ELSEVIER

Discrete Mathematics 223 (2000) 355–362

DISCRETE MATHEMATICS

www.elsevier.com/locate/disc

Note

Three distance theorem and grid graph

Nicolas Chevallier

Universite de Haute Alsace 4, rue des Freres Lumiere, 68093 Mulhouse, France

Received 5 May 1999; revised 3 August 1999; accepted 24 January 2000

Abstract

We will prove a d-dimensional version of the Geelen and Simpson theorem. © 2000 Elsevier Science B.V. All rights reserved.

Keywords: Irrational rotation; Three distance theorem; Graph; Connected component

1. Introduction

Let θ be in $\mathbb{T}^1=\mathbb{R}/\mathbb{Z}$, the points $0, \theta, 2\theta, \dots, n\theta$ divide $\mathbb{T}^1$ into $n+1$ intervals having at most three distinct lengths. This property is known as the three distance theorem conjectured by Steinhaus. A first generalisation was conjectured by Graham.

Let $\theta, \alpha_1, \dots, \alpha_d$ be in $\mathbb{T}^1$ and $n_1, \dots, n_d$ be positive integers. Then the points $\alpha_i + k\theta$, $i = 1, \dots, d$, $k = 0, \dots, n_i - 1$, divide $\mathbb{T}^1$ into intervals having at most $3d$ distinct lengths.

It was first proved by Chung and Graham in 1976 [4] but three years later Liang found a very simple proof [6]. In 1993, Geelen and Simpson [5] proved a two-dimensional version of the three distance theorem.

Let θ_1, θ_2 be in $\mathbb{T}^1$ and n_1, n_2 be positive integers. Then the points $k_1\theta_1 + k_2\theta_2$, $k_1 = 0, \dots, n_1 - 1$, $k_2 = 0, \dots, n_2 - 1$, divide $\mathbb{T}^1$ into intervals having at most $n_1 + 3$ distinct lengths.

Geelen and Simpson's proof is long and complex. They also conjectured the following d-dimensional generalisation.

E-mail address: n.chevallier@univ-mulhouse.fr (N. Chevallier).

PII: S0012-365X(00)00098-4

Let $\theta_1,\dots,\theta_d$ be in $\mathbb{T}^1$ and $n_1,\dots,n_d$ be positive integers. Then the points $k_1\theta_1+\cdots+k_d\theta_d$, $k_1=0,\dots,n_1-1,\dots,$ $k_d=0,\dots,n_d-1$, divide $\mathbb{T}^1$ into intervals having at most $\prod_{i=1}^{d-1} n_i + C_d$ distincts lengths where C_d depends only on d.

In this work we will prove a result which is close to this conjecture.

Theorem 1. *Suppose $d\geqslant 2$. Let $\theta_1,\dots,\theta_d$ be in $\mathbb{T}^1$ and $n_1,\dots,n_d$ be positive integers. Then the points $k_1\theta_1+\cdots+k_d\theta_d$, $k_1=0,\dots,n_1-1,\dots,$ $k_d=0,\dots,n_d-1$, divide $\mathbb{T}^1$ into intervals having at most $\prod_{i=1}^{d-1} n_i+3\prod_{i=1}^{d-2} n_i+1$ distinct lengths.*

If $d=2$ our result is not as good as Geelen and Simpson's; the upper bound is n_1+4 instead of n_1+3. Our proof decomposes in three steps. The first step (Lemma 2) uses the argument of Liang's proof of the three d distance theorem [6]. A combinatorial formulation of the same argument using the Rauzy graph of words, can be found in an article by Alessandri and Berthé [1]. In the d-dimensional torus, a partial extension of three distance theorem using Voronoï diagrams was proved in [2,3], the basic argument is also very similar to Liang's. The second step uses the symmetry of $\mathbb{T}^1$ $x\to -x+\sum_{i=1}^{d}(n_i-1)\theta_i$, this symmetry was already used in [1,2,5]. The aim of the third step is mainly to estimate the number of connected components of a subgraph of the standard grid graph on $\mathbb{Z}^d$.

2. Proof of Theorem 1

Notations. We choose the positive orientation on $\mathbb{T}^1=\mathbb{R}/\mathbb{Z}$. An interval $[a,b]$ of $\mathbb{T}^1$ is defined with respect to this orientation.

If A is a set, $|A|$ denotes its cardinality.

Put $E=\{\sum_{i=1}^{d} k_i\theta_i\colon 0\leqslant k_i<n_i,\ i=1,\dots,d\}$. For each $x\in E$ we call $\mathrm{suc}(x)$ the next element of E after x, i.e. such that $]x,\mathrm{suc}(x)[\,\cap E$ is empty.

First, we show we can suppose that $1,\theta_1,\dots,\theta_d$ are independent over $\mathbb{Q}$. Let δ be the smallest difference between the length of two intervals of different lengths. We can choose $\theta_1',\dots,\theta_d'$ sufficiently close to the θ_i such that the Hausdorff distance between E and $E'=\{\sum_{i=1}^{d} k_i\theta_i'\colon 0\leqslant k_i<n_i,\ i=1,\dots,d\}$ is smaller than $\delta/6$, we can also suppose that $1,\theta_1',\dots,\theta_d'$ are independent over $\mathbb{Q}$. In this case, if $]x,y[$ is an interval of $\mathbb{T}^1\setminus E$ there exists an interval $]x',y'[$ of $\mathbb{T}^1\setminus E'$ such that the difference between their lengths is smaller than $\delta/3$. Hence, if $\{]x_1,y_1[,\dots,]x_N,y_N[\}$ is a set of intervals of $\mathbb{T}^1\setminus E$ of different lengths then we can find a set $\{]x_1',y_1'[,\dots,]x_N',y_N'[\}$ of intervals of $\mathbb{T}^1\setminus E'$ whose lengths are all different.

From now, we assume that $1,\theta_1,\dots,\theta_d$ are independent over $\mathbb{Q}$. Put $f(x_1,\dots,x_d)=\sum_{i=1}^{d} x_i\theta_i$, the map f is one to one on the set

$$R=\{(x_1,\dots,x_d)\in\mathbb{Z}^d\colon 0\leqslant x_i<n_i,\ i=1,\dots,d\}$$

and $f(R)=E$. Therefore, the intervals $I(X)=]f(X),\mathrm{suc}(f(X))[$, $X\in R$, are all distinct. Let $e_1,\dots,e_d$ be the canonical basis of $\mathbb{R}^d$. Consider the non-oriented graph G whose vertices are the elements $X=(x_1,\dots,x_d)$ of R and whose edges are the pairs $\{X,X+e_i\}$ such that $I(X)+\theta_i=I(X+e_i)$. If X and Y are in the same connected component of G then $I(X)$ and $I(Y)$ have the same length so our aim is to estimate the number of connected component of G. The graph G is a subgraph of the grid graph R whose edges are all the pairs $\{X,X+e_i\}$. We will call a pair $\{X,X+e_i\}$ an i-edge. The graph G is obtained from the graph R by removing some of its edges. The number of removed edges is given by the following lemma whose proof used Liang's idea.

Lemma 2. *The number of removed i-edges is smaller than* $2\Pi_{j\neq i}n_j$.

Proof. Let X and $X'=x+e_i$ be two vertices of G. If $\{X,X'\}$ is not an edge of G then either $I(X)+\theta_i\nsubseteq I(X')$ or $I(X')-\theta_i\nsubseteq I(X)$. Suppose $I(X)+\theta_i\nsubseteq I(X')$. This means that the interval $]f(X)+\theta_i,\ \mathrm{suc}(f(X))+\theta_i[\,=\,]f(X'),\ \mathrm{suc}(f(X))+\theta_i[$ contains the point $\mathrm{suc}(f(X'))$. Then $\mathrm{suc}(f(X'))-\theta_i$ is in the open interval $I(X)$ and therefore is not in E. Call Y the element of R such that $f(Y)=\mathrm{suc}(f(X'))$. We have just shown that $f(Y-e_i)$ is not in E which means that $Y-e_i$ is not in R. Then $Y-e_i$ must belong to the face $F_i=\{(x_1,\dots,x_d)\in\mathbb{N}^d\colon x_i=-1 \textit{ and } 0\leqslant x_j<n_j,\ j\neq i\}$. The map $\{X,X'\}\to Y-e_i\in F_i$ is such that $f(Y-e_i)\in I(X)$. Since the intervals $I(X)$, $X\in R$, are all distinct, each point $f(Z)$, $Z\in F_i$, belongs to at most one interval $I(X)$, this shows that there is at most card $F_i=\Pi_{j\neq i}n_j$ vertices X of G such that, $X+e_i\in E$ and $I(X)+\theta\nsubseteq I(X+e_i)$. The same reasoning shows that there are at most $\Pi_{j\neq i}n_j$ points X of R such that, $X+e_i\in E$ and $I(X+e_i)-\theta\nsubseteq I(X)$. Therefore, the number of removed i-edges is at most $2\Pi_{j\neq i}n_j$. □

Some different components of G correspond to intervals of the same length because of the symmetry $s\colon x\in\mathbb{T}^1\to -x+\sum_{i=1}^{d}(n_i-1)\theta_i$, this is the content of the next lemma. Call $\mathscr{C}$ the set of connected components of G. For each $C\in\mathscr{C}$, let $l(C)$ be the cardinality of the projection of C on the axis $\mathbb{R}e_d$ and $A(C)$ be the cardinality of the projection of C on the hyperplane $\mathbb{R}e_1+\cdots+\mathbb{R}e_{d-1}$. Call $\mathscr{C}_1=\{C\in\mathscr{C}: A(C)=1\}$, $\mathscr{C}_2=\{C\in\mathscr{C}: A(C)\geqslant 2\}$, $N_1=|\mathscr{C}_1|$ and $N_2=|\mathscr{C}_2|$.

Lemma 3. *The number of lengths of the intervals of* $\mathbb{T}^1\backslash E$ *is smaller than* $2+\frac{1}{2}N_1+N_2$.

Proof. We reproduce an argument of Berthe and Allessandri (cf. [1, Theorem 18]). Let $C\in\mathscr{C}_1$, $C=\{X_0,X_1=X_0+e_d,\dots,X_k=X_0+ke_d\}$. Since $s(E)=E$ the intervals $s(I(X_0)),\dots,s(I(X_k))$ are intervals of $\mathbb{T}^1\setminus E$, then there exists $Y_0\in G$ such that $s(I(X_k))=I(Y_0)$, $s(I(X_{k-1}))=I(Y_0)+\theta_d,\dots,s(I(X_0))=I(Y_0)+k\theta_d$. Suppose $Y_0=X_0$ then there are only two possibilities:

(1) There exists i such that $s(I(X_i))=I(X_i)$. Therefore, the middle of the interval $I(X_i)$ must be a fixed point of s.
(2) If $s(I(X_i))\neq I(X_i)$ for all i then k is even and $s(I(X_{k/2}))=I(X_{k/2+1})=I(X_{k/2})+\theta_d$. Therefore, the middle of the interval $I(X_{k/2})$ must be a fixed point of the symmetry $s'(x)=s(x)-\theta_d$.

Since the symmetries s and s' have 2 fixed points each, there are at most 4 components in $\mathscr{C}_1$ such that $Y_0 = X_0$. Furthermore, if $Y_0 \neq X_0$ then the length of the intervals $I(X)$, $X \in C$, is the same than the length of the intervals of another component and the lemma follows. □

Let us now collect the constraints of the numbers $l(C)$ and $A(C)$.

First, we add to the graph R the boundary edges (i.e. the pairs $\{(x_1,\dots,x_d), (x_1,\dots,x_d) \pm e_i\}$ with $(x_1,\dots,x_d) \in R$ and $x_i = 0$ or $n_i - 1$, $i = 1,\dots,d$). With these new edges, to each vertex of R, there correspond two i-edges, and G is obtained from R by removing $4\Pi_{j\neq i} n_j$ i-edges instead of $2\Pi_{j\neq i} n_j$. Call $\mathscr{A}$ the set of edges of R. Consider the set

$$\mathscr{A}_i = \{(a, C) \in \mathscr{A} \times \mathscr{C}\colon a \text{ is a boundary } i\text{-edge of } C\}.$$

To each removed i-edge corresponds two vertices and therefore two components and to each new i-edge corresponds one component, thus

$$|\mathscr{A}_i| \leqslant 2(2\Pi_{j\neq i} n_j) + 2\Pi_{j\neq i} n_j.$$

Take $i = d - 1$. If $C \in \mathscr{C}$ contains the point $X = (a_1,\dots,a_d)$ then the hyperplan $x_d = a_d$ must contain at least two i-edges of the boundary of C. The number of such hyperplane is $l(C)$, then

$$\sum_{C\in\mathscr{C}} 2l(C) \leqslant |\mathscr{A}_{d-1}| \leqslant 6\, \Pi_{j\neq d-1}\, n_j,$$

thus

$$\sum_{C\in\mathscr{C}} l(C) \leqslant 3 \prod_{j\neq d-1} n_j.$$

Now take $i = d$. If $C \in \mathscr{C}$ contains the point $X = (a_1,\dots,a_d)$ then the line $x_1 = a_1,\dots,x_{d-1} = a_{d-1}$ must contain at least two d-edges of the boundary of C. The number of such lines is $A(C)$ then

$$\sum_{C\in\mathscr{C}} 2A(C) \leqslant |\mathscr{A}_d| \leqslant 6 \prod_{j<d} n_j.$$

Put

$$P = \prod_{j<d-1} n_j \quad (P = 1 \text{ if } d = 2),$$

we have

$$\sum_{C\in\mathscr{C}} A(C) \leqslant 3 \prod_{j<d} n_j,$$

$$\sum_{C\in\mathscr{C}} l(C) \leqslant 3Pn_d.$$

Secondly, if $C \in \mathscr{C}$ we have $|C| \leqslant l(C)A(C)$ then

$$\sum_{C\in\mathscr{C}} l(C)A(C) \geqslant \sum_{C\in\mathscr{C}} |C| = \prod_{i=1}^{d} n_i.$$

Thirdly, it is obvious that $l(C) \leqslant n_d$ for each $C \in \mathscr{C}$. The theorem will follow from the previous lemma and the following lemma.

Lemma 4. *Put* $N = N_1 + N_2$. *If* $s_1, \dots s_N$ *and* $h_1, \dots, h_N$ *are strictly positive integers such that*

(1) $\sum_{i=1}^{N} s_i h_i \geqslant \Pi_{i=1}^{d} n_i$,

(2) $\sum_{i=1}^{N} s_i \leqslant 3\Pi_{j=1}^{d-1} n_j$,

(3) $\sum_{i=1}^{N} h_i \leqslant 3Pn_d$,

(4) $\forall i \in \{1, \dots, N\}, \quad h_i \leqslant n_d$,

(5) $\begin{cases} \forall i \in \{1, \dots, N_1\}, & s_i = 1 \\ \forall i \in \{N_1 + 1, \dots, N\}, & s_i \geqslant 2 \end{cases}$,

then

$$\min\left(N, \left[\frac{1}{2}N_1 + N_2 + 2\right]\right) \leqslant \prod_{i=1}^{d-1} n_i + 3P + 1$$

Proof. If we order the s_i and the h_i in increasing order then $\sum_{i=1}^{N} s_i h_i$ will increase, so (1)–(5) are still verified. If $h_N < n_d$ and one of the h_i, $i < N$, is > 2, we can replace h_N by $h_N + 1$ and h_i by $h_i - 1$, because this increases $\sum_{i=1}^{N} s_i h_i$. Repeating this process we can assume that

$$h_1 = \cdots = h_{N-k-1} = 1, \quad h_{N-k} < n_d,$$
$$h_{N-k+1} = \cdots = h_N = n_d,$$

where k is an integer not exceeding $3P$. Moreover, if $N \leqslant 3P + 1$ there is nothing to prove so we can assume that $k \leqslant 3P \leqslant N - 2$. The idea of the proof lies in the following simple calculus. Put $S = \sum_{i=1}^{N-k-1} s_i$. By (1) we have

$$S + \sum_{i=N-k}^{N} h_i s_i \geqslant \prod_{i=1}^{d} n_i,$$

then by (2)

$$S + n_d \left(3 \prod_{i=1}^{d-1} n_i - S\right) \geqslant S + n_d \sum_{i=N-k}^{N} s_i \geqslant \prod_{i=1}^{d} n_i.$$

Solving this inequality we get

$$S \leqslant \frac{2}{n_d - 1} \prod_{i=1}^{d} n_i.$$

On the other hand, if $N_1 < N - k - 1$ we have

$$N_1 + 2N_2 \leqslant \sum_{i=1}^{N_1} s_i + \sum_{i=N_1+1}^{N-k-1} s_i + \sum_{i=N-k}^{N} 2 = S + 2(k+1)$$

and if $N_1 \geqslant N - k - 1 = N_1 + N_2 - k - 1$ we have

$$N_1 + 2N_2 = N + N_2 \leqslant N + k + 1 \leqslant S + 2(k+1).$$

In both cases

$$N_1 + 2N_2 \leqslant S + 2(k+1),$$

then

$$\frac{1}{2}N_1 + N_2 \leqslant \frac{1}{2}S + k + 1 \leqslant \frac{1}{n_d - 1}\prod_{i=1}^{d} n_i + 3P + 1 \approx \prod_{i=1}^{d-1} n_i + 3P + 1.$$

In order to prove the lemma we have to consider different cases.

If $k = 3P$ then by (3) we have $N = k = 3P$ and

$$N = 3P \leqslant \prod_{i=1}^{d-1} n_i + 3P + 1,$$

so we can assume $k < 3P$. If $s_{N-k} > 2$ we can replace s_N by $s_N + s_{N-k} - 2$ and s_{N-k} by 2, so we can assume $s_{N-k} = 1$ or 2 (the sequence (s_n) may now loose its monotonicity, but it does not matter for the proof).

Case 1: $s_{N-k} = 1$.

By (5) we have $N_1 \geqslant N - k$ and $S = N - k - 1$. Consequently, $N_2 \leqslant k$ and $N_1 + 2N_2 = N + N_2 \leqslant N + k = S + 1 + 2k$. By (1)

$$\prod_{i=1}^{d} n_i \leqslant S + h_{N-k} + n_d \sum_{i=N-k+1}^{N} s_i = S + n_d \sum_{i=N-k}^{N} s_i + h_{N-k} - n_d,$$

with (2) we get

$$\prod_{i=1}^{d} n_i \leqslant S + n_d \left(3\prod_{i=1}^{d-1} n_i - S \right) + h_{N-k} - n_d,$$

and

$$S(n_d - 1) \leqslant 2\prod_{i=1}^{d} n_i + h_{N-k} - n_d.$$

By (3) we have $h_{N-k} \leqslant 3Pn_d - kn_d - (N - k - 1)$. With the previous inequality this gives

$$\begin{aligned} S(n_d - 1) &\leqslant 2\prod_{i=1}^{d} n_i + 3Pn_d - kn_d - N + k + 1 - n_d \\ &= 2\prod_{i=1}^{d} n_i + n_d(3P - k - 1) + k + 1 - N, \end{aligned}$$

but $S+k+1=N$ then

$$S\leqslant\frac{1}{n_d}\left(2\prod_{i=1}^{d}n_i+n_d(3P-k-1)\right)=2\prod_{i=1}^{d-1}n_i+3P-k-1,$$

and

$$\begin{aligned}\frac{1}{2}N_1+N_2+2 &\leqslant \frac{1}{2}S+\frac{1}{2}+k+2\leqslant\prod_{i=1}^{d-1}n_i+\frac{3}{2}P+\frac{k}{2}+2\\ &\leqslant\prod_{i=1}^{d-1}n_i+\frac{3}{2}P+\frac{3P-1}{2}+2=\prod_{i=1}^{d-1}n_i+3P+\frac{3}{2}\end{aligned}$$

for $k<3P$.

Case 2: $s_{N-k}=2$.

We have $\frac{1}{2}N_1+N_2\leqslant\frac{1}{2}S+k+1$. Since

$$\prod_{i=1}^{d}n_i\leqslant S+s_{N-k}h_{N-k}+n_d\sum_{i=N-k+1}^{N}s_i=S+n_d\sum_{i=N-k}^{N}s_i+2(h_{N-k}-n_d),$$

by (2) we get

$$\prod_{i=1}^{d}n_i\leqslant S+n_d\left(3\prod_{i=1}^{d-1}n_i-S\right)+2(h_{N-k}-n_d),$$

and

$$S(n_d-1)\leqslant 2\prod_{i=1}^{d}n_i+2(h_{N-k}-n_d).$$

By (3) we have $h_{N-k}\leqslant 3Pn_d-kn_d-(N-k-1)$. With the previous inequality this gives

$$S(n_d-1)\leqslant 2\left(\prod_{i=1}^{d}n_i+3Pn_d-kn_d-N+k+1-n_d\right).$$

If $N\geqslant\prod_{i=1}^{d-1}n_i+3P+1$ then

$$\begin{aligned}S(n_d-1) &\leqslant 2\left(\prod_{i=1}^{d}n_i+3Pn_d-kn_d-\left(\prod_{i=1}^{d-1}n_i+3P+1\right)+k+1-n_d\right)\\ &\leqslant 2(n_d-1)\left(\prod_{i=1}^{d-1}n_i+3P-k-1\right)-2\end{aligned}$$

and

$$\begin{aligned}\frac{1}{2}N_1+N_2+2 &\leqslant \frac{1}{2}S+k+3\leqslant\prod_{i=1}^{d-1}n_i+3P-k-1-\frac{1}{n_d-1}+k+3\\ &<\prod_{i=1}^{d-1}n_i+3P+2. \qquad \square\end{aligned}$$

In the two-dimensional case the upper bound $\prod_{i=1}^{d-1} n_i + 3P + 1 = \prod_{i=1}^{d-1} n_i + 4$ is not the best. Our method of proof cannot reach the optimal bound given by Geelen and Simpson. To see this, let us take an example.

$n_1 = 4$ and $n_2 = 5$ with the following removed edges: removed 1-edges = all the 1-edges inside $\{0\} \times \{0,\dots,3\}$ and all the 1-edges $\{(1,k),(2,k)\}$ $k = 1,\dots,4$; removed 2-edges = all the 2-edges between $\{k\} \times \{0,\dots,3\}$ and $\{k+1\} \times \{0,\dots,3\}$, $k = 0$ and 2. The inequalities are verified and there are 8 connected components with $N_1 = N_2 = 4$ and $\frac{1}{2}N_1 + N_2 + 2 = 8 > 4 + 3$.

When $d \geqslant 3$ our proof uses only the constraints on the $(d-1)$-edges and the d-edges, so our bound $\prod_{i=1}^{d-1} n_i + 3P + 1$ is probably not optimal.

References

[1] P. Allessandri, V. Berthé, Three distance theorem and combinatorics on words, Enseign. Math. (1998).
[2] N. Chevallier, Meilleures approximations d'un élément du tore $\mathbb{T}^2$ et géométrie de la suite des multiples de cet élément, Acta Arith. 78 (1996) 19–35.
[3] N. Chevallier, Géométrie des suites de Kronecker, Manuscripta Math. 94 (1997) 231–241.
[4] F.R.K. Chung, R.L. Graham, On the set of distances determined by the union of arithmetic progressions, Ars Combin. 1 (1976) 57–76.
[5] J.F. Geelen, R.J. Simpson, A two dimensional Steinhaus theorem, Australas. J. Combin. 8 (1993) 169–197.
[6] F.M. Liang, A short proof of the 3d distance theorem, Discrete Math. 28 (1979) 325–326.