A proof of Erdős's lower bound for unit distances

We prove the following standard lower bound.

Theorem. For infinitely many n, there is a set of n points in the plane with at least

$$n^{1+\Omega(1/\log\log n)}$$

unit distances. In particular, after changing the implicit constant in the exponent, this gives a lower bound of the form

$$n^{1+1/\log\log n}$$

up to the usual convention that the constant in the numerator is not important. More precisely, the construction gives

$$n^{1+c/\log\log n}$$

for some absolute constant $c > 0$.

A warning about constants: the exact statement with numerator literally equal to 1 depends on the base of the logarithm and on hidden constants. The classical theorem is normally stated as

$$u(n) \geq n^{1+c/\log\log n}$$

for some constant $c > 0$, not with a canonical constant 1.

Step 1: Many integer vectors of the same length

Let

$$m = p_1 p_2 \cdots p_k,$$

where $p_1, \ldots, p_k$ are distinct primes congruent to $1 \pmod 4$.

A classical theorem on sums of two squares says that the number of integer solutions to

$$a^2 + b^2 = m$$

is

$$r_2(m) = 4\prod_{i=1}^{k}(1+1) = 4 \cdot 2^k.$$

Thus there are $4 \cdot 2^k$ integer vectors (a, b) of squared length m.

After scaling the plane by the factor $1/\sqrt{m}$, each such vector becomes a unit vector.

Step 2: Put many lattice points in a square

Let

$$L = \lfloor \sqrt{n} \rfloor.$$

Take the grid

$$P = \{1, 2, \dots, L\} \times \{1, 2, \dots, L\}.$$

Then

$$|P| = L^2 \le n.$$

We will count pairs of grid points whose difference is one of the vectors (a, b) satisfying

$$a^2 + b^2 = m.$$

Choose m small enough compared with L^2, say

$$m \le L^2/100.$$

Then every vector (a, b) with $a^2 + b^2 = m$ has

$$|a|, |b| \le \sqrt{m} \le L/10.$$

For each such vector (a, b), at least

$$(L - |a|)(L - |b|) \ge (9L/10)^2$$

ordered pairs $(x, y), (x + a, y + b)$ remain inside the grid.

Therefore, before scaling, the number of ordered pairs of grid points at squared distance m is at least

$$r_2(m)(9L/10)^2.$$

Dividing by 2 to pass from ordered pairs to unordered pairs, the number of distances equal to $\sqrt{m}$ is at least

$$CL^2 r_2(m)$$

for some absolute constant $C > 0$.

Since $L^2 \asymp n$, after scaling by $1/\sqrt{m}$, these become unit distances, and the number of unit distances is at least

$$C'nr_2(m).$$

So it remains to choose m with $r_2(m) = n^{\Omega(1/\log\log n)}$.

Step 3: Choosing m with many representations

We want $m \le L^2/100 \asymp n$, and we want m to have many prime factors, all congruent to $1 \pmod 4$.

Let $p_1 < p_2 < \cdots < p_k$ be the first k primes congruent to $1 \pmod 4$, and put

$$m = p_1 p_2 \cdots p_k.$$

By the prime number theorem for arithmetic progressions, the i-th prime congruent to $1 \pmod 4$ has size about

$$i \log i.$$

Hence

$$\log m = \sum_{i=1}^{k} \log p_i = O(k \log k).$$

Choose

$$k = c_0 \frac{\log n}{\log\log n}$$

with $c_0 > 0$ small enough. Then

$$\log m = O(k \log k) = O(\log n),$$

and, by taking c_0 sufficiently small, we can ensure

$$m \le n/100.$$

For this choice of m,

$$r_2(m) = 4 \cdot 2^k.$$

Now

$$2^k = \exp(k \log 2) = \exp\left(\Theta\left(\frac{\log n}{\log\log n}\right)\right) = n^{\Theta(1/\log\log n)}.$$

Therefore

$$r_2(m) = n^{\Omega(1/\log\log n)}.$$

Step 4: Final count

From Step 2, the number of unit distances in the scaled grid is at least

$$C' n r_2(m).$$

Using Step 3,

$$C' n r_2(m) \geq n \cdot n^{c/\log\log n} = n^{1+c/\log\log n}$$

for some absolute constant $c > 0$.

Thus, for infinitely many n, there is a set of n points in the plane with at least

$$n^{1+c/\log\log n}$$

unit distances.

This is Erdős's classical lower-bound construction.

Padding to get exactly n points

The construction above gives $L^2 \leq n$ points. If one wants exactly n points, add arbitrary extra points very far away from the grid and from each other, so that they create no unintended unit distances. Adding points cannot decrease the number of unit distances already present.

Therefore, for all sufficiently large n in the corresponding range, we get an n-point set with at least

$$n^{1+c/\log\log n}$$

unit distances, after adjusting constants.

What this proves, and what it does not prove

The robust and standard statement is

$$u(n) \geq n^{1+c/\log\log n}$$

for some absolute constant $c > 0$.

If the exponent is written as

$$1 + \frac{1}{\log \log n},$$

one should interpret that as suppressing constants in the numerator. The construction proves the right qualitative bound, but not a canonical literal constant 1 in the numerator.