THE LONGEST INCREASING SUBSEQUENCE

MIHIR GUPTA
mihirgupta292@gmail.com

Abstract. We look at the Longest Increasing Subsequence (LIS) in the context of the Ulam-Hammersley problem. We introduce fundamental concepts such as permutations, increasing and decreasing subsequences, and the definitions of LIS and Longest Decreasing Subsequence (LDS). The historical development is traced from the Erdős-Szekeres theorem to the results of Baik, Deift, and Johansson (BDJ). We detail the Robinson-Schensted-Knuth (RSK) correspondence, proving the equivalence of the LIS length to the length of the first row in the resulting Standard Young Tableaux (SYT). Key results, including the Erdős-Szekeres theorem, Hook-Length Formula, and Plancherel measure, are discussed. We then focus on the Ulam-Hammersley problem, analyzing the asymptotic behavior of the LIS length in random permutations, with results by Logan-Shepp and Vershik-Kerov confirming Hammersley's conjecture. Finally, we present the BDJ theorem, which describes the limiting distribution of the LIS length and its connection to the Tracy-Widom distribution with a small connection to random matrix theory.

1. Introduction

This paper discusses the Longest Increasing subsequence in the context of the Ulam-Hammersley problem and various surrounding theorems. Let S_n denote the symmetric group of all permutations of n distinct numbers. We write permutation $\sigma \in S_n$ as a sequence, $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$. An increasing subsequence of σ is a subsequence $\sigma_{i_1}, \sigma_{i_2}, \ldots, \sigma_{i_k}$ that satisfies $\sigma_{i_1} < \sigma_{i_2} < \ldots < \sigma_{i_k}$ and $i_1 < i_2 < \ldots < i_k$. Similarly, a decreasing subsequence of σ is a subsequence $\sigma_{j_1}, \sigma_{j_2}, \ldots, \sigma_{j_l}$ that satisfies $\sigma_{j_1} > \sigma_{j_2} > \ldots > \sigma_{j_l}$ and $j_1 < j_2 < \ldots < j_l$. For example for $\sigma = (5, 6, 4, 2, 7, 1, 3)$, $(5, 6, 7)$ is an increasing subsequence and $(5, 4, 3)$ is a decreasing subsequence.

We denote the length of the Longest Increasing Subsequence (LIS) and the Longest Decreasing Subsequence (LDS) of σ as $L(\sigma)$ and $D(\sigma)$, respectively. For $\sigma = (5, 6, 4, 2, 7, 1, 3)$, $L(\sigma) = 3$ and $D(\sigma) = 3$.

The study of longest increasing subsequence began with Erdős and Szekeres [3] in 1935, who proved that any sequence of $n^2 + 1$ distinct numbers contains either an increasing or a decreasing subsequence of length $n + 1$. Ulam [11] later conjectured about the expected length of the longest increasing subsequence in a random permutation, leading to the Ulam-Hammersley problem. In the 1970s, Logan-Shepp [6] and Vershik-Kerov [12] independently showed that the expected length of LIS converges asymptotically to $2\sqrt{n}$. In 1999, Baik, Deift, and Johansson [2] determined the limiting distribution of the LIS length, linking it to the Tracy-Widom distribution and providing a precise description of the fluctuations around $2\sqrt{n}$.

In the second section of this paper we will discuss the RSK correspondance and the bijection that Schensted presented between random permutations in S_n and pairs of standard

Young Tableaux. In this section, we will also discuss the Erdős-Szekeres Theorem, Hook-Length Formula, and the Plancherel Measure. In the third section we will discuss the Ulam-Hammersley problem, and various advancements made since Ulam posed the question about the distribution of the LIS length in a random permutation. In the fourth section we will discuss the distribution of LIS length.

1.1. **Background.**

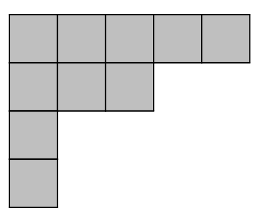
1.1.1. *Longest Increasing Subsequence.* The *Longest Increasing Subsequence* (LIS) of a sequence is defined as the longest such subsequence of the original sequence where the elements are in strictly increasing order. For a given sequence $A = \{a_1, a_2, \ldots, a_n\}$, a subsequence $A' = \{a_{i_1}, a_{i_2}, \ldots, a_{i_k}\}$ is an increasing subsequence if $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ and $a_{i_1} < a_{i_2} < \ldots < a_{i_k}$. Longest subsequence is the longest form of such subsequence.

1.1.2. *Longest Decreasing Subsequence.* The *Longest Decreasing Subsequence* (LDS) of a sequence is defined analogously to the LIS, but with elements in strictly decreasing order. For a given sequence $A = \{a_1, a_2, \ldots, a_n\}$, a subsequence $A' = \{a_{i_1}, a_{i_2}, \ldots, a_{i_k}\}$ is a decreasing subsequence if $1 \leq i_1 < i_2 < \ldots < i_k \leq n$ and $a_{i_1} > a_{i_2} > \ldots > a_{i_k}$.

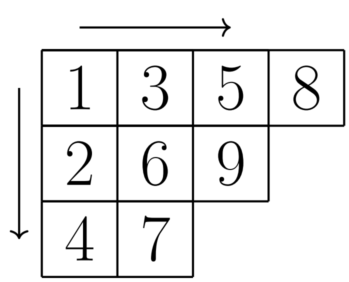
1.1.3. *Permutation.* A *permutation* of a set is an arrangement of its elements in a specific order. For instance, possible permutations of the set $\{1, 2, 3\}$ are:

$$(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1).$$

1.1.4. *Young Diagram.* For $n \in N$, we define a partition of n is a way to represent n as a sum of positive integers. Various summands represent the partition of n. When the summands are arranged in a decreasing order, the summands constitute parts of a *Young Diagram.* For instance, one possible partition when $n = 10$ is (5, 3, 1, 1) and the corresponding Young Diagram is shown below.



1.1.5. *Young Tableaux.* A *Young Tableaux*, referred to as Standard Young Tableux or SYT, is a Young Diagram where the cells of the diagram have been filled with distinct numbers. These numbers are arranged in an increasing order in each row and column. For instance, below we have a Young Tableaux where the cells have been filled with 9 consecutive numbers arranged in a non-consecutive permutation.



1.1.6. *Robinson-Schensted-Knuth (RSK) Correspondence.* The *RSK* correspondence is a combinatorial algorithm that converts any permutation to a pair of Standard Young Tableaux (SYT). The length of the LIS of the permutation is equal to the length of the first row of the SYT and length of LDS is equal to the length of the first column of the SYT.

Example 1.1.1. Consider the permutation $\sigma = (4, 3, 1, 2)$. Using the RSK correspondence, we obtain the Young tableaux:

1	2
3	
4	

The length of the LIS is 2, corresponding to the first row of the tableau. The length of the LDS is 3, corresponding to the first column of the tableau.

1.2. **Airplane Boarding.** In the context of airplane boarding, a naive model can be constructed where passengers board the plane in a specific order, represented by the sequence $w = a_1 a_2 \cdots a_n$ for seats 1, 2, ..., n. Each passenger takes one time unit to be seated after arriving at their seat. The boarding process can be analyzed using the concepts of increasing and decreasing subsequences. For instance, the total waiting time for all passengers is the length of the longest increasing subsequence (LIS) of the boarding sequence.

For a more sophisticated analysis, [1] demonstrated that the usual system of boarding from back-to-front does not significantly outperform a random boarding process. Instead, a more efficient method involves boarding passengers in the order of window seats first, followed by center seats, and finally aisle seats. This method minimizes the waiting time and congestion during the boarding process.

1.3. **Notations.** We use the following notations used in this paper.

- Sequences are denoted by σ.
- The length of a sequence σ is denoted by $|\sigma|$.
- Subscripts are used to indicate elements of a sequence, e.g., σ_i is the i-th element of the sequence σ.
- The Length of the LIS of a sequence σ is denoted by $L(\sigma)$.
- The Longest Decreasing Subsequence (LDS) of a sequence σ is denoted by $D(\sigma)$.
- The expected length of the LIS for a random permutation of n elements is denoted by l_n.
- S_n denotes the symmetric group of all permutations on n elements.
- λ represents shape of a Young Tableau.

2. The RSK Correspondence

The Robinson-Schensted-Knuth (RSK) algorithm constructs a bijection between a permutation and a pair of standard Young tableaux (SYT) that have the same shape.

2.1. **Algorithm.** Given a permutation $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$, the RSK algorithm works as follows:

(1) Start with an empty tableau.

(2) For each element σ_i in the permutation, insert it into the tableau using the following *row insertion* method:
   (a) Place σ_i in the first row, replacing the first element greater than σ_i. If σ_i is the largest element, place it at the end of the row.
   (b) The replaced element is bumped to the next row.
   (c) Continue the above two steps using the bumped element to be inserted in the subsequent rows.

   This process constructs the P-tableau (insertion tableau).
(3) Simultaneously, construct the Q-tableau (recording tableau) to record the order of insertions. Each square in the Q-tableau is given the number of the time step the square was created.

2.1.1. *Example Execution.* Consider the permutation $\sigma = (3, 5, 4, 9, 8, 2, 7)$:

Time Step	1	2	3	4	5	6	7
Insertion	3	5	4	9	8	2	7
P-tableau	3	3 5	3 4 5	3 4 9 5	3 4 8 5 9	2 4 8 3 9 5	2 4 7 3 8 5 9
Q-tableau	1	1 2	1 2 3	1 2 4 3	1 2 4 3 5	1 2 4 3 5 6	1 2 4 3 5 6 7

The length of the first row of the P-tableau is 3, which corresponds to the LIS of the permutation σ. The length of the first column of the P-tableau is 3, which corresponds to the LDS of the permutation σ.

2.2. **Bijection and SYT.**

The RSK algorithm ensures that both the P-tableau and Q-tableau are standard Young tableaux (SYT) of the same shape. The length of the first row of the P-tableau corresponds to the length of the LIS of the permutation and the height of the first column corresponds to the length of the LDS of the permutation.

Now we prove that the RSK correspondance results in a SYT whose first row's length is equal to the length of LIS. [8]

Definition 2.2.1. The *jth* basic subsequence of a permutation is the collection of numbers which are inserted into the *jth* column in the first row of the P-tablaeu

Lemma 2.2.2. *Each basic subseqence is a decreasing sequence*

Proof. Every time a number is inserted in the jth place in the first row of the P-tablaeu it must be less than the number that it displaces. ■

Lemma 2.2.3. *For any element in the j-th basic subsequence, there exists an element in the (j-1)-th basic subsequence that is smaller and appears earlier in the given sequence.*

Proof. The number in the (j-1)-th position of the first row, after inserting the given element of the j-th basic subsequence, is that element of the (j-1)-th basic subsequence. ∎

Theorem 2.2.4. *The number of columns in the P-tableau matches the length of the longest increasing subsequence of the related sequence.*

Proof. The number of columns is equal to the number of basic subsequences. As per Lemma 2.2.2, an increasing subsequence can have no more than one element from each basic subsequence. Lemma 2.2.3 shows that we can form an increasing subsequence with one element from each basic subsequence. ∎

Theorem 2.2.5. *The number of rows in the P-tableau matches the length of the longest decreasing subsequence of the related sequence.*

This theorem can be proven using symmetrical property of RSK, which states that running the RSK algorithm on the reverse sequence of a permutation of n numbers will result in a SYT that is the transpose of the SYT generated by running RSK algorithm on the original permutation.

2.3. **Erdős-Szekeres Theorem.** Erdős and Szekeres produced the first results concerning increasing and decreasing subsequences.

Theorem 2.3.1. *Let $\sigma \in S_n$ where $n > r \cdot s$ and $r, s \in \mathbb{N}$. Then, either the length of the longest increasing subsequence $L(\sigma)$ exceeds r or the length of the longest decreasing subsequence $D(\sigma)$ exceeds s.*

Proof. Suppose σ is a permutation of n elements, and assume for contradiction that both $L(\sigma) \leq r$ and $D(\sigma) \leq s$. Consider a Young tableau constructed from σ such that the tableau fits within a rectangle of width r and height s, thus containing $r \cdot s$ boxes.

We illustrate the SYT below:

$\sigma(1)$	$\sigma(2)$	$\cdots$	$\sigma(r)$
$\sigma(r+1)$	$\sigma(r+2)$	$\cdots$	$\sigma(2r)$
$\vdots$	$\vdots$	$\ddots$	$\vdots$
$\sigma((s-1)r+1)$	$\sigma((s-1)r+2)$	$\cdots$	$\sigma(rs)$

Since $n > r \cdot s$, there must be at least $r \cdot s + 1$ elements to place in the tableau. By the pigeonhole principle, inserting one more element into this already full tableau forces an increase in either the number of rows or the number of columns.

This implies that the tableau must have either more than r columns or more than s rows, which contradicts our assumption. Therefore, we must have either $L(\sigma) > r$ or $D(\sigma) > s$. ∎

2.4. **The Hook Length Formula.**

2.4.1. *Definition.* Let f^λ denote the number of Standard Young Tableaux (SYT) of shape λ. Where λ represents a partition of a positive integer n, we can express this as:

$$\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_k)$$

where $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_k \geq 1$ and $\sum_{i=1}^{k} \lambda_i = n$.

The RSK correspondence provides a bijection from permutations to pairs of Standard Young Tableaux (SYT), P and Q, both of which have the same shape represented using λ. The term f^λ represents the number of such tableaux, and therefore $(f^\lambda)^2$ represents all the permutations covered by λ. The summation over all these, for all possible shapes λ, is equal to the total number of permutations: $n!$:

$$\sum_{\lambda \vdash n} (f^\lambda)^2 = n!. \tag{2.4.1}$$

MacMahon initially provided a formula for f^λ (the Young–Frobenius formula at that time) in 1916 using difference methods. This formula was later simplified by Frame, Robinson, and Thrall.

Let u represent a cell in the Young diagram of λ, denoted as $u \in \lambda$. The hook length $h(u)$ of a cell u is defined as the total number of cells that are directly to the right of u and directly below u, including the cell u itself. For instance, for the partition $\lambda = (3, 2, 2)$, the hook lengths are:

5	4	1
3	2	
2	1	

Theorem 2.4.1. *The hook-length formula given by Frame, Robinson, and Thrall states that if $\lambda \vdash n$, then*

$$f^\lambda = \frac{n!}{\prod_{u\in\lambda} h(u)} = \frac{n!}{H(\lambda)}$$

Proof. For $\lambda = (3, 2, 2)$, we have

$$f^{(3,2,2)} = \frac{7!}{5 \cdot 4 \cdot 1 \cdot 3 \cdot 2 \cdot 2 \cdot 1} = 21.$$

For $f^{(3,2)}$ here are the possible SYTs:

1	2	3
4	5	

1	3	5
2	4	

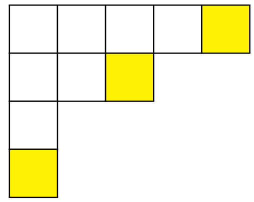
1	2	4
3	5	

1	3	4
2	5	

1	2	5
3	4	

$$f^{(3,2)} = \frac{5!}{4 \cdot 3 \cdot 2 \cdot 1 \cdot 1} = 5.$$

Greene, Nijenhuis, and Wilf, [4], proved the hook length formula using probabilistic hook walks. Given a standard young tableau with n boxes, n has to be in one of corner boxes (v) that have the hook length of 1. Removing this corner box leaves us with another standard young tableau with 1 less box. For instance, in the tableau below n can only appear in one of the yellow boxes.



$$f^{(5,3,1,1)} = f^{(4,3,1,1)} + f^{(5,2,1,1)} + f^{(5,3,1)}$$

We can prove the hook length formula using induction for all corners v of λ.

$$f^{\lambda} = \sum_{v} f^{(\lambda - v)}$$

For the base case when there are no boxes or just one box, $f_{\emptyset} = 1$ and $f_1 = 1$. For the inductive step, we need to demonstrate that

$$\frac{n!}{H(\lambda)} = \sum_{v \text{ corner of } \lambda} \frac{(n-1)!}{H(\lambda - v)}. \tag{2.4.2}$$

Or

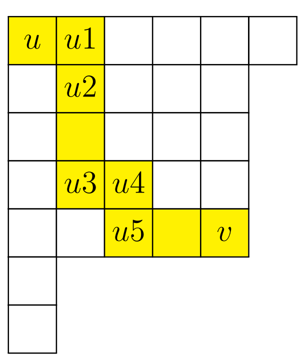
$$1 = \sum_{v \text{ corner}} \frac{1}{n} \frac{H(\lambda)}{H(\lambda - v)}.$$

The above can be thought of as a probability problem where several non-negative real numbers sum to 1. The following represents probability of "hook walk" to the corner v of λ from any point in λ.

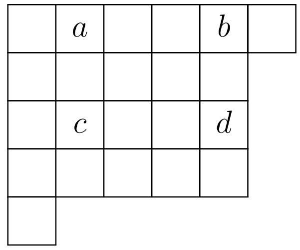
$$\sum_{u} P(u, v) = \frac{H(\lambda)}{H(\lambda - v)}$$

In developing the proof for the hook length formula, we referenced [9].

Hook Walk: Starting at u in the following SYT, jump from u to any square in the hook of u with equal probability. Repeating this process until you reach a corner box, in this case v, represents hook path from u to v.



Observation: For a fixed v, the hook walk stays within the top left corner and v. Also, for any rectangle with corners a, b, c, d, similar to that shown in the picture below

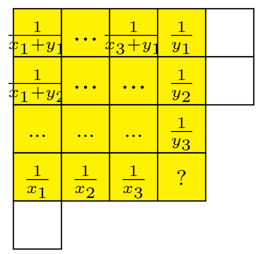


following is true.

$$h(a) + h(d) = h(b) + h(c) \implies h(a) - 1 + h(d) - 1 = h(b) - 1 + h(d) - 1$$

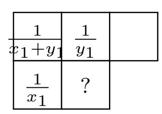
Specifically, if d is a corner cell $h(d) - 1 = 0$.

We will represent $h(i) - 1$ as x_i along x axis and as y_i along the y axis. The diagram below shows probability of a random hook step in each box. For instance, each box in the hook path of top-left box has $\frac{1}{x_1+y_1}$ probability.



Before calculating the total probability of all hook-walks, we will calculate the probabilities of all "lattice-paths", that involves moving down or moving right one box in each step.

For a Young diagram involving 2 by 2 grid, the boxes have probabilities as follows.



The total probability of using all lattice path from top-left to the bottom-right box is

$$\frac{1}{x_1 + y_1}.\frac{1}{y_1} + \frac{1}{x_1 + y_1}.\frac{1}{x_1} = \frac{1}{x_1.y_1}$$

By induction, for a Young diagram λ involving (k + 1) x (l + 1) rectangle, the sum of all the probabilities using all lattice paths from top-left (A) to bottom-right (B) is

$$\sum_{P:A->B} w(P) = \frac{1}{x_1.x_2...x_k.y_1.y_2...y_l}$$

Replacing "lattice path" with "hook-walk" in the above equation, gives us sum of all the probabilities of hook-walks from any box in λ to B.

$$\sum_{P:hook-walk->B} w(P) = (1+\frac{1}{x_1}).(1+\frac{1}{x_2})...(1+\frac{1}{x_k}).(1+\frac{1}{y_1}).(1+\frac{1}{y_2})...(1+\frac{1}{y_l})$$

Expanding the product shows us various hook paths from various boxes in λ to B. For instance 1 represents the start at B, and $\frac{1}{x_1.y_1}$ represents start from top-left to reach B in two hops via boxes labeled either x_1 or y_1.

Calling the boxes above or to the left of v the "co-hook of v", the above formulation for hook-walk can also be expressed as

$$\sum_u P(u,v) = \prod_{t\in co-hook(B)} (1+\frac{1}{h(t)-1})$$

Since on removal of v, all the hook length for boxes in co-hook of v decrease by 1 and other terms in the rectangle (λ - (all the boxes in co-hook)) cancel out, the above is equal to

$$\sum_u P(u,v) = \frac{H(\lambda)}{H(\lambda - v)}$$

For any fixed u, since any hook walk will end at some corner.

$$\sum_v P(u,v) = 1$$

Adding over all $u \in \lambda$ and all corner boxes v, since there are n boxes in λ.

$$\sum_{u\in\lambda}\sum_v P(u,v) = n$$

Or

$$\sum_v\sum_{u\in\lambda} P(u,v) = n$$

Substituting

$$\sum_v \frac{H(\lambda)}{H(\lambda - v)} = n$$

Multiplying both sides by $(n-1)!$ and rearranging, we get

$$\frac{n!}{H(\lambda)} = \sum_{v \text{ corner of } \lambda} \frac{(n-1)!}{H(\lambda - v)}$$

The above formula proves Equation 2.4.2 the inductive relationship we needed to prove the hook length formula. ■

2.5. **Plancherel Measure.** Plancherel measure links the Longest Increasing Subsequence (LIS) problem to the distribution of Young diagrams. For a partition λ of n, it's defined as:

$$P(\lambda^{(n)} = \lambda) = \frac{(f^\lambda)^2}{n!} \tag{2.5.1}$$

This measure describes the probability distribution of Young diagrams obtained from applying the Robinson-Schensted algorithm to random permutations in S_n.

For a random permutation $\sigma_n \in S_n$, the length of its LIS $(L(\sigma_n))$ has the same distribution as the length of the first row $(\lambda_1^{(n)})$ of a random Young diagram $\lambda^{(n)}$ chosen under Plancherel measure.

Studying $\lambda_1^{(n)}$ under Plancherel measure is thus equivalent to analyzing $L(\sigma_n)$, providing a new approach to the Ulam-Hammersley problem.

3. Ulam-Hammersley Problem

The Ulam-Hammersley problem is concerned with finding the asymptotic behavior of the expected length of the longest increasing subsequence (LIS) in a random permutation. For a permutation $\sigma \in S_n$, let $L(\sigma)$ denote the length of the LIS of σ. Define:

$$l_n = \frac{1}{n!} \sum_{\sigma \in S_n} L(\sigma). \tag{3.0.1}$$

The first few values of l_n are:

$$l_1 = 1.00, \quad l_2 = 1.50, \quad l_3 = 2.00, \quad l_4 = 2.41, \quad l_5 = 2.79$$

$$l_6 = 3.14, \quad l_7 = 3.47, \quad l_8 = 3.77, \quad l_9 = 4.06, \quad l_{10} = 4.33$$

In this section, we are interested in determining the asymptotic behavior of l_n as n becomes very large. Ulam briefly discussed the idea of studying the statistical distribution of l_n in 1961. John M. Hammersley undertook the first serious study of the Ulam problem in 1970. Henceforth, the problem is referred to as the Ulam-Hammersley problem.

3.1. **First Bounds - Hammersley.** In his seminal 1972 paper [5], Hammersley provided the first significant bounds on the expected value of the Longest Increasing Subsequence (LIS). He showed that for a random permutation $\sigma \in S_n$, the normalized expected length of the LIS must fall within certain bounds:

$$\frac{\pi}{2} \leq \lim_{n\to\infty} \frac{l_n}{\sqrt{n}} \leq e \tag{3.1.1}$$

To prove the lower bound $\frac{\pi}{2}$, we consider the points of a Poisson process with unit parameter on a square of area N. We referenced [5] while working on this proof.

A Poisson process with a unit rate parameter on a square of area N is considered. For any point P in the square, let $Q(P)$ denote the point of the Poisson process which is northeast of P and as close to P as possible. Let Q_0 be the southwest corner of the square. The sequence of points $\{Q_i\}_{i=0}^{\infty}$ is defined recursively by:

$$Q_{i+1} = Q(Q_i), \quad i = 0, 1, 2, \ldots$$

The expected value of the horizontal (or vertical) projection of the distance between consecutive points Q_i and Q_{i+1} is derived through integration:

$$\int_0^\infty \int_0^{\frac{\pi}{2}} re^{-\frac{\pi r^2}{4}} r \cos\theta \, d\theta \, dr = \frac{2}{\pi}. \tag{3.1.2}$$

This integral accounts for the projected length on one axis, taking into consideration the exponential decay in density of points as the distance increases, reflecting the properties of the Poisson process.

The projected lengths are independent; hence, the strong law of large numbers can be applied. As the area N becomes large, the sum of $\frac{\pi}{2}\sqrt{N}$ terms from the sequence, which are the projections, approaches the boundary of the square. It can be shown that:

$$\frac{\pi}{2}\sqrt{N} + o(\sqrt{N})$$

captures the behavior of the sequence, ensuring that the chain formed by these points does not exceed the boundary before reaching a significant length. This confirms the lower bound.

Therefore, the expected length of the LIS in a random permutation of n elements satisfies:

$$\lim_{n\to\infty} \frac{l_n}{\sqrt{n}} = E[L_n] \geq \frac{\pi}{2}\sqrt{n}.$$

We will now prove for the upper bound e. We referenced [7] while working on this proof.

Let $X_{n,k}$ denote the number of increasing subsequences of length k in the permutation σ_n.

The expected value of $X_{n,k}$ is calculated by considering all $\binom{n}{k}$ subsequences, where each has a $\frac{1}{k!}$ probability of being increasing. This yields:

$$E(X_{n,k}) = \frac{1}{k!}\binom{n}{k}$$

To find the probability that the longest increasing subsequence $L(\sigma_n)$ has at least k elements, use Markov's inequality:

$$P(L(\sigma_n) \geq k) = P(X_{n,k} \geq 1) \leq E(X_{n,k}) = \frac{1}{k!}\binom{n}{k}$$

Simplifying $\binom{n}{k}$ as $\frac{n^k}{k!}$, and then rearranging and bounding, the expression simplifies to:

$$\frac{n^k}{(k!)^2}$$

By fixing $k = [(1+\delta)e\sqrt{n}]$ (where $\delta > 0$ is small), we can further simplify the probability expression. The key here is to note that this probability converges to zero as $n \to \infty$

exponentially fast, making use of Stirling's approximation and bounds for the binomial coefficient.

This setup ensures that $L(\sigma_n)$ typically does not exceed $(1+\delta)e\sqrt{n}$, reinforcing that the rate of growth of the longest increasing subsequence is fundamentally rooted in the square root of n.

The proof concludes that the expectation ℓ_n (the expected length of $L(\sigma_n)$) is at most $(1+\delta)e\sqrt{n}$ plus terms that vanish faster than $\sqrt{n}$, and this bound holds not only for the expected value but also typically (in a probabilistic sense).

Hammersley further conjectured the following limit exists.

$$c = \lim_{n\to\infty} \frac{E(L(\sigma))}{\sqrt{n}} \tag{3.1.3}$$

He further conjectured that c equals 2. This conjecture turned out to be correct, as later proved by Logan and Shepp (1977) and Vershik and Kerov (1977) independently.

3.2. **Confirming Hammersley's Conjecture.** In 1977, two independent teams proved Hammersley's conjecture about the limit of the expected length of the Longest Increasing Subsequence (LIS). We present their results via details from [10].

Expected Length of LIS:

$$E(n) = \frac{1}{n!}\sum_{\lambda\vdash n} \lambda_1 (f^\lambda)^2.$$

Here, λ is a partition of n, λ_1 is the length of the longest increasing sequence in λ, and f^λ is the number of standard Young tableaux of shape λ.

The RSK correspondence shows that:

$$n! = \sum_{\lambda\vdash n} (f^\lambda)^2.$$

Given that the number of terms in the sum of the squares of f^λ is very small compared to $n!$, the maximum value of f^λ is close to $\sqrt{n!}$.

Using the above insights $E(n)$ approximates as:

$$E(n) \approx \frac{1}{n!}(\lambda_1^n (f^{\lambda^n}))^2 \approx (\lambda^n)_1,$$

where λ^n is the partition that maximizes f^λ. In order to maximize f^λ we need to minimize the product of various hook lengths in λ, $(\prod_{u\in\lambda} h(u))$, via Theorem 2.4.1.

Since we are interested in the behavior of λ^n as $n \to \infty$ we will normalize the Young diagram of any partition λ to have area one ($\implies$ each square of the diagram has length $1/\sqrt{n}$). The upper boundary or the y-axis of the diagram is directed to the right, and the left boundary or the x-axis is directed downwards.

Limiting Curve $\Psi(x)$: As $n \to \infty$, it is reasonable to assume that the boundary of the partition λ^n approaches some limiting curve $y = \Psi(x)$. Assuming this curve intersects the x-axis at $x = b$:

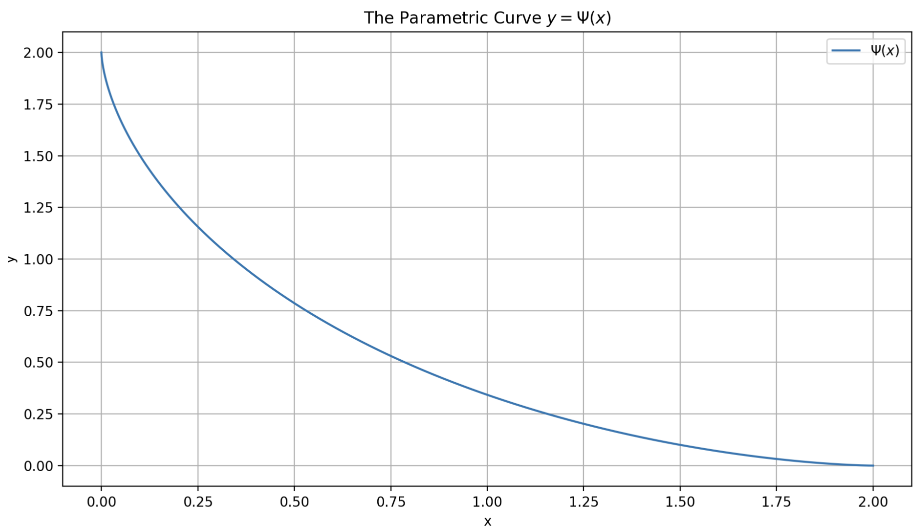
$$c := \lim_{n\to\infty} \frac{E(n)}{\sqrt{n}} \geq b.$$

Logan-Shepp noted that c could be larger than b if the first few parts of λ^n stretch out along the x-axis.

Logan-Shepp and Vershik-Kerov Result: Logan-Shepp [6] and Vershik-Kerov [12], independently, derived the curve $y = \Psi(x)$ as a solution to a variational problem. The normalized hook-length at (x, y) is given by $f(x)-y+f^{-1}(y)-x$. They minimize the functional $I(f) = \iint_A \log(f(x) - y + f^{-1}(y) - x)\,dx\,dy$, subject to the normalization $\iint_A dx\,dy = 1$.

Parametric Form of $\Psi(x)$: The curve $y = \Psi(x)$ is given parametrically by:

$$x = y + 2\cos\theta, \quad y = \frac{2}{\pi}(\sin\theta - \theta\cos\theta) \quad \text{for} \quad 0 \leq \theta \leq \pi.$$



This curve intersects the x-axis at $x = 2$, suggesting that $c \geq 2$.

The equation shows $E(n) \sim 2\sqrt{n}$, summarizing that the expected length of the longest increasing subsequence for a random permutation of n elements asymptotically approaches $2\sqrt{n}$.

Lemma 3.2.1 (Logan and Shepp, 1977, [6])**.** *For random permutations in S_n,*

(3.2.1)
$$\liminf_{n\to\infty} \frac{l_n}{\sqrt{n}} \geq 2$$

Vershik and Kerov, in their paper, showed that $c \leq 2$ through a clever use of the RSK algorithm.

Lemma 3.2.2 (Vershik and Kerov, 1977, [12])**.** *For random permutations in S_n,*

(3.2.2)
$$\limsup_{n\to\infty} \frac{l_n}{\sqrt{n}} \leq 2$$

Theorem 3.2.3. *The limit of the expected length of the LIS for random permutations exists and equals 2:*

$$\lim_{n\to\infty} \frac{l_n}{\sqrt{n}} = 2 \tag{3.2.3}$$

Proof. Follows directly from the combination of the lemmas by Logan-Shepp and Vershik-Kerov. ■

This result marked a significant advancement in the study of Longest Increasing Subsequences and proved that Hammersley's conjecture is correct.

4. The Distribution of LIS

4.1. **BDJ Theorem.** Let is_n denote the function $is : S_n \to \mathbb{Z}$. In 1999, Baik, Deift, and Johansson determined the entire limiting distribution of is_n, significantly extending earlier results on its expectation.

Theorem 4.1.1 (Baik-Deift-Johansson)**.** *For random (uniform) $\sigma \in S_n$ and all $t \in \mathbb{R}$,*

$$\lim_{n\to\infty} Prob\left(\frac{is_n(w) - 2\sqrt{n}}{n^{1/6}} \leq t\right) = F(t),$$

where $F(t)$ is the Tracy-Widom distribution.

The Tracy-Widom distribution $F(t)$ is defined as:

$$F(t) = \exp\left(-\int_t^\infty (x-t)u(x)^2 dx\right),$$

where $u(x)$ is the solution to the Painlevé II equation:

$$u''(x) = 2u(x)^3 + xu(x),$$

subject to the condition $u(x) \sim -\frac{e^{-\frac{2}{3}x^{3/2}}}{2\sqrt{\pi}x^{1/4}}$ as $x \to \infty$.

This theorem confirms the $2\sqrt{n}$ leading term in $E(n)$ found by Vershik-Kerov and Logan-Shepp, and provides the scale of fluctuations around this value, which are of order $n^{1/6}$.

The BDJ Theorem allows computation of limiting moments of $is_n(w)$. For instance, the variance of is_n as $n \to \infty$ is:

$$\lim_{n\to\infty} \frac{\mathrm{Var}(is_n)}{n^{1/3}} = \int t^2 dF(t) - \left(\int t dF(t)\right)^2 = 0.8131947928...$$

The expectation $E(n)$ can be expressed more precisely as:

$$E(n) = 2\sqrt{n} + \alpha n^{1/6} + o(n^{1/6}),$$

where $\alpha = \int t dF(t) = -1.7710868074...$.

This result provides the second term in the asymptotic behavior of $E(n)$, refining our understanding of the expected LIS length.

Acknowledgments

The author would like to thank Simon Rubenstein-Salzedo and Carson Mitchell for their guidance and support in the writing and research of this paper.

References

[1] E Bachmat, D Berend, L Sapir, S Skiena, and N Stolyarov. Analysis of aeroplane boarding via spacetime geometry and random matrix theory. *Journal of Physics A: Mathematical and General*, 39(29):L453–L459, July 2006.

[2] Jinho Baik, Percy Deift, and Kurt Johansson. On the distribution of the length of the longest increasing subsequence of random permutations. *Journal of the American Mathematical Society*, 12(4):1119–1178, 1999.

[3] Paul Erdös and George Szekeres. A combinatorial problem in geometry. *Compositio mathematica*, 2:463–470, 1935.

[4] Curtis Greene, Albert Nijenhuis, and Herbert S Wilf. A probabilistic proof of a formula for the number of young tableaux of a given shape. In *Young Tableaux in Combinatorics, Invariant Theory, and Algebra*, pages 17–22. Elsevier, 1982.

[5] John M Hammersley. A few seedlings of research. In *Proc. Sixth Berkeley Symp. Math. Statist. and Probability*, volume 1, pages 345–394, 1972.

[6] Benjamin F Logan and Larry A Shepp. A variational problem for random young tableaux. *Advances in mathematics*, 26(2):206–222, 1977.

[7] Dan Romik. *The surprising mathematics of longest increasing subsequences.* Number 4. Cambridge University Press, 2015.

[8] Craige Schensted. Longest increasing and decreasing subsequences. *Canadian Journal of mathematics*, 13:179–191, 1961.

[9] Shiyue and Andrew. Young tableaux and probability. Mathcamp 2019, August 2019.

[10] Richard P Stanley. Increasing and decreasing subsequences of permutations and their variants. *arXiv preprint math/0512035*, 2005.

[11] Stanislaw M Ulam. Monte carlo calculations in problems of mathematical physics. *Modern Mathematics for the Engineers*, 261:281, 1961.

[12] Anatolii Moiseevich Vershik and Sergei Vasil'evich Kerov. Asymptotics of the plancherel measure of the symmetric group and the limiting form of young tableaux. In *Doklady akademii nauk*, volume 233, pages 1024–1027. Russian Academy of Sciences, 1977.