

Санкт-Петербургский государственный университет
Кафедра теории вероятностей и математической статистики
Группа 22.Б13-мм

Агирре Кастильо Николас Франсиско

Реализация стохастической процедуры роста диаграмм Юнга

Отчёт по учебной практике
в форме «Производственное задание»

Научный руководитель:
Доцент кафедры теории вероятностей и математической статистики Якубович Ю.В.

Санкт-Петербург
2025

Оглавление

Введение	3
1. Постановка задачи	4
2. Обзор	5
2.1. Основные понятия	5
2.2. Классические модели	5
2.3. Выбранный модель	7
3. Реализация	8
3.1. Библиотеки	8
3.2. Диаграммы	8
3.3. Процесс	9
4. Результаты	11
4.1. Визуализация результатов	11
4.2. Анализ	16
Заключение	17
Список литературы	18

Введение

Диаграммы Юнга являются важным комбинаторным объектом, который активно изучается в различных областях математики. Они находят применение в теории представлений, статистической физике, теории вероятностей и ряде других дисциплин. В частности, изучение случайных процессов роста диаграмм Юнга позволяет глубже понять закономерности формирования сложных структур и их асимптотическое поведение.

Модели роста диаграмм Юнга представляют собой вероятностные процессы, в которых новая клетка добавляется по определённым стохастическим правилам. Такие модели находят аналогии в реальных физических и социальных системах, включая динамику роста кристаллов, эволюцию городских агломераций, а также развитие популяций. Анализ их предельных форм и статистических характеристик помогает выявить универсальные закономерности, характерные для широкого класса случайных процессов.

Настоящее исследование посвящено изучению одной из таких моделей роста, основанной на вероятностном добавлении клеток в зависимости от геометрических характеристик диаграммы. Рассматриваемый подход позволяет выявить свойства случайных диаграмм Юнга.

1. Постановка задачи

Цель настоящей работы — анализировать поведение процесса роста диаграммы Юнга, имеющей вероятность добавления новой клетки пропорциональную периметру прямоугольника, созданного от начала координат до верхней правой вершины новой клетки. Для реализации цели были установлены следующие задачи:

1. Проанализировать известные математические модели роста диаграммы Юнга.
2. Выбрать технологии и программированные библиотеки, наиболее подходящие для реализации работы.
3. Создать программу для визуализации разных характеристик прогресса роста диаграммы.
4. Провести анализ полученных результатов.

2. Обзор

2.1. Основные понятия

2.1.1. Диаграмма Юнга

Диаграммы Юнга являются важным инструментом в комбинаторике и теории представлений, позволяя наглядно представлять разбиения натуральных чисел и изучать их свойства. Они находят применение в различных областях математики, включая теорию представлений симметрических и полных линейных групп.

Диаграмма Юнга[6], соответствующая разбиению λ натурального числа n — это множество клеток, расположенных в k строках, где:

- $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_k)$
- $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_k > 0$
- $\lambda_1 + \lambda_2 + \dots + \lambda_k = n$
- i -я строка содержит λ_i клеток, выровненных по левой границе.

2.1.2. Модели роста диаграмм Юнга

Введение в стохастические модели роста диаграмм Юнга позволяет понять, как случайные процессы влияют на форму и структуру диаграмм, а также выявить связи с другими разделами математики, такими как теория случайных матриц и ортогональные полиномы.

Стохастические модели роста диаграмм Юнга описывают случайные процессы, в которых диаграммы изменяются во времени путем добавления ячеек по определённым вероятностным правилам.

2.2. Классические модели

Существует несколько известных моделей роста диаграмм Юнга [2], каждая из которых описывает динамику изменения диаграмм по определённым правилам.

2.2.1. Модель Роста

Рост [3] исследовал поведение одномерной асимметричной системы частиц в предельном случае. В дискретные моменты времени такую систему можно описать как случайно увеличивающуюся диаграмму Юнга, где на каждом шаге случайно выбирается внутренний угол и заполняется новым квадратом. Согласно результатам Роста, этот процесс стремится к предельной форме, заданной уравнением $\sqrt{x} + \sqrt{y} = 6^{1/4}$, если применяется масштабирование $a_n = n^{1/2}$.

2.2.2. Модель Морана

Модель Морана играет важную роль в математической популяционной генетике с бесконечным числом аллелей. Она представляет собой процесс рождения и гибели, который можно описать через диаграммы Юнга: на каждом этапе происходит чередование рождения и удаления квадратов, так что размер диаграммы остается фиксированным и равным n .

Правило рождения в этой модели совпадает с таковым в модели Саймона (с параметром μ), а удаление происходит во внешнем углу, связанном с длиной строки случайно выбранного квадрата. Было установлено, что стационарное распределение этого процесса соответствует распределению Эвенса с параметром $\theta = n\mu/(1 - \mu)$.

При фиксированном θ (то есть, если $\mu \approx 1/n$) это распределение не приводит к предельной форме. Однако такая форма существует при масштабировании $a_n = 1/\mu$, когда $n\mu \rightarrow \infty$.

2.2.3. Модель Саймона

Герберта Саймона исследовал математическую модель городского роста. Она была предложена для объяснения универсального наблюдения, согласно которому размеры городов имеют тенденцию следовать степенному закону, известному как «закон Ципфа для городов» [4]. Данная модель многократно переоткрывалась и фигурирует под названиями «бесконечная урновая модель Пойа», «процесс Юла» и «модель

предпочтительного присоединения».

Модель Саймона можно представить как случайно растущую диаграмму Юнга. На каждом шаге новый квадрат создает новую строку с вероятностью μ , а в противном случае он добавляется во внутренний угол, который выбирается случайным образом среди существующих квадратов в зависимости от длины строки. Согласно результату Саймона, при неограниченном увеличении числа квадратов ожидаемое количество $E[r_k]$ строк длины k стремится к степенному закону.

2.2.4. Модель случайного добавления ячеек

В этой модели каждая ячейка добавляется в диаграмму Юнга случайным образом с определённой вероятностью, при этом учитываются ограничения, связанные с формой диаграммы.

2.3. Выбранный модель

В данной работе предложена модель роста, основанная на случайном добавлении ячеек. Вероятность присоединения новой ячейки определяется пропорционально периметру прямоугольника, образуемого из исходной точки до предложенной ячейки.

$$P(c) = (S(c)^\alpha)/S$$

, где:

- $P(c)$ — вероятность клетки.
- $S(c)$ — периметр созданного прямоугольника.
- α — параметр.
- S — сумма периметров (всех предлагаемых ячеек на данном шаге), возведенных в степень α .

3. Реализация

3.1. Библиотеки

Для выполнения работы [1] были выбраны разные библиотеки Python, которые облегчают визуализацию результатов и процесс разработки программы.

matplotlib.pyplot[5] — это набор функций, упрощающих создание графиков. Она является частью более широкой библиотеки Matplotlib, которая очень популярна среди сообщества ученых данных, аналитиков и разработчиков благодаря своей способности создавать графики высокого качества простым и эффективным способом. Эта библиотека была выбрана для визуализации диаграммы.

random используется для генерации случайных чисел и выполнения случайных выборок. Она предоставляет функции для создания случайных чисел, выбора случайных элементов из последовательностей и выполнения других операций, связанных со случайностью.

3.2. Диаграммы

Для наблюдения различных характеристик и свойств, которые можно получить при построении диаграммы Юнга с выбранной моделью, было решено создать различные диаграммы:

- Общая диаграмма: Диаграмма Юнга через определённого количества шагов.
- Диаграмма частоты: Повторяем процесс построения диаграммы n раз и показываем их в одном графике. Цвет клетки зависит от ее частоты в диаграммах.
- Диаграмма на определенном шаге: Процесс построения диаграммы останавливается для показа графика в данном моменте. Далее построение продолжается и в конце можно визуализировать результаты.

3.3. Процесс

На этапе разработки программы существуют различные ключевые моменты функционирования кода и алгоритма, который используется для построения диаграмм работы.

3.3.1. Класс `Diagram`

Этот класс позволяет построить общую диаграмму (`main.py`). Важные методы, которые он включает:

- **`get_possible_moves`**: возвращает массив клеток, которых могут быть добавлены. Способ определения этой возможности заключается в проверке наличия соседней клетки слева (или оси ординат) и под ней (или оси абсцисс).
- **`get_S`**: вычисляет $S(c)$ (вес клетки).
- **`simulate`**: строит диаграмму Юнга.
- **`visualize`**: создает и показывает график диаграммы для ее визуализации.

3.3.2. Класс `Frequency_Diagram`

Этот класс позволяет построить диаграмму частоты (`main_freq.py`). Важные методы, которые он включает:

- **`create_diagrams`**: создает n диаграммы для сравнения частоты клеток.
- **`count_reps`**: вычисляет частоту каждой клетки.
- **`generate_colors`**: добавляет цвет для каждой клетки зависимо от частоты.
- **`visualize`**: создает и показывает график диаграммы для ее визуализации.

3.3.3. Важные изменения

- Диаграмма частоты (`main_freq.py`) использует похожий класс `Diagram`, но без метода `visualize`, так как не нужно показывать отдельные диаграммы.
- Для построения диаграммы на определенном шаге (`main_prog.py`) используется класс `Diagram`, но с добавлением остановки на определенном шаге. Можно добавить несколько остановок.

4. Результаты

4.1. Визуализация результатов

4.1.1. Общая диаграмма

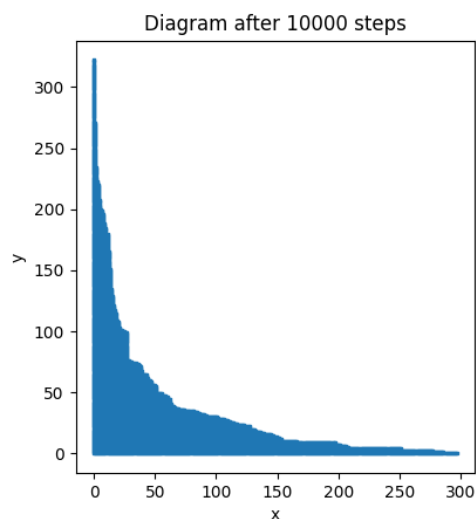


Рис. 1: Общая диаграмма при $\alpha = 0.5$

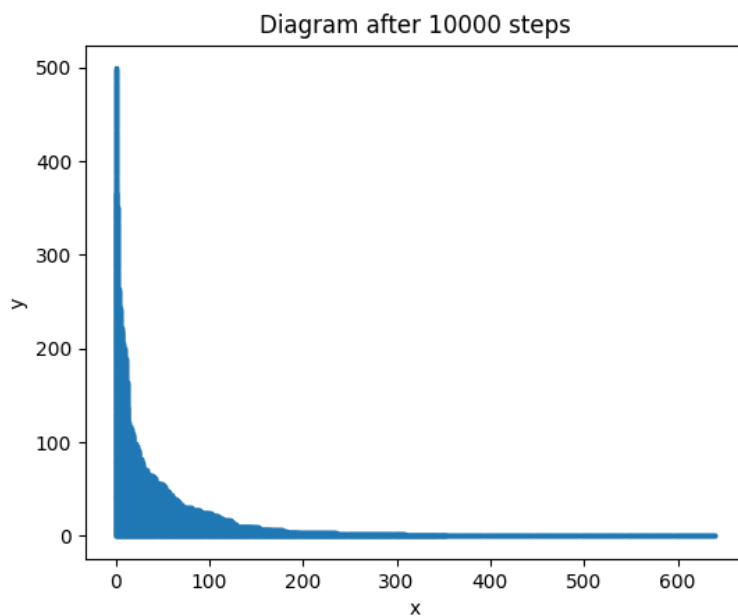


Рис. 2: Общая диаграмма при $\alpha = 0.8$

4.1.2. Диаграммы частоты

Следующие диаграммы частот получены путем повторения эксперимента k раз с изменением параметра α . После генерации k диаграмм с n ячейками они были наложены друг на друга. На представленных гистограммах показана частота появления определённой ячейки во время экспериментов. Цветовая гамма варьируется от красного цвета (обозначающего более высокую частоту) до зелёного (меньшая частота появления). Были зафиксированы параметры $n = 10000$ и $k = 10$.

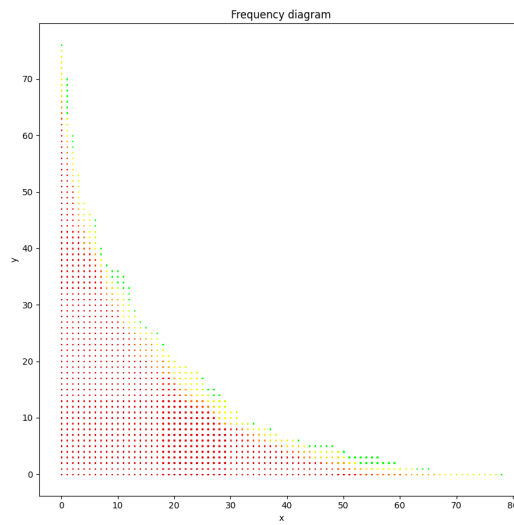


Рис. 3: Диаграмма частоты при $\alpha = 0$

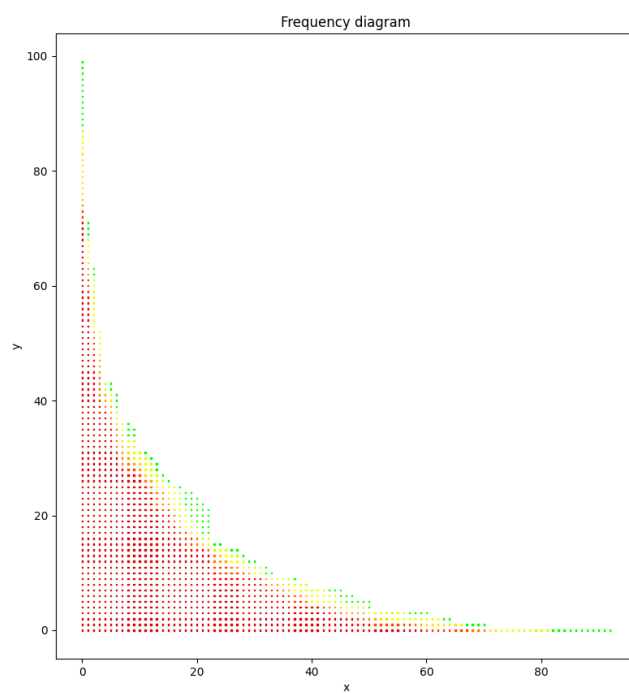


Рис. 4: Диаграмма частоты при $\alpha = 0.25$

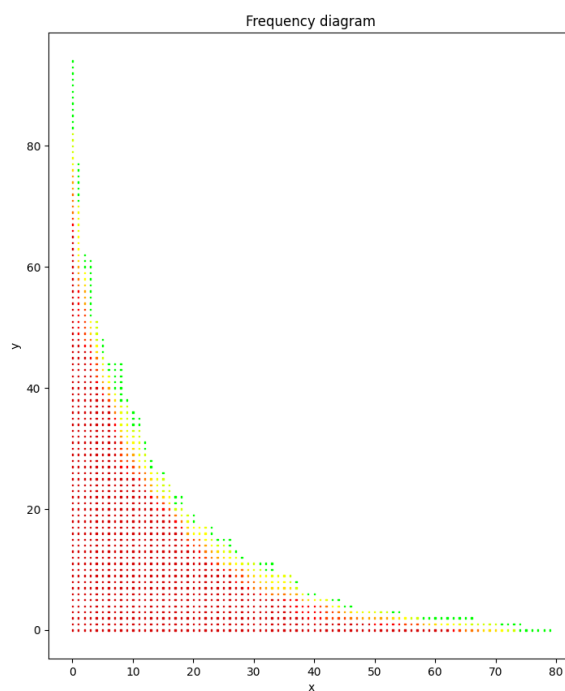
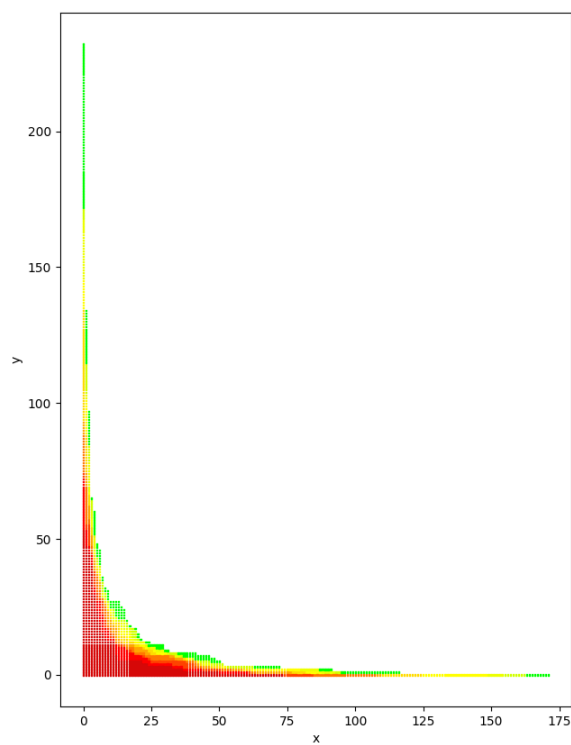
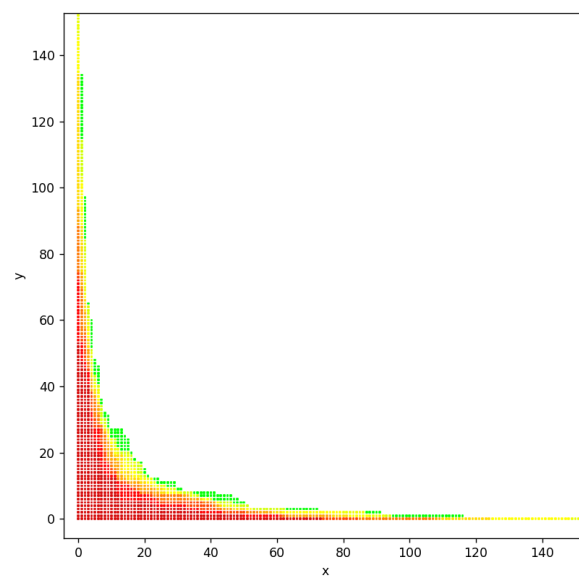


Рис. 5: Диаграмма частоты при $\alpha = 0.5$

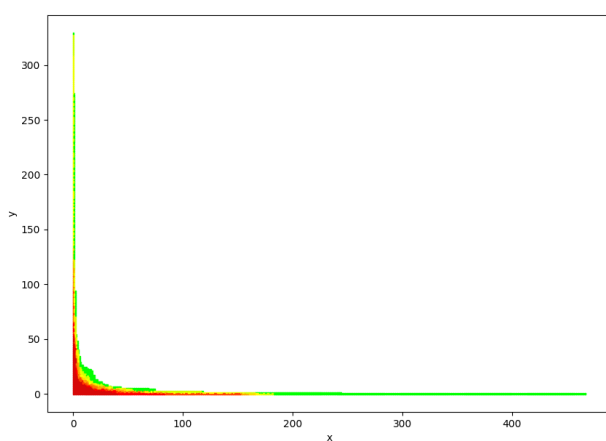


(a) Полная диаграмма

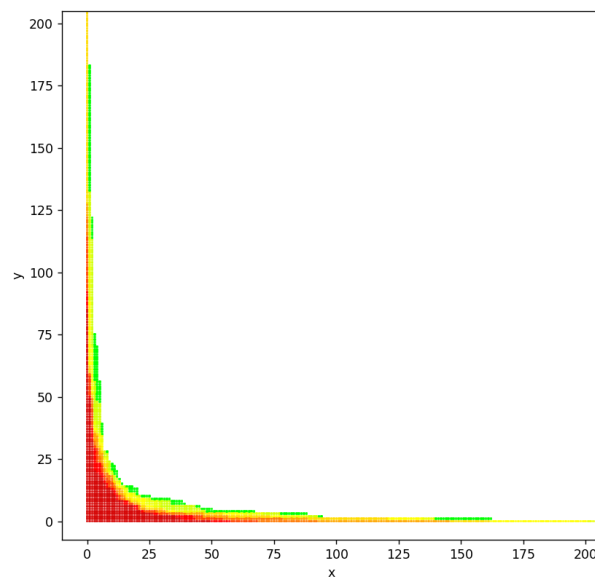


(b) Диаграмма ближе

Рис. 6: Диаграмма частоты при $\alpha = 0.75$

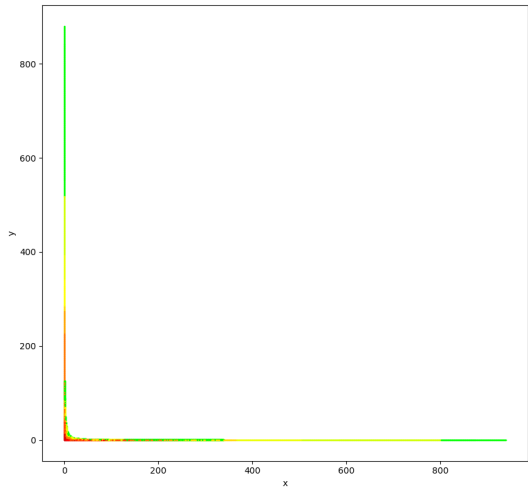


(a) Полная диаграмма

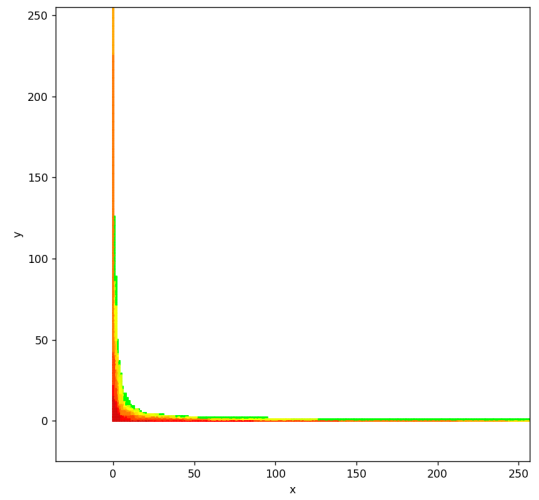


(b) Диаграмма ближе

Рис. 7: Диаграмма частоты при $k = 10$ и $\alpha = 1$



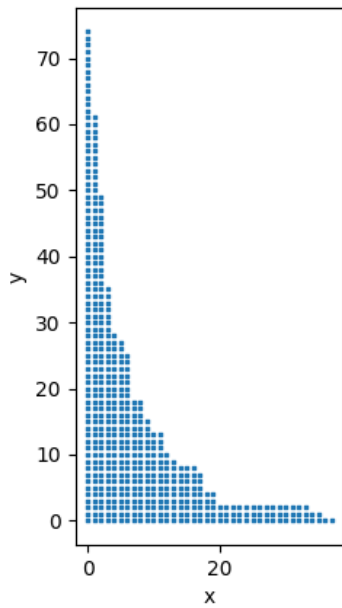
(a) Полная диаграмма



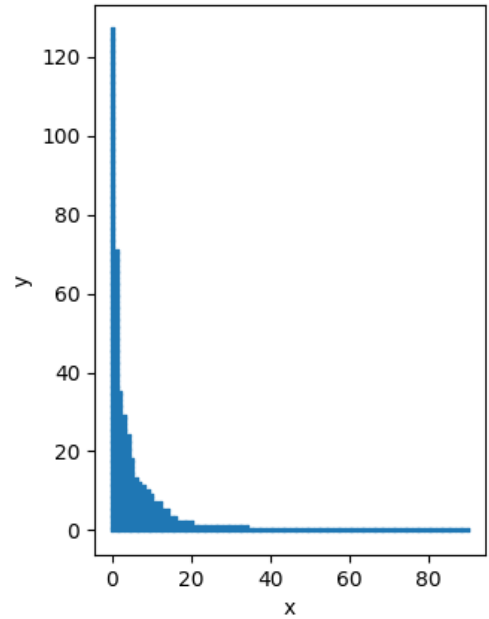
(b) Диаграмма ближе

Рис. 8: Диаграмма частоты $\alpha = 1.2$

4.1.3. Диаграмма на определенном шаге



(a) Диаграмма на 500-ом шаге



(b) Диаграмма на 1000-ом шаге

Рис. 9: Общая диаграмма при $n = 1000$ и $\alpha = 0.75$

4.2. Анализ

В ходе данной работы было продемонстрировано, как параметр α влияет на процесс стохастического роста диаграмм Юнга. Разработанная модель и визуализация результатов дают возможность исследовать предельные формы диаграмм и анализировать статистические свойства случайных процессов.

Визуализация общих диаграмм и диаграмм на определённом шаге показывает зависимости от параметра представлена визуализация диаграмм роста для различных значений параметра α и n . На Рис. 1 Рис. 2 можно видеть, как изменяются диаграммы по мере увеличения α . При $\alpha = 0.5$ диаграмма имеет определённую симметричную структуру, где клетки добавляются ближе к начальной точке. Однако, при $\alpha = 0.8$ виден большой разброс клеток, что указывает на влияние параметра на распределение и порядок добавления новых элементов. Это показывает, что увеличение α приводит к более равномерному распределению клеток по диаграмме. То же самое поведение можно наблюдать на Рис. 9, но в этот раз зависит от времени (шагов) и к чему стремится при $n \rightarrow \infty$.

Диаграммы частот, представленные на рисунках 3-8, демонстрируют, как часто клетки появляются в определённых позициях после нескольких симуляций. Чем выше α , тем более равномерно распределяются клетки по диаграмме. Например, при $\alpha = 0$ клетки сконцентрированы ближе к начальной точке, а при $\alpha = 1.2$ клетки появляются более равномерно по всей диаграмме, то есть площадь, где диаграммы часто встречаются, уменьшается. Это отражает то, как параметр α влияет на вероятность добавления клеток в различные области диаграммы: при больших значениях α диаграмма заполняется быстрее и более сбалансированно.

Заключение

По результатам выполнения научной работы были решены следующие задачи:

1. Проанализированы известные математические модели роста диаграммы Юнга.
2. Выбраны технологии и программированные библиотеки, наиболее подходящие для реализации работы.
3. Создана программа для визуализации разных характеристик прогресса роста диаграммы.
4. Проведен анализ полученных результатов.

Программы и выводы, полученные в данной работе, могут быть использованы в будущем исследовании по сравнению различных моделей роста, их изменений в зависимости от соответствующих параметров и количества ячеек.

Список литературы

- [1] Github repository. — <https://github.com/NicoAguirre1105/KP>. — Accessed: 06.03.2025.
- [2] Kimmo Eriksson Jonas Sjöstrand. Limiting shapes of birth-and-death processes on Young diagrams. — Advances in Applied Mathematics, 2011.
- [3] Rost H. Non-Equilibrium Behaviour of a Many Particle Process: Density Profile and Local Equilibria. — Zeitschrift für Wahrscheinlichkeitstheorie und verwandte Gebiete, 1981.
- [4] X. Gabaix Q. J. Econ. Zipf 's law for cities: An explanation. — The Quarterly Journal of Economics, 1999.
- [5] matplotlib. — <https://matplotlib.org/>. — Accessed: 06.03.2025.
- [6] Ю.Смирнов Е. Диаграммы Юнга, плоские разбиения и знакопередающиеся матрицы. — Издательство МЦНМО, 2013.