

Федеральное государственное бюджетное образовательное учреждение
высшего образования "Вологодский государственный университет"

На правах рукописи

Ковалёв Иван Александрович

**ПОЛУЧЕНИЕ ОЦЕНОК И ПОСТРОЕНИЕ ПРЕДЕЛЬНЫХ
ХАРАКТЕРИСТИК ДЛЯ НЕКОТОРЫХ СИСТЕМ МАССОВОГО
ОБСЛУЖИВАНИЯ С ОСОБЕННОСТЯМИ**

Специальность 1.2.3. Теоретическая информатика, кибернетика
(по физико-математическим наукам)

ДИССЕРТАЦИЯ

на соискание ученой степени кандидата
физико-математических наук

Научный руководитель,
доктор физико-математических
наук, профессор
Зейфман Александр Израилевич

Вологда – 2024

Оглавление

Введение	5
1 Основные понятия	13
1.1 Пространство l_1	13
1.2 Дифференциальные уравнения в пространстве l_1	15
1.3 Логарифмическая норма оператора	16
1.4 Марковские цепи	17
1.4.1 Основные понятия	17
1.4.2 Основные преобразования системы	20
1.4.3 Возмущенные процессы	22
2 Построение предельных характеристик, оценки эргодичности и устойчивости системы обслуживания типа $M_t/M_t/1$ с отказами, катастрофами, сбоями и ремонтами сервера	25
2.1 Описание модели	26
2.2 Оценки скорости сходимости	28
2.3 Оценки устойчивости	35
2.4 Примеры	38
2.4.1 Пример 1	38
2.4.2 Пример 2	44
3 Исследование модели с одним сервером, специальным групповым поступлением и со специальной политикой пропуска очереди	49
3.1 Описание модели	51

3.2	Получение оценок скорости сходимости	52
3.3	Оценки устойчивости	58
3.4	Получение границ для некоторых параметров модели	61
3.5	Примеры	64
3.5.1	Пример 1	65
3.5.2	Пример 2	70
4	Исследование модели с одним сервером, специальным группо- вым поступлением и со специальной политикой пропуска оче- реди с катастрофами	76
4.1	Описание модели	76
4.2	Получение оценок скорости сходимости	77
4.2.1	Получение оценок скорости сходимости в случае, когда интенсивности поступления групп требований экспоненци- ально убывают при увеличении размера группы.	78
4.2.2	Получение оценок скорости сходимости в случае, когда ин- тенсивности убывают со степенной скоростью.	81
4.2.3	Аппроксимация усечениями	83
4.3	Примеры	85
4.3.1	Пример 1	85
4.3.2	Пример 2	88
5	Исследование моделей, связанных с процессами рождения и гибели	95
5.1	Нестационарная марковская система с групповым поступлением и групповым обслуживанием требований с управлением, завися- щим от состояния	95
5.1.1	Описание модели	96
5.1.2	Оценки скорости сходимости	96
5.1.3	Оценки устойчивости	100
5.1.4	Пример	101

5.2	Оценки скорости сходимости и устойчивости для систем с нетерпеливыми клиентами	104
5.2.1	Описание модели	105
5.2.2	Скорость сходимости	105
5.2.3	Устойчивость	108
5.2.4	Пример	109
5.3	Исследование системы массового обслуживания с эластичным трафиком и нестационарной интенсивностью	114
5.3.1	Описание модели	114
5.3.2	Получение оценок скорости сходимости	117
5.3.3	Получение оценок устойчивости	121
5.3.4	Пример	123
	Заключение	130
	Приложение	132
	Описание программы	132
	Список литературы	133

Введение

Актуальность темы. Первоначальные исследования в области теории массового обслуживания были проведены А. К. Эрлангом почти столетие назад [43]. Однако до сих пор эта тема остается активно развивающимся разделом теории вероятностей, так как методы и модели массового обслуживания играют важную роль в исследовании телекоммуникационных систем, экономических и производственных процессов.

В этой области было проведено множество исследований, и российские и зарубежные ученые внесли большой вклад в ее развитие. Среди них следует отметить В.В. Анисимова, Л.Г. Афанасьеву, Г.П. Башарина, Ю.В. Гайдамаку, А.К. Горшенина, А.А. Назарова, А.Н. Моисеева, С.П. Моисееву, К. Е. Самуйлова, В. М. Вишневого, А.А. Боровкова, П.П. Бочарова, Р.Л. Добрушина, А.Н. Дудина, А.И. Зейфмана, В.В. Калашникова, Н.В. Карташова, В.Ю. Королева, Е.В. Морозова, А.В. Печинкина, В.В. Рыкова, О.В. Семенову, В.Г. Ушакова, С.Г. Фосса, E. Van Doorn, M. Neuts, R.L. Tweedie, W. Whitt и других (см. [1]–[10], [13]–[21], [24]–[67]).

Несмотря на это, вопросы свойств эргодичности и устойчивости для неоднородных марковских цепей с непрерывным временем и их применение к моделям массового обслуживания до сих пор остаются открытыми. Именно в этом направлении были проведены первоначальные исследования А.И. Зейфмана (см. [16, 68, 69]). Кроме того, задачи устойчивости стохастических моделей изучались также В.В. Анисимовым, А.Ю. Митрофановым, В.В. Калашниковым, Н.В. Карташовым (см. [2, 22, 49, 57]).

Интерес к исследованию нестационарных (неоднородных по времени) марковских цепей постоянно увеличивается, в связи с чем является актуальной задача получения оценок скорости сходимости, устойчивости и погрешности

аппроксимации для различных классов моделей, а также применение полученных оценок для построения основных предельных характеристик конкретных систем массового обслуживания.

Цель диссертационной работы. Целью работы является получение оценок вероятностных характеристик (скорости сходимости к предельному режиму, устойчивости) и построение предельных характеристик для некоторых систем массового обслуживания с особенностями.

Основные задачи. Для достижения заявленной цели решены следующие задачи:

1. Получены оценки скорости сходимости и устойчивости для систем массового обслуживания
 - типа $M_t/M_t/1$ с отказами, катастрофами, сбоями и ремонтами сервера;
 - с одним сервером, специальными групповыми поступлениями требований и специальной политикой пропуска очереди;
 - с одним сервером, специальными групповыми поступлениями требований, специальной политикой пропуска очереди и катастрофами;
 - с групповым поступлением и групповым обслуживанием требований с управлением, зависящим от состояния;
 - с нетерпеливыми клиентами;
 - с эластичным трафиком и нестационарной интенсивностью.
2. Для системы массового обслуживания с одним сервером, специальными групповыми поступлениями требований и специальной политикой пропуска очереди получены
 - оценки мощности сервера и мощности потока, при которой среднее число требований в системе не превышает заданного числа;
 - границы интенсивности обслуживания и интенсивности поступления требований, чтобы среднее оставалось в заданных границах.

Кроме того, для каждой из моделей с помощью разработанных алгоритмов и программ с применением численных методов проведены вычислительные эксперименты: построены предельные характеристики СМО, показаны зависимости реальных характеристик СМО от интенсивностей систем. Экспериментально проиллюстрированы свойства политики пропуска очереди.

Отметим, что перечисленные системы имеют широкое применение во многих областях: область сетевых систем связи, для моделирования сценария скачивания файла, в моделировании дорожного движения, бизнесе и отраслях промышленности, компьютерных коммуникациях, здравоохранении и медицинских науках, сервисных системах, розничных магазинах, а задачи, связанные с получением оценок скорости сходимости и устойчивости, решаются для них впервые.

Положения, выносимые на защиту.

1. Установление аналитических свойств новых систем массового обслуживания марковского типа, получение оценок скорости сходимости и устойчивости с помощью метода логарифмической нормы.
2. Нахождение зависимостей среднего числа требований в системе с одним сервером, специальными групповыми поступлениями требований и специальной политикой пропуска очереди от интенсивностей обслуживания и поступления требований.
3. Построение предельных характеристик для каждой из изученных систем массового обслуживания.

Научная новизна.

1. Для процессов, описывающих число требований в системе
 - типа $M_t/M_t/1$ с отказами, катастрофами, сбоями и ремонтами сервера;
 - с одним сервером, специальными групповыми поступлениями требований и специальной политикой пропуска очереди;
 - с одним сервером, специальными групповыми поступлениями требований, специальной политикой пропуска очереди и катастрофами;

- с групповым поступлением и групповым обслуживанием требований с управлением, зависящим от состояния;
- с нетерпеливыми клиентами;
- с эластичным трафиком и нестационарной интенсивностью.

получены новые оценки скорости сходимости к предельному режиму и предельному среднему, оценки устойчивости.

2. Для системы массового обслуживания с одним сервером, специальными групповыми поступлениями требований и специальной политикой пропуска очереди получены

- оценки мощности сервера и мощности потока, при которой среднее число требований в системе не превышает заданного числа;
- границы интенсивности обслуживания и интенсивности поступления требований, чтобы среднее оставалось в заданных границах.

Личное участие автора заключается в исследовании рассматриваемых моделей, получении новых оценок для них, а также разработке алгоритмов и программ для проведения вычислительных экспериментов.

Методы исследования. Для решения вышеописанных задач используется оператор Коши дифференциального уравнения в банаховом пространстве и оценки его нормы. Вопросы, связанные с вычислением требуемых параметров сводятся к изучению бесконечных систем дифференциальных уравнений на множестве стохастических векторов. Основным инструментом исследования и получения соответствующих оценок является метод, базирующийся на двух моментах: оценках, основанных на применении логарифмической нормы линейной операторной функции и специальных преобразованиях редуцированной матрицы интенсивностей марковской цепи. Для проведения вычислительных экспериментов используется программа на языке Java, которая решает задачу Коши методом Адамса–Мултона 4-го порядка.

Достоверность и обоснованность полученных результатов. Достоверность полученных результатов следует из строгих математических доказательств.

Теоретическая и практическая значимость. Полученные результаты могут быть использованы в исследовании конкретных систем линейных дифференциальных уравнений, стохастических моделей в технике, химии, биологии, физике и других отраслях. Описанные подходы могут быть применены в моделировании потоков информации, связанных с высокопроизводительными вычислениями.

Соответствие паспорту специальности. Диссертация выполнена в соответствии с паспортом специальности 1.2.3 – «Теоретическая информатика, кибернетика» и включает оригинальные результаты, направленные на развитие методов оценки и расчета вероятностных характеристик телекоммуникационных сетей. В соответствии с п. 9 «Математическая теория исследования операций» паспорта исследованы новые типы систем массового обслуживания, являющихся адекватными моделями реальных информационно-телекоммуникационных систем.

Содержание работы.

Во введении дается обоснование актуальности темы диссертации, приводится краткий обзор работ по данной тематике, сформулированы результаты, полученные в работе.

В главе 1 приводится вспомогательный математический аппарат, необходимый для дальнейшего исследования. Вводятся определения основных понятий: предельного среднего, пространства l_1 , логарифмической нормы оператора, а также некоторые понятия и методы, важные для дальнейшего исследования.

В следующих главах рассмотрено шесть новых моделей с использованием метода логарифмической нормы, получены оценки скорости сходимости, устойчивости и на основе этих оценок построены основные предельные характеристики. Для изучения каждой модели использован следующий алгоритм действий:

1. получение верхних оценок скорости сходимости, то есть нахождение момента времени t^* , начиная с которого вероятностные характеристики процесса $X(t)$ с заданной погрешностью не зависят от начальных условий;
2. получение оценок устойчивости для возмущенного процесса с близкими инфинитезимальными характеристиками;
3. в случае большой размерности исходного процесса (или счетного числа состояний) аппроксимации с помощью процессов меньшей размерности;

В каждом из примеров для проведения вычислительных экспериментов рассматривается соответствующая модель с 1-периодическими по времени интенсивностями:

- в случае большой размерности исходного процесса (или счетного числа состояний) выбираем размерность усеченного процесса N ;
- на основании полученных оценок определяем интервал на котором достигается желаемая точность $[0, t^* + 1]$;
- решаем прямую систему Колмогорова с простейшими начальными условиями $X(0) = 0$ и $X(0) = N$ для исходной (в случае необходимости, для усеченной) системы на отрезке $[0, t^* + 1]$;
- на отрезке $[t^*, t^* + 1]$ получаем с требуемой погрешностью все основные предельные характеристики как самого процесса $X(t)$, так и близких ему «возмущенных» процессов.

Заметим, что исследование каждой из этих моделей соответствует указанным разделам паспорта специальности.

В главе 2 рассмотрена система массового обслуживания типа $M_t/M_t/1$ с отказами, катастрофами, сбоями и ремонтами сервера.

В главе 3 рассматривается модель массового обслуживания с одним сервером, специальными групповыми поступлениями требований и специальной политикой пропуска очереди.

В главе 4 исследуется модель массового обслуживания с одним сервером, специальными групповыми поступлениями требований и специальной политикой пропуска очереди при наличии катастроф.

В §1 главы 5 рассмотрена система массового обслуживания с групповым поступлением и групповым обслуживанием требований с управлением, зависящим от состояния.

В §2 главы 5 рассмотрена система массового обслуживания с нетерпеливыми клиентами.

В §3 главы 5 рассмотрена система массового обслуживания с эластичным трафиком и нестационарной интенсивностью.

Для процессов, описывающих число требований в рассматриваемых системах получены оценки скорости сходимости к предельному режиму и предельному среднему, оценки устойчивости. Проведены вычислительные эксперименты. Для модели из главы 3 сформулированы утверждения об "управлении" интенсивностями поступления и обслуживания требований.

В заключении описаны и сформулированы основные результаты, полученные в ходе диссертационного исследования для рассмотренных моделей массового обслуживания.

В приложении приведено описание программы, с помощью которой выполняются построения основных характеристик марковского процесса.

Апробация результатов. Результаты работы докладывались на:

- семинарах кафедры прикладной математики ВоГУ "Современные методы стохастического моделирования сложных систем" (2020-2023),
- XXXVI Международном семинаре по проблемам устойчивости стохастических моделей (Петрозаводск, Россия, 2021),
- 19-й Международной конференции по численному анализу и прикладной математике (Родос, Греция, 2021),
- 20-й Международной конференции по численному анализу и прикладной математике (Родос, Греция, 2022).

Основные результаты опубликованы в [125]-[134], в том числе работы в журналах, рекомендованных ВАК. Программа для проведения вычислитель-

ных экспериментов [137] имеет свидетельство о государственной регистрации в Реестре программ для ЭВМ за номером 2020615415 от 22 мая 2020 года.

Глава 1

ОСНОВНЫЕ ПОНЯТИЯ

Глава является вспомогательной. Здесь вводится основной математический аппарат и необходимые для дальнейшего исследования понятия.

1.1 Пространство l_1

Рассмотрим множество всех абсолютно суммируемых последовательностей $x = \{x_1, x_2, \dots\}$, $x_i \in R$. Другими словами, должно выполняться условие $\sum_{i=1}^{\infty} |x_i| < \infty$.

Нормой называется величина $\|x\| = \sum_{i=1}^{\infty} |x_i|$. Множество всех таких последовательностей с введенной нормой называется пространством последовательностей l_1 . Векторами называются элементы этого пространства. l_1 является линейным пространством, полным относительно метрики $\rho(x, y) = \|x - y\|$ (т.е. банаховым). Векторное (линейное) пространство — математическая структура, которая формируется набором элементов - векторов, для которых определены операции сложения друг с другом и умножения на число — скаляр. Введённые операции подчинены восьми аксиомам. А скаляром может являться элемент вещественного, комплексного или любого другого поля чисел.

Банахово пространство — нормированное векторное пространство, полное по метрике, порождённой нормой. В свою очередь, полное пространство — это метрическое пространство, в котором каждая фундаментальная последовательность сходится к элементу этого же пространства. Единичные векторы (орты) пространства l_1 будем обозначать через e_i . Итак, e_i - вектор, у которого i -й

член соответствующей последовательности равен 1, а остальные – нули. Каждый вектор $x \in l_1$ представляем в виде $x = \sum_{i=1}^{\infty} a_i \cdot e_i$, причем $\sum_{i=1}^{\infty} |a_i| < \infty$. Рассмотрим отображение A из l_1 в себя. Тогда этот оператор однозначно определяется матрицей $(a_{ij})_{i,j=1}^{\infty}$. Норма оператора вычисляется по формуле:

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\| / \|x\| = \sup_{\|x\|=1} \|Ax\| = \sup_j \sum_i |a_{ij}|.$$

Рассматриваются только ограниченные операторы, то есть такие, для которых

$$\|A\| = \sup_j \sum_i |a_{ij}| < \infty.$$

Пусть каждому $t \geq 0$ ставится в соответствие вектор $\mathbf{x} \in l_1$. Тогда задана вектор-функция $\mathbf{x}(t)$. Вектор-функция называется непрерывной (в точке t_0), если при $t \rightarrow t_0$

$$\|\mathbf{x}(t) - \mathbf{x}(t_0)\| \rightarrow 0.$$

Дифференцируемость в точке и понятие интеграла от вектор-функции вводятся соответственно через предел отношения и интегральные суммы. Понятие оператор-функции, её непрерывности, дифференцируемости и интегрируемости вводятся аналогично.

Рассмотрим понятие показательной функции вида e^{tF}

$$e^{tF} = I + (tF) + (tF)^2/2! + (tF)^3/3! + \dots = \sum_{i=0}^{\infty} \frac{(tF)^i}{i!}.$$

Ряд в правой части сходится при любом t , что следует из сходимости следующего ряда

$$\sum_{i=0}^{\infty} \left\| \frac{(tF)^i}{i!} \right\| \leq \sum_{i=0}^{\infty} \frac{t^i \|F\|^i}{i!} = e^{t\|F\|}.$$

Получаем, что

$$\|e^{tF}\| \leq e^{t\|F\|}.$$

Таким образом, при любых действительных t, s справедливо

$$e^{(t+s)F} = e^{tF} \cdot e^{sF}.$$

Откуда при $s = -t$ вытекает, что оператор e^{tF} обратим при любом t .

1.2 Дифференциальные уравнения в пространстве l_1

Рассмотрим дифференциальное уравнение в пространстве последовательностей l_1

$$\frac{dy}{dt} = A(t)y(t) + f(t) \quad (1.2.1)$$

и соответствующее однородное уравнение

$$\frac{dx}{dt} = A(t)x(t),$$

где $x(t)$, $y(t)$, $f(t)$ - вектор-функции из R^+ в l_1 , а $A(t)$ - оператор l_1 в l_1 .

Назовем $U(t, \tau)$ - оператором Коши дифференциальных уравнений, где

$$U(t, \tau) = I + \int_{\tau}^t A(s_1)ds_1 + \int_{\tau}^t A(s_1) \int_{\tau}^{s_1} A(s_2)ds_2ds_1 + \dots,$$

причем ряд в правой части сходится равномерно на любом конечном отрезке и

$$U(t, s) = U(t, \tau)U(\tau, s).$$

Теорема 1. Пусть $A(t)$, $f(t)$ - непрерывны, $\tau \geq 0$ и $y^* \in l_1$. Тогда существует единственная $y(t)$, определенная на $[\tau, \infty)$, такая, что:

- 1) $y(\tau) = y^*$;
- 2) $y(t)$ непрерывна и дифференцируема при всех $t \geq \tau$.

Теорема 2. Пусть $A(t)$, $f(t)$ - непрерывны, $\tau \geq 0$ и x^* , $y^* \in l_1$. Тогда существуют единственные $x(t)$, $y(t)$, определенные на $[\tau, \infty)$, такие, что:

$$x(\tau) = x^*, \quad y(\tau) = y^*,$$

$$x(t) = U(t, \tau)x(\tau),$$

$$y(t) = U(t, \tau)y(\tau) + \int_{\tau}^t U(t, s)f(s)ds. \quad (1.2.2)$$

1.3 Логарифмическая норма оператора

Понятие логарифмической нормы для конечных матриц было введено и, в дальнейшем, изучено Лозинским (см. [23]). А также обобщено на случай оператор-функций в [12]. Рассмотрим само понятие и важные оценки, связанные с ним.

Определение 1. Логарифмической нормой $\gamma(A(t))$ оператора A называется число

$$\gamma(A(t)) = \lim_{h \rightarrow +0} \frac{\|U(t+h, t)\| - \|U(t, t)\|}{h}.$$

Кроме того, справедливо следующее утверждение.

Теорема 3. При всех $t > 0$ существует $\gamma(A(t))$, причем

$$\gamma(A(t)) = \lim_{h \rightarrow +0} \frac{\|I + hA(t)\| - 1}{h}. \quad (1.3.3)$$

Следствие 1.

$$\gamma(A(t)) = \sup_j \left(a_{jj} + \sum_{i \neq j} |a_{ji}(t)| \right).$$

Теорема 4. Для любых t, s ($t \geq s \geq 0$) выполняется:

$$e^{-\int_s^t \gamma(-A(\tau))d\tau} \leq \|U(t, s)\| \leq e^{\int_s^t \gamma(A(\tau))d\tau}. \quad (1.3.4)$$

Тогда

$$\frac{d}{dt} \|x(t)\| \leq \gamma(A(t)) \|x(t)\|. \quad (1.3.5)$$

Таким образом, выполняется следующая оценка

$$\|x(t)\| \leq e^{\int_0^t \gamma(A(u))du} \|x(0)\|. \quad (1.3.6)$$

Теорема 5. Свойство неотрицательности: $U(t, s) \geq 0$ при всех $t \geq s \geq 0$ - равносильно тому, что $a_{ij}(u) \geq 0$ при всех i, j таких, что $i \neq j$, и любом $u \geq 0$.

Рассмотрим случай подпространства в l_1 . Пусть матрица D образована элементами последовательности $\{d_i\}$ и $d = \inf_{i \geq 0} d_i > 0$.

Пусть l_{1D} - пространство последовательностей $\mathbf{z} = (p_0, p_1, p_2 \dots)$ таких, что $\|\mathbf{z}\|_{1D} = \|D\mathbf{z}\|_1 < \infty$.

Теорема 6.

Пусть $B : l_1 \rightarrow l_1$ - линейный оператор. Пусть B действует на векторы из l_{1D} , тогда

$$\|B\|_{l_{1D}} = \|DBD^{-1}\|_{l_1}, \quad (1.3.7)$$

$$\gamma(B)_{l_{1D}} = \gamma(DBD^{-1})_{l_1}. \quad (1.3.8)$$

1.4 Марковские цепи

1.4.1 Основные понятия

Рассмотрим систему S , способную в момент времени t находиться в одном из состояний с номерами $0, 1, \dots, N$. Множество $E_N = \{0, 1, \dots, N\}$ называется пространством состояний стохастической системы S . Через $X(t)$ обозначим состояние системы в момент времени t и предположим, что если $X(t) = i$, то при $h > 0$ $X(t+h) = j$ с вероятностью

$$\begin{cases} q_{ij}(t)h + o_{ij}(h), & j \neq i, \\ 1 - \sum_{k \neq i} q_{ik}(t)h + o_i(h), & j = i, \end{cases} \quad (1.4.9)$$

где все $o_i(h)$ равномерны по i , то есть $\sup_i |o_i(h)| = o(h)$.

Данное условие является жестким. Будем рассматривать только те процессы, которые удовлетворяют этому условию. Данные процессы будем называть марковскими цепями с непрерывным временем и счетным пространством состояний. Функция $q_{ij}(t)$ - интенсивность перехода из состояния i в состояние j . Марковская цепь $X(t)$ называется стационарной, если все $q_{ij}(t) = q_{ij}$ и нестационарной - в противном случае. Будем использовать термины "стационарный" и "нестационарный".

Положим $q_{ii}(t) = -\sum_{k \neq i} q_{ik}(t)$ и назовем матрицу $Q(t) = (q_{ij}(t))_{i,j=0}^{\infty}$ матрицей интенсивностей для марковской цепи $X(t)$. Введем в рассмотрение переходные вероятности $p_{ij}(t, s) = Pr(X(t) = j | X(s) = i)$, вероятности состояний $p_i(t) = Pr(X(t) = i)$ и вектор-столбец вероятностей состояний $\mathbf{p}(t) = (p_0(t), p_1(t) \dots)^T$. Положим $a_{ij}(t) = q_{ji}(t)$ и рассмотрим матрицу $A(t) = (a_{ij}(t))_{i,j=0}^{\infty} = Q^T(t)$. Тогда получим

$$\mathbf{p}(t+h) = \mathbf{p}(t) + Ah\mathbf{p}(t) + o(h), \quad (1.4.10)$$

откуда вытекает прямая система Колмогорова в виде дифференциального уравнения:

$$\frac{d}{dt}\mathbf{p}(t) = A(t)\mathbf{p}. \quad (1.4.11)$$

Пусть $U(t, s)$ - оператор Коши данного дифференциального уравнения, тогда $P(s, t) = U^T(t, s) = (p_{ij}(s, t))_{i,j=0}^{\infty}$ называется матрицей перехода $X(t)$. Обозначим через Ω множество всех стохастических векторов $\mathbf{x} = (x_0, x_1 \dots)^T \in \Omega$, то есть $\mathbf{x} \geq 0$ и $\|\mathbf{x}\| = 1$.

Теорема 7.

- (i) При каждом $s \geq 0$, $t \geq s$ и любом $\mathbf{p} \in l_1$ существует единственное $\mathbf{p}(t)$ такое, что $\mathbf{p}(s) = \mathbf{p}$. При этом $\mathbf{p}(t) = U(t, s)\mathbf{p}(s)$;
- (ii) если $\mathbf{p}(s) \in \Omega$, то и $\mathbf{p}(t) \in \Omega$ при $t \geq s$;
- (iii) уравнение 1.4.11 устойчиво, а $\|\mathbf{p}^1(t) - \mathbf{p}^2(t)\|$ монотонно не возрастает при любых начальных условиях, где $\mathbf{p}^1(t)$, $\mathbf{p}^2(t)$ являются решениями, соответствующими начальным условиям $\mathbf{p}^1(s)$, $\mathbf{p}^2(s)$.

Определение 2. Матрица $H = (h_{ij})_0^{\infty}$ стохастическая, если все ее элементы неотрицательны, а сумма элементов каждого столбца равна единице.

Определение 3. Для любых $s \geq 0$, $t \geq s$ матрица Коши $U(t, s)$ является стохастической.

Определение 4. Марковская цепь $X(t)$ называется нуль-эргодичной, если при любом $i \lim_{t \rightarrow \infty} p_i(t) = 0$.

Определение 5. Марковская цепь $X(t)$ называется слабо эргодичной, если $\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \rightarrow 0$ при $t \rightarrow \infty$ для любых начальных условий $\mathbf{p}^*(s), \mathbf{p}^{**}(s)$ и любом $s \geq 0$.

Определение 6. Марковскую цепь $X(t)$ назовем сильно эргодичной, если существует вектор $\pi \in \Omega$ такой, что $\lim_{t \rightarrow \infty} \|\mathbf{p}(t) - \pi\| = 0$ при любом $\mathbf{p}(0) = \mathbf{p} \in \Omega$. При этом вектор π называется стационарным распределением марковской цепи $X(t)$.

Обозначим через $E(t, k) = E\{X(t) | X(s) = k\}$ математическое ожидание процесса (среднее число требований) в момент времени t при условии, что в момент s он находится в состоянии k . Кроме того, введем более общее обозначение $E_{\mathbf{p}}(t)$ - математическое ожидание процесса в момент времени t при начальном распределении вероятностей состояний $\mathbf{p}(0) = \mathbf{p}$. Кроме того, если положим $E_k(t) = E\{X(t) | X(0) = k\}$, тогда соответствующее начальное условие системы (1.4.11) - это k -й единичный вектор \mathbf{e}_k .

Определение 7. Пусть $X(t)$ - марковский процесс. Тогда $\varphi(t)$ - предельное среднее процесса $X(t)$, если

$$\lim_{t \rightarrow \infty} (\varphi(t) - E_k(t)) = 0$$

для любого k .

Определение 8. Если предел

$$\mathbf{E} = \lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t E\{X(u) | X(0) = k\} du \quad (1.4.12)$$

существует и не зависит от k , то \mathbf{E} - двойное среднее для цепи $X(t)$.

Предельное среднее процесса показывает среднее количество требований в момент времени (при достаточно больших t). При этом начальное состояние системы не оказывает влияние на предельное среднее. Двойное среднее - некоторая средняя характеристика системы на всем промежутке ее существования.

1.4.2 Основные преобразования системы

Рассмотрим основные преобразования системы

1. Исключение нулевого состояния

Используя свойство $\mathbf{p}(t) \in \Omega$, положим

$$p_0(t) = 1 - \sum_{i \geq 1} p_i(t),$$

тогда получаем систему

$$\frac{d}{dt} \mathbf{z}(t) = B(t) \mathbf{z}(t) + \mathbf{f}(t), \quad (1.4.13)$$

где

$$B(t) = \begin{pmatrix} a_{11}(t) - a_{10}(t) & a_{12}(t) - a_{10}(t) & a_{13}(t) - a_{10}(t) & \cdots \\ a_{21}(t) - a_{20}(t) & a_{22}(t) - a_{20}(t) & a_{23}(t) - a_{20}(t) & \cdots \\ a_{31}(t) - a_{30}(t) & a_{32}(t) - a_{30}(t) & a_{33}(t) - a_{N0}(t) & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix}, \quad (1.4.14)$$

$$\mathbf{z} = (p_1(t), p_2(t), p_3(t), \dots)^T, \quad \mathbf{f}(t) = (a_{10}(t), a_{20}(t), a_{30}(t), \dots)^T \quad (1.4.15)$$

2. Вычитание положительного элемента

Положим $\gamma^*(t) = \inf_n a_{0n}(t) > 0$ и перепишем прямую систему Колмогорова как

$$\frac{d}{dt} \mathbf{p}(t) = A^*(t) \mathbf{p}(t) + \mathbf{g}(t), \quad t \geq 0, \quad (1.4.16)$$

где $\mathbf{g}(t) = (\gamma^*(t), 0, 0, \dots)^T$, $A^*(t) = (a_{ij}^*(t))_{i,j \geq 0}$ и

$$a_{ij}^*(t) = \begin{cases} a_{0j}(t) - \gamma^*(t), & \text{если } i = 0, \\ a_{ij}(t), & \text{если } i > 0. \end{cases} \quad (1.4.17)$$

Теперь можно исследовать более удобную «редуцированную» систему (1.4.13) или (1.4.16).

3. Треугольное преобразование

Рассмотрим следующее вспомогательное преобразование. Положим

$$t_{ij} = \begin{cases} 1 & \text{при } j \geq i \\ 0 & \text{при } j < i \end{cases}$$

и

$$T = (t_{ij})_{i,j=1}^{\infty}. \quad (1.4.18)$$

То есть T верхняя треугольная матрица вида

$$T = \begin{pmatrix} 1 & 1 & 1 & 1 & \cdots \\ 0 & 1 & 1 & 1 & \cdots \\ 0 & 0 & 1 & 1 & \cdots \\ 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (1.4.19)$$

Тогда

$$T^{-1} = \begin{pmatrix} 1 & -1 & 0 & 0 & 0 & \cdots \\ 0 & 1 & -1 & 0 & 0 & \cdots \\ 0 & 0 & 1 & -1 & 0 & \cdots \\ 0 & 0 & 0 & 1 & -1 & \cdots \\ 0 & 0 & 0 & 0 & 1 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \end{pmatrix}. \quad (1.4.20)$$

4. Диагональное преобразование

Пусть теперь $1 \leq d_1 \leq d_2 \leq \dots$, $i = 1, 2, \dots$ и $D = \text{diag}(d_1, d_2, \dots)$ диагональная матрица:

$$D = \begin{pmatrix} d_1 & 0 & 0 & 0 & \cdots \\ 0 & d_2 & 0 & 0 & \cdots \\ 0 & 0 & d_3 & 0 & \cdots \\ 0 & 0 & 0 & d_4 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (1.4.21)$$

Тогда

$$D^{-1} = \begin{pmatrix} \frac{1}{d_1} & 0 & 0 & 0 & \dots \\ 0 & \frac{1}{d_2} & 0 & 0 & \dots \\ 0 & 0 & \frac{1}{d_3} & 0 & \dots \\ 0 & 0 & 0 & \frac{1}{d_4} & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (1.4.22)$$

Использование треугольной и диагональной матриц позволяет в некоторых случаях привести матрицу к существенно-неотрицательному виду.

1.4.3 Возмущенные процессы

Пусть процесс $X(t), t \geq 0$, является неоднородной непрерывной марковской цепью с матрицей интенсивностей $Q(t)$ и транспонированной матрицей интенсивностей $A(t) = Q^T(t)$.

Определение 9. Процесс $X(t)$ назовем экспоненциально эргодичным, если найдутся такие константы $b > 0$, $c > 1$, что для всех s и t ($0 \leq s \leq t$) выполнено неравенство

$$\| \mathbf{p}^*(t) - \mathbf{p}^{**}(t) \| \leq c \cdot e^{-b \cdot (t-s)}, \quad (1.4.23)$$

где $\mathbf{p}^*(t)$, $\mathbf{p}^{**}(t)$ – решения системы (1.4.11).

Пусть $\bar{X}(t), t \geq 0$ – возмущенный процесс с инфинитезимальной матрицей $\bar{Q}(t)$ и соответствующей транспонированной матрицей $\bar{A}(t)$.

Обозначим символом $\hat{A}(t)$ разность матриц $A(t) - \bar{A}(t)$ и назовем матрицей возмущений, а условие

$$\| \hat{A}(t) \| \leq \varepsilon, \quad (1.4.24)$$

выполненное почти для всех t , будем трактовать как условие малости возмущений матрицы интенсивностей.

Определение 10. Марковская цепь $X(t)$ называется устойчивой, если для любого $\varepsilon > 0$ найдется $\delta > 0$ такое, что из условия $\sup_{t \geq 0} \| \hat{A}(t) \| < \delta$ следует неравенство $\| \mathbf{p}(t) - \bar{\mathbf{p}}(t) \| < \varepsilon$ для всех $\mathbf{p}(0) = \bar{\mathbf{p}}(0) = \mathbf{p} \in \Omega$.

Теорема 8. Если марковская цепь такова, что выполнены условия (1.4.23) и (1.4.24), то для любых начальных условий $\mathbf{p}(s)$, $\bar{\mathbf{p}}(s)$ справедлива оценка

$$\|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\| \leq \begin{cases} \|\mathbf{p}(s) - \bar{\mathbf{p}}(s)\| + (t-s) \cdot \varepsilon, & 0 < t-s < b^{-1} \ln \frac{c}{2} \\ \frac{c}{2} \cdot e^{-b(t-s)} \|\mathbf{p}(s) - \bar{\mathbf{p}}(s)\| + \varepsilon \cdot b^{-1} (\ln \frac{c}{2} + 1 - c \cdot e^{-b(t-s)}), & t-s \geq b^{-1} \ln \frac{c}{2}. \end{cases} \quad (1.4.25)$$

Также имеет место оценка устойчивости решения

$$\overline{\lim}_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\| \leq b^{-1} (\ln \frac{c}{2} + 1) \cdot \varepsilon. \quad (1.4.26)$$

Если, кроме того, пространство состояний конечно $\{0, 1, \dots, N\}$, то оценка устойчивости для математического ожидания имеет вид

$$\overline{\lim}_{t \rightarrow \infty} |E_0(t) - \bar{E}_0(t)| \leq \frac{N}{b} (\ln \frac{c}{2} + 1) \cdot \varepsilon. \quad (1.4.27)$$

Рассмотрим пространство

$$l_{1D} = \left\{ z = (p_1, p_2, \dots)^T : \|z\|_{1D} \equiv \|Dz\| < \infty \right\}.$$

Здесь нормы определены следующим образом:

$$\|A\|_{1D} = \|D \cdot A \cdot D^{-1}\|, \quad \|\mathbf{p}\|_{1D} = \|z\|_{1D}.$$

И пусть для системы (1.4.13)

$$\|B(t)\|_{1D} \leq \mathfrak{B} < \infty, \quad \|f(t)\|_{1D} \leq \mathfrak{f} < \infty$$

и соответствующей возмущенной системы

$$\|\bar{B}(t)\|_{1D} \leq \bar{\mathfrak{B}} < \infty, \quad \|\bar{f}(t)\|_{1D} \leq \bar{\mathfrak{f}} < \infty$$

почти для всех t .

Определение 11. Процесс $X(t)$ назовем $1D$ -экспоненциально эргодичным, если найдутся такие константы $M > 0$, $a > 0$, что для всех s и t ($0 \leq s \leq t$) выполнено неравенство

$$\| \mathbf{p}^*(t) - \mathbf{p}^{**}(t) \|_{1D} \leq M \cdot e^{-a \cdot (t-s)} \cdot \| \mathbf{p}^*(s) - \mathbf{p}^{**}(s) \|_{1D} \quad (1.4.28)$$

для любых начальных условий $\mathbf{p}^*(s), \mathbf{p}^{**}(s) \in l_{1D}$.

Теорема 9. Если марковские цепи $X(t), \bar{X}(t)$ $1D$ -экспоненциально эргодичны, то выполнено следующее неравенство

$$\overline{\lim}_{t \rightarrow \infty} \| \mathbf{p}(t) - \bar{\mathbf{p}}(t) \|_{1D} \leq \frac{M \cdot (M |\mathfrak{B} - \bar{\mathfrak{B}}| \cdot \mathfrak{f} + a |\mathfrak{f} - \bar{\mathfrak{f}}|)}{a \cdot (a - M |\mathfrak{B} - \bar{\mathfrak{B}}|)}. \quad (1.4.29)$$

Кроме того, оценка устойчивости для математического ожидания имеет вид

$$\overline{\lim}_{t \rightarrow \infty} |E_0(t) - \bar{E}_0(t)| \leq \frac{M \cdot (M |\mathfrak{B} - \bar{\mathfrak{B}}| \cdot \mathfrak{f} + a |\mathfrak{f} - \bar{\mathfrak{f}}|)}{a \cdot W \cdot (a - M |\mathfrak{B} - \bar{\mathfrak{B}}|)}, \quad (1.4.30)$$

где $W = \inf \{ \frac{d_k}{k} \}$.

Глава 2

Построение предельных характеристик, оценки эргодичности и устойчивости системы обслуживания типа $M_t/M_t/1$ с отказами, катастрофами, сбоями и ремонтами сервера

Изучаемая система имеет широкое применение во многих областях, одной из которых является область сетевых систем связи. Применимость этой модели можно увидеть в системах коммуникационных сетей. Если в системе выстроено множество пакетов, локальные пакеты всегда принимаются, а удаленные пакеты представляют собой пакеты с пороговым значением, ожидающие обработки в узле. Затем вновь прибывший либо решает не присоединяться к системе, либо уходит после присоединения к системе. Если эта сеть была заражена вирусом, это может привести к потере некоторых пакетов в результате повторного перезапуска сети или передачи этих пакетов в другую сеть. Кроме того, в компьютерных системах, где в системе есть несколько клиентов (данных), выстроенных в очередь до определенного порогового значения, вновь прибывший может решить не входить в систему после этого значения. В случае заражения вирусом данные будут уничтожены или переданы другим процессорам. Такие системы можно описать как модели массового обслуживания с катастрофами и отказами. Эти системы могут быть представлены в виде предлагаемой модели массового обслуживания.

Большая часть литературы по этой теме посвящена исключительно изучению частного случая стационарного поведения. В [70] автор обсуждал стационарное поведение двухпроцессорной гетерогенной системы с катастрофами, отказами серверов и ремонтами. В [77] проанализировано стационарное поведение очереди $M/M/1$ с катастрофами, сбоями и ремонтами сервера, а в [80] расширены результаты для очереди $M/M/1$ с блокировкой, катастрофами, сбоями и ремонтами сервера, где блокировка происходит тогда и только тогда, когда размер системы равен или превышает пороговое значение k . В [79] получены явные выражения для стационарных вероятностей очереди $M/M/1$ с отказами, катастрофами, сбоями и ремонтами сервера. В [72] изучено стационарное поведение двусторонней очереди с катастрофами и ремонтами.

С другой стороны, в [71] исследована нестационарность двухпроцессорной гетерогенной системы с катастрофами, отказами серверов и ремонтами. В [73] описана неоднородность по времени для двусторонней очереди, подверженной катастрофам и ремонтам (продолжение предыдущей работы [72]). Некоторые другие нестационарные модели были изучены рядом авторов, см., например, [74, 75, 76, 54, 78, 125].

В данной главе рассмотрена система массового обслуживания типа $M_t/M_t/1$ с отказами, катастрофами, сбоями и ремонтами сервера в случае нестационарного поведения. Получены оценки скорости сходимости и устойчивости для соответствующего процесса длины очереди. Такие оценки дают возможность найти предельные оценки для класса близких к этой очереди марковских моделей. В случае постоянных параметров результаты согласуются с результатами, полученными [80].

Результаты главы опубликованы в [123].

2.1 Описание модели

Рассмотрим систему массового обслуживания с одним сервером с отказами, катастрофами, сбоями и возможными ремонтами сервера. Обозначим интенсивности: поступления требований $\lambda(t)$, обслуживания требований $\mu(t)$, катастрофы $\gamma_i(t)$, восстановления вышедшего из строя сервера $\eta(t)$. При поступ-

лении требования, если впереди него есть k (пороговое значение) или больше требований, то оно присоединяется к очереди с вероятностью $\beta(t)$ и может отказаться с вероятностью $1 - \beta(t)$, $r(t)$ - вероятность того, что сервер находится в ремонте в момент времени t с $r(0) = 0$.

Данная модель может быть описана марковским процессом $X(t), t > 0$, где $X(t)$ обозначает количество требований в системе в момент времени t (процесс длины очереди). Обозначим через $p_n(t) = P(X(t) = n)$, $n = 0, 1, 2, 3, \dots$

Исходя из вышеизложенных предположений результирующее поведение вероятностей состояний описывается прямой системой Колмогорова:

$$r'(t) = -\eta(t)r(t) + \sum_{i=0}^{\infty} \gamma_i(t)p_i(t) \quad (2.1.1)$$

$$p'_0(t) = \mu(t)p_1(t) - (\lambda(t) + \gamma_0(t))p_0(t) + \eta(t)r(t) \quad (2.1.2)$$

$$p'_n(t) = \lambda(t)p_{n-1}(t) - (\lambda(t) + \gamma_n(t) + \mu(t))p_n(t) + \mu(t)p_{n+1}(t), 1 \leq n \leq k-1 \quad (2.1.3)$$

$$p'_k(t) = \lambda(t)p_{k-1}(t) - (\lambda(t)\beta(t) + \gamma_k(t) + \mu(t))p_k(t) + \mu(t)p_{k+1}(t), n = k \quad (2.1.4)$$

$$p'_n(t) = \lambda(t)\beta(t)p_{n-1}(t) - (\lambda(t)\beta(t) + \gamma_n(t) + \mu(t))p_n(t) + \mu(t)p_{n+1}(t), n > k. \quad (2.1.5)$$

Теперь рассмотрим соответствующую нестационарную ситуацию. А именно, предполагаем, что процесс длины очереди $\{X(t), t \geq 0\}$ является неоднородной Марковской цепью непрерывного времени. Предполагаем, что все возможные интенсивности переходов, скажем $q_{ij}(t)$, являются неслучайными функциями времени и все функции интенсивности неотрицательны и локально интегрируемы на $[0, \infty)$.

Обозначим через $\mathbf{p}(t) = (r(t), p_0(t), p_1(t), \dots)^T$ вектор вероятностей состояний в данный момент времени t . Положим $a_{ij}(t) = q_{ji}(t)$ для $j \neq i$ и $a_{ii}(t) = -\sum_{j \neq i} a_{ji}(t) = -\sum_{j \neq i} q_{ij}(t)$.

Предполагаем, что матрица интенсивности существенно ограничена, т.е.

$$\sup_i |a_{ii}(t)| \leq L < \infty \quad (2.1.6)$$

почти для всех $t \geq 0$. Тогда имеем $\|A(t)\| = 2 \sup_k |a_{kk}(t)| \leq 2L$ почти для всех $t \geq 0$.

Можем рассматривать прямую систему Колмогорова (2.1.1)–(2.1.5) как дифференциальное уравнение (1.4.11) в пространстве последовательностей l_1 , где $A(t)$ является ограниченным почти для всех $t \geq 0$ линейным оператором в l_1 и задается соответствующей транспонированной матрицей интенсивности:

$$\left(\begin{array}{cccccccc} -\eta(t) & \gamma_0(t) & \gamma_1(t) & \cdots & \gamma_{k-1}(t) & \gamma_k(t) & \gamma_{k+1}(t) & \gamma_{k+2}(t) & \cdots \\ \eta(t) & -(\lambda(t) + \gamma_0(t)) & \mu(t) & \cdots & 0 & 0 & 0 & 0 & \cdots \\ 0 & \lambda(t) & -(\lambda(t) + \gamma_1(t) + \mu(t)) & \cdots & 0 & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda(t) & -(\lambda(t)\beta(t) + \gamma_k(t) + \mu(t)) & \mu(t) & 0 & \cdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda(t)\beta(t) & -(\lambda(t)\beta(t) + \gamma_{k+1}(t) + \mu(t)) & \mu(t) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right) \quad (2.1.7)$$

2.2 Оценки скорости сходимости

Метод получения оценок скорости сходимости основан на понятии логарифмической нормы и соответствующих оценках для оператора Коши. Кроме того, для рассматриваемой модели рассматриваем два подхода [71, 85]. Опишем эти подходы.

Первый подход, см. [71]. Положим $\gamma^*(t) = \inf_n \gamma_n(t)$ и перепишем прямую систему Колмогорова (1.4.11) как

$$\frac{d\mathbf{p}}{dt} = A^*(t) \mathbf{p} + \mathbf{g}(t), \quad t \geq 0, \quad (2.2.8)$$

где $\mathbf{g}(t) = (\gamma^*(t), 0, 0, \dots)^T$, $A^*(t) = (a_{ij}^*(t))_{i,j=0}^\infty$, и

$$a_{ij}^*(t) = \begin{cases} a_{0j}(t) - \gamma^*(t), & \text{если } i = 0, \\ a_{ij}(t), & \text{если } i > 0. \end{cases} \quad (2.2.9)$$

Тогда

$$A^*(t) =$$

$$\left(\begin{array}{cccccccc} -(\eta(t) + \gamma^*(t)) & \gamma_0(t) - \gamma^*(t) & \gamma_1(t) - \gamma^*(t) & \cdots & \gamma_{k-1}(t) - \gamma^*(t) & \gamma_k(t) - \gamma^*(t) & \gamma_{k+1}(t) - \gamma^*(t) & \cdots \\ \eta(t) & -(\lambda(t) + \gamma_0(t)) & \mu(t) & \cdots & 0 & 0 & 0 & \cdots \\ 0 & \lambda(t) & -(\lambda(t) + \gamma_1(t) + \mu(t)) & \cdots & 0 & 0 & 0 & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots \\ 0 & 0 & 0 & \cdots & \lambda(t) & -(\lambda(t)\beta(t) + \gamma_k(t) + \mu(t)) & \mu(t) & \cdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda(t)\beta(t) & -(\lambda(t)\beta(t) + \gamma_{k+1}(t) + \mu(t)) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{array} \right) \quad (2.2.10)$$

Решение этого уравнения можно записать в виде

$$\mathbf{p}(t) = U^*(t, 0) \mathbf{p}(0) + \int_0^t U^*(t, \tau) \mathbf{g}(\tau) d\tau, \quad (2.2.11)$$

где $U^*(t, s)$ является оператором Коши соответствующей однородной системы

$$\frac{d\mathbf{x}}{dt} = A^*(t) \mathbf{x}. \quad (2.2.12)$$

Теорема 10 Пусть интенсивности катастроф существенны, т.е.

$$\int_0^\infty \gamma^*(t) dt = +\infty. \quad (2.2.13)$$

Тогда процесс $X(t)$ слабо эргодичен (в равномерной операторной топологии) и имеет место следующая оценка скорости сходимости

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq e^{-\int_0^t \gamma^*(\tau) d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\| \leq 2e^{-\int_0^t \gamma^*(\tau) d\tau} \quad (2.2.14)$$

для любых начальных условий $\mathbf{p}^*(0), \mathbf{p}^{**}(0)$ и всех $t \geq 0$.

Доказательство. Это утверждение вытекает из равенства

$$\gamma(A^*(t))_1 = \sup_i \left(a_{ii}^*(t) + \sum_{j \neq i} a_{ji}^*(t) \right) = -\gamma^*(t), \quad (2.2.15)$$

отсюда получаем

$$\begin{aligned} \|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| &\leq \|U^*(t, 0)\| \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\| \leq \\ &\leq e^{-\int_0^t \gamma^*(\tau) d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\| \leq 2e^{-\int_0^t \gamma^*(\tau) d\tau} \end{aligned} \quad (2.2.16)$$

для любых начальных условий $\mathbf{p}^*(0)$, $\mathbf{p}^{**}(0)$ и всех $t \geq 0$.

Рассмотрим теперь получение оценок скорости сходимости во «взвешенных» нормах. Рассмотрим неубывающую последовательность $1 = d_0 \leq d_1 \leq \dots$ и диагональную матрицу $D = \text{diag}(d_0, d_1, d_2, \dots)$. Обозначим через $l_{1D} = \{\mathbf{p} / \|\mathbf{p}\|_{1D} = \|D\mathbf{p}\|_1 < \infty\}$. $A^*(t)$ является существеннонеорцицательной матрицей, т.е. все внедиагональные элементы $A^*(t)$ неотрицательны для всех $t \geq 0$.

Положим

$$\gamma_{**}(t) = \inf_i \left(|a_{ii}^*(t)| - \sum_{j \neq i} \frac{d_j}{d_i} a_{ji}^*(t) \right). \quad (2.2.17)$$

Рассмотрим (2.2.8) как дифференциальное уравнение в пространстве последовательностей l_{1D} . Тогда

$$\begin{aligned} \|A^*(t)\|_{1D} &= \|DA^*(t)D^{-1}\| = \sup_i \left(|a_{ii}^*(t)| + \sum_{j \neq i} \frac{d_j}{d_i} a_{ji}^*(t) \right) = \\ &= \sup_i \left(2|a_{ii}^*(t)| + \sum_{j \neq i} \frac{d_j}{d_i} a_{ji}^*(t) - |a_{ii}^*(t)| \right) \leq \\ &\leq 2 \sup_i |a_{ii}^*(t)| - \gamma_{**}(t) \leq 2L - \gamma_{**}(t), \end{aligned} \quad (2.2.18)$$

следовательно, операторная функция $A^*(t)$ ограничена в пространстве l_{1D} , и можно применить тот же подход к уравнению (2.2.8) в пространстве l_{1D} . Получаем

$$\gamma(A^*(t))_{1D} = \gamma(DA^*(t)D^{-1}) = \sup_i \left(a_{ii}^*(t) + \sum_{j \neq i} \frac{d_j}{d_i} a_{ji}^*(t) \right) = -\gamma_{**}(t) \quad (2.2.19)$$

и следующее утверждение.

Теорема 11 Пусть для некоторой последовательности $\{d_i\}$, $1 = d_0 \leq d_1 \leq d_2 \leq \dots$ выполнено условие

$$\int_0^\infty \gamma_{**}(t) dt = +\infty. \quad (2.2.20)$$

Тогда $X(t)$ слабо эргодичен и выполняется следующая оценка скорости сходимости:

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{1D} \leq e^{-\int_0^t \gamma_{**}(\tau) d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1D} \quad (2.2.21)$$

для любых начальных условий $\mathbf{p}^*(0), \mathbf{p}^{**}(0)$ и всех $t \geq 0$.

Пусть $l_{1E} = \{\mathbf{p} = (r, p_0, p_1, p_2, \dots)\}$ пространство последовательностей таких что $\|\mathbf{p}\|_{1E} = \sum_{k \geq 0} k|p_k| < \infty$. $\|\mathbf{p}\|_{1D} = \|D\mathbf{p}\| = \|(d_0r, d_1p_0, d_2p_1, \dots)^T\| = d_0r + \sum_{k \geq 0} d_{k+1}p_k \geq \sum_{k \geq 1} k \frac{d_{k+1}}{k} p_k$. $W = \inf_{k \geq 1} \frac{d_{k+1}}{k}$. Тогда $W\|\mathbf{p}\|_{1E} \leq \|\mathbf{p}\|_{1D}$.

Следствие 1. Пусть последовательность $\{d_i\}$ такая, что выполняется (2.2.20) и существует $W > 0$. Тогда $X(t)$ имеет предельное среднее значение, скажем $\phi(t) = E(t, 0)$, и выполняется следующая оценка:

$$|E(t, j) - E(t, 0)| \leq \frac{d_{j+1}}{W} e^{-\int_0^t \gamma_{**}(\tau) d\tau} \quad (2.2.22)$$

для всех j и всех $t \geq 0$.

Пусть $d_0 = 1$ и $d_{n+1} = (1 + \varepsilon)d_n$ для $n \geq 0$ и положительного ε .

Тогда, также аналогично [71], можем получить следующие явные оценки.

Утверждение 1. Пусть существует $\varepsilon > 0$ такое что

$$\int_0^\infty (\gamma^*(t) - \varepsilon v(t)) dt = +\infty, \quad (2.2.23)$$

где $v(t) = \max(\eta(t), \lambda(t))$. Тогда

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{1D} \leq e^{-\int_0^t (\gamma^*(\tau) - \varepsilon v(\tau)) d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1D} \quad (2.2.24)$$

и

$$|E(t, j) - E(t, 0)| \leq \frac{d_{j+1}}{W} e^{-\int_0^t (\gamma^*(\tau) - \varepsilon v(\tau)) d\tau} \quad (2.2.25)$$

для соответствующего l_{1D} и W .

Второй подход, см. также [85]. Рассмотрим частный случай, при котором скорости катастроф одинаковые, а именно, предположим, что все $\gamma_n(t) =$

$\gamma^*(t)$. В этой ситуации уравнение (2.1.1) будет выглядеть следующим образом $r'(t) = -\eta(t)r(t) + \gamma^*(t)$, следовательно, его можно решить:

$$r(t) = \int_0^t e^{-\int_{\tau}^t \eta(u) du} \gamma^*(\tau) d\tau, \quad (2.2.26)$$

т.к. $r(0) = 0$.

Рассмотрим теперь систему Колмогорова (2.2.8) в виде

$$\frac{d\mathbf{z}}{dt} = B(t)\mathbf{z} + \mathbf{f}(t), \quad t \geq 0, \quad (2.2.27)$$

где $\mathbf{f}(t) = (\eta(t)r(t), 0, 0, \dots)^T$, $\mathbf{z}(t) = (p_0(t), p_1(t), \dots)^T$, и

$$B(t) = \begin{pmatrix} -(\lambda(t) + \gamma^*(t)) & \mu(t) & 0 & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots \\ \lambda(t) & -(\lambda(t) + \gamma^*(t) + \mu(t)) & \mu(t) & 0 & \dots & 0 & 0 & 0 & 0 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & \lambda(t) & -(\lambda(t)\beta(t) + \gamma^*(t) + \mu(t)) & \mu(t) & 0 & \dots \\ 0 & 0 & 0 & 0 & \dots & 0 & 0 & \lambda(t)\beta(t) & -(\lambda(t)\beta(t) + \gamma^*(t) + \mu(t)) & \mu(t) & \dots \\ \vdots & \vdots & \vdots & \vdots & \dots & \vdots & \vdots & \vdots & \vdots & \vdots & \dots \end{pmatrix}.$$

Решение уравнения (2.2.27) может быть записано в виде

$$\mathbf{z}(t) = U_B(t, 0)\mathbf{z}(0) + \int_0^t U_B(t, \tau)\mathbf{f}(\tau) d\tau, \quad (2.2.28)$$

где $U_B(t, s)$ оператор Коши соответствующего однородного уравнения

$$\frac{d\mathbf{v}}{dt} = B(t)\mathbf{v}. \quad (2.2.29)$$

Заметим, что в этом случае оценка будет полностью аналогична теореме 10, только с заменой в левой части \mathbf{p} на \mathbf{z} .

Существенно иная ситуация с этим подходом возникает, когда рассматриваем общий случай при получении взвешенных оценок.

Теперь не можем найти $r(t)$ в замкнутой форме, как в (2.2.26). Вместо этого используем равенство $r(t) = 1 - \sum_{i \geq 0} p_i(t)$. Затем снова получаем уравнение (2.2.27) с другим $B(t)$,

$$B(t) = \begin{pmatrix} -(\lambda(t) + \gamma_0(t) + \eta(t)) & \mu(t) - \eta(t) & -\eta(t) & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \lambda(t) & -(\lambda(t) + \gamma_1(t) + \mu(t)) & \mu(t) & \cdots & 0 & 0 & 0 & 0 & \cdots & \cdots \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \vdots & \cdots & \cdots \\ 0 & 0 & 0 & 0 & \cdots & \lambda(t) & -(\lambda(t)\beta(t) + \gamma_k(t) + \mu(t)) & \mu(t) & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 0 & \lambda(t)\beta(t) & -(\lambda(t)\beta(t) + \gamma_{k+1}(t) + \mu(t)) & \mu(t) & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}.$$

Более того, теперь $\mathbf{f}(t) = (\eta(t), 0, 0, \dots)^T$ и $\mathbf{z}(t) = (p_0(t), p_1(t), \dots)^T$.

Рассмотрим

$$\mathfrak{D} = \begin{pmatrix} d_0 & d_0 & d_0 & \cdots \\ 0 & d_1 & d_1 & \cdots \\ 0 & 0 & d_2 & \cdots \\ \cdots & \cdots & \cdots & \cdots \end{pmatrix} \quad (2.2.30)$$

и $B^*(t) = \mathfrak{D}B(t)\mathfrak{D}^{-1}(t) =$

$$\begin{pmatrix} -(\eta(t) + \gamma_0(t)) & \frac{d_0}{d_1}(\gamma_0(t) - \gamma_1(t)) & \frac{d_0}{d_2}(\gamma_1(t) - \gamma_2(t)) & \cdots & \cdots & \cdots \\ \frac{d_1}{d_0}\lambda(t) & -(\lambda(t) + \gamma_1(t) + \mu(t)) & \frac{d_1}{d_2}(\mu(t) + \gamma_1(t) - \gamma_2(t)) & \frac{d_1}{d_3}(\gamma_2(t) - \gamma_3(t)) & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \end{pmatrix}. \quad (2.2.31)$$

Положим

$$\gamma_B(t) = \inf_i \left(|b_{ii}^*(t)| - \sum_{j \neq i} b_{ji}^*(t) \right). \quad (2.2.32)$$

Рассмотрим (2.2.27) как дифференциальное уравнение в пространстве последовательностей $l_{1\mathfrak{D}}$. Имеем

$$\begin{aligned} \|B^*(t)\| &= \|B(t)\|_{1\mathfrak{D}} = \|\mathfrak{D}B(t)\mathfrak{D}^{-1}\| = \sup_i \left(|b_{ii}^*(t)| + \sum_{j \neq i} b_{ji}^*(t) \right) = \quad (2.2.33) \\ &= \sup_i \left(2|b_{ii}^*(t)| + \sum_{j \neq i} b_{ji}^*(t) - |b_{ii}^*(t)| \right) \leq 2 \sup_i |b_{ii}^*(t)| - \gamma_B(t) \leq 2L - \gamma_B(t), \end{aligned}$$

следовательно, операторная функция $B(t)$ ограничена в пространстве $l_{1\mathfrak{D}}$, и можно применить тот же подход к уравнению (2.2.27) в пространстве $l_{1\mathfrak{D}}$. Теперь равенство

$$\gamma(B(t))_{1\mathfrak{D}} = \gamma(\mathfrak{D}B(t)\mathfrak{D}^{-1}) = \sup_i \left(b_{ii}^*(t) + \sum_{j \neq i} b_{ji}^*(t) \right) = -\gamma_B(t) \quad (2.2.34)$$

подразумевает следующее утверждение.

Теорема 12 Пусть $\{d_i\}$, $1 = d_0 \leq d_1 \leq \dots$ неубывающая последовательность, такая что,

$$\int_0^\infty \gamma_B(t) dt = +\infty. \quad (2.2.35)$$

Тогда $X(t)$ слабо эргодичен и выполняется следующая оценка:

$$\|z^*(t) - z^{**}(t)\|_{1D} \leq e^{-\int_0^t \gamma_B(\tau) d\tau} \|z^*(0) - z^{**}(0)\|_{1D} \quad (2.2.36)$$

для любых начальных условий $z^*(0), z^{**}(0)$ и всех $t \geq 0$.

Пусть $l_{1E} = \{\mathbf{z} = (p_0, p_1, p_2, \dots)\}$ пространство последовательностей таких, что $\|\mathbf{z}\|_{1E} = \sum_{k \geq 1} k|p_k| < \infty$. Пусть $W = \inf_{k \geq 1} \frac{d_k}{k}$. Тогда $W\|\mathbf{z}\|_{1E} \leq \|\mathbf{z}\|_{1D}$.

Следствие 2. Пусть последовательность $\{d_i\}$ такова, что (2.2.35) выполняется, а $W > 0$. Тогда $X(t)$ имеет предельное среднее значение, скажем $\phi(t) = E(t, 0)$, и выполняется следующая оценка:

$$|E(t, j) - E(t, 0)| \leq \frac{1 + d_j}{W} e^{-\int_0^t \gamma_B(\tau) d\tau} \quad (2.2.37)$$

для всех j и всех $t \geq 0$.

Замечание 1. Можно взять $d_0 = 1$, $d_1 = \epsilon$, $d_{k+1} = (1 + \epsilon)d_k$ для $k \geq 1$ и получить аналог утверждения 1 для второго подхода.

Замечание 2. Во всех наших утверждениях можно заменить условие монотонности последовательности $\{d_k\}$ условием $d = \inf_k d_k > 0$, с соответствующим изменением оценок; см., например, [87].

2.3 Оценки устойчивости

Здесь рассмотрим применение общего метода оценки устойчивости (см. в [86]) для исследуемой модели.

Рассмотрим «возмущенный» процесс длины очереди $\bar{X}(t)$, $t \geq 0$ с соответствующей транспонированной матрицей интенсивности $\bar{A}(t)$, где матрица «возмущения» $\hat{A}(t) = A(t) - \bar{A}(t)$ в некотором смысле мала. А именно, предполагаем, что возмущенная очередь имеет ту же природу, что и исходная. Следовательно, матрица возмущенной интенсивности также имеет ту же структуру, с соответствующими возмущенными интенсивностями $\bar{\eta}(t)$, $\bar{\gamma}_n(t)$, $\bar{\lambda}(t)$, $\bar{\mu}(t)$, $\bar{\beta}(t)$.

Пусть

$$\begin{aligned} |\eta(t) - \bar{\eta}(t)| = |\hat{\eta}(t)| &\leq \hat{\epsilon}, & |\gamma_n(t) - \bar{\gamma}_n(t)| = |\hat{\gamma}_n(t)| &\leq \hat{\epsilon}, \\ |\lambda(t) - \bar{\lambda}(t)| = |\hat{\lambda}(t)| &\leq \hat{\epsilon}, & |\mu(t) - \bar{\mu}(t)| = |\hat{\mu}(t)| &\leq \hat{\epsilon}, \\ |\beta(t) - \bar{\beta}(t)| = |\hat{\beta}(t)| &\leq \hat{\epsilon}. \end{aligned} \quad (2.3.38)$$

Следовательно

$$|\lambda(t)\beta(t) - \bar{\lambda}(t)\bar{\beta}(t)| \leq \lambda(t)|\hat{\beta}(t)| + \bar{\beta}(t)|\hat{\lambda}(t)| \leq (L + 1)\hat{\epsilon}. \quad (2.3.39)$$

Тогда получаем из (2.1.7) следующее ограничение

$$\begin{aligned} \|\hat{A}(t)\| &= 2 \sup_k |\hat{a}_{kk}(t)| = 2 \max(|\hat{\eta}(t)|, |\hat{\lambda}(t)| + |\hat{\gamma}(t)|), \\ |\hat{\lambda}(t)| + |\hat{\gamma}_n(t)| + |\hat{\mu}(t)|, &(L + 1)\hat{\epsilon} + |\hat{\gamma}_n(t)| + |\hat{\mu}(t)| \leq (2L + 6)\hat{\epsilon}. \end{aligned} \quad (2.3.40)$$

Во-первых, сформулируем оценки возмущения для вектора вероятности состояния в ситуации теоремы 10.

Следующее утверждение непосредственно вытекает из теоремы 10.

Теорема 13 Пусть в предположении теоремы 10 интенсивность катастрофы $\gamma(t)$ будет такой, что

$$e^{-\int_s^t \gamma(\tau) d\tau} \leq N e^{-\gamma_0(t-s)}, \quad (2.3.41)$$

для некоторых положительных N, γ_0 . Тогда выполняется следующая оценка возмущения:

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\| \leq \frac{\hat{\epsilon}(2L + 6)(1 + \log(N/2))}{\gamma_0} \quad (2.3.42)$$

для любой возмущенной очереди с соответствующими возмущенными интенсивностями, удовлетворяющими (2.3.38).

Оценка теоремы 13 основана на результатах устойчивости из [82, 83].

Обратите внимание, что (2.2.10), (2.3.38) и (2.3.40) предполагает неравенство:

$$\|\hat{A}^*(t)\| = 2 \sup_k |\hat{a}_{kk}^*(t)| \leq \|\hat{A}(t)\| \leq (2L + 6) \hat{\epsilon}. \quad (2.3.43)$$

С другой стороны, у нас есть $\|\mathbf{g}(t)\|_{1D} = \gamma^*(t) \leq L$ почти для всех $t \geq 0$.

Тогда Теорема 4 из [83] подразумевает следующее утверждение.

Теорема 14 Пусть при допущениях теоремы 11 выполняются следующие оценки:

$$e^{-\int_s^t \gamma_{**}(\tau) d\tau} \leq N^{**} e^{-\gamma_0^{**}(t-s)} \quad (2.3.44)$$

для некоторых положительных N^{**}, γ_0^{**} и

$$H = \sup_{|i-j|=1} \frac{d_i}{d_j} < \infty. \quad (2.3.45)$$

Тогда

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\|_{1D} \leq \frac{(4L + 12) \hat{\epsilon} H L (N^{**})^2}{\gamma_0^{**} (\gamma_0^{**} - (4L + 12) \hat{\epsilon} H)}. \quad (2.3.46)$$

Более того, если $W = \inf_{k \geq 1} \frac{d_{k+1}}{k} > 0$, тогда

$$\limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq \frac{(4L + 12) \hat{\epsilon} H L (N^{**})^2}{W \gamma_0^{**} (\gamma_0^{**} - (4L + 12) \hat{\epsilon} H)}. \quad (2.3.47)$$

Наконец, получаем оценки возмущений, основанные на оценках эргодичности теоремы 12.

Теорема 15 Пусть при допущениях теоремы 12 выполняются следующие оценки:

$$e^{-\int_s^t \gamma_B(\tau) d\tau} \leq N^B e^{-\gamma_0^B (t-s)} \quad (2.3.48)$$

для некоторых положительных N^B, γ_0^B и

$$H = \sup_{|i-j|=1} \frac{d_i}{d_j} < \infty. \quad (2.3.49)$$

Тогда

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\|_{1D} \leq \frac{\hat{\epsilon} N^B (L + 1) (6H L N^B + \gamma_0^B)}{\gamma_0^B (\gamma_0^B - 12\hat{\epsilon} H N^B (L + 1))}. \quad (2.3.50)$$

Кроме того, если $W = \inf_{k \geq 1} \frac{d_k}{k} > 0$, тогда

$$\limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq \frac{\hat{\epsilon} N^B (L + 1) (6H L N^B + \gamma_0^B)}{W \gamma_0^B (\gamma_0^B - 12\hat{\epsilon} H N^B (L + 1))}. \quad (2.3.51)$$

Доказательство. Достаточно отметить, что

$$\|B(t) - \bar{B}(t)\|_{1D} \leq H \|B(t) - \bar{B}(t)\|_1 \leq H \|A(t) - \bar{A}(t)\|_1 \leq (2L + 6) H \hat{\epsilon} \quad (2.3.52)$$

и

$$\|\mathbf{f}(t) - \bar{\mathbf{f}}(t)\|_{1D} = \|\eta(t)r(t) - \bar{\eta}(t)\bar{r}(t)\| \leq (L + 1) \hat{\epsilon}. \quad (2.3.53)$$

тогда наше утверждение следует из теоремы 2 [86].

2.4 Примеры

Для подтверждения теоретических оценок проведём вычислительные эксперименты. Рассмотрим, в качестве примеров, две модели массового обслуживания с 1-периодическими интенсивностями и с пороговым значением равным 100 ($k = 100$).

Сразу отметим, что для построения вероятностных характеристик моделей, представленных в диссертационной работе, используется алгоритм, заключающийся в усечении исходного процесса, использовании скорости сходимости для оценивания оператора Коши, решении системы дифференциальных уравнений с помощью метода Адамса-Мултона IV порядка, нахождении основных вероятностных характеристик и, наконец, построении графиков этих характеристик с помощью разработанных программ, позволяющих с оптимальной скоростью находить значения в случае достаточно большого количества уравнений N в усеченной системе и при достаточно больших t .

2.4.1 Пример 1

Пусть очередь имеет следующие интенсивности перехода:

$$\eta(t) = 3 + \sin 2\pi t, \quad \gamma_k(t) = \gamma(t) = 2 + 0.5 \cos 2\pi t, \quad \lambda(t) = 10 + 10 \sin 2\pi t, \quad \mu(t) = 2 + \cos 2\pi t, \quad \beta(t) = 0.7, \quad k = 100.$$

Применим все исследования для этой конкретной ситуации.

Для теоремы 10 и соответствующей теоремы 13 об «устойчивости» нам нужны L , N и γ_0 . Очевидно, что у нас есть $L \leq 25.5$. Рассмотрим теперь

$$e^{-\int_s^t \gamma(\tau) d\tau} = e^{-2(t-s) - \frac{\sin 2\pi t - \sin 2\pi s}{4\pi}} \leq e^{-2(t-s) + \frac{1}{2\pi}} \leq 2e^{-2(t-s)},$$

следовательно, можно взять $N = 2$ и $\gamma_0 = 2$ в (2.3.41).

Для применения теорем 11, 14 возьмем $\varepsilon = 0.05$, $d_0 = 1$ и $d_{k+1} = (1 + \varepsilon)d_k$ для $k \geq 0$. Тогда $H = 1 + \varepsilon < 2$, $v(t) = \max(\eta(t), \lambda(t)) = 10 + 10 \sin 2\pi t$, и $\gamma_{**}(t) = \gamma_\varepsilon(t) = 1.5 + 0.5 \cos 2\pi t - 0.5 \sin 2\pi t$.

$$e^{-\int_s^t \gamma_{**}(\tau) d\tau} \leq e^{-1.5(t-s) + \frac{1}{4\pi}} \leq 2e^{-1.5(t-s)},$$

поэтому можно взять $N^{**} = 2$ и $\gamma_0^{**} = 1.5$ в (2.3.44), $W = 0.13$.

Наконец, для применения теорем 12, 15 возьмем $\varepsilon = 0.05$, $d_0 = 1$, $d_1 = \varepsilon$, и $d_{k+1} = (1 + \varepsilon)d_k$ для $k \geq 1$. Тогда $H = \frac{1}{\varepsilon}$, и $\gamma_B(t) = \gamma_{**}(t) = \gamma_\varepsilon(t)$, поэтому можно взять $N^B = 2$, $\gamma_0^B = 1.5$ в (2.3.48).

Получаем следующие оценки скорости сходимости:

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 4e^{-2t} \quad (2.4.54)$$

из теоремы 10;

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{1D} \leq 2e^{-1.5t} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1D}, \quad (2.4.55)$$

$$|E(t, j) - E(t, 0)| \leq \frac{2 \cdot 1.05^{j+1}}{0.13} e^{-1.5t} \quad (2.4.56)$$

из теоремы 11 и следствия 1, и почти то же самое из теоремы 15.

Соответствующие оценки возмущений имеют вид:

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\| \leq 30\hat{\varepsilon} \quad (2.4.57)$$

из теоремы 13;

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\|_{1D} \leq \frac{4 \cdot 10^5 \hat{\varepsilon}}{1 - 4 \cdot 10^3 \hat{\varepsilon}} \quad (2.4.58)$$

и

$$\limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq \frac{4 \cdot 10^5 \hat{\varepsilon}}{0.13(1 - 4 \cdot 10^3 \hat{\varepsilon})} \quad (2.4.59)$$

из теоремы 14; а оценки теоремы 15 гораздо хуже.

Для подтверждения теоретических оценок и построения предельного режима выполним численный эксперимент:

1. на основании оценки, в данном случае (2.4.54), получаем значение t° , гарантирующее сходимость решения $\mathbf{p}(\mathbf{t})$ на отрезке $[0; t^\circ]$;
2. численно решаем задачу Коши при начальных условиях $X(0) = 0$ и $X(0) = 200$: вычисляем вероятность пустой очереди $p_0(t)$ и математическое ожидание числа требований в очереди $E(t, k)$;
3. на основании построенных графиков выделяем отрезок $[0; t^* + 1]$, который в нашем примере меньше чем гарантированный неравенством (2.4.54);
4. в итоге получаем предельный режим на отрезке единичной длины $[t^*; t^* + 1]$.

На рисунках 2.1-2.4 показано поведение вероятности пустой очереди и среднего значения соответственно. Заметим, что оценка (2.4.54) гарантирует совпадение вероятностных характеристик с различными начальными условиями и заданной точностью для $t \geq 5$. На самом деле, как показывают графики, разница становится довольно мала гораздо раньше.

На рис. 2.5-2.6 можно видеть оценки возмущений для соответствующих предельных характеристик с $\hat{\epsilon} = 10^{-3}$ для (2.4.57) и $\hat{\epsilon} = 10^{-7}$ для (2.4.59). Очевидно, что реальные возмущения вероятности пустой очереди и среднего числа требований в системе меньше чем гарантирует теоретические оценки (2.4.58) и (2.4.59).

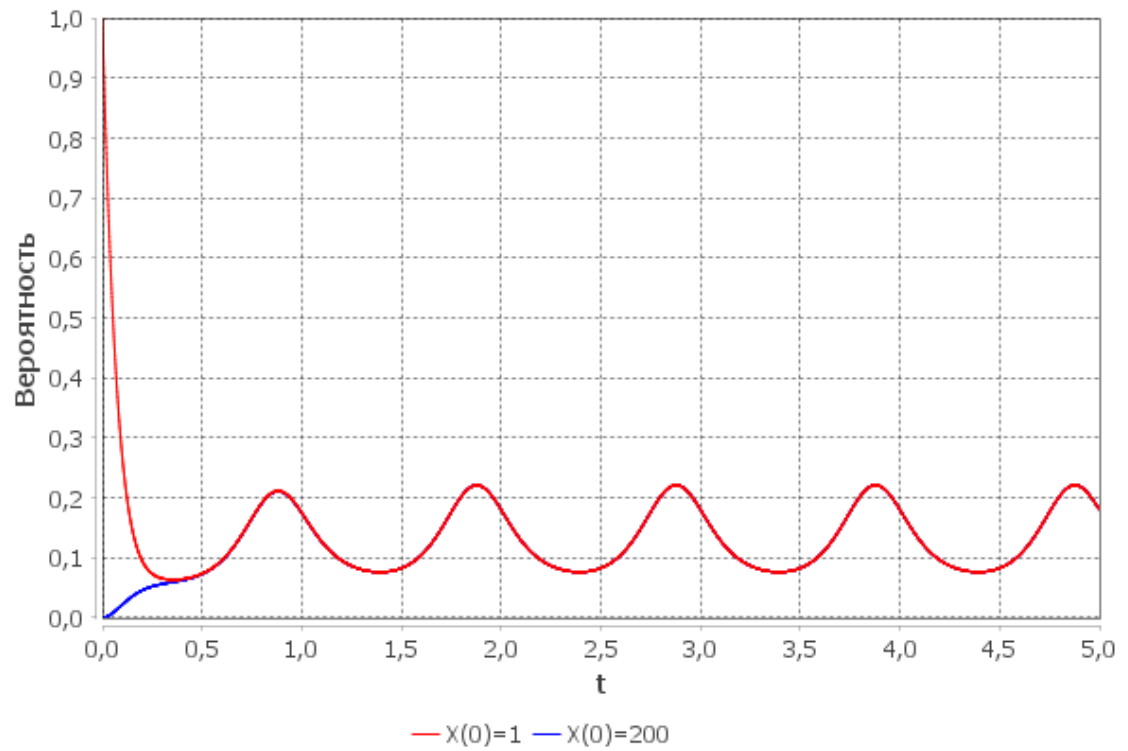


Рис. 2.1: Пример 1. Вероятность пустой очереди для $t \in [0, 5]$.

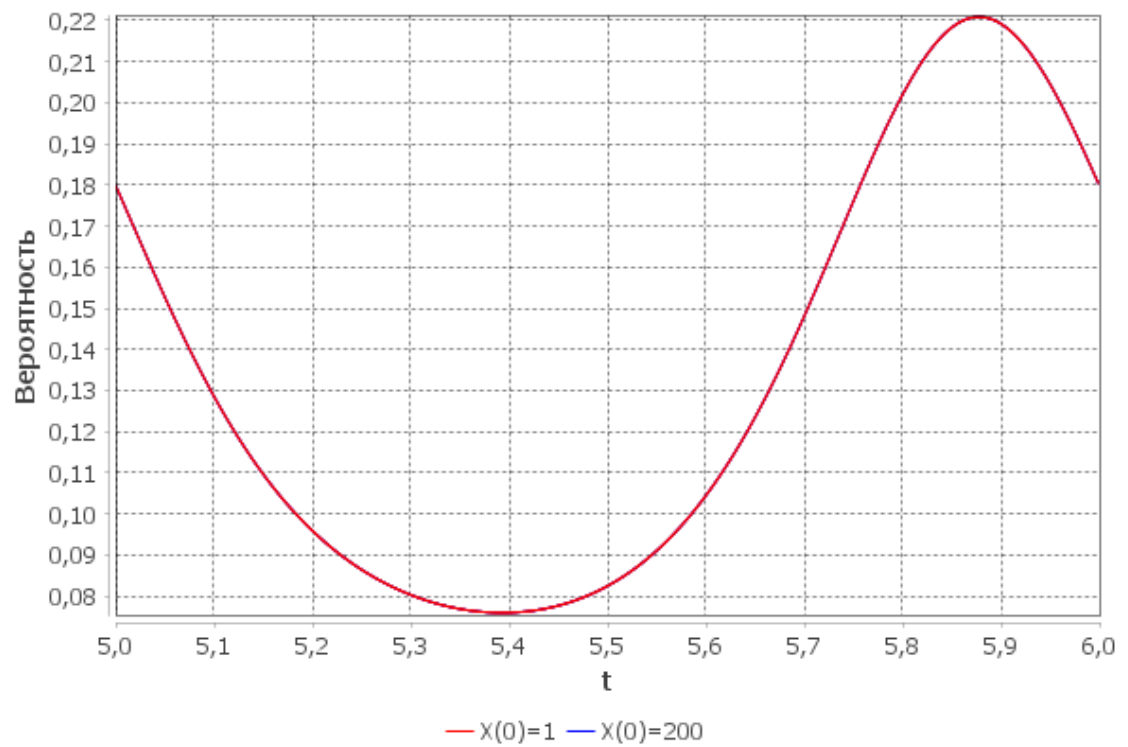


Рис. 2.2: Пример 1. Аппроксимация предельной вероятности пустой очереди для $t \in [5, 6]$.

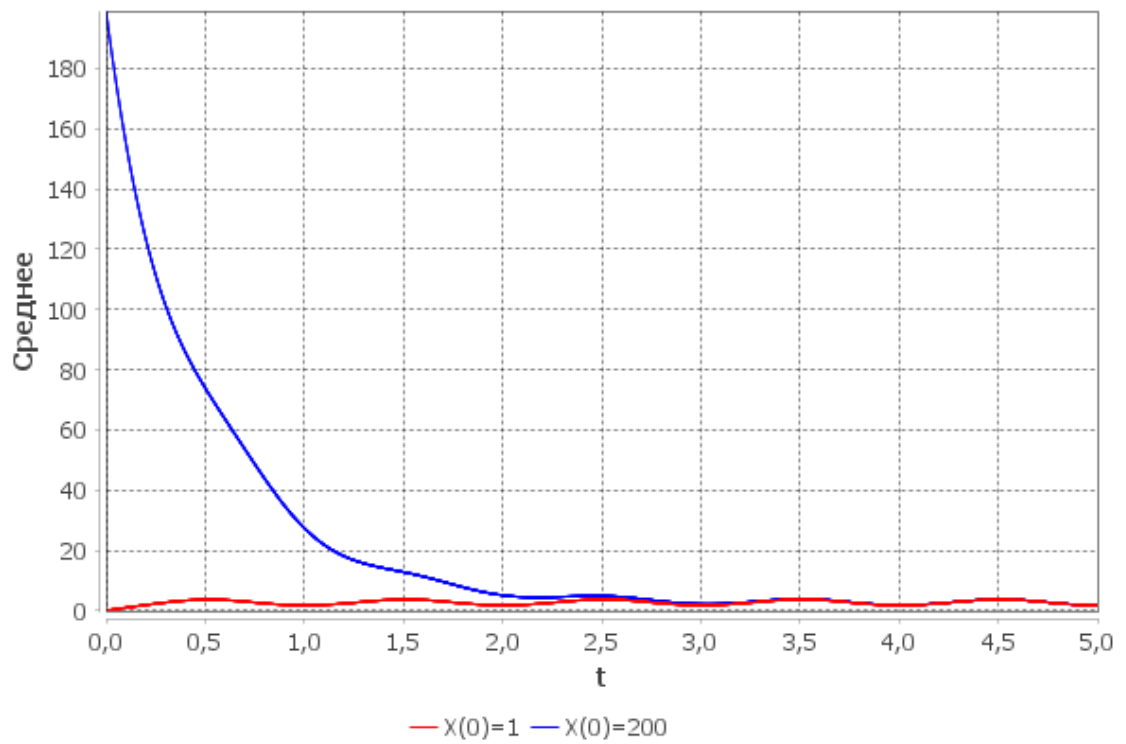


Рис. 2.3: Пример 1. Среднее $E(t, k)$ для $t \in [0, 5]$.

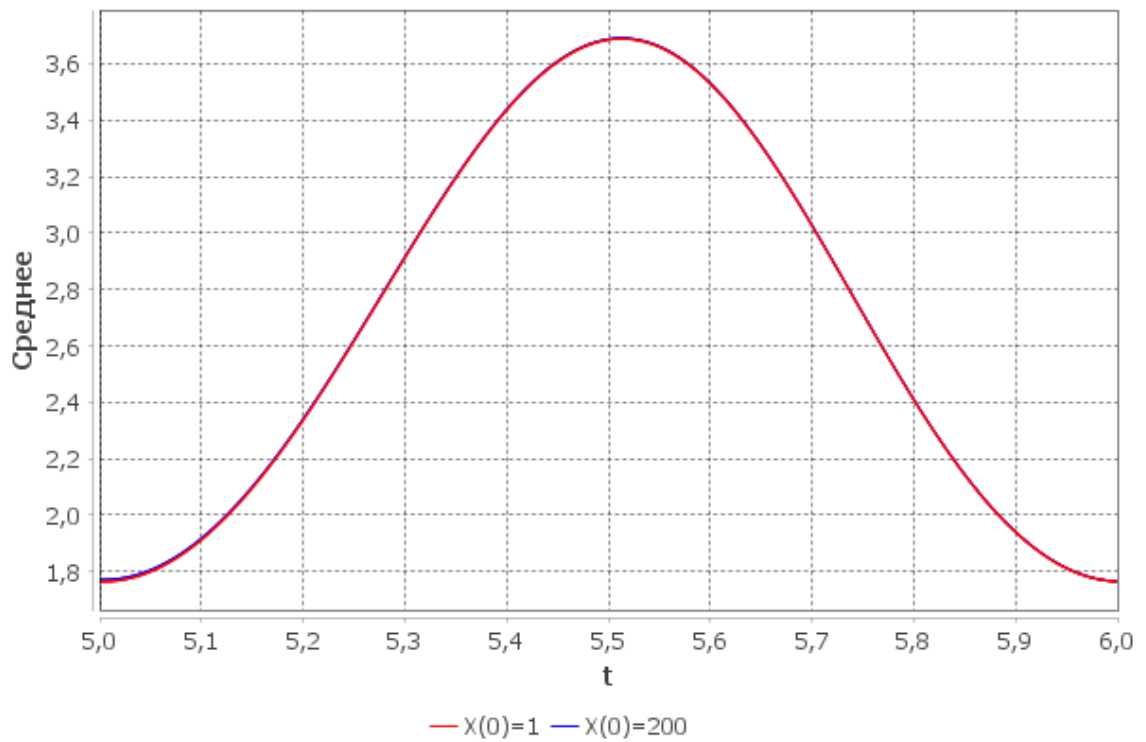


Рис. 2.4: Пример 1. Аппроксимация предельного среднего значения $E(t, k)$ для $t \in [5, 6]$.

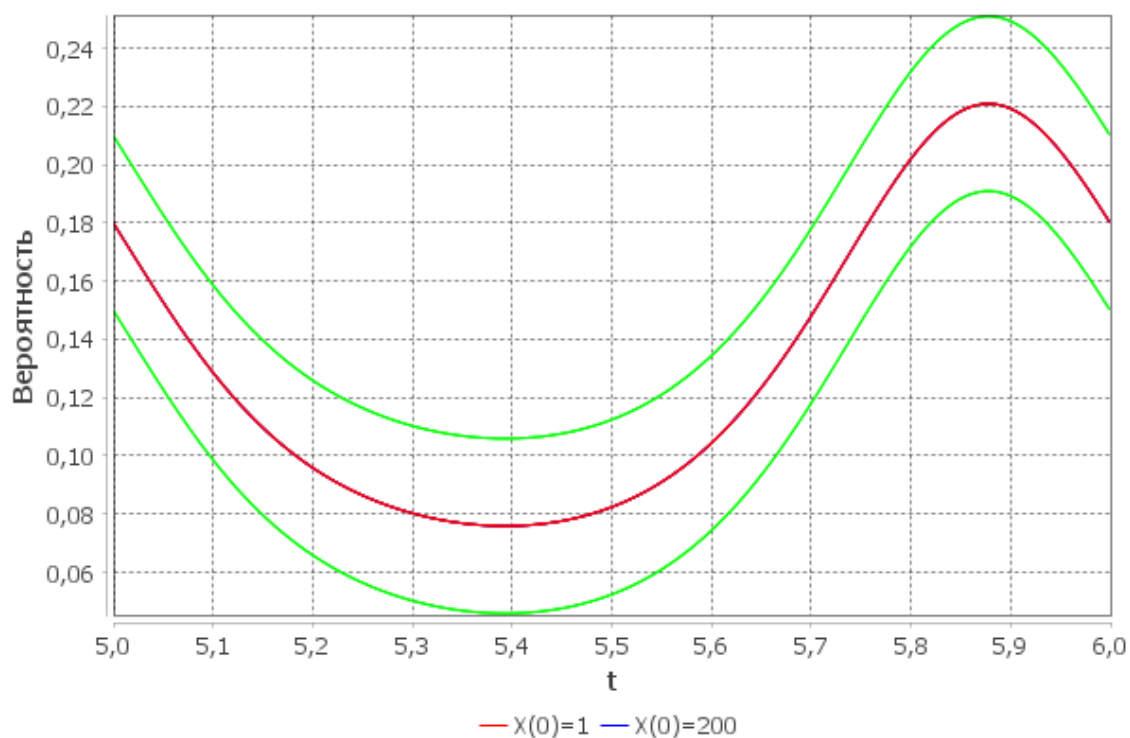


Рис. 2.5: Пример 1. Оценки возмущения для предельной вероятности пустой очереди для $t \in [5, 6]$.

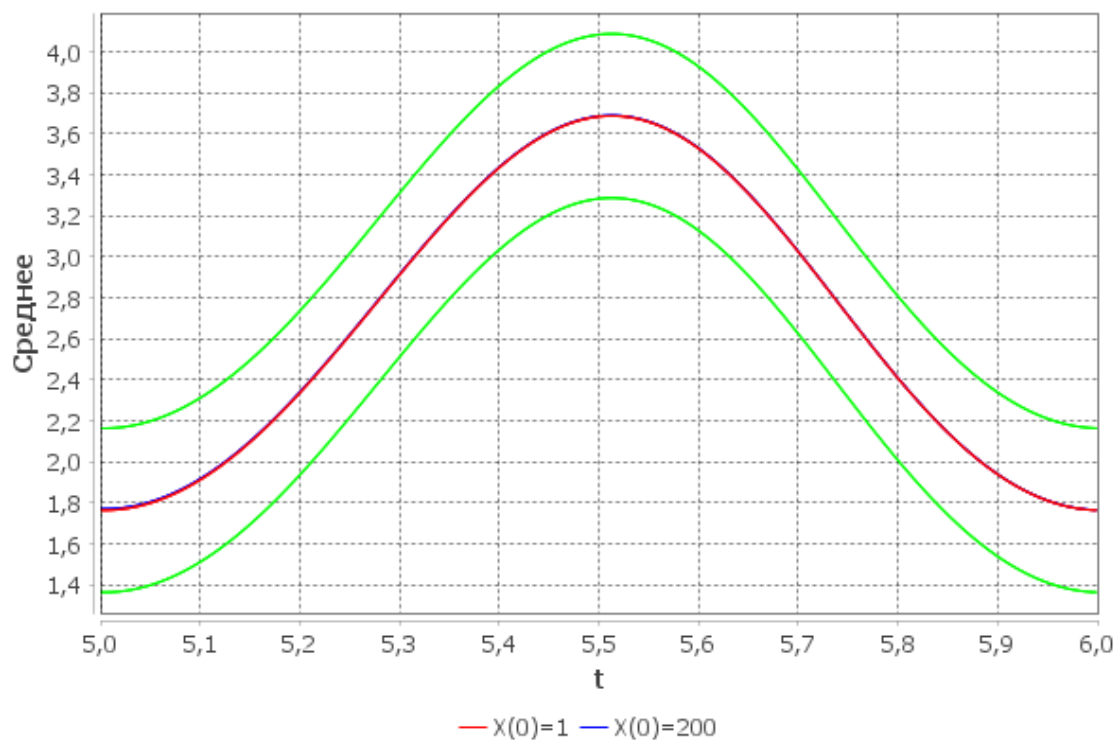


Рис. 2.6: Пример 1. Оценки возмущений для предельного среднего значения $E(t, 0)$ для $t \in [5, 6]$.

2.4.2 Пример 2

Рассмотрим модель со следующими интенсивностями: $\eta(t) = 3 + \sin 2\pi t$, $\gamma_0(t) = 2 + 0.5 \cos 2\pi t$, $\gamma_k(t) = 0$ при $k \geq 1$, $\lambda(t) = 1 + \sin 2\pi t$, $\mu(t) = 5 + \cos 2\pi t$, $\beta(t) = 0.7$.

Так как $\gamma^* = 0$, то первый подход применить нельзя.

Для использования теорем 12 и 15 положим $d_0 = 1$, $d_1 = 2.5$, и $d_{k+1} = (1.5)d_k$ для $k \geq 1$.

Получаем

$$|b_{00}^*(t)| - \sum_{j \neq 0} b_{j0}^*(t) = \eta(t) + \gamma_0(t) - \frac{5}{2}\lambda(t) \geq 0.5,$$

$$|b_{11}^*(t)| - \sum_{j \neq 1} b_{j1}^*(t) = \mu(t) - 0.5\lambda(t) - \frac{2}{5}\gamma_0(t) \geq 2,$$

$$|b_{ii}^*(t)| - \sum_{j \neq i} b_{ji}^*(t) = \frac{0.5}{1.5}\mu(t) - 0.5\lambda(t) \geq \frac{1}{3}, \quad 2 \leq i \leq k-2,$$

$$|b_{ii}^*(t)| - \sum_{j \neq i} b_{ji}^*(t) = \frac{0.5}{1.5}\mu(t) + \lambda(t)(1 - \beta(t)) - 0.5\lambda(t)\beta(t) \geq \frac{1}{3}, \quad i = k-1,$$

$$|b_{ii}^*(t)| - \sum_{j \neq i} b_{ji}^*(t) = \frac{0.5}{1.5}\mu(t) - 0.5\lambda(t)\beta(t) \geq \frac{1}{3}, \quad i \geq k.$$

Следовательно, в (2.2.34) получаем

$$\gamma_B(t) = \inf_i \left(|b_{ii}^*(t)| - \sum_{j \neq i} b_{ji}^*(t) \right) \geq \frac{1}{3},$$

и получаем из (2.2.36) и (2.2.37) следующие оценки:

$$\|z^*(t) - z^{**}(t)\|_{1D} \leq e^{-\frac{t}{3}} \|z^*(0) - z^{**}(0)\|_{1D}, \quad (2.4.60)$$

и

$$|E(t, j) - E(t, 0)| \leq \frac{1 + d_j}{W} e^{-\frac{t}{3}}. \quad (2.4.61)$$

Кроме того, теорема 15 дает нам соответствующие оценки устойчивости, а именно, $L = 8$, $H = 2$, $N^B = 1$, $\gamma_0^B = \frac{1}{3}$, и из (2.3.50) и (2.3.51) получаем следующие оценки:

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\|_{1D} \leq 10^4 \cdot \hat{\epsilon}, \quad (2.4.62)$$

и

$$\limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq 2 \cdot 10^4 \cdot \hat{\epsilon}. \quad (2.4.63)$$

Далее проводим вычислительные эксперименты, подтверждающие второй подход получения теоретических оценок.

На рисунках 2.7-2.10 показано поведение вероятности пустой очереди и среднего значения соответственно, а на рисунках 2.11-2.11 можно увидеть границы возмущений для соответствующих предельных характеристик при $\hat{\epsilon} = 10^{-7}$ для (2.4.62) и $\hat{\epsilon} = 5 \cdot 10^{-8}$ для (2.4.63).

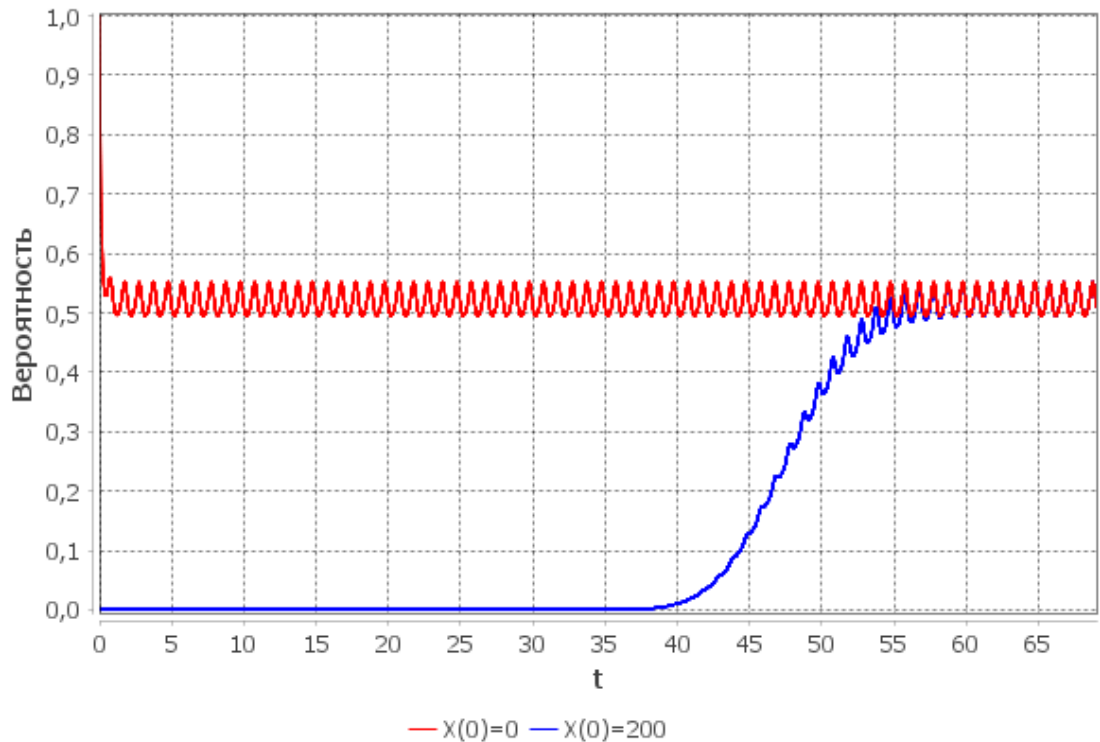


Рис. 2.7: Пример 2. Вероятность пустой очереди для $t \in [0, 69]$.

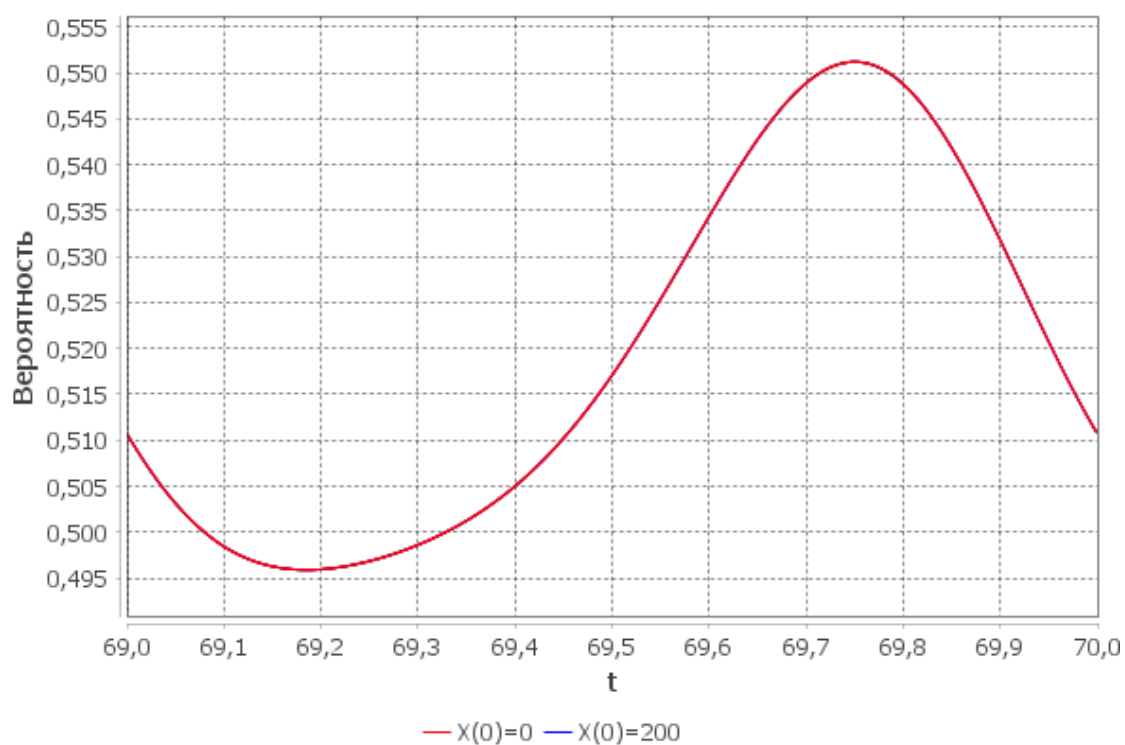


Рис. 2.8: Пример 2. Аппроксимация предельной вероятности пустой очереди для $t \in [69, 70]$.

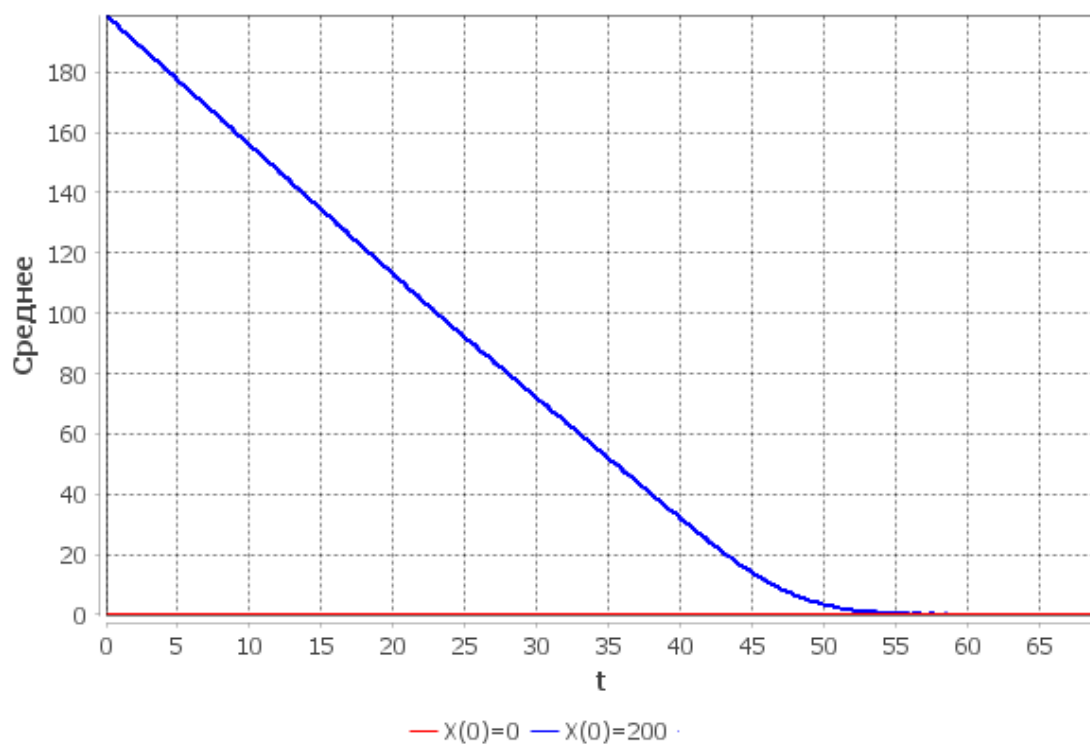


Рис. 2.9: Пример 2. Среднее $E(t, k)$ для $t \in [0, 69]$.

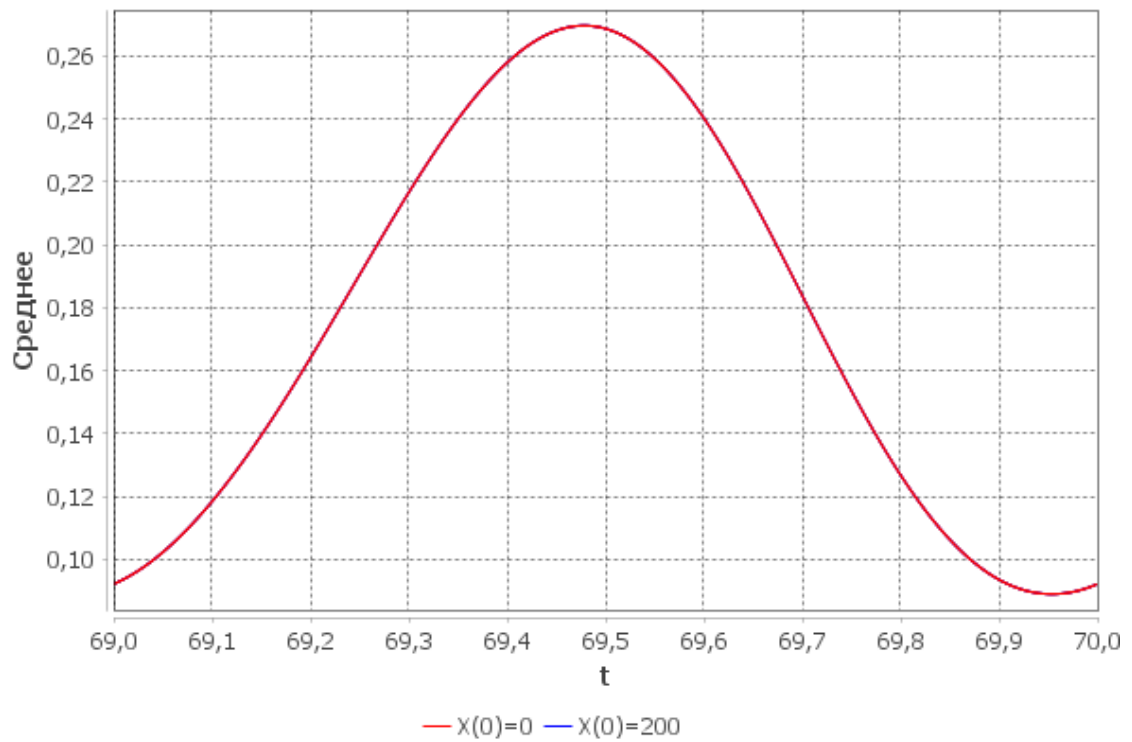


Рис. 2.10: Пример 2. Аппроксимация предельного среднего значения $E(t, k)$ для $t \in [69, 70]$.

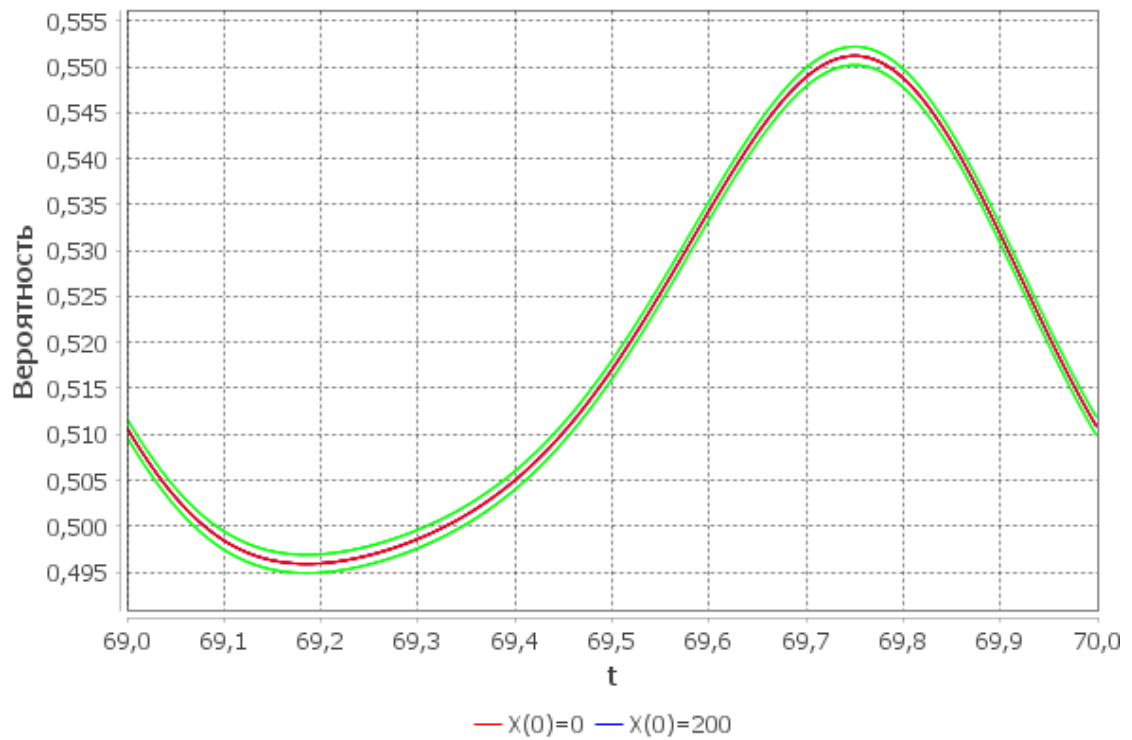


Рис. 2.11: Пример 2. Оценки возмущения для предельной вероятности пустой очереди для $t \in [69, 70]$.

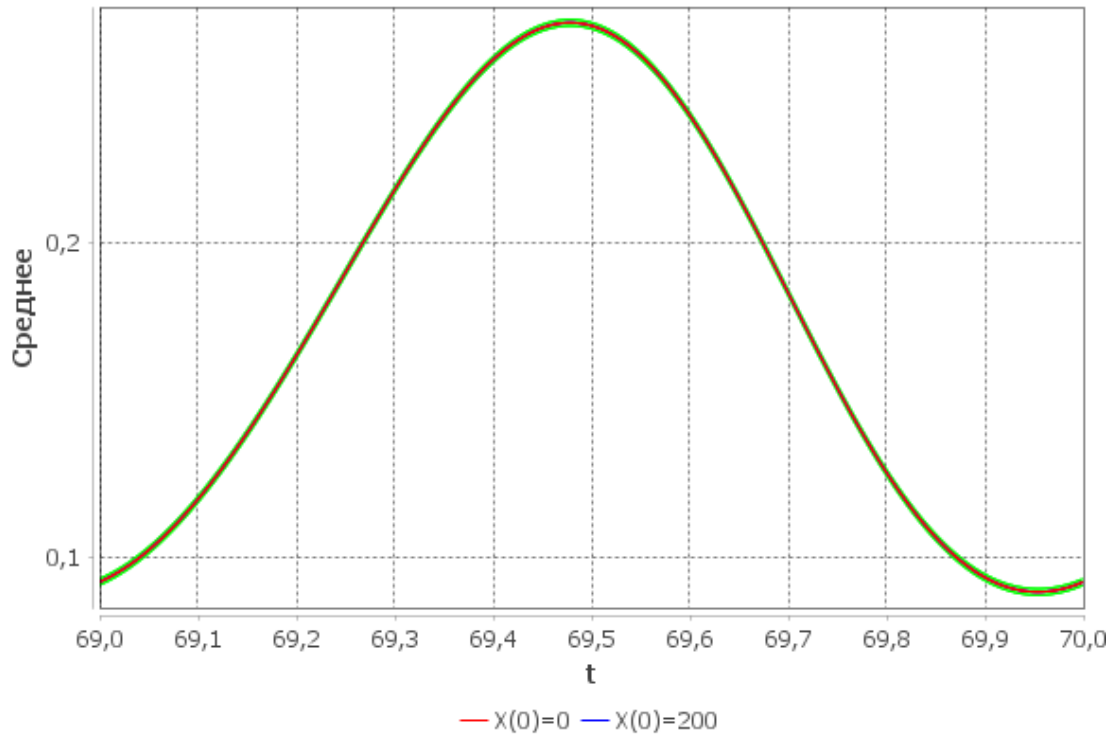


Рис. 2.12: Пример 2. Оценки возмущений для предельного среднего значения $E(t,0)$ для $t \in [69, 70]$.

Замечание. С другой стороны, имея неполную информацию о процессе, но зная его предельные характеристики, можно восстановить остальные параметры системы, например, пороговое значение k . И если знать границы для предельных характеристик для начального или возмущенного процесса, то можно управлять соответствующими границами для параметров системы.

Глава 3

Исследование модели с одним сервером, специальным групповым поступлением и со специальной политикой пропуска очереди

Две наиболее распространенные точки зрения на производительность системы массового обслуживания - это точка зрения владельца системы и клиентов системы. Обычно их цели противоречат друг другу. В то время как клиент стремится минимизировать одну (или более) характеристику заданий¹, которые он передает системе (например, среднее время отклика задания), владелец системы стремится максимально использовать ресурсы (например, загрузку процессора). Обе точки зрения привлекли внимание сообщества исследователей операций в последние десятилетия. Но оценка производительности систем массового обслуживания с точки зрения клиента, похоже, привлекла больше внимания.

Если ограничиться только очередями с одним сервером, то, вероятно, одним из наиболее известных результатов является оптимальность политики SRPT (кратчайшее оставшееся время обработки) по отношению к среднему времени отклика задания (или требования)². Как известно, в соответствии с SRPT сервер всегда работает над "кратчайшим" заданием. В [115] (на основе преды-

¹На протяжении всей этой статьи, когда говорим о характеристике производительности, имеем в виду ее долгосрочную перспективу, т.е. значение, когда система находится в стационарном или предельном режиме.

²См. [111] и [103] для получения последних результатов для многосерверных очередей.

дущих работ [99, 109]) было замечено, что несколько похожая идея может быть использована для построения политик³, которые увеличивают использование всех серверов в системе. Одна из таких политик, далее именуемая политикой "пропуска очереди", работает следующим образом (см. рис. 3.1). Предположим,

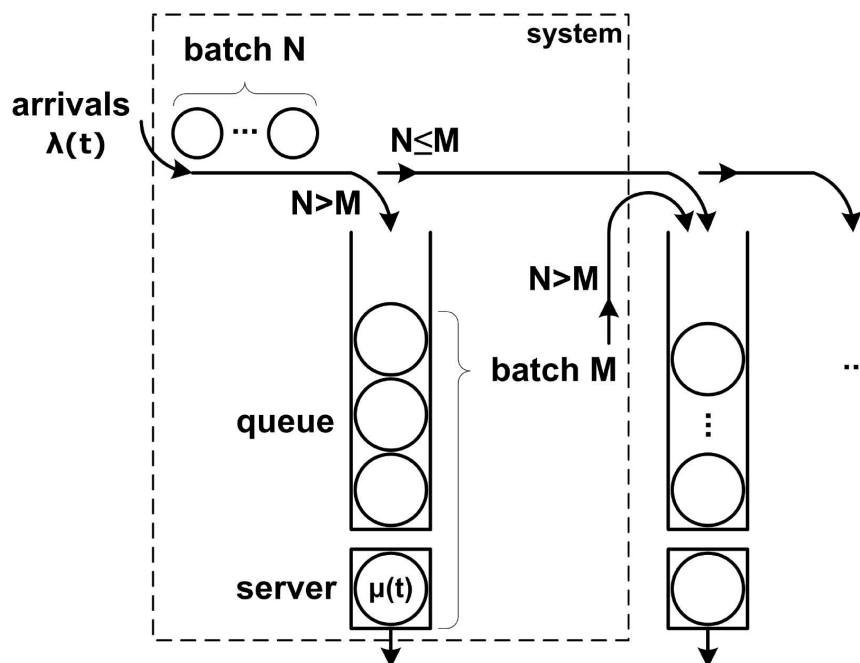


Рис. 3.1: Модель системы с политикой пропуска очереди. Можно видеть, что пакеты, удаленные из системы, (предполагается, что они не очищаются) выгружаются в следующую систему с той же политикой пропуска очереди и так далее.

что требования поступают в систему партиями, но обслуживаются по одному в любом порядке. Размер поступающей партии становится известен по ее прибытии, и в любое время также известен текущий размер системы (т.е. общее число требований в системе). Согласно политике пропуска очереди⁴ если в систему поступает пакет, размер которого превышает текущий размер системы, все текущие требования в системе удаляются и новая партия помещается в очередь. В противном случае новый пакет будет потерян и не окажет влияния на систему. В [115] авторы применили методы обратных по времени цепочек (разработанные в [104, 105, 106, 107]) для изучения производительности системы $M/M/1$ с

³Конечно, требование минимальности среднего времени отклика задания при такой политике отбрасывается.

⁴Как упоминалось выше, эта политика выгодна с точки зрения владельца системы, поскольку при применении к системам последовательно, как на рис. 3.1, она увеличивает использование серверов. Из рис. 3.1 также видно, что последовательные системы с такой политикой в некотором смысле похожи на очереди упорядоченного ввода, которые являются хорошо известными моделями конвейерных систем (открытого и замкнутого контура) с несколькими станциями разгрузки (см. [110, 101, 108]).

такой политикой пропуска очереди и обычно распределенным размером пакета, продемонстрировали результаты влияние политики на использование системы.

В этой главе исследуется система с изменяющейся во времени интенсивностью, т.е. когда интенсивность прибытия $\lambda(t)$ и/или интенсивность обслуживания $\mu(t)$ являются неслучайными функциями времени. Поскольку система является марковской, (изменяющаяся во времени) функция плотности вероятности общего числа $X(t)$ требований в системе развивается в соответствии с системой обыкновенных дифференциальных уравнений (ОДУ) - прямыми уравнениями Колмогорова. За исключением очень особых случаев (например, когда поступающая партия всегда имеет размер 1), эта система не может быть решена. Более того, если распределение размера пакета имеет бесконечную поддержку, в системе бесконечно много ОДУ, и точное аналитическое решение невозможно⁵. Здесь применён подход аппроксимации (обзор [113, Раздел 1] и [112]), который позволяет обойти трудности путем усечения системы ОДУ.

Оценки скорости сходимости опубликованы в [124], устойчивости в [134]. Результаты, связанные с оценками параметров системы, докладывались на конференции [135] и опубликованы в [134].

3.1 Описание модели

Рассмотрим очередь $M_t/M_t/1$, интенсивность которой является периодической функцией времени, и политику пропуска очереди. Требования поступают в систему партиями в соответствии с неоднородным пуассоновским процессом с интенсивностью $\lambda(t)$. Размер поступающей партии становится известен по ее прибытию и является случайной величиной с заданным распределением вероятностей $\{b_n, n \geq 1\}$, имеющей конечное среднее значение

$$\bar{b} = \sum_{k=1}^{\infty} B_k, \quad (3.1.1)$$

⁵Следует отметить, что для практических целей всегда возможно численное вычисление зависящих от времени (и предельных) плотностей. Действительно, $X(t)$ - это неоднородная непрерывная временная цепь рождения и гибели Маркова. Таким образом, всякий раз, когда его пространство состояний конечно (или каким-то образом усекается, чтобы стать конечным), можно применять различные алгоритмы униформизации (см., например, [100]).

$B_k = \sum_{n=k}^{\infty} b_n$. Принятая политика пропуска очереди подразумевает, что всякий раз, когда пакет поступает в систему, его размер, скажем, \widehat{B} , сравнивается с текущим общим количеством требований в системе, скажем, \widetilde{B} . Если $\widehat{B} > \widetilde{B}$, то все требования, которые в данный момент находятся в системе, мгновенно удаляются из нее, а весь пакет \widehat{B} помещается в очередь, и первое требование в пакете поступает на сервер. Если $\widehat{B} \leq \widetilde{B}$ новый пакет покидает систему, не оказывая на нее никакого влияния. Всякий раз, когда сервер становится свободным, одно требование из очереди (если таковое имеется) переходит на сервер⁶ и обслуживается в соответствии с экспоненциальным распределением с интенсивностью $\mu(t)$.

3.2 Получение оценок скорости сходимости

Пусть $X(t)$ - общее число требований в системе в момент времени t . Из описания системы следует, что $X(t)$ - это цепь Маркова с непрерывным временем и дискретным пространством состояний $\mathcal{X} = \{0, 1, 2, \dots, b^*\}$, где b^* - максимально возможный размер партии, т.е. индекс наибольшего ненулевого элемента в последовательности b_n . Если распределение размера пакета имеет бесконечную поддержку, то пространство состояний \mathcal{X} является счетным; в противном случае оно конечно.

Обозначим через $Q(t)$ матрицу интенсивности (бесконечно малый генератор) $X(t)$. $Q(t)$ имеет вид

$$Q(t) = \begin{pmatrix} -\lambda(t) & \lambda(t)b_1 & \lambda(t)b_2 & \dots \\ \mu(t) & -(\mu(t) + \lambda(t)B_2) & \lambda(t)b_2 & \dots \\ 0 & \mu(t) & -(\mu(t) + \lambda(t)B_3) & \dots \\ 0 & 0 & \mu(t) & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

⁶Поскольку в этой работе не изучаются характеристики времени ожидания, дисциплина обслуживания неважна, и для уверенности будем считать, что требования обслуживаются в FIFO или LIFO или в случайном порядке.

Предполагаем, что $\lambda(t)$ и $\mu(t)$ являются произвольными неслучайными функциями t , локально интегрируемыми на $[0, \infty)$ и что интенсивность поступления ограничена константой, т.е. существует $L > 0$ такое, что $\lambda(t) \leq L < \infty$ для $t \geq 0$.

Обозначим через $p_i(t) = \mathbf{P}(X(t) = i)$ вероятность того, что цепь Маркова $X(t)$ находится в состоянии i в момент времени t . Пусть $\mathbf{p}(t) = (p_0(t), p_1(t), \dots)^T$ - вектор распределения вероятностей в момент времени t . При любом надлежащем начальном условии $\mathbf{p}(0)$ вероятностная динамика цепи Маркова $X(t)$ описывается прямой системой дифференциальных уравнений Колмогорова

$$\frac{d}{dt}\mathbf{p}(t) = A(t)\mathbf{p}(t), \quad (3.2.2)$$

где $A(t) = Q^T(t)$ - транспонированная матрица интенсивности. Выбором векторных норм будет l_1 -норма, то есть $\|\mathbf{p}(t)\| = \sum_{i \in \mathcal{X}} |p_i(t)|$; операторной нормой будет та, которая индуцируется l_1 -норма для векторов строк, то есть $\|A(t)\| = \sup_{j \in \mathcal{X}} \sum_{i \in \mathcal{X}} |a_{ij}(t)|$.

Используя условие нормализации $p_0(t) = 1 - \sum_{i \geq 1, i \in \mathcal{X}} p_i(t)$, перепишем систему (3.2.2) следующим образом:

$$\frac{d}{dt}\mathbf{z}(t) = B(t)\mathbf{z}(t) + \mathbf{f}(t), \quad (3.2.3)$$

где $B(t) = (b_{ij}(t))_{i,j=1}^{\infty}$, $b_{ij}(t) = a_{ij}(t) - a_{i0}(t)$, и

$$\mathbf{f}(t) = (\lambda(t)b_1, \lambda(t)b_2, \dots)^T,$$

$$\mathbf{z}(t) = (p_1(t), p_2(t), \dots)^T,$$

$$B(t) = \begin{pmatrix} -(\mu(t) + \lambda(t)) & \mu(t) - \lambda(t)b_1 & -\lambda(t)b_1 & -\lambda(t)b_1 \cdots \\ 0 & -(\mu(t) + \lambda(t)b_2) & \mu(t) - \lambda(t)b_2 & -\lambda(t)b_2 \cdots \\ \vdots & \vdots & \vdots & \vdots \cdots \end{pmatrix}.$$

Отметим, что матрица $B(t)$ не имеет вероятностного значения. Пусть $\mathbf{z}^*(t)$ и $\mathbf{z}^{**}(t)$ являются решениями (3.2.3), соответствующими (различным) начальным условиям $\mathbf{z}^*(0)$ и $\mathbf{z}^{**}(0)$. Затем для вектора $\mathbf{y}(t) = \mathbf{z}^*(t) - \mathbf{z}^{**}(t) =$

$(y_1(t), y_2(t), \dots)^T$, который имеет координаты произвольных знаков, получаем

$$\frac{d}{dt}\mathbf{y}(t) = B(t)\mathbf{y}(t). \quad (3.2.4)$$

Таким образом, все оценки скорости сходимости к предельному режиму $X(t)$ соответствуют тем же оценкам скорости сходимости решений системы (3.2.4).

Удобнее изучать скорость сходимости, используя преобразованную матрицу $B^*(t)$ из $B(t)$, заданную $B^*(t) = TB(t)T^{-1}$, где T - верхняя треугольная матрица (1.4.19) и T^{-1} - обратная ей матрица (1.4.20).

Пусть $\mathbf{u}(t) = T\mathbf{y}(t) = (u_1(t), u_2(t), \dots)^T$. Затем, умножив (3.2.4) слева на T , получаем

$$\frac{d}{dt}\mathbf{u}(t) = B^*(t)\mathbf{u}(t), \quad (3.2.5)$$

где $\mathbf{u}(t)$ - вектор с координатами произвольных знаков, а матрица $B^*(t)$ имеет следующую структуру:

$$B^*(t) = \begin{pmatrix} -\mu(t) - \lambda(t) & \mu(t) & 0 & 0 & \dots \\ 0 & -\mu(t) - \lambda(t)B_2 & \mu(t) & 0 & \dots \\ 0 & 0 & -\mu(t) - \lambda(t)B_3 & \mu(t) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Матрица $B^*(t)$ существенно неотрицательна, т.е. все ее недиагональные элементы неотрицательны для любого $t \geq 0$. Из этого факта следует, что, если начальное условие $\mathbf{u}(s)$ неотрицательно, то любое решение $\mathbf{u}(t)$ из (3.2.5) неотрицательно для любого $0 \leq s \leq t$.

Пусть $\{d_i, i \geq 1\}$ - последовательность положительных чисел, такая, что $1 = d_1 \leq d_2 \leq \dots$ и пусть $D = \text{diag}(d_1, d_2, \dots)$ - диагональная матрица. Обозначив $\mathbf{w}(t) = D\mathbf{u}(t)$ в (3.2.5), получаем следующее уравнение

$$\frac{d}{dt}\mathbf{w}(t) = B^{**}(t)\mathbf{w}(t),$$

где $B^{**}(t) = DB(t)^*D^{-1} = (b_{ij}^{**}(t))_{i,j=1}^{\infty}$. Заметим, что $B^{**}(t)$ также неотрицательна для любого $t \geq 0$. Положим

$$\alpha_i(t) = - \sum_{j=1}^{\infty} b_{ji}^{**}(t), \quad i \geq 1,$$

и пусть $\alpha(t) = \inf_{i \geq 1} \alpha_i(t)$. Получили, что логарифмическая норма $\gamma(B^{**}(t)) = -\alpha(t)$ и

$$\|\mathbf{w}(t)\| \leq e^{-\int_s^t \alpha(u) du} \|\mathbf{w}(s)\|,$$

для любых s, t таких что $0 \leq s \leq t$.

Пусть $\delta < 1$ - положительное число, и $d_{k+1} = \delta^{-k}$, $k \geq 1$. Тогда значения α_i равны:

$$\alpha_1(t) = \lambda(t) + \mu(t),$$

$$\alpha_k(t) = \lambda(t)B_k + \mu(t)(1 - \delta), \quad k \geq 2,$$

и, следовательно:

$$\alpha(t) \geq \alpha^*(t) = (1 - \delta)\mu(t). \quad (3.2.6)$$

Наличие предельного режима $X(t)$ в рассматриваемой очереди зависит от формы распределения размера пакета $\{b_n, n \geq 1\}$. Покажем, что когда хвост распределения геометрический или более «легкий», то всегда существует предельный режим $X(t)$, тогда как для более тяжелых хвостов вопрос остается открытым. Начнем с анализа (3.2.3). Пусть $V(t, s)$ - оператор Коши для (3.2.3).

Тогда

$$\mathbf{z}(t) = V(t)\mathbf{z}(0) + \int_0^t V(t, \tau)\mathbf{f}(\tau) d\tau.$$

Пусть $\mathbf{r}(t) = DT\mathbf{z}(t)$. Тогда вместо (3.2.3) получаем:

$$\frac{d}{dt}\mathbf{r}(t) = B^{**}(t)\mathbf{r}(t) + \mathbf{f}^{**}(t), \quad (3.2.7)$$

где $\mathbf{f}^{**}(t) = DT\mathbf{f}(t)$, и

$$\mathbf{r}(t) = V^{**}(t)\mathbf{r}(0) + \int_0^t V^{**}(t, \tau)\mathbf{f}^{**}(\tau) d\tau, \quad (3.2.8)$$

где $\|V^{**}(t, s)\| \leq e^{-\int_s^t \alpha^*(u) du}$ из (3.2.6). Если существуют $N > 0$ и $a > 0$ такие, что

$$e^{-\int_s^t \alpha^*(u) du} \leq Ne^{-a(t-s)}, \quad (3.2.9)$$

для любого $0 \leq s \leq t$, то имеем экспоненциальную слабую эргодичность в соответствующей норме.

Пусть существуют $0 < q < 1$ и $C > 0$ такие, что $b_k \leq Cq^k$ для всех $k \geq 1$. Тогда

$$\begin{aligned} \|\mathbf{r}(t)\| &\leq \|V^{**}(t)\| \|\mathbf{r}(0)\| + \\ &+ \int_0^t \|V^{**}(t, \tau)\| \|\mathbf{f}^{**}(\tau)\| d\tau \leq Ne^{-at} \|\mathbf{r}(0)\| + \\ &+ \int_0^t Ne^{-a(t-\tau)} K d\tau \leq \frac{NK}{a} + o(1), \end{aligned} \quad (3.2.10)$$

где

$$\begin{aligned} \|\mathbf{f}^{**}(t)\| &= \|DT\mathbf{f}(t)\| = \\ &= C\lambda(t) \left(\frac{d_1 q}{1-q} + \frac{d_2 q^2}{1-q} + \dots \right) \leq \\ &\leq CL \left(\frac{q}{1-q} + \frac{\delta^{-1} q^2}{1-q} + \frac{\delta^{-2} q^3}{1-q} + \dots \right) = \\ &= \frac{CL\delta q}{(\delta - q)(1 - q)} = C^*L = K, \end{aligned} \quad (3.2.11)$$

для $\delta > q$.

Предположим теперь, что $b_k \rightarrow 0$ медленнее, скажем $b_k \geq k^{-s}$ для некоторых $s > 1$. Отметим, что существенная неотрицательность $B^{**}(t)$ подразумевает неотрицательность матрицы $V^{**}(t, s)$ для любого $0 \leq s \leq t$. Если $\mathbf{r}(s) \geq \mathbf{0}$, то $\mathbf{r}(t) \geq \mathbf{0}$ для любого $t \geq s$. Это следует из неотрицательности $\mathbf{f}^{**}(t)$ и $V^{**}(t, s)$ в (3.2.8).

Пусть $\mathbf{r}(0) \geq \mathbf{0}$. Тогда $\|\mathbf{r}(t)\| = \sum_{k \in \mathcal{X}} r_k(t)$ для любого $t \geq 0$. Из (3.2.7) для любого $t \geq 0$ и для любого $0 < \delta < 1$ получаем

$$\begin{aligned} \frac{d}{dt} \|\mathbf{r}(t)\| &= \sum_{k \in \mathcal{X}} \frac{d}{dt} r_k(t) \geq \|\mathbf{f}^{**}(t)\| = \\ &= \lambda(t) \left((b_1 + b_2 + b_3 + \dots) + \delta^{-1} (b_2 + b_3 + \dots) + \right. \\ &\quad \left. + \delta^{-2} (b_3 + \dots) + \dots \right) \geq \lambda(t) \sum_{k \geq 1} \delta^{-k} k^{-s} = \infty. \end{aligned}$$

Следовательно, соответствующее уравнение (3.2.7) не имеет предельного решения в соответствующей норме. Таким образом, слабая эргодичность и наличие ограничивающих характеристик гарантированы только в том случае, если хвост распределения порций по размерам является геометрическим или более «легким».

Теорема 16 *Предположим, что существуют $0 < q < 1$ и $C > 0$ такие, что $b_k \leq Cq^k$ для всех $k \geq 1$, и*

$$\int_0^\infty \mu(t) dt = +\infty. \quad (3.2.12)$$

Тогда цепь Маркова $X(t)$ слабо эргодична, и для любого начального условия $\mathbf{w}(0)$ и любого $t \geq 0$ верна следующая оценка

$$\|\mathbf{w}(t)\| \leq e^{-\int_0^t (1-\delta)\mu(u) du} \|\mathbf{w}(0)\|,$$

для всех $\delta \in (q, 1)$. Если (3.2.9) выполняется для некоторых $N > 0$ и $a > 0$, то $X(t)$ экспоненциально слабо эргодичен.

Теперь получим оценки в более естественных нормах. Отметим, что $\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 2\|\mathbf{z}^*(t) - \mathbf{z}^{**}(t)\| \leq 4\|\mathbf{w}(t)\|$ и $\|\mathbf{z}(t)\|_{1E} \leq W^{-1}\|\mathbf{w}(t)\|$ (см. [114]), где $l_{1E} = \{z(t) = (p_1(t), p_2(t), \dots)^T : \|z(t)\|_{1E} \equiv \sum_{n \in \mathcal{X}} n|p_n(t)| < \infty\}$ и $W = \inf_{k \geq 1} \frac{d_k}{k} > 0$.

Следствие 1. *В соответствии с предположениями теоремы цепь Маркова $X(t)$ имеет предельное среднее значение $\phi(t)$ и выполняются следующие*

оценки скорости сходимости:

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 4e^{-\int_0^t (1-\delta)\mu(u) du} \|\mathbf{w}(0)\|, \quad (3.2.13)$$

$$|E(t, k) - \phi(t)| \leq \frac{4}{W} e^{-\int_0^t (1-\delta)\mu(u) du} \|\mathbf{w}(0)\|. \quad (3.2.14)$$

Следствие 2. Пусть $X(t)$ - однородная цепь Маркова, т.е. $\lambda(t) = \lambda$ и $\mu(t) = \mu$. Тогда $X(t)$ сильно эргодично, и для любого начального условия $\mathbf{w}(0)$ и любого $t \geq 0$ верны следующие оценки:

$$\|\mathbf{w}(t)\| \leq e^{-(1-\delta)\mu t} \|\mathbf{w}(0)\|,$$

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 4e^{-(1-\delta)\mu t} \|\mathbf{w}(0)\|,$$

$$|E(t, k) - \phi(t)| \leq \frac{4}{W} e^{-(1-\delta)\mu t} \|\mathbf{w}(0)\|.$$

Следствие 3. Пусть интенсивность поступления $\lambda(t)$ и интенсивность обслуживания $\mu(t)$ будут 1-периодическими. Тогда предположения теоремы эквивалентны неравенству $\int_0^1 \mu(t) dt > 0$. Более того, предельное распределение вероятностей цепи Маркова $X(t)$ является 1-периодическим, предельное среднее значение также является 1-периодическим.

Из теоремы и следствий 1-3 следует, что оценки скорости сходимости справедливы для общих функций интенсивности. Если последние периодичны во времени, то предельные вероятностные характеристики $X(t)$ также являются периодическими.

3.3 Оценки устойчивости

Пусть $\bar{X}(t), t \geq 0$ - "возмущенный" процесс с бесконечно малой матрицей $\bar{Q}(t)$ и соответствующей транспонированной матрицей $\bar{A}(t)$, где матрица возмущений $\bar{A}(t) = A(t) - \bar{A}(t)$ в некотором смысле мала.

Рассматриваем оценки устойчивости основных характеристик процесса $X(t)$ при таких возмущениях, предполагая дополнительно, что процесс $X(t)$

экспоненциально эргодичен, то есть, что для некоторых положительных M , a и всех s, t , $0 \leq s \leq t$ выполняется неравенство

$$4e^{-\int_s^t (1-\delta)\mu(u) du} \leq Me^{-a(t-s)}. \quad (3.3.15)$$

Рассмотрим простейшую ситуацию, когда возмущенный процесс $\bar{X}(t)$ таков, что соответствующая матрица интенсивности $\bar{Q}(t)$ имеет ту же структуру.

Пусть возмущенные интенсивности, равные $\bar{\lambda}(t)$, $\bar{\mu}(t)$, такие что $|\lambda(t) - \bar{\lambda}(t)| \leq \hat{\epsilon}_\lambda$, $|\mu(t) - \bar{\mu}(t)| \leq \hat{\epsilon}_\mu$ для всех $t \geq 0$. Тогда

$$\|B(t)\|_{1D} \leq |\lambda(t)| + (1 + \delta)|\mu(t)| \leq (2 + \delta)L,$$

$$\|f(t)\|_{1D} \leq C^*L,$$

$$\|B(t) - \bar{B}(t)\|_{1D} \leq |\lambda(t) - \bar{\lambda}(t)| + (1 + \delta)|\mu(t) - \bar{\mu}(t)| \leq \hat{\epsilon}_\lambda + (1 + \delta)\hat{\epsilon}_\mu,$$

$$\|f(t) - \bar{f}(t)\|_{1D} \leq C^*\hat{\epsilon}_\lambda$$

для всех $t \geq 0$.

Теорема 17 Пусть $X(t)$ - $1D$ -экспоненциально слабо эргодическая цепь Маркова. Тогда возмущенный процесс $\bar{X}(t)$ также является $1D$ -экспоненциально слабо эргодичным для достаточно малых $\hat{\epsilon}_\lambda, \hat{\epsilon}_\mu$ и выполняется следующая оценка возмущения:

$$\limsup_{t \rightarrow \infty} \|p(t) - \bar{p}(t)\|_{1D} \leq \frac{MC^*(ML((1 + \delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda) + a\hat{\epsilon}_\lambda)}{a(a - M((1 + \delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda))} \quad (3.3.16)$$

где M, a определяются (3.3.15). Более того, если $W = \inf_{k \geq 1} \frac{d_k}{k} > 0$, то обе цепи Маркова $X(t)$ и $\bar{X}(t)$ имеют предельные математические ожидания и

$$\limsup_{t \rightarrow \infty} |\phi(t) - \bar{\phi}(t)| \leq \frac{MC^*(ML((1 + \delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda) + a\hat{\epsilon}_\lambda)}{Wa(a - M((1 + \delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda))}. \quad (3.3.17)$$

Доказательство. Обозначим через $V(t, s)$ оператор Коши уравнения. (3.2.3), и через $\bar{V}(t, s)$ оператор Коши соответствующего уравнения для возмущенного процесса. Следовательно,

$$\|V(t, s)\|_{1D} \leq Me^{-a(t-s)}, \quad (3.3.18)$$

в соответствии с (3.3.15). Тогда лемма 3.2.3 из [116] дает нам оценку

$$\|\bar{V}(t, s)\|_{1D} \leq M e^{-(a-M((1+\delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda))(t-s)}, \quad (3.3.19)$$

для всех $t \geq s \geq 0$; следовательно $\bar{X}(t)$ - $1D$ -экспоненциально слабо эргодичный для достаточно малых $\hat{\epsilon}_\mu, \hat{\epsilon}_\lambda$. Перепишем систему (3.2.3) как

$$\frac{d}{dt}\mathbf{z}(t) = \bar{B}(t)\mathbf{z}(t) + \mathbf{f}(t) + (B(t) - \bar{B}(t))\mathbf{z}(t), \quad (3.3.20)$$

Тогда получаем

$$\mathbf{z}(t) = \bar{V}(t, 0)\mathbf{z}(0) + \int_0^t \bar{V}(t, \tau)\mathbf{f}(\tau)d\tau + \int_0^t \bar{V}(t, \tau)(B(\tau) - \bar{B}(\tau))\mathbf{z}(\tau)d\tau \quad (3.3.21)$$

и

$$\bar{\mathbf{z}}(t) = \bar{V}(t, 0)\mathbf{z}(0) + \int_0^t \bar{V}(t, \tau)\bar{\mathbf{f}}(\tau)d\tau \quad (3.3.22)$$

Следовательно, в любой норме получаем ограниченную

$$\|\mathbf{z}(t) - \bar{\mathbf{z}}(t)\| \leq \int_0^t \|\bar{V}(t, \tau)\|(\|B(\tau) - \bar{B}(\tau)\|\|\mathbf{z}(\tau)\| + \|\mathbf{f}(\tau) - \bar{\mathbf{f}}(\tau)\|)d\tau \quad (3.3.23)$$

если начальные условия для обоих процессов одинаковы.

Из (3.2.3) получаем неравенство

$$\begin{aligned} \|\mathbf{z}(t)\|_{1D} &= \|\bar{V}(t, 0)\|_{1D}\|\mathbf{z}(0)\|_{1D} + \int_0^t \|\bar{V}(t, \tau)\|_{1D}\|\mathbf{f}(\tau)\|_{1D}d\tau \leq \quad (3.3.24) \\ &\leq M e^{-at}\|\mathbf{z}(0)\|_{1D} + \frac{M}{a}C^*L \end{aligned}$$

для всех $t \geq 0$. Следовательно, неравенство (3.3.23) подразумевает следующую оценку в $1D$ -норме:

$$\begin{aligned} \|\mathbf{z}(t) - \bar{\mathbf{z}}(t)\| &\leq \quad (3.3.25) \\ &\leq M \left(((1+\delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda) \frac{M}{a}C^*L + C^*\hat{\epsilon}_\lambda \right) \int_0^t e^{-(a-M((1+\delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda))(t-\tau)}d\tau + \\ &\quad + M \int_0^t e^{-(a-M((1+\delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda))(t-\tau)}((1+\delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda)M e^{a\tau}\|\mathbf{z}(0)\|d\tau \leq \\ &\leq \frac{M(MC^*L((1+\delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda) + aC^*\hat{\epsilon}_\lambda)}{a(a - M((1+\delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda))} + o(1), \end{aligned}$$

для любого $\mathbf{z}(0) \in l_{1D}$ (где $o(1) \rightarrow 0$ при $t \rightarrow \infty$) и верна оценка (3.3.16).

Используя хорошо известные соотношения между нормами: $\|p^* - p^{**}\| \leq 2\|z^* - z^{**}\| \leq 4\|z^* - z^{**}\|_{1D}$, где z^* и z^{**} являются соответствующими векторами для p^* и p^{**} соответственно, получаем соответствующую оценку возмущения в "естественной" l_1 -норме.

Следствие 1. Пусть $X(t)$ - неоднородная $1D$ -экспоненциально слабо эргодическая цепь Маркова. Тогда выполняется следующая оценка возмущения:

$$\limsup_{t \rightarrow \infty} \|p(t) - \bar{p}(t)\| \leq \frac{4MC^*(ML((1 + \delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda) + a\hat{\epsilon}_\lambda)}{a(a - M((1 + \delta)\hat{\epsilon}_\mu + \hat{\epsilon}_\lambda))} \quad (3.3.26)$$

Замечание. При исследовании модели Марина на устойчивость по отношению к возмущениям b_k нужны дополнительные требования на малость возмущений во взвешенной норме. В то же время просто малые изменения b_k могут все сломать. Например, если взять $\lambda = \mu = 1$, а $b_1 = 1 - \varepsilon$, $b_k = 0$ при $1 < k \leq N$, $b_k = \frac{\varepsilon}{(k-N)(k-N+1)}$ при $k > N$, то при возрастании N норма возмущения в l_1 будет ε . Однако при этом уже не выполнено условие $\sum_k B_k < \infty$.

3.4 Получение границ для некоторых параметров модели

В этом разделе рассмотрим задачи на управление.

Задача 1. Считая, что известны среднее число требований в системе $E(t, 0)$ с границами и интенсивность поступления требований $\lambda(t)$, получить границы для интенсивности обслуживания требований $\mu(t)$. Решив эту задачу, можно понять, насколько надо увеличить интенсивность обслуживания, чтобы среднее уменьшилось на $v\%$ (или, наоборот, насколько можно уменьшить интенсивность обслуживания, чтобы среднее не выросло больше чем на $v\%$).

Для её решения полезны оценки устойчивости, а именно теорема 17. Сформулируем из неё утверждение:

Утверждение 1. Пусть задано среднее число требований в системе $E(t, 0)$ с некоторыми границами, т.е.

$$\limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq h \quad (3.4.27)$$

для некоторого положительного h . Тогда границы для интенсивности обслуживания будут иметь вид

$$|\mu(t) - \bar{\mu}(t)| \leq \frac{Wa^2h}{(1 + \delta)(C^*LM^2 + MhWa)}. \quad (3.4.28)$$

Задача 2. Считая, что известны среднее число требований в системе $E(t, 0)$ с границами и интенсивность обслуживания требований $\mu(t)$, получить границы для интенсивности поступления требований $\lambda(t)$. Решив эту задачу, можно понять насколько надо уменьшить интенсивность поступления, чтобы среднее уменьшилось на $v\%$ (или, наоборот, насколько можно увеличить интенсивность поступления, чтобы среднее не выросло больше чем на $v\%$).

Утверждение 2. Пусть задано среднее число требований в системе $E(t, 0)$ с некоторыми границами, т. е.

$$\limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq h \quad (3.4.29)$$

для некоторого положительного h . Тогда границы интенсивности поступления требований можно оценить следующим образом:

$$|\lambda(t) - \bar{\lambda}(t)| \leq \frac{Wa^2h}{MC^*(ML + a) + WahM}. \quad (3.4.30)$$

Задача 3. Определить мощность сервера, которая гарантирует, что в долгосрочной перспективе среднее число требований в системе не будет превышать N^* .

Будем считать, что интенсивность обслуживания задается выражением $\mu(t) = \mu g(t)$, где $g(t)$ известно, $e^{-\int_{\tau}^t g(u) du} \leq He^{-v(t-\tau)}$, а μ (мощность сервера) можем управлять. Оценим параметры, связанные с обслуживанием, а именно мощность μ сервера, считая, что остальные параметры заданы, и среднее не должно превосходить N^* , т.е. $\limsup_{t \rightarrow \infty} E(t, 0) \leq N^*$.

По определению $E(t, 0) = \sum_{n \in \mathcal{X}} np_n(t)$. Поскольку все $p_i(t)$ неотрицательны, имеем

$$\|\mathbf{z}\|_{1D} = \sum_{n \geq 1} p_n \sum_{k \geq 1} d_k \geq \sum_{n \geq 1} d_n p_n \geq W \sum_{n \geq 1} np_n, \quad (3.4.31)$$

где $W = \inf_{k \geq 1} \frac{d_k}{k} > 0$. Тогда $E(t, 0) \leq \frac{1}{W} \|\mathbf{z}(t)\|_{1D}$.

С другой стороны, получаем следующее неравенство:

$$\begin{aligned} \|\mathbf{z}\|_{1D} &\leq \|V(t)\|_{1D} \|\mathbf{z}(0)\|_{1D} + \int_0^t \|V(t, \tau)\|_{1D} \|\mathbf{f}(\tau)\|_{1D} d\tau \leq \\ &\leq 4e^{-(1-\delta)\mu \int_0^t g(u) du} \|\mathbf{z}(0)\|_{1D} + 4C^*L \int_0^t e^{-(1-\delta)\mu \int_\tau^t g(u) du} d\tau \leq \\ &\leq 4(He^{-vt})^{(1-\delta)\mu} \|\mathbf{z}(0)\|_{1D} + 4C^*L \int_0^t \left(He^{-v(t-\tau)}\right)^{(1-\delta)\mu} d\tau \leq \frac{4C^*LH^{(1-\delta)\mu}}{v(1-\delta)\mu} \end{aligned}$$

Тогда $\limsup_{t \rightarrow \infty} E(t, 0) \leq \frac{4C^*LH^{(1-\delta)\mu}}{Wv(1-\delta)\mu} \leq N^*$ и получаем следующее утверждение.

Теорема 18 Пусть задана интенсивность поступления требований $\lambda(t)$, известны $\{b_k\}$. Интенсивность обслуживания имеет вид $\mu(t) = \mu g(t)$, причем $g(t)$ известна и выполняется неравенство $e^{-\int_\tau^t g(u) du} \leq He^{-v(t-\tau)}$, а μ можем управлять. Тогда для того, чтобы

$$\limsup_{t \rightarrow \infty} E(t, 0) \leq N^*, \quad (3.4.32)$$

достаточно, чтобы

$$\frac{H^{(1-\delta)\mu}}{\mu} \leq \frac{v(1-\delta)N^*W}{4C^*L}. \quad (3.4.33)$$

Задача 4. Определить мощность потока требований, которая гарантирует, что в долгосрочной перспективе среднее число требований в системе не будет превышать N^* .

Будем считать, что интенсивность поступления задается выражением $\lambda(t) = \lambda g(t)$, в котором $g(t)$ известна и $|g(t)| \leq H$, а λ (мощность потока поступающих требований) можем управлять. Оценим параметры, связанные с поступлением, а именно мощность потока λ , считая, что остальные параметры заданы и среднее не должно превосходить N^* , т.е. $\limsup_{t \rightarrow \infty} E(t, 0) \leq N^*$.

С одной стороны имеем (3.4.31), а с другой стороны, верно следующее неравенство:

$$\begin{aligned} \|z\|_{1D} &\leq \|V(t)\|_{1D} \|z(0)\|_{1D} + \int_0^t \|V(t, \tau)\|_{1D} \|f(\tau)\|_{1D} d\tau \leq \\ &\leq 4e^{-(1-\delta)Lt} \|z(0)\|_{1D} + 4C^* H \lambda \int_0^t e^{-(1-\delta)L(t-\tau)} d\tau \leq \frac{4C^* H \lambda}{(1-\delta)L}. \end{aligned} \quad (3.4.34)$$

Следовательно

$$E(t, 0) \leq \frac{4C^* H \lambda}{(1-\delta)LW} \leq N^* \quad (3.4.35)$$

и

$$\lambda \leq \frac{(1-\delta)LN^*W}{4C^*H}. \quad (3.4.36)$$

Теорема 19 Пусть задана интенсивность обслуживания требований $\mu(t)$. Интенсивность поступления имеет вид $\lambda(t) = \lambda g(t)$, причем $g(t)$ известна и $|g(t)| \leq H$. Тогда для того, чтобы

$$\limsup_{t \rightarrow \infty} E(t, 0) \leq N^*, \quad (3.4.37)$$

достаточно, чтобы

$$\lambda \leq \frac{(1-\delta)LN^*W}{4C^*H}. \quad (3.4.38)$$

3.5 Примеры

Во всех примерах, представленных ниже, предполагается, что распределение пакетов по размерам $\{b_k, k \geq 1\}$ является геометрическим, т.е. $b_k = (1-q)q^{k-1}$, $k \geq 1$, $0 < q < 1$. Также предполагается, что интенсивности прихода и/или перехода являются периодическими функциями. Учитывая, что $\{b_k, k \geq 1\}$ является геометрическим, независимо от параметра геометрического распределения, периодические интенсивности гарантируют существование (периодического) предельного распределения $X(t)$.

Когда интенсивности являются периодическими, но размер партии имеет общее распределение, до сих пор не смогли найти даже достаточных условий существования предельного распределения $X(t)$.

Ниже описаны результаты трёх вычислительных экспериментов: первый посвящен подтверждению полученных теоретических оценок сходимости и построению предельных характеристик системы, второй показывает влияние интенсивностей системы на реальные значения характеристик СМО, а третий (пример 2) иллюстрирует свойства политики пропуска очереди.

3.5.1 Пример 1

В этом примере показано, как точно может быть вычислена верхняя оценка скорости сходимости $X(t)$. Пусть как интенсивность прибытия, так и интенсивность обслуживания являются периодическими и равны $\lambda(t) = 1 + \sin(2\pi t)$ и $\mu(t) = 1 + \cos(2\pi t)$. Пусть $q = \frac{2}{3}$, т.е. распределение размера партии будет $b_k = \frac{2^{k-1}}{3^k}$, $k \geq 1$, т.е. средний размер партии $\sum_{k=1}^{\infty} kb_k$ равен 3. Из (3.2.10) следует, что для вычисления верхней оценки скорости сходимости необходимо выбрать, во-первых, $\delta \in (q, 1)$ и, во-вторых, $\|\mathbf{w}(0)\|$. А именно, с одной стороны, для лучшей скорости сходимости должны выбрать наименьшую возможную δ , а с другой стороны, для лучшего ограничения $\|\mathbf{w}(0)\|$ должны выбрать как можно больше δ . Пусть $\delta = \frac{5}{6}$. Таким образом, $\alpha^*(u) = \frac{1}{6}\mu(t)$ и из (3.2.9) следует, что

$$e^{-\int_s^t \alpha^*(u) du} = e^{-\frac{1}{6} \int_s^t (1 + \cos(2\pi u)) du} \leq 2e^{-\frac{1}{6}(t-s)},$$

следовательно, в правой части (3.2.9) можем положить $a = \frac{1}{6}$ и $N = 2$. Теперь рассмотрим выбор $\|\mathbf{w}(0)\|$.

Рассмотрим (3.2.11). У нас есть $L = 2$ и $C = \frac{1}{3}$. Таким образом, $C^* = \frac{10}{3}$, $K = \frac{20}{3}$ в (3.2.11) и из (3.2.10) следует, что $\limsup_{t \rightarrow \infty} \|\mathbf{r}(t)\| \leq 80$. Поскольку $\mathbf{w}(t) = DT\mathbf{y}(t)$, неравенство (3.2.10) гарантирует, что l_1 -норма предельного распределения $X(t)$ не превышает 80, т.е. $\|\mathbf{w}(0)\| \leq 80$. Таким образом (3.2.13) дает

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 160e^{-\frac{t}{6}} \quad (3.5.39)$$

для любых начальных условий $\mathbf{p}^*(0)$ и $\mathbf{p}^{**}(0)$. Например, если $t = 80 = t^*$, то правая часть (3.5.39) не превышает 10^{-3} , т.е. начиная с $t > t^*$ система

"забывает" свое начальное состояние и распределение вероятностей $X(t)$ для $t > t^*$ можно рассматривать как предельное распределение $X(t)$. Ошибка (в l_1 -норме), которая при этом допустима, не превышает 10^{-3} . Более того, поскольку предельное распределение $X(t)$ является периодическим, нам разрешено решать систему ОДУ только в интервале $[0, t^* + 1]$. Распределение вероятностей $X(t)$ в интервале $[t^*, t^* + 1]$ является оценкой (с ошибкой, не превышающей 10^{-3} в l_1 -норме) предельного распределения вероятностей $X(t)$. Следует отметить, что поскольку $b_k > 0$ для всех k , система ОДУ содержит бесконечное число уравнений. Таким образом, для того, чтобы решить ее численно, необходимо провести усечение системы. Выполним его в соответствии с методом в [84].

Верхняя оценка скорости сходимости среднего числа требований в системе $E(t, k)$ к ее предельному значению $\phi(t)$ вычисляется таким же образом. Во-первых, напомним, что $d_{k+1} = \delta^{-k}$ и поскольку $\delta = \frac{5}{6}$ было исправлено выше, то $d_{k+1} = \delta^{-k} = \left(\frac{6}{5}\right)^k$. Таким образом, $W = \inf_{k \geq 1} \frac{d_k}{k} = \frac{3}{4}$. Теперь рассмотрим (3.2.14). Таким образом, $\|\mathbf{w}(0)\| \leq 80$ и из (3.2.14) следует, что

$$|E(t, k) - \phi(t)| \leq \frac{640}{3} e^{-\frac{t}{6}} \quad (3.5.40)$$

для любого начального условия $X(0) = k$, $k \geq 0$. Таким образом, для $t > t^*$ значение $E(t, k)$ можно рассматривать как предельное значение среднего числа требований и содержит ошибку (в l_1 -норма), не превышающую 10^{-3} .

На рис. 3.2 и рис. 3.4 можно увидеть графики $p_0(t)$ и $E(t, 0)$ в интервале $[0, t^* = 80]$. ОДУ решаются с начальным условием $X(0) = 0$, т.е. система изначально пуста. Можно видеть, что полученные верхние оценки (3.5.39) и (3.5.40) не являются жесткими: системы входят в периодический режим ограничения до t^* .

Из теоремы 17 и следствия из нее получаем соответствующие оценки устойчивости

$$\limsup_{t \rightarrow \infty} \|p(t) - \bar{p}(t)\|_{1D} \leq \frac{40(44\hat{\epsilon}_\mu + 25\hat{\epsilon}_\lambda)}{1 - 22\hat{\epsilon}_\mu - 12\hat{\epsilon}_\lambda}; \quad (3.5.41)$$

$$\limsup_{t \rightarrow \infty} |\phi(t) - \bar{\phi}(t)| \leq \frac{160(44\hat{\epsilon}_\mu + 25\hat{\epsilon}_\lambda)}{3(1 - 22\hat{\epsilon}_\mu - 12\hat{\epsilon}_\lambda)}; \quad (3.5.42)$$

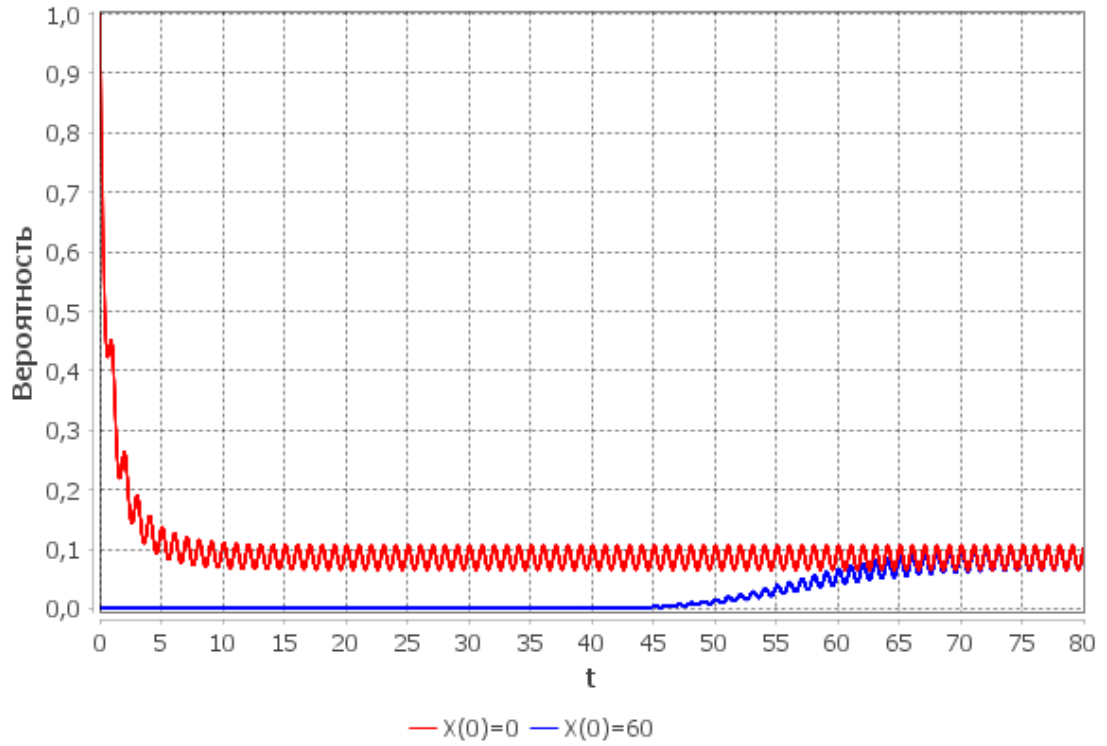


Рис. 3.2: Пример 1. Вероятность пустой очереди для $t \in [0, 80]$.

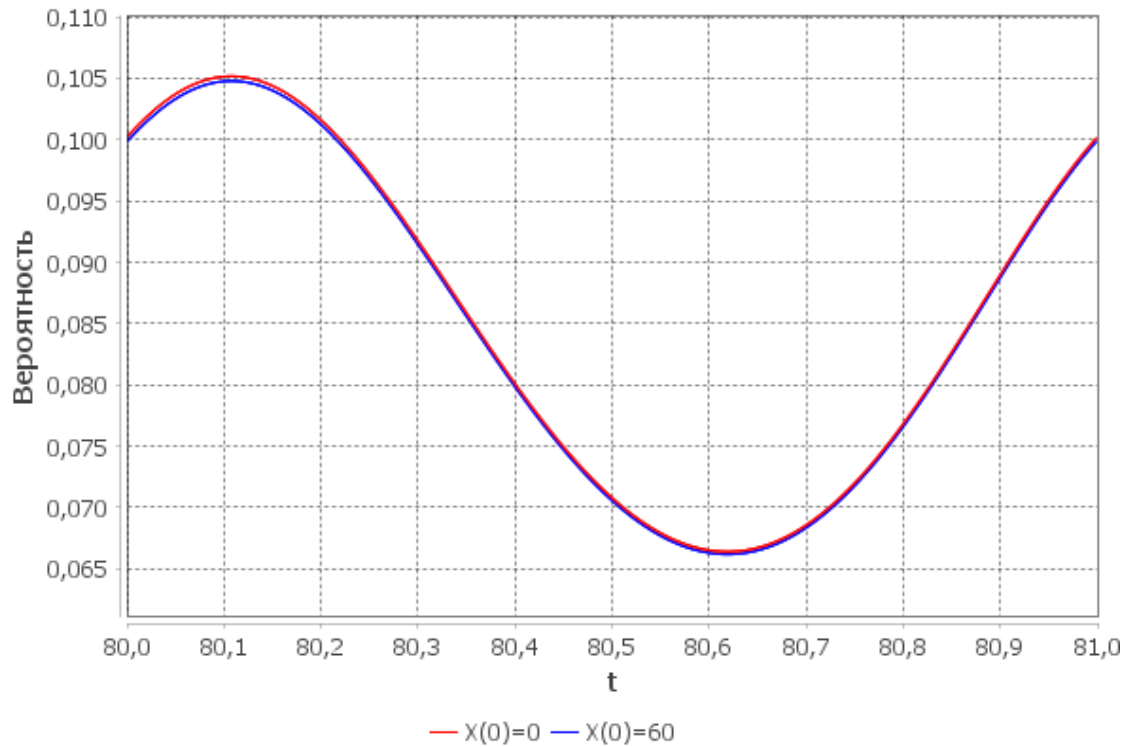


Рис. 3.3: Пример 1. Аппроксимация предельной вероятности пустой очереди для $t \in [80, 81]$.

$$\limsup_{t \rightarrow \infty} \|p(t) - \bar{p}(t)\| \leq \frac{160(44\hat{\epsilon}_\mu + 25\hat{\epsilon}_\lambda)}{1 - 22\hat{\epsilon}_\mu - 12\hat{\epsilon}_\lambda}. \quad (3.5.43)$$

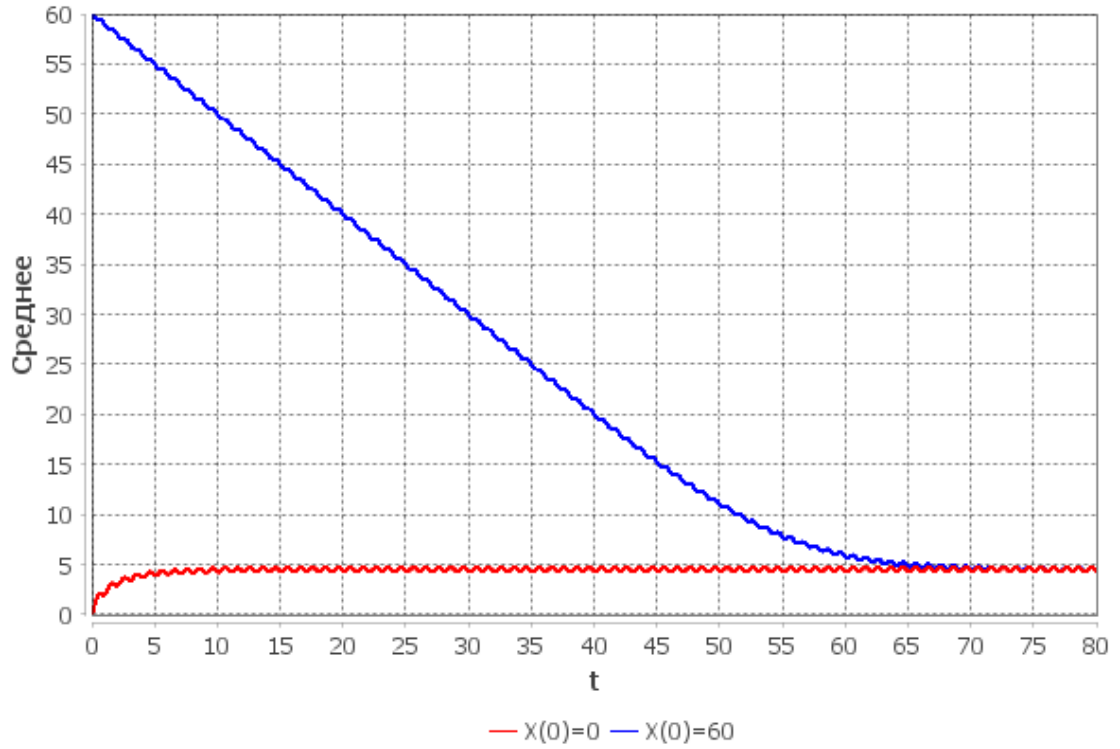


Рис. 3.4: Пример 1. Среднее $E(t, k)$ для $t \in [0, 80]$.

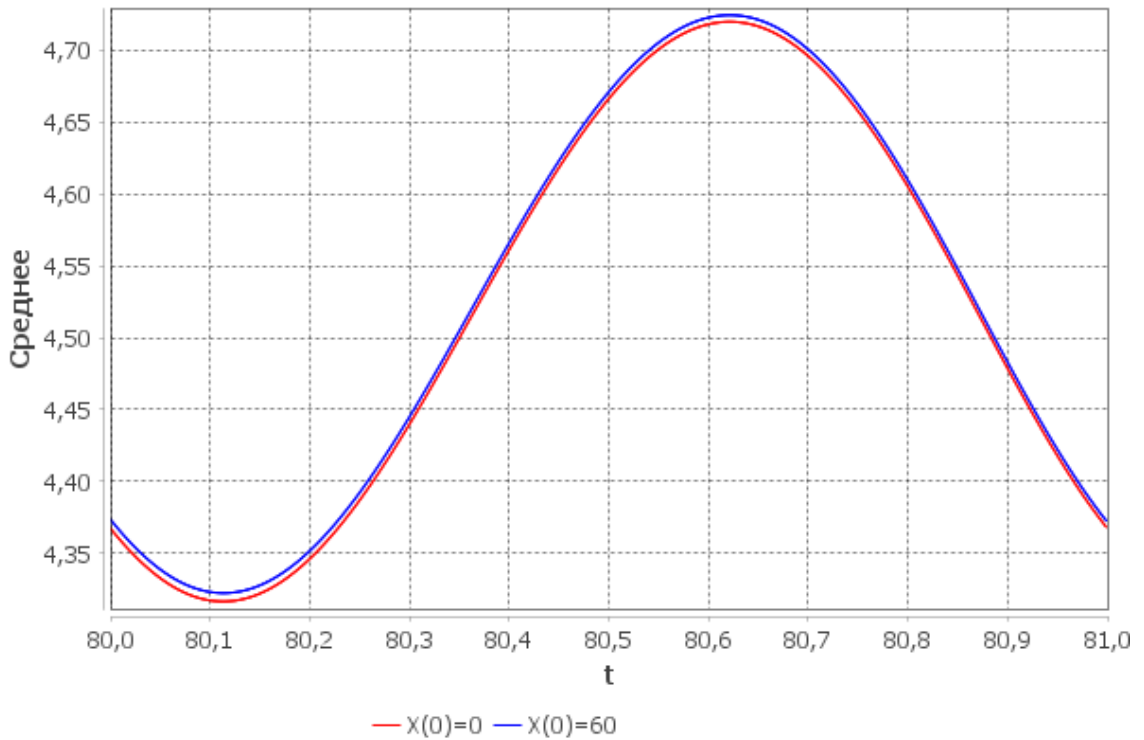


Рис. 3.5: Пример 1. Аппроксимация предельного среднего значения $E(t, k)$ для $t \in [80, 81]$.

Пусть $|E(t, 0) - \bar{E}(t, 0)| \leq h = 3$. Тогда из утверждения 1 границы интенсивности обслуживания будут иметь вид $|\mu(t) - \bar{\mu}(t)| \leq \hat{\epsilon}_\mu \leq \frac{3h}{7040+66h} \leq 10^{-2}$.

Неравенство $E(t, 0) \leq \frac{4C^*LH^{(1-\delta)\mu}}{Wv(1-\delta)\mu}$ гарантирует, что среднее число требований в системе не превышает 1437. Из теоремы 18 следует, что для того, чтобы среднее число требований в системе не превышало, например, 500, достаточно увеличить мощность сервера в 5 раз.

Пусть $|E(t, 0) - \bar{E}(t, 0)| \leq h = 5$. Тогда из утверждения 2 границы для интенсивности поступления будут иметь вид $|\lambda(t) - \bar{\lambda}(t)| \leq \hat{\epsilon}_\lambda \leq 10^{-2}$.

Неравенство $E(t, 0) \leq \frac{4C^*H\lambda}{(1-\delta)LW}$ гарантирует, что среднее число требований в системе не превышает 107. Из теоремы 19 следует, что для того, чтобы среднее число требований в системе не превышало, например, 60, достаточно уменьшить мощность потока поступающих требований в 2 раза.

На рис. 3.6 показано поведение среднего значения числа требований при $\mu(t) = 0,99(1 + \cos(2\pi t))$. На рис. 3.7 показано поведение среднего числа требований в системе, когда $\lambda(t) = 0,99(1 + \sin(2\pi t))$.

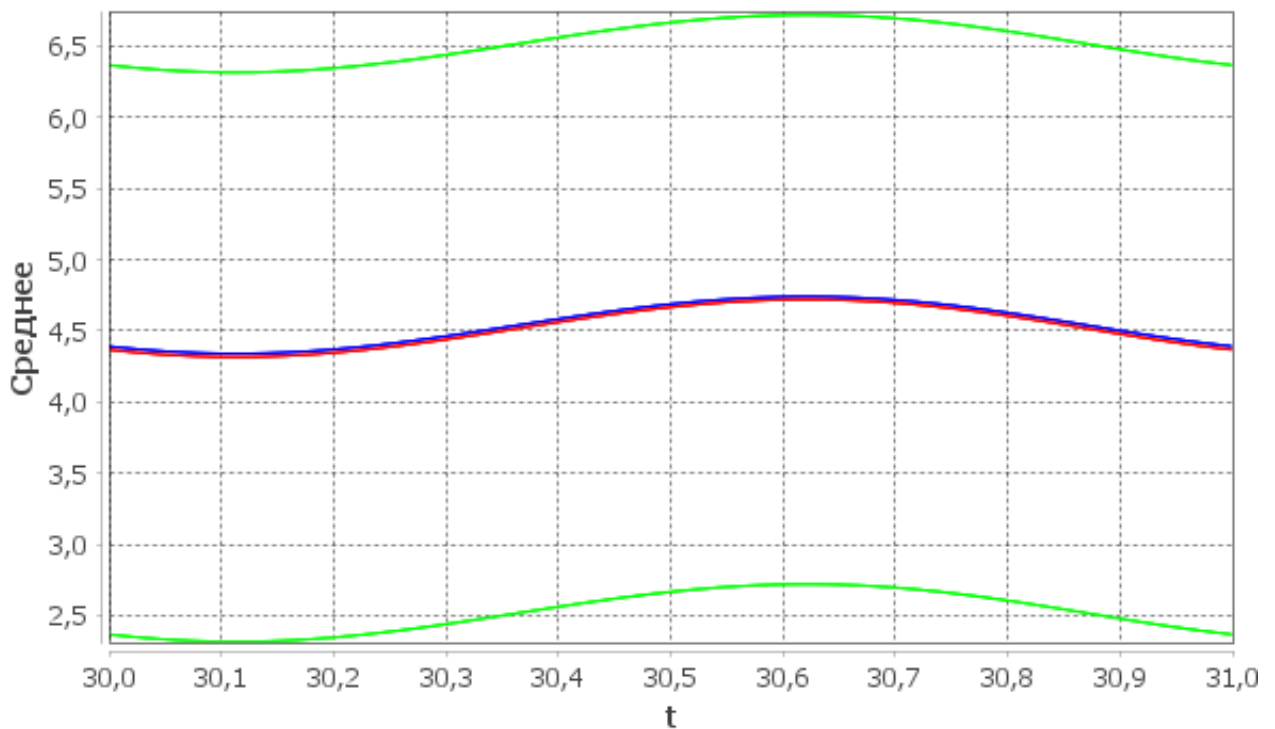


Рис. 3.6: Аппроксимация предельного среднего значения $E(t; 0)$ с границами для $\mu = 1$ (красная линия) и $\mu = 0,99$ (синяя линия) для $t \in [30, 31]$

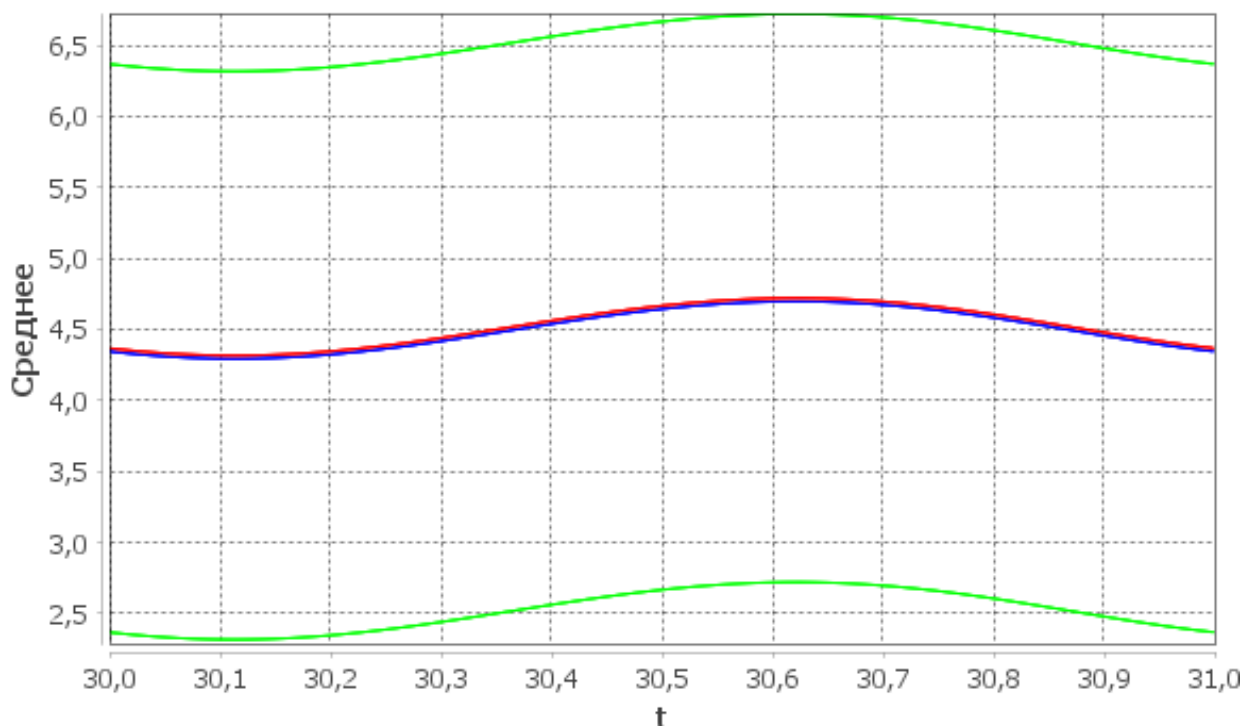


Рис. 3.7: Аппроксимация предельного среднего значения $E(t, 0)$ с границами для $\lambda = 1$ (красная линия) и $\lambda = 0,99$ (синяя линия) на $t \in [30, 31]$

3.5.2 Пример 2

Этот пример посвящен иллюстрации свойств политики пропуска очереди в случае, когда интенсивности перехода зависят от времени (чисто марковский случай изучен в [115]). Для простоты предполагаем, что только интенсивность прибытия $\lambda(t)$ зависит от времени, а интенсивность обслуживания постоянна, т.е. $\mu(t) = \mu$, $t \geq 0$. Поскольку предельный режим существует, когда распределение партий по размерам является геометрическим, никаких ограничений (дополнительных к тем, которые требуются теоремой) на интенсивность поступления $\lambda(t)$ не накладывается.

На рис. 3.8, 3.9 и 3.10 можно увидеть, как предельные вероятности $p_0(t)$, $p_1(t)$, $p_2(t)$ и $p_3(t)$ ведут себя в зависимости от интенсивности обслуживания μ . Предполагается, что $\lambda(t) = 0,8 + 0,1 \sin(2\pi t)$, $b_k = 2^{-k}$, $k \geq 1$, т.е. средний размер партии равен 2. Из-за низкой амплитуды интенсивности прибытия $\lambda(t)$ амплитуда предельных вероятностей также мала и зависит от интенсивности обслуживания.

Как и во многих других системах массового обслуживания, вероятность простоя $p_0(t)$ является одним из ключевых показателей эффективности. На рис. 3.11 и 3.12 можно увидеть поведение предельного значения $p_0(t)$, когда интенсивность обслуживания фиксирована ($\mu(t) = \mu = 1$). Из рисунков видно, что, как и ожидалось, предельная вероятность простоя стремится к 0 при увеличении размера партии или интенсивности поступления.

Наконец, также представляет интерес сравнить предельную вероятность простоя $p_0(t)$ рассматриваемой системы с политикой пропуска очереди с предельной вероятностью простоя в системе чистой блокировки, т.е. очереди $M_t/M/1/0$ при той же интенсивности прибытия $\lambda(t)$. Поскольку поступающая партия имеет средний размер $(1 - q)^{-1}$, должны изменить интенсивность обслуживания в системе блокировки на $\mu(1 - q)$. На рис. 3.13 и 3.14 можно увидеть графики $p_0(t)$ в этих двух системах, учитывая, что $\lambda(t) = 0.8 + 0.1 \sin(2\pi t)$ и $\mu = 1$.

Наблюдаем, что даже неоднородная по времени система с политикой пропуска очереди (точно так же, как однородная система, изученная в [115]) дает гораздо лучшее использование, чем система, неоднородная по времени блокировки, когда средний размер пакета большой (т.е. q близок к 1) и интенсивность прибытия высока.

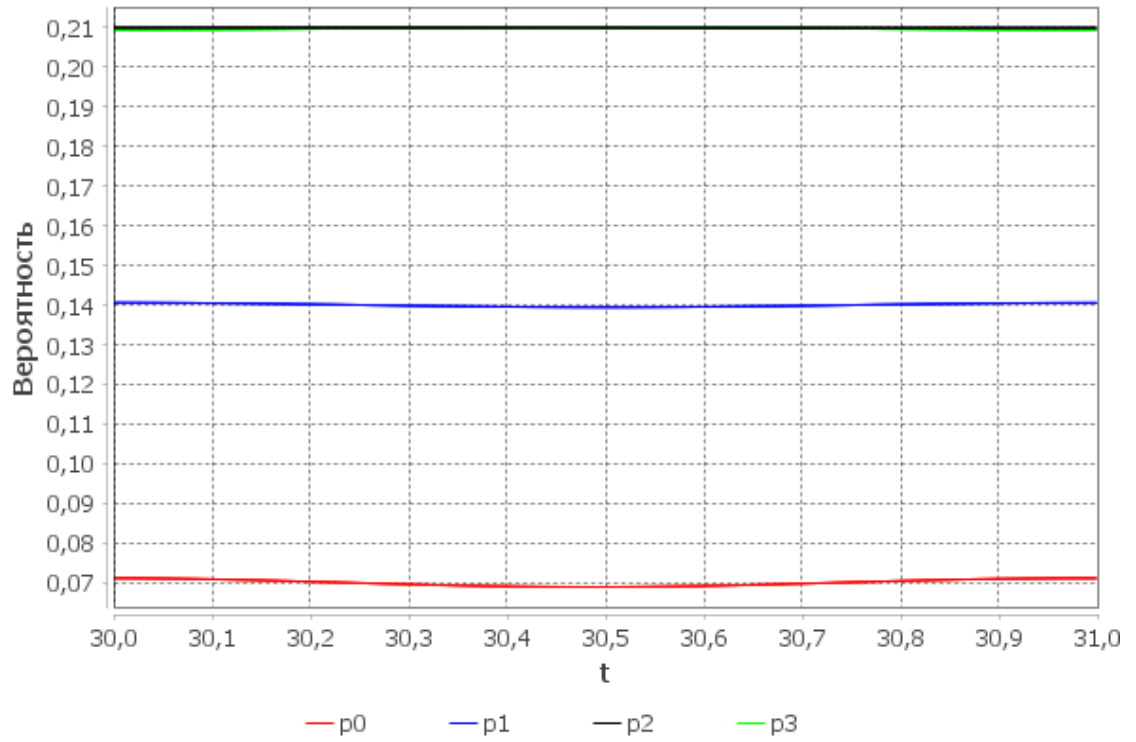


Рис. 3.8: Предельные вероятности $p_i(t)$, $0 \leq i \leq 3$, для $\mu = 0.4$.

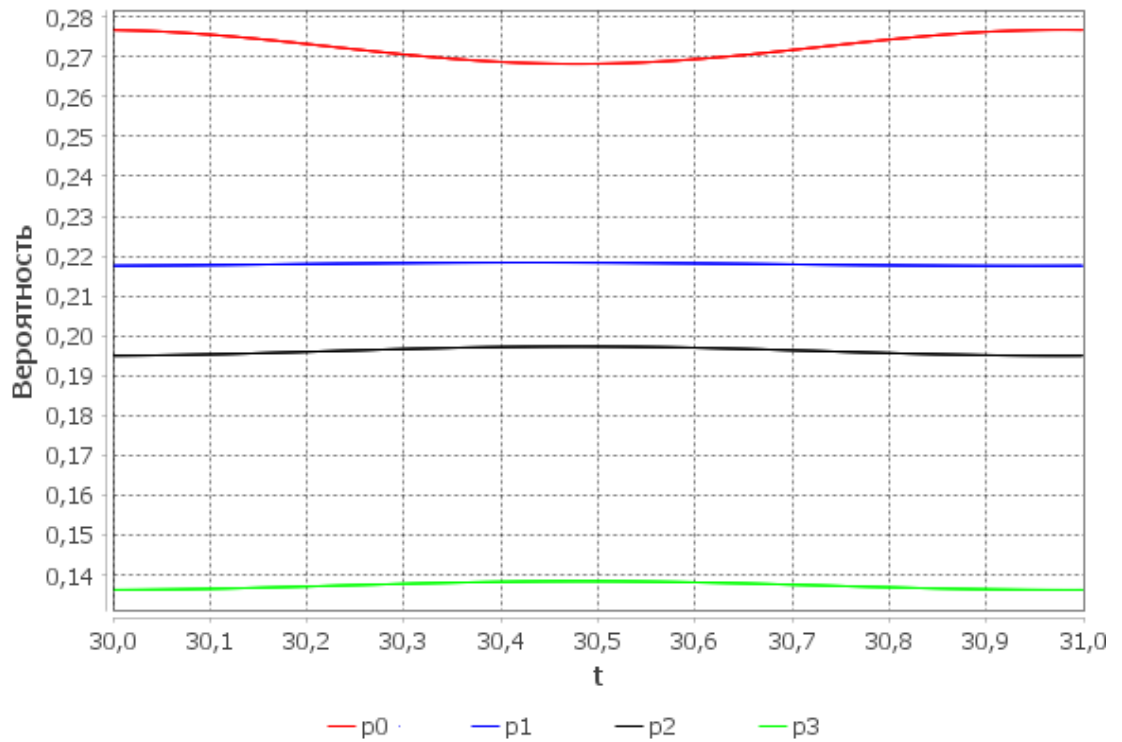


Рис. 3.9: Предельные вероятности $p_i(t)$, $0 \leq i \leq 3$, для $\mu = 1$.

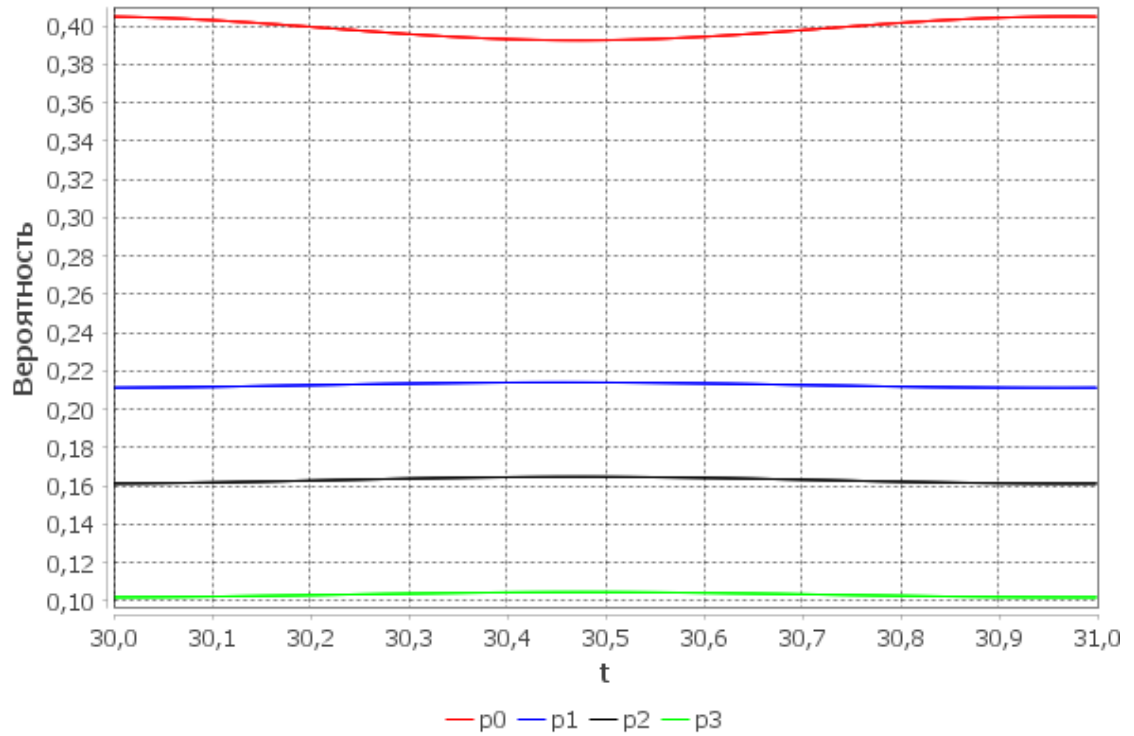


Рис. 3.10: Предельные вероятности $p_i(t)$, $0 \leq i \leq 3$, для $\mu = 1.5$.

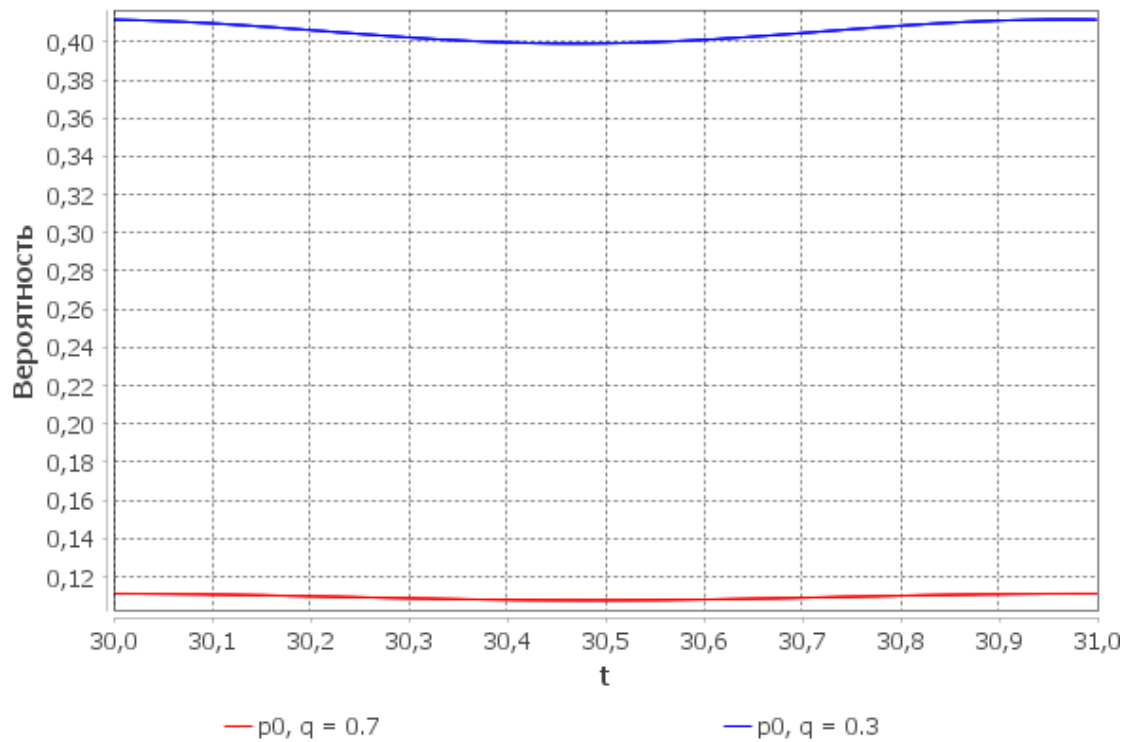


Рис. 3.11: Предельные вероятности $p_0(t)$ для $q = 0.7$ и для $q = 0.3$, $\lambda(t) = 0.8 + 0.1 \sin(2\pi t)$.

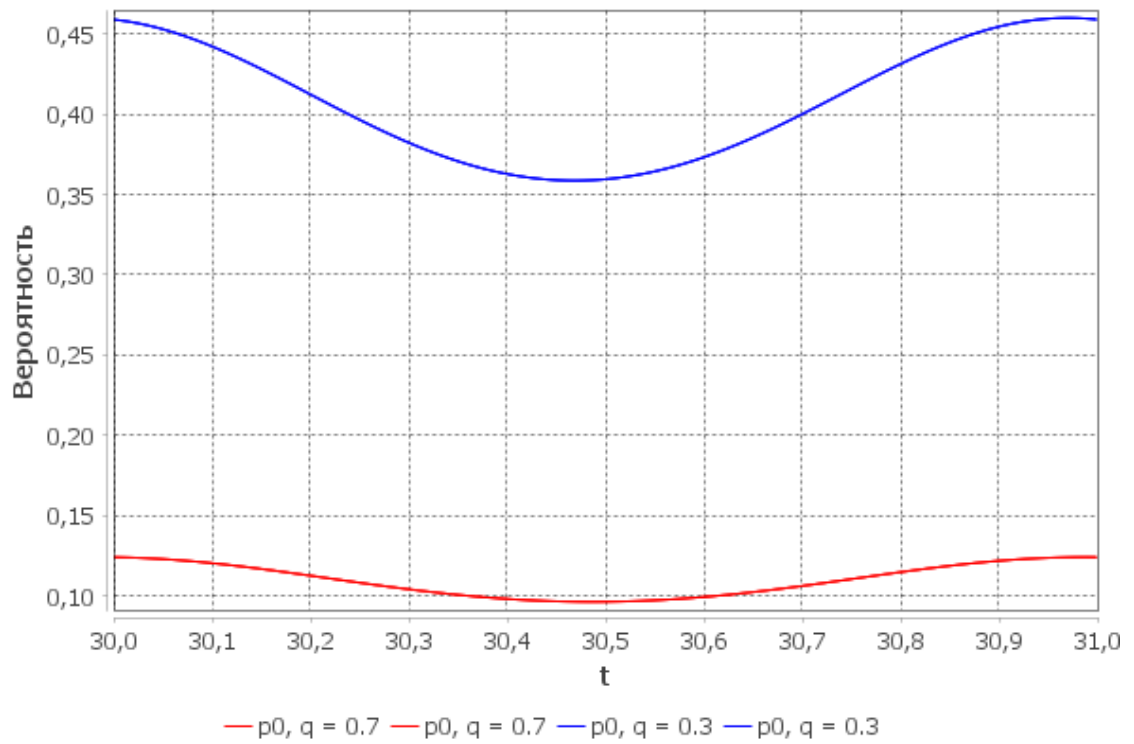


Рис. 3.12: Предельные вероятности $p_0(t)$ для $q = 0.7$ и для $q = 0.3$, $\lambda(t) = 0.8 + 0.8 \sin(2\pi t)$.

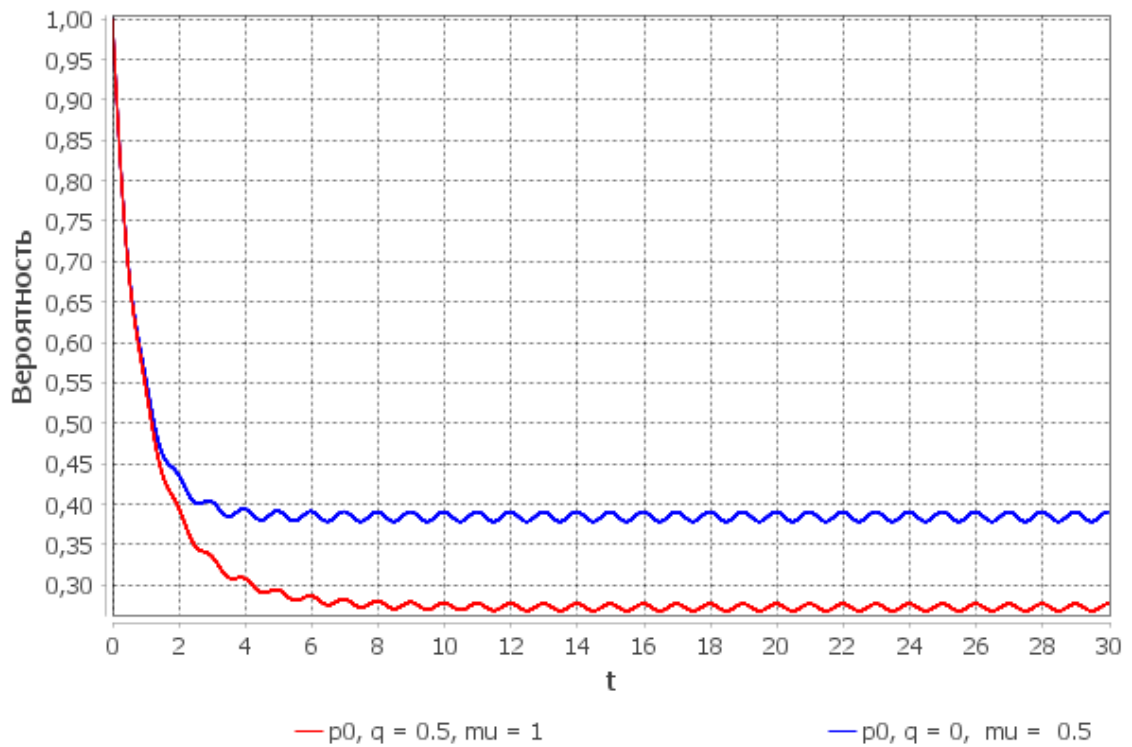


Рис. 3.13: Скорость сходимости вероятности пустой системы $p_0(t)$ для $q = 0.5, \mu = 1$ и для $q = 0, \mu = 0.5$ в интервале $[0, 30]$.

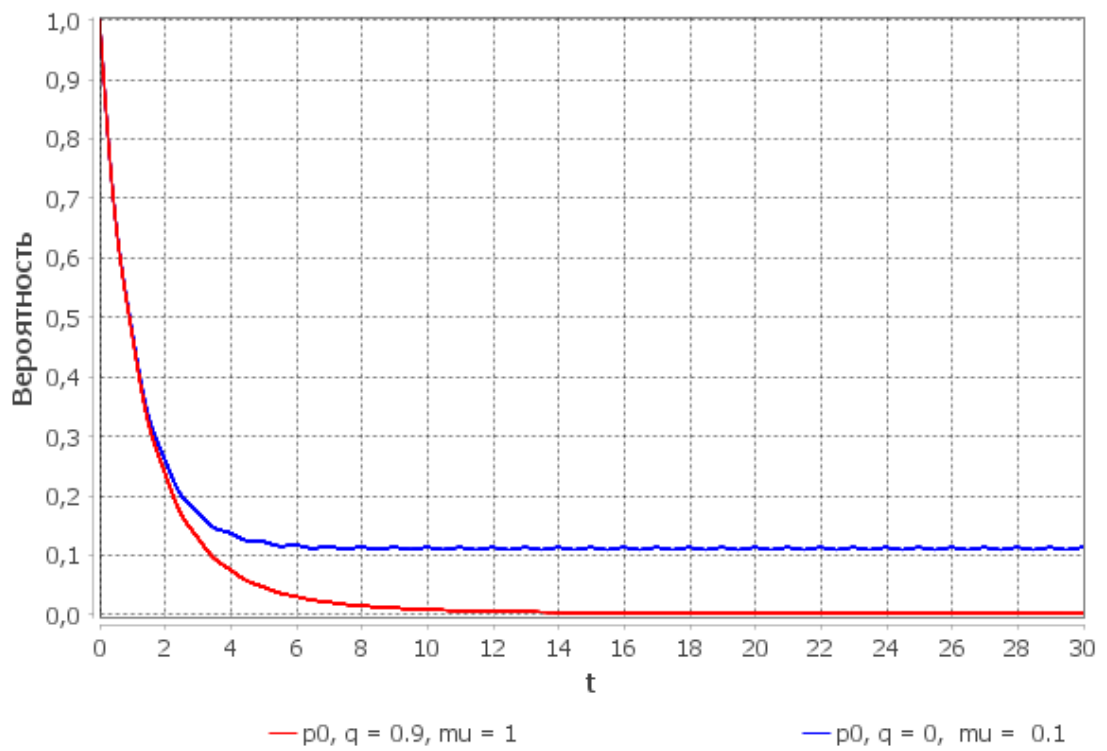


Рис. 3.14: Скорость сходимости вероятности пустой системы $p_0(t)$ для $q = 0.9, \mu = 1$ и для $q = 0, \mu = 0.1$ в интервале $[0, 30]$.

Глава 4

Исследование модели с одним сервером, специальным групповым поступлением и со специальной политикой пропуска очереди с катастрофами

Здесь рассмотрим модель из предыдущей главы в случае наличия катастрофических сбоев системы.

Получение оценок скорости сходимости в случае, когда интенсивности поступления групп требований экспоненциально убывают при увеличении размера группы опубликовано в [125]. Получение оценок скорости сходимости в случае, когда интенсивности убывают со степенной скоростью опубликовано в [126].

4.1 Описание модели

Подробное описание модели приведено в предыдущей главе, здесь же отметим, что транспонированная матрица интенсивностей $A(t) = (a_{ij}(t))_{i,j=0}^{\infty}$ имеет вид

$$A(t) = \begin{pmatrix} -\lambda(t) & \mu(t) + \gamma(t) & \gamma(t) & \gamma(t) & \dots \\ \lambda(t)b_1 - (\lambda(t)B_2 + \mu(t) + \gamma(t)) & \mu(t) & 0 & 0 & \dots \\ \lambda(t)b_2 & \lambda(t)b_2 & -(\lambda(t)B_3 + \mu(t) + \gamma(t)) & \mu(t) & \dots \\ \lambda(t)b_3 & \lambda(t)b_3 & \lambda(t)b_3 & -(\lambda(t)B_4 + \mu(t) + \gamma(t)) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix},$$

а $\gamma(t)$ – интенсивность катастрофы (одномоментной потери всех требований в системе).

4.2 Получение оценок скорости сходимости

Представим распределение $X(t)$ в виде вектора вероятности $\mathbf{p}(t)$, где $\mathbf{p}(t) = \sum_{k=0}^{b^*} \mathbf{P}(X(t) = k) \mathbf{e}_k$ для всех $t \geq 0$. Вероятностная динамика $X(t)$ описывается прямым уравнением Колмогорова $\frac{d}{dt} \mathbf{p}(t) = A(t) \mathbf{p}(t)$, которое может быть переписано в форме

$$\frac{d}{dt} \mathbf{p}(t) = A^*(t) \mathbf{p}(t) + \mathbf{g}(t), \quad t \geq 0, \quad (4.2.1)$$

где $\mathbf{g}(t) = (\gamma(t), 0, 0, \dots)^T$ и $A^*(t)$ матрица с $a_{ij}^*(t)$ равными

$$a_{ij}^*(t) = \begin{cases} a_{0j}(t) - \gamma(t), & \text{если } i = 0, \\ a_{ij}(t), & \text{если } i > 0. \end{cases} \quad (4.2.2)$$

Из-за ограничений, наложенных на $\gamma(t)$, имеем, что $\int_0^\infty \gamma(t) dt = \infty$. Таким образом, $X(t)$ не может быть нуль-эргодичным независимо от значений $\lambda(t)$ и $\mu(t)$.

Теорема 20 Пусть интенсивность катастрофы $\gamma(t)$ такая, что $\int_0^\infty \gamma(t) dt = \infty$. Тогда цепь Маркова $X(t)$ слабо эргодична и для любых двух начальных условий $\mathbf{p}^*(0)$ и $\mathbf{p}^{**}(0)$ верна оценка

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq e^{-\int_0^t \gamma(u) du} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\| \leq 2e^{-\int_0^t \gamma(u) du}, \quad t \geq 0. \quad (4.2.3)$$

Доказательство.

Легко проверить, что логарифмическая норма оператора $A^*(t)$ равна $-\gamma(t)$. Обозначим теперь через $U^*(t, s)$ оператор Коши уравнения (4.2.1). Тогда утверждение теоремы следует из неравенств $\|U^*(t, s)\| \leq e^{-\int_s^t \gamma(u) du}$ и

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq \|U^*(t, 0)\| \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|.$$

4.2.1 Получение оценок скорости сходимости в случае, когда интенсивности поступления групп требований экспоненциально убывают при увеличении размера группы.

Несмотря на то, что (4.2.3) является допустимой оценкой эргодичности для $X(t)$, это мало помогает всякий раз, когда пространство состояний $X(t)$ счетно и нужно выполнить численное решение (3.2.2). Это связано с тем, что оценка (4.2.3) находится в равномерной операторной топологии, что не позволяет использовать аналитические фреймворки (например, [81]) для нахождения правильных усечений бесконечной системы ОДУ. Для последней задачи требуются оценки эргодичности для $X(t)$ в более сильных (чем l_1) "взвешенных" нормах. Можно сказать, что при таких ограничениях каждому начальному состоянию присваивается вес, и, таким образом, процедура усечения становится чувствительной к количеству состояний. Ниже, в теореме 21, получаем такую оценку при дополнительном предположении (используемое определение см. [117]; соответствующий тест для монотонных функций можно найти в [Предложении 1] [118]). Хотя приведенная выше теорема 20 справедлива для любого распределения $\{b_n, n \geq 1\}$, это предположение существенно для теоремы 21. Для распределений с хвостами, более тяжелыми, чем геометрическое распределение, не смогли найти условия, гарантирующие существование предельного режима процесса размера очереди даже для периодических интенсивностей). Распределение размера партии $\{b_n, n \geq 1\}$ такое, что $\sum_{j=k}^{\infty} B_{j+1} \leq \bar{b} (1 - \bar{b}^{-1})^k$ для всех $k \geq 0$.

Используя условие нормализации $p_0(t) = 1 - \sum_{i \geq 1} p_i(t)$ прямая система Колмогорова $\frac{d}{dt} \mathbf{p}(t) = A(t) \mathbf{p}(t)$ может быть переписана как

$$\frac{d}{dt} \mathbf{z}(t) = A^{**}(t) \mathbf{z}(t) + \mathbf{f}(t), \quad t \geq 0, \quad (4.2.4)$$

где $\mathbf{f}(t) = (\lambda(t)b_1, \lambda(t)b_2, \lambda(t)b_3, \lambda(t)b_4, \dots)^T$, $\mathbf{z}(t) = (p_1(t), p_2(t), \dots)^T$ и

$$A^{**}(t) = \begin{pmatrix} -(\lambda(t)+\mu(t)+\gamma(t)) & \mu(t)-\lambda(t)b_1 & -\lambda(t)b_1 & -\lambda(t)b_1 & \dots \\ 0 & -(\lambda(t)B_2+\mu(t)+\gamma(t)) & \mu(t)-\lambda(t)b_2 & -\lambda(t)b_2 & \dots \\ 0 & 0 & -(\lambda(t)B_3+\mu(t)+\gamma(t)) & \mu(t)-\lambda(t)b_3 & \dots \\ 0 & 0 & 0 & -(\lambda(t)B_4+\mu(t)+\gamma(t)) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (4.2.5)$$

Пусть $\mathbf{z}^*(t)$ и $\mathbf{z}^{**}(t)$ два решения системы (4.2.4) с соответствующими двумя различными начальными состояниями $\mathbf{z}^*(0)$ и $\mathbf{z}^{**}(0)$. Тогда для вектора $\mathbf{y}(t) = \mathbf{z}^*(t) - \mathbf{z}^{**}(t) = (y_1(t), y_2(t), \dots)^T$, с произвольными координатами получаем систему

$$\frac{d}{dt}\mathbf{y}(t) = A^{**}(t)\mathbf{y}(t). \quad (4.2.6)$$

Матрица $A^{**}(t)$ в (4.2.6) может иметь отрицательные недиагональные элементы. Возьмём $d \in (1, 1 + (\bar{b} - 1)^{-1}]$ и возрастающую последовательность положительных чисел $\{\delta_n, n \geq 0\}$ на $\delta_n = d^{n-1}$. Пусть $D = \text{diag}(\delta_1, \delta_2, \dots)$ - диагональная матрица. Положим $\mathbf{w}(t) = D\mathbf{T}\mathbf{u}(t)$ в (4.2.6), где T верхнетреугольная матрица вида

$$T = \begin{pmatrix} 1 & 1 & 1 & \dots \\ 0 & 1 & 1 & \dots \\ 0 & 0 & 1 & \dots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Получаем систему уравнений

$$\frac{d}{dt}\mathbf{w}(t) = \tilde{A}(t)\mathbf{w}(t), \quad (4.2.7)$$

где матрица $\tilde{A}(t) = DTA^{**}(t)T^{-1}D^{-1}$ имеет неотрицательные недиагональные элементы:

$$\tilde{A}(t) = \begin{pmatrix} -(\lambda(t)+\mu(t)+\gamma(t)) & \frac{1}{d}\mu(t) & 0 & 0 & \dots \\ 0 & -(\lambda(t)B_2+\mu(t)+\gamma(t)) & \frac{1}{d}\mu(t) & 0 & \dots \\ 0 & 0 & -(\lambda(t)B_3+\mu(t)+\gamma(t)) & \frac{1}{d}\mu(t) & \dots \\ 0 & 0 & 0 & -(\lambda(t)B_4+\mu(t)+\gamma(t)) & \dots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (4.2.8)$$

Положим $-\alpha_k(t)$ равной сумме элементов k -го столбца матрицы $\tilde{A}(t)$ то есть

$$\begin{aligned}\alpha_1(t) &= \lambda(t) + \mu(t) + \gamma(t), \\ \alpha_k(t) &= \gamma(t) + \left(1 - \frac{1}{d}\right) \mu(t) + \lambda(t)B_k, \quad k \geq 2\end{aligned}$$

и логарифмическая норма $\tilde{A}(t)$ равна

$$-\beta(t) = \sup_i \left\{ \tilde{a}_{ii}(t) + \sum_{j \neq i} \tilde{a}_{ji}(t) \right\} = \inf_i \{ \alpha_k(t) \} = -\gamma(t) - \left(1 - \frac{1}{d}\right) \mu(t),$$

получаем

$$\|\mathbf{z}^*(t) - \mathbf{z}^{**}(t)\|_{1D} \leq e^{-\int_0^t (\gamma(u) + (1-d^{-1})\mu(u)) du} \|\mathbf{z}^*(0) - \mathbf{z}^{**}(0)\|_{1D}. \quad (4.2.9)$$

Теперь, помня, что $\mathbf{w}(t) = D\mathbf{u}(t) = DT\mathbf{y}(t)$, верхняя оценка для $\mathbf{y}(t) = \mathbf{z}^*(t) - \mathbf{z}^{**}(t)$ во взвешенной норме имеет вид

$$\|DT\mathbf{y}(t)\| \leq e^{-\int_0^t \beta(u) du} \|DT\mathbf{y}(0)\|. \quad (4.2.10)$$

Верхняя оценка для $\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|$ получена из (4.2.10). Во-первых, заметим, что $\|\mathbf{y}(t)\| \leq 2\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|$ поскольку $\mathbf{y}(t)$ является решением (4.2.6) — система с исключенным нулевым состоянием. Во-вторых, можно доказать (это показано, например, в [уравнении (18)] из [81]), что $\|\mathbf{x}\| \leq 2\|DT\mathbf{x}\|$ для любого вектора \mathbf{x} . Следовательно,

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 4e^{-\int_0^t (\gamma(u) + (1-d^{-1})\mu(u)) du} \|\mathbf{z}^*(0) - \mathbf{z}^{**}(0)\|_{1D}, \quad (4.2.11)$$

$$|E(t, k) - E(t, 0)| \leq \frac{1 + d^{k-1}}{W} e^{-\int_0^t (\gamma(u) + (1-d^{-1})\mu(u)) du}, \quad k \geq 1, \quad W = \inf_n \frac{d^n}{n+1}. \quad (4.2.12)$$

Эти результаты можно свести воедино в одной теореме.

Теорема 21 *Предположим, что распределение размера партии $\{b_n, n \geq 1\}$ с конечным средним \bar{b} такое, что $\sum_{j=k}^{\infty} B_{j+1} \leq \bar{b} (1 - \bar{b}^{-1})^k$ для всех $k \geq 0$. То-*

где, если $\int_0^\infty (\gamma(t) + (1 - d^{-1})\mu(t)) dt = +\infty$ для некоторого $d \in (1, 1 + (\bar{b} - 1)^{-1}]$, тогда цепь Маркова $X(t)$ слабо эргодична и выполняется оценка эргодичности (4.2.11).

4.2.2 Получение оценок скорости сходимости в случае, когда интенсивности убывают со степенной скоростью.

Будем предполагать, что найдется $\varepsilon > 0$ такое, что

$$\int_0^\infty (\gamma(t) - \varepsilon\lambda(t)) dt = \infty. \quad (4.2.13)$$

Если, в частности, интенсивности постоянны, то (4.2.13) выполнено при положительном γ , а если 1-периодичны, то для выполнения (4.2.13) достаточно чтобы $\int_0^1 \gamma(t) dt > 0$.

Как показано выше, выполнение условия (4.2.13) гарантирует слабую эргодичность $X(t)$ в равномерной операторной топологии и оценку

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq e^{-\int_0^t \gamma(u) du} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\| \leq 2e^{-\int_0^t \gamma(u) du}, \quad t \geq 0, \quad (4.2.14)$$

справедливую при любых начальных условиях $\mathbf{p}^*(0)$ и $\mathbf{p}^{**}(0)$.

Однако, как и в предыдущих работах, нас больше интересует не само наличие предельного режима, а возможность его построения. Для получения нужных свойств и оценок потребуются некоторые вспомогательные «взвешенные» нормы.

Положим $d_0 = 1$, и пусть $\{d_k\}$ – неубывающая последовательность, $k \geq 0$. Рассмотрим диагональную матрицу $\Lambda = \text{diag}(d_0, d_1, d_2, \dots)$.

Тогда из (4.2.1) получим уравнение

$$\frac{d}{dt} \tilde{\mathbf{p}}(t) = \tilde{A}^*(t) \tilde{\mathbf{p}}(t) + \tilde{\mathbf{g}}(t), \quad (4.2.15)$$

где $\tilde{\mathbf{p}}(t) = \Lambda \mathbf{p}(t)$, $\tilde{A}(t) = \Lambda A^*(t) \Lambda^{-1}$, а $\tilde{\mathbf{g}}(t) = \Lambda \mathbf{g}(t)$.

Далее будем оценивать логарифмическую норму оператора $\tilde{A}(t)$. Если обозначить через $-\tilde{\alpha}_k(t)$ сумму всех элементов k -того столбца матрицы $\tilde{A}(t)$,

то получим

$$\begin{aligned}\tilde{\alpha}_0(t) &\geq \gamma(t) - \lambda(t) \sum_{j=1}^{\infty} b_j \left(\frac{d_j}{d_0} - 1 \right) := \beta(t), \\ \tilde{\alpha}_k(t) &\geq \gamma(t) - \lambda(t) \sum_{j=k+1}^{\infty} b_j \left(\frac{d_j}{d_k} - 1 \right) \geq \tilde{\alpha}_0(t) = \beta(t), \quad k \geq 1.\end{aligned}$$

Из условия (3.1.1) вытекает, что для любого $\varepsilon > 0$ найдется натуральное N такое, что

$$\sum_{k \geq N} (k-1) b_k < \varepsilon. \quad (4.2.16)$$

Положим теперь $d_k = 1$, если $k < N$, и $d_k = k$ при $k \geq N$.

Тогда логарифмическая норма оператора $\tilde{A}(t)$ равна

$$-\beta^*(t) = \sup_i \left\{ \tilde{a}_{ii}(t) + \sum_{j \neq i} \tilde{a}_{ji}(t) \right\} = -\beta(t) \leq -(\gamma(t) - \varepsilon \lambda(t)), \quad (4.2.17)$$

Следовательно, вместо (4.2.14) получаем

$$\|\tilde{\mathbf{p}}^*(t) - \tilde{\mathbf{p}}^{**}(t)\| \leq e^{-\int_0^t \beta(u) du} \|\tilde{\mathbf{p}}^*(0) - \tilde{\mathbf{p}}^{**}(0)\|, \quad t \geq 0. \quad (4.2.18)$$

Далее, сравнивая соответствующие нормы и математические ожидания, получаем такое утверждение.

Теорема 22 Пусть выполнены условия (3.1.1) и (4.2.13). Тогда $X(t)$ слабо эргодичен, имеет предельное среднее, и справедливы следующие оценки скорости сходимости:

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq e^{-\int_0^t \beta(u) du} \|\tilde{\mathbf{p}}^*(0) - \tilde{\mathbf{p}}^{**}(0)\|, \quad t \geq 0, \quad (4.2.19)$$

и

$$|E(t, k) - E(t, 0)| \leq kN e^{-\int_0^t \beta(u) du}, \quad t \geq 0, \quad (4.2.20)$$

где $E(t, j)$ – математическое ожидание (среднее число требований) для $X(t)$ при условии, что $X(0) = j$.

Оценим теперь само предельное среднее. Дополнительно предположим выполнение условий

$$e^{\int_s^t \beta(u) du} \leq R e^{-a(t-s)}, \quad \gamma(t) \leq \theta, \quad (4.2.21)$$

для всех $0 \leq s \leq t$, при некоторых положительных R, a, θ .

Обозначим через $\tilde{U}(t, s)$ оператор Коши уравнения (4.2.15). Тогда получаем

$$\tilde{\mathbf{p}}(t) = \tilde{U}(t, 0)\tilde{\mathbf{p}}(0) + \int_0^t \tilde{U}(t, \tau)\tilde{\mathbf{g}}(\tau) d\tau, \quad (4.2.22)$$

откуда

$$\limsup_{t \rightarrow \infty} \tilde{\mathbf{p}}(t) \leq \limsup_{t \rightarrow \infty} \int_0^t R e^{-a(t-\tau)} \theta d\tau \leq \frac{R\theta}{a}. \quad (4.2.23)$$

А тогда имеем

Следствие 2 Пусть выполнены условия (3.1.1) и (4.2.21). Тогда при любом k справедлива следующая оценка:

$$\limsup_{t \rightarrow \infty} E(t, k) \leq \frac{NR\theta}{a}. \quad (4.2.24)$$

4.2.3 Аппроксимация усечениями

Рассмотрим вопрос о построении предельного режима и предельного среднего с помощью аппроксимации усеченными процессами. Получение не зависящих от времени оценок при таких аппроксимациях описано в [81, 84].

Аналогично этим работам, будем отождествлять конечные векторы и счетные векторы с теми же ненулевыми координатами. Рассмотрим "усеченную" матрицу (для краткости зависимость от t не записываем)

$$A_K^* = \begin{pmatrix} -\lambda B_1^* - \gamma & \mu & 0 & 0 & \dots & 0 \\ \lambda b_1 & -(\lambda B_2^* + \mu + \gamma) & \mu & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \lambda b_K & \lambda b_K & \dots & \dots & \dots & -(\lambda B_K^* + \mu + \gamma) \end{pmatrix},$$

где $B_k^* = \sum_{j=k}^K b_j$.

Запишем аналогичную (4.2.1) систему для усеченного процесса в виде:

$$\frac{d}{dt} \mathbf{p}_K(t) = A^*(t) \mathbf{p}_K(t) + \mathbf{g}(t) + (A_K^*(t) - A^*(t)) \mathbf{p}_K(t), \quad t \geq 0. \quad (4.2.25)$$

Тогда

$$\begin{aligned} \mathbf{p}_K(t) &= U(t, 0) \mathbf{p}(0) + \int_0^t U(t, \tau) \mathbf{g}(\tau) d\tau + \\ &+ \int_0^t U(t, \tau) (A_K^*(\tau) - A^*(\tau)) \mathbf{p}_K(\tau) d\tau = \\ &= \mathbf{p}(t) + \int_0^t U(t, \tau) (A_K^*(\tau) - A^*(\tau)) \mathbf{p}_K(\tau) d\tau, \end{aligned} \quad (4.2.26)$$

где $U(t, s)$ - оператор Коши уравнения (4.2.1).

Следовательно, в любой норме справедлива оценка:

$$\|\mathbf{p}(t) - \mathbf{p}_K(t)\| \leq \int_0^t \|U(t, \tau)\| \|(A_K^*(\tau) - A^*(\tau)) \mathbf{p}_K(\tau)\| d\tau. \quad (4.2.27)$$

Рассмотрим норму $\|\mathbf{x}\|_\Lambda = \|\Lambda \mathbf{x}\|$, тогда $\|\tilde{U}(t, s)\| = \|U(t, s)\|_\Lambda \leq R e^{-a(t-s)}$.

Для оценки второго множителя под знаком интеграла в (4.2.27) отметим, что в левом верхнем квадрате матрицы $A_K^* - A^*$ (то есть если оба индекса не превосходят K) ненулевыми являются только диагональные элементы, каждый из которых равен $-\lambda B_K$.

Значит, $(A_K^*(\tau) - A^*(\tau)) \mathbf{p}_K(\tau) = -\lambda(\tau) B_K (p_0(\tau), \dots, p_K(\tau))^T$. А тогда, предполагая, что $\lambda(t) \leq \theta$ при всех t , получаем $\|(A_K^*(\tau) - A^*(\tau)) \mathbf{p}_K(\tau)\| \leq B_K \theta \sum_{k \leq K} d_k p_k(\tau) \leq B_K N \theta$.

Тогда правая часть в (4.2.27) в Λ норме не превосходит $\frac{B_K N R \theta}{a}$ и получаем следующее утверждение.

Теорема 23 Пусть выполнены условия (3.1.1) и (4.2.21). Тогда при $X(0) = 0$ справедливы следующие оценки:

$$\|\mathbf{p}(t) - \mathbf{p}_K(t)\| \leq \frac{B_K N R \theta}{a} \rightarrow 0 \text{ при } k \rightarrow \infty, \quad (4.2.28)$$

u

$$|E(t, 0) - E_K(t, 0)| \leq \frac{B_K N^2 R \theta}{a} \rightarrow 0 \text{ при } k \rightarrow \infty. \quad (4.2.29)$$

4.3 Примеры

С помощью вычислительных экспериментов подтвердим теоретические выводы.

4.3.1 Пример 1

Рассмотрим пример, показывающий зависимость от t величин $p_0(t)$ и $E(t, k)$. Предположим, что $b_k = \frac{1}{3} \left(\frac{2}{3}\right)^{k-1}$, $\lambda(t) = 9(1 + \sin 2\pi t)$, $\mu(t) = 8(1 + \cos 2\pi t)$ и $\gamma(t) = 1$, т.е. интенсивность катастрофы постоянна и средний размер \bar{b} прибывающей партии не превосходит 3. Можно проверить, что $d = \frac{3}{2}$ удовлетворяет условиям теоремы 21. Из (4.2.11) и (4.2.12) получаем оценки

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 4e^{-\frac{5}{3}t} \|\mathbf{z}^*(0) - \mathbf{z}^{**}(0)\|_{1D}, \quad (4.3.30)$$

$$|E(t, k) - E(t, 0)| \leq \frac{1 + \left(\frac{3}{2}\right)^{k-1}}{\frac{9}{8}} e^{-\frac{5}{3}t}, \quad k \geq 0. \quad (4.3.31)$$

На рисунке 4.1 показано, как $p_0(t)$ ведет себя при увеличении t , а на рисунке 4.2 показано его предельное значение. Если $t \geq 60$, то правая часть (4.3.30) не превышает $3 \cdot 10^{-2}$, т.е. начиная с момента $t = 60 = t^*$ система "забывает" свое начальное состояние и распределение $X(t)$ для $t > t^*$ можно рассматривать как предельное. Более того, поскольку предельное распределение $X(t)$ является периодическим, достаточно решить (численно, (следует отметить, что, поскольку $b_k > 0$ для всех k , система ОДУ содержит бесконечное число уравнений. Таким образом, для того, чтобы решить ее численно, нужно ее усечь. Выполняем это усечение в соответствии с методом в [119])). систему ОДУ только в интервале $[0, t^* + T]$, где T - наименьшее общее кратное периодов $\lambda(t)$ и $\mu(t)$, т.е. $T = 1$. Распределение вероятностей $X(t)$ в интервале $[t^*, t^* + T]$ является оценкой (с ошибкой, не превышающей $3 \cdot 10^{-2}$ в l_1 -норме) предельного распределения вероятностей $X(t)$. Верхняя оценка скорости сходимости условного среднего числа

клиентов в системе $E(t, k)$ приведена в (4.3.31). Если $t \geq t^*$, то правая часть не превышает $0,3$, т.е. начиная с момента $t = t^*$ система "забывает" свое начальное состояние и значение $E(t, k)$ можно рассматривать как предельное значение от среднего числа требований с погрешностью не более $0,3$. Скорость сходимости $E(t, k)$ и поведение ее предельного значения можно увидеть на рисунках 4.3 и 4.4. Полученные верхние оценки не являются жесткими: система переходит в периодический режим ограничения до момента $t = t^*$.

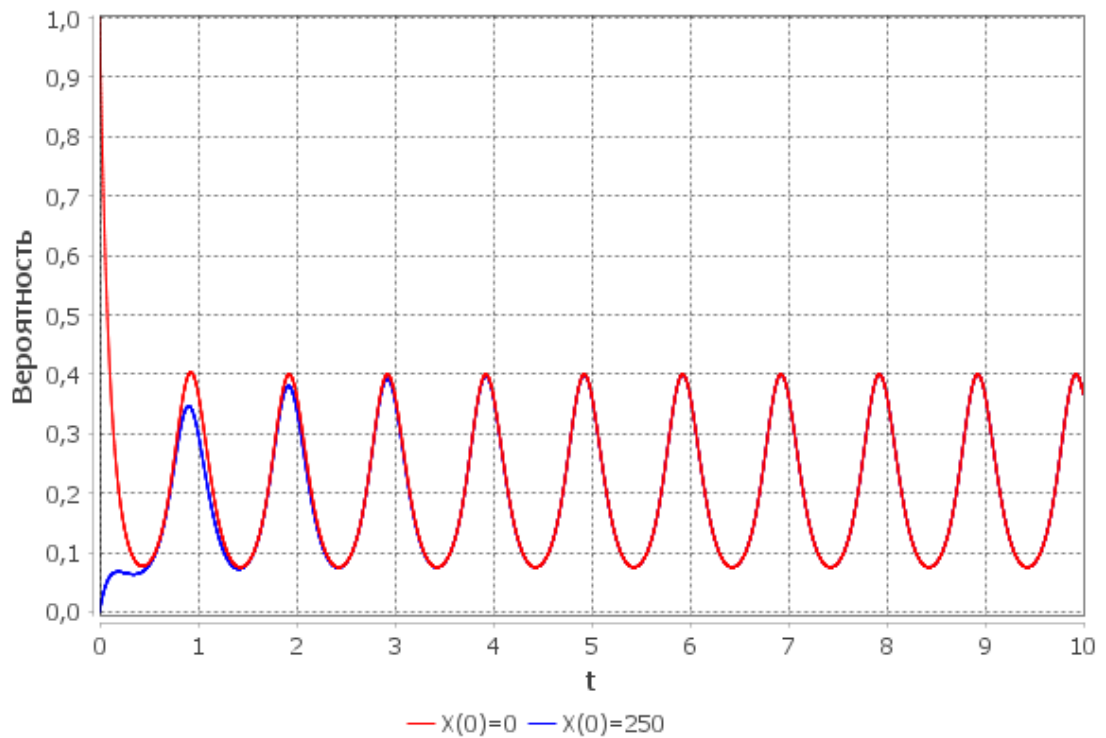


Рис. 4.1: Пример 1. Скорость сходимости вероятности пустой очереди $p_0(t)$ в интервале $[0, 10]$.

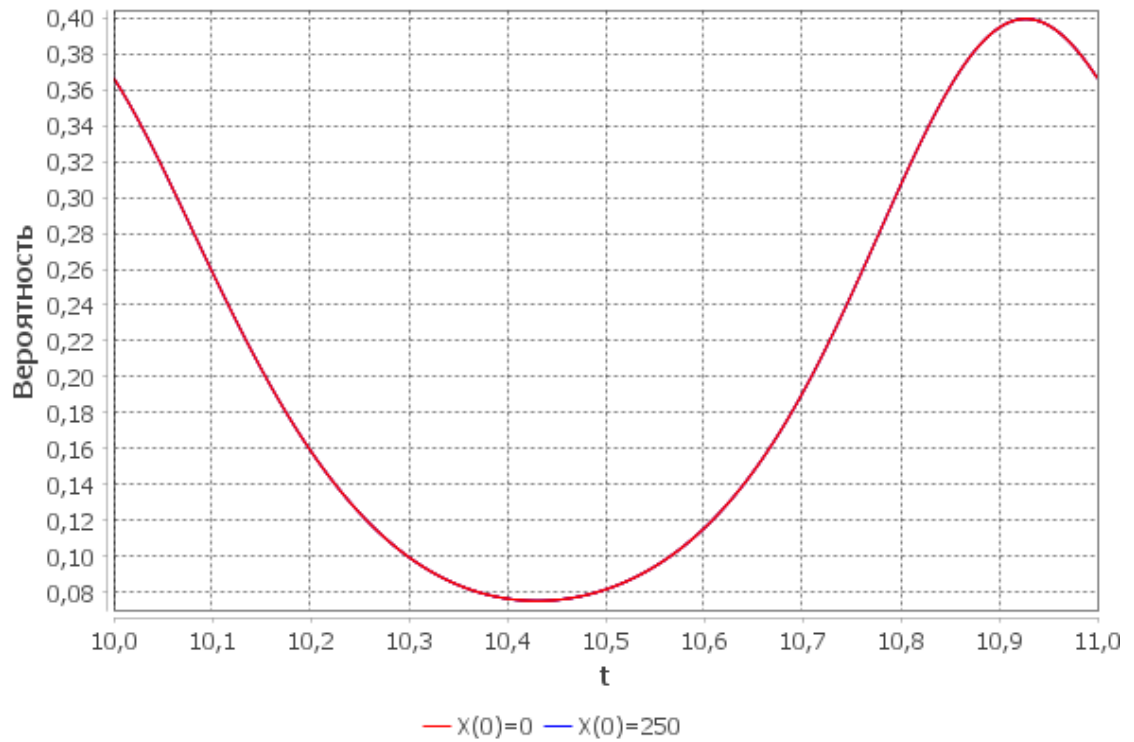


Рис. 4.2: Пример 1. Предельная вероятность $p_0(t)$ пустой очереди.

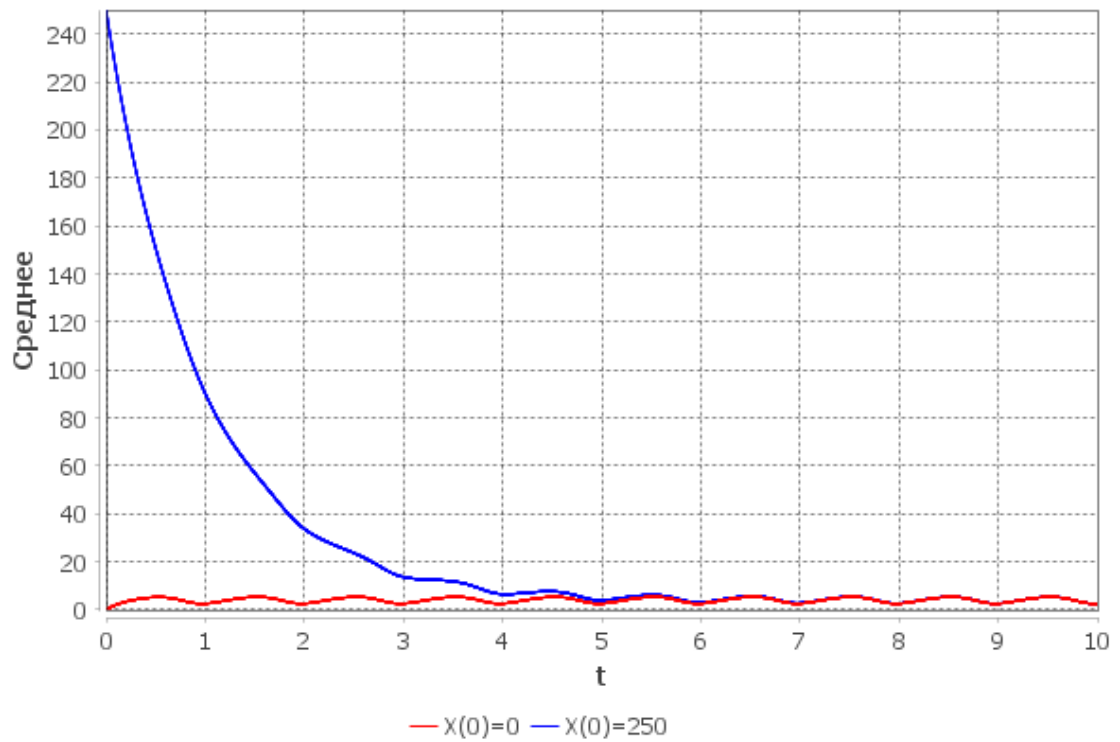


Рис. 4.3: Пример 1. Скорость сходимости среднего $E(t, k)$ числа требований в системе для $t \in [0, 10]$.

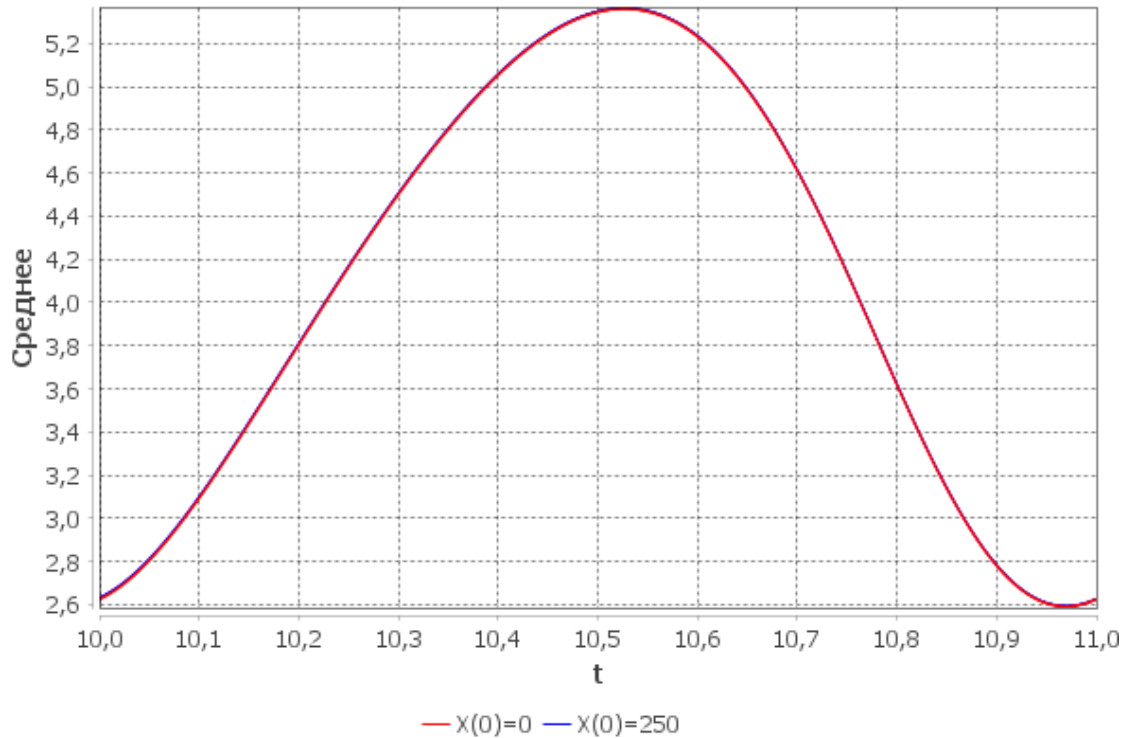


Рис. 4.4: Пример 1. Аппроксимация предельного среднего значения $E(t, k)$ для $t \in [19, 20]$.

4.3.2 Пример 2

Рассмотрим описанную модель с интенсивностями $\gamma = 1$, $\lambda(t) = 1 + \sin 2\pi t$, $\mu(t) = 1 + \cos 2\pi t$, предполагая при этом, что $b_k = \frac{4}{k(k+1)(k+2)}$, то есть убывание имеет степенной характер.

На приведенных рисунках показано поведение вероятности отсутствия требований в системе $p_0(t)$ и среднего числа требований в системе $E(t, k)$ для усеченных процессов с числом состояний 100, 200, 300. При этом можно отметить, что погрешность вектора вероятностей состояний при усечениях, соответствующих $K = 100$ и $K = 200$ получается $1.2 \cdot 10^{-2}$ и $1.2 \cdot 10^{-3}$ соответственно, а для средних – $7.2 \cdot 10^{-2}$ и $7.2 \cdot 10^{-3}$ соответственно.

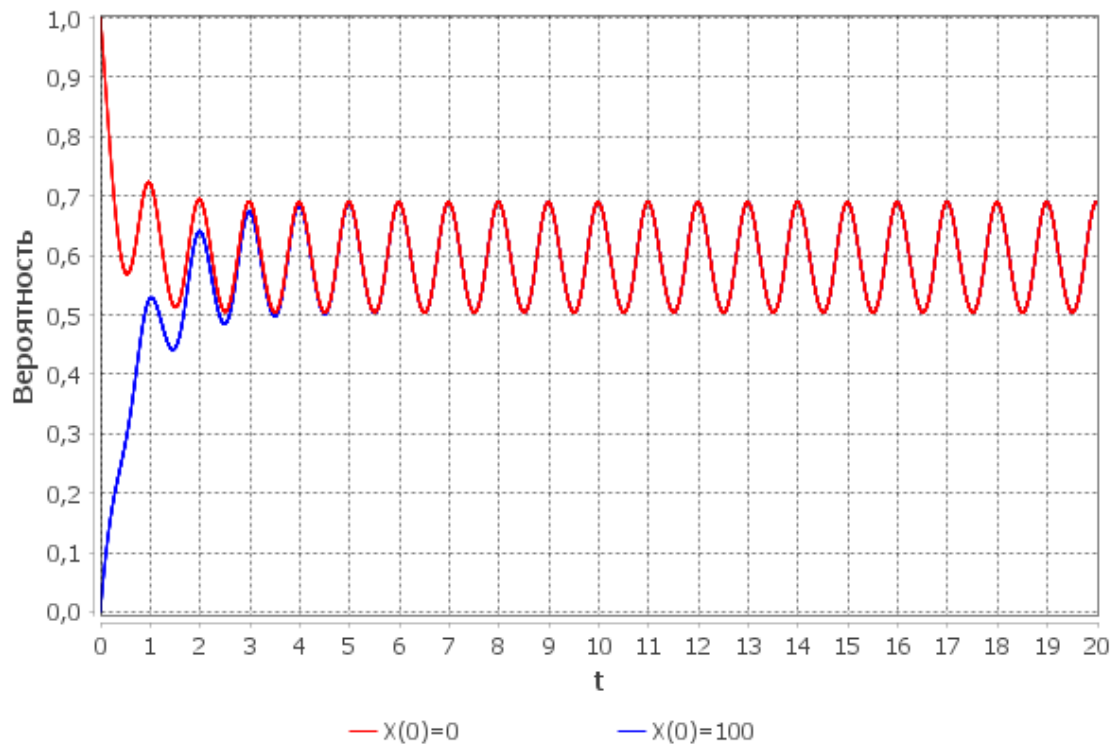


Рис. 4.5: Пример 2. Поведение вероятности $p_0(t)$ для усеченного процесса, 100 состояний, интервал $[0, 20]$.

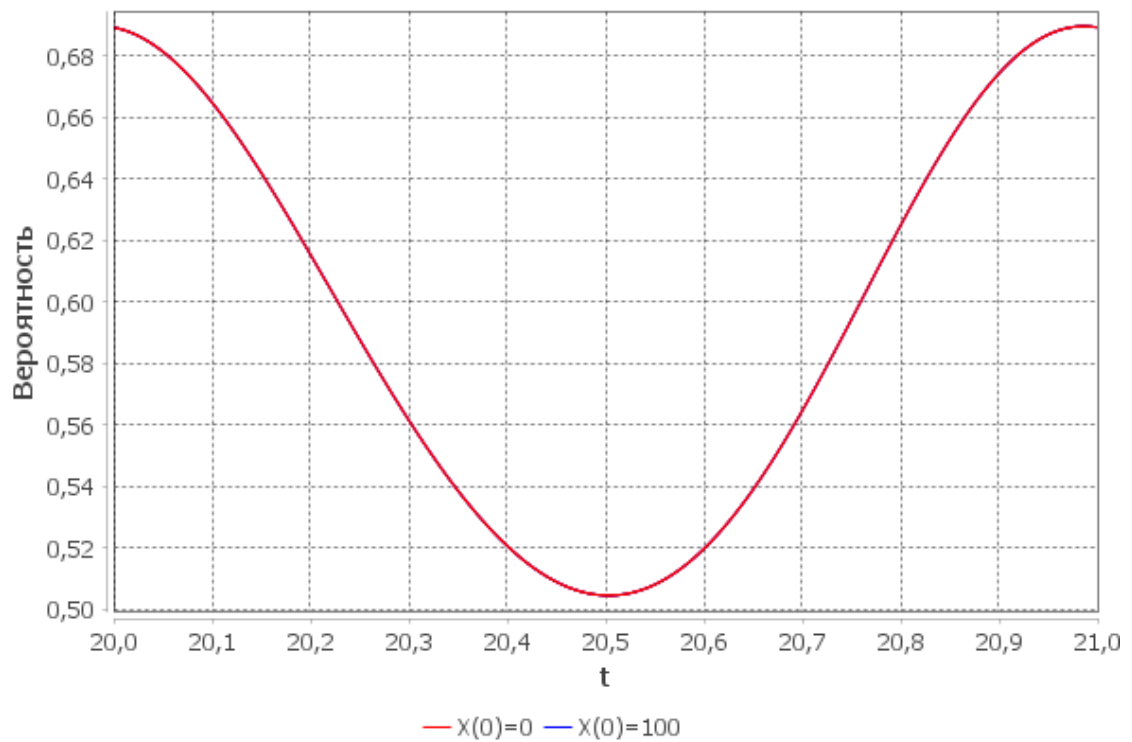


Рис. 4.6: Пример 2. Предельная вероятность $p_0(t)$ для усеченного процесса, 100 состояний, интервал $[20, 21]$.

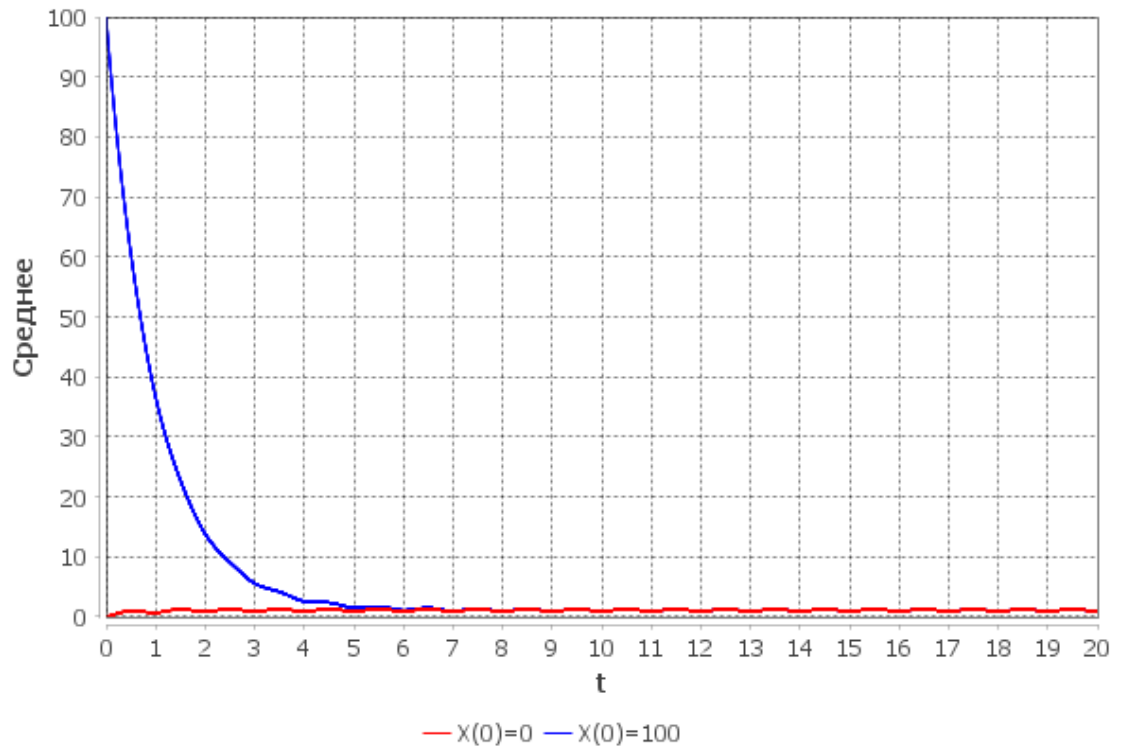


Рис. 4.7: Пример 2. Поведение среднего $E(t, k)$ для усеченного процесса, 100 состояний, интервал $[0, 20]$.

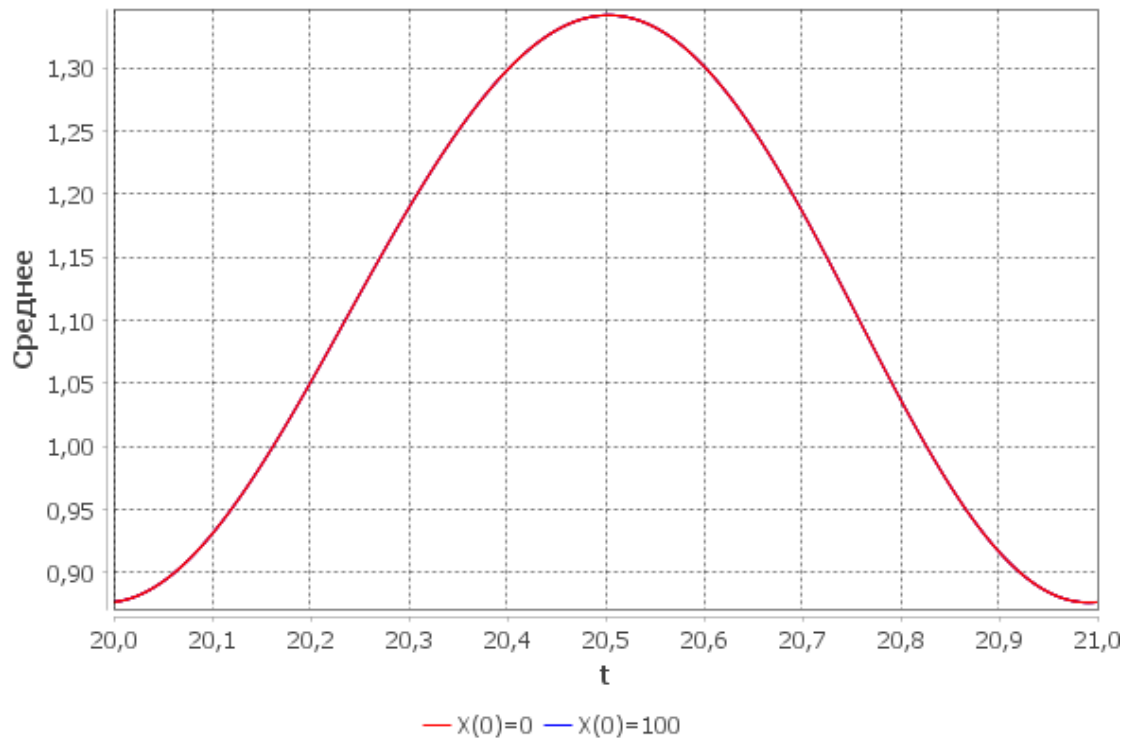


Рис. 4.8: Пример 2. Предельное среднее $E(t, k)$ для усеченного процесса, 100 состояний, интервал $[20, 21]$.

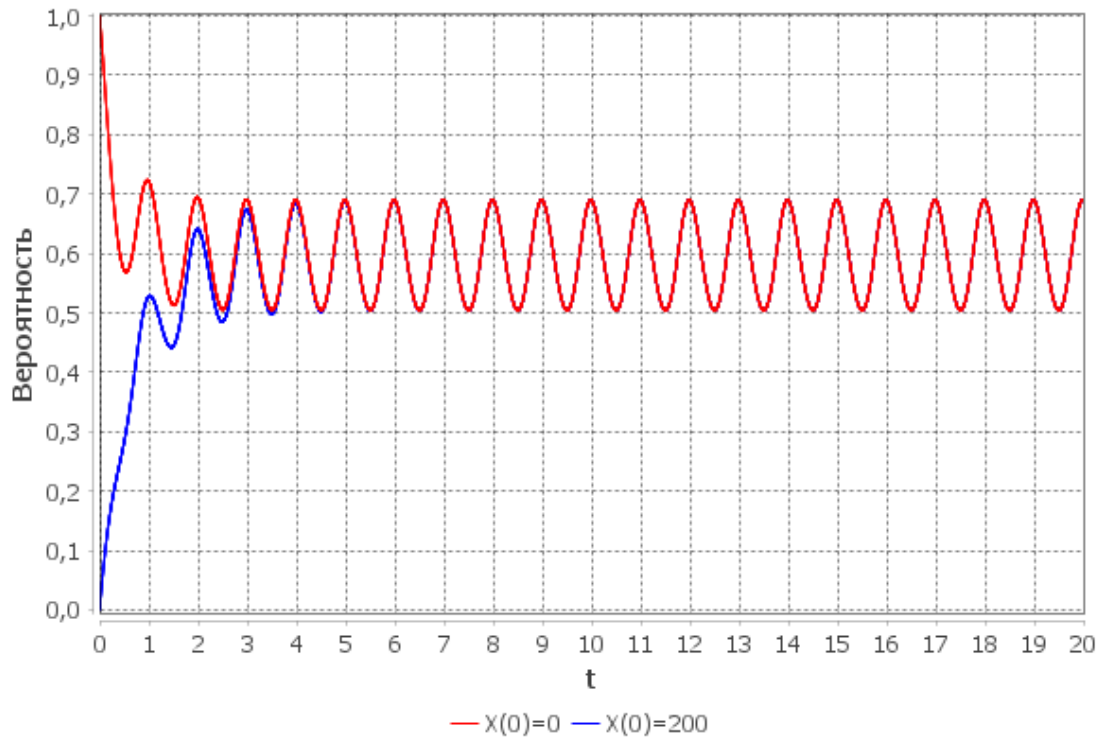


Рис. 4.9: Пример 2. Поведение вероятности $p_0(t)$ для усеченного процесса, 200 состояний, интервал $[0, 20]$.

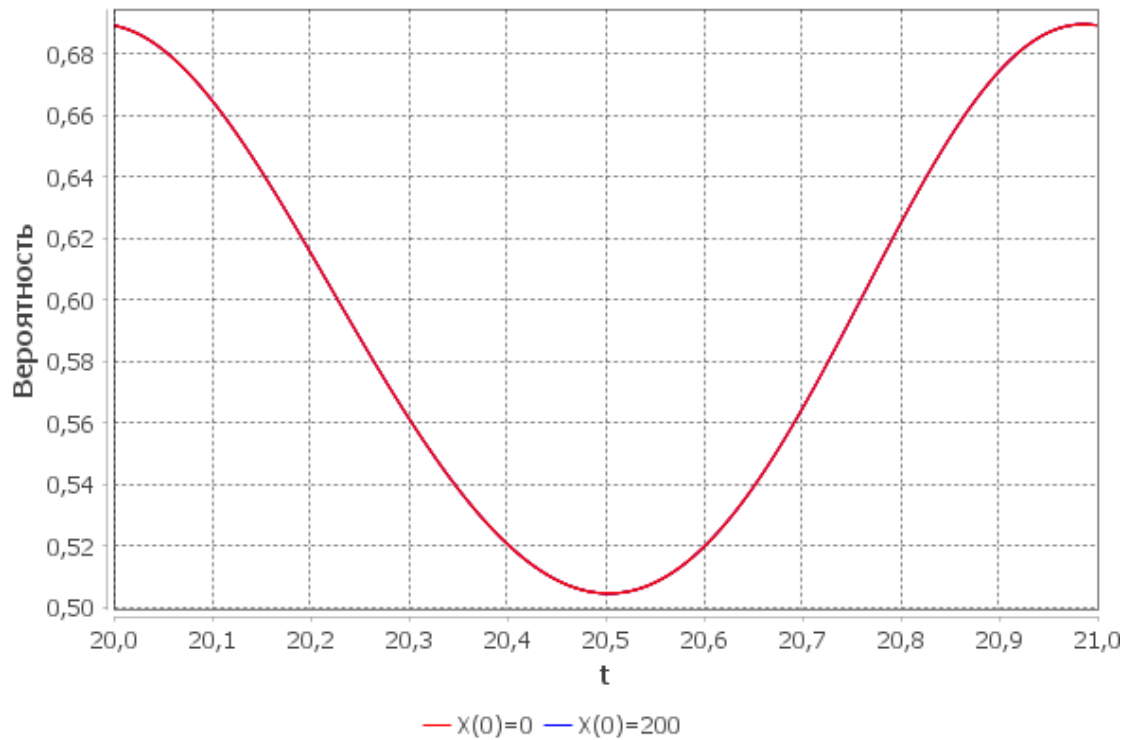


Рис. 4.10: Пример 2. Предельная вероятность $p_0(t)$ для усеченного процесса, 200 состояний, интервал $[20, 21]$.

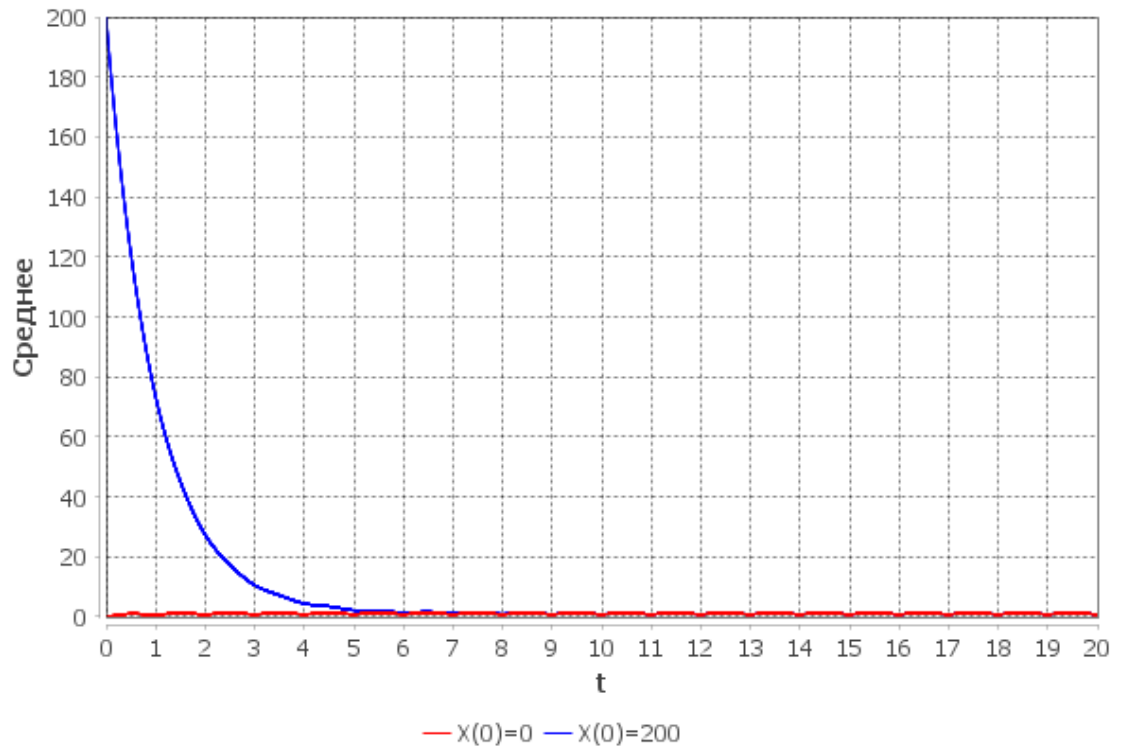


Рис. 4.11: Пример 2. Поведение среднего $E(t, k)$ для усеченного процесса, 200 состояний, интервал $[0, 20]$.

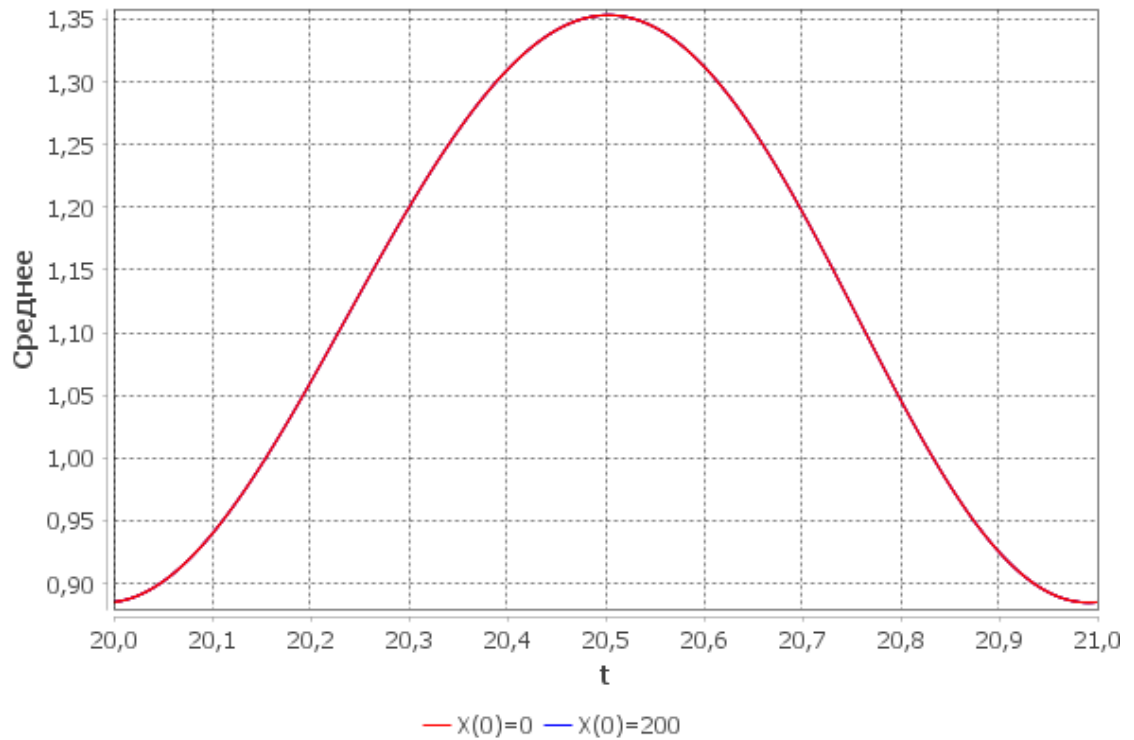


Рис. 4.12: Пример 2. Предельное среднее $E(t, k)$ для усеченного процесса, 200 состояний, интервал $[20, 21]$.

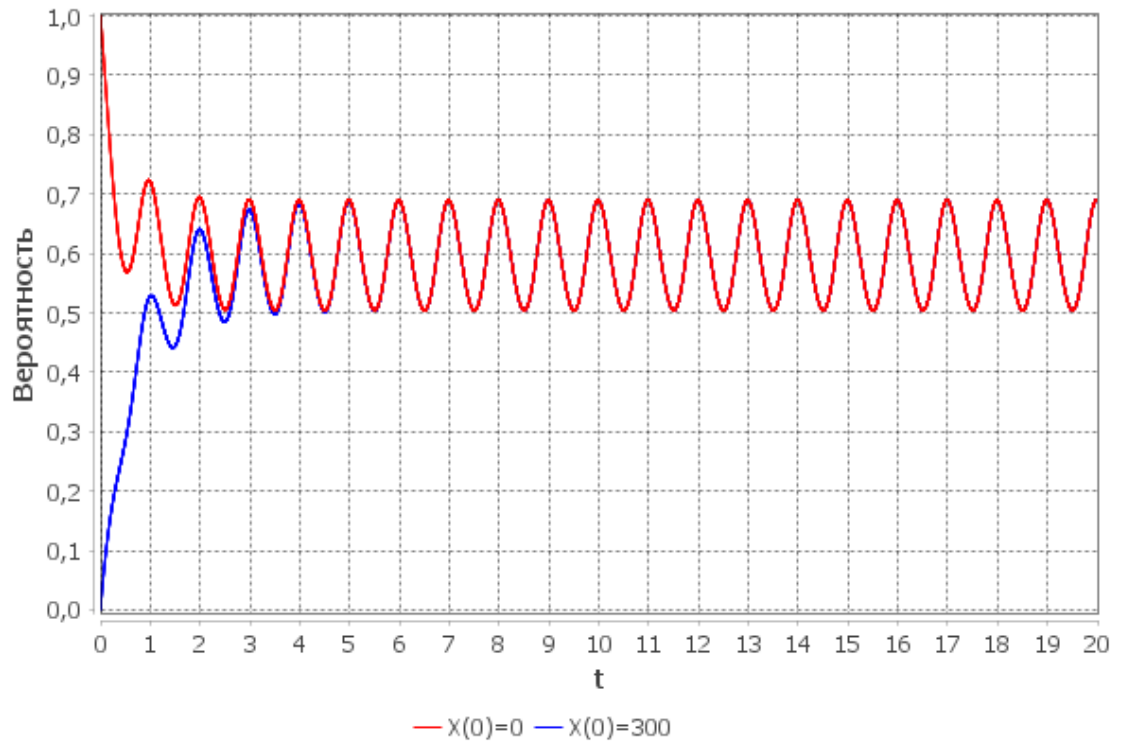


Рис. 4.13: Пример 2. Поведение вероятности $p_0(t)$ для усеченного процесса, 300 состояний, интервал $[0, 20]$.

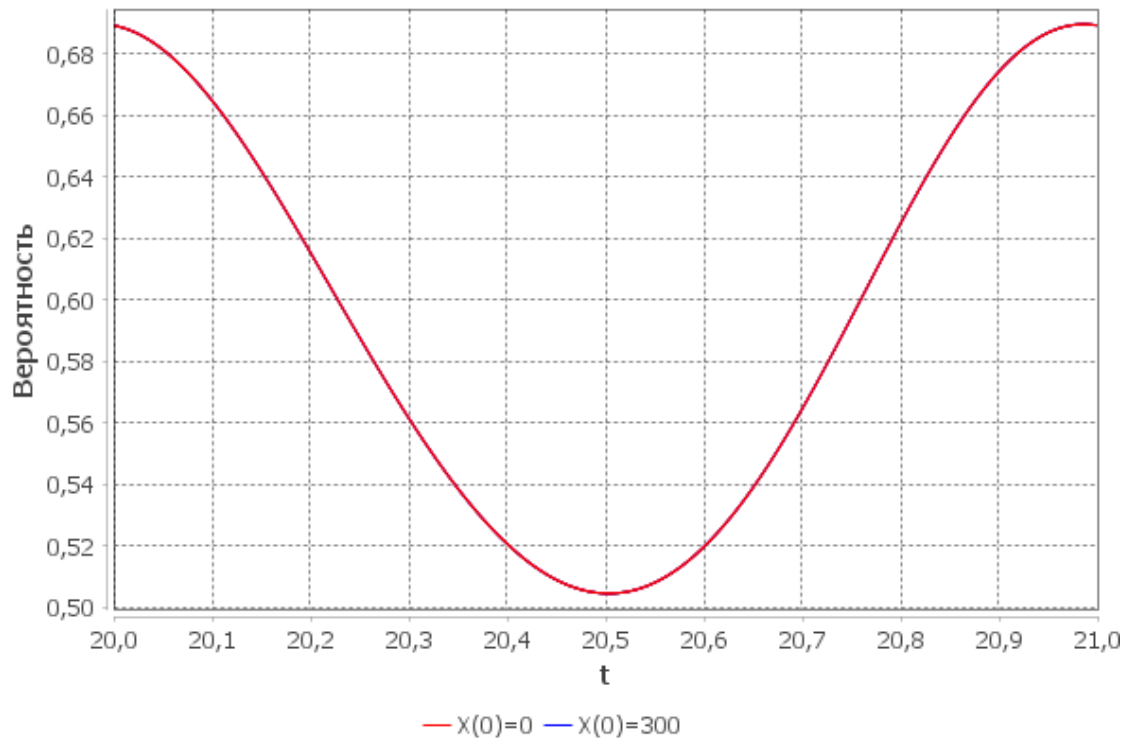


Рис. 4.14: Пример 2. Предельная вероятность $p_0(t)$ для усеченного процесса, 300 состояний, интервал $[20, 21]$.

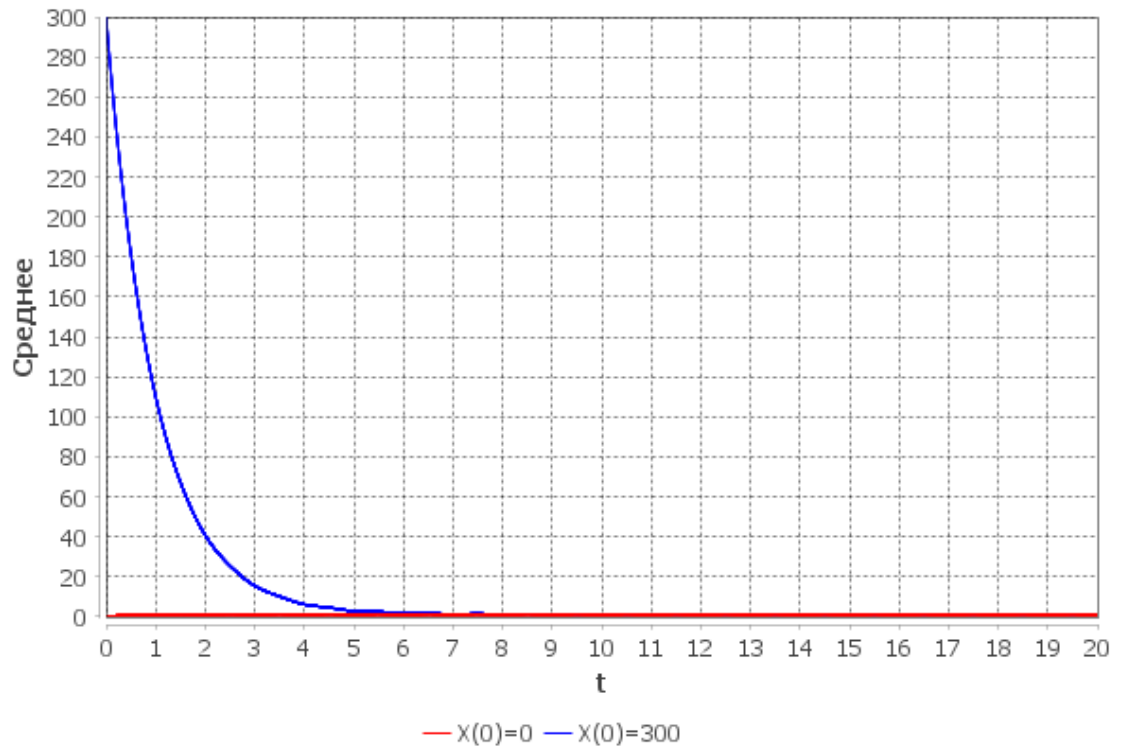


Рис. 4.15: Пример 2. Поведение среднего $E(t, k)$ для усеченного процесса, 300 состояний, интервал $[0, 20]$.

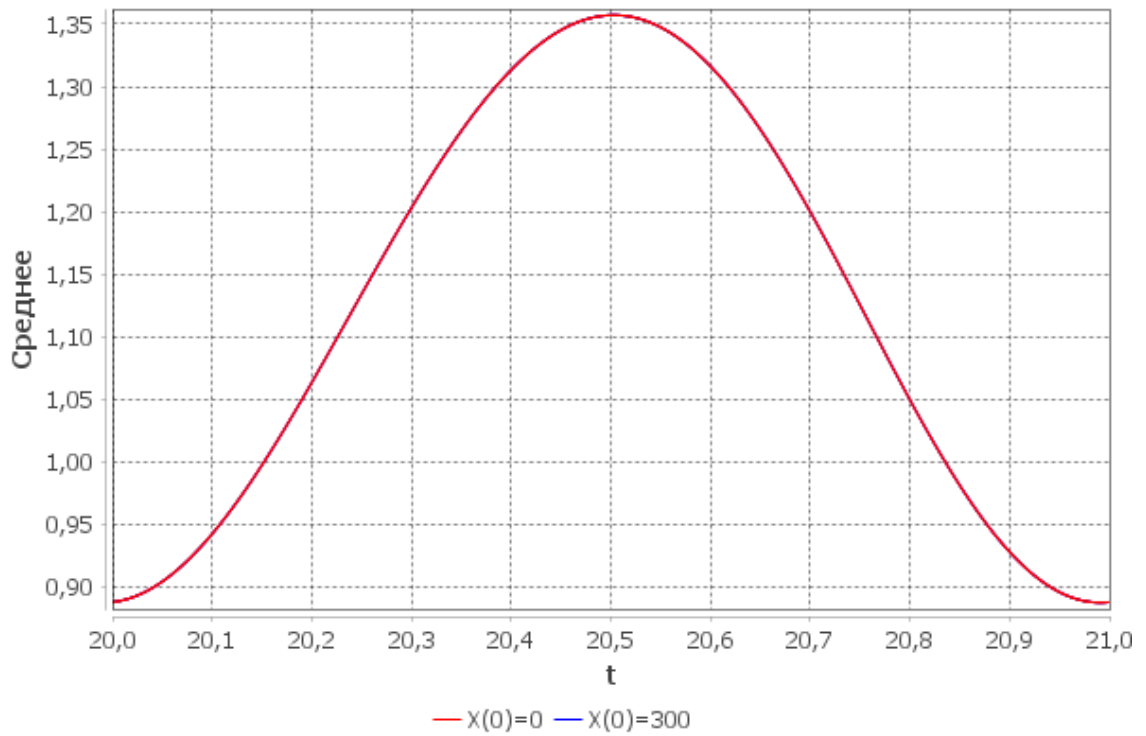


Рис. 4.16: Пример 2. Предельное среднее $E(t, k)$ для усеченного процесса, 300 состояний, интервал $[20, 21]$.

Глава 5

Исследование моделей, связанных с процессами рождения и гибели

5.1 Нестационарная марковская система с групповым поступлением и групповым обслуживанием требований с управлением, зависящим от состояния

В статье [120] авторы рассмотрели марковскую систему массового поступления и массового обслуживания с общим управлением, зависящим от состояния (см. также [37, 38, 121, 120, 122]). Общее число $X(t)$ требований в момент времени t в этой системе составляет цепь Маркова с непрерывным временем и пространством состояний $\{0, 1, 2, \dots\}$. Матрица интенсивностей $Q(t) = (q_{ij}(t))_{i,j=0}^{\infty}$ имеет довольно специфическую структуру:

$$q_{ij}(t) = \begin{cases} h_{ij}(t), & \text{если } 0 \leq i \leq k-1, j \geq 0, \\ b_{i-j+k}(t), & \text{если } i \geq k, j \geq i-k, \\ 0 & \text{иначе,} \end{cases} \quad (5.1.1)$$

где

$$\begin{aligned} h_{ij}(t) &\geq 0 \quad (i \neq j), \\ 0 \leq -h_{ii}(t) &= \sum_{j \neq i} h_{ij}(t) < +\infty \quad (0 \leq i \leq k-1), \\ b_0(t) &> 0, \quad b_j(t) \leq 0 \quad (j \neq k), \end{aligned}$$

$$\sum_{j=N+1}^{\infty} b_j(t) > 0 \text{ и } 0 \leq -b_k(t) = \sum_{j \neq k} b_j(t) < +\infty$$

для некоторого фиксированного целого $k \geq 1$.

Покажем, что по крайней мере для одного частного случая этой системы, даже когда интенсивности зависят от времени, можно получить оценки скорости сходимости и устойчивости. В частности, возьмем пример из раздела 7 [120], за исключением того, что все интенсивности перехода зависят от времени.

Результаты исследования этой модели, а именно, оценки скорости сходимости опубликованы в [125], устойчивости в [134].

5.1.1 Описание модели

Элементы матрицы интенсивности $Q(t)$ следующие: $h_{i,i-1}(t) = \mu(t)$, $h_{i,i+1}(t) = \lambda(t)$, $h_{i,i}(t) = -(\lambda(t) + \mu(t))$, $b_0 = \mu$, $b_{k+1} = \lambda(t)$, $b_k = -(\lambda(t) + \mu(t))$ и $k = 3$, где $\lambda(t)$ и $\mu(t)$ неотрицательны и локально интегрируемы по $[0, \infty)$. Тогда транспонированная матрица интенсивностей $A(t) = (a_{ij}(t))_{i,j=0}^{\infty} = Q^T(t)$ имеет вид

$$A(t) = \begin{pmatrix} -\lambda(t) & \mu(t) & 0 & \mu(t) & 0 & 0 & \dots \\ \lambda(t) & -(\lambda(t) + \mu(t)) & \mu(t) & 0 & \mu(t) & 0 & \dots \\ 0 & \lambda(t) & -(\lambda(t) + \mu(t)) & 0 & 0 & \mu(t) & \dots \\ 0 & 0 & \lambda(t) & -(\lambda(t) + \mu(t)) & 0 & 0 & \dots \\ 0 & 0 & 0 & \lambda(t) & -(\lambda(t) + \mu(t)) & 0 & \dots \\ 0 & 0 & 0 & 0 & \lambda(t) & -(\lambda(t) + \mu(t)) & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (5.1.2)$$

5.1.2 Оценки скорости сходимости

Рассмотрим вопросы, связанные с нуль-эргодичностью и слабой эргодичностью для марковской цепи $X(t)$.

Обозначим распределение $X(t)$ через $\mathbf{p}(t)$ т.е. $\mathbf{p}(t) = (p_0(t), p_1(t), \dots)^T$ и рассмотрим дифференциальное уравнение

$$\frac{d\mathbf{p}(t)}{dt} = A(t) \mathbf{p}(t). \quad (5.1.3)$$

Нуль-эргодичность

Сформулируем и докажем условие нуль-эргодичности.

Теорема 24 Если $\int_0^\infty (\lambda(t)(1 - \sigma) + \mu(t)(1 - \sigma^{-3})) dt = +\infty$ для некоторого $\sigma \in (0, 1)$, то Марковская цепь $X(t)$ нуль-эргодична,

$$\sum_{i=0}^{\infty} \sigma^i p_i(t) \leq e^{-\int_0^t (\lambda(u) + \mu(u) - \sigma\lambda(u) - \sigma^{-3}\mu(u)) du} \sum_{i=0}^{\infty} \sigma^i p_i(0), \quad t \geq 0, \quad (5.1.4)$$

и для всех $n \geq 0$ и $N \geq 0$ выполнено неравенство

$$\mathbf{P}(X(t) > n | X(0) = N) \geq 1 - \sigma^{N-n} e^{-\int_0^t (\lambda(u) + \mu(u) - \sigma\lambda(u) - \sigma^{-3}\mu(u)) du}. \quad (5.1.5)$$

Доказательство. Зафиксируем $\sigma > 0$ и определим убывающую последовательность положительных чисел $\{\delta_n, n \geq 0\}$ на $\delta_n = \sigma^n$. Положим $\tilde{\mathbf{p}}(t) = \Lambda \mathbf{p}(t)$, где $\Lambda = \text{diag}(\delta_0, \delta_1, \dots)$. Тогда получим

$$\frac{d}{dt} \tilde{\mathbf{p}}(t) = \tilde{A}(t) \tilde{\mathbf{p}}(t), \quad (5.1.6)$$

где $\tilde{A}(t) = \Lambda A(t) \Lambda^{-1}$.

Обозначим через $-\tilde{\alpha}_i(t)$ сумму всех элементов в i -м столбце $\tilde{A}(t)$, т.е.

$$\begin{aligned} \tilde{\alpha}_0(t) &= (1 - \sigma) \lambda(t), \\ \tilde{\alpha}_i(t) &= (1 - \sigma) (\lambda(t) + \mu(t) - \sigma^{-1}\mu(t)), \quad i = 1, 2, \\ \tilde{\alpha}_i(t) &= \underbrace{\lambda(t) + \mu(t) - \sigma\lambda(t) - \sigma^{-3}\mu(t)}_{=\beta(t)}, \quad i \geq 3. \end{aligned}$$

Если $0 < \sigma < 1$, то $\tilde{\alpha}_0(t) \geq \beta(t)$, $\tilde{\alpha}_1(t) \geq \beta(t)$, $\tilde{\alpha}_2(t) \geq \beta(t)$ и верхняя оценка следует из (1.3.6), применяемого к (5.1.6):

$$\sum_{i=0}^{\infty} \sigma^i p_i(t) \leq e^{-\int_0^t \beta(u) du} \sum_{i=0}^{\infty} \sigma^i p_i(0). \quad (5.1.7)$$

Если σ выбрано так, что $\int_0^t (\lambda(u) + \mu(u) - \sigma\lambda(u) - \sigma^{-3}\mu(u)) du = +\infty$, то из (5.1.7) следует, что $p_i(t) \rightarrow 0$ при $t \rightarrow \infty$ для каждого $i \geq 0$ и, следовательно, $X(t)$ является нуль эргодичным. В таком случае можно извлечь дополнитель-

ную информацию из (5.1.7). Заметим, что для любого фиксированного $n \geq 0$ справедливо что

$$\sigma^n \sum_{i=0}^n p_i(t) \leq \sum_{i=0}^n \sigma^i p_i(t) \leq \sum_{i=0}^{\infty} \sigma^i p_i(t).$$

Таким образом, если $X(0) = N$, т.е. $p_N(0) = 1$, то для любого $n \geq 0$ выполняется следующая верхняя оценка условной вероятности $P(X(t) \leq n | X(0) = N)$, $N \geq 0$:

$$P(X(t) \leq n | X(0) = N) \leq \sigma^{N-n} e^{-\int_0^t \beta(u) du}. \quad (5.1.8)$$

Слабая эргодичность

Используя условие нормализации $p_0(t) = 1 - \sum_{i \geq 1} p_i(t)$ перепишем систему (5.1.3) следующим образом:

$$\frac{d}{dt} \mathbf{z}(t) = B(t) \mathbf{z}(t) + \mathbf{f}(t), \quad (5.1.9)$$

где матрица $B(t)$ с элементами $b_{ij}(t) = a_{ij}(t) - a_{i0}(t)$ не имеет вероятностного значения, а векторы $\mathbf{f}(t)$ и $\mathbf{z}(t)$ равны

$$\mathbf{f}(t) = (\lambda(t), 0, 0, \dots)^T, \quad \mathbf{z}(t) = (p_1(t), p_2(t), \dots)^T.$$

Пусть $\mathbf{z}^*(t)$ и $\mathbf{z}^{**}(t)$ будут двумя решениями (5.1.9), соответствующими двум различным начальным условиям $\mathbf{z}^*(0)$ и $\mathbf{z}^{**}(0)$. Тогда для вектора $\mathbf{y}(t) = \mathbf{z}^*(t) - \mathbf{z}^{**}(t) = (y_1(t), y_2(t), \dots)^T$, с произвольными элементами получаем систему

$$\frac{d}{dt} \mathbf{y}(t) = B(t) \mathbf{y}(t). \quad (5.1.10)$$

Матрица $B(t)$ в (5.1.10) может иметь отрицательные недиагональные элементы.

Возьмём $\sigma > 1$ и возрастающую последовательность положительных чисел $\{d_n, n \geq 1\}$ как $d_n = \sigma^{n-1}$. Легко заметить, что преобразование подобия

$DB(t)D^{-1} = B^{**}(t)$, где D - матрица вида

$$D = \begin{pmatrix} d_1 & d_1 & d_1 & \cdots \\ 0 & d_2 & d_2 & \cdots \\ 0 & 0 & d_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (5.1.11)$$

даёт матрицу $B^{**}(t)$:

$$\begin{pmatrix} -(\lambda(t)+\mu(t)) & \mu(t)\frac{d_1}{d_2} & -\mu(t)\frac{d_1}{d_3} & \mu(t)\frac{d_1}{d_4} & 0 & 0 & \cdots \\ \lambda(t)\frac{d_2}{d_1} & -(\lambda(t)+\mu(t)) & 0 & 0 & \mu(t)\frac{d_2}{d_5} & 0 & \cdots \\ 0 & \lambda(t)\frac{d_3}{d_2} & -(\lambda(t)+\mu(t)) & 0 & 0 & \mu(t)\frac{d_3}{d_6} & \cdots \\ 0 & 0 & \lambda(t)\frac{d_4}{d_3} & -(\lambda(t)+\mu(t)) & 0 & 0 & \cdots \\ 0 & 0 & 0 & \lambda(t)\frac{d_5}{d_4} & -(\lambda(t)+\mu(t)) & 0 & \cdots \\ 0 & 0 & 0 & 0 & \lambda(t)\frac{d_6}{d_5} & -(\lambda(t)+\mu(t)) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}. \quad (5.1.12)$$

Положим $\mathbf{w}(t) = D\mathbf{y}(t)$ в (5.1.10), получим систему уравнений

$$\frac{d}{dt}\mathbf{w}(t) = B^{**}(t)\mathbf{w}(t). \quad (5.1.13)$$

Обозначим через $-\alpha_i(t)$ сумму всех элементов в i -м столбце $B^{**}(t)$, т.е.,

$$\alpha_i(t) = \begin{cases} \lambda(t) + \mu(t) - \lambda(t)\frac{d_2}{d_1}, & i = 1, \\ \lambda(t) + \mu(t) - \lambda(t)\frac{d_3}{d_2} - \mu(t)\frac{d_1}{d_2}, & i = 2, \\ \lambda(t) + \mu(t) - \lambda(t)\frac{d_4}{d_3} - \mu(t)\frac{d_1}{d_3}, & i = 3, \\ \lambda(t) + \mu(t) - \lambda(t)\frac{d_{i+1}}{d_i} - \mu(t)\frac{d_{i-3}}{d_i}, & i \geq 4. \end{cases} \quad (5.1.14)$$

Учитывая, что $d_n = \sigma^{n-1}$, получим $\alpha_i(t) =$

$$\begin{cases} \lambda(t) + \mu(t) - \lambda(t)\sigma, \\ \lambda(t) + \mu(t) - \lambda(t)\sigma - \mu(t)\sigma^{-1}, \\ \lambda(t) + \mu(t) - \lambda(t)\sigma - \mu(t)\sigma^{-2}, \\ \lambda(t) + \mu(t) - \lambda(t)\sigma - \mu(t)\sigma^{-3}, \end{cases} \geq \begin{cases} \mu(t) - \lambda(t)(\sigma - 1), \\ \mu(t)(1 - \sigma^{-1}) - \lambda(t)(\sigma - 1), \\ \mu(t)(1 - \sigma^{-2}) - \lambda(t)(\sigma - 1), \\ \mu(t)(1 - \sigma^{-3}) - \lambda(t)(\sigma - 1) \end{cases} \quad (5.1.15)$$

и логарифмическая норма $B^{**}(t)$ равна $-\beta(t) = -\min(\min_{1 \leq k \leq 3} \alpha_k(t), \inf_{k \geq 4} \alpha_k(t)) \geq -(\mu(t)(1 - \sigma^{-1}) - \lambda(t)(\sigma - 1))$.

Помня, что $\mathbf{w}(t) = D\mathbf{y}(t)$, верхняя оценка для $\mathbf{y}(t) = \mathbf{z}^*(t) - \mathbf{z}^{**}(t)$ во «взвешенной» норме из (1.3.6) имеет вид

$$\|D\mathbf{y}(t)\| \leq e^{-\int_0^t \beta(u) du} \|D\mathbf{y}(0)\|. \quad (5.1.16)$$

Верхняя оценка для $\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|$ получена из (5.1.16). Заметим, что $\|\mathbf{y}(t)\| \leq 2\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|$ поскольку $\mathbf{y}(t)$ является решением (5.1.10) — система с исключенным нулевым состоянием. Во-вторых, в [уравнении (18)] в [81] показано, что $\|\mathbf{x}\| \leq 2\|DT\mathbf{x}\|$ для любого вектора \mathbf{x} . Следовательно,

Теорема 25 *Если $\int_0^\infty (\mu(t)(1 - \sigma^{-1}) - \lambda(t)(\sigma - 1)) dt = +\infty$ для некоторого $\sigma > 1$, тогда цепь Маркова $X(t)$ слабо эргодична, и выполняется оценка*

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 4e^{-\int_0^t (\mu(u)(1 - \sigma^{-1}) - \lambda(u)(\sigma - 1)) du} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1D}. \quad (5.1.17)$$

5.1.3 Оценки устойчивости

Рассмотрим применение общего метода оценки устойчивости (см. обзор в [86]) для исследуемых моделей.

Рассмотрим процесс $\bar{X}(t), t \geq 0$, описывающий число требований в «возмущенной» системе обслуживания, с соответствующей транспонированной матрицей интенсивностей $\bar{A}(t)$, где матрица «возмущения» $\hat{A}(t) = A(t) - \bar{A}(t)$ в некотором смысле мала, и имеет ту же структуру, с соответствующими возмущенными интенсивностями $\bar{\lambda}(t)$ и $\bar{\mu}(t)$.

Пусть

$$|\lambda(t) - \bar{\lambda}(t)| = |\hat{\lambda}(t)| \leq \hat{\epsilon}, \quad |\mu(t) - \bar{\mu}(t)| = |\hat{\mu}(t)| \leq \hat{\epsilon}. \quad (5.1.18)$$

Тогда получаем из (5.2.25) ограничение

$$\|\hat{A}(t)\| = 2 \sup_k |\hat{a}_{kk}(t)| \leq 4\hat{\epsilon} \quad (5.1.19)$$

и

$$\begin{aligned}\|B(t)\|_{1D} &\leq (1 + \sigma)|\lambda(t)| + \left(1 + \frac{1}{\sigma}\right)|\mu(t)| \leq (3 + \sigma)L, \\ \|f(t)\|_{1D} &\leq L, \\ \|B(t) - \bar{B}(t)\|_{1D} &\leq (1 + \sigma)|\hat{\lambda}(t)| + \left(1 + \frac{1}{\sigma}\right)|\hat{\mu}(t)| \leq (3 + \sigma)\hat{\epsilon}, \\ \|f(t) - \bar{f}(t)\|_{1D} &\leq \hat{\epsilon}\end{aligned}$$

для всех $t \geq 0$. Из теоремы 2 статьи [86] следующие оценки

Теорема 26 *Если процесс $X(t)$ $1D$ -экспоненциально слабо эргодичен, то $\bar{X}(t)$ также $1D$ -экспоненциально слабо эргодичен т.е.*

$$e^{-\int_s^t \beta(\tau) d\tau} \leq M e^{-\alpha(t-s)}, \quad (5.1.20)$$

для некоторых положительных M, α . Тогда выполняется следующая оценка устойчивости

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\|_{1D} \leq \frac{M\hat{\epsilon}(M(\sigma + 3)L + a)}{a(a - M\hat{\epsilon}(\sigma + 3))}. \quad (5.1.21)$$

Более того, если $W = \inf_{k \geq 1} \frac{d_{k+1}}{k} > 0$, тогда

$$\limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq \frac{M\hat{\epsilon}(M(\sigma + 3)L + a)}{W a (a - M\hat{\epsilon}(\sigma + 3))}. \quad (5.1.22)$$

Замечание 1. *Из теоремы 26 можно показать, что верна оценка*

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\| \leq \frac{4M\hat{\epsilon}(M(\sigma + 3)L + a)}{a(a - M\hat{\epsilon}(\sigma + 3))}. \quad (5.1.23)$$

5.1.4 Пример

Для вычислительных экспериментов рассмотрим периодический случай с интенсивностями $\lambda(t) = 3(1 + \sin 2\pi t)$ и $\mu(t) = 4(1 + \cos 2\pi t)$.

Последовательность $\{d_n, n \geq 1\}$, определенная $d_n = \left(\frac{10}{9}\right)^{n-1}$, приводит к

$$\begin{aligned}\beta(t) &= 4(1 + \cos 2\pi t) \left(1 - \left(\frac{10}{9}\right)^{-1}\right) - 3(1 + \sin 2\pi t) \left(\frac{10}{9} - 1\right) = \\ &= \frac{1}{15} + 0.4 \cos 2\pi t - \frac{1}{3} \sin 2\pi t\end{aligned}$$

, неравенству

$$e^{-\int_0^t \beta(\tau) d\tau} \leq 22e^{-0.066t}.$$

Тогда положим $M = 22$ и $\alpha = 0.066$ и из теоремы 25 получим оценку

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 88e^{-0.066t} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1D}.$$

Предварительные и предельные значения величины $p_0(t)$ и $E(t, k)$ показаны на рисунках 5.1–5.4.

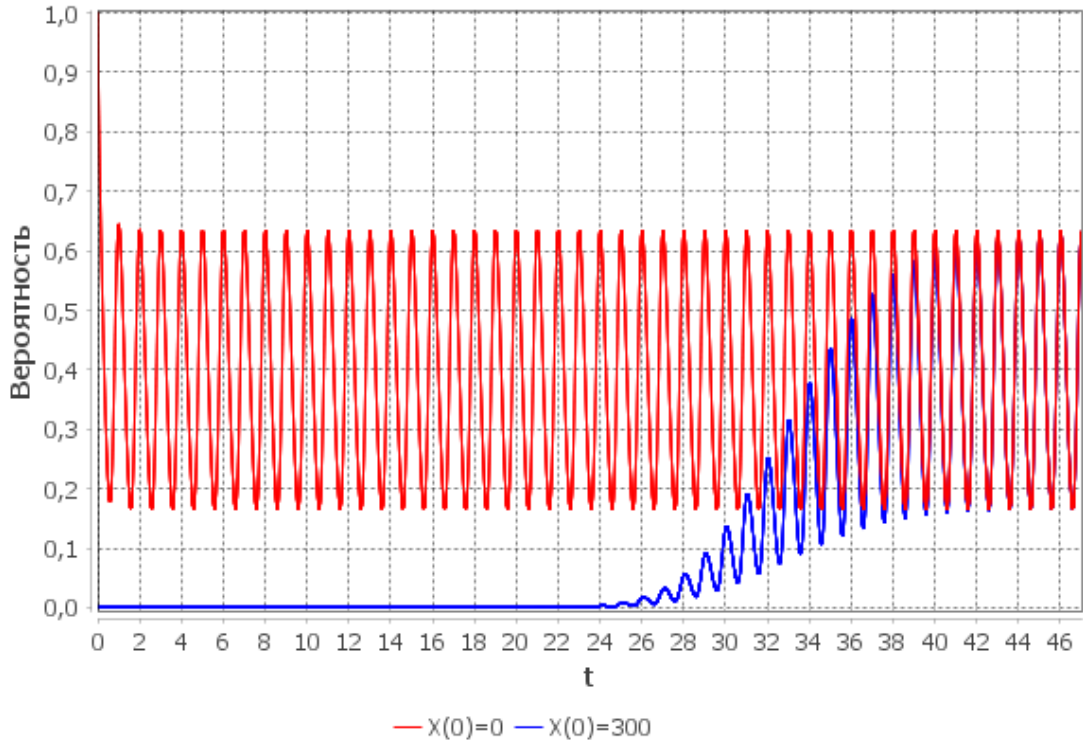


Рис. 5.1: Скорость сходимости вероятности пустой очереди $p_0(t)$ в интервале $[0, 47]$.

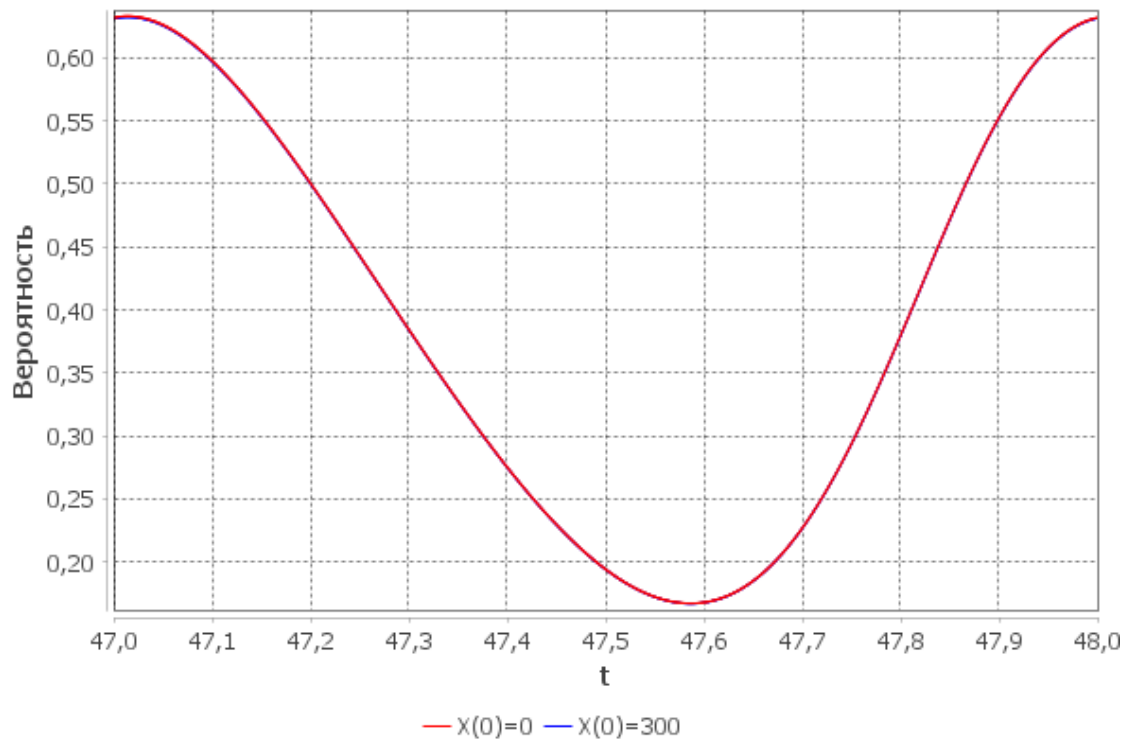


Рис. 5.2: Предельная вероятность $p_0(t)$ пустой очереди.

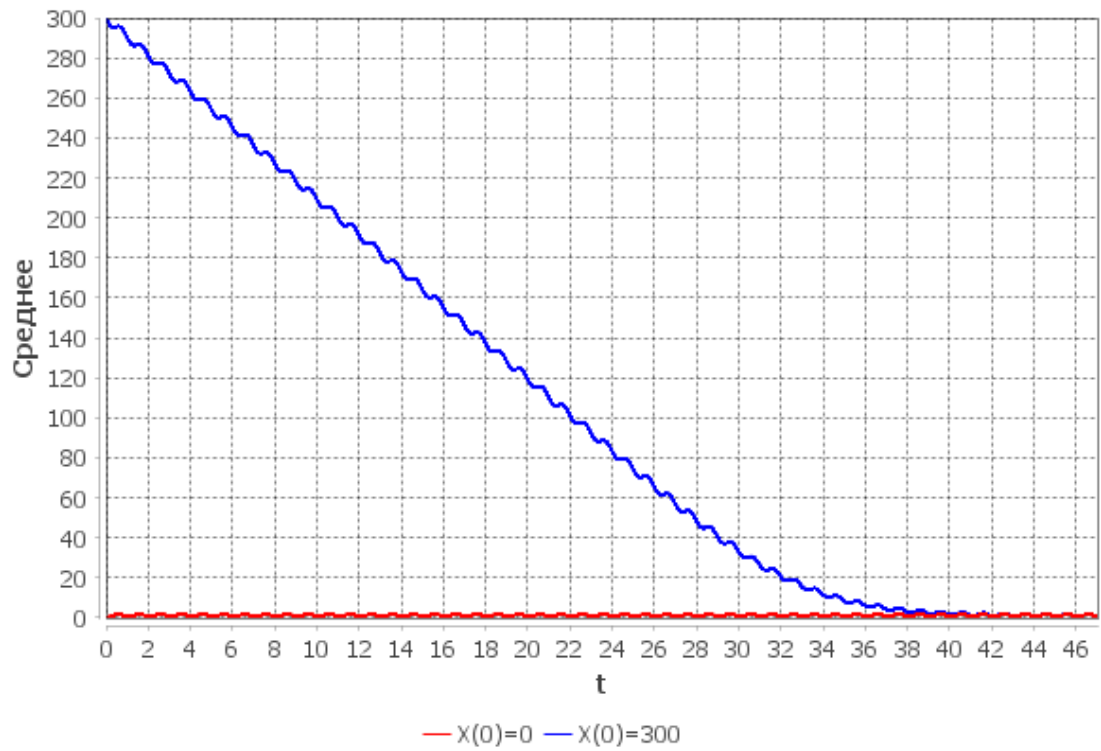


Рис. 5.3: Скорость сходимости среднего $E(t, k)$ числа требований в системе в интервале $[0, 47]$.

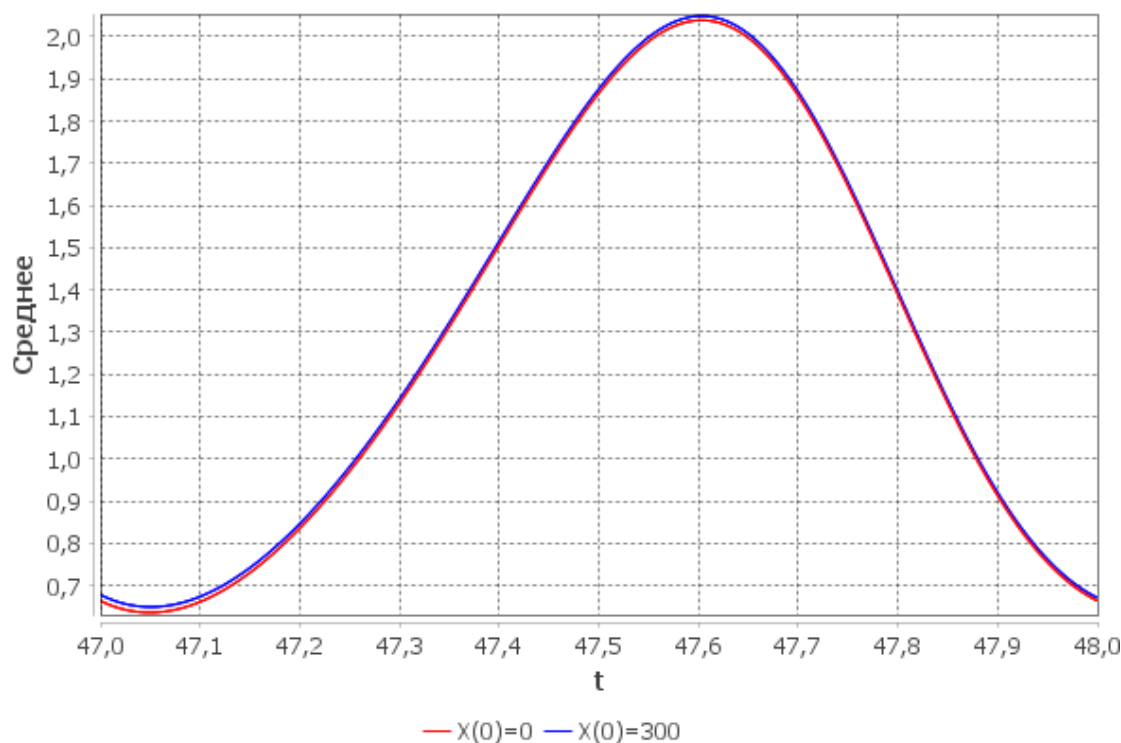


Рис. 5.4: Предельное среднее $E(t, k)$ число требований в систем.

5.2 Оценки скорости сходимости и устойчивости для систем с нетерпеливыми клиентами

Применение многосерверных систем организации очередей с нетерпеливыми клиентами можно увидеть в моделировании дорожного движения, бизнесе и отраслях промышленности, компьютерных коммуникациях, здравоохранении и медицинских науках, сервисных системах, розничных магазинах и т.д. Нетерпение влияет на скорость поступления в систему массового обслуживания. Важно отметить, что распространенность феномена нетерпеливости клиентов оказывает крайне негативное влияние на исследуемую систему массового обслуживания. Нетерпеливость клиентов стала насущной проблемой как частных, так и государственных предприятий. Если говорить с точки зрения бизнеса, фирмы теряют своих потенциальных клиентов из-за нетерпеливости клиентов, что сказывается на бизнесе фирм в целом.

Здесь рассматривается нестационарная система массового обслуживания с S серверами и нетерпеливыми клиентами, в предположении, что интенсивно-

сти поступления новых требований уменьшаются с ростом очереди. Рассмотрен процесс $X(t)$, описывающий число требований в такой системе, доказано существование предельного режима распределения вероятностей состояний и предельного среднего для $X(t)$, получены оценки скорости сходимости к предельному режиму и предельному среднему. Получены оценки устойчивости. Для исследования применяется подход, основанный на понятии логарифмической нормы операторной функции. В качестве примера рассмотрена простая модель нестационарной системы, в которой интенсивности требований линейно убывают с увеличением длины очереди.

Результаты исследования модели опубликованы в [132].

5.2.1 Описание модели

Рассмотрим модель массового обслуживания с S серверами, интенсивности поступления требований в которую уменьшаются при увеличении длины очереди, такого рода модели изучались, например, в [88, 89, 90, 91, 92, 93, 94] и называются там системами, в которых потенциальным клиентам не нравятся большие очереди. Заявки поступают в систему с интенсивностями $\phi_i \lambda(t)$ (где коэффициенты ϕ_i монотонно убывают т.е. интенсивность поступления требований уменьшается с ростом очереди), а обслуживаются с интенсивностями $\mu_k(t) = \mu(t) \min(k, S)$.

Пусть $X(t), t \geq 0$ - процесс, описывающий число требований в рассматриваемой системе массового обслуживания. Это процесс рождения и гибели (ПРГ) с интенсивностями рождения и гибели $\lambda_i(t) = \phi_i \lambda(t)$, если $i \geq 0$ и $\mu_i(t) = i \mu(t)$, если $k \leq S$ или $\mu_i(t) = S \mu(t)$, если $k > S$ соответственно. Предполагаем, что все функции, описывающие интенсивности переходов, неотрицательны и локально интегрируемы на $[0, \infty)$.

5.2.2 Скорость сходимости

Для получения оценки скорости сходимости применим подход, изложенный в [95]. Обозначим через $\mathbf{p}(t) = (p_0(t), p_1(t), \dots)^T$ вектор вероятностей состояний в данный момент времени t для процесса $X(t)$, получаем прямую

систему Колмогорова

$$\frac{d\mathbf{p}(t)}{dt} = A(t) \mathbf{p}(t), \quad (5.2.24)$$

где $A(t) =$

$$\begin{pmatrix} -\phi_0\lambda(t) & \mu(t) & 0 & 0 & \cdots \\ \phi_0\lambda(t) & -(\phi_1\lambda(t) + \mu(t)) & \min(2, S)\mu(t) & 0 & \cdots \\ 0 & \phi_1\lambda(t) & -(\phi_2\lambda(t) + \min(2, S)\mu(t)) & \min(3, S)\mu(t) & \cdots \\ 0 & 0 & \phi_2\lambda(t) & -(\phi_3\lambda(t) + \min(3, S)\mu(t)) & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{pmatrix}, \quad (5.2.25)$$

Предполагаем, что матрица интенсивности существенно ограничена, т.е.

$$\sup_i |a_{ii}(t)| \leq L < \infty, \quad (5.2.26)$$

почти для всех $t \geq 0$.

Обозначим через $\|\cdot\|$ l_1 -норму вектора, $\|x\| = \sum |x_i|$, $\|B\| = \sup_j \sum_i |b_{ij}|$, если $B = (b_{ij})_{i,j=0}^{\infty}$, и обозначим через Ω множество всех векторов из l_1 с неотрицательными координатами и единичной нормой. Тогда имеем $\|A(t)\| = 2 \sup_k |a_{kk}(t)| \leq 2L$ почти для всех $t \geq 0$.

Исключим нулевое состояние, полагая

$$p_0(t) = 1 - \sum_{i \geq 1} p_i(t),$$

получаем из (5.2.24)

$$\frac{d\mathbf{z}}{dt} = B(t)\mathbf{z}(t) + \mathbf{f}(t), \quad (5.2.27)$$

где

$$B(t) = \begin{pmatrix} -(\phi_0\lambda(t) + \phi_1\lambda(t) + \mu(t)) & \min(2, S)\mu(t) - \phi_0\lambda(t) & -\phi_0\lambda(t) & \cdots \\ \phi_1\lambda(t) & -(\phi_2\lambda(t) + \min(2, S)\mu(t)) & \min(3, S)\mu(t) & \cdots \\ 0 & \phi_2\lambda(t) & -(\phi_3\lambda(t) + \min(3, S)\mu(t)) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (5.2.28)$$

и $\mathbf{f}(t) = (\phi_0\lambda(t), 0, 0, \dots)^T$.

Для получения нужных свойств и оценок потребуются некоторые вспомогательные «взвешенные» нормы. Рассмотрим последовательность положительных чисел $\{d_i\}$, такую что $d = \inf d_i > 0$, матрицу

$$D = \begin{pmatrix} d_1 & d_1 & d_1 & \cdots \\ 0 & d_2 & d_2 & \cdots \\ 0 & 0 & d_3 & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix} \quad (5.2.29)$$

и пространство последовательностей $l_{1D} = \{\mathbf{z} / \|\mathbf{z}\|_{1D} = \|\mathbf{Dz}\|_1 < \infty\}$.

Тогда из (5.2.27) получим уравнение:

$$\frac{d\mathbf{u}}{dt} = W(t)\mathbf{u}(t) + \mathbf{g}(t), \quad (5.2.30)$$

где $\mathbf{u}(t) = \mathbf{Dz}(t)$; $\mathbf{g}(t) = \mathbf{Df}(t)$; $W(t) = \mathbf{DB}(t)\mathbf{D}^{-1} =$

$$\begin{pmatrix} -(\phi_0\lambda(t) + \mu(t)) & \frac{d_1}{d_2}\mu(t) & 0 & \cdots \\ \frac{d_2}{d_1}\phi_1\lambda(t) & -(\phi_1\lambda(t) + \min(2, S)\mu(t)) & \frac{d_2}{d_3}\min(2, S)\mu(t) & \cdots \\ 0 & \frac{d_3}{d_2}\phi_2\lambda(t) & -(\phi_2\lambda(t) + \min(3, S)\mu(t)) & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{pmatrix}.$$

Положим

$$\gamma_*(t) = \inf_i \left(|w_{ii}(t)| - \sum_{j \neq i} |w_{ji}(t)| \right). \quad (5.2.31)$$

Тогда

$$\begin{aligned} \|B(t)\|_{1D} &= \|\mathbf{DB}(t)\mathbf{D}^{-1}\| = \|W(t)\| = \sup_i \left(|w_{ii}(t)| + \sum_{j \neq i} |w_{ji}(t)| \right) = \\ &= \sup_i \left(2|w_{ii}(t)| + \sum_{j \neq i} |w_{ji}(t)| - |w_{ii}(t)| \right) \leq 2 \sup_i |w_{ii}(t)| - \gamma_*(t) \leq 2L - \gamma_*(t), \end{aligned}$$

следовательно, операторная функция $B(t)$ ограничена в пространстве l_{1D} . Получаем

$$\gamma(B(t))_{1D} = \gamma(W(t)) = \sup_i \left(w_{ii}(t) + \sum_{j \neq i} |w_{ji}(t)| \right) = -\gamma_*(t). \quad (5.2.32)$$

Далее оценим логарифмическую норму оператора $W(t)$. Обозначим через $-\alpha_k(t)$ сумму всех элементов k -го столбца матрицы $W(t)$

$$\alpha_k(t) = \begin{cases} \mu(t) - \left(\frac{d_2}{d_1}\phi_1 - \phi_0\right) \lambda(t), & k = 1, \\ \left(\min(k, S) - \frac{d_{k-1}}{d_k} \min(k-1, S)\right) \mu(t) - \left(\frac{d_{k+1}}{d_k}\phi_k - \phi_{k-1}\right) \lambda(t), & k > 1. \end{cases} \quad (5.2.33)$$

и $\gamma_*(t) = \inf_k \alpha_k$. Пусть $d_k = \delta \in (1; \frac{S}{S-1})$ при $1 \leq k \leq S$ и $d_k = \delta d_{k-1}$ при $k > S$ и $\delta > 1$. Получаем

$$\alpha_k(t) = \begin{cases} \mu(t) + (\phi_{k-1} - \phi_k) \lambda(t), & 1 \leq k < S, \\ \mu(t) - (\delta\phi_k - \phi_{k-1}) \lambda(t), & k = S, \\ (S - \frac{1}{\delta}S) \mu(t) - (\delta\phi_k - \phi_{k-1}) \lambda(t), & k > S, \end{cases} \geq \quad (5.2.34)$$

$$\geq \begin{cases} \mu(t) - (\delta\phi_S - \phi_{S-1}) \lambda(t), \\ (S - \frac{1}{\delta}S) \mu(t) - (\delta\phi_k - \phi_k) \lambda(t), \end{cases} \geq \begin{cases} \mu(t) - (\delta - 1) \phi_S \lambda(t), \\ (S - \frac{1}{\delta}S) \mu(t) - (\delta - 1) \phi_S \lambda(t), \end{cases} \quad (5.2.35)$$

$$\geq \begin{cases} \mu(t) - (\delta - 1) \phi_S \lambda(t), \\ \frac{\delta-1}{\delta} (S\mu(t) - \delta\phi_S \lambda(t)), \end{cases} \quad (5.2.36)$$

а значит $\gamma_*(t) \geq \frac{\delta-1}{\delta} (S\mu(t) - \delta\phi_S \lambda(t))$ и справедливо следующее утверждение.

Теорема 27 Пусть найдется такое $\delta \in (1; \frac{S}{S-1})$, что $\int_0^\infty (S\mu(t) - \delta\phi_S \lambda(t)) dt = +\infty$. Тогда $X(t)$ слабо эргодична, причем

$$\|\mathbf{p}^1(t) - \mathbf{p}^2(t)\|_{1D} \leq e^{-\frac{\delta-1}{\delta} \int_0^t (S\mu(u) - \delta\phi_S \lambda(u)) du} \|\mathbf{p}^1(0) - \mathbf{p}^2(0)\|_{1D} \quad (5.2.37)$$

для любых начальных условий $\mathbf{p}^1(0) \in \Omega$, $\mathbf{p}^2(0) \in \Omega$ и всех $t \geq 0$.

5.2.3 Устойчивость

Применим подход такой же, как и в предыдущей части.

Рассмотрим процесс $\bar{X}(t), t \geq 0$, описывающий число требований в "возмущенной" системе обслуживания, с соответствующей транспонированной матрицей интенсивностей $\bar{A}(t)$, где матрица "возмущения" $\hat{A}(t) = A(t) - \bar{A}(t)$ в некотором смысле мала, и имеет ту же структуру, с соответствующими возмущенными интенсивностями $\bar{\lambda}_k(t)$ и $\bar{\mu}_k(t)$.

Пусть

$$|\lambda_k(t) - \bar{\lambda}_k(t)| = |\hat{\lambda}_k(t)| \leq \hat{\epsilon}, \quad |\mu_k(t) - \bar{\mu}_k(t)| = |\hat{\mu}_k(t)| \leq \hat{\epsilon}. \quad (5.2.38)$$

Тогда мы получаем из (5.2.25) следующее ограничение

$$\|\hat{A}(t)\| = 2 \sup_k |\hat{a}_{kk}(t)| \leq 4\hat{\epsilon} \quad (5.2.39)$$

и из теоремы 2 статьи [86] следующие оценки:

Теорема 28 *Если процесс $X(t)$ 1D-экспоненциально слабо эргодичен, то $\bar{X}(t)$ также 1D-экспоненциально слабо эргодичен т.е.*

$$e^{-\int_s^t \gamma_*(\tau) d\tau} \leq M e^{-\alpha(t-s)}, \quad (5.2.40)$$

для некоторых положительных M, α и

$$H = \sup_{|i-j|=1} \frac{d_i}{d_j} < \infty. \quad (5.2.41)$$

Тогда выполняется следующая оценка устойчивости

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\|_{1D} \leq \frac{M\hat{\epsilon}(2M(H+1) + a)}{a(a - 2M\hat{\epsilon}(H+1))}. \quad (5.2.42)$$

Более того, если $W = \inf_{k \geq 1} \frac{d_{k+1}}{d_k} > 0$, тогда

$$\limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq \frac{M\hat{\epsilon}(2M(H+1) + a)}{W a (a - 2M\hat{\epsilon}(H+1))}. \quad (5.2.43)$$

5.2.4 Пример

Пусть число серверов в системе $S = 5$, а интенсивности имеют вид: $\lambda(t) = 2 + 0.5 \sin 2\pi t$, $\mu(t) = 3 + 2 \cos 2\pi t$, $\phi_k = 3$ при $k < S$ и $\phi_k = \frac{3}{k-S+2}$ при $k \geq S$.

Для теоремы 27 и соответствующей теоремы 28 об "устойчивости" возьмем $\delta = \frac{6}{5}$. Очевидно, что $L \leq 65$. Тогда $H = \delta < 1.25$,

$$e^{-\int_0^t \frac{\delta-1}{\delta}(S\mu(\tau) - \delta\phi_S\lambda(\tau)) d\tau} \leq 7e^{-1.9t},$$

поэтому можно взять $a = 1.9$ и $M = 7$.

Получаем следующие оценки скорости сходимости:

$$\|\mathbf{p}^1(t) - \mathbf{p}^2(t)\|_{1D} \leq 7e^{-1.9t} \|\mathbf{p}^1(0) - \mathbf{p}^2(0)\|_{1D}, \quad (5.2.44)$$

из теоремы 27.

Соответствующие оценки возмущений имеют вид:

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\|_{1D} \leq \frac{124\hat{\epsilon}}{1.9 - 31.5\hat{\epsilon}}, \quad (5.2.45)$$

и

$$\limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq \frac{124\hat{\epsilon}}{0.24(1.9 - 31.5\hat{\epsilon})}, \quad (5.2.46)$$

из теоремы 28.

Далее проводим описанные в §2.4 эксперименты для подтверждения полученных результатов и построения предельных характеристик.

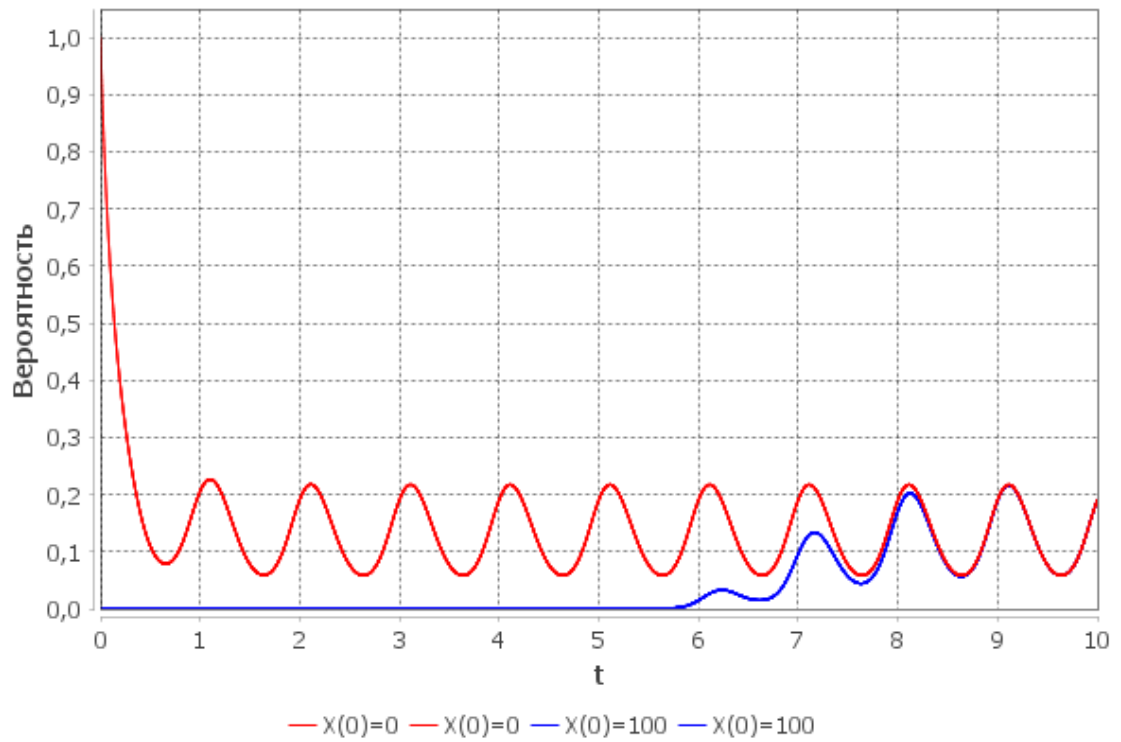


Рис. 5.5: Вероятность пустой очереди для $t \in [0, 10]$.

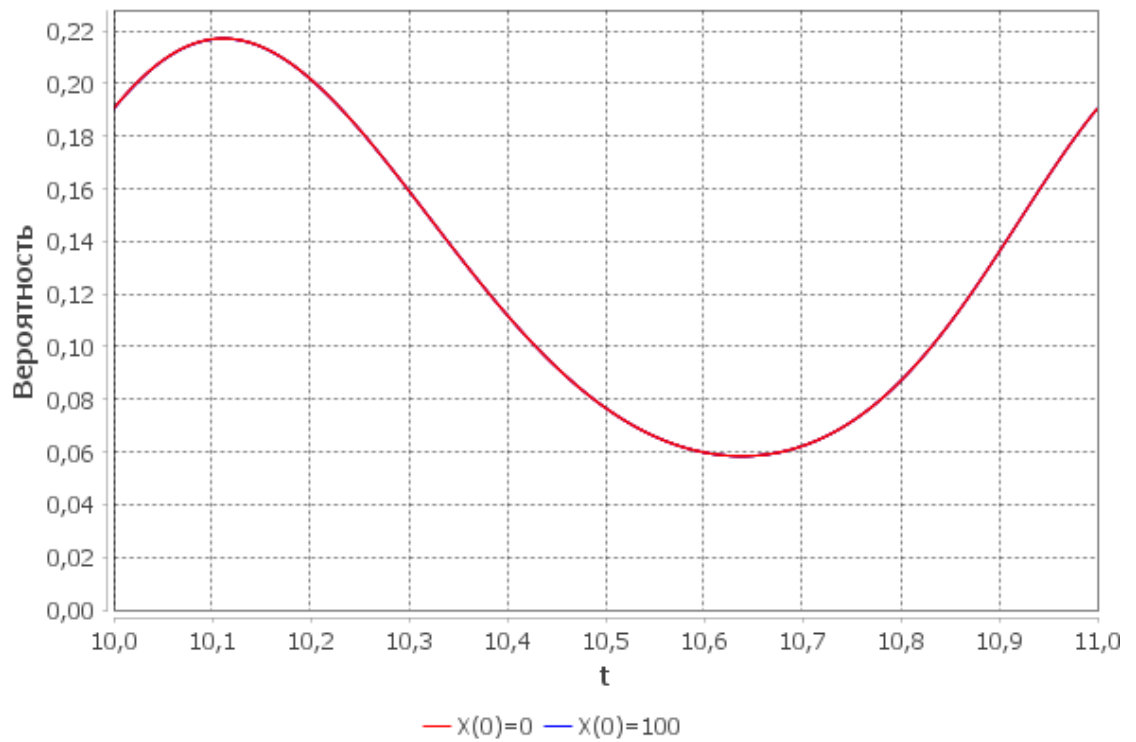


Рис. 5.6: Аппроксимация предельной вероятности пустой очереди для $t \in [10, 11]$.

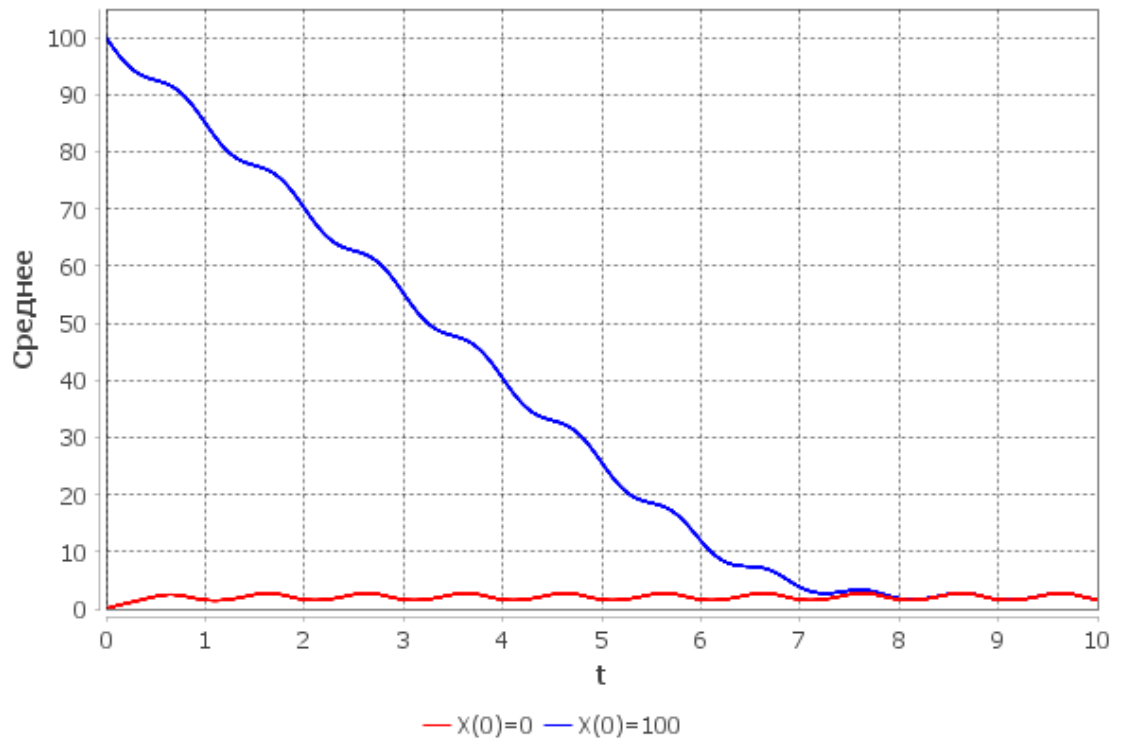


Рис. 5.7: Среднее $E(t, k)$ для $t \in [0, 10]$.

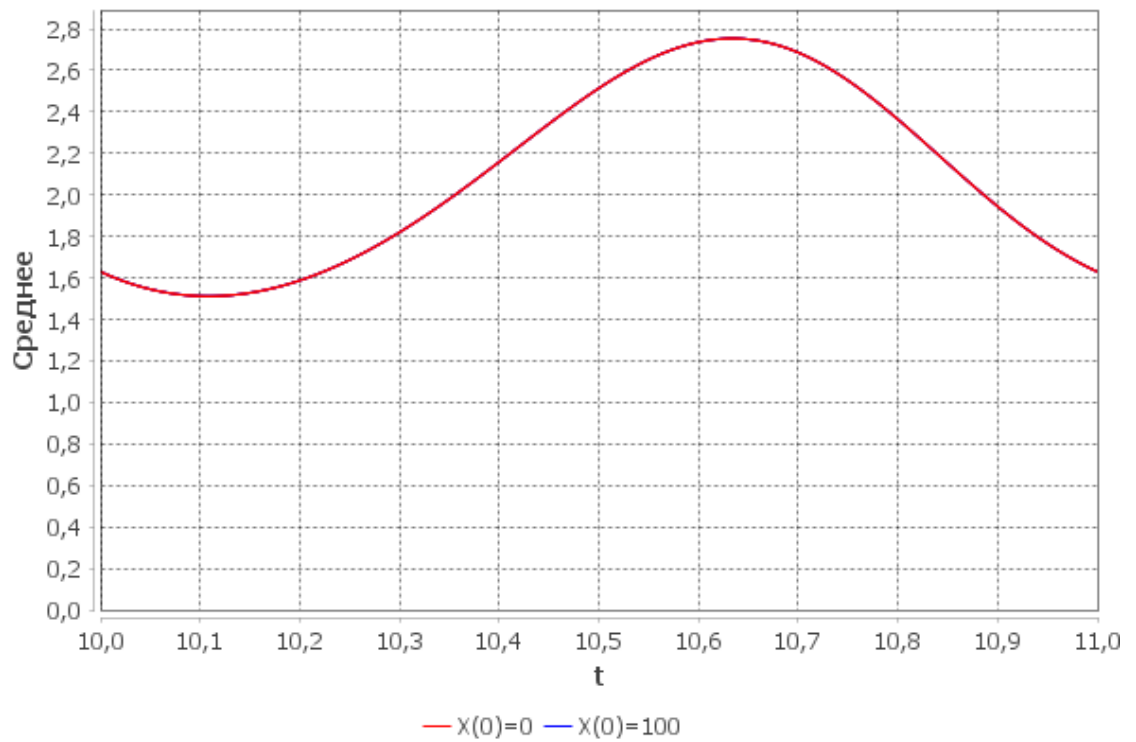


Рис. 5.8: Аппроксимация предельного среднего значения $E(t, k)$ для $t \in [10, 11]$.

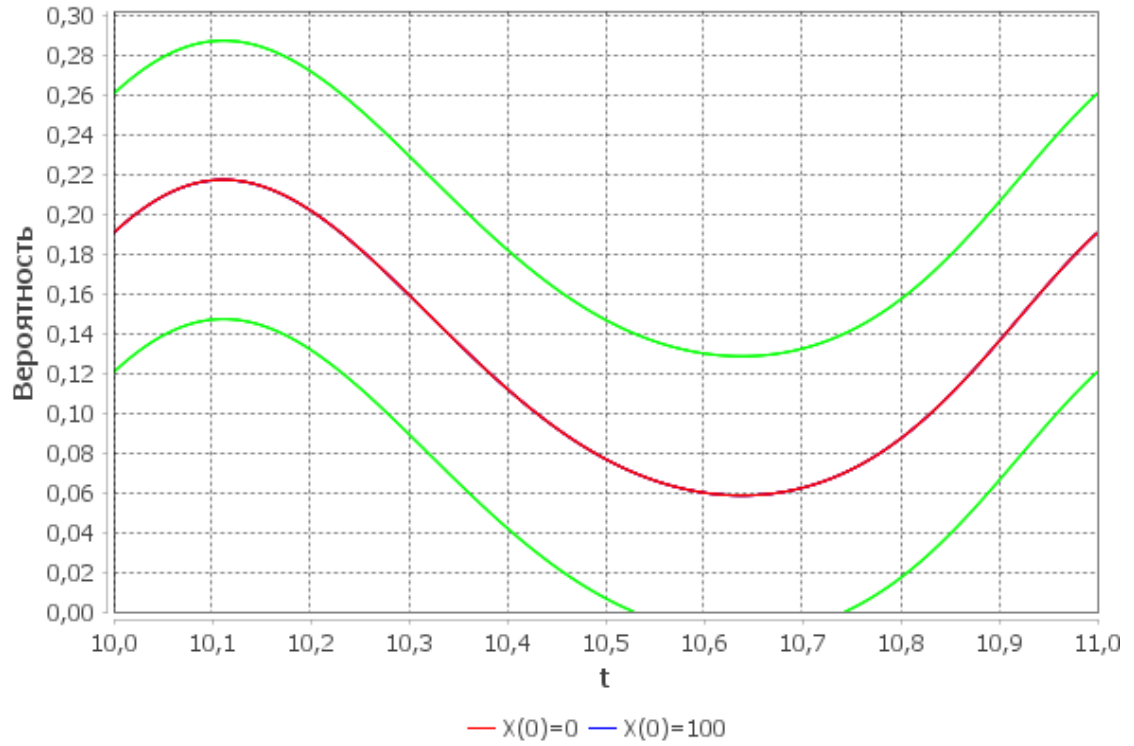


Рис. 5.9: Оценки устойчивости для предельной вероятности пустой очереди для $t \in [10, 11]$, $\hat{\epsilon} = 0.001$.

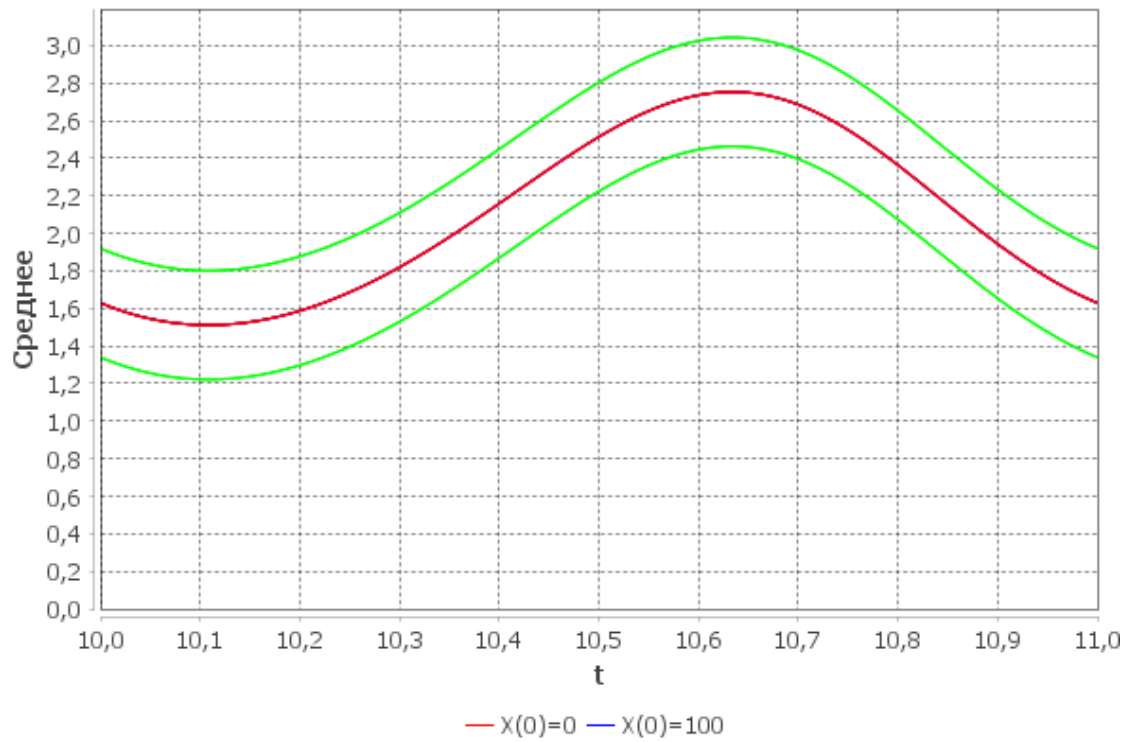


Рис. 5.10: Оценки устойчивости для предельного среднего значения для $t \in [10, 11]$, $\hat{\epsilon} = 0.001$.

5.3 Исследование системы массового обслуживания с эластичным трафиком и нестационарной интенсивностью

Исследуемая система массового обслуживания используется для моделирования сценария скачивания файла в пределах одного сегмента сети с нестационарными скоростями - по поступлениям, размерам файлов и потерям из-за нетерпения. Это особый, но достаточно нестандартный, гетерогенный процесс рождения и гибели, для которого, в принципе, можно применить методы, описанные, например, в статьях [86, 87], но модель имеет свою специфику, в частности, медленную скорость сходимости. Обычно униформизация используется как метод расчета переходных вероятностей для цепей Маркова, как в [96, 97]. Однако, основанные на ней методы очень плохо работают в случае медленной сходимости. Кроме того, без предварительного понимания момента достижения предельного режима требуются значительные вычислительные усилия, чтобы быть хоть в какой-то степени уверенным в том, что полученное решение является искомым [98].

Результаты исследования модели опубликованы в [132].

5.3.1 Описание модели

Для описания потока запросов с переменным числом пользователей подходит пуассоновский поток первого рода со следующими параметрами: интенсивность поступления $\lambda(t)$, минимальное требование к ресурсу b и длина передаваемого блока данных - $\theta(t)$. В таблице 5.1 отражены основные параметры системы.

В рассматриваемой системе имеется ресурс объема C , запоминающее устройство с конечной емкостью r . Кроме того, запросы обладают свойством нетерпения - они покидают очередь с интенсивностью $\gamma(t)$. Будем считать, что длина блока равна некоторому значению $\theta(t)$. Весь объем C делится поровну между заказами, т.е. если число клиентов равно 1, то весь ресурс потребляется этим клиентом, а стоимость обслуживания равна $\frac{\theta(t)}{C}$; если число клиентов

Таблица 5.1: Параметры системы

Параметр	Описание
$\lambda(t)$	Интенсивность потока запросов на передачу эластичных данных
b	Минимальная гарантированная эластичная скорость передачи блоков данных
$\theta(t)$	Среднее значение длины блока данных
r	Очередь заявок на передачу блока эластичных данных
$\gamma(t)$	Интенсивность потерь
C	Пропускная способность сети (скорость обслуживания)

равно 2, то стоимость обслуживания составляет $\frac{2\theta(t)}{C}$ - объем ресурса делится пополам. В случае, когда C не может быть разделен поровну между клиентами с обеспечением минимального гарантированного порога b , в очередь заходит новый клиент.

Пусть $N(t) \in \{1, \dots, \lfloor \frac{C}{b} \rfloor\}$ - число обработанных заказов на данный момент $t \geq 0$. Следовательно, число $\lfloor \frac{C}{b} \rfloor = N$ - это максимальное число запросов, которые устройство может обрабатывать одновременно. Пространство состояний системы выглядит следующим образом $X = \{n \in 0, \dots, N, \dots, N + r : c(n) \leq C\}$

Легко видеть, что данная модель может быть описана марковским процессом $X(t), t > 0$, где $X(t)$ обозначает число клиентов в системе в момент времени t (процесс длиной очереди). Обозначим через $p_n(t) = P(X(t) = n)$, $n = 0, 1, 2, 3, \dots, N + r$.

Исходя из приведенных выше предположений, результирующее поведение вероятностей состояний описывается прямой системой Колмогорова:

$$p'_0(t) = -\lambda(t)p_0(t) + \frac{C}{\theta(t)}p_1(t) \quad (5.3.47)$$

$$p'_n(t) = \lambda(t)p_{n-1}(t) - \left(\frac{C}{\theta(t)} + \lambda(t) \right) p_n(t) + \frac{C}{\theta(t)}p_{n+1}(t), 1 \leq n < N \quad (5.3.48)$$

$$p'_N(t) = \lambda(t)p_{N-1}(t) - \left(\frac{C}{\theta(t)} + \lambda(t) \right) p_N(t) + \left(\frac{C}{\theta(t)} + \gamma(t) \right) p_{N+1}(t) \quad (5.3.49)$$

$$p'_n(t) = \lambda(t)p_{n-1}(t) - \left(\frac{C}{\theta(t)} + (n - N) \gamma(t) + \lambda(t) \right) p_n(t) +$$

$$+ \left(\frac{C}{\theta(t)} + (n+1-N)\gamma(t) \right) p_{n+1}(t), N < n < N+r \quad (5.3.50)$$

$$p'_{N+r}(t) = \lambda(t)p_{N+r-1}(t) - \left(\frac{C}{\theta(t)} + r\gamma(t) + \lambda(t) \right) p_{N+r}(t) \quad (5.3.51)$$

Теперь рассмотрим соответствующую нестационарную ситуацию. А именно, предполагаем, что процесс длиной в очередь $\{X(t), t \geq 0\}$ является неоднородной цепью Маркова с непрерывным временем. Предполагается, что все возможные интенсивности перехода, скажем, $q_{ij}(t)$, являются неслучайными функциями времени. Предполагаем, что все функции интенсивности неотрицательны и локально интегрируемы на $[0, \infty)$.

Пусть $\mathbf{p}(t) = (p_0(t), p_1(t), p_2(t), \dots, p_{N+r}(t))^T$ вектор вероятностей состояний в момент времени t . Положим $a_{ij}(t) = q_{ji}(t)$ для $j \neq i$ и $a_{ii}(t) = -\sum_{j \neq i} a_{ji}(t) = -\sum_{j \neq i} q_{ij}(t)$.

Можем рассматривать прямую систему Колмогорова (5.3.47) – (5.3.51) как дифференциальное уравнение

$$\frac{d\mathbf{p}(t)}{dt} = A(t)\mathbf{p}(t), \quad (5.3.52)$$

в пространстве последовательностей l_1 , где $A(t)$ является ограниченным почти для всех $t \geq 0$ линейным оператором в l_1 с матрицей интенсивностей

$$\begin{pmatrix} -\lambda(t) & \mu_1(t) & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \lambda(t) & -(\mu_1(t) + \lambda(t)) & \cdots & 0 & 0 & \cdots & 0 & 0 \\ 0 & \lambda(t) & \cdots & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & -(\mu_N(t) + \lambda(t)) & \mu_{N+1}(t) & \cdots & 0 & 0 \\ 0 & 0 & \cdots & \lambda(t) & -(\mu_{N+1}(t) + \lambda(t)) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & \lambda(t) & -\mu_{N+r}(t) \end{pmatrix}, \quad (5.3.53)$$

где $\mu(t) = \frac{C}{\theta(t)}$, и $\mu_n(t) = \mu(t) + \max(0, n - N)\gamma(t)$ для $n \geq 1$.

Показатели эффективности

Для анализа системы рассмотрим некоторые характеристики модели. А именно, вероятность блокировки входящего приложения

$$P_{block}(t) = p_{N+r}(t), \quad (5.3.54)$$

среднее число обслуживаемых заявок

$$\bar{C}(t) = \sum_{i=1}^N ip_i(t) + N \cdot \sum_{i=1}^r p_{N+i}(t), \quad (5.3.55)$$

среднее число заявок в очереди

$$Q(t) = \sum_{i=N+1}^{N+r} (i - N)p_i(t), \quad (5.3.56)$$

среднее время обслуживания заявок

$$T(t) = \frac{\bar{C}(t)}{\lambda(t)(1 - P_{block}(t))}, \quad (5.3.57)$$

среднее время пребывания заявок в очереди

$$W(t) = \frac{Q(t)}{\lambda(t)(1 - P_{block}(t))}. \quad (5.3.58)$$

5.3.2 Получение оценок скорости сходимости

Предполагая, что $p_0 = 1 - \sum_{i=1}^{N+r} p_i(t)$, тогда из (5.3.52) получаем следующее уравнение

$$\frac{d\mathbf{p}}{dt} = B(t)\mathbf{p} + \mathbf{g}(t), \quad t \geq 0, \quad (5.3.59)$$

где $\mathbf{g}(t) = (\lambda(t), 0, 0, \dots, 0)^T$ и $B(t)$ равна

$$\begin{pmatrix} -(\mu_1(t) + 2\lambda(t)) & \mu_2(t) - \lambda(t) & -\lambda(t) & -\lambda(t) & \cdots & -\lambda(t) & -\lambda(t) \\ \lambda(t) & -(\mu_2(t) + \lambda(t)) & \mu_3(t) & 0 & \cdots & 0 & 0 \\ 0 & \lambda(t) & -(\mu_3(t) + \lambda(t)) & \mu_4(t) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & -(\mu_{N+r-1}(t) + \lambda(t)) & \mu_{N+r}(t) \\ 0 & 0 & 0 & 0 & \cdots & \lambda(t) & -\mu_{N+r}(t) \end{pmatrix}. \quad (5.3.60)$$

Решение этого уравнения может быть представлено в следующем виде

$$\mathbf{p}(t) = U^*(t, 0) \mathbf{p}(0) + \int_0^t U^*(t, \tau) \mathbf{g}(\tau) d\tau, \quad (5.3.61)$$

где $U^*(t, s)$ оператор Коши соответствующего однородного уравнения

$$\frac{d\mathbf{x}}{dt} = B(t) \mathbf{x}. \quad (5.3.62)$$

Далее рассмотрим оценки в "взвешенных" нормах. Предположим, что d_1, d_2, \dots, d_{N+r} являются положительными числами. Тогда

$$D = \begin{pmatrix} d_1 & d_1 & \cdots & d_1 \\ 0 & d_2 & \cdots & d_2 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & d_{N+r} \end{pmatrix} \quad (5.3.63)$$

Обозначим через $\|\mathbf{z}\|_{1D} = \|D\mathbf{z}\|_1$. Тогда получаем $B^{**}(t) = DB(t)D^{-1} =$

$$\begin{pmatrix} -(\mu_1(t) + \lambda(t)) & \frac{d_1}{d_2} \mu_1(t) & 0 & \cdots & 0 & 0 \\ \frac{d_2}{d_1} \lambda(t) & -(\mu_2(t) + \lambda(t)) & \frac{d_2}{d_3} \mu_2(t) & \cdots & 0 & 0 \\ 0 & \frac{d_3}{d_2} \lambda(t) & -(\mu_3(t) + \lambda(t)) & \cdots & 0 & 0 \\ 0 & 0 & \frac{d_4}{d_3} \lambda(t) & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -(\mu_{N+r-1}(t) + \lambda(t)) & \frac{d_{N+r-1}}{d_{N+r}} \mu_{N+r-1}(t) \\ 0 & 0 & 0 & \cdots & \frac{d_{N+r}}{d_{N+r-1}} \lambda(t) & -(\mu_{N+r}(t) + \lambda(t)) \end{pmatrix}. \quad (5.3.64)$$

Положим

$$\gamma_{**}(t) = \inf_i \left(|b_{ii}(t)| - \sum_{j \neq i} \frac{d_j}{d_i} b_{ji}(t) \right). \quad (5.3.65)$$

Тогда получим

$$\gamma(B(t))_{1D} = \gamma(DB(t)D^{-1}) = \sup_i \left(b_{ii}(t) + \sum_{j \neq i} \frac{d_j}{d_i} b_{ji}(t) \right) = -\gamma_{**}(t). \quad (5.3.66)$$

Пусть $-\alpha_k(t)$ – это сумма элементов k -столбца, тогда

$$\alpha_1(t) = \mu_1(t) + \lambda(t) - \frac{d_2}{d_1} \lambda(t) = \mu(t) - \left(\frac{d_2}{d_1} - 1 \right) \lambda(t),$$

$$\begin{aligned}
\alpha_k(t) &= \mu_k(t) + \lambda(t) - \frac{d_{k-1}}{d_k} \mu_{k-1}(t) - \frac{d_{k+1}}{d_k} \lambda(t) = \\
&= \left(1 - \frac{d_{k-1}}{d_k}\right) \mu(t) - \left(\frac{d_{k+1}}{d_k} - 1\right) \lambda(t), \quad 2 \leq k \leq N, \\
\alpha_{N+1}(t) &= \mu_{N+1}(t) + \lambda(t) - \frac{d_N}{d_{N+1}} \mu_N(t) - \frac{d_{N+2}}{d_{N+1}} \lambda(t) = \\
&= \left(1 - \frac{d_N}{d_{N+1}}\right) \mu(t) + \gamma(t) - \left(\frac{d_{N+2}}{d_{N+1}} - 1\right) \lambda(t), \\
\alpha_k(t) &= \mu_k(t) + \lambda(t) - \frac{d_{k-1}}{d_k} \mu_{k-1}(t) - \frac{d_{k+1}}{d_k} \lambda(t) = \\
&= \left(1 - \frac{d_{k-1}}{d_k}\right) (\mu(t) + (k - N)\gamma(t)) + \frac{d_{k-1}}{d_k} \gamma(t) - \left(\frac{d_{k+1}}{d_k} - 1\right) \lambda(t), \\
&\quad N + 2 \leq k < N + r, \\
\alpha_{N+r}(t) &= \mu_{N+r}(t) - \frac{1}{\delta} \mu_{N+r-1}(t) + \lambda(t) = \\
&= \left(1 - \frac{d_{k-1}}{d_k}\right) (\mu(t) + r\gamma(t)) + \frac{d_{k-1}}{d_k} \gamma(t) + \lambda(t)
\end{aligned} \tag{5.3.67}$$

и $\gamma_{**}(t) = \inf_k \alpha_k$.

Оценки скорости сходимости для больших интенсивностей обслуживания

Для некоторого положительного $\delta > 1$ положим $d_k = \delta^{k-1}$ для $k \geq 1$. Тогда получим следующее

$$\begin{aligned}
\alpha_1(t) &= \mu_1(t) + \lambda(t) - \delta \lambda(t) = \mu(t) - (\delta - 1) \lambda(t), \\
\alpha_k(t) &= \mu_k(t) + \lambda(t) - \frac{1}{\delta} \mu_{k-1}(t) - \delta \lambda(t) \geq \left(1 - \frac{1}{\delta}\right) \mu_{k-1}(t) - (\delta - 1) \lambda(t) \geq \\
&\quad \left(1 - \frac{1}{\delta}\right) (\mu(t) - \delta \lambda(t)), \quad 2 \leq k < N + r, \\
\alpha_{N+r}(t) &= \mu_{N+r}(t) - \frac{1}{\delta} \mu_{N+r-1}(t) + \lambda(t) \geq \left(1 - \frac{1}{\delta}\right) (\mu(t) - \delta \lambda(t)).
\end{aligned} \tag{5.3.68}$$

Следовательно

$$\gamma_{**}(t) = \min(\alpha_i(t)) = \left(1 - \frac{1}{\delta}\right) (\mu(t) - \delta \lambda(t)). \tag{5.3.69}$$

Теорема 29 Пусть существует положительное число $\delta > 1$ такое, что

$$\int_0^\infty (\mu(t) - \delta \lambda(t)) dt = +\infty. \tag{5.3.70}$$

Тогда цепь Маркова $X(t)$ слабо эргодична и имеет следующие оценки скорости сходимости:

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{1D} \leq e^{-\int_0^t (1-\frac{1}{\delta})(\mu(t)-\delta\lambda(t)) d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1D}, \quad (5.3.71)$$

$$\begin{aligned} \|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| &\leq 4\delta^{N+r} e^{-\int_0^t (1-\frac{1}{\delta})(\mu(t)-\delta\lambda(t)) d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\| \leq \\ &\leq 8\delta^{N+r} e^{-\int_0^t (1-\frac{1}{\delta})(\mu(t)-\delta\lambda(t)) d\tau}, \end{aligned} \quad (5.3.72)$$

для любых начальных условий $\mathbf{p}^*(0), \mathbf{p}^{**}(0)$ и любого $t \geq 0$.

Положим $W = \min_{k \geq 1} \frac{d_k}{k} = \min_{k \geq 0} \frac{\delta^k}{k+1}$. Тогда получим $W \|\mathbf{p}\|_{1E} \leq \|\mathbf{p}\|_{1D}$.

Следствие 1. Из условий теоремы 29 $X(t)$ имеет предельное среднее значение, тогда $\phi(t) = E(t, 0)$, и получаем, что следующая оценка верна для любого j и любого $t \geq 0$:

$$|E(t, j) - E(t, 0)| \leq \frac{1 + \delta^{j-1}}{W} e^{-\int_0^t (1-\frac{1}{\delta})(\mu(t)-\delta\lambda(t)) d\tau}. \quad (5.3.73)$$

Оценки скорости сходимости для больших интенсивностей поступления

Теперь рассмотрим случай $\delta < 1$, предположим, что $\delta \in [\frac{r-1}{r}, 1)$ и последовательность $d_k = \delta^{k-1}$ для $1 \leq k \leq N$ и $d_k = \delta^{N-1}$ для $k > N$. В этом случае

$$\begin{aligned} \alpha_1(t) &\geq \mu(t) + (1 - \delta)\lambda(t), \\ \alpha_k(t) &= (1 - \frac{1}{\delta}) \mu(t) - (\delta - 1) \lambda(t) \geq (\frac{1}{\delta} - 1) (\delta\lambda(t) - \mu(t)), \quad 2 \leq k \leq N, \\ \alpha_k(t) &= \gamma(t), \quad N + 1 \leq k < N + r, \\ \alpha_{N+r}(t) &= \gamma(t) + \lambda(t) \geq \gamma(t) \end{aligned} \quad (5.3.74)$$

Тогда из этого следует, что

$$\gamma_{**}(t) = \min(\alpha_i(t)) = \min\left(\left(\frac{1}{\delta} - 1\right) (\delta\lambda(t) - \mu(t)), \gamma(t)\right) \quad (5.3.75)$$

Теорема 30 Пусть

$$\int_0^{\infty} \gamma_{**}(t) dt = +\infty, \quad (5.3.76)$$

для некоторого $\delta \in [\frac{r-1}{r}, 1)$. Тогда цепь Маркова $X(t)$ слабо эргодична и имеет следующие оценки скорости сходимости:

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{1D} \leq e^{-\int_0^t \gamma_{**}(\tau) d\tau} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1D}, \quad (5.3.77)$$

и

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq \frac{8}{\delta^{N-1}} e^{-\int_0^t \gamma_{**}(\tau) d\tau}, \quad (5.3.78)$$

для любых начальных условий $\mathbf{p}^*(0), \mathbf{p}^{**}(0)$ и любого $t \geq 0$. Кроме того, также существует предельное среднее и оценка (5.3.73).

Замечание 1. Если процесс однороден (т.е. все интенсивности постоянны), то условия теоремы 29 и теоремы 30 эквивалентны неравенствам $\mu > \lambda$ и $\mu < \lambda$ соответственно.

Замечание 2. Если все интенсивности процесса являются 1-периодическими, то слабая эргодичность $X(t)$ и оценки теоремы 29 или теоремы 30 существуют если

$$\int_0^1 \lambda(t) dt \neq \int_0^1 \mu(t) dt. \quad (5.3.79)$$

5.3.3 Получение оценок устойчивости

Рассмотрим "возмущенный" процесс длины очереди $\bar{X}(t), t \geq 0$ с соответствующей транспонированной матрицей интенсивности $\bar{A}(t)$, где "возмущающая" матрица $\hat{A}(t) = A(t) - \bar{A}(t)$ в некотором смысле мала. То есть предполагаем, что возмущенная очередь имеет ту же природу, что и исходная. Тогда матрица возмущенной интенсивности также имеет ту же структуру с соответствующими возмущенными интенсивностями $\bar{\theta}(t), \bar{\gamma}(t), \bar{\lambda}(t)$. Предполагаем, что $\bar{\mu}_n(t) = \frac{C}{\bar{\theta}(t)} + \max(0, n - N)\bar{\gamma}(t)$ для $n \geq 1$.

Предполагаем, что

$$\left| \frac{1}{\theta(t)} - \frac{1}{\hat{\theta}(t)} \right| = \left| \frac{1}{\hat{\theta}(t)} \right| \leq \hat{\epsilon}, |\gamma(t) - \bar{\gamma}(t)| = |\hat{\gamma}(t)| \leq \hat{\epsilon}, |\lambda(t) - \bar{\lambda}(t)| = |\hat{\lambda}(t)| \leq \hat{\epsilon}. \quad (5.3.80)$$

Следовательно, получим следующее

$$\begin{aligned} |\mu_n(t) - \bar{\mu}_n(t)| &= |\hat{\mu}_n(t)| = \left| \frac{C}{\theta(t)} + \max(0, n - N)\gamma(t) - \frac{C}{\hat{\theta}(t)} - \right. \\ &\quad \left. - \max(0, n - N)\bar{\gamma}(t) \right| \leq \\ &\leq \left| \frac{C}{\theta(t)} - \frac{C}{\hat{\theta}(t)} \right| + \max(0, n - N) |(\gamma(t) - \bar{\gamma}(t))| \leq \\ &\leq C\hat{\epsilon} + r\hat{\epsilon} = (C + r)\hat{\epsilon} \end{aligned} \quad (5.3.81)$$

Затем из (5.3.53) получаем следующую оценку

$$\begin{aligned} \|\hat{A}(t)\| &= 2 \sup_k |\hat{a}_{kk}(t)| = 2 \max \left(|\hat{\lambda}(t)|, |\hat{\lambda}(t)| + |\hat{\mu}_n(t)|, |\hat{\mu}_{N+r}(t)| \right) \leq \\ &\leq 2(C + r + 1)\hat{\epsilon}. \end{aligned} \quad (5.3.82)$$

Теперь из теоремы 1 и следствия 1 статьи [86] следуют следующие оценки возмущения.

Теорема 31 *Предположим, что по условиям теоремы 29 или теоремы 30 цепь Маркова $X(t)$ экспоненциально эргодична, т.е.*

$$e^{-\int_s^t \gamma_{**}(\tau) d\tau} \leq K e^{-\gamma_0(t-s)}, \quad (5.3.83)$$

для положительных K, γ_0 . Тогда имеют место следующие оценки возмущения:

$$\begin{aligned} \limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\| &\leq \\ &\leq \frac{2\hat{\epsilon}(C + r + 1)(1 + \log(4K) + (N + r)|\log(\delta)|)}{\gamma_0} \end{aligned} \quad (5.3.84)$$

u

$$\begin{aligned} & \limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq \\ & \leq \frac{2(N+r)\hat{\epsilon}(C+r+1)(1+\log(4K) + (N+r)|\log(\delta)|)}{\gamma_0}, \end{aligned} \quad (5.3.85)$$

для любой возмущенной очереди с соответственно замкнутыми интенсивностями, удовлетворяющими (5.3.80).

5.3.4 Пример

Рассмотрим пример, в котором объем ресурсного блока равен $C = 100$ Мбит/с, а накопитель конечной емкости равен $r = 100$. Размер передаваемого файла равен $\theta(t) = \theta = 10$ Мбайт, то есть 80 Мбит, а минимальная скорость передачи составляет $b = 1$ Мбит/с. Интенсивность поступления равна $\lambda(t) = 3 \cdot a(t)$, где

$$\begin{aligned} a(t) = & 0.6521 - 0.08888 \cdot \cos(wt) - 0.3963 \cdot \sin(wt) + \\ & + 0.1384 \cdot \cos(2wt) - 0.1411 \cdot \sin(2wt) - \\ & - 0.01446 \cdot \cos(3wt) + 0.05875 \cdot \sin(3wt) \end{aligned} \quad (5.3.86)$$

с $w = 0.2603$ и интенсивность потери нетерпеливых заявок $\gamma(t) = \gamma = 0.01$.

Положим $\delta = 0.99$, $d_k = \delta^{k-1}$ для $1 \leq k \leq N$ и $d_k = \delta^{N-1}$ для $k > N$. Очевидно, что $\gamma_{**}(t) = \gamma(t) = 0.01$ и верно неравенство

$$e^{-\int_s^t \gamma_{**}(\tau) d\tau} \leq e^{-0.016(t-s)}, \quad (5.3.87)$$

следовательно, можно получить $K = 1$ и $\gamma_0 = 0.01$ в (5.3.83).

Тогда из теоремы 30 получаем оценки скорости сходимости

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\|_{1D} \leq e^{-0.01t} \|\mathbf{p}^*(0) - \mathbf{p}^{**}(0)\|_{1D}$$

и

$$\|\mathbf{p}^*(t) - \mathbf{p}^{**}(t)\| \leq 22e^{-0.01t}.$$

Для среднего из следствия 1 теоремы 30 верна оценка

$$|E(t, j) - E(t, 0)| \leq \frac{1 + 0.99^{j-1}}{W} e^{-0.01t}$$

Из теоремы 31 оценки возмущения имеют вид

$$\limsup_{t \rightarrow \infty} \|\mathbf{p}(t) - \bar{\mathbf{p}}(t)\| \leq 10^5 \hat{\epsilon}$$

и

$$\limsup_{t \rightarrow \infty} |E(t, 0) - \bar{E}(t, 0)| \leq 2 * 10^7 \hat{\epsilon}.$$

На рисунках 5.11-5.19 показаны графики сходимости характеристик, указанных в разделе 5.3.1 для двух начальных условий $p_0(0) = 1$ ($X(0) = 0$) и $p_{200}(0) = 1$ ($X(0) = 200$).

На рисунках 5.11 и 5.12 показаны графики вероятности блокировки. Отметим, что правая часть неравенства 5.3.78 в момент времени $t = 1000$ превышает 10^{-3} , а период колебания составляет 24. Далее, на рисунках 5.13 и 5.14 заметим, что среднее число запросов на обслуживание сходится на уровне 100, это может указывать на высокую загрузку устройств в системе. Однако вероятность блокировки невелика.

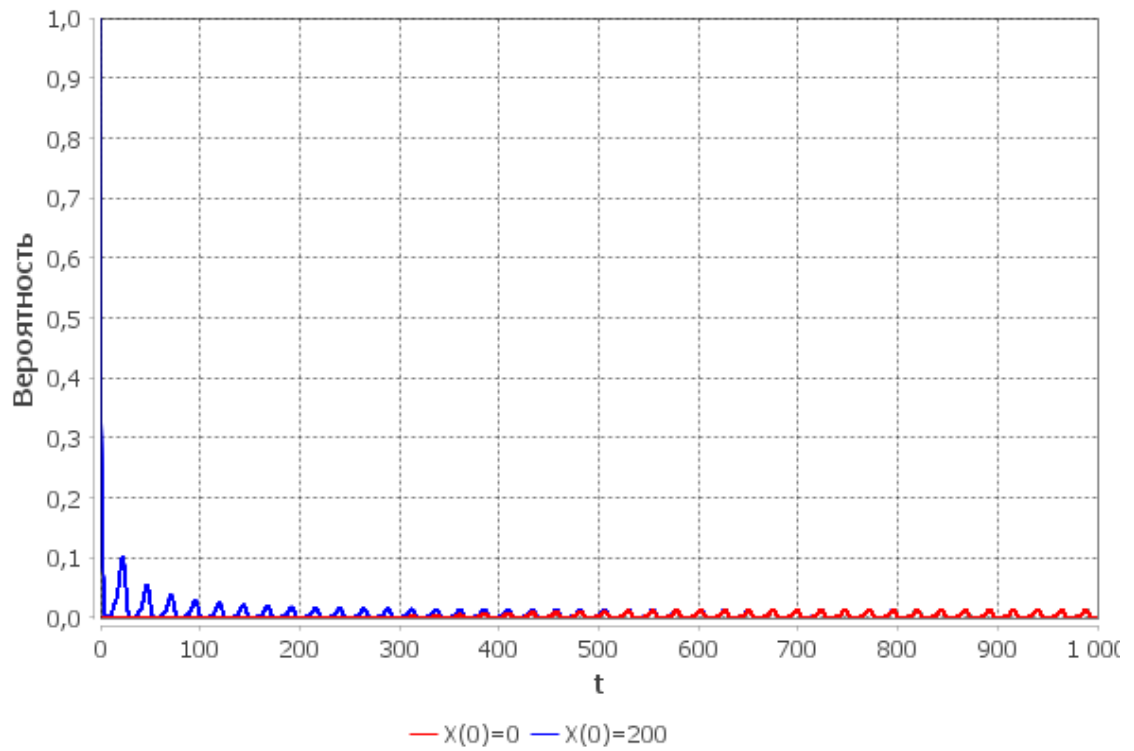


Рис. 5.11: Вероятность блокировки для $t \in [0, 1000]$.

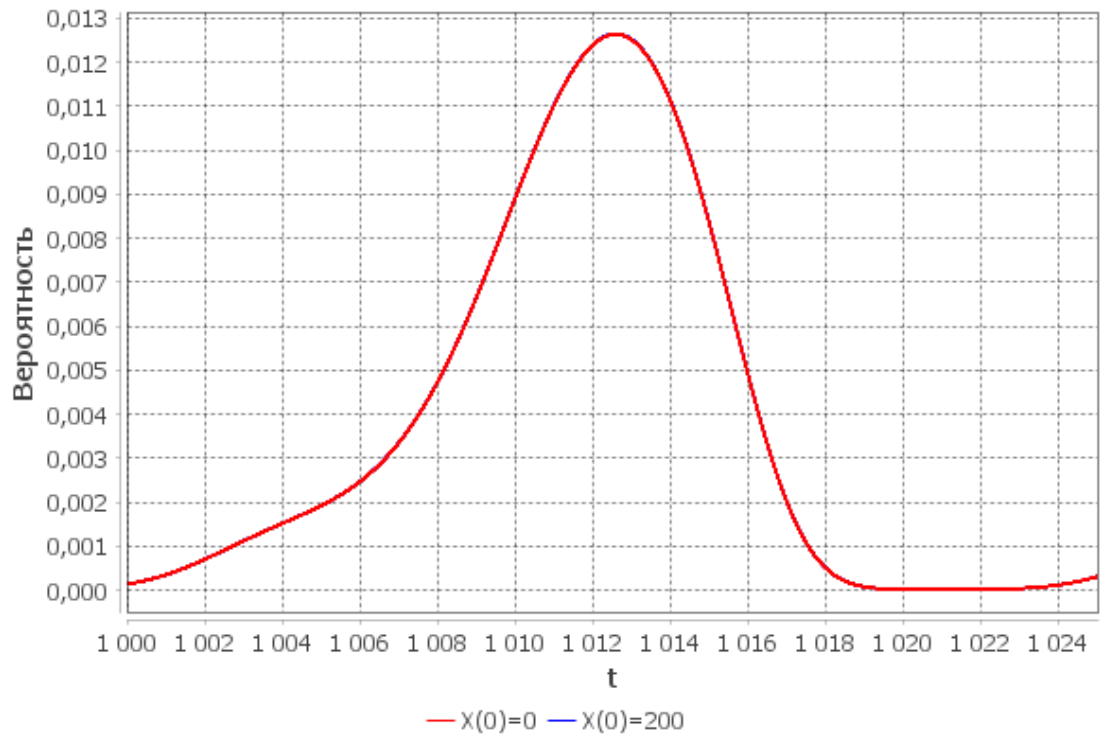


Рис. 5.12: Аппроксимация вероятности блокировки для $t \in [1000, 1025]$.

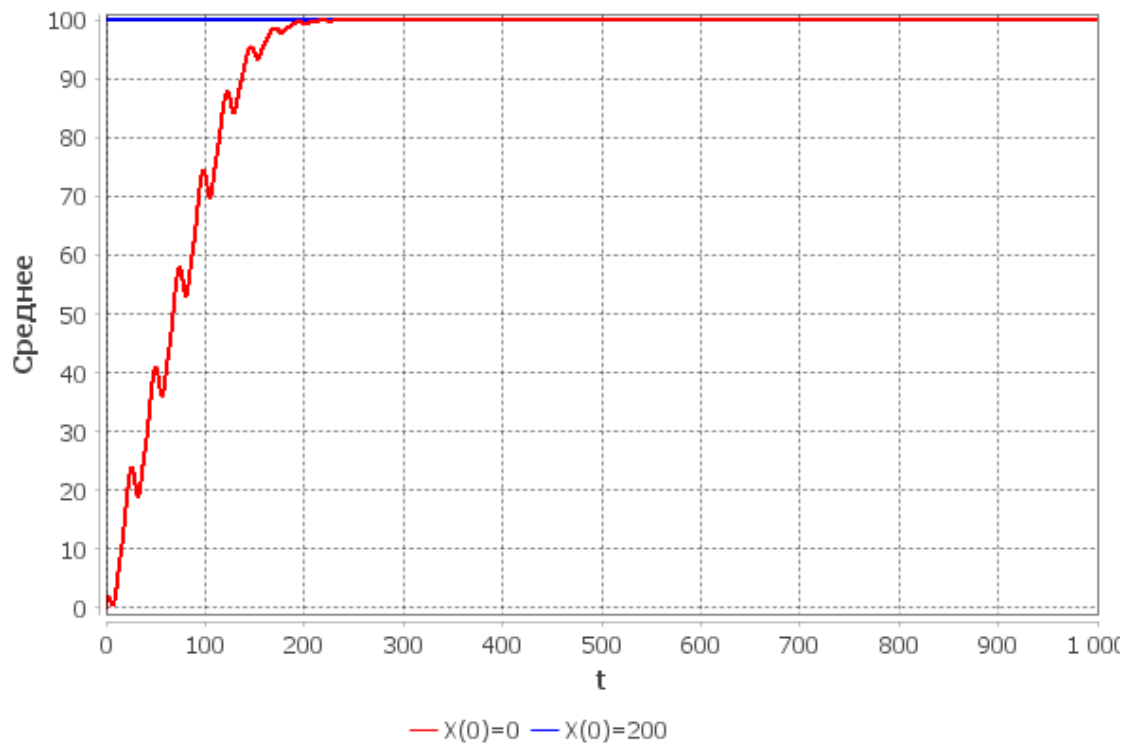


Рис. 5.13: Среднее число поданных заявок $E(t, k)$ для $t \in [0, 1000]$.

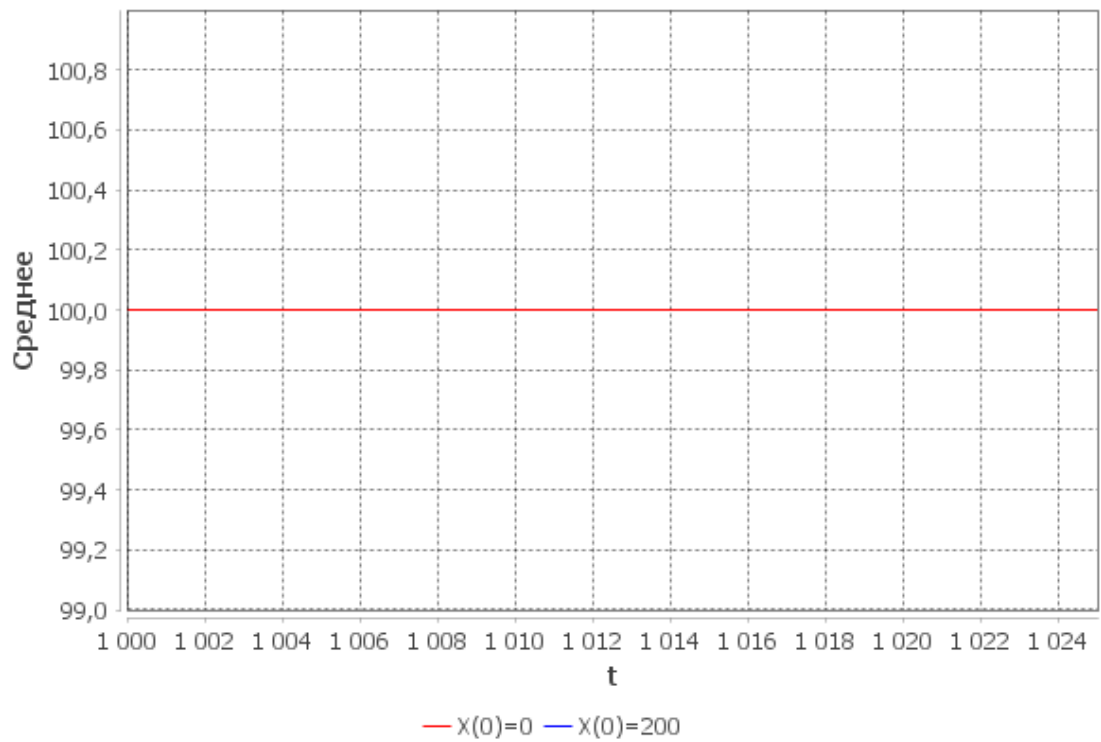


Рис. 5.14: Аппроксимация среднего числа поданных заявок $E(t, k)$ для $t \in [1000, 1025]$.

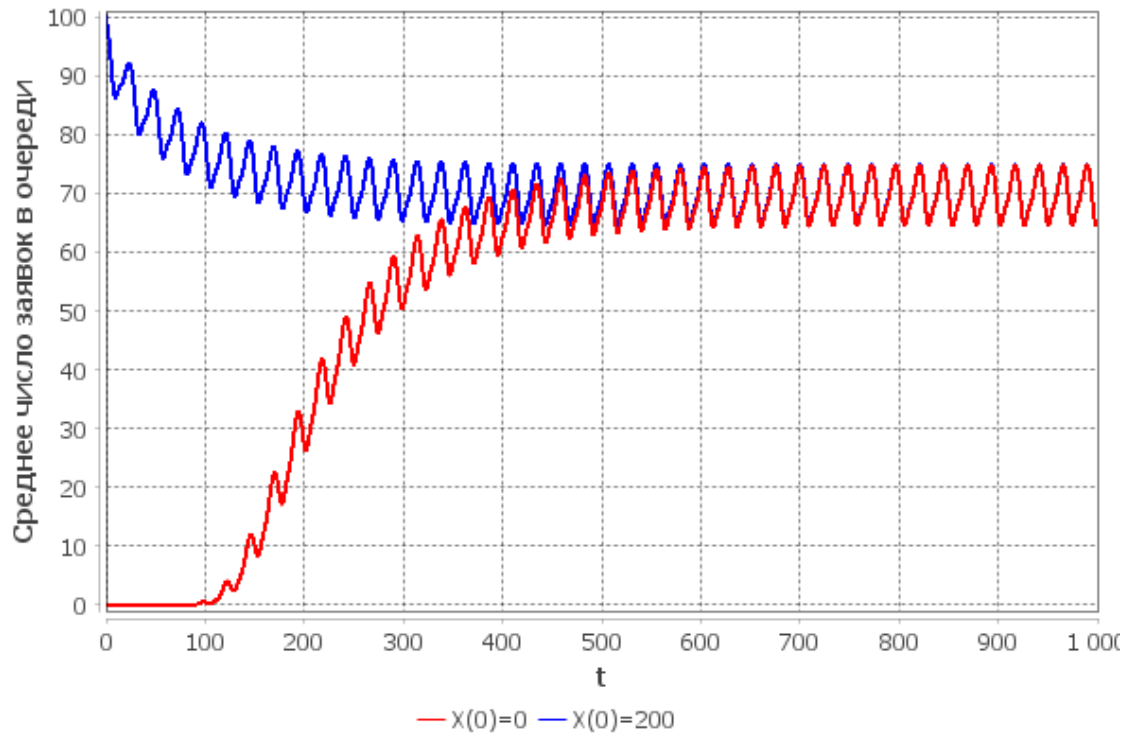


Рис. 5.15: Среднее число заявок в очереди $Q(t)$ для $t \in [0, 1000]$.

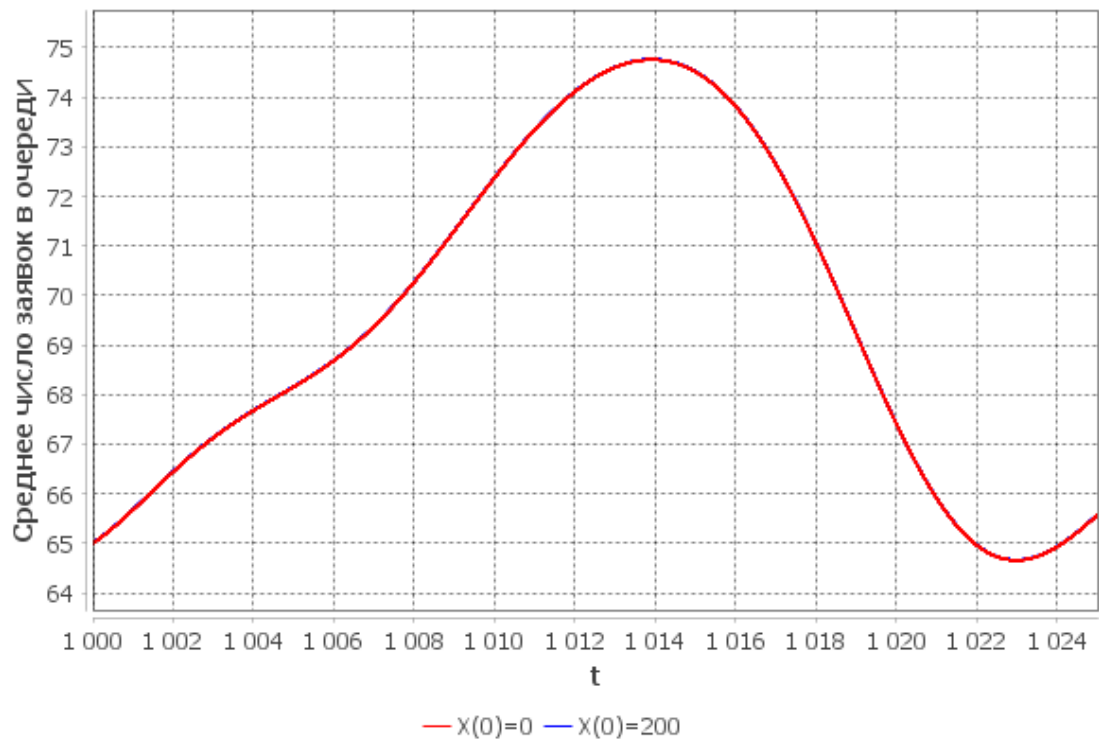


Рис. 5.16: Аппроксимация среднего числа заявок в очереди $Q(t)$ для $t \in [1000, 1025]$.

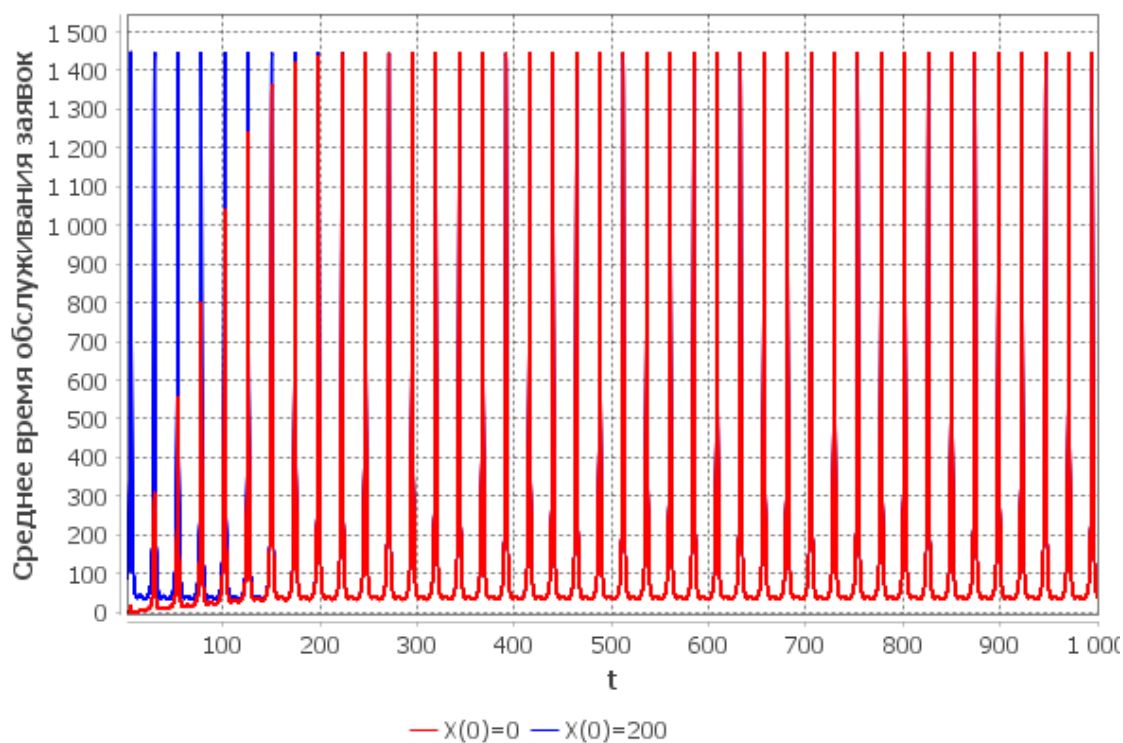


Рис. 5.17: Среднее время обслуживания заявок $Q(t)$ для $t \in [0, 1000]$.

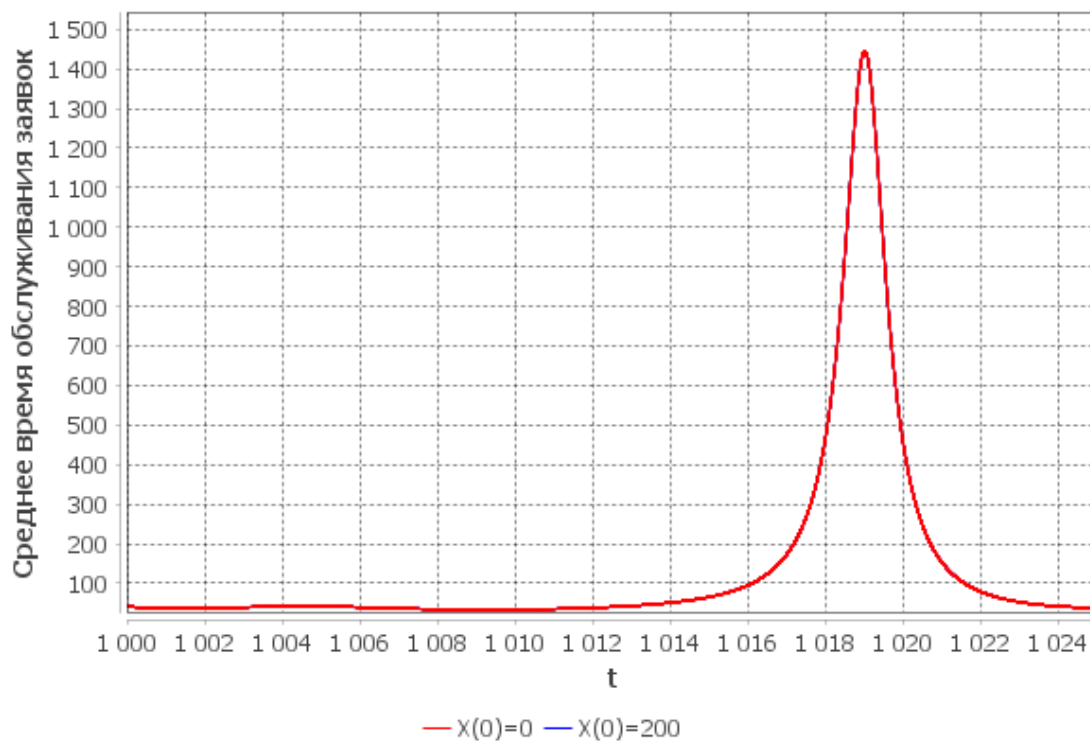


Рис. 5.18: Аппроксимация среднего времени обслуживания заявок $Q(t)$ для $t \in [1000, 1025]$.

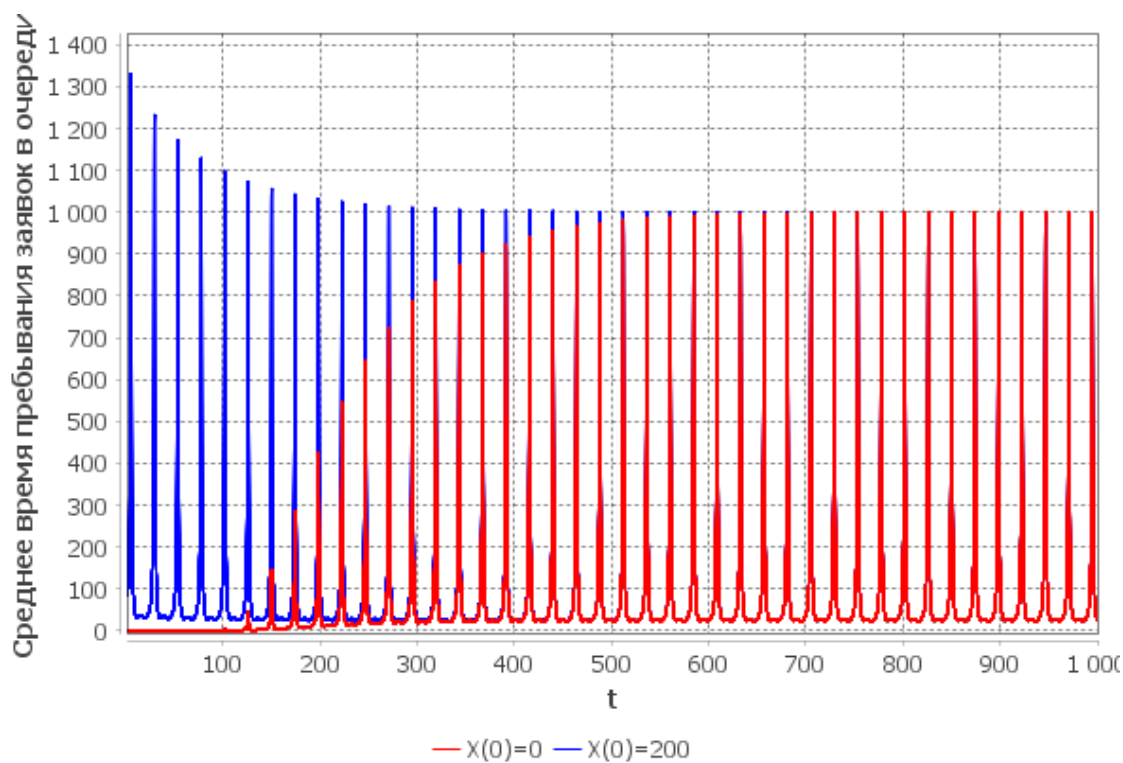


Рис. 5.19: Среднее время пребывания заявок в очереди $Q(t)$ для $t \in [0, 1000]$.

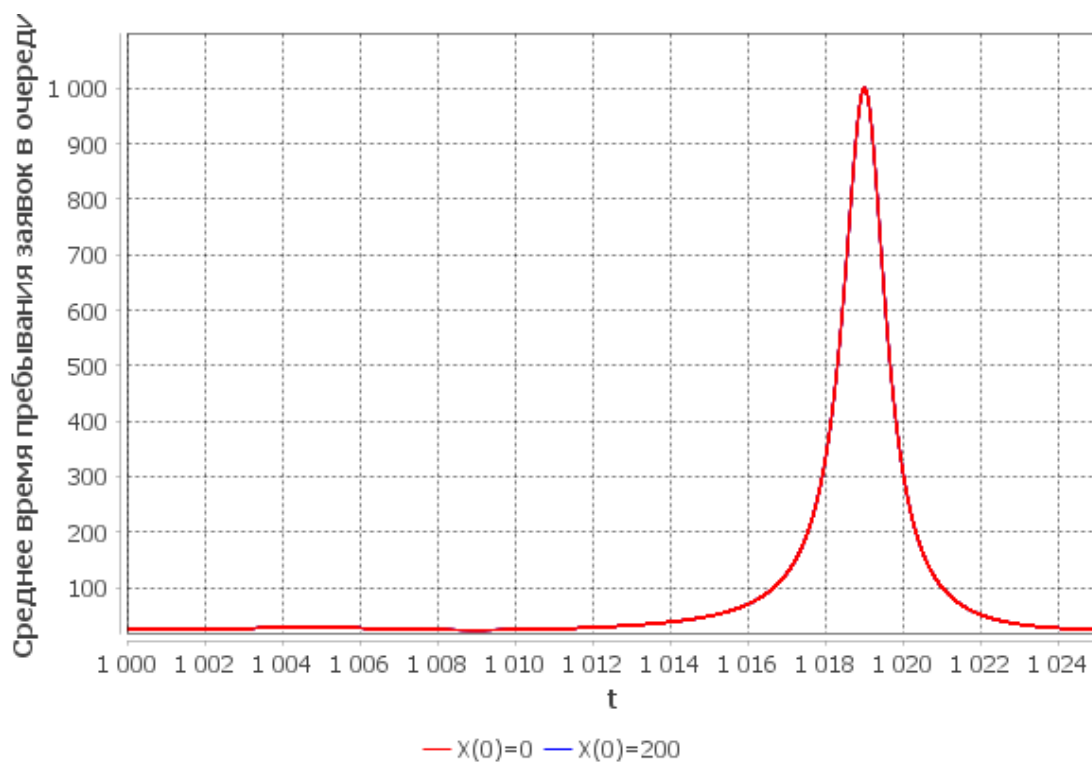


Рис. 5.20: Аппроксимация среднего времени пребывания заявок в очереди $Q(t)$ для $t \in [1000, 1025]$.

Заключение

Основные результаты

В ходе решения поставленных в диссертационной работе задач получены оценки скорости сходимости и устойчивости для

- системы массового обслуживания типа $M_t/M_t/1$ с отказами, катастрофами, сбоями и ремонтами сервера;
- системы массового обслуживания с одним сервером, специальными групповыми поступлениями требований и специальной политикой пропуска очереди;
- системы массового обслуживания с одним сервером, специальными групповыми поступлениями требований, специальной политикой пропуска очереди и катастрофами;
- системы массового обслуживания с групповым поступлением и групповым обслуживанием требований с управлением, зависящим от состояния;
- системы массового обслуживания с нетерпеливыми клиентами;
- системы массового обслуживания с эластичным трафиком и нестационарной интенсивностью.

Оценки применены для построения предельных характеристик.

Для системы массового обслуживания с одним сервером, специальными групповыми поступлениями требований и специальной политикой пропуска очереди получены

- оценки мощности сервера и мощности потока, при которой среднее число требований в системе не превышает заданного числа;

- границы интенсивности обслуживания и границы интенсивности поступления требований, чтобы среднее оставалось в заданных границах.

Для каждой из моделей с помощью разработанных алгоритмов и программ с применением численных методов проведены вычислительные эксперименты, подтверждающие теоретические выводы.

Результаты, полученные в данной диссертационной работе, позволяют расширить круг практических задач, допускающих аналитическое исследование важнейших вероятностных характеристик.

Описание программы

Описание программы для решения задачи Коши для системы однородных дифференциальных уравнений 1-го порядка методом Адамса-Мултона 4-го порядка. Функционал программы позволяет находить решения задачи Коши методом Адамса-Мултона 4-го порядка по следующим входным данным: матрица размерности $N + 1$, соответствующая системе однородных дифференциальных уравнений I -ого порядка, пример: матрица $A(t)$; пара начальных условий, например, два начальных ($t = 0$) N -мерных вектора: $p^1(0) = (1, 0, \dots, 0)$ и $p^2(0) = (0, \dots, 0, 1)$; отрезок времени и временной шаг, необходимый для схемы вычисления метода Рунге-Кутты, пример: 0.001. По найденным результатам для обоих начальных условий: вектор-решение по определенной координате и математическое ожидание, строятся графики с целью иллюстрирования предполагаемой сходимости решений. Программа реализована на языке Java с применением функционала библиотеки JFreeChart, используемого для построения различного вида графиков.

Включает следующие функциональные возможности:

1. решение систем ОДУ 1-го порядка с достаточно большим количеством уравнений,
2. Построение графиков вероятностных характеристик на основании результатов, полученных в процессе решения системы.

Программа для проведения вычислительных экспериментов [137] имеет свидетельство о государственной регистрации в Реестре программ для ЭВМ за номером 2020615415 от 22 мая 2020 года.

Литература

1. Андреев, Д. Эргодичность и устойчивость нестационарных систем обслуживания / Д. Андреев, М. Елесин, А. Кузнецов, Е. Крылов, А. Зейфман // Теория вероятностей и математическая статистика. – 2003. – т. 68. – с.1-11.
2. Анисимов, В.В. Оценки отклонений переходных характеристик неоднородных марковских процессов / В.В. Анисимов // Укр.матж.ж. – 1988. – 40. – с.699-706.
3. Афанасьева, Л.Г., Булинская, Е.В. Случайные процессы в теории массового обслуживания / Л.Г. Афанасьева, Е.В. Булинская. – М.: Изд-во МГУ. – 1980.
4. Башарин, Г.П., Харкевич, А.Д., Шнепс, М.А. Массовое обслуживание в телефонии / Г.П. Башарин, А.Д. Харкевич, М.А. Шнепс, – Наука. – 1968.
5. Башарин, Г.П., Самуйлов, К.Е., Яркина, Н.В., Гудкова, И.А. Новый этап развития математической теории телетрафика/ Г.П. Башарин, К.Е. Самуйлов, Н.В. Яркина, И.А. Гудкова // Автоматика и телемеханика. - 2009. – N12. – 16-28.
6. А. К. Горшенин, С. А. Горбунов, Д. Ю. Волканов. О кластеризации объектов сетевой вычислительной инфраструктуры на основе анализа статистических аномалий в трафике, Информ. и её примен., 17:3 (2023), 76–87
7. Nazarov, A., Moiseev, A., Moiseeva, S. (2021). Mathematical Model of Call Center in the Form of Multi-Server Queueing System. Mathematics, 9(22), 2877.
8. Basharin G. P., Gaidamaka Y. V., Samouylov K. E. (2013) Mathematical theory of teletraffic and its application to the analysis of multiservice communication of next generation networks, Autom. Control Comput. Sci., 47(2), 62–69.

9. Боровков, А.А. Эргодичность и устойчивость случайных процессов / А.А. Боровков. – М.: Эдиториал УРСС. – 1999.
10. Бочаров, П.П., Печинкин, А.В. Теория массового обслуживания / П.П. Бочаров, А.В. Печинкин. – М.: Изд-во Рос. ун-та дружбы народов. – 1995.
11. Гнеденко, Б. В., Макаров, И. П. Свойства решений задачи с потерями в случае периодических интенсивностей / Б.В. Гнеденко, И.П. Макаров // Дифференциальные уравнения. – 1971. – №9. – с. 1696–1698.
12. Далецкий, Ю.Л. Устойчивость решений дифференциальных уравнений в банаховом пространстве / Ю.Л. Далецкий, М.Г. Крейн. – М.: Наука. – 1970.
13. Зейфман, А. И. О погрешности усечения системы рождения и гибели // Ж. вычисл. матем. и матем. физ. – 1988. – 28, №12, 1906-1907.
14. Зейфман, А. И. Некоторые свойства системы с потерями в случае переменных интенсивностей // Автоматика и телемеханика. – 1989. – №1, с. 107–113.
15. Зейфман, А.И. Стохастические модели. Процессы рождения и гибели / А.И. Зейфман. – Вологда: Издательство "Русь". – 1994.
16. Зейфман, А. И., Бенинг, В. Е., Соколов И. А. Марковские цепи и модели с непрерывным временем. М.: Элекс-КМ. – 2008.
17. Зейфман, А.И. О нестационарной модели Эрланга // Автоматика и телемеханика. – 2009. – №12.– с.71–80.
18. Зейфман, А.И., Королев, В.Ю., Коротышева, А.В., Шоргин, С.Я. Общие оценки устойчивости для нестационарных марковских цепей с непрерывным временем. Информатика и ее применения. – 2014. – 8, вып. 1, 106–117.
19. Зейфман, А.И., Коротышева, А.В., Киселева, К.М., Королев, В.Ю., Шоргин, С.Я. Об оценках скорости сходимости и устойчивости для некоторых моделей массового обслуживания.// Информатика и ее применения. – 2014. – 8, вып. 3, 19-27.

20. Зейфман, А.И., Сатин, Я.А., Коротышева, А.В., Королев, В.Ю., Сатин, Я.А. Оценки погрешности аппроксимаций неоднородных марковских цепей с непрерывным временем // Теория вероятностей и ее применения. – 2016. – 61, вып., 563-569.
21. Зейфман, А., Коротышева, А., Сатин, Я., Киселева, К., Разумчик, Р., Королев, В., Шоргин, С. Оценки погрешности аппроксимации для марковских систем обслуживания, описываемых процессами рождения и гибели с дополнительными переходами. // Системы и средства информатики. – 2017. – 27.
22. Калашников, В.В. Качественный анализ сложных систем методом пробных функций / В.В. Калашников. – М.: Наука. – 1978.
23. Лозинский, С.М. Оценка погрешности численного интегрирования обыкновенных дифференциальных уравнений / С.М. Лозинский // Изв. ВУ-Зов. Матем. – 1958. – №5. – с.52-90.
24. Семенова, О.В., Дудин, А.Н. Система массового обслуживания $M|M|N$ с управляемым режимом обслуживания и катастрофическими сбоями // Автоматика и вычислительная техника. – 2007. – №6. С. 72-80.
25. Степанов, С.Н., Цитович, И.И. Некоторые аспекты исследования систем с повторными вызовами качественными методами // Сб. «Методы теории телетрафика в децентрализованных системах управления. М.: Наука. – 1986. – С.68-90.
26. Степанов, С.Н., Цитович, И.И. Качественные методы исследования систем с повторными вызовами // Проблемы передачи информации. Т.23. №2. – 1987. – С.92-112.
27. Ушаков, В.Г., Ушаков, Н.Г. О длине очереди в системе обслуживания с эрланговским входящим потоком // Вестник Московского университета. Т.40, №3. – 2016. – С. 118-122.
28. Чегодаев, А.В. Математические модели и методы оценки характеристик стохастических систем, близких к поглощающим: дисс. ... канд.ф.-м. наук / А.В. Чегодаев. – Вологда. – 2009. – 127 с.

29. Штойян, Д. Качественные свойства и оценки стохастических моделей / Д. Штойян. – М.: Мир. – 1979.
30. Artalejo, J.R. Stationary analysis of the characteristics of the $M/M/2$ queue with constant repeated attempts. *Opsearch*. – 1996. – 33, 83–95.
31. Artalejo, J.R., Gómez-Corral, A., Neuts, M.F. Analysis of multiserver queues with constant retrial rate. – 2001. – *European Journal of Operational Research*, 135, 569–581.
32. Avrachenkov, K., Yechiali, U. Retrial networks with finite buffers and their application to Internet data traffic. *Probability in the Engineering and Informational Sciences*. – 2008. – 22. – P. 519–536.
33. Avrachenkov, K., Yechiali, U. On tandem blocking queues with a common retrial queue. *Computers and Operations Research*. – 2010. – 37(7). – P.1174–1180.
34. Avrachenkov, K., Morozov, E.V. Stability analysis of $GI/G/c/K$ retrial queue with constant retrial rate. – 2014. – *Math. Meth. Oper. Res.*, 79, 273–291.
35. Avrachenkov, K., Nekrasova, E., Morozov, E., Steyaert, B. Stability analysis and simulation of N -class retrial system with constant retrial rates and Poisson inputs. – 2014. – *Asia-Pacific Journal of Operational Research*, 31, No.2.
36. Chen, A.Y., Renshaw, E. The $M/M/1$ queue with mass exodus and mass arrives when empty. – 1997. – *J. Appl. Prob.* **34**, pp. 192–207.
37. Chen, A.Y., Renshaw, E. Markov bulk-arriving queues with state-dependent control at idle time. *Adv. Appl. Prob.* 36. – 2004. – P. 49-524.
38. Chen, A.Y., Pollet, P., Li, J., Zhang, H. Markovian bulk-arrival and bulk-service queues with state-dependent control. *Queueing Systems*, 64. – 2010, P.267-304.
39. Ching, Wai-K., Ng, Michael, K. Markov chains: models, algorithms and applications / Wai-Ki Ching, K. Michael. – International Series in Operations Research & Management Science 83. – 2006. – New York, NY: Springer.

40. Choi, B.D., Shin, Y.W., Ahn, W.C. Retrial queues with collision arising from unslotted CSMA/CD protocol. – 1992. – *Queueing Systems*, 11, 335–356.
41. Choi, B.D., Park, K.K., Pearce, C.E.M. An $M/M/1$ retrial queue with control policy and general retrial times. – 1993. – *Queueing Systems*, 14, 275–292.
42. Choi, B.D., Rhee K.H., Park, K.K. The $M/G/1$ retrial queue with retrial rate control policy. – 1993. – *Probability in the Engineering and Informational Sciences*, 7, 29–46.
43. Erlang, A. K. Løsning af nogle Problemer fra Sandsynlighedsregningen af Betydning for de automatiske Telefoncentraler. *Elektroteknikeren*. – 1917. – 13, 5–13.
44. Fayolle, G. A simple telephone exchange with delayed feedback. In Boxma, O.J., Cohen J.W., and Tijms, H.C. (eds.), *Teletraffic Analysis and Computer Performance Evaluation*. – 1986. – 7, 245–253.
45. Foss, S.G., Kalashnikov, V.V. Regeneration and renovation in queues / S.G. Foss, V.V. Kalashnikov // *Queueing Systems*. - 1991. - 8(1). - p. 211-223.
46. Fricker, C., Robert, P., Tibi, D.: On the rate of convergence of Erlang's model. *J. Appl. Probab.* – 1999. – 36, 1167–1184.
47. Halfin, S., Whitt, W. Heavy-traffic limits for queues with many exponential servers / S. Halfin, W. Whitt // *Oper. Res.*. - 1981. - 29. - p. 567-588.
48. Islam, M.A. A Birth-Death Process Approach to Constructing Multistate Life Tables / M.A. Islam // *Bull. Malaysian Math. Sc. Soc. (Second Series)*. – 2003. - 26. - p. 101-108.
49. Kartashov, N.V. Strong stable Markov chains / Kartashov N. V. - Kiev: Utrecht, VSP, TBiMC. – 1996.
50. Kijima, M.: On the largest negative eigenvalue of the infinitesimal generator associated with $M/M/n/n$ queues. *Oper. Res. Let.* – 1990. – 9, 59–64.

51. Klimenok, V., Dudin, A. Multi-dimensional asymptotically quasi-Toeplitz Markov chains and their application in queueing theory / Valentina Klimenok, Alexander Dudin // Queueing Systems. - 2006. - 54. - p. 245-259.
52. Lillo, R.E. A $G/M/1$ queue with exponential retrial. – 1996. – TOP, 4, 99–120.
53. Massey, W. A., Whitt, W.: On analysis of the modified offered-load approximation for the nonstationary Erlang loss model. Ann. Appl. Probab. – 1994. – 4, 1145–1160.
54. Mandelbaum A., Massey W. Strong approximations for time-dependent queues// Math. Oper. Res. – 1995. – No.20, p. 33-64.
55. Margolius, B.H. The matrices R and G of matrix analytic methods and the time-inhomogeneous periodic Quasi-Birth-and-Death process / // Queueing Systems. - 2008. - 60(1-2). - p. 131-151.
56. Meyn, Sean, P., Robert, L. Tweedie. Computable bounds for geometric convergence rates of Markov chains / Meyn, Sean P., Robert L. // The Annals of Applied Probability. - 1994. - p. 981-1011.
57. Mitrophanov, A. Stability and exponential convergence of continuous-time Markov chains// J. Appl. Probab. – 40. – 2003. – P. 970–979.
58. Morozov, E. The stability of a non-homogeneous queueing system with regenerative input / Evsei Morozov // Journal of Mathematical Sciences 83.3. - 1997. - p. 407-421.
59. Morozov, E. A multiserver retrial queue: regenerative stability analysis / Evsey Morozov // Queueing Systems 56.3-4. - 2007. - p. 157-168.
60. Parthasarathy, P. R., Krishna Kumar, B. Density-dependent birth and death processes with state-dependent immigration. – 1991. – Mathematical and Computer Modelling **15**, p. 11–16.
61. Rykov, V. On Markov Reliability Model of a System, Operating in Markov Random Environment. XXXI ISSPSM Conference, April. – 2013. – PFUR. Moscow. (jointly with Tran Ahn Ngia).

62. Semenova, O., Dudin, A.N., Karolik, A.V., Maslakova, O.V. Investigation of a BMAP/SM/1 Retrial System with Markovian Arrival Input of Disasters and Non-instantaneous Recovery of the Server // “Computer Data Analysis and Modeling” (Proceedings of the 6th International Conference). – Minsk. – 2001. – V.1. P. 128 – 131.
63. Stepanov, S.N. Markov Models with Retrials: The Calculation of Stationary Performance Measures Based on the Concept of Truncation // Mathematical and Computer Modelling. – 1999. – 30. – P. 207-228.
64. Van Doorn, E. A., Zeifman, A. I.: On the speed of convergence to stationarity of the Erlang loss system. – 2009. – Queueing Syst. 63, 241–252.
65. E. A. Van Doorn, A. I. Zeifman, T. L. Panfilova. Bounds and asymptotics for the rate of convergence of birth-death processes // Th. Prob. Appl. – 2010. – **54**, 97–113.
66. Wong, E.W.M., Andrew, L.L.H., Cui, T., Moran, B., Zalesky, A., Tucker, R.S., Zukerman, M. Towards a bufferless optical internet. – 2009. – Journal of Lightwave Technology, 27, 2817–2833.
67. Yao, S., Xue, F., Mukherjee, B., Yoo, S.J.B., Dixit, S. Electrical ingress buffering and traffic aggregation for optical packet switching and their effect on TCP-level performance in optical mesh networks. – 2002. – IEEE Communications Magazine, 40(9), 66–72.
68. Zeifman, A. I. Stability for continuous-time nonhomogeneous Markov chains. Lect. Notes Math. – 1985. – 1155, 401–414.
69. Zeifman, A. I. Upper and lower bounds on the rate of convergence for nonhomogeneous birth and death processes. – 1995. – Stoch. Proc. Appl. 59, 157–173.
70. Ammar S. I. Transient behavior of a two-processor heterogeneous system with catastrophes, server failures and repairs // Applied Mathematical Modelling, 2014, 38(7-8), 2224-2234.

71. Ammar, S. I., Zeifman, A., Satin, Y., Kiseleva, K., Korolev, V. (2020). On limiting characteristics for a non-stationary two-processor heterogeneous system with catastrophes, server failures and repairs // Journal of Industrial & Management Optimization, 2021, 17(3), 1057.
72. Di Crescenzo, A., Giorno, V., Krishna Kumar, B. et al. A Double-ended Queue with Catastrophes and Repairs, and a Jump-diffusion Approximation // Methodol. Comput. Appl. Probab., 2012, 14, 937–954.
73. Di Crescenzo, A., Virginia Giorno. Balasubramanian Krishna Kumar, Amelia Nobile. 2018. A time-non-homogeneous double-ended queue with failures and repairs and its continuous approximation // Mathematics 2018, 6, 1–23.
74. Heidergott, B., Leahu, H., Lpker, A., Pflug, G. Perturbation analysis of inhomogeneous finite Markov chains // Advances in Applied Probability, 2016, 48(1), 255–273.
75. Knessl, C., Tier, C. Applications of singular perturbation methods in queueing // Advances in queueing theory, methods, and open problems, probability and stochastics series, 1995, 311–336.
76. Knessl, C., Yang, Y. P. An exact solution for an $M(t)/M(t)/1$ queue with time-dependent arrivals and service // Queueing systems, 2002, 40(3), 233–245.
77. Krishna Kumar, B., Krishnamoorthy, A., Pavai Madheswari, S. et al. Transient analysis of a single server queue with catastrophes, failures and repairs // Queueing Syst 2007, 56, 133–141.
78. Masuyama, H., Takine, T. Algorithmic computation of the time-dependent solution of structured Markov chains and its application to queues // Stochastic models, (2005), 21(4), 885–912.
79. Suranga Sampath, M.I.G., Liu, J. Transient Analysis of an $M/M/1$ Queue with Reneging, Catastrophes, Server Failures and Repairs // Bull. Iran. Math. Soc. 2018, 44, 585–603.

80. Tarabia, A. M. Transient and steady-state analysis of an M/M/1 queue with balking, catastrophes, server failures and repairs // J. Ind. Manage. Optim, 2011, 7, 811-823.
81. Zeifman, A., Satin, Ya., Korolev, V., Shorgin, S. On truncations for weakly ergodic inhomogeneous birth and death processes // Int. J. Appl. Math. Comp. Sci. 2014, 24, 503–518.
82. Zeifman, A., Y. Satin, A. Korotysheva, V. Korolev, S. Shorgin, R. Razumchik. Ergodicity and perturbation bounds for inhomogeneous birth and death processes with additional transitions from and to origin // Int. J. Appl. Math. Comput. Sci, 2015, 25(4), 503–518.
83. Zeifman A. I., Korolev V. Yu., Korotysheva A. V., Satin Ya. A. Bounds for inhomogeneous Markovian queueing systems with particularities in zero. Moscow, 2016. IPI FRC CSC RAS.
84. Zeifman, A. I., Korotysheva, A. V., Korolev, V. Y., Satin, Y. A. Truncation Bounds for Approximations of Inhomogeneous Continuous-Time Markov Chains // Theory of Probability & Its Applications, 2017, 61(3), 513–520.
85. Zeifman, A., Satin, Y., Kiseleva, K., Korolev, V., Panfilova, T. On limiting characteristics for a non-stationary two-processor heterogeneous system // Applied Mathematics and Computation, 2019, 351, 48–65.
86. Zeifman, A., Korolev, V., Satin, Y. Two approaches to the construction of perturbation bounds for continuous-time Markov chains // Mathematics, 2020, 8(2), 253.
87. Zeifman, A., Satin, Y., Kryukova, A., Razumchik, R., Kiseleva, K., Shilova, G. On the Three Methods for Bounding the Rate of Convergence for some Continuous-time Markov Chains // Int. J. Appl. Math. Comput. Sci, 2020, 30 (2), 251–266.
88. E. van Doorn. The transient state probabilities for a queueing model where potential customers are discouraged by queue length // J. Appl. Probab., 1981. Vol. 18. P. 499–506.

89. Natvig B. On the transient state probabilities for a queueing model where potential customers are discouraged by queue length // J. Appl. Probab., 1974. Vol. 11. P. 345–354.
90. Reynolds J.E., The stationary solution of a multiserver queueing model with discouragement // Oper. Res., 1968. Vol. 16. P. 64–71.
91. Srivastava H.M., Kashyap B.R.K., Special Functions in Queueing Theory and Related Stochastic Processes // Bull. Amer. Math. Soc. (N.S.), 1984. Vol. 10. No. 1. P. 139–141.
92. Granovsky B.L., Zeifman A., Nonstationary Queues: Estimation of the Rate of Convergence // Queueing Systems, 2004. Vol. 46. P. 363–388.
93. Parthasarathy P. R., Selvaraju N., Transient analysis of a queue where potential customers are discouraged by queue length // Mathematical Problems in Engineering, 2001. Vol. 7. P. 433–454.
94. Jain M., Singh M., Transint analysis of a Markov queueing model with feedback, discouragement and disaster // International Journal of Applied and Computational Mathematics, 2020. Vol. 6. No. 2. P. 1–14.
95. Zeifman A., On the Study of Forward Kolmogorov System and the Corresponding Problems for Inhomogeneous Continuous-Time Markov Chains // Springer Proceedings in Mathematics and Statistics, 2020. Vol. 333. P. 21–39.
96. Arns, M., Buchholz, P., Panchenko, A. On the numerical analysis of inhomogeneous continuous-time Markov chains // Informs Journal on Computing, 2010, 22, 416–432.
97. Andreychenko, A., Sandmann, W., Wolf, V. Approximate adaptive uniformization of continuous-time Markov chains // Applied Mathematical Modelling, 2018, 61, 561–576.
98. Katoen, J.-P., Zapreev, I.S. Safe on-the-fly steady-state detection for time-bounded reachability // In Proceedings of the 3rd International Conference on

the Quantitative Evaluation of Systems, California, CA, USA, 11–14 September 2016.

99. Balsamo, S., Harrison, P. G., Marin, A. (2010). A unifying approach to product-forms in networks with finite capacity constraints. *ACM SIGMETRICS Performance Evaluation Review*, 38(1), 25–36.
100. Burak, M. R., Korytkowski, P. (2020). Inhomogeneous CTMC Birth-and-Death Models Solved by Uniformization with Steady-State Detection. *ACM Transactions on Modeling and Computer Simulation (TOMACS)*, 30(3), 1–S18.
101. Disney, R. L. (1962). Some multichannel queueing problems with ordered entry. *Journal of Industrial Engineering*, 13(1), 46–48.
102. Elsayed, E. A. (1983). Multichannel queueing systems with ordered entry and finite source. *Computers & Operations Research*, 10(3), 213–222.
103. Grosz, I., Scully, Z., Harchol-Balter, M. (2018). SRPT for multiserver systems. *Performance Evaluation*, 127, 154–175.
104. Harrison, P. G. (2003). Turning back time in Markovian process algebra. *Theor. Comput. Sci.*, 290(3), 1947–1986.
105. Harrison, P. G., Marin, A. (2014). Product-forms in multi-way synchronizations. *The Computer Journal*, 57(11), 1693–1710.
106. Kelly, F. P. (2011). *Reversibility and stochastic networks*. Cambridge University Press.
107. Marin, A., Balsamo, S., Fourneau, J. M. (2017). LB-networks: a model for dynamic load balancing in queueing networks. *Performance Evaluation*, 115, 38–53.
108. Matsui, M. (2009). *Manufacturing and service enterprise with risks. A Stochastic Management Approach*. International Series in Operations Research & Management Science book series (ISOR, volume 125).

109. Pittel, B. (1979). Closed exponential networks of queues with saturation: the Jackson-type stationary distribution and its asymptotic analysis. *Mathematics of Operations Research*, 4(4), 357–378.
110. Razumchik, R., Zaryadov, I. (2016). Stationary blocking probability in multi-server finite queuing system with ordered entry and Poisson arrivals. *Communications in Computer and Information Science book series (CCIS, volume 601)*, 344–357. Springer, Cham.
111. Schrage, L. (1968). A proof of the optimality of the shortest remaining processing time discipline. *Operations Research*, 16(3), 687–690.
112. Schwarz, J. A., Selinka, G., Stolletz, R. (2016). Performance analysis of time-dependent queueing systems: Survey and classification. *Omega*, 63, 170–189.
113. Whitt, W., You, W. (2019). Time-varying robust queueing. *Operations Research*, 67(6), 1766–1782.
114. Zeifman, A., Leorato, S., Orsingher, E., Satin, Y., Shilova, G. (2006). Some universal limits for nonhomogeneous birth and death processes. *Queueing systems*, 52(2), 139–151.
115. Marin, A., Rossi, S. A Queueing Model that Works Only on the Biggest Jobs // *Lecture Notes in Computer Science book series (LNCS)*, 2020. Vol. 12039. P. 118–132.
116. Daleckij, Ju.L., Krein, M.G., 1974. Stability of solutions of differential equations in Banach space. *Amer. Math. Soc. Transl.* 43.
117. Klefsjö, B. The hnbue and hnwue classes of life distributions. *Nav. Res. Logist.* **1982**, 29, 331–344.
118. Conti, P.L.J. An asymptotic test for a geometric process against a lattice distribution with monotone hazard. // *Ital. Statist. Soc.* **1997**, 6, 213–231.
119. Zeifman, A., Korolev, V., Korotysheva, A., Satin, Ya., Kiseleva, K., Shorgin, S. Bounds for Markovian queues with possible catastrophes. – 2017. – Proceedings

31st European Conference on Modeling and Simulation, ECMS, Budapest, Hungary, p.628-634.

120. Chen, A., Wu, X., Zhang, J. Markovian bulk-arrival and bulk-service queues with general state-dependent control. // *Queueing Syst.* **2020**, 1–48, doi:10.1007/s11134-020-09660-0.
121. Chen, A., Li, J., Hou, Z., Ng, K.W. Decay properties and quasi-stationary distributions for stopped Markovian bulk-arrival and bulk-service queues. *Queueing Syst.* **2010**, *66*, 275–311.
122. Li, J., Chen, A. Decay property of stopped Markovian bulk-arriving queues. // *Adv. Appl. Probab.* **2008**, *40*, 95–121.
123. Zeifman, A. I., Satin, Y. A., Kovalev, I. A., Sherif I. Ammar (2021). Ergodicity and Perturbation Bounds for Mt/Mt/1 Queue with Balking, Catastrophes, Server Failures and Repairs. *Rairo – Operations Research*, 55, 2223-2240.
124. Zeifman, A. I., R. V. Razumchik, Y. A. Satin, and I. A. Kovalev. Ergodicity bounds for the Markovian queue with time-varying transition intensities, batch arrivals and one queue skipping policy. // *Applied Mathematics and Computation* 395:125846 (2021).
125. Zeifman, A., Satin, Y., Kovalev, I., Razumchik, R., & Korolev, V. Facilitating Numerical Solutions of Inhomogeneous Continuous Time Markov Chains Using Ergodicity Bounds Obtained with Logarithmic Norm Method. *Mathematics*, 2021, 9(1), 42.
126. А. И. Зейфман, Я. А. Сатин, И. А. Ковалёв (2021). Об одной нестационарной модели обслуживания с катастрофами и тяжелыми хвостами. *Информатика и ее применения*. 2021. Т. 15. № 2, 20-25
127. Kochetkova I., Satin Y., Kovalev I., Makeeva E., Chursin A., Zeifman A. (2022). Convergence Bounds for Limited Processor Sharing Queue with Impatience for Analyzing Non-Stationary File Transfer in Wireless Network, 10(1), *Mathematics* 2022, 30;

128. И. А. Усов, И. А. Ковалёв, А. И. Зейфман (2022) Оценка погрешности аппроксимации неоднородных марковских цепей с непрерывным временем и катастрофами. Системы и средства информатики. 2022. Т. 32 №1, 34-45
129. А. И. Зейфман, В. Ю. Королев, Р. В. Разумчик, Я. А. Сатин, И. А. Ковалев, “О предельных характеристиках для систем обслуживания с исчезающими возмущениями”, Докл. РАН. Матем., информ., проц. упр., 506 (2022), 83–88; Dokl. Math., 106:2 (2022), 375–379
130. И. А. Ковалёв, Я. А. Сатин, А. В. Синицина А. И. Зейфман (2022) Об одном подходе к оцениванию скорости сходимости нестационарных марковских моделей систем обслуживания 16(3), 75-82 DOI: 10.14357/19922264220310
131. A. Zeifman, Y. Satin, I. Kovalev, A. Kryukova, G. Shilova and I. Usov (2022) On the study of forward Kolmogorov system: the corresponding problems and bounds for inhomogeneous continuous-time Markov chains and models. IEEE Xplore: 2022 International Conference on Information, Control, and Communication Technologies (ICCT) 3-7 окт DOI: 10.1109/icct56057.2022.9976673
132. 10. И. А. Ковалёв, Я. А. Сатин, А. И. Зейфман (2022) Оценки скорости сходимости и устойчивости для одного класса нестационарных марковских моделей систем с нетерпеливыми клиентами. Системы и средства информатики. 32(4) 21-31
133. Satin Y.A., Kovalev I.A., Razumchik R.V., Zeifman A.I. Upper bound on the rate of convergence and truncation bound for non-homogeneous birth and death processes on Z // Applied Mathematics and Computation. 2022. Т. 423. С. 127009.
134. Ковалёв И.А. Об оценках устойчивости и их применении для некоторых моделей массового обслуживания Системы и средства информатики. 2023. Т. 33 №1, 90-104
135. I. Kovalev, Y. Satin, A. Zeifman About Service Intensity Bounds for a Queuing Model. – 2021. – 19th international conference of numerical analysis and applied

mathematics September 2021, Rhodes, Greece. AIP Conf. Proc. 2849, 100004 (2023)

136. A. Zeifman, R. Razumchik, Y. Satin, I. Kovalev, I. Usov, G. Shilova, A. Kryukova Ergodicity Bounds for One Nonstationary Markovian Model with Impatience – 2022. – 20th international conference of numerical analysis and applied mathematics 19-25 September 2022, Heraklion, Crete, Greece
137. Ковалёв И.А., Сатин Я.А., Зейфман А.И., Программа для решения задачи Коши для системы однородных дифференциальных уравнений 1-го порядка методом Адамса-Мултона 4-го порядка // Свидетельство №2020615415 (РФ; Программа). – 2020 год.