Nikolas Mavrogeneiadis
gravitorious
These notes represent only my thoughts about the problem and it doesn't contain rigorous proofs.
March 14, 2022

Closest pair

1. Given $n \geqslant 2$ points in the plane, print the pair of points with smallest Euclidean distance.

Let P be the set of all given points $p_i = (x_i, y_i)$. We can easily make (with $O(nlogn)$) two more sets P_x and P_y that contains all pairs sorted by x_i (x - coordinate) and y_i (y - coordinate) respectively (ascending sorting). Let $Q \subseteq P_x$ be the first half of P_x with the first $\lceil \frac{n}{2} \rceil$ elements and $R \subseteq P_x$ with the next $\lfloor \frac{n}{2} \rfloor$ elements. Let $Q_x = Q$ and $R_x = R$. Let Q_y and R_y be the elements from Q and R respectively sorted by the y-coordinate. We can easily create these sets (with $O(n)$) from P_y in the following way. For each element from P_y, compare the x - coordinate of this element with the x - coordinate from the last point of Q_x (median point of P_x, x^*). If it is lower or equal (therefore, on the first half), then put that element to Q_y. Else, put that element to R_y (it is on the second half). With those created sets, we can easily find recursively the smallest distance from the first (left) half (d_1) and the second (right) half (d_2).
Now, we need to check the pairs of points that are not on the same side. But lets do it cleverly. Let $d = \min\{d_1, d_2\}$. Let $q(q_x, q_y) \in Q$ and $r(r_x, r_y) \in R$ with $d(q, r) < d$.
It is obvious that

$$q_x \leqslant x^* \leqslant r_x \Longrightarrow x^* - q_x \leqslant r_x - q_x \quad (1)$$

and that means

$$(r_x - q_x)^2 < (r_x - q_x)^2 + (r_y - q_y)^2 \Longrightarrow |r_x - q_x| < \sqrt{(r_x - q_x)^2 + (r_y - q_y)^2} = d \quad (2)$$

Therefore we have that

$$(x^* - q_x < d) \wedge (r_x - x_* < d) \quad (3)$$

Now let $S_y \subseteq P_y$ be the set with all points that (3) holds, sorted by y-coordinate (we can easily create that from P_y). If there exists such a pair, then it must be on this set. But for each pair, we do not need to compute the distance with all of the other pairs but only seven pairs above it. So let's see what happens when we check the first point p that $p \in S_y$ (for the other points it is exactly the same). The next figure showing eight boxes with side length $\frac{d}{2}$. The bottom of the boxes is aligned in the y-coordinate of point that we are checking.

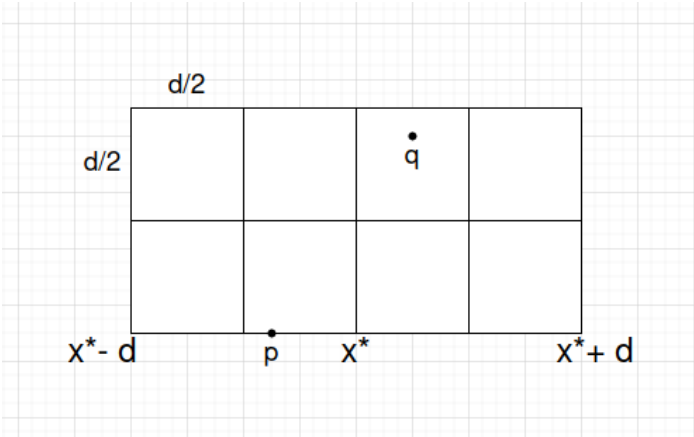

Figure 1: Graphical representation

If there exists a point q that is outside of these boxes, then we are sure that $d(p, q) > d$. We can say that each box have at most one point. Remember that we have found that there is no pair of points on the first or second half that has distance lower than d. Each box is either on the first or second half. Imagine two points that are on the same box. Then the maximum distance that they can have is the length of diagonal that it is equal to $\frac{d}{\sqrt{2}}$. But this contradicts of what we found earlier (that the distance is at least d). This implies that each of the eight boxes has at most one point. So we need to compute the distance with at most seven points above the p.