MA/CSSE 473 Day 13

Brute Force

Divide and Conquer

MA/CSSE 473 Day 13

- **Student Questions**
- Brute force algorithms
- Divide and Conquer

What is the brute force approach to

1. Calculate the n^{th} Fibonacci number?
2. Compute the n^{th} power of an integer?
3. Search for a particular value in a sorted array?
4. Sort an array?
5. Search for a substring of a string?
6. Find the maximum contiguous subsequence in an array of integers?
7. Find the largest element in a Binary Search Tree?
8. Find the two closest points among N points in the plane?
9. Find the convex hull of a set of points in the plane?
10. Find the shortest path from vertex A to vertex B in a weighted graph?
11. Solve the traveling salesman problem?
12. Solve the knapsack problem?
13. Solve the assignment problem?
14. Solve the $n \times n$ non-attacking chess queens problem?
15. Other problems that you can think of?

DIVIDE AND CONQUER

Divide-and-conquer algorithms

- Definition
- Examples seen prior to this course or so far in this course

Q1

Closest Points problem

- Given a collection, S, of N points, find the minimum distance between two points in S.
- Running time for brute force algorithm?

Closest Points divide phase

- Given a collection, S, of N points, find the minimum distance between two points in S.
- For simplicity, we assume $N = 2^k$ for some k.
- Sort the points by x-coordinate.
  - If we use merge sort, the worst case is $\Theta(N \log N)$
- If two points have the same x-coordinate, order them by y-coordinate.
- Let c be the median x-value of the points
- Let S_1 be $\{(x, y): x \leq c\}$, and S_2 be $\{(x, y): x \geq c\}$
  - adjust so we get exactly N/2 points in each subset

Q2

Closest Points problem

- Assume that the points of S are sorted by x-coordinate.
- Let d_1 be the minimum distance between two points in S_1 (the set of "left half" points).
- Let d_2 be the minimum distance between two points in S_2 (the set of "right half" points).
- Let d = min(d_1, d_2). Is d the minimum distance for S?
- What else do we have to consider? **Q3**
- Suppose we needed to compare every point in S_1 to every point in S_2. What would the running time be? **Q4**
- How can we avoid doing so many comparisons?

Quick Review: The Master Theorem

- The Master Theorem for Divide and Conquer recurrence relations:
- Consider the recurrence $T(n) = aT(n/b) + f(n)$, $T(1)=c$, where $f(n) = \Theta(n^k)$ and $k \geq 0$,
- The solution is
  - $\Theta(n^k)$ if $a < b^k$
  - $\Theta(n^k \log n)$ if $a = b^k$
  - $\Theta(n^{\log_b a})$ if $a > b^k$

For details, see Levitin pages 483-485 or Weiss section 7.5.3.

Grimaldi's Theorem 10.1 is a special case of the Master Theorem.

We will use this theorem often. You should review its proof soon (Weiss's proof is a bit easier than Levitin's).

After recursive calls on S_1 and S_2

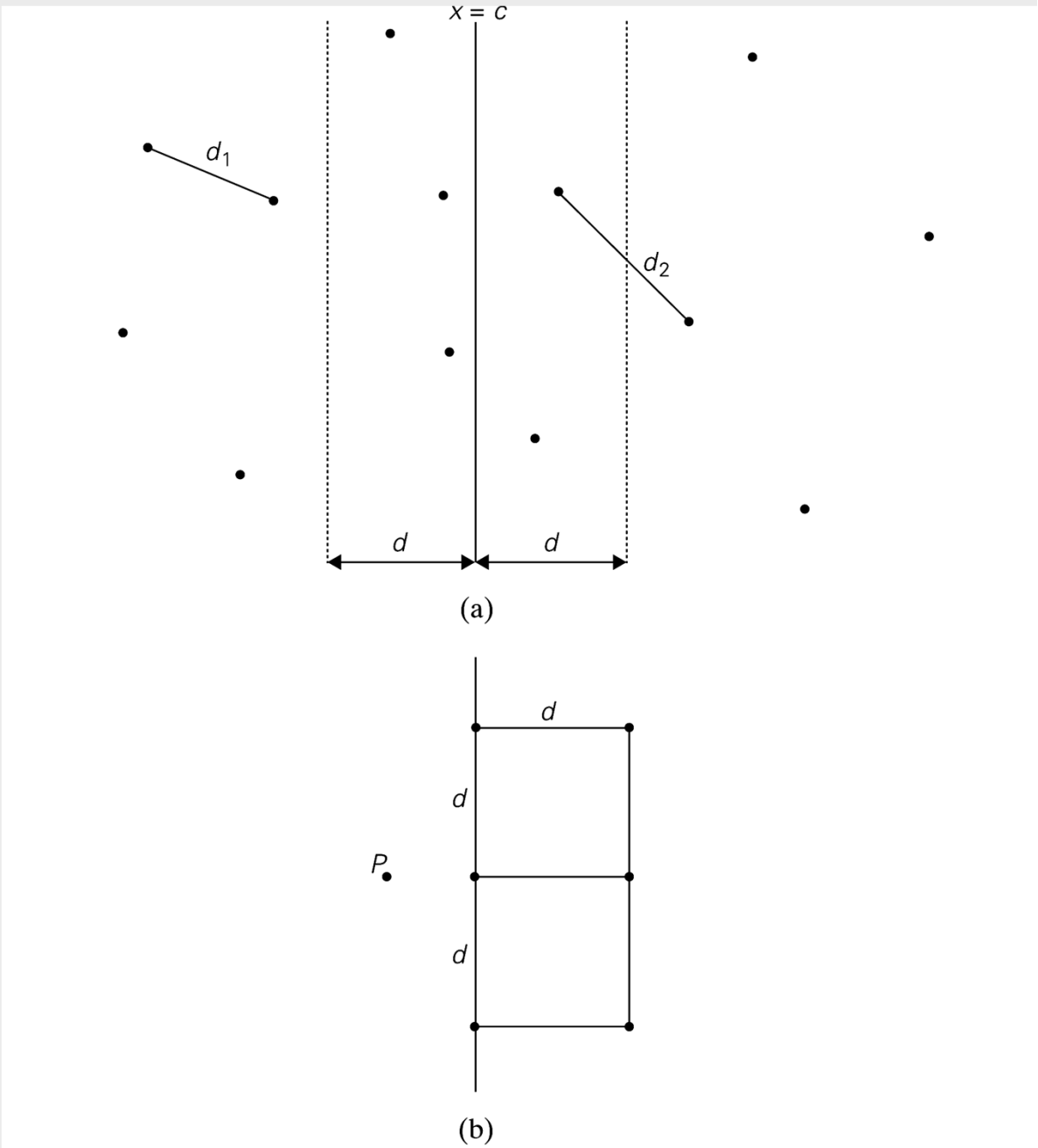

FIGURE 4.7 (a) Idea of the divide-and-conquer algorithm for the closest-pair problem. (b) The six points that may need to be examined for point P.

Convex Hull Problem

- Again, sort by x-coordinate, with tie going to larger y-coordinate.

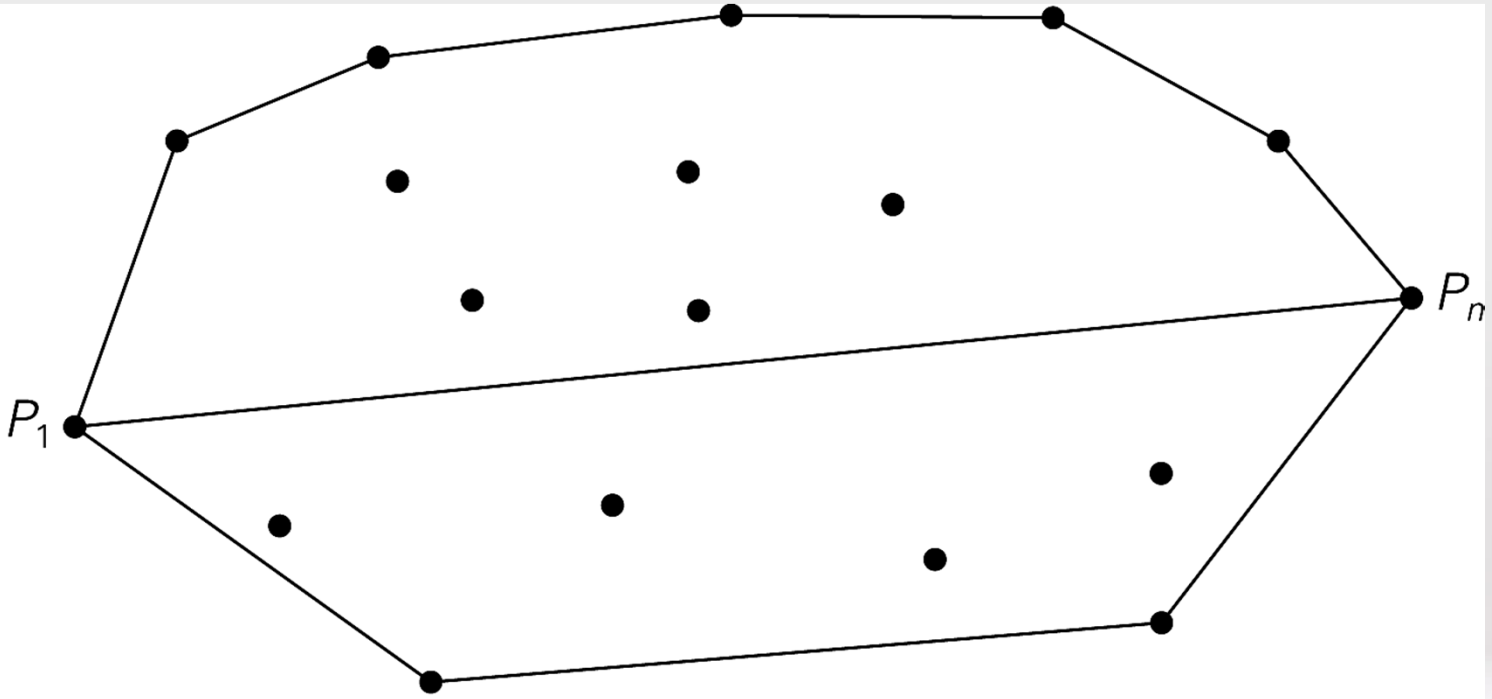

FIGURE 4.8 Upper and lower hulls of a set of points

Recursive calculation of Upper Hull

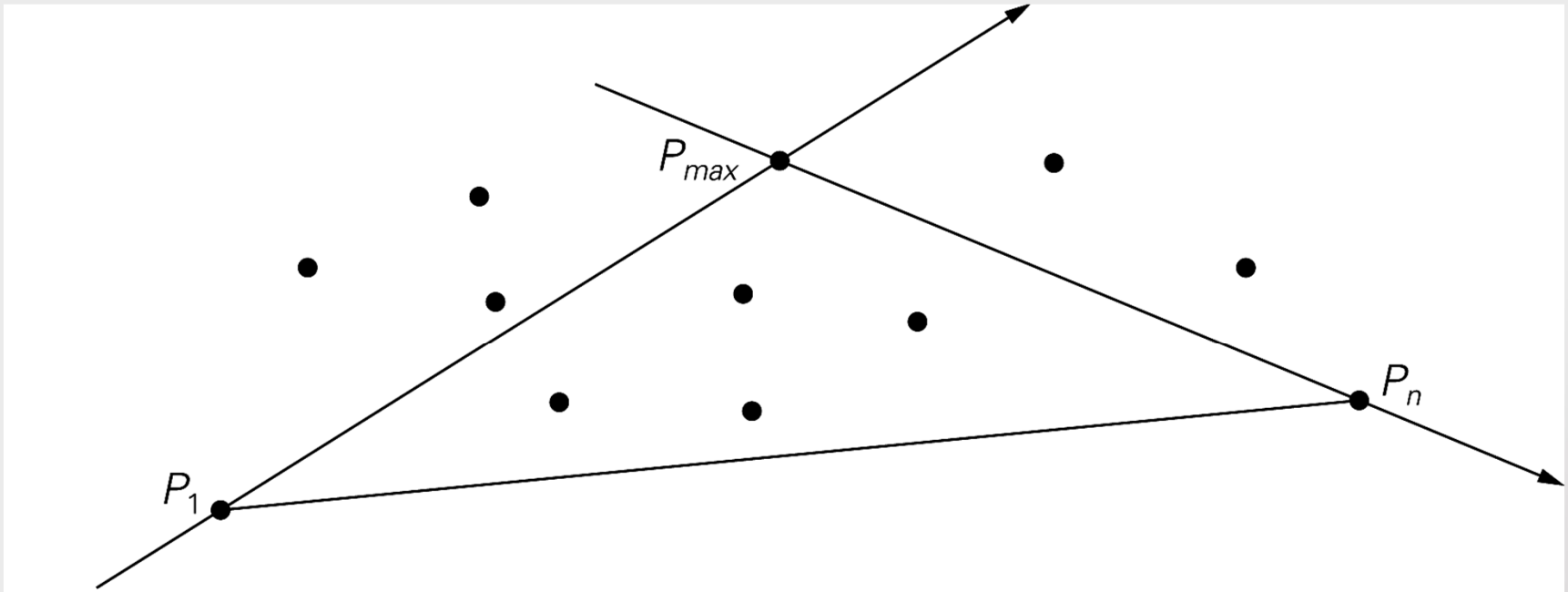

FIGURE 4.9 The idea of quickhull

Simplifying the Calculations

- The maximum distance of P from line P_1P_2, and
- Determining whether P is "to the left" of P_1P_2
  - The area of the triangle through $P_1=(x_1,y_1)$, $P_2=(x_2,y_2)$, and $P_3=(x_3,y_e)$ is ½ of the absolute value of the determinant

$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = x_1y_2 + x_3y_1 + x_2y_3 - x_3y_2 - x_2y_1 - x_1y_3$$

  - The sign of the determinant is positive if the order of the three points is clockwise, and negative if it is counter-clockwise
  - For a proof of this property, see http://mathforum.org/library/drmath/view/55063.html

- Clockwise means that P_3 is "to the left" of directed line segment P_1P_2
- What about max distance?

Efficiency of quickhull algorithm

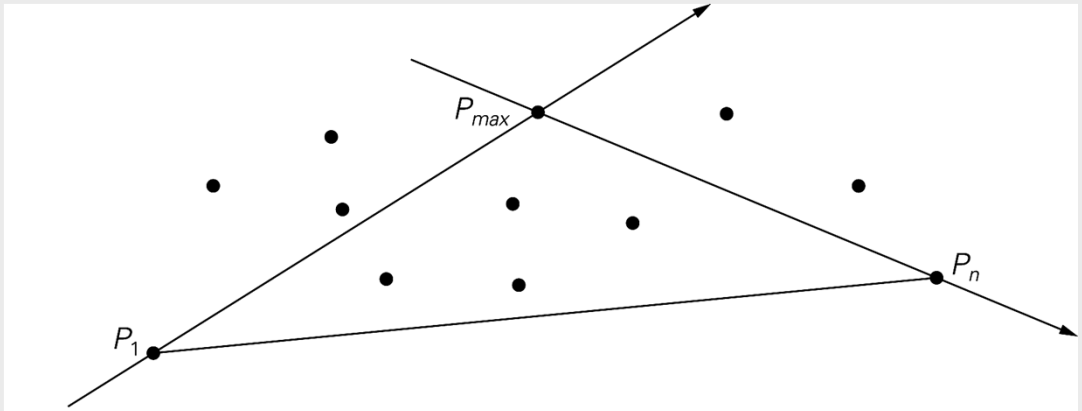

FIGURE 4.9 The idea of quickhull

- What arrangements of points is worst case?
- Average case is much better. Why?