# เรียงลำดับเทปปริศนา

## 1 บทนำ ข้อตกลง และข้อสังเกตเบื้องต้น

**ข้อตกลง 1.1.** เพื่อให้เข้าใจตรงกัน เราจะพูดคุยถึงปัญหาที่ลดรูปมาแล้ว นั่นคือให้ลำดับ $A$ ของสิ่งของ $n$ ชิ้นมา ที่เรียงแล้ว และให้ลำดับ $B$ ของสิ่งของ $n$ ชิ้นมา ที่เรียงแล้วเช่นกัน เราต้องการเรียงลำดับ $A$ กับ $B$ รวมกัน โดยแท้จริงแล้วสมาชิกแต่ละตัวของ $A$ และ $B$ อยู่ในตำแหน่งบางตำแหน่งที่ไม่ได้เรียงอยู่ (สมมติว่า $A[i]$ อยู่ตำแหน่ง $p_i$ และ $B[i]$ อยู่ตำแหน่ง $q_i$ โดยทั้ง $p_0, p_1, \ldots, p_{n-1}$ และ $q_0, q_1, \ldots, q_{n-1}$ เป็นการเรียงสับเปลี่ยนของ $\{0, 1, \ldots, n-1\}$)

### 1.1 Naive Merge

พิจารณาอัลกอริทึมดังต่อไปนี้

**Algorithm 1** Naive Merge

**Require:** $n \geq 0, A$ is sorted, $B$ is sorted
**Ensure:** The content of $C$ is the content of concatenation of $A$ and $B$, and $C$ is sorted

```
C ← empty list
i ← 0
j ← 0
while i < n and j < n do
    if A[i] ⪯ B[j] then
        C ← C append A[i]
        i ← i + 1
    else
        C ← C append B[j]
        j ← j + 1
    end if
end while
while i < n do
    C ← C append A[i]
    i ← i + 1
end while
while j < n do
    C ← C append B[j]
    j ← j + 1
end while
```

สังเกตว่าอัลกอริทึมนี้จะทำการเทียบ $A[i] \preceq B[j]$ ไม่เกิน $2n$ ครั้ง แต่ละครั้งจะทำการเคลื่อนตัวชี้จากตำแหน่ง $p_{i-1}$ ไปยัง $p_i$ และ $q_{j-1}$ ไปยัง $q_j$ (ซึ่งหมายความว่าเลื่อนไม่เกิน $2n$ ตำแหน่ง) จึงใช้ค่าใช้จ่ายรวมไม่เกิน $4n^2$ จะทำให้ได้คะแนนประมาณ 20 คะแนน

คำถามคือ จะมีวิธีที่ดีกว่านี้มั้ย? ก่อนจะเข้าสู่ส่วนถัดไป จะขอกล่าวถึงทฤษฎีบทจากแคลคูลัสที่มีประโยชน์ ดังต่อไปนี้

**ทฤษฎีบท 1.2.** ให้ $f\colon (a,b) \to \mathbb{R}$ เป็นฟังก์ชัน ถ้า $f$ มี local minimum ที่ $x_0 \in (a,b)$ และ $f$ หาอนุพันธ์ได้ที่ $x_0$ จะได้ว่า $f'(x_0) = 0$

สำหรับบทพิสูจน์ จะต้องอาศัยความรู้พื้นฐานทาง analysis จึงขอละไว้ (แต่จะมีบทพิสูจน์โดยคร่าวในภาคผนวก)

สำหรับส่วนถัดไป จะเป็นรายการคำใบ้ทั้งหมดที่ค่อย ๆ นำไปสู่เฉลย

# 2 คำใบ้

## 2.1 การแบ่งส่วน

หากเราหั่น $A$ ออกเป็นสองส่วน คือส่วนซ้ายกับส่วนขวา จะเกิดอะไรขึ้น? จะทำอะไรกับคำถามได้มั้ย? และถ้าหั่นออกเป็น $k > 2$ ส่วนล่ะ? จะทำอะไรได้บ้าง?

## 2.2 การถามคำถามลักษณะ $(A[i] \preceq B[j])$?

หากเราถามหนึ่งคำถาม เราก็จะต้องเลื่อนตัวชี้แรกไปอยู่ที่ตำแหน่ง $p_i$ และเลื่อนตัวชี้ที่สองไปอยู่ที่ตำแหน่ง $q_j$ ทำให้ต้องเลื่อน $\Omega(n)$ ตำแหน่งในกรณีเลวร้าย (และกรณีเฉลี่ย)
แต่หากเรามีคำถามหลายคำถามพร้อมกันล่ะ กล่าวคือ อยากรู้ว่า $A[i_0] \preceq B[j_0]$ มั้ย, $A[i_1] \preceq B[j_1]$ มั้ย, $\ldots$ ไปจนถึง $A[i_{Q-1}] \preceq B[j_{Q-1}]$ มั้ย จะมีวิธีที่เสียค่าใช้จ่ายน้อยกว่า $\sim 2Qn$ มั้ย?

## 2.3 การเรียงลำดับของคำถาม

หาก $p_0 \leq p_1 \leq p_2 \leq \cdots \leq p_{Q-1}$ และ $q_0 \leq q_1 \leq q_2 \leq \cdots \leq q_{Q-1}$ เราก็ไม่จำเป็นต้องทำอะไรเลย การเลื่อนจาก $p_{i-1}$ ไป $p_i$ และ $q_{i-1}$ ไป $q_i$ ทำได้โดยง่าย และค่าใช้จ่ายรวมคือ

$$\sum_{i=1}^{Q-1} |p_i - p_{i-1}| + |q_i - q_{i-1}| = p_{Q-1} - p_0 + q_{Q-1} - q_0 \leq 2n$$

คำถามคือถ้ามันไม่ใช่กรณีนี้ล่ะ เราเริ่มต้นจากการทดลอง เรียง $((p_i, q_i))_i$ ตามลำดับ $p_i$ จากน้อยไปมาก จะได้ว่า $p_i$ เรียงแล้ว แต่ $q_i$ จะเป็นยังไงก็ได้ เท่านี้ก็จะลด bound จาก $2Qn$ เหลือ $Qn$ ได้แล้ว ต่อมาจะลดอย่างไรได้อีก?

## 2.4 ท่าแปลก หรือ ธรรมชาติของ $\mathbb{R}^Q$

การจัดการเรียงลำดับคำถามเพื่อลดค่าใช้จ่ายนั้น ทำได้หลายแบบ วิธีหนึ่งคือใช้ท่า ที่เรียกว่า "Mo's Algorithm"[a] นอกจากนี้ยังมีวิธีที่ดีกว่านี้อีก ที่อาศัยบทพิสูจน์จาก "Dilworth's Theorem" เพื่อสังเกตธรรมชาติบางอย่างของลำดับของคู่อันดับของจำนวนจริง

[a] ขอแนะนำผู้สนใจให้หาอ่านเพิ่มเติมเอาด้วยตนเอง อย่างไรก็ตาม จะมีคำอธิบายย่อในส่วนท้าย

# 3 แนวคิดหลัก

แนวคิดหลักจะประกอบด้วยสองส่วน คือการแบ่งส่วนโดเมนที่จะถามออกเป็นหลาย ๆ ส่วน กับการจัดเรียงคำถามที่ถามพร้อมกันได้ โดยจะกล่าวในส่วนนี้ (ส่วนที่ 3) และส่วนถัดไป (ส่วนที่ 4)

## 3.1 วิธีที่แย่กว่า: $n^2$ คำถาม

หากเราไล่ถาม $A[i] \preceq B[j]$ สำหรับทุก $(i, j)$ ที่เป็นไปได้ ย่อมสามารถให้คำตอบเราได้ (แต่จะเสียค่าใช้จ่ายเพิ่มมากขึ้น) อย่างไรก็ตามหากถามมั่วจะเสียค่าใช้จ่าย $\approx 4n^3$ แต่หากไล่เรียงคำถามดี ๆ จะเสียเพียง $n^2$ พอดี (ไล่ $i = 0, j = 0, 1, 2, \ldots, n-1$ แล้วต่อมาย้อนจากหลังมาหน้า เป็น $i = 1, j = n-1, n-2, \ldots, 0$ แล้วสลับไปมาจนครบ)

ความรู้ที่ว่า "การถาม $n^2$ ครั้งนั้นเพียงพอที่จะให้คำตอบได้" จะมีประโยชน์ในกระบวนการต่อไป

## 3.2 2-decomposition

สมมติแบ่ง $A$ ออกเป็นส่วนซ้าย $A_L$ ค่ากลาง $A[m]$ และส่วนขวา $A_R$ สิ่งที่เราสามารถสังเกตได้คือ หากเรารู้ว่าค่าใน $B$ ที่ใกล้เคียงกับ $A[m]$ อยู่ที่ $B[m']$ และแบ่ง $B$ ออกเป็น $B_L$ และ $B_R$ เราก็จะรู้ด้วยว่า ไม่มีประโยชน์ใด ๆ ที่จะเทียบค่าใน $A_L$ กับ $B_R$ และไม่มีประโยชน์ใด ๆ ที่จะเทียบค่า $A_R$ กับ $B_L$ (เพราะ $A_L \preceq B_R$ ทุกตัวอยู่แล้ว และ $B_L \preceq A_R$ ทุกตัวอยู่แล้ว)

เท่ากับว่าเราสามารถถาม $n$ คำถามในตอนแรก (เพื่อหาว่า $m'$ คือค่าไหน) แล้วถามอีก $|A_L||B_L| + |A_R||B_R|$ คำถาม (ไล่ถามทุกคู่ใน $(A_L, B_L)$ กับทุกคู่ใน $(A_R, B_R)$) สังเกตว่า

$$|A_L||B_L| + |A_R||B_R| \leq \frac{n}{2}|B_L| + \frac{n}{2}|B_R| = \frac{n}{2}(|B_L| + |B_R|) = \frac{n}{2}n = \frac{n^2}{2}$$

จึงใช้ค่าใช้จ่าย $\frac{n^2}{2}$

## 3.3 $k$-decomposition

หากคิดในลักษณะเดียวกัน แต่เปลี่ยนจาก $k = 2$ ส่วน เป็น $k > 2$ ส่วน จะเกิดอะไรขึ้น ในตอนแรกเราจะต้องหาว่า $A[m_1], \ldots, A[m_{k-1}]$ จะไปเชื่อมต่อกับ $B[m'_1], \ldots, B[m'_{k-1}]$ ตรงไหนบ้าง ซึ่งเราสามารถใช้ $(k-1)n$ คำถามในส่วนนี้ได้ นอกจากนี้ ที่สำคัญคือเมื่อเราหยุดอยู่ที่ $A[m_1]$ เราสามารถไล่หาทุก $B$ โดยเคลื่อนเพียงหนึ่งตำแหน่งต่อคำถาม และเมื่อย้ายไป $A[m_2]$ ก็ย้ายตัวชี้ฝั่ง $B$ ย้อนกลับมา สังเกตว่าฝั่งซ้ายเลื่อนรวมกันไม่เกิน $n-1$ ตำแหน่ง และฝั่งขวาเลื่อนรวมกัน $(k-1)(n-1)$ ตำแหน่งพอดี จึงเสียค่าใช้จ่ายรวมไม่เกิน $k(n-1)$ (เพื่อความง่ายจะกล่าวว่า ไม่เกิน $kn$)

ถัดมาเราจะแบ่ง $A$ และ $B$ ได้ แบ่งออกเป็นอย่างละ $k$ ส่วน โดยขนาดของแต่ละส่วนของ $A$ จะประมาณ $\frac{N}{k}$ ทั้งหมด แต่

ขนาดของ $B$ อาจเป็นอะไรก็ได้ หลังจากนั้นเราทำการถาม $|A_1||B_1| + |A_2||B_2| + \cdots + |A_k||B_k|$ ครั้งเพื่อให้ได้คำตอบ

สังเกตได้ว่า

$$\sum_{i=1}^{k} |A_i||B_i| \leq \sum_{i=1}^{k} \frac{n}{k}|B_i| = \frac{n}{k}\sum_{i=1}^{k} |B_i| = \frac{n}{k}n = \frac{n^2}{k}$$

จึงต้องถามทั้งหมดอีก $Q = \frac{n^2}{k}$ ครั้ง

หากทำการถามแบบปกติทั่วไป (ไม่เรียงหรือแปลงชุดคำถามจากทั้ง $Q$ ครั้ง) จะได้ว่าเสียค่าใช้จ่ายทั้งหมดไม่เกิน $kn + Qn = kn + \frac{n^3}{k}$ ครั้ง ซึ่งดูจะไม่เป็นประโยชน์อะไรนัก แต่หากการถามคำถาม $Q$ คำถาม สามารถใช้ค่าใช้จ่ายน้อยกว่า $\Omega(Qn)$ ก็จะน่าสนใจ

สำหรับส่วนนี้ เราจะสมมติว่าเราสามารถถามคำถาม $Q$ คำถามโดยใช้ค่าใช้จ่ายไม่เกิน $c_n(Q)$ เมื่อ $c_n(Q) = \lambda n^a Q^b$ สำหรับบาง $a, b \in \mathbb{R}_+$

จะได้ว่าค่าใช้จ่ายรวมมีค่าไม่เกิน

$$\begin{aligned} nk + c_n(Q) &= nk + \lambda n^a \left(\frac{n^2}{k}\right)^b \\ &= nk + \lambda n^{a+2b} k^{-b} \end{aligned}$$

โดยเพื่อความง่ายเราจะให้ $f(k) := nk + \lambda n^{a+2b} k^{-b}$ แทนขอบเขตของค่าใช้จ่ายเมื่อให้ค่า $k \in \mathbb{N}_{>0}$ เป็นตัวแปร สังเกตว่าเราสามารถทำการขยาย $f$ ไปในโดเมนของ $\mathbb{R}_+$ ได้ นิยามโดย

$$f(k) = nk + \lambda n^{a+2b} k^{-b}$$

เหมือนเดิม

จากทฤษฎีบท 1.2 สมมติว่ามีค่า $\hat{k}$ ที่ $f(\hat{k})$ มีค่าน้อยสุด เนื่องจาก $f \in \mathcal{C}^2(0, +\infty)$ อย่างชัดเจน จึงได้ว่า $f'(\hat{k}) = 0$ แต่เราทราบว่า

$$f'(k) = n - b\lambda \frac{n^{a+2b}}{k^{b+1}}$$

สมมติให้ $f'(\hat{k}) = 0$ จะได้

$$\begin{aligned} 0 &= n - b\lambda \frac{n^{a+2b}}{\hat{k}^{b+1}} \\ b\lambda n^{a+2b-1} &= \hat{k}^{b+1} \\ b^{\frac{1}{b+1}} n^{\frac{a+2b-1}{b+1}} &= \hat{k} \end{aligned}$$

จะได้ $\hat{k} = b^{\frac{1}{b+1}} n^{\frac{a+2b-1}{b+1}}$ เนื่องจาก $f'(\hat{k}) = 0$ สังเกตว่าหาก $\tilde{k} > \hat{k}$ แล้ว $f'(\tilde{k}) = n - b\lambda \frac{n^{a+2b}}{\tilde{k}^{b+1}} > n - b\lambda \frac{n^{a+2b}}{\hat{k}^{b+1}} = 0$ ในทำนองเดียวกัน สังเกตว่าหาก $0 < \tilde{k} < \hat{k}$ แล้ว $f'(\tilde{k}) = n - b\lambda \frac{n^{a+2b}}{\tilde{k}^{b+1}} < n - b\lambda \frac{n^{a+2b}}{\hat{k}^{b+1}} = 0$ แสดงว่า $f'(\tilde{k}) < 0$ สำหรับทุก $\tilde{k} < \hat{k}$ (นั่นคือ $f$ เป็นฟังก์ชันไม่เพิ่มในระยะ $(0, \hat{k})$) และ $f'(\tilde{k}) > 0$ สำหรับทุก $\tilde{k} > \hat{k}$ (นั่นคือ $f$ เป็นฟังก์ชันไม่ลดในระยะ $(\hat{k}, +\infty)$) จึงมั่นใจได้ว่า $\hat{k}$ เป็น global minimum ของ $f$ บน $(0, +\infty)$

เราจึงได้ข้อสรุปค่าใช้จ่ายด้วยวิธีนี้เป็น

$$\begin{aligned} f(\hat{k}) &= n\hat{k} + \lambda n^{a+2b} \hat{k}^{-b} \\ &= b^{\frac{1}{b+1}} n^{\frac{a+2b-1}{b+1}+1} + \lambda n^{a+2b} (b^{\frac{1}{b+1}} n^{\frac{a+2b-1}{b+1}})^{-b} \end{aligned}$$

**หมายเหตุ.** การพยายามสังเกตอะไรจากพจน์ในลักษณะนี้ ในขณะนี้ เป็นสิ่งที่ค่อนข้างยาก จึงจะขอละนิพจน์เอาไว้ดังนี้ก่อน หากสนใจสามารถทดลองแทนค่า $a = 1, b = 1$ (แบบปกติ) หรือ $a = 0.5, b = 1$ กับ $a = 1, b = 0.5$ ได้ เพื่อสังเกตผลลัพธ์โดยคร่าว

**หมายเหตุ.** $f(\hat{k})$ ยังเอาไปใช้ไม่ได้ทันทีเพราะเราบอกไม่ได้ว่า $\hat{k} \in \mathbb{N}_{>0}$ แต่เรารู้ว่าสำหรับฟังก์ชัน $f$ บนจำนวนจริง $\hat{k}$ ให้ค่าน้อยสุดแล้ว การใช้จริงจึงสามารถอาศัยค่า $\lfloor \hat{k} \rfloor$ กับ $\lceil \hat{k} \rceil$ ไปใช้แทนได้

## 3.4 Divide and Conquer

หากทำการลดรูปปัญหา เป็นการแก้ปัญหาเมื่อให้ $A$ มาขนาด $n_A$ และให้ $B$ มาขนาด $n_B$ เราสามารถทำการแบ่งครึ่งในครั้งแรกได้ เลื่อนตัวชี้บน $A$ ไปเสียค่าใช้จ่าย $n_A$ ก่อน ต่อมาถาม $n_B$ ครั้งไล่ตัวชี้บน $n_B$ ได้ (แต่ในชั้นที่ไม่ใช่ชั้นแรกเราไม่สามารถมั่นใจได้ว่าจะเสียค่าใช้จ่าย $1$ ครั้งต่อการไล่หรือไม่ จึงให้ใช้ค่าใช้จ่าย $c_n(n_B)$ ไปก่อน) หลังจากนั้นจะรู้ว่าจะต้องแบ่ง $B$ ออกเป็น $B_L$ กับ $B_R$ ยังไง แล้วก็ลงไปถามต่อได้ว่า $A_L$ กับ $B_L$ จะรวมออกมาเป็นอย่างไร และ $A_R$ กับ $B_R$ จะรวมออกมาเป็นอย่างไร

สามารถคำนวณค่าใช้จ่ายได้โดย ให้ $C(n_A, n_B)$ แทนค่าใช้จ่ายสำหรับ $A$ ขนาด $n_A$ และ $B$ ขนาด $n_B$ จะได้

$$C(n_A, n_B) \leq n_A + c_n(n_B) + C\left(\left\lceil \frac{n_A - 1}{2} \right\rceil, n_{B_L}\right) + C\left(\left\lceil \frac{n_A - 1}{2} \right\rceil, n_{B_R}\right)$$

สังเกตได้ว่าในต้นไม้การเรียกซ้ำ ในแต่ละชั้นจะมีผลรวมของ $n_B$ รวมกันเท่ากับขนาดในตอนแรก นั่นคือ $n$ นอกจากนี้ผลรวมของ $n_A$ ก็มีค่าเท่ากับ $n$ ด้วยในแต่ละชั้น จึงได้ว่าแต่ละชั้นจะเสียค่าใช้จ่าย $\sum n_A + c_n(n_B)$ แต่เราสังเกตว่า

คำถามแต่ละชุด $n_B$ ในแต่ละรอบนั้นเป็นอิสระต่อกัน เราจึงสามารถนำมารวมกันก่อนแล้วค่อยถามทีเดียว จะเสียค่าใช้จ่าย $(\sum n_A) + c_n(\sum n_B)$ แทน ซึ่งมีค่าเท่ากับ $n + c_n(n)$ ต่อชั้น หากคุมให้แต่ละชั้น $n_A$ แบ่งครึ่งพอดี จะได้จำนวนชั้นไม่เกิน $\lceil \log_2 n \rceil + 1$ จึงได้ค่าใช้จ่ายรวมไม่เกิน

$$(2 + \log_2 n)(n + c_n(n))$$

จากสมมติฐานที่ว่า $c_n(Q) = \lambda n^a Q^b$ เราสามารถแทนค่าลงไป ได้ค่าใช้จ่ายรวมไม่เกิน

$$(2 + \log_2 n)(n + \lambda n^{a+b})$$

# 4 ค่าใช้จ่ายสำหรับการถาม $Q$ คำถามที่เป็นอิสระต่อกันพร้อมกัน

## 4.1 วิธีพื้นฐาน

จากส่วนที่แล้ว เราสมมติให้ $c_n(Q) = \lambda n^a Q^b$ เป็นขอบเขตบนของค่าใช้จ่ายสำหรับการถาม $Q$ คำถามที่เป็นอิสระต่อกันพร้อมกันบน $A, B$ ขนาดอย่างละ $n$

จากวิธีการปกติทั่วไป เราสามารถถามได้โดยเสียค่าใช้จ่ายไม่เกิน $Qn$ จึงได้ $c_n(Q) = nQ$ นั่นคือ $\lambda = 1, a = 1, b = 1$ โดยง่าย คำถามคือจะมีวิธีที่ดีกว่านี้มั้ย?

## 4.2 Mo's Algorithm

วิธีการยอดนิยมสำหรับการถามคำถามประเภทที่ต้องอาศัยตัวเลื่อน และสามารถตอบแบบ offline (รับคำถามมาให้หมดก่อนแล้วค่อยตอบทีเดียว) ได้ คือ Mo's Algorithm โดยจะมีวิธีการดำเนินการดังนี้

สังเกตว่าการขยับทั้งหมดในแต่ละบล็อก $Q_i$ จะเสียค่าใช้จ่ายของการขยับ $l$ ไม่เกิน $|Q_i|k$ และค่าใช้จ่ายของการขยับ $r$ ไม่เกิน $N$ (เพราะเรียงจากน้อยไปมากแล้ว ก็ขยับจากซ้ายไปขวาได้เลย) จึงใช้ค่าใช้จ่ายรวมเท่ากับ

$$\sum_{i=0}^{\lceil \frac{N}{k} \rceil} |Q_i|k + N = k|Q| + \frac{N^2}{k}$$

เมื่อแทนค่า $k = \lfloor \frac{N}{\sqrt{|Q|}} \rfloor$ จะได้ค่าใช้จ่ายไม่เกิน $3N\sqrt{|Q|}$ จึงได้ $\lambda = 3, a = 1, b = 0.5$ สำหรับวิธีการนี้ (จริง ๆ แล้ว $\lambda$ เกิน 2 ไปแค่นิดเดียว ในทางปฏิบัติสามารถใช้ $\lambda = 2$ ได้ แต่เนื่องจากการวิเคราะห์ส่วนนี้จะทำให้ซับซ้อนมากเกินไป จึงขอใช้ $\lambda = 3$ ไปเลย)

**Algorithm 2** Mo's Algorithm

รับ $Q$ ซึ่งเป็นรายการคำถามในรูปแบบ $(L, R)$ เมื่อ $0 \leq L \leq R < N$
$k \leftarrow \lfloor \frac{N}{\sqrt{|Q|}} \rfloor$
แบ่งส่วน $Q$ ออกเป็น $Q_0, Q_1, \ldots, Q_{\lceil \frac{N}{k} \rceil}$ โดยที่ $Q_i$ จะมีทุกคำถาม $(L, R)$ ใน $Q$ ที่ $ik \leq L < (i+1)k$
$i \leftarrow 0$
$l \leftarrow 0$
$r \leftarrow 0$
**while** $i \leq \lceil \frac{n}{k} \rceil$ **do**
    เรียง $Q_i$ โดยพิกัด $R$ จากน้อยไปมาก
    $j \leftarrow 0$
    **while** $j < |Q_i|$ **do**
        $(L, R) = Q_i[j]$
        ขยับ $l$ ไปยัง $L$ ▷ เสียค่าใช้จ่าย $|l - L - 1|$
        ขยับ $r$ ไปยัง $R$ ▷ เสียค่าใช้จ่าย $|r - R - 1|$
        $j \leftarrow j + 1$
    **end while**
    $i \leftarrow i + 1$
**end while**

### 4.3 Iteration บน Erdős–Szekeres theorem

สำหรับส่วนนี้ จะอาศัยผลลัพธ์จาก Dilworth's theorem ดังนี้

**ทฤษฎีบท 4.1.** ให้ลำดับของจำนวนจริงใด ๆ ความยาว $n$ จะได้ว่ามีลำดับย่อยไม่เพิ่ม (nonincreasing subsequence) ความยาวอย่างน้อย $\lfloor \sqrt{n} \rfloor$ อยู่ หรือ มีลำดับย่อยไม่ลด (nondecreasing subsequence) ความยาวอย่างน้อย $\lfloor \sqrt{n} \rfloor$ อยู่

ทฤษฎีบทดังกล่าว คล้ายคลึงกับ Erdős–Szekeres theorem มาก และสามารถพิสูจน์ด้วย Erdős–Szekeres theorem เช่นกัน โดย Erdős–Szekeres theorem เป็นดังต่อไปนี้

**ทฤษฎีบท 4.2** (Erdős–Szekeres, 1935)**.** ให้ $r, s \in \mathbb{N}_{>0}$ สำหรับลำดับของจำนวนจริงที่ไม่มีจำนวนซ้ำกัน ที่มีความยาวอย่างน้อย $(r-1)(s-1)+1$ จะได้ว่ามีลำดับย่อยขนาด $r$ ที่เป็นลำดับเพิ่ม หรือ มีลำดับย่อยขนาด $s$ ที่เป็นลำดับลด

ก่อนอื่นสมมติว่าทฤษฎีบททั้งคู่เป็นจริง เราจะได้ว่าสำหรับคำถาม $Q$ คำถาม (ในรูปแบบ $(L_i, R_i)$ ที่เป็นอิสระต่อกัน เราสามารถเริ่มจากการเรียงตาม $L_i$ ก่อนเลย แล้วพิจารณาว่าจะสามารถแบ่งเป็นกลุ่ม ๆ ที่ $R_i$ เรียงจากน้อยไปมากหรือมากไปน้อย ได้หรือไม่

เราจึงสร้างอัลกอริทึมดังนี้

จะได้ว่าเราสามารถแบ่งคำถามใน $Q$ ออกเป็นกลุ่ม $Q_1, Q_2, \ldots, Q_k$ โดยที่แต่ละกลุ่มจะมีสมบัติว่า เมื่อนำ $L_i$ มาเรียงจากน้อยไปมากแล้ว $R_i$ จะเรียงจากน้อยไปมากหรือไม่ก็มากไปน้อย ทำให้เวลาถามคำถามในกลุ่มหนึ่ง สามารถเลื่อนตัวชี้ทางซ้ายไปที่น้อยสุด แล้วค่อย ๆ เลื่อนขึ้น และเลื่อนตัวขวาไปที่น้อย/มากสุด แล้วค่อย ๆ เลื่อน ขึ้น/ลง (ขึ้นอยู่กับ

**Algorithm 3** Erdős–Szekeres Iteration

รับ $Q$ ซึ่งเป็นรายการคำถามในรูปแบบ $(L, R)$ เมื่อ $0 \leq L \leq R < N$
เรียง $Q$ จากน้อยไปมาก
ต่อมาจะแบ่ง $Q$ ออกเป็นกลุ่ม ๆ เป็นกลุ่ม $Q_1, Q_2, \ldots$
$i \leftarrow 1$
**while** $Q$ ยังไม่ว่าง **do**
    หา longest nonincreasing subsequence หรือ longest nondecreasing subsequence บน $R_i$ จาก $Q$ แล้วหยิบอันที่ยาวที่สุด เรียกสิ่งนี้ว่า $Q_i$
    เอา $Q_i$ ออกจาก $Q$
    $i \leftarrow i + 1$
**end while**

ทิศของการเรียง) จึงสามารถตอบ $|Q_i|$ คำถามด้วยการเลื่อนตัวชี้ ค่าใช้จ่ายไม่เกิน $3N$ (หากไม่ได้รับประกันอะไรเกี่ยวกับจุดเริ่มต้น อาจจะต้องใช้ $4N$ แต่สำหรับอัลกอริทึม 3 ทุกครั้งที่เปลี่ยนกลุ่ม ปลายตัวชี้จะอยู่ที่สุดขอบแล้ว จึงจำเป็นจะต้องเลื่อนเพียงตัวชี้ฝั่ง $B$ จะต้องใช้ $3N$ แต่หากเรียงลำดับกลุ่มให้ทิศสลับกันไปมา จะดีขึ้นไปอีก ทำให้เหลือ $2N$ ได้หากกลุ่มที่หันขึ้นกับกลุ่มที่หันลงมีจำนวนเท่ากัน)

จึงได้ว่าเสียค่าใช้จ่ายรวม $3Nk$ ได้ คำถามคือ $k$ มีค่าเท่าไหร่? สังเกตว่าทุก ๆ ครั้งเราสามารถลดของจาก $Q$ ไปด้วยขนาดอย่างน้อย $\lfloor\sqrt{|Q|}\rfloor$ จึงได้

$$T(n) = 1 + T(n - \lfloor\sqrt{n}\rfloor)$$

เมื่อ $T(1) = 1$

**บทตั้ง 4.3.** สำหรับทุก $k \in \mathbb{N}_{>1}$ สำหรับทุกจำนวนเต็ม $1 \leq i \leq 2k+1$ จะได้ว่า

$$T(k^2 + i) = \begin{cases} T(k^2) + 1 & \text{ถ้า } 1 \leq i \leq k \\ T(k^2) + 2 & \text{ถ้า } k+1 \leq i \leq 2k+1 \end{cases}$$

*พิสูจน์.* จะพิสูจน์ด้วยอุปนัยเชิงคณิตศาสตร์ ก่อนอื่นพิจารณาขั้นฐานคือ $k = 1$ จะได้ $T(k) = 1$ ต่อมาจำนวนเต็ม $i$ ระหว่าง $1$ ถึง $2k+1$ จะมีกรณี $i = 1, 2, 3$ พิจารณา $T(k^2+1) = T(2) = 1 + T(2-1) = 2$, $T(k^2+2) = T(3) = 1 + T(3-1) = 3$ และ $T(k^2+3) = T(4) = 1 + T(4-2) = 3$ ซึ่งสอดคล้องกับข้อความที่ต้องการจะพิสูจน์

ถัดมาสมมติให้ $k > 1$ เป็นจำนวนเต็มใด ๆ และให้ $1 \leq i \leq 2k+1$ สำหรับกรณี $1 \leq i \leq k$ จะได้ว่า

$$\begin{aligned} T(k^2 + i) &= 1 + T(k^2 + i - \lfloor\sqrt{k^2+i}\rfloor) \\ &= 1 + T(k^2 + i - k) \\ &= 1 + T((k-1)^2 + 2k - 1 + i - k) \\ &= 1 + T((k-1)^2 + k + i - 1) \\ &= 1 + T((k-1)^2 + j) \end{aligned}$$

เนื่องจาก $2(k-1)+1 \geq j = k+i-1 \geq (k-1)+1$ จากสมมติฐานอุปนัยบน $k-1$ จะได้ว่า $T((k-1)^2+j) = T((k-1)^2)+2$ จึงได้ว่า $T(k^2+i) = 3+T((k-1)^2)$ แต่เราทราบว่า $T((k-1)^2+2(k-1)+1) = 2+T((k-1)^2)$ จากสมมติฐานอุปนัยบน $k-1$ เมื่อแทน $i$ ด้วย $2(k-1)+1$ จึงสรุปได้ว่า $T(k^2) = T((k-1)^2)+2$ ทำให้

$$T(k^2+i) = 1 + (T((k-1)^2)+2) = 1 + T(k^2)$$

ต่อมาจะพิจารณากรณี $k+1 \leq i \leq 2k$ จะได้ว่า

$$\begin{aligned} T(k^2+i) &= 1 + T(k^2+i-\lfloor\sqrt{k^2+i}\rfloor) \\ &= 1 + T(k^2+i-k) \\ &= 1 + T(k^2+j) \end{aligned}$$

เมื่อ $j = i-k \in [1,k]$ จากที่ได้พิสูจน์ไปในกรณีก่อนหน้า จึงสรุปได้ว่า

$$T(k^2+i) = 1 + (T(k^2)+1) = T(k^2)+2$$

สุดท้ายจะพิจารณากรณี $i = 2k+1$ จะได้ว่า

$$\begin{aligned} T(k^2+i) &= T(k^2+2k+1) \\ &= 1 + T(k^2+2k+1-\lfloor\sqrt{k^2+2k+1}\rfloor) \\ &= 1 + T(k^2+2k+1-(k+1)) \\ &= 1 + T(k^2+k) \end{aligned}$$

จากที่พิสูจน์ไปในกรณีแรก จะได้ว่า $T(k^2+k) = T(k^2)+1$ จึงสรุปได้ว่า

$$T(k^2+2k+1) = T(k^2)+2$$

จากการอุปนัยเชิงคณิตศาสตร์ จึงสามารถสรุปได้ว่า

$$T(k^2+i) = \begin{cases} T(k^2)+1 & \text{ถ้า } 1 \leq i \leq k \\ T(k^2)+2 & \text{ถ้า } k+1 \leq i \leq 2k+1 \end{cases}$$

สำหรับทุกจำนวนเต็ม $k \geq 1$ □

**บทตั้ง 4.4.** สำหรับ $k \in \mathbb{N}_{>0}$ ใด ๆ จะได้ว่า $T(k^2) = 2k-1$

*พิสูจน์.* จะอุปนัยเชิงคณิตศาสตร์ สังเกตว่า $T(1^2) = 1 = 2(1)-1$ ต่อมาสำหรับจำนวนเต็มบวก $k$ ใด ๆ สมมติว่า $T(k^2) = 2k-1$ แล้ว จากบทตั้ง 4.3 จะได้ว่า $T((k+1)^2) = T(k^2+(2k+1)) = T(k^2)+2 = 2k-1+2 =$

$2k+1 = 2(k+1)-1$ จากการอุปนัยเชิงคณิตศาสตร์จึงสรุปได้ว่า $T(k^2) = 2k-1$ $\square$

**ทฤษฎีบท 4.5.** สำหรับ $n \in \mathbb{N}_{>0}$ ใด ๆ จะได้ว่า $T(n) \leq 2\sqrt{n}$

*พิสูจน์.* ให้ $n \in \mathbb{N}_{>0}$ และสมมติ $k = \lfloor\sqrt{n}\rfloor$ แล้วให้ $i = n - k^2$ จะได้ $n = k^2 + i$ เมื่อ $i \in [0, 2k]$ หาก $i = 0$ จะได้ว่า $n = k^2$ จึงได้ $T(n) = T(k^2) = 2k - 1 \leq 2k = 2\sqrt{n}$ ตามต้องการ แต่หาก $i \neq 0$ จะแยกได้เป็นสองกรณีคือ $i \in [1, k]$ กับ $i \in [k+1, 2k]$

กรณี $i \in [1, k]$ จะได้ $T(n) = T(k^2+i) = T(k^2) + 1 = 2k - 1 + 1 = 2k \leq 2\sqrt{n}$ ตามต้องการ

กรณี $i \in [k+1, 2k]$ จะได้ $T(n) = T(k^2+i) = T(k^2) + 2 = 2k - 1 + 2 = 2k + 1$ จึงเหลือเพียงการพิสูจน์ว่า $k + \frac{1}{2} \leq \sqrt{n} = \sqrt{k^2+i}$ เนื่องจาก

$$\begin{aligned}(k+\frac{1}{2})^2 &= k^2 + k + \frac{1}{4} \\ &\leq k^2 + i\end{aligned}$$

เพราะ $i \geq k + 1 \geq k + \frac{1}{4}$

จึงได้ $k + \frac{1}{2} \leq \sqrt{k^2+i} = \sqrt{n}$ จึงได้ $2\sqrt{n} \geq 2k+1 = T(n)$ ตามต้องการ $\square$

จึงได้ว่า $k = T(|Q|) \leq 2\sqrt{|Q|}$ จากทฤษฎีบท 4.5 ได้ค่าใช้จ่ายรวมไม่เกิน $3Nk \leq 6N\sqrt{|Q|}$ จึงได้ $\lambda = 6, a = 1, b = 0.5$

ถัดมาจะเป็นการพิสูจน์ทฤษฎีบท 4.1

**ทฤษฎีบท 4.1.** *ให้ลำดับของจำนวนจริงใด ๆ ความยาว $n$ จะได้ว่ามีลำดับย่อยไม่เพิ่ม (nonincreasing subsequence) ความยาวอย่างน้อย $\lfloor\sqrt{n}\rfloor$ อยู่ หรือ มีลำดับย่อยไม่ลด (nondecreasing subsequence) ความยาวอย่างน้อย $\lfloor\sqrt{n}\rfloor$ อยู่*

*พิสูจน์.* ให้ $(a_k)_{k=1}^n$ เป็นลำดับของจำนวนจริงความยาว $n$ และนิยามเซต $S = \{1, 2, 3, \ldots, n\}$ ประกอบกับความสัมพันธ์อันดับ $\preceq = \{(i,j) \in S \times S : i \leq j$ และ $a_i \leq a_j\}$ หากโซ่ที่ยาวที่สุดของ $\preceq$ ยาวอย่างน้อย $\lfloor\sqrt{n}\rfloor$ ก็จะได้ตามต้องการ แต่หากโซ่นั้นยาวน้อยกว่า $\lfloor\sqrt{n}\rfloor$ เราจะสร้างความสัมพันธ์อันดับใหม่คือ $\preceq' = \{(i,j) \in S \times S : i \leq j$ และ $a_i < a_j\} \cup \{(i,i) : i \in S\}$ จะได้ว่าโซ่ที่ยาวที่สุดของ $\preceq'$ ยาวน้อยกว่า $\sqrt{n}$ ด้วย แสดงว่าสามารถแบ่ง $(S, \preceq')$ ออกเป็น antichain ได้มากที่สุดไม่เกิน $\sqrt{n}$ antichain (จาก Dilworth's theorem) จึงได้ว่ามีอย่างน้อยหนึ่ง antichain ที่ขนาดอย่างน้อย $\sqrt{n}$ เพราะหากทุก antichain ขนาดน้อยกว่า $\sqrt{n}$ และมีจำนวน antichain ทั้งหมดไม่เกิน $\sqrt{n}$ จะได้จำนวนสมาชิกทั้งหมดน้อยกว่า $\sqrt{n}\sqrt{n} = n$ ขัดแย้งกับที่เราพูดไว้ว่า antichain นั้นเป็นส่วนแบ่งคั่นของสมาชิก $n$ ตัวพอดี

จาก antichain ที่มีขนาดอย่างน้อย $\sqrt{n}$ นั้น เราจะเรียก antichain นี้ว่า $B$ (ซึ่ง $B \subseteq S$) จะได้ว่า หากเรียง $B$ จากน้อยไปมากตามปกติแล้วจะสามารถเขียน $B$ เป็นลำดับ $b_1, b_2, \ldots, b_k$ เมื่อ $k \geq \sqrt{n}$ โดยที่ $b_1 < b_2 < \cdots < b_k$ และ $a_{b_1} \geq a_{b_2} \geq a_{b_3} \geq \cdots \geq a_{b_k}$ (เพราะหากมี $a_{b_i} < a_{b_{i+1}}$ แล้วจะได้ว่า $b_i \preceq' b_{i+1}$ ขัดแย้งกับสมบัติของ antichain) ทำให้ได้ว่า $a_{b_1}, a_{b_2}, \ldots, a_{b_k}$ เป็นลำดับย่อยไม่เพิ่มขนาด $k \geq \sqrt{n}$ ของลำดับ $(a_i)_{i=1}^n$ $\square$

# A ภาคผนวก: เครื่องมือจากคณิตวิเคราะห์

**บทตั้ง A.1** (อสมการสามเหลี่ยม)**.** สำหรับจำนวนจริง $a, b$ ใด ๆ จะได้ว่า $|a + b| \leq |a| + |b|$

*พิสูจน์.* ให้ $c \in \mathbb{R}$ ก่อนอื่นจะแสดงว่า $|1 + c| \leq 1 + |c|$ โดยสมมติ $c \geq 0$ จะได้ $|1 + c| = 1 + c = 1 + |c|$ จึงได้ $|1 + c| \leq 1 + |c|$ ด้วย ตามต้องการ

หาก $c < -1$ จะได้ $|1 + c| = -(1 + c) = -1 - c = -1 + |c| \leq 1 + |c|$ ตามต้องการ

หาก $-1 \leq c < 0$ จะได้ว่า $|1+c| = 1+c \leq 1-c = 1+|c|$ ตามต้องการ จากทุกกรณีจึงสรุปได้ว่า $|1+c| \leq 1+|c|$ สำหรับทุก $c \in \mathbb{R}$ ต่อมาจึงแทนค่า $c$ ด้วย $\frac{b}{a}$ กรณี $a \neq 0$ จึงได้

$$\left|1 + \frac{b}{a}\right| \leq 1 + \left|\frac{b}{a}\right|$$

เมื่อนำ $|a|$ มาคูณทั้งสองข้าง จะได้

$$\left||a| + \frac{b}{a}|a|\right| \leq |a| + |b|$$

ถ้า $a > 0$ จะได้ $|a + b| \leq |a| + |b|$ ทันที แต่ถ้า $a < 0$ จะได้ $|-a - b| \leq |a| + |b|$ แต่ $|-a - b| = |a + b|$ จึงได้ $|a + b| \leq |a| + |b|$

ส่วนกรณี $a = 0$ จะแยกออกไปต่างหาก เพราะไม่สามารถเขียน $\frac{b}{a}$ ได้ สังเกตได้ว่า $|a+b| = |0+b| = |b| = 0+|b| = |a| + |b|$ จึงได้ $|a + b| \leq |a| + |b|$ ตามต้องการ $\square$

**นิยาม A.2** (ลิมิตของฟังก์ชันค่าจริงของตัวแปรจำนวนจริง)**.** ให้ $f\colon I \to \mathbb{R}$ นิยามบนช่วงเปิด $I$ สำหรับจำนวนจริง $\ell$ เราจะกล่าวว่า

$$\lim_{x \to a} f(x) = \ell$$

ก็ต่อเมื่อสำหรับทุก $\varepsilon > 0$ จะมี $\delta > 0$ ที่ $(a - \delta, a + \delta) \subseteq I$ และสำหรับทุก $x' \in I$ ที่ $0 < |x' - a| < \delta$ นั้น $|f(x') - \ell| < \varepsilon$ และหากมี $\ell$ ดังกล่าว จะเรียกว่า $\ell$ เป็นลิมิตของ $f$ ที่ $a$

**ทฤษฎีบทประกอบ A.3.** หาก $f\colon I \to \mathbb{R}$ บนช่วงเปิด $I$ มีลิมิตที่ $a$ แล้ว จะมีลิมิตเพียงค่าเดียวเท่านั้น

*พิสูจน์.* สมมติ $\lim_{x \to a} f(x) = \ell_1$ และ $\lim_{x \to a} f(x) = \ell_2$ ด้วย จะได้ว่า สำหรับทุก $\varepsilon > 0$ จะมี $\delta_1, \delta_2 > 0$ ที่ $(a-\delta_1, a+\delta_1) \subseteq I$ และ $(a-\delta_2, a+\delta_2) \subseteq I$ และสำหรับทุก $x' \in (a-\delta_1, a+\delta_1) - \{a\}$ จะได้ $|f(x') - \ell_1| < \varepsilon_n$ และสำหรับทุก $x' \in (a - \delta_2, a + \delta_2) - \{a\}$ จะได้ $|f(x') - \ell_2| < \varepsilon_n$ จึงได้ว่า สำหรับทุก $x' \in (a - \delta, a + \delta) - \{a\}$ จะได้ $|f(x') - \ell_1| < \varepsilon_n$ และ $|f(x') - \ell_2| < \varepsilon_n$ เมื่อ $\delta = \min(\delta_1, \delta_2)$

พิจารณา $|\ell_1 - \ell_2|$ จากอสมการสามเหลี่ยมจะได้ว่า เราสามารถเลือก $x' \in \mathbb{R}$ สำหรับแต่ละ $\varepsilon_n$ ที่ทำให้ $|\ell_1 - \ell_2| = |\ell_1 - f(x') + f(x') - \ell_2| \leq |\ell_1 - f(x')| + |f(x') - \ell_2| < \varepsilon + \varepsilon = 2\varepsilon$ ได้เสมอ

เนื่องจาก ไม่ว่าจะเลือก $\varepsilon > 0$ เป็นอะไรก็ตาม จะมี $\delta_1, \delta_2 > 0$ เสมอ ทำให้ $\delta = \min(\delta_1, \delta_2) > 0$ มีอยู่เสมอ จึงได้ว่ามี $x' \in (a - \delta_1, a + \delta_1) - \{a\}$ เสมอ ทำให้ $|\ell_1 - \ell_2| < \varepsilon$

จึงสรุปได้ว่า $|\ell_1 - \ell_2| = 0$ (เพราะหากไม่เท่ากับ $0$ กล่าวคือให้ $d = |\ell_1 - \ell_2| > 0$ จะสามารถเลือก $\varepsilon = \frac{d}{4}$ ที่ทำให้ $|\ell_1 - \ell_2 < 2\varepsilon = \frac{d}{2} < d = |\ell_1 - \ell_2|$ ได้ จึงเกิดข้อขัดแย้ง) จากข้อสรุปจึงได้ว่า $|\ell_1 - \ell_2| = 0$ นั่นคือ $\ell_1 = \ell_2$ □

**นิยาม A.4** (ฟังก์ชันต่อเนื่อง)**.** เราเรียกฟังก์ชัน $f\colon I \to \mathbb{R}$ ว่าฟังก์ชันต่อเนื่องบน $I$ เมื่อ $I$ เป็นช่วงเปิด และ $\lim_{x\to a} f(x) = f(a)$ สำหรับทุก $a \in I$ และเขียนแทนด้วย $f \in \mathcal{C}^0(I)$

**นิยาม A.5** (อนุพันธ์ของฟังก์ชันของจำนวนจริง)**.** สำหรับ $f \in \mathcal{C}^0(I)$ ใด ๆ เรานิยามอนุพันธ์ของ $f$ ที่ $a$ ว่า

$$\lim_{h\to 0} \frac{f(a+h) - f(a)}{h}$$

หากลิมิตนี้ไม่มีอยู่จริง จะกล่าวว่า $f$ ไม่สามารถหาอนุพันธ์ที่ $a$ ได้ หากมีอยู่ จะเขียนแทนด้วย $f'(a)$

เราจะกล่าวว่าฟังก์ชัน $f$ หาอนุพันธ์ได้ (differentiable) บน $(L, R)$ ก็ต่อเมื่อ สำหรับทุก $a \in (L, R)$ นั้น $f$ สามารถหาอนุพันธ์ที่ $a$ ได้

**ทฤษฎีบท A.6.** หาก $f\colon (L, R) \to \mathbb{R}$ สามารถหาอนุพันธ์บน $(L, R)$ ได้ จะได้ว่า $f$ ต่อเนื่องบน $(L, R)$

*พิสูจน์.* สมมติ $f$ หาอนุพันธ์บน $(L, R)$ ได้ จะได้ว่าสำหรับทุก $a \in (L, R)$, $f'(a) = \lim_{h\to 0} \frac{f(a+h)-f(a)}{h}$ แต่เนื่องจาก $0 = \lim_{h\to 0} h$ จึงได้ว่า

$$\begin{aligned} f'(a) \cdot 0 &= \left(\lim_{h\to 0} \frac{f(a+h) - f(a)}{h}\right)\left(\lim_{h\to 0} h\right) \\ 0 &= \lim_{h\to 0}(f(a+h) - f(a)) \\ f(a) &= \lim_{h\to 0} f(a+h) \end{aligned}$$

จึงได้ว่า $f$ ต่อเนื่องที่ $a$ □

**นิยาม A.7.** สมมติ $A \subseteq \mathbb{R}$ ให้ $f\colon A \to \mathbb{R}$ เราจะกล่าวว่า $f$ มี local minimum ที่ $x_0$ เมื่อมี $\delta > 0$ ที่ $f(x) \geq f(x_0)$ สำหรับทุก $x \in A \cap (x_0 - \delta, x_0 + \delta)$

**นิยาม A.8.** เราจะกล่าวว่า $f\colon A \to \mathbb{R}$ มี global minimum ที่ $x_0$ เมื่อ $f(x) \geq f(x_0)$ สำหรับทุก $x \in A$ สังเกตว่าหาก $x_0$ เป็นจุดที่มี global minimum แล้ว $x_0$ ก็จะเป็นจุดที่มี local minimum ด้วย

เราจะได้บทพิสูจน์ของทฤษฎีบท 1.2 ดังต่อไปนี้

**ทฤษฎีบท 1.2.** *ให้ $f\colon (a, b) \to \mathbb{R}$ เป็นฟังก์ชัน ถ้า $f$ มี local minimum ที่ $x_0 \in (a, b)$ และ $f$ หาอนุพันธ์ได้ที่ $x_0$ จะได้ว่า $f'(x_0) = 0$*

*พิสูจน์.* สมมติ $f\colon (a,b) \to \mathbb{R}$ มี local minimum ที่ $x_0$ และ $f'(x_0)$ มีค่าอยู่จริง จะแสดงว่า $f'(x_0) = 0$ จากนิยาม local minimum จะได้ว่ามี $\delta > 0$ ที่ $f(x_0) \leq f(x)$ สำหรับทุก $x \in (x_0 - \delta, x_0 + \delta)$ จึงได้ว่า $f(x) - f(x_0) \geq 0$ สำหรับทุก $x \in (x_0 - \delta, x_0 + \delta)$ สังเกตว่าสำหรับ $x \leq x_0$ จะได้ว่า $\frac{f(x)-f(x_0)}{x-x_0} \leq 0$ และสำหรับ $x \geq x_0$ จะได้ว่า $\frac{f(x)-f(x_0)}{x-x_0} \geq 0$

เนื่องจาก $f'(x_0) = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = \lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}$ แต่

$$\lim_{x\to x_0^-} \frac{f(x)-f(x_0)}{x-x_0} \leq 0 \text{ และ } \lim_{x\to x_0^+} \frac{f(x)-f(x_0)}{x-x_0} \geq 0$$

และเนื่องจากลิมิตนี้มีอยู่ แสดงว่าลิมิตทางซ้ายและลิมิตทางขวาจะต้องเท่ากัน กล่าวคือ มีค่าเท่ากับศูนย์ จึงสรุปได้ว่า $f'(x_0) = 0$ $\square$