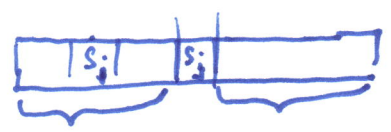


BEWEIS:

Wir betrachten eine zufällige Eingabepermutation, d.h. alle $n!$ Permutationen sind gleich wahrscheinlich, jeweils mit Wahrscheinlichkeit $\frac{1}{n!}$.

- Zwei Elemente s_i und s_j werden
- höchstens einmal verglichen
 - genau dann verglichen, wenn eines das Referenzelement ist.
- (Danach sind beide in verschiedenen Teilarrays!)



Wir betrachten eine Zufallsvariable

$$X_{ij} = \begin{cases} 1 & \text{falls } s_i \text{ und } s_j \text{ verglichen werden} \\ 0 & \text{sonst} \end{cases}$$

Dann ist

$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^n X_{ij}$$

die gesamte Zahl der Vergleiche.

Damit ergibt sich:

$$\begin{aligned}
 E[X] &= E \left[\sum_{i=1}^{n-1} \sum_{j=i+1}^n X_{ij} \right] \\
 \text{„Erwartungswert“} &= \sum_{i=1}^{n-1} \sum_{j=i+1}^n E[X_{ij}] \quad (*) \\
 &= 1 \cdot P(X_{ij}=1) + 0 \cdot P(X_{ij}=0)
 \end{aligned}$$

Wir müssen also die Wahrscheinlichkeit bestimmen, dass s_i und s_j verglichen werden.

Dafür betrachten wir

$$S_{ij} := \{s_i, \dots, s_j\} \quad \boxed{|s_i| \dots |s_j|}$$

Dann ist

$$P(s_i \text{ wird mit } s_j \text{ verglichen}) = P(s_i \text{ oder } s_j \text{ ist erster Pivot in } S_{ij}).$$

Denn wird s_i oder s_j gewählt, wird es mit allen anderen verglichen; wird aber zuerst ein Pivot dazwischen gewählt, werden s_i und s_j getrennt und nicht mehr verglichen!

Da nur einer erster Pivot sein kann, ist

$$\begin{aligned} & P(s_i \text{ oder } s_j \text{ erster Pivot in } S_{ij}) \\ &= P(s_i \text{ erster Pivot in } S_{ij}) \\ &+ P(s_j \text{ erster Pivot in } S_{ij}) \end{aligned}$$

$$= \frac{1}{j-i+1} + \frac{1}{j-i+1} = \frac{2}{j-i+1}, \quad (**)$$

denn die Pivots werden zufällig und unabhängig gewählt, und die Menge S_{ij} hat $j-i+1$ Elemente

Jetzt kombinieren wir (*) und (**):

$$E[X] = \sum_{i=1}^{n-1} \sum_{j=i+1}^n \frac{2}{j-i+1}$$

↓ Setze $k := j-i$

$$= \sum_{i=1}^{n-1} \sum_{k=1}^{n-i} \frac{2}{k+1}$$

$$< \sum_{i=1}^{n-1} \sum_{k=1}^n \frac{2}{k}$$

Jetzt schätzen wir die harmonische Reihe ab:

$$\begin{aligned} \sum_{k=1}^n \frac{1}{k} &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \dots + \frac{1}{n} \\ &\leq \underbrace{1}_1 + \underbrace{\frac{1}{2} + \frac{1}{2}}_1 + \underbrace{\frac{1}{4} + \frac{1}{4} + \frac{1}{4}}_1 + \dots + \frac{1}{2^{\log_2 n}} \end{aligned}$$

Das sind $1 + \log_2 n$ Einsen, also

$$\sum_{k=1}^n \frac{1}{k} \leq 1 + \log_2 n, \text{ also}$$

$$\sum_{k=1}^n \frac{2}{k} \in \Theta(\log n)$$

und damit

$$E[X] \in O(n \log n).$$

