)$.

Lemma 3. *Suppose that $a_1, ..., a_n$ are vectors in $\mathbf{R}^n$ and $\max_{i=1}^{n} \|a_i\| \leq M$. Then there are linearly independent elements $b_1, ..., b_n \in L(a_1, ..., a_n)$ so that $\max_{i=1}^{n} \|b_i\| \leq (n^3 + \frac{1}{2}n)M$ and the volume of $\mathcal{P}(b_1, ..., b_n)$ is between $\frac{1}{2}(n^3M)^n$ and $2(n^3M)^n$, its surface area is at most $6n(n^3M)^{n-1}$ and its minimal height is at least $\frac{2}{3}n^3M$. Moreover if $a_1, ..., a_n \in \mathbf{Z}^n$ then $b_1, ..., b_n$ can be computed in time polynomial in $\sum_{i=1}^{n} \mathrm{size}(a_i)$.*

Proof. The assumption about the lengths of the basis vectors a_i imply that for each vector v there is a $v' \in L(a_1, ..., a_n)$ so that $\|v - v'\| \leq \frac{1}{2}Mn$. Indeed we may get such a v' by expressing v as a linear combination of the vectors a_i with real coefficients end then rounding off each coefficient to the closest integer. Assume now that $f_1, ..., f_n$ are pairwise orthogonal n-dimensional vectors with length exactly n^3M. For each $i = 1, ..., n$ let b_i be a lattice vector so that $\|f_i - b_i\| \leq \frac{1}{2}nM$. (Clearly this construction which only involves the solution of a linear system of equations and rounding can be completed in polynomial time.) Let $Q = \mathcal{P}(f_1, ..., f_n)$, $Q' = \mathcal{P}(b_1, ..., b_n)$. The distance of each vertex of Q' from the corresponding vertex of Q is at most $\frac{1}{2}n^2M$. Therefore if we enlarge the cube Q from its center by a factor of $1+\frac{1}{2n}$ then it will contain Q'. Q_0 will denote the enlarged cube. In a similar way if we reduce it into a cube Q_1 by the same factor than it will be contained in Q'. volume$(Q_1) \leq$ volume$(Q') \leq$ volume(Q_0) and the inequalities $\frac{1}{2} \leq (1+\frac{1}{2n})^{-n}$ and $(1+\frac{1}{2n})^n \leq 2$ imply our assertion about the volume. $Q_1 \subseteq \mathcal{P}(b_1,...,b_n)$ therefore $\mathcal{P}(b_1,...,b_n)$ contains a sphere of radius at least $\frac{1}{2}(n^3M(1-\frac{1}{2n})) \geq \frac{1}{3}n^3M$ and so the minimal height of $\mathcal{P}(b_1,...,b_n)$ is at least $\frac{2}{3}n^3M$. We get the upper bound on the surface area by estimating the area of each face using the upper bound $(n^3+\frac{1}{2}n)M$ on the lengths of their edge vectors. These yields the upper bound $2n(n^3+\frac{1}{2}n)^{n-1}M^{n-1} = 2n(n^3M)^{n-1}(1+\frac{1}{2n^2})^{n-1} \leq 6n(n^3M)^{n-1}$. *Q.E.D.*(Lemma 3)

Definitions. 1. Suppose that n, m and q are positive integers. Let $\mathcal{V}_{n,m,q}$ be the set of all sequences $v_1,...,v_m$ so that each v_i is an n dimensional vector whose coefficients are nonnegative integers in the interval $[0,q)$. If $s \in \mathcal{V}$ then we will denote by $\Lambda(s,q) = \Lambda(s)$ the set of all sequences of integers $h_1,...,h_m$ so that each coordinate of the n-dimensional vector $\sum_{i=1}^m h_iv_i$ is divisible by q. For any choice of $s \in \mathcal{V}_{n,m,q}$ the set $\Lambda(s)$ is a lattice in $\mathbf{R}^m$.

2. $Z_{n,m,q}$ will be a random variable which takes its values with uniform distribution from $\mathcal{V}_{n,m,q}$. This definition implies that $\Lambda(Z_{n,m,q},q)$ is a random variable which takes its values on certain lattices in $\mathbf{R}^m$. With the notation of the extended abstract $\lambda'_{n,c_1,c_2} = Z_{n,[c_1 n\log n],[n^{c_2}]}$.

Lemma 4. *Assume that $a_1,...,a_n \in \mathbf{R}^n$ are linearly independent vectors, $d_1,...,d_n \in L(a_1,...,a_n)$ are also linearly independent and $\|d_i\| \leq M$. Then there is a basis of $L(a_1,...,a_n)$ consisting of vectors no longer than nM. Moreover if a_i, d_i are integers for $i=1,...,n$ then the required basis can be found in time polynomial in $\sum_{i=1}^n(\text{size}(a_i)+\text{size}(d_i))$*

We prove the lemma by induction on n. The $n=1$ case is trivial. Suppose that our assertion holds for lattices of dimension $n-1$. Let F be the hyperplane generated by $a_1,...,a_{n-1}$ and let $L' = L(a_1,...,a_n) \cap F$. Since $a_1,...,a_n$ are linearly independent we have $L' = L(a_1,...,a_{n-1})$. According to our inductive assumption L' has a basis $d_1,...,d_{n-1}$ with $\max_{i=1}^{n-1}\|d_i\| \neq (n-1)M$. Clearly $d_1,...,d_{n-1},a_n$ is a basis of L. Let a' be the vector that we get from a_n by projecting it orthogonally to F. By expressing a' as a linear combination of the vectors $d_1,...,d_n$, then rounding off the coefficients to the nearest integer we may write a in the form of $w+a''$, where $w \in L(d_1,...,d_{n-1})$ and $\|a''\| \leq \sum_{i=1}^{n-1}\|d_i\| \leq (n-1)M$. Therefore $d_1,...,d_i,a_n-w$

is a basis of $L(a_1, ..., a_n)$ and $\|a_n - w\| \leq (\|(a_n - a'\|^2 + \|a' - w\|^2)^{1/2} \leq (\|a_n\|^2 + \|a''\|)^{1/2} \leq (1 + (n-1)^2)^{1/2} M < nM$ implies that every element of this basis is of length at most $(n-1)M$. The inequality $|a_n - w| \leq (n^2 - 2n)^{1/2} M < nM$ shows that even if we compute a' only approximately with a precision greater than, say, $\frac{1}{n^2} M$ the vector $a_n - w \in L$ that we get from this approximate value will be shorter than $(n-1)M$. Q.E.D.(4)

We need the following lemma to show that if a parallelepiped W is not very skewed and it is large with respect to $\mathrm{bl}(L)$ (e.g. the one constructed in lemma 3), then the number of lattice points in all of the parallelepipeds $b + W$, $b \in \mathbf{R}^n$ is about the same and is roughly proportional to the volume of the parallelepiped. Moreover for any fixed hyperplane F the number of lattice points in $F \cap (b + W)$ is small with respect to the number of lattice points in $b + W$.

Lemma 5. *Assume that $L = L(a_1, ..., a_n)$ is a lattice in $\mathbf{R}^n$, where $|a_i| \leq M$, $i = 1, ..., n$ and $g_1, ..., g_n$ are linearly independent vectors in $\mathbf{R}^n$ (not necessarily in L) and $b \in \mathbf{R}^n$. Let k_0 resp. k_1 be the number of lattice points in the closed set $b + \mathcal{P}(g_1, ..., g_n)$ resp. in its interior. Let H be the minimal height, let V be the volume and let S be the surface area of $\mathcal{P}(g_1, ..., g_n)$). Then*

(a) $(\det L)^{-1}(1 - \frac{2Mn}{H})^n V \leq k_j \leq (\det L)^{-1}(1 + \frac{2Mn}{H})^n V$, $j = 0, 1$

(b) If F is a hyperplane then the number of lattice points in $F \cap (b + \mathcal{P}(g_1, ..., g_n))$ is at most $2SMn(1 + \frac{2Mn}{H})^{n-1}(\det L)^{-1}$.

Proof. (a) Let $W = b + \mathcal{P}(g_1, ..., g_n)$, let W' be the set that we get from W by enlarging it from its center by a factor of $1 + \frac{2Mn}{H}$ and W'' be the set that we get from it by reducing it by $1 - \frac{2Mn}{H}$. Let B be the set of all parallelepipeds of the form $v + \mathcal{P}(a_1, ..., a_n)$, where v is a lattice point and $(v + \mathcal{P}(a_1, ..., a_n) \cap W$ is non-empty. The definitions of W', W'' imply that every element of B is contained in W' and every element of B intersecting W'' is contained in W. Therefore we get the upper bounds from the fact that the number of elements of B contained in W' can be at most $\mathrm{volume}(W')/\det(L)$. We get the lower bound on k_0 in the following way. Let D be the set of those elements of B that intersect W''. Clearly $|D| \leq k_0$. The definition of W'' implies that the elements of D cover W'' so $|D| \geq \mathrm{volume}(W'')(\det L)^{-1}$. To get the lower bound on k_1, we may repeat our argument for each $\epsilon > 0$ with W''_ϵ instead of W'' where we get W''_ϵ by reducing W with a factor of $1 - \frac{2Mn}{H} - \epsilon$. This way the elements of the set D will be in the interior of W. Taking the limit for all of the resulting lower bounds for k_1 we get (a).

(b). Let G be the set of those elements of B which intersect F. The defintion of W' implies that the distance of $F \backslash W'$ from $F \cap W$ is at least Mn. (Any pair of points from them are separated by a pair of corresponding parallel faces of W and W' whose distance is at least Mn.) Therefore if π is the orthogonal projection of $\mathbf{R}^n$ to F and $T \in G$ then $\pi(T)$ is in $F \cap W'$. Consequently each $T \in G$ is contained in the body that consist of all points x with $\pi x \in W' \cap F$ whose distance from F is at most Mn. The volume of this body is 2area$(W' \cap F)Mn$ and area$(W' \cap F)$ is at most the surface area of W' which implies our inequality. *Q.E.D.*(Lemma 5)

Definition. If $a_1, ..., a_n \in \mathbf{R}^n$ are linearly independent vectors then $\mathcal{P}^-(a_1, ..., a_n)$ will denote the set $\{\sum_{i=1}^n \gamma_i b_i | 0 \le \gamma_j < 1\}$.

Lemma 6. *Assume that $L = L(a_1, ..., a_n)$ is a lattice in $\mathbf{R}^n$, $\|a_i\| \le M$ for $i = 1, ..., n$, $b_1, ..., b_n$ are linearly independent elements of L, $\|b_i\| \le Y$ for $i = 1, ..., n$, V is the volume, S is the surface area and H is the minimial height of $\mathcal{P}(b_1, ..., b_n)$, q is a positive integer and the following inequalities hold*

(i) $\frac{M}{H} \le \frac{1}{4n^4}$

(ii) $5SMn \le V$.

Suppose further that ξ is a random variable that takes its values with uniform distribution on the set R of lattice points of $\mathcal{P}^-(b_1, ..., b_n)$. Then there are random variables ζ, η with $\xi = \zeta + \eta$ so that ζ has uniform distribution on $E = \{\sum_{i=0}^n \kappa_i b_i | \kappa_i \in \{0, \frac{1}{q}, ..., \frac{q-1}{q}\}, i = 1, ..., n\}$, and for each fixed $t \in E$ the conditional distribution of η with the condition $\zeta = t$ meets the following requirements:

(a) $P(\eta \in \mathcal{P}^-(\frac{1}{q}b_1, ..., \frac{1}{q}b_n) | \zeta = t) > 1 - \frac{1}{n^2}$

(b) *for any fixed hyperplane F in $\mathbf{R}^n$, $P(\eta \in F | \zeta = t) < 1/2$*

Proof. Let T be the set of all sequences $t_1, ..., t_n$ so that $t_i \in \{0, 1, ..., q-1\}$ and for each $t = \langle t_1, ..., t_n \rangle \in T$ let $W_t = \mathcal{P}(\frac{1}{q}b_1, ..., \frac{1}{q}b_n) + \sum_{i=1}^n \frac{t_i}{q} b_i$. Lemma 5 gives the following estimate on w_t the number of lattice points in W_t:

$(\det L)^{-1}(1 - \frac{2Mn}{H})^n V \le w_t \le (\det L)^{-1}(1 + \frac{2Mn}{H})^n V$.

Inequality (i) implies that $1 - \frac{1}{3n^2} \le (1 - \frac{2Mn}{H})^n \le 1 \le (1 + \frac{2Mn}{H})^n \le 1 + \frac{1}{3n^2}$ and so

(4) $(1 - \frac{1}{3n^2})(\det L)^{-1} V \le w_t \le (1 + \frac{1}{3n^2})(\det L)^{-1} V$.

Let $\alpha = [(1 - \frac{1}{3n^2})(\det L)^{-1} V]$ and for each $t \in X$ let W'_t be an arbitrary but fixed subset of W_t with exactly α elements. For the defintion of ζ we will use another random variable ρ which is independent of ξ and has uniform distribution on E.

Suppose that both ξ and ρ has been randomized. If $\xi \in \bigcup_{t\in T} W'_t$ then there is a unique $t = \langle t_1, ..., t_n\rangle \in T$ with $\xi \in W'_t$. In this case let $\zeta = \sum_{i=1}^n \frac{t_i}{q} b_i$. If ξ is outside of $\bigcup_{t\in T} W'_t$ then let $\zeta = \rho$. Since $|W'_t|$ does not depend on t and ξ, ρ are independent, we have that ζ has uniform distribution on E.

(a) (4) and the definition of α implies that the probability of $\xi \in \bigcup_{t\in T} W'_t$ is greater than $1 - \frac{1}{n^2}$. In this case the definition of ζ implies that if $\xi \in W_t$ then then $W_t = \zeta + \mathcal{P}(\frac{1}{q}b_1, ..., \frac{1}{q}b_n)$, and so $\eta = \xi - \zeta \in \mathcal{P}(\frac{1}{q}b_1, ..., \frac{1}{q}b_n)$.

(b) According to (a) it is enough to show that $P(\eta \in F | \xi = t, \xi = \zeta) < \frac{1}{2} - \frac{1}{n^2}$. By Lemma 5 and inequalities (i),(ii), the number of lattice points on $F \cap W'_t \subseteq F \cap W_t$ is at most $\frac{2}{5} V(\det L)^{-1}$. Therefore the definition of $\alpha = |W_t|$ and the fact that with the condition $\xi = \zeta$, ζ is uniform on W_t implies (b). *Q.E.D.*(Lemma 6)

Lemma 7. *Assume that $a_1, ..., a_n \in \mathbf{R}^n$ are linearly independent. Then, for each $b \in \mathbf{R}^n$, there is a unique $b' \in \mathcal{P}^-(a_1, ..., a_n)$ so that $b - b' \in L(a_1, ..., a_n)$ moreover, if $b \in \mathbf{Z}^n$ and $a_i, \in \mathbf{Z}^n$, $i = 1, ..., n$ then b' can be computed in polynomial time in* $\mathrm{size}(b) + \sum_{i=1}^n \mathrm{size}(a_i)$

Proof. We express b as a linear combination of the vectors a_i then take the integral part of the coefficients. Assume that we get the vector $v = \sum_{i=1}^n r_i a_i$. $b' = b - v$ will satisfy our requirement. The uniqueness of b' is trivial. *Q.E.D.*(Lemma 7)

Definition. Assume that $a_1, ..., a_n, b$ are as in lemma 6. We will denote the unique b' described in the lemma by $b_{(\mathrm{mod}\ \ a_1, ..., a_n)}$.

Lemma 8. *For all $c_1 > 0$ there is a $c_2 > 0$ so that the following holds. Assume that $d_1, ..., d_n$ are linearly independent vectors in $\mathbf{Z}^n$, $\sigma \geq n$ and $a_1, ..., a_n \in L = L(d_1, ..., d_n)$ is a set of linearly independent vectors as well, with $\max_{i=1}^n \|a_i\| \leq 2^{\sigma^{c_1}}$ and $\max_{i=1}^n \|d_i\| \leq 2^{\sigma^{c_1}}$. Suppose further that $\mu_1, ..., \mu_n$ are independent random variables which take their values with uniform distribution on the integers in the interval $[0, 2^{\sigma^{c_2}}]$. Let $\chi = (\sum_{i=1}^n \mu_i d_i)_{(\mathrm{mod}\ \ a_1, ..., a_n)}$. Then the distribution of χ on the points of $L \cap \mathcal{P}^-(a_1, ..., a_n)$ is almost uniform in the following sense:*

if for each $v \in \mathcal{P}^-(a_1, ..., a_n)$, $p_v = P(\chi = v)$ and k is the number of lattice points in $\mathcal{P}^-(a_1, ..., a_n)$, then

$\sum_{v\in\mathcal{P}^-(a_1, ..., a_n)} |p_v - \frac{1}{k}| \leq 2^{-\sigma^{c_1}}$.

Proof. We will need the following observations in the proof. For each real number α let $W_\alpha = \mathcal{P}^-(\alpha d_1, ..., \alpha d_n)$. Since $d_1, ..., d_n$ is a basis of L we have that if α is a positive integer then the number of lattice points in W_α is α^n. Since the volume of W_1 is at least 1, (the value of a nonzero determinant with integer entries) and the area of any face of it is at most $\prod_{i=1}^n \|d_i\|$ we have that the minimal height H of W_1 is at least $(\prod_{i=1}^n d_i)^{-1} \geq 2^{-\sigma^{c_1+1}}$.

Let $t = [\sigma^{c_2}]$. Let X' be the set of all parallelepipeds J of the form $J = u + \mathcal{P}^-(a_1, ..., a_n)$ with $u \in L$ and $J \cap W_t \neq \emptyset$. Let X be the set of all sets $J \in X'$ with $J \subseteq W_t$. If we enlarge W_t from its center by a factor of $\gamma = 1 + \frac{22^{\sigma^{c_1+1}}}{tH}$ then the resulting set W' will contain every element of X'. By lemma 5 the number of lattice points in $W' - W$ is at most $(\det L)^{-1}((1 + \frac{22^{\sigma^{c_1}} n}{tH})^n \gamma^n t^n - (1 - \frac{22^{\sigma^{c_1}} n}{tH})^n t^n)$. If c_2 is sufficiently large with respect to c_1 then this is at most $2^{-\sigma^{2c^1+1}} t^n$.

Let τ be the unique element of X' containing χ. The elements of X are disjoint, so $p_v = (\sum_{J \in X} P(\chi = v | \tau \in J) P(\tau \in J)) + P(\chi \in V | \tau \notin \bigcup X) P(\tau \notin \bigcup X)$. The distribution of χ is uniform on $\mathcal{P}^-(a_1, ..., a_n)$ with the condition $\chi \in J$ for each fixed $J \in X$ therefore the first term is $\frac{1}{k} \frac{|X| k}{t^n}$ which does not depend on v.

The second term is at most $P(\tau \notin \bigcup X)$. This is smaller than the number of lattice points in $\bigcup X' \setminus \bigcup X$ divided by t^n that is smaller than $2^{-\sigma^{2c_1+1}}$. Since the number of lattice points in $\mathcal{P}^-(a_1, ..., a_n)$ is at most volume$(a_1, ..., a_n)(\det L)^{-1} \leq 2^{\sigma^{c_1+1}}$ this implies our statement. *Q.E.D.*(Lemma 8)

Definitions. 1. c_M will denote a fixed positive real number so that *for all* $n = 1, 2, ...$ and for all lattice L in R^n there exists a $v \in L$, $v \neq 0$ with $\|v\| \leq c_M n^{\frac{1}{2}} (\det L)^{\frac{1}{n}}$. Minkowski's theorem about closed, convex, central-symmetric bodies applied to a sphere implies the existence of such a constant.

2. If L is a lattice in $\mathbf{R}^n$ then unit(L) will denote the number $(\det L)^{\frac{1}{n}}$.

3. Suppose that L is a lattice in $\mathbf{R}^n$ and H is a k-dimensional subspace of R^n so that $L' = H \cap L$ is a (k-dimensional) lattice in H. The factor lattice L/L' will be the lattice that we get from L by orthogonally projecting it onto $H^\perp$. (We have to prove that L/L' is indeed a lattice, that is, it has a basis consisting of $n - k$ elements (over the integers). We may pick a basis $a_1, ..., a_n$ for L so that $a_1, ..., a_k$ is in L' (the assumption that $H \cap L$ is a k-dimensional lattice implies the existence of such a basis). If π is the orthogonal projection of $\mathbf{R}^n$ onto $H^\perp$ then $\pi a_{k+1}, ..., \pi a_n$ will be the required basis of L/L'.)

Lemma 9 . *Suppose that L is a lattice in $\mathbf{R}^n$ and $K > 0$. Then either L has a factor lattice L_1 with* unit$(L_1) \geq K$ *or L_1 has a basis whose each vector is not longer than $c_M K \sum_{i=1}^n i^{\frac{1}{2}}$.*

Proof. It is enough to prove the lemma for $K = 1$ since we may replace L by $\frac{1}{K}L$. We prove the lemma by induction on n. For $n = 1$, unit(L) is the length of a shortest vector and so $c_M \geq 1$, therefore our statement trivially holds.

Assume now that the lemma holds for $n-1$. If unit$(L) \geq 1$, then our statement holds with $L_1 = L$. Suppose that unit$(L) < 1$, then by Minkowski's theorem there is a $v \in L$, $v \neq 0$ so that $\|v\| \leq c_M n^{1/2}\text{unit}(L) < c_M n^{1/2}$. Let W be the subspace orthogonal to v. Let L_v be the one dimensional lattice generated by v and L_1 be the factor lattice L/L_v. According to the inductive assumption either L_1 has a factor lattice L_1' with unit$(L_1') \geq 1$ or L_1 has a basis B' with vector lengths no longer then $c_M \sum_{i=1}^{n-1} i^{1/2}$. In the former case we are done since a factor lattice of L_1 is also a factor lattice of L. In the latter case we may construct a basis B of L in the following way. B will contain v and for each element $b' \in B$ we take an element b of L so that $b - b' \neq 0$ is in the one dimensional vectorspace generated by v and $\|b - b'\|$ is minimal with this condition. We may pick such a b from those elements whose image is b' under the orthogonal projection of L onto $v^{\perp}$. Moreover we may asssume that $\|b - b'\| \leq \|v\|$. Therefore the length of each element of B is at most $\|v\| + c_M \sum_{i=1}^{n-1} i^{1/2} < c_M \sum_{i=1}^{n} i^{1/2}$.

Definitions. 1. With each $v \in \mathbf{R}^n$ we associate a linear functional ϕ_v on $\mathbf{R}_n$ defined by $\phi_v(u) = v \cdot u$, for all $u \in \mathbf{R}^n$, where $\cdot$ is the inner product defined on $\mathbf{R}^n$ in the usual way.

2. Let L be a lattice in $\mathbf{R}^n$. We define a subset $L^* \subseteq \mathbf{R}^n$ in the following way: $v \in L^*$ iff the functional ϕ_v takes integer values on every element of L. It is easy to see that L^* is a lattice in $\mathbf{R}^n$. If $a_1, ..., a_n$ is basis of L then the set of those functionals which take the value 1 on exactly one a_i and the value 0 on all of the others form a basis of L^*. This is called the dual basis of $a_1, ..., a_n$). This construction also shows that $(\det L)(\det L^*) = 1$ and so unit(L)unit$(L^*) = 1$.

3. If L is a lattice in $\mathbf{R}^n$, then sh(L) will denote the length of the shortest non-zero vector in L and bl(L) will be the smallest real number K so that L has a basis $a_1, ..., a_n$ with $\max_{i=1}^{n} \|a_i\| = K$.

Lemma 10. *If L is a lattice in $\mathbf{R}^n$ then*
$1 \leq \text{sh}(L^*)\text{bl}(L) \leq c_M^2 n^{1/2} \sum_{i=1}^{n} i^{1/2} \leq cn^2$, *where c is an absolute constant.*

Proof of the lower bound. Assume that $v \in L^*$, $\|v\| = \text{sh}(L^*)$ and $a_1, ..., a_n$ is a basis of L with $\max_{i=1}^{n} \|a_i\| = \text{bl}(L)$. Since $v^{\perp}$ is an $n-1$-dimensional subspace, there is an a_j so that a_j and v are not orthogonal and so $a_j \cdot v \neq 0$. By the defintion of L^*, $a_j \cdot v$ is an integer and therefore $|a_j \cdot v| \geq 1$ and so $\|a_j\|\|v\| \geq 1$ and bl(L)sh$(L^*) \geq 1$.

Proof of the upper bound. For the proof we need the following trivial observation: the dual space of the factorspace (L/L') is a subspace of L^*. Indeed assume that $u \in (L/L')^*$. Then u is a vector in $\mathbf{R}^n$, it is orthogonal to L' and for each $v \in L/L'$, $u \cdot v$ is an integer. Let $w \in L$ be arbitrary. By the definition of L/L', w can be written in the form of $v + v'$, where $v \in L/L'$ and v' is in the real vectorspace generated by L'. Therefore $u \cdot w = u \cdot v + u \cdot v' = u \cdot v$ is an integer and so $u \in L^*$.

Suppose that $c_M K \sum_{i=1}^{n} i^{\frac{1}{2}} = \mathrm{bl}(L)$. Then by Lemma 9 for any $K' < K$, $K' > 0$ there is a factor lattice L_1 of L so that $\mathrm{unit}(L_1) \geq K'$. Assume that the dimension of L_1 is $m \leq n$. Since $\mathrm{unit}(L_1^*)\mathrm{unit}(L_1) = 1$, we have $\mathrm{unit}(L_1^*) \leq \frac{1}{K'}$ and so Minkowski's theorem implies that there is a non-zero vector $v \in L_1^*$ so that $\|v\| \leq c_M \frac{1}{K'} m^{1/2}$. As we have seen $L_1^* \subseteq L^*$, therefore $\mathrm{sh}(L^*)\mathrm{bl}(L) \leq \frac{K}{K'} c_M n^{1/2} c_M \sum_{i=1}^{n} i^{i/2}$. This holds for any $K' < K$, which implies our upper bound. *Q.E.D.*(Lemma 10)

Lemma 11 . *There are absolute constants c_1, c_2 with the following property. Suppose that there is a probabilistic polynomial time algorithm $\mathcal{A}$ which given a value of the random variable $Z_{n,[c_1 n \log n],[n^{c_2}]}$ as input with a probability of at least $1/2$ outputs a vector of $\Lambda(Z_{n,[c_1 n \log n],[n^{c_2}]})$ of length at most n. Then there is a probabilistic algorithm $\mathcal{B}$ which given the linearly independent vectors $a_1, ..., a_n \in \mathbf{Z}^n \subseteq \mathbf{R}^n$ as input will output an integer z in time polynomial in $\sigma = \sum_{i=1}^{n} \mathrm{size}(a_i)$ so that if v is the shortest non-zero vector in $L(a_1, ..., a_n)$ then with a probability greater than $1 - 2^{-\sigma}$ we have $z \leq \|v\| \leq n^{c_3} z$.*

Remarks. 1. The probability $1/2$ in the theorem can be replaced by n^{-c}. This will increase the running time of $\mathcal{B}$ by a factor of at most n^c but does not affect the constants c_1 and c_2.

Lemma 12. *There are absolute constants c_1, c_2, c_3 with the following property. Suppose that there is a probabilistic polynomial time algorithm which given a value of the random variable $Z_{n,[c_1 n \log n],[n^{c_2}]}$ as input with a probability of at least $1/2$ outputs a vector of $\Lambda(Z_{n,[c_1 n \log n],[n^{c_2}]})$ of length at most n. Then there is a probabilistic algorithm which given the linearly independent vectors $a_1, ..., a_n \in \mathbf{Z}^n \subseteq \mathbf{R}^n$ as input will output a basis $b_1, ..., b_n$ of $L(a_1, ..., a_n)$ in time polynomial in $\sigma = \sum_{i=1}^{n} \mathrm{size}(a_i)$ so that if $d_1, ..., d_n$ is an arbitrary set of linearly independent vectors in $L(a_1, ..., a_n)$ then with a probability greater than $1 - 2^{-s}$ we have $\max_{i=1}^{n} \|b_i\| \leq n^{c_3} \max_{i=1}^{n} \|d_i\|$.*

Lemma 13 . *There are absolute constants c_1, c_2 with the following property. Suppose that there is a probabilistic polynomial time algorithm $\mathcal{A}$ which given a value of the random variable $Z_{n,[c_1 n \log n],[n^{c_2}]}$ as input with a probability of at least $1/2$ outputs a vector of $\Lambda(Z_{n,[c_1 n \log n],[n^{c_2}]})$ of length at most n. Then there is a probabilistic algorithm which given two sets of linearly independent vectors $a_1, ..., a_n \in \mathbf{Z}^n \subseteq \mathbf{R}^n$, $u_1, ..., u_n \in L(a_1, ..., a_n)$ as input will output n linearly independent vector $b_1, ..., b_n$ of $L(a_1, ..., a_n)$ in time polynomial in $\sigma = \sum_{i=1}^{n}(\text{size}(a_i) + \text{size}(u_i))$ so that with a probability of greater than $1 - 2^{-\sigma}$, either $b_1, ..., b_n$ meets the requirement of lemma 12 or $\max_{i=1}^{n} \|b_i\| \leq \frac{1}{2} \max_{i=1}^{n} \|u_i\|$.*

Proof. First we describe the algorithm.

Using lemma 3 with $a_i \to u_i$ and $M \to \max_{i=1}^{n} \|u_i\|$ we construct a set of linearly independent vectors $v_1, ..., v_n \in L(a_1, ..., a_n)$ so that $\max_{i=1}^{n} \|v_i\| \leq (n^3 + \frac{1}{2}n)M$ and for the volume V, surface area S and minimal height H of $\mathcal{P}(v_1, ..., v_n)$ we have certain bounds. Now we take a random point of $L(a_1, ..., a_n)$ with almost uniform distribution in $W = \mathcal{P}^{-}(v_1, ..., v_n)$. More precisely lemma 8 guarantees that we can compute in polynomial time the value of a random variable χ which takes its values from R, the set of lattice points in W and has the property $\sum_{v \in R} |P(\chi = v) - \frac{1}{R}| \leq 2^{-n^{c'}}$. We may write χ in the form of $\sum_{i=1}^{n} \beta_i v_i$ where $0 \leq \beta_i < 1$. By solving a system of linear equations we may find the rational numbers β_i in polynomial time. Let $q = [n^{c_2}]$ and $t_i = [q\beta_i]$, $i = 1, ..., n$ and $\sigma = \langle t_1, ..., t_n \rangle$. Repeating this procedure with independent values of χ we get a sequence of values χ_j, σ_j, $j = 1, ..., m$, where $m = [c_1 n \log n]$. Let L_1 be the lattice of m dimensional integer vectors $\langle h_1, ..., h_m \rangle$ so that $q | \sum_{i=1}^{m} h_i \sigma_i$. Now we apply our probabilistic algorithm $\mathcal{A}$, whose existence was assumed, with the lattice L_1 and in polynomial time we either get a vector $s_1 \in L_1$ with $\|s_1\| \leq n^2$ or we recognize that the algorithm failed to produce the required result. In this case let $s_1 = 0 \in \mathbf{R}^m$. In either case $s_1 = \langle z_1, ..., z_m \rangle$ is a sequence of integers. Next we find the vector $f_1 = \sum_{i=1}^{m} z_i \chi_i$ and $g_1 = (f_1)_{(\text{mod } v_1, ..., v_n)}$. (That is g_1 is the unique element of $\mathcal{P}^{-}(v_1, ..., v_n)$ with $f_1 - g_1 \in L(v_1, ..., v_n)$). We repeat this whole procedure $3n$ times and get a sequence of vectors $g_1, ..., g_{3n}$. Let G be the set of those vectors g_i, $i = 1, ..., 3n$ which are nonzero and are shorter than $(n^3 + \frac{1}{2}n)M\frac{n}{q} \leq \frac{M}{2}$. We try to select n linearly independent vectors from G. If we succeed then the set of these vectors $b_1, ..., b_n$ is the output. If we do not succeed then we apply the algorithm given in lemma 4 with $d_i \to u_i$ and we get a basis $b_1, ..., b_n$ with $\max_{i=1}^{n} \|b_i\| \leq n \max_{i=1}^{n} \|u_i\|$. In this case the sequence $b_1, ..., b_n$ defined in this shorter alternative way will be the output.

Now we prove the correctness of our algorithm. If for any basis $d_1, ..., d_n$ of $L(a_1, ..., a_n)$ we have $\max_{i=1}^n \|u_i\| \leq \max_{i=1}^n n^{c_3+1}\|d_i\|$ then the vectors $b_1, ..., b_n$ defined by the short alternative way using lemma tv (described at the very end of the algorithm) satisfy the requirements of the lemma. Therefore we may assume in the following that there is a basis $d_1, ..., d_n \in L(a_1, ..., a_n)$ so that $\max_{i=1}^n \|u_i\| > \max_{i=1}^n n^{c_3+1}\|d_i\|$.

When we sart the algorithm we have n linearly independent vector $u_1, ..., u_n$ in the lattice $L(a_1, ..., a_n)$. We try to construct from them an other set of vectors whose maximal norm is smaller by a factor of two. To start our construction we replace $u_1, ..., u_n$ by an other set of vectors $v_1, ..., v_n$ which are not essentially longer (only by about a factor of n^3) but whose prallelepiped $\mathcal{P}(v_1, ..., v_n)$ is as close to a cube as possible. Lemma 3 with $a_i \to u_i$ gives such a construction. Therefore we get a set of vectors $v_1, ..., v_n \in L(a_1, ..., a_n)$ so that if $\max_{i=1}^n \|u_i\| = M$ then $\max_{i=1}^n \|v_i\| \leq (n^3 + \frac{1}{2}n)M$ and if V is the volume, S is the surface area and H is the minimal height of $\mathcal{P}(v_1, ..., v_n)$ then $\frac{1}{2}(n^3M)^n \leq V \leq 2(n^3M)^n$, $S \leq 6n(n^3M)^{n-1}$ and $H \geq \frac{2}{3}n^3M$. The role of these inequalities will be that they guarantee that if we take parallelepipeds $x + \mathcal{P}(v_1, ..., v_n)$ for different elements $x \in \mathbf{R}^n$ then the number of lattice points in them will be about the same in the sense that the differnces will be small relative to the total number of lattice points. Another consequence of the inequalities that there will be realtively few lattice points in a parallelepiped of this type which lies on any single fixed hyperplane. These properties do not necessarily hold if the the parallelepiped is either small relative to the maximal length of any basis of the lattice, or it is very much distorted e.g. one of its heights is very small. Actually we will need these properties in the case of parallelepipeds of the form $\mathcal{P}(\frac{1}{q}v_1, ..., \frac{1}{q}v_n)$ where $q = [n^{c_2}]$.

For the next step we need the following observation. Lemma 8 gives a random variable χ which has only an almost uniform distribution on the set R. However in our proof we may assume that the distribution of χ is actually uniform. Indeed we know that $\sum_{v \in R} |P(\chi = v) - \frac{1}{|R|}| \leq 2^{-n^{c'}}$. This means that there is a random variable χ' so that χ' has uniform distribution and $P(\chi \neq \chi') \leq 2^{-n^{c'}}$. Therefore we may assume that we work with χ' and with high probability its value is the same as χ. This will lead only to a $2^{-n^{c'}}$ failure rate in the algorithm. (Even if the failure rate would be higher we may decrease it exponentially by repeating the algorithm).

Assume now that the vectors $g_1, ..., g_j$ has been already constructed for some $0 \leq j < c_4n$ and we now start the conrstructcion of g_{j+1}. Let G_j be a maximal subset of linearly independent vectors of $\{g_1, ..., g_j\}$ with the property that for all

$g \in G$ we have $g \neq 0$ and $\|g\| < (n^3 + \frac{1}{2}n)M\frac{n}{q}$. Let F be a hyperplane in $\mathbf{R}^n$ containing G_j. We will prove that (for the randomizations involved in the selection of g_{j+1} only and considering F as fixed), we have

(5) $\quad P(g_{j+1} \notin F \text{ and } \|g_{j+1}\| \leq (n^3 + \frac{1}{2}n)M\frac{n}{q}) \geq \frac{1}{2} - \frac{2m}{n^2} \geq \frac{1}{3}.$

First we notice that (5) implies the lemma. Indeed (5) and Chernoff's inequality imply that the set G as defined in the algorithm will contain n elements.

Now we prove (5). First we prove that

(6) $\quad P(\|g_{j+1}\| \leq (n^3 + \frac{1}{2}n)M\frac{n}{q}) \geq 1 - \frac{m}{n^2}.$

We apply lemma 6 with $b_1 \to v_1, ..., b_n \to v_n$ and $\xi \to \chi$. (As we have explained above we may assume that χ has uniform distribution on the set of lattice points in $\mathcal{P}^-(v_1, ..., v_n)$). According to lemma 6, χ can be written in the form of $\zeta + \eta$ where ζ is uniform on E and we also know something about the conditional distribution of η. We claim that if we repeat this process and get the sequences $\zeta_1, ..., \zeta_m$, $\eta_1, ..., \eta_m$ then with a probability of at least $1 - \frac{m}{n^2}$,

(7) $\quad \zeta_1 = \sigma_1, ..., \zeta_m = \sigma_m$ and $\|\eta_i\| \leq n^2(n^3 + \frac{1}{2}M)\frac{n}{q}$ for $i = 1, ..., m$.

Indeed, (a) of lemma 6 implies that for all $i = 1, ..., m$ with a probability of at least $1 - \frac{1}{n^2}$, we have $\zeta_i = \sigma_i$ and the vector η_i is inside the parallelepipedon $\mathcal{P}(\frac{1}{q}v_1, ..., \frac{1}{q}v_n)$ and so the upper bound on the vectors $v_1, ..., v_n$ imply the required upper bound on η_i. The vector $z = \langle z_1, ..., z_n \rangle$ is no longer than n. We show that (7) implies that $\|g_j\| \leq (n^3 + \frac{1}{2}n)M\frac{n}{q}$. Indeed by (7) the definition of f_j we have $f_j = \sum_{i=1}^m z_i\chi_i = (\sum z_i\zeta_i) - \sum z_i\eta_i = (\sum z_i\sigma_i) - \sum z_i\eta_i$. We know that either $z = 0$ or we get z as the output of $\mathcal{A}$. In either case we have $\|z\| \leq n$ and $q | \sum_{i=1}^m z_i\sigma_i$. The latter relation and the definition of σ implies that $\sum_{i=1}^m z_i\zeta_i \in L(v_1, ..., v_n)$ and so $g_j = (f_j)_{(\text{mod } v_1, ..., v_n)} = -\sum_{i=1}^m z_i\eta_i \leq (n^3 + \frac{1}{2}n)M\frac{n}{q}$ which completes the proof of (6).

We continue the proof of (5) by showing that

(8) $\quad P(g_{j+1} \notin F) \geq \frac{1}{2} - \frac{2m}{n^2}.$

As we have seen the probability of $\sigma_1 = \zeta_1, ..., \sigma_m = \zeta_m$ is at least $1 - \frac{m}{n^2}$. Therefore it is enough to show that if we change our algorithm so that instead of σ_i, $i = 1, ..., m$ we use ζ_i, $i = 1, ..., m$ in the defintion of the vector $h_1, ..., h_m$ and so in the defintion of z, f_{j+1} and g_{j+1} then (8) holds if we change the right-hand side into $\frac{1}{2} - \frac{m}{n^2}$.

We may randomize all of the random variables $\chi_1, ..., \chi_m$ by first randomizing $\zeta_1, ..., \zeta_m$ and then $\eta_1, ..., \eta_m$. Since the definition of the number h_i depend only on ζ_i (and not on η_i), the values $\zeta_1, ..., \zeta_n$ already determine whether algorithm $\mathcal{A}$ succeeds in finding a short vector. The probability (for the randomization of $\zeta_1, ..., \zeta_m$ only) that it does not succeed is at most $1/2$. Therefore it is sufficient to show that for any possible values $t^{(1)}, ..., t^{(m)}$ of the sequence $\zeta_1, ..., \zeta_m$, if $\zeta_1 = t^{(1)}, ..., \zeta_n = t^{(m)}$ implies that if $\mathcal{A}$ finds a short vector then

(9) $\quad P(g_{j+1} \notin F | \zeta_1 = t^{(1)}, ..., \zeta^{(m)} = t^{(m)}) \geq \frac{1}{2} - \frac{2m}{n^2}$.

Assume now that $\zeta_1 = t^{(1)}, ..., \zeta^{(m)} = t^{(m)}$ for such a sequence $t^{(1)}, ..., t^{(n)}$. Since $\mathcal{A}$ finds a short vector we have $z \neq 0$. Let ρ be the smallest positive integer with $z_\rho \neq 0$. We consider ρ as a random variable, it determined by ζ_i and by the randomization included in $\mathcal{A}$. Now we randomize η_ρ. (b) of Lemma 6 implies for any fixed r we have $P(\eta_\rho \in F | \zeta = t^{(1)}, ..., \zeta^{(m)} = t^{(m)}, \rho = r) < 1/2$ Since this is true for any choice of r we have (9). *Q.E.D.*(Lemma 13)

Proof of lemma 12. Assume that $\max \|a_i\| = M$. If we apply the algorithm whose existence is stated in lemma 4 then we either get the required output immediately or get a linearly independent system $u_1^{(1)}, ..., u_n^{(1)}$ in polynomial time with $\max_{i=1}^n \|u_i^{(1)}\| \leq \frac{M}{2}$. Iterating this procedure we get a sequence of linearly independent sets of vectors $u_1^{(j)}, ..., u_n^{(j)}$ so that $\max_{i=1}^n \|u_i^{(j)}\| \leq \frac{M}{2^{(i)}}$. Since $\log_2 M \leq \sum_{i=1}^n \mathrm{size}(a_i)$ we get the output after a polynomial number of iterations, that is in polynomial time. *Q.E.D.*(Lemma 12)

Proof of lemma 11. Let L^* be the dual lattice of $L = L(a_1, ..., a_n)$. We can get a basis of L be taking the dual of the basis $(a_1, ..., a_n)$ that is a set of vectors $d_1, ..., d_n$ so that for all $1 \leq i, j \leq n$ $a_i \cdot d_j = \delta_{i,j}$. The coordinates of the vectors $d_1, ..., d_n \in \mathbf{R}^n$ are rationals since they are the unique solution of a linear system of equations with rational coefficients. Since the number of unknowns in this system is n^2 we get that the number of bits in the value of determinant of the system remains below a polynomial bound (in the size of our input). Therefore all of the coordinates in the vectors d_i $i = 1, ..., n$ can be written as fractions whith the same common denominator r where $\mathrm{size}(r)$ is polynomial in the size of the input. (The numerators are also of polynomial lengths.)

Consequently we may apply the algortihm lemma 12 to the vectors $rd_1, ..., rd_n \in \mathbf{R}^n$. The ouptut of this algorithm determines $\mathrm{bl}(L(rd_1, ..., rd_n))$ and so $\mathrm{bl}(L(d_1, ..., d_n))$ upto a factor of n^{c_3}. According to Lemma 10 this gives the required estimate on the length of a shortest vector in $(L(d_1, ..., d_n))^* = L(a_1, ..., a_n)$. *Q.E.D.*(Lemma 11)

Lemma 14. *Assume that c_1, c_2, c_3 are the constants given in lemma 13. Then there is an absolute constant c with the following property. Suppose that there is* *a probabilistic polynomial time algorithm which given a value of the random variable $Z_{n,[c_1 n \log n],[n^{c_2}]}$ as input with a probability of at least 1/2 outputs a vector of $\Lambda(Z_{n,[c_1 n \log n],[n^{c_2}]})$ of length at most n. Then there are probabilistic algorithms $\mathcal{B}_1$, $\mathcal{B}_2$ with the following properties:*

(a) assume that $a_1, ..., a_n \in \mathbf{Z}^n$ and $v \in L(a_1, ..., a_n)$, $v \neq 0$ and for all $w \in L$ we have that if w is not in the subspace generated by v then $\|w\| \geq n^c \langle v \rangle$.

Then given $a_1, ..., a_n$ as input, $\mathcal{B}_1$ will output a vector $\tilde{v}$ in time polynomial in $\sigma = \sum_{i=1}^{n} \mathrm{size}(a_i)$ so that with a probability greater than $1 - 2^{-s}$, $\tilde{v}$ is either v or $-v$.

(b) assume that $a_1, ..., a_n \in \mathbf{Z}^n$ and there is a basis $g_1, ..., g_n$ of $L(a_1, ..., a_n)$ so that $\max_{i=1}^{n-1} \|g_i\| \leq M$ and the distance of g_n from the hyperplane F generated by $g_1, ..., g_{n-1}$ is at least $n^c M$.

Then, given $a_1, ..., a_n$ as input, $\mathcal{B}_2$ finds a basis $d_1, ..., d_{n-1}$ of $F \cap L(a_1, ..., a_n)$ in time polynomial in $\sigma = \sum_{i=1} n\mathrm{size}(a_i)$ and with a probability of at least $1 - 2^{-\sigma}$.

Proof. (b). Let $K = \max_{i=1}^{n} \|a_i\|$. By Lemma 12 we may assume that $K \leq n^{c_3} \mathrm{bl}(L)$. If D is the distance of g_n from D, then $\mathrm{bl}(L) \leq D + (n-1)M$ and so $K \leq n^{c_4} D$ for some absolute constant c_4. (We will assume that c is sufficiently large with respect to c_4.) According to Lemma 4 it is enough to find $n-1$ linearly independent elements $d_1, ..., d_{n-1}$ in F. We choose the elements d_k $k = 1, ..., n-1$ by recursion on k with the additional property that $\|d_k\| \leq 2n^{c_4+5} D$. Assume that the linearly independent elements $d_1, ..., d_k \in F$, $\|d_i\| \leq 2nK$ has been already selected for some $0 \leq k \leq n-2$ (that is, we include the $\{d_1, ..., d_k\} = \emptyset$ case). We may pick a basis $d_1, ..., d_k, b_1, ..., b_{n-k}$ of $L(a_1, ..., a_n)$ so that $\{b_1, ..., b_{n-k}\} \subseteq \{a_1, ..., a_n\}$. Let $N = n^{c_4+4} D$. We consider the set Y_N of all linear combinations $\sum_{j=1}^{n-k} \beta_k b_k$, where β_j, $j = 1, ..., n-k$ are integers with $0 \leq \beta_j \leq N$. The assumption that $d_1, ..., d_k, b_1, ..., b_{n-k}$ is a basis implies that if F_k is the vectorspace generated by $d_1, ..., d_k$ over $\mathbf{R}$, then all of the elements of Y_N are in different cosets of F_k. Clearly $|Y_N| \geq |N|^{n-k} \geq (n^{c_4+3} D)^{n-k}$. For each $u \in Y_N$ we have $\|u\| \leq (n-k)N$. Therefore Y_N is contained in a sphere S with radius $(n-k)N$. Since the distance between the neighboring cosets of F (which has nonempty intersection with L) is D we have that the number of cosets of F which intersects $S \cap L$ is at most $1 + 2(n-k)ND^{-1} < 2n^{2+c_4}$. Since $Y_N \geq n^{3+c_4}$ if we start to list the points of Y_N in some arbitrary order, then we will not run out of points in the first $2n^{2+c_4}$ steps and actually among these points there will be two that are in the same coset of F. Suppose that $y_1, ..., y_s$, $s = n^{2+c_4}$ are the list of these points and for some $k \neq l$ $y_k - y_l \in F$. (Later we will show that we can actually decide in polynomial time whether a $v \in L$ is also an element in F if $\mathrm{size}(v)$ is polynomial in the input.) We claim that $d_{k+1} = y_k - y_l$ meets our requirement. Indeed $d_{k+1} \in F$ and since y_k and y_l are in different cosets of F_k we have $d_{k+1} \notin F_k$ and so $d_1, ..., d_k, d_{k+1}$ are linearly independent. By the defintion of Y_N we have $\|d_{k+1}\| \leq 2(n-k)N \leq 2n^{c_4+5}D$.

Finally we show how is it possible to decide whether a $v \in L(a_1, ..., a_n)$ is also an element of F, provided that size$(v) \leq U$ where U is polynomial in the size of the input. Let t be a prime in the interval $= [2^U, 2^{U+1}]$ where c_3 is the constant given in lemma 11. (We can find such a number t so that with a probability exponentially close to 1 it meets this requirements.) We may assume that $U > n^{c_3}$ and $2^U > 2nND^{-1}$. Let $w = \frac{1}{t}v$. We consider the $\mathbf{Z}$-module A generated by the vectors $a_1, ..., a_n, w$. Since $tA \subseteq \mathbf{Z}^n$, A, as a $\mathbf{Z}$-module, can be generated by n elements so it is a lattice. By lemma 4 we can give an estimate z_A on $\mathrm{bl}(A) = \frac{1}{t}\mathrm{bl}(tA)$ in polynomial time with an error not greater then a factor n^{c_3} (in the sense of lemma 4). We may get a similar esitmate z_L for $\mathrm{bl}(L)$. We claim that if $v \in F$ then $z_L/z_A \leq n^{c^3}$ and if $v \notin F$ then $z_L/z_A > n^{c_3}$.

Indeed, if $v \in F$ and D is the distance of the hyperlane F from g_n then

(10) $\quad D \leq \mathrm{bl}(A) \leq D + nM$

Since $D \geq n^c M$ where c is sufficiently large with respect to c_3, this implies $z_L/z_A \leq n^{c_3}$.

Assume now that $v \notin F$ and that e.g. v and g_n are in the same halfspace determined by the hyperplane F. Since $g_1, ..., g_n$ is a basis of L and $\{g_1, ..., g_{n-1}\} \subseteq F$, we may write each vector iw, $i = 1, ..., t$ in the form $x_i + \tau_i v$ where $0 \leq \tau_i < 1$ and $x_i \in jg_n + F$ for some positive integer j. Since $v \in kg_n + F$ for some integer k. The choice of U and t imply that $t > k$ and so the primality of t implies that $\tau_i > 0$ for $i = 1, ..., t-1$ and trivially $\tau_t = 0$. Since τ_i is the fractional part of $i\tau_1$ this implies that $\tau_1 = s/t$ for some integer s and therefore there is a j, $0 < j < t$ with $\tau_j = \frac{1}{t}$. Let $x_j \in k'g_n + F$ and let u be the point that we get from jw by orthogonally projecting it on $k'g_n + F$. Clearly $\|jv - u\| \leq \frac{1}{t}D$. Since $\|g_i\| \leq M$, $i = 1, ..., n-1$, there is a $y \in k'g_n + F$ so that $\|u - y\| nM$. $g_1, ..., g_{n-1}, jw - y$ are linearly independent vectors in A, $\|jw - y\| \leq nM + \frac{1}{t}D$, $\|g_i\| \leq M$ for $i = 1, ..., n-1$ therefore lemma 4 implies that $\mathrm{bl}(A) \leq n^2 M + \frac{n}{t}D$. This togehter with (10) and $t \geq n^{2c_3}$ imply that $z_L/z_A > n^{c_3}$. *Q.E.D.*(b)

The only probabilistic step involved in this proof was the choice of the prime t. Even this can be avoided if we perform the described test for all $t = r^{nc'}$, $r = 1, ..., n^{c''}$. If $v \notin F$ for at least one value of t, (when k is not divisible by t) the test will show this fact.

(a). Let L^* be the dual lattice of $L(a_1, ..., a_n)$. We will show that L^* satisfies the assumption of (b) with a suitable choice of $g_1, ..., g_n \in L^*$. First we note that the assumption about the element v implies that if L_v is the one dimensional lattice generated by v then

(11) the factor lattice L/L_v has no shorter non-zero vector than $(n^c - 1)\|v\|$

Let $v = v_1, v_2, ..., v_n$ be a basis of L, let $h_1, ..., h_n$ be the corresponding dual basis of L^* and let $g_n = h_1$. This definition of g_n implies that $v \cdot g_n = 1$. Let F be the hyperplane orthogonal to v. $v \cdot g_n = 1$ implies that the distance of g_n from F is $\|v\|^{-1}$. We claim that $F \cap L^* = L(h_2, ..., h_n)$ has a basis whose elements are shorter then $n^{-c'}\|v\|^{-1}$. Indeed, this lattice is the dual of L/L_v therefore lemma 10 and property (11) implies our claim. Let $g_1, ..., g_{n-1}$ be an arbitrary basis of $F \cap L^*$ with elements no longer than $n^{-c'}\|v\|^{-1}$. This way (b) is satisfied with $M = n^{-c'}\|v\|^{-1}$. Therefore using the algorithm whose existence was stated in (b) we are able to find a basis $u_1, ..., u_{n-1}$ for $F \cap L^*$ in polynomial time, if $a_1, ..., a_n$ given as an input. The computational problem that the vectors in the dual space may have non-integer coefficient, can be handled in the same way as in the proof of lemma 11. We may pick a u_n so that $u_1, ..., u_n$ is a basis of L^*. Let $d_1, ..., d_n$ be the dual basis in L. We claim that d_1 is v or $-v$. Indeed d_1 is orthogonal to $u_1, ..., u_{n-1}$ therefore it is parallel to v. Since v is a shortest vector in L we have $d_1 = kv$ for some integer k. k must be 1 or -1 otherwise $L(d_1, ..., d_n)$ could not contain v. *Q.E.D.*(Lemma 14)

The following lemma is not necessary for the proof of our main result. It shows that the random lattice has a short basis with high probability. (To make the proof simpler we prove it only the case when q is odd, but it is easy to modify the proof for an arbitrary q. The smallest prime p which is not a divisor of q may take the role of the number 2.)

Lemma 15 . *For each positive integer u and $\epsilon > 0$ there is a $c > 0$ so that if n is a positive integer and $q \geq n^{\epsilon}$ is an odd positive integer then the probability of the following event is at least $1 - n^{-u}$:*

$\Lambda(Z_{n,[cn \log n],q})$ has a basis $d_1, ..., d_n$ so that $\|d_i\| \leq n^2$ for all $i = 1, ..., n$

Proof. Assume that c is sufficiently large with respect to the constant c of lemma 1 and let $m = cn \log n$. We define d_i in the following way. We will write d_i in the form of $d_i = \langle h_1^{(i)}, ..., h_m^{(i)} \rangle$. Let $h_i^{(i)} = 1$. Now we apply lemma 1 with $k = m - 1$ and $\langle b_1, ..., b_k \rangle \to \langle v_1, ..., v_{i-1}, v_i, ..., v_{i+1} \rangle$. According to lemma 1 with a probability of at least $1 - 2^{-c(m-1)}$ the sequence v_j has the following property: if we take the numbers $\delta_j \in \{0, 1\}$ independently and with uniform distribution then the distribution of $\sum_{1 \leq j \leq m, j \neq i} v_j$ is almost uniform on $(\mathbf{Z}/(q))^m$, where $\mathbf{Z}/(q)$ is the ring of residue classes mod q. (We apply lemma 1 so that A is the additive group of this structure). According to the lemma there is a choice for the sequence δ_j so that $\sum_{1\le j\le m, j\ne i} v_j \equiv -\frac{v}{2} \pmod q$. (Since q is odd $\frac{v}{2}$ is uniquely defined mod q). Let $h_j^{(i)} = 2\delta_j$ for all $j = 1, ..., i-1, i+1, ..., m$. The element b_i defined by this sequence is certainly in $\Lambda(Z_{n,m,q})$ since our definition implies that $\sum_{j=1}^{m} h_j^{(i)} v_j \equiv 0, \pmod q$. The definition also implies that $\|b_i\| \le (4m+1)^{1/2}$. We claim that the m vectors $b_1, ..., b_m$ are linearly independent in $\mathbf{R}^m$. This is an immediate consequence of the fact that their matrix is the unit matrix $\pmod 2$ and therefore their determinant cannot be 0. Finally according to Lemma 4 the existence of a linearly independent system of dimension m and of length at most $M = (4m+1)^{1/2}$ implies the existence of a basis in the lattice whose vectors have length at most nM. $Q.E.D.$(Lemma 15)