A Simple Proof of the Joints Problem

Xiaoqing Tang

May 9, 2012

Abstract

After Zeev Dvir proved the Kakeya conjecture over finite fields [1] using polynomial method, people have been inspired and applied this method to other problems. Shortly after Dvir's proof, Larry Guth and Nets Hawk Katz proved the optimal upper bound for the joints problem [2]. However, the proof uses rather complicated combinatorics to prepare for the polynomial method. In this paper, we present a simpler approach using less combinatorics without introducing more algebraic geometry background.

1 Introduction

Joints problem is formulated as follows. Given N lines in $\mathbb{R}^3$ space. Define a joint to be a point where at least 3 non-colinear intersects. The joints problem asks, what is the upper bound for the number of joints. As mentioned in [2], if we consider a cube of length $\sqrt{\frac{N}{3}}$, picking the N lines to be parallel to one of the 3 axes and to intersect the integer lattice points, then all integer latice points in the cube are joints, which gives $c'N^{\frac{3}{2}}$ joints for a constant c'.

The theorem which Guth and Katz presents in [2] concludes that, this is the optimal case, and one can never exceed this quantity.

Theorem 1.1. *Any set of N lines in $\mathbb{R}^3$ form at most $O(N^{\frac{3}{2}})$ joints.*

Our strategy is similar as what Guth and Katz did. If L is a set of lines, we first try to pick a subset L' nice enough, so that we can construct a polynomial f which vanishes at all lines in L'. Then we apply the polynomial method. More precisely, after picking a L' nice enough, we want to construct a square-free polynomial (it's irreducible in Guth and Katz's proof [2]) f which vanishes on all lines and has some degree d low enough. Since each joint is the intersection of non-colinear lines, it would follow that these joints are critical points. With enough joints on each line, any line is actually a critical line of f. However, if $deg(p) = d$ and f square-free, there could be at most $d(d-1)$ such critical lines, which results in a contradiction. We present the detailed proof in section 3, after we introduce some basic lemmas.

2 Background

We state some necessary facts and lemmas about polynomials before we prove the theorem.

For the following lemma, we follow the proof presented in [3].

Lemma 2.1. *Let k be a field, and $E \subseteq k^n$ such that $|E| < \binom{n+d}{n}$ for some d. Then there exists a non-zero polynomial f in n variables of degree at most d such that f vanishes at all points of E.*

Proof. Now, let V be the vector space of all polynomials in $F[x_1, x_2, \cdots, c_n]$ of degree at most d over F. We claim that $dim_F(V) = \binom{n+d}{n}$. Clearly $S = \{\prod_{i=1}^{n} x_i^{a_i} | \sum_{i=1}^{n} a_i \leq d\}$ is a basis of V over F, so we only need to count $|S|$, i.e. how many non-negative integer solutions to $\sum_{i=1}^{n} a_i \leq d$. We represent a solution by a configuration of n black marbles and d white marbles put in a line, while a_i is denoted by the number of consecutive white marbles on the left of the i-th black marble. It's easy to check that this correspondence is one-to-one and onto, so $|S| = \binom{n+d}{n}$ since there're exactly $\binom{n+d}{n}$ such marble configurations.

Knowing $dim_F(V)$, we set an evaluation map $V \to F^E$ to be $P \mapsto (P(x))_{x \in E}$. Clearly $dim_F(F^E) = |E| < \binom{n+d}{n}$. This implies that this linear map has a nonzero kernel. QED □

Next, we want to prove that square-free polynomials cannot have too many critical lines on it.

Proposition 2.2. *Suppose $f \in \mathbb{R}[x_1, \cdots, x_n]$ is square-free. Then there exists an $\mathbb{R}$-linear combination $q = \sum_{i=1}^{n} c_i \frac{\partial}{\partial x_i} f$ where $c_i \in \mathbb{R}$ such that q and f don't have any common factor.*

Proof. First, we write $f_i = \frac{\partial}{\partial x_i} f$. Suppose there exists an irreducible $p \in \mathbb{R}[x_1, \cdots, x_n]$ such that $p|f$ and $p|f_i$ for all i. Write $p_i = \frac{\partial}{\partial x_i} p$, $f = pg$, $g_i = \frac{\partial}{\partial x_i} g$, and $f_i = ph_i$. We get $ph_i = \frac{\partial}{\partial x_i}(pg) = p_i g + pg_i$. This gives

$$p(h_i - g_i) = p_i g \tag{1}$$

Since p is irreducible, p is not unit $\mathbb{R}[x_1, \cdots, x_n]$, so $deg(p) \geq 1$, so there exists an i such that p has positive degree with respect to x_i. We now focus on this i. Clearly $p_i \neq 0$ and $deg(p_i) \leq deg(p) - 1$. These imply $p \nmid p_i$. Combine this with equation (1), we conclude that $p|g$. This implies $p^2|f$, a contradiction. This means there there is no single irreducible that divides f and all f_i's simultaneously.

Now, factorize $f = \prod_{i=1}^{k} p_i$. Let V be the $\mathbb{R}$-vector space generated by all f_i. It's easy to check that for any irreducible polynomial p, $\{g \in V | p \text{ divides } g\}$ is a subspace of V. Let $V_i = \{g \in V | p_i \text{ divides } g\}$

where $i \in \{1, 2, \cdots, k\}$. By what we have just proved, p_i does not divide all f_i for any i, so V_i is a proper subspace of V. Since a finite union of proper subspace over infinite field is always proper, we can find a polynomial $q \in V$ such that $q \notin V_i$ for all i. This means q does not have any factor in common with f. □

We now quote a corollary from [2] as a proposition here.

Proposition 2.3. *Let $f, g \in \mathbb{R}[x_1, x_2, x_3]$, and $deg(f) = l, deg(g) = m$. Suppose that f, g simultaneously vanish on more than lm lines, then f, g have a common factor.*

Proof. See [2], corollary 2.5. □

Now we state a more generalized result of proposition 3.1 of [2]. Suppose $f \in \mathbb{R}[x_1, x_2, x_3]$, $deg(f) = d > 0$, $S = \{x \in \mathbb{R}^3 | f(x) = 0\}$. We call a point $a \in S$ is critical if $\nabla f(a) = 0$, and a line l is critical if for all $x \in l$, x is critical.

Lemma 2.4. *If f is a square-free polynomial in $\mathbb{R}[x_1, x_2, x_3]$ of degree d and S is its vanishing set. Then S contains at most $d(d-1)$ critical lines.*

Proof. By lemma 2.2, there exists an $\mathbb{R}$-linear combination q of components of ∇f that has no common factor with f. But any such linear combination will have degree at most $\max deg(\frac{\partial}{\partial x_i} f) \leq d - 1$. In particular, $deg(q) \leq d - 1$. But when we apply q and f to proposition 2.3, we get a common factor if S contains more than $d(d-1)$ critical lines. □

In addition, we prove one more proposition from [2] in full here. We call a point regular if it's in the vanishing set but not critical.

Proposition 2.5. *Let S be the vanishing set of polynomial f, and let a be a regular point in S. Suppose $a \in l \subseteq S$, then $l \subseteq T_a S$ where $T_a S$ is the tangent plane to S at a.*

Proof. Let the line be expressed as $a + xt$ where $x \in \mathbb{R}^3$ be the direction vector and $t \in \mathbb{R}$. Let $g(t) = f(a + xt)$. Clearly by chain rule, $g'(t) = x \cdot \nabla f(a)$. But since f vanishes on line $a + xt$, it follows that g is zero. So trivially $g' = 0$, which implies $x \cdot (\nabla f(a)) = 0$. Since $\nabla f(a)$ is nonzero, $\nabla f(a)$ denotes a normal vector of $T_a S$. But $x \cdot (\nabla f(a)) = 0$ would imply that x has to lie on the plane perpendicular to the normal vector $\nabla f(a)$, which gives the result. □

This has an immediate corollary which we will use in the main proof

Corollary 2.6. *Let S be the vanishing set of polynomial f, and let a be a joint point of three non-colinear lines l_1, l_2, l_3 where $l_1, l_2, l_3 \subseteq S$. Then a is a critical point.*

Proof. If a is regular, then by proposition 2.5, l_1, l_2, l_3 should be on the plane $T_a S$, which is a contradiction. □

3 Proof of Joints Problem

Now we are ready to present the proof for theorem 1.1. Suppose the theorem does not hold, then for any K, there exist N lines such that there are more than $KN^{\frac{3}{2}}$ joints defined by them. We'll first study arbitrary K, and then decide a large enough K to derive contradiction.

Instead of the long combinatorics done in [2], we do the following. Let L be the set of N lines. For any subset $S \subseteq L$, let $J(S)$ be the joints defined by lines in S. Since there are only finitely many subsets of L, we can find a subset L' that maximizes the quantity $\frac{|J(S)|}{|S|^{\frac{3}{2}}}$ among all non-empty subsets of L. Without loss of generality, we can assume $L = L'$, since otherwise we can consider the joint problem on L' instead, while the relation $|J(L')| > K|L'|^{\frac{3}{2}}$ still holds because of the maximality.

Proposition 3.1. *Under the above setting of L, for any $l \in L$, l is incident to at least $K\sqrt{N}$ joints in $J(L)$*

Proof. Pick a $l \in L$, and let $L_0 = L - \{l\}$. Clearly $|L_0| = N - 1$. Suppose l is incident to x joints, then $|J(L_0)| \geq |J(L)| - x$. Furthermore, by the maximality of L, we get $\frac{|J(L)|}{|L|^{\frac{3}{2}}} \geq \frac{|J(L_0)|}{|L_0|^{\frac{3}{2}}}$. Combine these together, we get

$$\frac{|J(L)|}{N^{\frac{3}{2}}} \geq \frac{|J(L_0)|}{(N-1)^{\frac{3}{2}}} \geq \frac{|J(L)| - x}{(N-1)^{\frac{3}{2}}}$$

This implies

$$x \geq |J(L)|(1 - (1 - \frac{1}{N})^{\frac{3}{2}}) \geq |J(L)|(1 - (1 - \frac{1}{N})) = \frac{|J(L)|}{N} > K\sqrt{N}$$

where the second inequality is justified since $1 - \frac{1}{N} \in (0, 1)$ and $\frac{3}{2} > 1$. □

Now we use a probabilistic method to construct a nice set X, using the idea in claim 9 of [5].

Proposition 3.2. *Under the above setting of L, and assume K large enough. Then there exists a subset $X \subseteq L$, such that $|X| \leq \frac{1}{6}N$, and that for any line $l \in L$, l intersects at least $\sqrt{N} + 1$ lines in X and the intersections are distinct.*

This proposition is similar to what's in [2]. But we'll present a full proof here.

Proof. We pick a random subset L'' of L by selecting each line at probability $\frac{1}{12}$, and let $J(L'')$ be the joint formed by lines in L. By lemma 3.1, each line in L has at least $K\sqrt{N}$ joints on it. On the other hand, the probability that any joint stays in $J(L'')$ is at least $\frac{1}{12^3}$, since we only need to choose at least 3 of the many lines that passes through the joint.

Now, let X_l be the event that line $l \in L$ has at least $\frac{K}{12^9}\sqrt{N}$ joints in $J(L'')$. By comparing our distribution with the binomial distribution on $\{j \in J(L) | j \in l\}$ with each joint chosen at probability

exactly $p = \frac{1}{12^3}$, we get the following (assuming that there are $u\sqrt{N}$ joints on l where $u > K$):

$$Pr[X_l] \geq 1 - \sum_{k=0}^{\frac{K}{12^9}\sqrt{N}} \binom{u\sqrt{N}}{k} p^k (1-p)^{n-k} \geq 1 - exp(-\frac{1}{2p}\frac{(u-\frac{K}{12^9})^2}{u}\sqrt{N}) \geq 1 - exp(-cK\sqrt{N})$$

where c is some universal constant. Now we claim that for any proper subset $A \subset L$ and $l_0 \in L - A$, the following holds:

$$Pr[X_{l_0} | \bigcap_{l \in A} X_l] \geq Pr[X_{l_0}]$$

(Proof?)

But then, by Bayes' rule and induction, we immediately get $Pr[\bigcap_{l \in L} X_l] \geq (1 - exp(-cK\sqrt{N}))^N \geq 1 - N \cdot exp(-cK\sqrt{N})$. Notice that the probability that $|L''| > \frac{1}{6}N$ is at most a constant (again by comparing it with the binomial distribution mentioned above). So by picking K large enough, the expression $N \cdot exp(-cK\sqrt{N}$ will be small enough regardless the choice of N. So we can have the event $\bigcap_{l \in L} X_l$ and $|L''| \leq \frac{1}{6}N$ all happen at a positive probability, so there exists one such set L''. $\square$

Now, by proposition 3.2, we pick a set X such that $|X| \leq \frac{1}{6}N$, and for any line $l \in L$, l intersects at least $\sqrt{N}+1$ lines in X and the intersections are distinct. By lemma 3.1, each line in X has at least $K\sqrt{N}$ joints, so we arbitrarily pick $\sqrt{N}+1$ joints and form a set of joints Y. Clearly $|Y| \leq (\sqrt{N}+1)|X| \leq \frac{N(\sqrt{N}+1)}{6}$. We now try to construct a polynomial f vanishing on all points of Y. By lemma 2.1, since $\binom{\sqrt{N}+3}{3} > \frac{N(\sqrt{N}+1)}{6} \geq |Y|$, we can actually construct f with degree at most $\sqrt{N}$. We can also assume f is square-free since we can simply cancel out the repeated factors without affecting the vanishing set.

Notice that on any line in $\mathbb{R}^3$, f is a univariate polynomial of degree at most $\sqrt{N}$. Since f vanishes on all points on Y, it must vanish on at least $\sqrt{N}+1$ points on any line $l \in X$. This implies that f actually vanishes on all lines on X. Now by proposition 3.2, each line l *not* in X also intersects at least $\sqrt{N}+1$ lines in X. So f also vanishes on those lines, which in turn implies f vanishes on l. Eventually, f has to vanish on all lines in L.

But now, for any joint in $J(L)$ formed by some lines, f vanish on all these lines. By corollary 2.6, any joint must be a critical point of f. Now, by proposition 3.1, for each line l, there exists at least $K\sqrt{N} > \sqrt{N}$ joints, so each component of ∇f vanish on at least $\sqrt{N}$ points on each line. But again since each component of ∇f has degree at most $\sqrt{N}-1$, it implies that ∇f vanishes on all lines in L, which means f has N critical lines.

However, by lemma 2.4, f should have only at most $\sqrt{N}(\sqrt{N}-1) < N$ critical lines. This is a contradiction.

References

[1] Zeev Dvir, On the size of Kakeya sets in finite fields, arXiv:0803.2336, 2008

[2] Larry Guth and Nets Hawk Katz, Algebraic Methods in Discrete Analogs of the Kakeya Problem, arXiv:0812.1043v1, 2008

[3] Terence Tao, Dvir's proof of the finite field Kakeya conjecture, http://terrytao.wordpress.com/2008/03/24/dvirs-proof-of-the-finite-field-kakeya-conjecture/, 2008

[4] D. Cox, J. Little, and D. O'Shea, Ideals, Varieties, and Algorithms: An Introduction to Computational Algebraic Geometry and Commutative Algebra, Springer Undergraduate Texts in Mathematics, first edition, 1997

[5] Zeev Dvir, Lecture notes on Incidence Theorem and Their Applications, http://www.cs.princeton.edu/~zdvir/teaching/incidence12/8.%20Guth-Katz%20Theorem.pdf, 2012