

(1)

Feb 2015

Letter to Scott Aaronson and Andy
Pollington on the Solovay-Kitaev Theorem and
Golden Gates.

Dear Scott and Andy,

Thanks for pointing me to the papers [K-M-M 1,2], [R-S]. I was not aware of these interesting developments in connection with the Solovay-Kitaev Theorem and the design of universal 1-qubit quantum gates. These papers refer to some others which are also very interesting such as [B-G-S]. All of these results can be understood in a unified way in terms of the arithmetic of quaternion algebras. Doing so clarifies (at least for me) the constructions and it also allows one to prove some fundamental properties for these gates as well as to relate them to some older and more recent developments. I explicate these points below, for myself as much as for the reader.

The problem is one of constructing an efficient universal set of quantum gates. That such exist is the Solovay-Kitaev Theorem [N-C]. The basic state in quantum computation is a single qubit, which is a unit vector in \mathbb{C}^2 . The construction of universal quantum gates reduces to finding efficient topological generators of $G = \text{PSU}(2)$ (or $\text{PSU}(2)$) [N-C]. G comes with a left and right

②

invariant metric:

$$d_G^2(x, y) = 1 - \frac{|\text{trace}(x^* y)|}{2} \quad (1)$$

$$= d_G^2(hx, hy) = d_G^2(xh, yh). \quad (2)$$

A set $S' = \{s_1, s_2, \dots, s_r\}$ of elements of G (the gates) are universal if Γ , the group generated by the s_j 's is topologically dense in G . Different gates might have different costs $w(s_j) \geq 0$ to be implemented. So define the height $h(\gamma)$, of $\gamma \in \Gamma$ to be

$$h(\gamma) = \min \left\{ \sum_{k=1}^l w(s_{j_k}) : \gamma = s_{j_1} s_{j_2} \dots s_{j_l} \right\} \quad (3)$$

The efficiency of the gate set S is measured by their ability to approximate any $x \in G$ by γ 's of small height (this corresponds to the size of the corresponding quantum circuit). Given $\epsilon > 0$ let t_ϵ be the least t for which

$$G \subset \bigcup_{\gamma \in V(t)} B_G(\gamma, \epsilon) \quad (4)$$

where

$$V(t) = \{\gamma \in \Gamma : h(\gamma) \leq t\} \quad (5)$$

and $B_G(x, r)$ is a ball centered x radius r in G .

(3)

Clearly

$$|V(t_\varepsilon)| \cdot \mu(B_G(\varepsilon)) \geq 1, \quad (6)$$

where μ is normalized Haar measure on G and $B_G(\varepsilon)$ any ball radius ε (they all have the same measure which is $\sim \varepsilon^3$ as $\varepsilon \rightarrow 0$).

To measure the covering efficiency of the points $V(t)$, we define the covering exponent $K(S)$ by

$$K(S) := \lim_{\varepsilon \rightarrow 0} \frac{\log |V(t_\varepsilon)|}{\log(1/\mu(B_G(\varepsilon)))}. \quad (7)$$

Clearly $K(S) \geq 1$ and if $K(S) = 1$ then the generating gates are optimally asymptotically efficient. For example if the points in $V(t)$ were behaving like a random set of $|V(t)|$ points, then we would have $K(S) = 1$.

The mathematical problems are:

(A) How small can we make $K(S')$ by choosing the gates in S' suitably?

(B) To give a $\text{poly}(\log 1/\varepsilon)$ time algorithm to find good approximations to these minimal cost circuits.

(C) The super-efficient continued fraction algorithm approximates a real number by rationals using $x \mapsto x+n$, $x \mapsto 1/x$, and one is seeking something like it for $PV(2)$.

(4)

The Solovay-Kitaev Theorem asserts that for any universal gate set S , given $\epsilon > 0$ and $x \in G$, an element $y \in \Gamma$ can be found in $O((\log^{1/\epsilon} n)^{2.71})$ steps for which $d_G(x, y)$ is at most ϵ and the height of y is $O(\log^{1/\epsilon} n)^{3.97}$; these explicit powers of $\log^{1/\epsilon} n$ being due to [D-N].

This is satisfactory from a theoretical point of view but in practice one would like to do much better. Indeed from the above it does not even follow that $K(S) < \infty$. The last is in fact true, at least if the entries in each S_j are algebraic numbers. This follows from the spectral gap theorem [B-G] which asserts that the self adjoint operator (assuming S is symmetric, $S_j \in S \Leftrightarrow S_j^{-1} \in S$ and given by $w(S_j) = w(S_j^{-1})$) $T_S : L^2(G) \rightarrow L^2(G)$

$$T_S f(x) = \sum_{j=1}^n w(S_j) f(S_j x), \quad (8)$$

has a gap in its spectrum below its top eigenvalue, which is $\sum_{j=1}^n w(S_j)$.

The proof of the spectral gap does not yield any feasible value for the gap and hence for $K(S)$. Moreover being a counting argument it offers nothing on problem (B). Still that $K(S) < \infty$ is a good step and an indication of what one might achieve.

(5)

by choosing the gates carefully.

All the known good gates Σ come from arithmetic and number theory and the proof that they are efficient uses the theory of automorphic forms on associated groups. We review the constructions, they all have the same pros and cons.

(i) $p=5$ ($[L-P-S]_{1,2}$, $[B-G-S]$):

$$\Sigma' = \{S_1, S_1^{-1}, S_2, S_2^{-1}, S_3, S_3^{-1}\} \text{ where}$$

$$S_1 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1+2i & 0 \\ 0 & 1-2i \end{bmatrix}, S_2 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2i \\ 2i & 1 \end{bmatrix}, S_3 = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}.$$

The weights $W(S_j)$ are all taken to be 1, Γ is a free group on S_1, S_2, S_3 and

$h(\gamma)$ is the reduced word length of γ .

$$|V(t)| = 6 \cdot 5^{t-1}, \text{ for } t \geq 1.$$

$$(a) \bullet \quad \frac{4}{3} \leq K(S) \leq 2$$

The last asserts (essentially) that for $\epsilon > 0$ and $x \in G$ there is a $\gamma \in \Gamma$ with $h(\gamma) \leq 6 \log_5(1/\epsilon)$ and $d_G(x, \gamma) < \epsilon$, while there are y 's in G for which $d_G(y, \gamma) < \epsilon$ $\Rightarrow h(\gamma) \geq 4 \log_5(1/\epsilon)$.

(6)

- (b) • For $\varepsilon > 0$ there is t_ε with

$$(\log|V_{t_\varepsilon}|)/(\log \varepsilon^{-3}) \rightarrow \pm 1 \text{ as } \varepsilon \rightarrow 0,$$

such that for most points $y \in G$ w.r.t Haar measure, there is an ε -approximation γ of y of height t_ε (so optimal for most y 's).

- (c) • There is a $\text{poly}(\log \frac{1}{\varepsilon})$ time algorithm which assuming some reasonable conjectures about the distribution of primes, produces for any $x \in G$, a $\gamma \in P$ with $h(\gamma) \leq 12 \log_5(1/\varepsilon)$ and $d_G(x, \gamma) < \varepsilon$. Moreover if x is diagonal then γ has $h(\gamma) \leq 4 \log_5(1/\varepsilon)$.

- (d) • There is a probabilistic algorithm with expected running time $\text{poly}(\log \frac{1}{\varepsilon})$ which for x diagonal produces (if it stops) the γ of smallest height in $B_G(x, \varepsilon)$ (assuming that one has a polynomial time factoring algorithm) [R-S].

(ii). $p=2$ (see [ch. 1]):

$$\mathcal{S} = \left\{ S_0, S_1, S_1^{-1} \right\}, \text{ where}$$

$$S_0 = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, S_1 = \frac{1}{\sqrt{32}} \begin{bmatrix} 2+i\sqrt{2} & \sqrt{26}i \\ \sqrt{26}i & 2-i\sqrt{2} \end{bmatrix}.$$

(7)

The weights $w(s_j)$ are all chosen to be 1. $\Gamma \cong \mathbb{Z}/2\mathbb{Z} \times \mathbb{Z}$ and $h(\gamma)$ is reduced word length in S_0 ($= s_0^{-1}$) and S_1 (in $\mathrm{PSU}(2)$). $|V(t)| = 3 \cdot 2^{t-1}$.

(a) • $\frac{4}{3} \leq K(S) \leq 2$.

The last asserts that for $\varepsilon > 0$ and $x \in G$ there is a $y \in \Gamma$ with $h(y) \leq 6 \log_2 \frac{1}{\varepsilon}$ and $d_G(x, y) < \varepsilon$, while there are y 's in Γ for which $d_G(y, x) < \varepsilon \Rightarrow h(y) \geq 4 \log_2 \frac{1}{\varepsilon}$.

(b) • For $\varepsilon > 0$ there is t_ε with

$(\log |V_{t_\varepsilon}|) / (\log \frac{1}{\varepsilon^3}) \rightarrow 1$ as $\varepsilon \rightarrow 0$, such that for most points $y \in G$ w.r.t Haar measure, there is an ε approximation of y of height t_ε (so optimal for most y 's).

(c) • There is a $\mathrm{poly}(\log \frac{1}{\varepsilon})$ time algorithm which assuming some reasonable conjectures about the distribution of primes, produces for any $x \in G$ a $y \in \Gamma$ with $h(y) \leq 12 \log_2 \frac{1}{\varepsilon}$, and $d_G(x, y) < \varepsilon$. Moreover if x is diagonal then y has $h(y) \leq 4 \log_2 \frac{1}{\varepsilon}$.

(d) • There is a probabilistic algorithm with expected running time $\mathrm{poly}(\log \frac{1}{\varepsilon})$ which for x diagonal produces (if it stops) the y of smallest height in $B_G(x, \varepsilon)$ (assuming that one has a polynomial time factoring algorithm).

(8)

(iii) H-T gates $([\bar{A} \cdot \bar{M}], [\bar{B} \cdot \bar{S}], [\bar{K} \cdot \bar{M} \cdot \bar{M}], [\bar{R} \cdot \bar{S}])$

$$S = \{H, T\}$$

$$H = \frac{i}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{The "Hadamard Gate"}$$

$$T = \begin{bmatrix} e^{i\pi/8} & 0 \\ 0 & e^{-i\pi/8} \end{bmatrix} \quad \text{The "}\frac{\pi}{8}\text{" gate".}$$

Gates such as H which lie in the finite

Clifford group C ($C = \langle H, T^2 \rangle \cong S_4$)

is a quotient of \mathbb{C} , the usual Clifford group
of order 192 , since we work here in $PSU(2)$
are apparently much easier to prepare. To obtain
a universal gate set one has to add an
element that is expensive. The T-gate
which is of order 8 is a popular one to add
and whose implementation has been studied.

Set $W(c) = 0$ for $c \in C$ and $W(T) = 1$
(The weights can be chosen as in $[\bar{B} \cdot \bar{S}]$),
so that $h(\gamma)$ only counts applications of the
T gates. $|V_T| \sim c^{2^\gamma}$.

$$(a) \bullet \quad \frac{4}{3} \leq K(S) \leq 2$$

The last asserts that for $\epsilon > 0$
and $x \in G$ there is a $\gamma \in \mathbb{N}$
with $h(\gamma) \leq 6 \log_2(1/\epsilon)$

(9)

And $d_G(x, y) < \varepsilon$, while there are y 's in G for which $d_G(y, x) < \varepsilon \Rightarrow h(y) \geq 4 \log_2^{1/\varepsilon}$.

- (b) • For $\varepsilon > 0$ There is t_ε with $(\log |V_{t_\varepsilon}|)/(\log \varepsilon^{-3}) \rightarrow 1$, as $\varepsilon \rightarrow 0$ such that for most points $y \in G$ w.r.t d_H , There is an ε -approximation of y of height t_ε (optimal).
- (c) • There is a $\text{poly}(\log^{1/\varepsilon})$ time algorithm which assuming some reasonable conjectures about the distribution of primes, produces for any $x \in G$ a $y \in \Gamma$ with $h(y) \leq 12 \log_2^{1/\varepsilon}$ and $d_G(x, y) < \varepsilon$. Moreover if x is diagonal then y has $h(y) \leq 4 \log_2^{1/\varepsilon}$.
- (d) • There is a probabilistic algorithm with expected running time $\text{poly}(\log^{1/\varepsilon})$ which for x diagonal produces (if it stops) The y of smallest height in $B_G(x, \varepsilon)$ (assuming that one has a polynomial time factoring algorithm).

Discussion:

The source for (i) is the Hamilton quaternion algebra D (that is generated by $1, i, j, k$ with $i^2 = j^2 = -1$, $ij = -ji = k$...) over \mathbb{Q} .

D is split at $p=5$ (one could use

(10)

any other $p \neq 2$ just as well) And is ramified at the real place and at ∞ .

So $D \otimes \mathbb{Q}_5 \cong \text{Mat}_{2 \times 2}(\mathbb{Q}_5)$ and if

$D^*(\mathbb{Z}[\frac{1}{5}])$ consists of all elements in D whose entries and those of their inverses are in $\mathbb{Z}[\frac{1}{5}]$, then the diagonal embedding corresponding to $p=5$ and ∞ ;

$$D^*(\mathbb{Z}[\frac{1}{5}]) \xrightarrow{\text{diag}} \text{PGL}_2(\mathbb{Q}_5) \times (\text{SU}(2) \times \mathbb{R}^*)$$

has its image in $\text{SU}(2)$ equal to Γ in (i). At the same time from the first factor $D^*(\mathbb{Z}[\frac{1}{5}])$ acts isometrically on the 6-regular tree $X = \text{PGL}_2(\mathbb{Q}_5) / \text{PGL}_2(\mathbb{Z}_5)$. In fact as shown in [L-P-S] it acts simply transitively on X and $\mu(\gamma)$ is simply the distance in X from $\gamma \bar{z}$ to \bar{z} , where $\bar{z} = \text{PGL}_2(\mathbb{Z}_5)$. This allows one to use automorphic forms to analyze the equidistribution in G of the points in V_t which corresponds to a ball in X of radius t [L-P-S]. The emphasis there was to establish sharp equidistribution, but for the purpose of the covering number one can improve the exponent by a positivity argument. This was done in the thesis [CH2] in a somewhat different setting. Since this leads to the best bound that we have for $K(S)$ namely $K(S) \leq 2$, we give

(11)

a variation of the argument in Appendix 1. The number 2 is a reflection of the square root cancellation coming from the Ramanujan Conjecture (Deligne's Theorem).

That is we need $|V(t_\varepsilon)|$ to be the square of $\|\mu(B_\varepsilon(\varepsilon))\|$ in order to ensure that V_{t_ε} meets every such ball. In Appendix 1 we also show that this square root feature leads to most points having optimally short circuit approximations.

That $K(S) \geq 4/3$ is more elementary and is a consequence of the points in V_t having big holes near the projections onto the unit sphere S^3 ($\cong \text{SU}(2)$ metrically) of integer points in \mathbb{Z}^4 . This big hole phenomenon was first observed in the context of approximating real 2×2 matrices of unit determinant by projections of ones of determinant p .

In our setting of (i) the points V_t correspond to integer solutions of

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 5^h, h \leq t. \quad (9)$$

If $0 \neq y \in \mathbb{Z}^4$ then the ball

$B_{S^3}(\frac{y}{\|y\|}) \frac{1}{10(\|y\|5^{t/2})^{1/2}}$ contains a ball in S^3 of radius $\frac{1}{20t(\|y\|5^{t/2})^{1/2}}$ which has no points of V_t . (10)

(12)

We give a proof of (10) in Appendix 2.
That $K(S) \geq 4/3$ follows immediately from (10).

The analysis of (ii) is similar. This time the gate set S' comes from the quaternion algebra D/\mathbb{Q} where $i^2 = -2$, $j^2 = -13$. D is ramified at ∞ and $p = 13$. So in this case we can localize at $p = 2$ and form the group $D^*(\mathbb{Z}[\frac{1}{2}])$. In [Ch1] it is shown that $D^*(\mathbb{Z}[\frac{1}{2}])$ acts simply transitively on the 3-regular tree $\mathrm{PGL}_2(\mathbb{Q}_2)/\mathrm{PGL}_2(\mathbb{Z}_2)$. The generating set S' correspond again to the 3-neighbours of $\mathrm{PGL}_2(\mathbb{Z}_2)$. The rest of the analysis is the same as in example (i).

We turn to (iii) which has been investigated directly and quite intensively recently ([K-M-H], [S], [R-S]). These H-T gates also come from a quaternion algebra and this allows us among other things, to conclude that $K(S) \leq 2$. The quaternion algebra is again the Hamilton quaternions, D , but this time we consider it over the field $k = \mathbb{Q}(\sqrt{2})$. k has two nonarchimedean places at both of which D is ramified, giving the dyadic $H(R)$. We fix one of these real places. The prime 2 is ramified in R , $(2) = (\sqrt{2})^2 := P^2$.

(13)

$k_{\mathbb{P}} = \mathbb{Q}_2(\sqrt{2})$, and D is split over $k_{\mathbb{P}}$ since

$(\sqrt{-7} + 2\underline{i} + \sqrt{2}\underline{j} + \underline{k})$ is a zero divisor in

$D_{\mathbb{P}} := D \otimes k_{\mathbb{P}}$; note $\sqrt{-7} \in \mathbb{Q}_2 \subset k_{\mathbb{P}}$). In fact

D/k is tamified only at its two infinite places.

Let \mathcal{O} be the ring of integers of k , and $\mathcal{O}_{\mathbb{P}}$ the

integers in $k_{\mathbb{P}}$. The key arithmetic group is

$\Delta = D^*(\mathcal{O}[\frac{1}{2}])$. Under the diagonal embedding

$$\Delta = D^*(\mathcal{O}[\frac{1}{2}]) \hookrightarrow \mathrm{PGL}_2(k_{\mathbb{P}}) \times (\mathrm{SU}(2) \times \mathbb{R}^*)$$

—(11)

(using our chosen real place), Δ is mapped to a subgroup of $\mathrm{SU}(2)$. We show that Δ is equal to $\Gamma = \langle H, T \rangle$.

Let $\tilde{H} = (\underline{i} + \underline{k})/\sqrt{2}$ and $\tilde{T} = \frac{\underline{2} + \sqrt{2}}{2} + \frac{\sqrt{2}}{2}\underline{i}$. —(12)

Then

$$\mathrm{Norm}(\tilde{H}) = 1 \text{ and } \mathrm{Norm}(\tilde{T}) = \sqrt{2}(1 + \sqrt{2}).$$

—(13)

Hence \tilde{H} and \tilde{T} are in Δ .

Also $\tilde{T}^2 = (1 + \sqrt{2})(1 + \underline{i})$ so that under

the usual identification of $H^*(\mathbb{R})/\mathbb{R}^*$ with $\mathrm{SU}(2)$ ($x_1 + x_2\underline{i} + x_3\underline{j} + x_4\underline{k} \rightarrow \begin{bmatrix} x_1 + ix_2 & x_2 + ix_3 \\ x_2 - ix_3 & x_1 - ix_2 \end{bmatrix}$)

\tilde{H} corresponds to H and \tilde{T} to T .

(14)

It follows that $\langle \tilde{H}, \tilde{T} \rangle \subset \Delta$. One way of seeing that these groups are equal is to consider their image under the usual double-cover $\pi: D^* \rightarrow SO_p$

(over k) where $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$ over k .

Now $\pi(\Delta) \subset SO_p(\mathcal{O}[\frac{1}{2}])$ and as shown in

[Ser1]** The last group is the quageneral

group $G(4, 8)$ generated by rotations R_4 and R_8 of orders 4 and 8 about orthogonal axes (and also $G(4, 8) \cong S_4 * D_8$).

Moreover, $\pi(\tilde{T}) = R_8$ and

$\pi(\tilde{H}\tilde{T}^2\tilde{H}^{-1}) = R_4$. Hence $G(4, 8) \subset \pi(\Delta)$

and so $\pi(\Delta) = SO_p(\mathcal{O}[\frac{1}{2}])$ and $\Gamma = \Delta$.

To exploit this realization of Γ as the arithmetic group $D^*(\mathcal{O}[\frac{1}{2}])$ we need to relate the height $h(\gamma)$ in terms of T counts to the action of γ on the \mathfrak{I} -regular tree $X = PGL_2(k_{\mathfrak{I}})/PGL_2(\mathcal{O}_{\mathfrak{I}})$ under the embedding (11). The elements in $D_{\mathfrak{I}}$ of the form

$$\theta_{\mathfrak{I}} \frac{i+j+k}{2} + \theta_{\mathfrak{I}} \frac{1+i}{2} + \theta_{\mathfrak{I}} + \theta_{\mathfrak{I}i} + \theta_{\mathfrak{I}j} + \theta_{\mathfrak{I}k}$$

form a maximal order $M_{\mathfrak{I}}$ of $D_{\mathfrak{I}}$. Note that

\tilde{H} and \tilde{T}^2 are units in $M_{\mathfrak{I}}$, so that the

Clifford group $C = \langle \tilde{H}, \tilde{T}^2 \rangle$ is contained in the unit group of $M_{\mathfrak{I}}$. It follows that C stabilizes $\xi = PGL_2(\mathcal{O}_{\mathfrak{I}})$ in X . On the other hand since $\text{Norm}(\tilde{T}) = (\sqrt{2})$ it follows that \tilde{T} moves ξ to one of its

(*) in answer to a question posed in [RJ].

neighbors in X and $\tilde{H}\tilde{T}$ and \tilde{T} move $\tilde{\gamma}$ to the other two nearest neighbors.

Thus Δ acts transitively on X and the height of any $\gamma \in \Delta$ in terms of \tilde{T} and \tilde{T}^{-1} counts, is at most the distance from $\tilde{\gamma}$ to $\delta\tilde{\gamma}$ in X plus $O(1)$. One can also use T -counts alone as is done in [B-S] and [R-S], where one treats the gates in C as being free. Again the distance in X from $\tilde{\gamma}$ to $\delta\tilde{\gamma}$ is essentially this T -count.

With this analysis we are in the same position as in examples (i') and (ii') and everything there applies equally well for the $H-T$ gates (The Ramanujan Conjectures for definite division algebras over totally real number fields are known).

One can also study this example (iii) in terms of an arithmetic unitary group. This view point is perhaps even more natural when examining gates sets such as the cyclotomic-Clifford gates in [F-G-K-M]. Let $n \geq 3$ and E the "CM-field" $\mathbb{Q}(\zeta_{2n})$ and F the corresponding totally real subfield $\mathbb{Q}(\zeta_{2n} + \zeta_{2n}^{-1})$, here ζ_{2n} is a primitive $2n$ -th root of 1. The unitary group U/F consists of all 2×2 matrices with entries in E , which preserve the hermitian form $h(u, w) = \bar{u}_1 w_1 + \bar{u}_2 w_2$, where $u = (u_1, u_2)$,

(15')

$\omega = (\omega_1, \omega_2)$ and $-$ is Galois conjugation E/F . As a group over F one can localize U to places v of F . Denote by F_v the completion of F at v and by U_v the corresponding unitary group over F_v (its elements are matrices in $E_v = E \otimes F_v$). If v is split in E then U_v is isomorphic to $GL_2(F_v)$, while if v is inert or tamified in E then U_v is a genuine unitary group over F_v with corresponding field extension E_v . At the archimedean places σ of F (all such places are real) U_σ is a definite (compact) unitary group. The group $U_2(R_n)$ in $[F-G-K-M]$ is the full S -arithmetic group $U(\mathcal{O}[\frac{1}{2}])$ where \mathcal{O} is the ring of integers of F . To see this as an S -arithmetic group, factor \mathfrak{Q} in \mathcal{O} : $(\mathfrak{Q}) = (P_1 P_2 \dots P_g)$ where $e_{fg} = \deg[F:\mathbb{Q}]$, e is the ramification and f the degree of the residue fields of any F_{P_i} . In this way $U_2(R_n)$ is the S -arithmetic group of U consisting of all elements in $U(F)$ which are integral outside of $S = \{P_1, \dots, P_g\}$. The local groups U_{P_i} , $i = 1, 2, \dots, g$, are noncompact and isomorphic. The diagonal embedding

$$U_2(R_n) \hookrightarrow U_{P_1} \times U_{P_2} \times \dots \times U_{P_g} \quad (*)$$

realizes $U_2(R_n)$ as a cocompact lattice in the latter group. The gate group G_n in $[F-G-K-M]$ is the group generated by the finite Clifford group C and the

15"

The element $T_n = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & e^{i\pi/n} \end{bmatrix}$. Clearly $g_n \in U_2(R_n)$

And we want to know when these are equal or at least when does g_n have finite index in $U_2(R_n)$; That is whether g_n is thin or not in the sense of [Sa].

The first thing to note is that if $g > 1$ then g_n is thin. Indeed as explained in [F-G-K-M], g_n has a presentation as a simple amalgamated product. It follows from rigidity theorems for lattices in different groups, namely ones in a rank one type group (such as an amalgamated product as above) and ones in a higher rank group, cannot be isomorphic. So if $g > 1$ so that the product group in (k) is of higher rank, then g_n cannot be a lattice and so it is thin.

One can analyze the factorization of 2 in O_F using Dirichlet characters and one finds that if $2n = 2^k \cdot 5$, with 5 odd, Then $g = 1$ iff $\{-1, 2\}$ generate $(\mathbb{Z}/5\mathbb{Z})^*$ and $f = |(\mathbb{Z}/5\mathbb{Z})/(\pm 1)|$. On the other hand [F-G-K-M] show that the intersection of g_n with the diagonal torus $A = \left\{ \begin{bmatrix} x & 0 \\ 0 & x \end{bmatrix} \in U_2 \right\}$ is finite, while that of $U_2(R_n)$ with A is finite iff -1 is in the group generated by 2 in $(\mathbb{Z}/5\mathbb{Z})^*$. Combining these we have that if g_n is arithmetic then 2 must generate $(\mathbb{Z}/5\mathbb{Z})^*$.

Restricting to the cases where g_n might be arithmetic, let p be the unique

(15th)

prime in F above 2 . U_p is a split 2-dimensional unitary group over F_p (since we are assuming that G_n is infinite, U_p cannot be definite). The projective group $P\Gamma U_p$ is isomorphic to $\mathrm{PGL}_2(F_p)$. Hence $P\Gamma U_2(R_n)$ is realized as a lattice in $\mathrm{PGL}_2(F_p)$ and it acts discontinuously and isometrically on the 2^{f+1} regular tree $X = \mathrm{PGL}_2(F_p) / \mathrm{PGL}_2(O_p)$. In this setting deciding anything about G_n such as a presentation, or whether it acts with a compact quotient (i.e. it is arithmetic since $\Gamma U_2(R_n)$ does so) can be done in any given instance using the techniques in [Ser2]. In fact using these he shows in [Ser1] that in the case $s=1$, for $n=2, 4, 8$, G_n is arithmetic (its generators are H and T_n but since H and T_2 generate C , this group is the same as G_n) and in fact $G_n = U_2(R_n)$, while if $n = 2^v$ with $v \geq 4$ then G_n is thin (for the latter he computes Euler characteristics and estimates these using Tamagawa numbers, i.e. arithmetic formulae for the volumes of $U_2(R_n) \backslash X$). Presumably one can proceed similarly in the remaining cases when $s \neq 1$. However given the normal forms for members of G_n in [F-G-K-M] it seems to me that one can complete the analysis by simply counting elements in G_n . Note that if G_n were arithmetic then $\sum_{e \in G_n} d_X(e, e) \leq C$ $\#G_n 2^{ft}$ as $t \rightarrow \infty$, with $C > 0$. This follows from a lattice point count in the tree X . However

(15^{IV})

according to the normal form Cor 4.2 [F-G-K-M] and relating the T-covert of $y \in G_n$ to the distance it moves e in X (I didn't check that the relation is similar to the $n=4$ case) one only has $O(2^t)$ such elements in G_n . So if G_n is to be arithmetic then we must have $f=1$ and hence $S=1$ or 3. If so the only possible new arithmetic cases are $n=3 \cdot 2^v$ with $v \geq 1$. According to [F-G-K-M] $v=1$ and 2 are in fact arithmetic (and again $G_n = U(R_n)$ in these cases). For v larger I expect that again by counting the number of elements in G_n with $d_X(y_e, e) \leq t$, the count will be too small (and certainly by estimating Tamagawa numbers).

In summary it is likely that the cases $n=3, 4, 8, 12$ from [F-G-K-M], are the full list of arithmetic clifford-cyclotomic gates. As for the case $n=4$ discussed in (iii) above, the Ramanujan Conjectures are known for these unitary groups and (a), (b), (c), (d) holds for them.

The basic algorithm

We review the algorithm ([K-M-M], [B-G-S], [S], [R-S]) for finding good approximations of small height. Consider the case (i°) (the same ideas work in (ii°) and (iii°)). The elements in V_L are solutions to (9). A key point is that if we have an integral solution x to (9) which approximates some $y \in G$ suitably, then finding the short circuit for x in terms of the gates S can be done

(16)

quickly (i.e. polynomial in $t = O(\log \gamma_2)$). The action of Γ on the 6-regular tree X tells us how to move γ of height h to the identity coset $\mathbb{1}$ using the gates; one simply navigates along the geodesics in X . An equivalent way of achieving this, pointed out in [B-G-S], is to use the left and right unique factorization theory in $D(\mathbb{Z})$. Thus the problem of finding a short circuit is reduced to finding solutions to (9) with $x/\gamma^{1/2}$ doing the required approximation. The idea is that for a diagonal matrix

$Z = \begin{pmatrix} \alpha & 0 \\ 0 & \bar{\alpha} \end{pmatrix}$, $|\alpha| = 1$ one can find a good approximation as follows:

We seek γ_C of height t such that γ_C satisfies (9) and

$$|\gamma_1 - \bar{\gamma}_1| \leq 5^{\frac{t}{2}} \varepsilon, \quad |\gamma_2 - \bar{\gamma}_2| \leq 5^{\frac{t}{2}} \varepsilon. \quad (14)$$

$$|\gamma_3| \leq 5^{\frac{t}{2}} \varepsilon, \quad |\gamma_4| \leq 5^{\frac{t}{2}} \varepsilon$$

where $\bar{\gamma}_1^2 + \bar{\gamma}_2^2 = 5^t$ ($\alpha = \frac{\bar{\gamma}_1 + i\bar{\gamma}_2}{5^{\frac{t}{2}}}$). (14)

Choose $\gamma_1 = \lfloor \bar{\gamma}_1 \rfloor - k_1$, $\gamma_2 = \lfloor \bar{\gamma}_2 \rfloor - k_2$

with $0 \leq k_1, k_2 \leq t^A$ for some fixed $A > 1$.

Then $0 < 5^t - (\gamma_1^2 + \gamma_2^2) \leq 2t^A 5^{\frac{t}{2}}$. (15)

(16)

(17)

For each choice of R_1, R_2 as above we check if $5^t - x_1^2 - x_2^2$ is a prime $P \equiv 1(4)$. This can be done in $\text{poly}(t)$ steps [AKS] and if this happens then

one can find x_3, x_4 in $\text{poly}(t)$ steps [Sc], satisfying:

$$x_3^2 + x_4^2 = 5^t - x_1^2 - x_2^2 \quad \dots \quad (17)$$

In this case we arrive at a solution to (14) with $\varepsilon = 2t^{A/2}/5^{t/4}$. That is we arrive at a point of height (essentially) $4\log_5(1/\varepsilon)$. If our choice of x_1, x_2 fails to produce such a P we repeat with different choices of R_1, R_2 and do so t^{2A} ($\text{poly}(t)$) times.

Given the density of the distribution of primes we expect that if A is big enough (but fixed) that one will always arrive at such a prime P .

So assuming this heuristic about the distribution of primes of this form, one will find (in $\text{poly}(t)$ steps) an ε approximation to \mathbb{Z} of height at most $4\log_5 1/\varepsilon$. To approximate the general yet one factors y as $T_1 T_2 T_3$ with the T_j 's diagonal about different orthogonal axes and approximates each T_j by the above algorithm. This leads to the $12\log_5 1/\varepsilon$ length circuit approximating y .

(18)

The probabilistic algorithm searches (quickly) for x_1, x_2 in the region defined by (14') and $5^t - (x_1^2 + x_2^2) \leq \varepsilon^2 5^t$. We won't describe it further except to point out that it is in solving (17) for x_3, x_4 in $\text{poly}(t)$ steps, that one needs to assume that the right hand side of (17) can be factored quickly. In the H-T gate case (iii) this involves a search for integers in \mathbb{R} in certain regions [R-S]. The probabilistic algorithm apparently works very well in practice as illustrated with some runs in [R-S].

Intrinsic Diophantine Approximation:

In these arithmetic cases the problem of finding a global solution to (9) which approximates a given $y \in G$, is an extension of the problem of approximating points on spheres by rational and 'S-integral' points of small height. These are problems of intrinsic diophantine approximation that have been studied recently.

(A) The classic case of rational numbers a/q in $[0, 1]$ with the height of a/q being q . It consists of the Farey fractions of denominator at most t . There are $\approx t^2$ such points and their spacings vary from $1/t$ to $1/t^2$.

(19)

Hence the exponent K is ;

$$K([0, 1], \text{Farey}) = 2.$$

The fact that most points in $[0, 1]$ can be approximated to within $1/t^2$ by these Farey points, while at the same time there are points that cannot be approximated so well (i.e. $K \neq 1$) is the source of the rich theory of diophantine approximation.

(b) Consider the rational points in $[0, 1]$ which are '2-integral' that is the points $\mathbb{Z}[\frac{1}{2}] \cap [0, 1]$.

This time the 2^t points of height at most t are exactly 2^{-t} apart in $[0, 1]$. Thus the exponent is

$$K([0, 1], \mathbb{Z}[\frac{1}{2}]) = 1.$$

The theory of diophantine approximation here is simply the truncation of the dyadic expansion of a real number. So the theory is very simple.

The following concern the standard unit sphere S^n in \mathbb{R}^{n+1} ($n > 0$) and subsets of rational points on S^n which we denote by Q .

(20)

$$(c) Q = \left\{ \frac{x}{q} : x \text{ is integral}, q \geq 1, \left| \frac{x}{q} \right|^2 = 1 \right\}$$

$h\left(\frac{x}{q}\right) = q$ if $(x_1, \dots, x_{n+1}, q) = 1$. The recent work [K-M] shows that

$$K(S^n, Q) \leq 2.$$

The repulsion property (10) for these points takes a simpler and transparent form

$$d_{S^n} \left(\frac{x}{q}, \frac{x'}{q'} \right) \geq \frac{1}{(qq')^{1/2}}. \quad (18)$$

From this it follows that

$$K(S^n, Q) = 2,$$

and here too there is a rich theory of diophantine approximation.

(d) Consider S^1 with rational points which are say 5-integral that is in $\mathbb{Z}[\frac{1}{5}]$. These correspond to solutions of

$$x^2 + y^2 = 5^{2k}, x, y \in \mathbb{Z}.$$

The question of what $K(S^1, \mathbb{Z}[\frac{1}{5}])$ is, reduces to the diophantine properties of the generator of these rational points, namely the numbers $m \alpha \pmod{2\pi\mathbb{Z}}$, $m \geq 1$ where

$$\alpha = \cos^{-1}(\frac{3}{5}).$$

(21)

If as we expect this number is typical irrational in terms of its diophantine properties then

$$K(S^1, \mathbb{Z}[\frac{1}{5}]) = 1.$$

On the other hand if α is not typical then $K > 1$. I am not sure what, if anything is known about the type of the number $\alpha/\alpha\pi$.

(E) S^2 with the points in $\mathbb{Z}[\frac{1}{5}]$.

This is a very interesting case and corresponds to solutions to

$$x_1^2 + x_2^2 + x_3^2 = 5^{2t} \quad — (19)$$

Using automorphic forms (again the Ramanujan Conjecture) exploiting that the right hand side of (19) is a square, one can show that

$$K(S^2, \mathbb{Z}[\frac{1}{5}]) \leq 2. \quad — (20)$$

In the recent paper [B-R-S] the local statistical of the distribution on S^2 of sums of 3-squares is investigated. The thesis there is that these behave like a random set of points placed on S^2 , at least if the r.h.s of (19) is

(22)

square free. I expect that this also holds for (19) and hence that

$$K(S^2, \mathbb{Z}[\frac{1}{5}]) = 1.$$

If true this would be striking.

(f) S^3 with points in $\mathbb{Z}[\frac{1}{5}]$. That is solutions to

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = 5^{2h}, h \leq t.$$

—(21)

This is very close to (1') except that the exponent of 5 in (21) is restricted to be even. As discussed in (1') we have

$$\frac{4}{3} \leq K(S^3, \mathbb{Z}[\frac{1}{5}]) \leq 2.$$

—(22)

So this is the smallest dimension for S^n for which there is definitely a 'rich theory of intrinsic diophantine approximation for '5-integers'. An in depth study of such problems in great generality has been carried out in [G-G-N]. In particular they define local covering exponents which in this case are given as follows:

For $x \in S^3$ and $\varepsilon > 0$ let $t_\varepsilon(x)$ be the least t such that

(23)

$B_3(x, \varepsilon)$ contains an element in S^3 with coordinates in $\mathbb{Z}[\frac{1}{5}]$ and of height 5^h with $h \leq t$. Set

$$K(x, \mathbb{Z}[\frac{1}{5}]) := \lim_{\varepsilon \rightarrow 0} \frac{-\log |N_{t_\varepsilon}(x)|}{\log(1/\varepsilon^3)}.$$

They show that for almost all x (w.r.t. Lebesgue measure on S^3)

$$K(x, \mathbb{Z}[\frac{1}{5}]) = 1. \quad \text{--- (23)}$$

This of course is optimal (again the Ramanujan Conjectures are an ingredient).

(g) For $n \geq 4$ the determination of K for S -arithmetic points on S^n becomes easier. As pointed out to me by Bourgain, an application of the circle method as done in the Appendix [B-R], can be promoted to an asymptotic (and not just an upper bound) when one has this many variables, for counting these points in small caps. As a consequence we have together with the relation (18) that for $n \geq 4$,

$$K(S^n, \mathbb{Z}[\frac{1}{5}]) = 2 - \frac{2}{n}.$$

--- (24)

We end with some basic problems:

(1) For the arithmetic S 's in (i), (ii) and (iii) to show that $K(S) \leq 2$ and perhaps even that $K(S) = 4/3$.

(2) Is there a set of gates with $K(S) = 1$?

Perhaps with a 'thin' subgroups (see

[5a]) where all we know is that

$K(S) < \infty$, there may some cases where

$$K(S) = 1.$$

(3) For the arithmetic gates S we know from (i), (ii), (iii) (b), that for most points $y \in G$ there are optimally short circuits approximating them. Find a probabilistic algorithm with expected running time $\text{poly} \log^{1/\epsilon}$, which finds such short circuits. So for the H-T gates this means an algorithm to find for most y 's in G and $\epsilon > 0$ a circuit of T-count at most $(1+\delta)3\log^{1/\epsilon}(y)$ which ϵ -approximates y (arbitrarily small). In practice achieving optimality for most y 's in G , while being only slightly worse for singular y 's, would render these arithmetic S 's to be Golden Gates.

(25)

Appendix 1

We work in the context of a Hecke orbit on S^2 coming from the points a ball in the p -regular tree, as in [L-P-S]. The analysis for $SL(2)$ or more generally the setting of Hecke orbits in cases where the Ramanujan Conjectures are known, is similar.

Let $V_t = \{ S \in O_p \}$ (with $f(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2$) be a set of representatives

for the Hecke points. One can find an orthonormal basis ϕ_j of $L^2(S^2)$ of Hecke eigenfunctions (each ϕ_j is also a spherical harmonic that is an eigenfunction of the Laplacian Δ_{S^2}). If the Hecke orbit in S^2 comes from V_t applied to a fixed point $x_0 \in S^2$ Then

$$\sum_{S \in V_t} \phi_j(Sx_0) = \lambda_j(t) \phi_j(x_0) \quad (25)$$

$$\phi_0 = \frac{1}{\sqrt{4\pi}} \quad \text{and} \quad \lambda_0(t) = |V_t|.$$

$|V_t| = 6 \cdot 5^{t-1}$ for case (i) and for simplicity we stick to this case.

The Ramanujan Conjectures imply that

(26)

$$|\lambda_j(t)| \leq t |V_t|^{1/2} \quad \text{for } j \neq 0 \quad \text{--- (26)}$$

Let $R_\varepsilon(x, y)$ be a point-pair invariant

on S^2 , that is $R_\varepsilon(\sigma x, \sigma y) = R_\varepsilon(x, y)$ for

$\sigma \in O_f$, which we take to be nonnegative and an approximation to the identity

under \star ($k_1 * k_2(z, y) = \int_{S^2} k_1(z, z) k_2(z, y) dA(y)$)

Specifically

$$R_\varepsilon(x, y) \geq 0$$

$$\int_{S^2} R_\varepsilon(x, y) dA(y) = 1.$$

$$R_\varepsilon(x, y) = 0 \quad \text{if } d_{S^2}(x, y) > \varepsilon.$$

We can also choose k to be positive-

definite (i.e. $h_k(t) \geq 0$ where

h_k is the spherical transform of k

at a nonnegative eigenvalue t of Δ_{S^2})

by taking $k = R_{\varepsilon/2} * k_{\varepsilon/2}$ for example.

Moreover we can choose R_ε so that

(27)

$$k_\varepsilon(x, x) \leq \frac{c}{\varepsilon^2} \quad \text{for a fixed constant } c.$$

Expanding $R_\varepsilon(x, y)$ in the basis ϕ_j

yields [Sel]:

$$k_\varepsilon(x, y) = \sum_{j=0}^{\infty} h_{k_\varepsilon}(t_j) \phi_j(x) \phi_j(y) \quad (28)$$

Applying (25) with $x = x_0$ gives.

$$\sum_{S \in V_t} k_\varepsilon(Sx_0, y) = \frac{|V_t|}{\sqrt{4\pi}} + \sum_{j=1}^{\infty} h(t_j) \lambda_j(t) \phi_j(x_0) \phi_j(y)$$

If for some $y \in S^2$

$$\sum_{S \in V_t} k_\varepsilon(Sx_0, y) = 0, \quad \text{then from (29) and (26)}$$

$$\frac{|V_t|}{\sqrt{4\pi}} \leq \sum_{j=1}^{\infty} h(t_j) |\lambda_j(t)| |\phi_j(x_0)| |\phi_j(y)|$$

$$\leq |V_t|^{\frac{1}{2}} \cdot t \sum_{j=1}^{\infty} h(t_j) \frac{|\phi_j(x_0)|^2 + |\phi_j(y)|^2}{2}$$

(28)

$$\leq |V_t| \cdot t \cdot k_{\varepsilon}(z, z)$$

$$\leq |V_t|^{\frac{1}{2}} \cdot t \cdot \frac{C}{\varepsilon^2} \quad \text{--- (30)}$$

using (27).

That is

$$|V_t| \leq \frac{4\pi C t^2}{\varepsilon^4}. \quad \text{--- (31)}$$

Hence if $|V_t| > 4\pi C t / \varepsilon^4$ then forevery y , $\sum_{S \in V_t} k_{\varepsilon}(Sx_0, y) > 0$ and inparticular $d_{S^2}(Sx_0, y) < \varepsilon$ for some $S \in V_t$.This prove that $K \leq 2$ in this case,

and in fact this sharper and non-asymptotic form.

To see that most y 's have optimally good approximation by membersof V_t we compute the variance over y in (29)

$$\int_{S^2} \left| \sum_{S \in V_t} k_{\varepsilon}(Sx_0, y) - \frac{|V_t|}{\sqrt{4\pi}} \right|^2 dA(y)$$

(29)

$$= \sum_{j=1}^{\infty} h_{\varepsilon}(t_j)^2 |\lambda_j(t)|^2 |k_j(x_0)|^2$$

$$\leq t^2 |V_t| \sum_{j=1}^{\infty} h_{\varepsilon}(t_j)^2 |\phi_j(x_0)|^2.$$

$$\leq t^2 |V_t| \int_{S^2} |k_{\varepsilon}(x_0, y)|^2 dA(y)$$

(using (28)).

— (32)

Now from (27) we get that

$$\int_{S^2} \left| \sum_{S \in V_t} k_{\varepsilon}(Sx_0, y) - \frac{|V_t|}{\sqrt{4\pi}} \right|^2 dA(y) \leq \frac{c^2 t^2 |V_t|}{\varepsilon^2}$$

— (33)

If $B = \left\{ y : \sum_{S \in V_t} k_{\varepsilon}(Sx_0, y) = 0 \right\}$ Then $B \supset \left\{ y : d_{S^2}(Sx_0, y) = 0 \text{ for all } S \in V_t \right\}.$

According to (33)

$$\mu_{S^2}(B) \cdot \frac{|V_t|^2}{4\pi} \leq \frac{c^2 t^2 |V_t|}{\varepsilon^2}$$

(30)

or

$$\mu_{S^2}(B) \leq \frac{4\pi c^2 t^2}{\varepsilon^2 |V_t|} \quad (34)$$

This is a quantitative ^{and} explicit form of the fact that B has small measure if $|V_t|$ is somewhat larger than $\frac{1}{\mu_{S^2}(B)}$.



Appendix 2

Let $x, y \in \mathbb{Z}^4$ both not 0 and

$\frac{x}{|x|} \neq \frac{y}{|y|}$. Set $|x|^2 = M$ and $|y|^2 = N$. Then

$$d_{S^3}^2 \left(\frac{x}{|x|}, \frac{y}{|y|} \right) = 2 - \frac{2 \langle x, y \rangle}{\sqrt{MN}}$$

$$= 2 \left(\frac{\sqrt{MN} - \langle x, y \rangle}{\sqrt{MN}} \right)$$

(34)

If $\frac{x}{|x|}$ and $\frac{y}{|y|}$ are close then

$\langle x, y \rangle > 0$ and

$$d_{S^3}^2 \left(\frac{x}{|x|}, \frac{y}{|y|} \right) = \frac{2(MN - \langle x, y \rangle^2)}{\sqrt{MN}(\sqrt{MN} + \langle x, y \rangle)}$$

$$> \frac{MN - \langle x, y \rangle^2}{MN} \quad (35)$$

(31)

In the interval $[MN - \sqrt{MN}, MN]$ there is at most one square number t_{MN} . It follows that if

$$\langle x, y \rangle^2 \neq t_{MN} \text{ Then}$$

$$d_{S^3}^2 \left(\frac{x}{|x|}, \frac{y}{|y|} \right) \geq \frac{1}{\sqrt{t_{MN}}} \quad (36)$$

So given a y as above then for

$$x \text{ with } |x|^2 = M, \text{ either } d \left(\frac{y}{|y|}, \frac{x}{|x|} \right) = \xi_{MN}$$

$$\text{with } 0 < \xi_{MN} < \frac{1}{(MN)^{1/4}} \text{ or } d \left(\frac{x}{|x|}, \frac{y}{|y|} \right) \geq \frac{1}{(MN)^{1/4}}$$

(37)

For $h \leq t$ and x varying over

$$|x|^2 = 5^h,$$

and our given y , we have either

$$d \left(\frac{x}{|x|}, \frac{y}{|y|} \right) = \xi_h, \quad 0 < \xi_h < \frac{1}{(|y|^2 5^h)^{1/4}}$$

$$\text{or } d \left(\frac{x}{|x|}, \frac{y}{|y|} \right) \geq \frac{1}{(|y|^2 5^h)^{1/4}}.$$

Hence there is an annulus $A_{\alpha, \beta}$ about the point $y/|y|$ in S^3

(32)

$$A_{\alpha, \beta} = \left\{ z \in S^3 \mid \alpha \leq d\left(\frac{y}{|y|}, z\right) \leq \beta \right\}$$

with $\beta \leq \frac{1}{(|y|^2 5^t)^{1/4}}$ and $\beta - \alpha \geq \frac{1}{2t \cdot (|y|^2 5^t)^{1/4}}$

which is free of any $x \in V_t$.

Choosing a maximal radius ball B in
this annulus yields (10).



REFERENCES:

[A-K-S] M. Agrawal, N. Kayal, N. Saxena, Ann of Math, 160 (2004) 781-793.

[A-m] K. AMANO, K. MATSUMOTO arXiv 0806.3834

[B-G] J. Bourgain and A. Gamburd, Inv. Math 171 (2008) 83-121

[B-R] J. Bourgain and Z. Rudnick, GAFA, Vol 22 (2012)

[B-R-S] J. Bourgain, Z. Rudnick, P. Sarnak arXiv 1204.0134

[B-S] A. BOCHAROV, K. SVORE arXiv 1206.3223

[B-G-S] A. BOCHAROV, Y. GUREVICH, K. SVORE arXiv 1303.1411

(CH1] P. CHIU COMBINATORICA 12(3) (1992) 275-285

(CH2] P. CHIU JNL NUMBER THEORY 53, 25-44 (1995)

[D-N] C. DAWSON, M. NIELSEN; QUANT INF COMP 6 (2006) 81-95.

[F-G-K-M] Forest, Gosset, Kliuchnikov, Mckinnon, arXiv 1501.04944 33

[G-N-N] A.Ghosh, A. Gorodnik, A. Nevo
arXiv 1205.4426

[K-M] D. Kleinbock, K. Merrill, arXiv 1301-0989

[HA] G. Harmon Jel No' Th. 34, 63-81 (1990)

[K-M-M 1,2] V. Kliuchnikov, D. Maslov, M. Mosca
arXiv 1212.6964, arXiv 1212.0822.

[L-P-S 1,2] A. Lubotzky, R. Phillips and P. Sarnak
COMM PURE AND APPL. MATH XXXIX, 149-186 (1986)
XL, 401-420 (1987).

[N-S] M. Nielsen and I. Chuang "Quantum computation and quantum information" CUP
[R] G. Robinson, Jul of Alg. 306 (2006) 201-207. (2000)

[R-S] N. Ross, P. Selinger arXiv 1403.2975

[Sa] P. Sarnak "Notes on THIN MATRIX GROUPS" MSRI
PUBL. 61 (2014) 343-362.

[Sc] R. Schoof Math. Comp. 44 (1985) 483-494

[S] P. Selinger arXiv 1212.6253

[Sel] A. Selberg J. Ind. Math. Soc (1956) 20, 47-87,

[Ser1] J.-P. Serre "Le groupe fondamental, un groupe arithmétique"
Oberwolfach Report 2011?

[Ser2] J.-P. Serre "Arbres, amalgames, SL_2 ", Asterisque 46 (1977)
Regards Peter Sarnak