ИПМ им. М.В. Келдыша РАН • Электронная библиотека
Математические вопросы кибернетики • Выпуск 21

А. Д. Яшунский

Некоторые вопросы вероятности на конечных квазигруппах

Рекомендуемая форма библиографической ссылки: Яшунский А. Д. Некоторые вопросы вероятности на конечных квазигруппах // Математические вопросы кибернетики. Вып. 21. — М.: ФИЗМАТЛИТ, 2023. — С. 194–205. URL: https://library.keldysh.ru/mvk.asp?id=2023-194 DOI: 10.20948/mvk-2023-194

НЕКОТОРЫЕ ВОПРОСЫ ВЕРОЯТНОСТИ НА КОНЕЧНЫХ КВАЗИГРУППАХ

А.Д. ЯШУНСКИЙ

(МОСКВА / ЛИМАСОЛ / НЕТАНИЯ)

Введение

Квазигрупповые операции на конечных множествах, а также связанные с ними латинские квадраты (и обобщающие их латинские гиперкубы), — достаточно естественные математические объекты. Они обладают множеством интересных свойств и привлекают внимание исследователей по крайней мере со времен Эйлера.

Одно из свойств квазигрупповых операций, которое можно неформально назвать «выравниванием», заключается в том, что результат применения квазигрупповой операции к случайным величинам имеет «более равномерное» распределение, чем исходные случайные величины. Это свойство уже некоторое время рассматривается как одно из обоснований для использования квазигрупповых операций в криптографических схемах. При этом суть упомянутого выравнивающего действия с вероятностной точки зрения зачастую исследуется недостаточно глубоко или же не исследуется вовсе. Создается впечатление, что исследователи, занятые алгебраическими и комбинаторными аспектами квазигрупп, готовы принять их выравнивающие свойства «на веру» с минимальными обоснованиями. Так, например, в работе [1] главным аргументом в пользу выравнивающих свойств (и, как следствие, возможности криптографических приложений) выступают несколько примеров последовательностей символов, преобразованных с помощью квазигрупповых операций: читателю предлагается убедиться в том, что в результате преобразования частоты символов приблизились к равномерному распределению.

Современные применения квазигрупп в криптографии (см. обзоры [3, 11, 23]) в качестве вероятностных предпосылок подобных применений практически неизменно отсылают к серии работ С. Марковски и Д. Глигороски с соавторами [13–18]. В этих работах рассматриваются потоковые фильтры для строк — криптографические преобразования, основанные на квазигруппах, — и приводится анализ возникающей вероятностной модели. Природа рассматриваемых преобразований позволяет использовать для их анализа математический аппарат конечных цепей Маркова, что успешно эксплуатируется авторами. Однако в действительности выравнивающие свойства квазигрупп продолжают проявляться и в тех случаях, когда рассматриваемые

DOI: 10.20948/mvk-2023-194

объекты перестают непосредственно описываться конечными марковскими цепями.

В настоящей работе рассматриваются некоторые аспекты формализации выравнивающих свойств квазигрупп в достаточно общей постановке задачи.

§ 1. Квазигрупповые свертки

Напомним, что конечное множество Q с бинарной операцией $x \circ y$ называется *квазигруппой*, если для любых $a, b \in Q$ уравнения $a \circ y = b$ и $x \circ a = b$ однозначно разрешимы относительно y и x. Эти уравнения, таким образом, задают на множестве Q операции правого и левого деления, соответствующие операции $\circ$: $x = b/a$, $y = a\backslash b$. Таблица умножения квазигруппы представляет собой латинский квадрат, т. е. в каждой строке и каждом столбце встречаются все $|Q|$ элементов из множества Q.

В дальнейшем для удобства будем считать, что $Q = \{1, 2, \ldots, k\}$. Пусть X — случайная величина на множестве Q. Обозначая вероятность $\mathsf{P}\{X = i\}$ через p_i для каждого $i \in Q$, получим набор вероятностей $\boldsymbol{p} = (p_1, \ldots, p_k)$: *распределение вероятностей* случайной величины X. Всевозможные распределения случайных величин на Q образуют *стохастический симплекс*:

$$\boldsymbol{S}^{(k)} = \{\boldsymbol{p} = (p_1, \ldots, p_k) \colon p_1 \geqslant 0, \ldots, p_k \geqslant 0, \sum_{i \in Q} p_i = 1\}.$$

Его элементы — распределения, которые мы также будем называть *стохастическими векторами*. *Носителем* распределения $\boldsymbol{p}$ называется множество $\mu(\boldsymbol{p}) = \{i \in Q \colon p_i > 0\} \subseteq Q$. Введем специальные обозначения для вырожденных распределений $\boldsymbol{e}^{(1)} = (1, 0, \ldots, 0)$, $\boldsymbol{e}^{(2)} = (0, 1, 0, \ldots, 0)$, $\ldots$ $\ldots$, $\boldsymbol{e}^{(k)} = (0, \ldots, 0, 1)$ и для *равномерного распределения* $\boldsymbol{u} = (1/k, \ldots, 1/k)$. Отметим, что для каждого распределения, если рассматривать его как элемент $\mathbb{R}^k$, имеет место представление $\boldsymbol{p} = \sum_{i \in Q} p_i \boldsymbol{e}^{(i)}$, т. е. $\boldsymbol{p}$ представляется в виде выпуклой комбинации векторов $\boldsymbol{e}^{(1)}, \ldots, \boldsymbol{e}^{(k)}$. Хотя линейная комбинация распределений, вообще говоря, не является распределением, рассмотрение выпуклых (а в некоторых случаях даже аффинных) комбинаций не выводит за пределы стохастического симплекса.

Если X и Y — независимые случайные величины на множестве Q с распределениями $\boldsymbol{p}$ и $\boldsymbol{q}$ соответственно, то $Z = X \circ Y$ — также случайная величина на множестве Q. Компоненты ее распределения $\boldsymbol{r}$ могут быть вычислены следующим образом: $r_i = \sum_{j \in Q} p_j q_{j\backslash i}$, $i \in Q$. Эти соотношения определяют на множестве стохастических векторов операцию *квазигрупповой свертки* $*$, соответствующую квазигрупповой операции $\circ$. Для распределений $\boldsymbol{p}, \boldsymbol{q}, \boldsymbol{r}$ будем записывать $\boldsymbol{r} = \boldsymbol{p} * \boldsymbol{q}$.

Непосредственно из определения квазигрупповой свертки вытекает, что для любого стохастического вектора $\boldsymbol{p}$ выполнено $\boldsymbol{p} * \boldsymbol{u} = \boldsymbol{u} * \boldsymbol{p} = \boldsymbol{u}$. Кроме того, квазигрупповая свертка билинейная, если рассматривать стохастические векторы как элементы $\mathbb{R}^k$. Это свойство может быть использовано для преобразований сверток выпуклых комбинаций стохастических векторов, а именно

$$\left(\sum_s \alpha_s \boldsymbol{p}^{(s)}\right) * \left(\sum_t \beta_t \boldsymbol{q}^{(t)}\right) = \sum_s \sum_t \alpha_s \beta_t (\boldsymbol{p}^{(s)} * \boldsymbol{q}^{(t)}).$$

§ 2. Меры неравномерности

Для того, чтобы формализовать «выравнивание», осуществляемое квазигрупповыми преобразованиями случайных величин, требуется некоторый способ сравнивать «равномерность» различных распределений. Ответ на этот вопрос в терминах частичного предпорядка на множестве распределений дает теория мажоризации (см. [4]). Для рассматриваемого в ней отношения мажорирования распределений $\preceq$ равномерное распределение $\boldsymbol{u}$ — наименьший элемент. При этом для квазигрупповой свертки распределений $\boldsymbol{r}=\boldsymbol{p}*\boldsymbol{q}$ несложно показать выполнение отношений $\boldsymbol{r}\preceq\boldsymbol{p}$ и $\boldsymbol{r}\preceq\boldsymbol{q}$. Однако они не влекут непосредственно приближение распределения $\boldsymbol{r}$ к равномерному распределению по сравнению с $\boldsymbol{p}$ и $\boldsymbol{q}$, а говорят лишь о неудалении от него.

При рассмотрении вычислений, содержащих многократное применение квазигрупповых операций, для доказательства сходимости результата к равномерному распределению, а также оценки скорости этой сходимости предпочтительнее иметь некоторую числовую функцию $f(\boldsymbol{p})$ на множестве распределений, характеризующую их близость к равномерному распределению.

Выбор конкретного числового выражения для равномерности (или неравномерности) распределения $\boldsymbol{p}$ во многом определяется традициями рассматриваемых задач. Так, например, при исследовании экстракторов случайности (см. [22]) получение более равномерного распределения часто понимается как повышение мин-энтропии $H_\infty(\boldsymbol{p})=\min\limits_{p_i>0}(-\log p_i)$. Однако точно так же можно было бы использовать в качестве меры равномерности энтропию Шеннона $H(\boldsymbol{p})=\sum\limits_i -p_i\log p_i$ или энтропию Реньи $H_\alpha(\boldsymbol{p})=\dfrac{1}{1-\alpha}\log\sum\limits_i p_i^\alpha$, с некоторым значением $\alpha\geqslant 0$. Последняя обобщает энтропии $H_\infty(\boldsymbol{p})$ и $H(\boldsymbol{p})$: она сходится к мин-энтропии при $\alpha\to\infty$ и к шенноновской энтропии при $\alpha\to 1$.

Другой естественный способ оценки близости распределения — вычисление расстояния от распределения $\boldsymbol{p}$ до равномерного распределения $\boldsymbol{u}$ в некоторой метрике. Наиболее распространен выбор *метрики полной вариации*, также известной как *статистическое расстояние*, определяемой для распределений $\boldsymbol{p}$ и $\boldsymbol{q}$ как $\rho_{TV}(\boldsymbol{p},\boldsymbol{q})=\dfrac{1}{2}\sum\limits_i|p_i-q_i|$. Однако возможно использование и других метрик. Так, например, при исследовании случайных блужданий на конечных группах (см. [21]) рассматривается целое семейство метрик $\rho_s(\boldsymbol{p},\boldsymbol{q})=\left(\sum\limits_i|p_i-q_i|^s\right)^{1/s}$ для $s\in[1,\infty)$. Отметим, что метрика ρ_1 отличается лишь множителем от ранее упомянутой метрики полной вариации: $\rho_1(\boldsymbol{p},\boldsymbol{q})=2\rho_{TV}(\boldsymbol{p},\boldsymbol{q})$.

Приведенные выше способы численного выражения (не)равномерности распределения оказываются связанными друг с другом неравенствами. В частности, величина неравномерности $d_{TV}(\boldsymbol{p})=\rho_{TV}(\boldsymbol{p},\boldsymbol{u})$, определяемой через метрику полной вариации, может быть оценена через энтропию распределения $\boldsymbol{p}$, и наоборот — энтропия может быть оценена через величину $d_{TV}(\boldsymbol{p})$. Подробнее см., например, [20].

В действительности для целей доказательства приближения распределений результатов квазигрупповых преобразований к равномерному распре-

делению можно выбирать меру неравномерности $f(\boldsymbol{p})$ из достаточно широкого класса функций.

Естественным требованием к функции f будет сохранение отношения мажорирования, т.е. выпуклость по Шуру*) (см. [4]). В случае строгой выпуклости по Шуру функция f будет иметь в точке $\boldsymbol{u}$ глобальный минимум, который можно, прибавляя подходящую константу к f, сделать равным нулю.

Наконец, непрерывность функции f относительно топологии, заданной стандартной евклидовой (или любой эквивалентной) метрикой, позволят трактовать сходимость значений f к нулю как критерий сходимости последовательности распределений к равномерному.

Теорема 1. *Пусть $f(\boldsymbol{p})$ — непрерывная, строго выпуклая по Шуру функция на $\boldsymbol{S}^{(k)}$, удовлетворяющая равенству $f(\boldsymbol{u})=0$, а $\boldsymbol{p}^{(n)}$, $n\in\mathbb{N}$, — некоторая последовательность распределений. Тогда сходимость $f(\boldsymbol{p}^{(n)})\to 0$ равносильна сходимости $\boldsymbol{p}^{(n)}\to\boldsymbol{u}$.*

Доказательство. С одной стороны, если $\boldsymbol{p}^{(n)}\to\boldsymbol{u}$, то из непрерывности f получаем $f(\boldsymbol{p}^{(n)})\to f(\boldsymbol{u})=0$.

С другой стороны, пусть $f(\boldsymbol{p}^{(n)})\to 0$. Поскольку последовательность $\boldsymbol{p}^{(n)}$ лежит в $\boldsymbol{S}^{(k)}$, она ограничена и, следовательно, имеет по крайней мере одну предельную точку**). Пусть далее $\boldsymbol{q}$ — какая-то предельная точка последовательности $\boldsymbol{p}^{(n)}$.

Из непрерывности f следует, что для подпоследовательности $\boldsymbol{p}^{(n_i)}$, $i\in\mathbb{N}$, сходящейся к $\boldsymbol{q}$, выполнено $f(\boldsymbol{p}^{(n_i)})\to f(\boldsymbol{q})$. При этом $f(\boldsymbol{p}^{(n_i)})$ — подпоследовательность последовательности $f(\boldsymbol{p}^{(n)})$, а значит,

$$f(\boldsymbol{q})=\lim_{i\to\infty} f(\boldsymbol{p}^{(n_i)})=\lim_{n\to\infty} f(\boldsymbol{p}^{(n)})=0.$$

Из строгой выпуклости по Шуру функции f вытекает, что значение 0 функция f принимает только на распределении $\boldsymbol{u}$. Таким образом, необходимо $\boldsymbol{q}=\boldsymbol{u}$, т.е. $\boldsymbol{u}$ — единственная предельная точка последовательности $\boldsymbol{p}^{(n)}$, а значит $\boldsymbol{p}^{(n)}\to\boldsymbol{u}$. Теорема доказана.

Функцию $f(\boldsymbol{p})$, удовлетворяющую условиям теоремы 1, мы будем называть *мерой неравномерности* или просто *неравномерностью* распределения $\boldsymbol{p}$. Все вводимые нами в данной статье неравномерности оказываются эквивалентными***), поэтому выбор конкретной неравномерности не влияет ни на факт сходимости, ни на оценку порядка ее скорости.

§ 3. Преобразования квазигрупповыми формулами

Частным случаем квазигрупповых преобразований случайных величин можно считать случайные блуждания на конечных группах, поскольку рассматриваемые в них групповые операции — частный случай квазигрупповых, дополнительно обладающих ассоциативностью.

---

*) Отметим, что энтропийные функции выгнуты по Шуру, поскольку фактически измеряют равномерность, но переход к противоположным функциям приводит к выпуклым по Шуру функциям, измеряющим неравномерность.

**) Точка называется *предельной* для последовательности, если в любой ее окрестности имеются элементы указанной последовательности.

***) Неравномерности $d'(\boldsymbol{p})$ и $d''(\boldsymbol{p})$ *эквивалентны*, если для некоторых констант $C_1, C_2>0$ и любых $\boldsymbol{p}$ выполнено $C_1 d'(\boldsymbol{p})\leqslant d''(\boldsymbol{p})\leqslant C_2 d'(\boldsymbol{p})$.

Для групповой операции $\circ$ задача о предельном распределении произведений $Y_n = X_1 \circ X_2 \circ \ldots \circ X_n$ растущего числа независимых одинаково распределенных случайных величин достаточно подробно исследована. Одна из первых работ в этой области — статья Н. Н. Воробьева [2], обобщающая центральную предельную теорему на сложение случайных величин в конечных абелевых группах.

Для квазигрупповой операции $\circ$ невозможно рассматривать аналогичные произведения непосредственно. В силу неассоциативности операции $\circ$ различные расстановки скобок в выражении приводят к различным величинам. Для ассоциативной операции расстановка скобок, естественно, не влияет на результат и может быть выбрана произвольно.

Один из вариантов постановки задачи в случае произвольной квазигрупповой операции — рассмотрение выражений с некоторой фиксированной расстановкой скобок. Именно таким образом устроены *потоковые фильтры* из работ [14,16–18]. В них рассматриваются квазигрупповые произведения вида:

$$Y_n = (((X_1 \circ X_2) \circ X_3) \circ \cdots \circ X_{n-1}) \circ X_n. \tag{1}$$

Легко видеть, что $Y_n = Y_{n-1} \circ X_n$, и величины Y_n можно рассматривать как состояния некоторой цепи Маркова. Если величины X_i для всех $i \in \mathbb{N}$ имеют распределение $\boldsymbol{p}$, а распределения величин Y_n обозначить через $\boldsymbol{q}^{(n)}$, то выполнено равенство

$$\begin{pmatrix} q_1^{(n)} \\ q_2^{(n)} \\ \vdots \\ q_k^{(n)} \end{pmatrix} = \begin{pmatrix} p_{1\backslash 1} & p_{2\backslash 1} & \ldots & p_{k\backslash 1} \\ p_{1\backslash 2} & p_{2\backslash 2} & \ldots & p_{k\backslash 2} \\ \vdots & \vdots & \ddots & \vdots \\ p_{1\backslash k} & p_{2\backslash k} & \ldots & p_{k\backslash k} \end{pmatrix} \begin{pmatrix} q_1^{(n-1)} \\ q_2^{(n-1)} \\ \vdots \\ q_k^{(n-1)} \end{pmatrix}.$$

Матрица переходов $P = (p_{j\backslash i})_{i,j}$ в силу квазигрупповых свойств операции $\circ$ (и ее левой обратной операции $\backslash$) — дважды стохастическая. Отсюда следует равномерность предельного распределения при условии, что соответствующая цепь Маркова эргодична. Для групповой операции $\circ$ это равносильно тому, что носитель $\mu(\boldsymbol{p})$ распределения $\boldsymbol{p}$ не лежит ни в каком смежном классе рассматриваемой группы.

Для групповой операции $\circ$ приведенное выше построение практически закрывает вопрос о предельном распределении величин Y_n, а также говорит, что сходимость распределений к предельному экспоненциальная: расстояние от $\boldsymbol{q}^{(n)}$ до $\boldsymbol{u}$ убывает как $C\lambda^n$, где $C > 0$ и $0 \leqslant \lambda < 1$ — некоторые постоянные. Основным предметом исследований случайных блужданий на конечных группах в настоящее время выступает феномен, называемый cut-off («обрыв»). Оказывается, что в некоторых группах приближение к предельному распределению происходит скачкообразно после относительно небольшого числа шагов цепи Маркова — быстрее, чем это следовало бы из экспоненциальной оценки сходимости (см. [21]).

В отличие от группового случая, где расстановка скобок без ограничения общности может считаться такой, как в (1), для квазигрупповой операции $\circ$ различные расстановки скобок в произведении $X_1 \circ \ldots \circ X_n$ могут давать различные результаты. Однако при достаточно общих условиях

можно показать, что, независимо от расстановки скобок, квазигрупповые произведения независимых случайных величин имеют распределения, приближающиеся к равномерному с ростом числа сомножителей. Это доказано в работе автора [6, теорема 3] с использованием меры неравномерности $d_\delta(\boldsymbol{p}) = \max\limits_i p_i - \min\limits_i p_i$.

Теорема 2 [6]. *Пусть $X_1, \ldots, X_n$ — независимые одинаково распределенные случайные величины на Q, $|Q| = k$, с распределением $\boldsymbol{p}$, удовлетворяющим $|\mu(\boldsymbol{p})| > k/2$. Тогда найдется такое число $\beta > 0$, что для любой правильной расстановки скобок в произведении $X_1 \circ \ldots \circ X_n$ распределение результата $\boldsymbol{r}$ удовлетворяет неравенству $d_\delta(\boldsymbol{r}) \leqslant n^{-\beta}$.*

Условие $|\mu(\boldsymbol{p})| > k/2$, фигурирующее в теореме, достаточно естественно: если носитель распределения $\boldsymbol{p}$ содержит не более $k/2$ элементов, он может целиком лежать в подквазигруппе рассматриваемой квазигруппы. В этом случае все квазигрупповые преобразования случайных величин с распределением $\boldsymbol{p}$ также будут принимать значения только в этой подквазигруппе и их распределения не будут приближаться к равномерному.

Однако заменить условие $|\mu(\boldsymbol{p})| > k/2$ на порождение всего множества Q носителем $\mu(\boldsymbol{p})$ относительно операции $\circ$, как это обычно делается для групповой операции $\circ$ (см. [21]), не представляется возможным. Это демонстрируется следующим примером из работы [6].

Рассмотрим квазигрупповые преобразования случайных величин на множестве $Q = \{1, 2, 3, 4, 5, 6\}$ с операцией $\circ$, заданной таблицей.

Таблица

Таблица умножения квазигруппы

$\circ$	1	2	3	4	5	6
1	1	3	2	4	6	5
2	3	2	5	1	4	6
3	2	4	6	5	1	3
4	4	1	3	6	5	2
5	5	6	1	2	3	4
6	6	5	4	3	2	1

Пусть распределение $\boldsymbol{p}$ таково, что $\mu(\boldsymbol{p}) = \{1, 2\}$. Тогда носителем распределения $(\boldsymbol{p} * \boldsymbol{p}) * (\boldsymbol{p} * \boldsymbol{p})$ будет все множество Q. Как следствие, любая формула со случайными величинами, содержащая в качестве подформулы $((X_j \circ X_{j+1}) \circ (X_{j+1} \circ X_{j+3}))$, будет также иметь распределение с носителем Q. Вместе с тем при всех n распределения $(((\boldsymbol{p} * \boldsymbol{p}) * \boldsymbol{p}) * \cdots * \boldsymbol{p}) * \boldsymbol{p}$ случайных величин $(((X_1 \circ X_2) \circ X_3) \circ \cdots \circ X_{n-1}) \circ X_n$ будут иметь носитель, лежащий в множестве $\{1, 2, 3, 4\}$.

Таким образом, хотя носитель $\{1, 2\}$ и порождает всю квазигруппу, среди выражений, содержащих сколь угодно большое число случайных величин, встречаются как те, распределение которых имеет носитель, совпадающий с Q, так и те, у которых носитель распределения лежит в $\{1, 2, 3, 4\}$. Очевидно, что все эти распределения не могут стремиться к равномерному с ростом числа случайных величин в выражении.

Утверждение теоремы 2 может быть распространено на случай, когда в выражении со случайными величинами используется не одна,

а произвольный набор квазигрупповых операций на заданном множестве Q. Частный случай обобщения (с фиксированной расстановкой скобок (см. (1)) и дополнительными ограничениями на порядок использования различных квазигрупповых операций) доказывается в работе [14], однако во всей общности оно приведено в работе [6].

Теорема 3 [6]. *Пусть $X_1, \ldots, X_n$ — независимые одинаково распределенные случайные величины на Q, $|Q| = k$, с распределением $\boldsymbol{p}$, удовлетворяющим $|\mu(\boldsymbol{p})| > k/2$. Тогда найдется такое число $\beta > 0$, что для любой правильной расстановки скобок в произведении $X_1 \circ_1 \ldots \circ_{n-1} X_n$, где $\circ_1, \ldots, \circ_{n-1}$ — какие-то квазигрупповые операции на Q, распределение результата $\boldsymbol{r}$ удовлетворяет неравенству $d_\delta(\boldsymbol{r}) \leqslant n^{-\beta}$.*

Доказанная в теоремах 2 и 3 оценка $d_\delta(\boldsymbol{r}) \leqslant n^{-\beta}$ для неравномерности распределения $\boldsymbol{r}$ является «полиномиальной» по n, что существенно слабее, чем экспоненциальная оценка $C\lambda^n$, имеющая место для групповых преобразований случайных величин. Доказательство экспоненциальных оценок для квазигрупп оказывается осуществимым при использовании вместо неравномерности $d_\delta(\boldsymbol{p})$ других мер неравномерности.

§ 4. Скорость сходимости

В работе У. Маурера, К. Пьетрзака и Р. Реннера [19], посвященной криптографическим преобразованиям, доказано следующее неравенство для неравномерности d_{TV} квазигрупповой свертки распределений $\boldsymbol{p}$ и $\boldsymbol{q}$:

$$d_{TV}(\boldsymbol{p} * \boldsymbol{q}) \leqslant 2 d_{TV}(\boldsymbol{p}) d_{TV}(\boldsymbol{q}).$$

В общем случае это неравенство неулучшаемо, так как для распределений на двухэлементной квазигруппе (для $|Q| = 2$) оно обращается в равенство.

Домножив неравенство на 2, его можно переписать в более симметричной форме с помощью неравномерности $d_1(\boldsymbol{p}) = \rho_1(\boldsymbol{p}, \boldsymbol{u}) = \sum\limits_i |p_i - 1/k|$:

$$d_1(\boldsymbol{p} * \boldsymbol{q}) \leqslant d_1(\boldsymbol{p}) d_1(\boldsymbol{q}). \tag{2}$$

Выполнение подобного неравенства для какой-либо меры неравеномерности влечет единую оценку неравномерности распределений квазигрупповых преобразований независимых одинаково распределенных случайных величин.

Теорема 4 [8]. *Пусть мера неравномерности $d(\boldsymbol{p})$ такова, что для любой квазигрупповой свертки $*$ выполнено $d(\boldsymbol{p} * \boldsymbol{q}) \leqslant d(\boldsymbol{p}) d(\boldsymbol{q})$.*

Пусть $X_1, \ldots, X_n$ — независимые одинаково распределенные случайные величины на Q, $|Q| = k$, с распределением $\boldsymbol{p}$. Тогда для любой правильной расстановки скобок в произведении $X_1 \circ_1 X_2 \circ_2 \ldots \circ_{n-1} X_n$, где $\circ_1, \ldots, \circ_{n-1}$ — какие-то квазигрупповые операции на Q, распределение результата $\boldsymbol{r}$ удовлетворяет неравенству $d(\boldsymbol{r}) \leqslant (d(\boldsymbol{p}))^n$.

Доказательство легко осуществляется индукцией по структуре выражения, получающегося расстановкой скобок в $X_1 \circ_1 X_2 \circ_2 \ldots \circ_{n-1} X_n$.

В силу теоремы 4 неравенство (2) может быть использовано для получения экспоненциальной скорости сходимости квазигрупповых сверток

к равномерному распределению. Для этого потребуется выполнение условия $d_1(\boldsymbol{p})<1$. Однако это условие не тождественно фигурировавшему в теореме 2 условию $|\mu(\boldsymbol{p})|>k/2$, а слабее, как будет показано ниже. То есть неравенство (2) позволяет доказать экспоненциальную скорость сходимости *не для всех* распределений $\boldsymbol{p}$, для которых имеет место сходимость к равномерному распределению.

Покажем сначала, что из неравенства $d_1(\boldsymbol{p})<1$ вытекает неравенство $|\mu(\boldsymbol{p})|>k/2$. Действительно, пусть $|\mu(\boldsymbol{p})|=m$. Не ограничивая общности можем считать, что $p_1+\ldots+p_m=1$ и выполнены неравенства:

$$p_1\geqslant p_2\geqslant\ldots\geqslant p_s\geqslant\frac{1}{k}>p_{s+1}\geqslant\ldots\geqslant p_m>0.$$

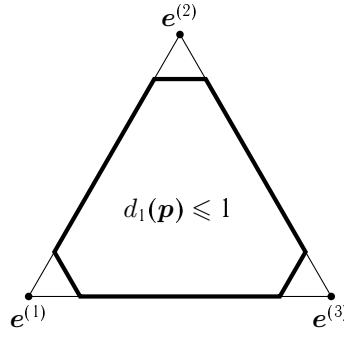
Тогда из условия $d_1(\boldsymbol{p})<1$ выводим следующую цепочку соотношений:

$$1>d_1(\boldsymbol{p})=\sum_{i=1}^{k}\left|p_i-\frac{1}{k}\right|=\sum_{i=1}^{s}\left(p_i-\frac{1}{k}\right)+\sum_{i=s+1}^{m}\left(\frac{1}{k}-p_i\right)+\sum_{i=m+1}^{k}\frac{1}{k}\geqslant$$
$$\geqslant\sum_{i=1}^{s}\left(p_i-\frac{1}{k}\right)-\sum_{i=s+1}^{m}\left(\frac{1}{k}-p_i\right)+\frac{k-m}{k}=\sum_{i=1}^{m}\left(p_i-\frac{1}{k}\right)+\frac{k-m}{k}=$$
$$=\sum_{i=1}^{m}p_i-\frac{m}{k}+1-\frac{m}{k}=2\left(1-\frac{m}{k}\right).$$

Итак, $\frac{1}{2}>1-\frac{m}{k}$, откуда очевидно $m>k/2$. Следовательно, неравенство (2) и теорема 4 не расширяют условия сходимости относительно теоремы 3, а лишь улучшают оценку скорости сходимости для некоторых распределений $\boldsymbol{p}$.

Убедиться в том, что условие $d_1(\boldsymbol{p})<1$ не выполняется даже для некоторых распределений с носителем $\mu(\boldsymbol{p})=Q$, можно, рассмотрев область $d_1(\boldsymbol{p})\leqslant 1$ в симплексе $\boldsymbol{S}^{(3)}$ (см. рисунок).

Для произвольного $k>3$ это также верно: как несложно заметить, $d_1(\boldsymbol{e}^{(i)})=2(k-1)/k>1$ для $k>2$. В силу непрерывности d_1 как функции от $\boldsymbol{p}$ ее значение будет превышать 1 в некоторой окрестности каждого из вырожденных распределений $\boldsymbol{e}^{(i)}$, в том числе на распределениях с носителем, равным Q.



Множество $\{\boldsymbol{p}\colon d_1(\boldsymbol{p})\leqslant 1\}\subset\boldsymbol{S}^{(3)}$

Таким образом, условие $d_1(\boldsymbol{p})<1$, позволяющее вывести из неравенства (2) экспоненциальную сходимость квазигрупповых сверток к равномерному распределению, сужает множество распределений по сравнению с условием $|\mu(\boldsymbol{p})|>k/2$, которое фигурирует в теореме 3.

Экспоненциальная сходимость для всех распределений, удовлетворяющих условию $|\mu(\boldsymbol{p})|>k/2$, вытекает из теоремы, доказанной в работе [8]. В ней используется неравномерность $d_\Delta(\boldsymbol{p})$, определяемая следующим образом. Следуя [4], для распределения $\boldsymbol{p}=(p_1,\ldots,p_k)$ определим набор $(p_{[1]},p_{[2]},\ldots,p_{[k]})$ как перестановку компонент распределения $\boldsymbol{p}$, удовлетворяющую условиям $p_{[1]}\geqslant p_{[2]}\geqslant\ldots\geqslant p_{[k]}$. Положим

$$d_\Delta(\boldsymbol{p})=p_{[1]}+\ldots+p_{[\lfloor k/2\rfloor]}-p_{[\lceil k/2\rceil+1]}-p_{[\lceil k/2\rceil+2]}-\ldots-p_{[k]}.$$

В работе [8] доказано неравенство

$$d_\Delta(\boldsymbol{p}*\boldsymbol{q}) \leqslant d_\Delta(\boldsymbol{p})d_\Delta(\boldsymbol{q}). \tag{3}$$

В силу теоремы 4 оно влечет экспоненциальную сходимость сверток к равномерному распределению для всех распределений $\boldsymbol{p}$, удовлетворяющих условию $d_\Delta(\boldsymbol{p})<1$. Это условие, как несложно проверить, равносильно условию $|\mu(\boldsymbol{p})|>k/2$.

Доказательство неравенства (3) чрезвычайно трудоемко. Хотя оно, по-видимому, дает самые общие на данный момент условия экспоненциальной сходимости для квазигрупповых сверток, некоторый интерес представляют и более слабые, но менее трудоемкие оценки. Покажем, что условиям теоремы 4 удовлетворяет мера неравномерности $d_0(\boldsymbol{p})$, определяемая как $d_0(\boldsymbol{p}) = 1 - k\min\limits_i p_i$.

Т е о р е м а 5. *Для любых распределений $\boldsymbol{p}$ и $\boldsymbol{q}$ на k-элементном множестве и любой квазигрупповой свертки $*$ выполнено* $d_0(\boldsymbol{p}*\boldsymbol{q}) \leqslant d_0(\boldsymbol{p})d_0(\boldsymbol{q})$.

Д о к а з а т е л ь с т в о. Представим распределения $\boldsymbol{p}$ и $\boldsymbol{q}$ в следующем виде:

$$\boldsymbol{p} = k\min_i p_i\boldsymbol{u} + \sum_j (p_j - \min_i p_i)\boldsymbol{e}^{(j)},$$
$$\boldsymbol{q} = k\min_i q_i\boldsymbol{u} + \sum_j (q_j - \min_i q_i)\boldsymbol{e}^{(j)}.$$

Положим $a = k\min\limits_i p_i$ и $b = k\min\limits_i q_i$. Тогда для $\boldsymbol{p}*\boldsymbol{q}$ имеем следующее представление:

$$\begin{aligned}\boldsymbol{p}*\boldsymbol{q} &= \left(a\boldsymbol{u} + \sum_j (p_j - \min_i p_i)\boldsymbol{e}^{(j)}\right) * \left(b\boldsymbol{u} + \sum_j (q_j - \min_i q_i)\boldsymbol{e}^{(j)}\right) = \\ &= ab(\boldsymbol{u}*\boldsymbol{u}) + a\sum_j (p_j - \min_i p_i)(\boldsymbol{u}*\boldsymbol{e}^{(j)}) + b\sum_j (q_j - \min_i q_i)(\boldsymbol{u}*\boldsymbol{e}^{(j)}) + \\ &\quad + \left(\sum_j (p_j - \min_i p_i)\boldsymbol{e}^{(j)}\right) * \left(\sum_j (q_j - \min_i q_i)\boldsymbol{e}^{(j)}\right).\end{aligned}$$

Используя равенства $\boldsymbol{u}*\boldsymbol{u} = \boldsymbol{u}*\boldsymbol{e}^{(j)} = \boldsymbol{u}$, выполненные для любого j, и принимая во внимание неотрицательность всех компонент в свертке $\left(\sum\limits_j (p_j - \min\limits_i p_i)\boldsymbol{e}^{(j)}\right) * \left(\sum\limits_j (q_j - \min\limits_i q_i)\boldsymbol{e}^{(j)}\right)$, получаем неравенство:

$$\begin{aligned}(\boldsymbol{p}*\boldsymbol{q})_i &\geqslant \frac{1}{k}\left(ab + b\sum_j (p_j - \min_i p_i) + a\sum_j (q_j - \min_i q_i)\right) = \\ &= \frac{1}{k}(ab + b(1 - k\min_i p_i) + a(1 - k\min_i q_i)) = \frac{1}{k}(ab + b(1-a) + a(1-b)) = \\ &= \frac{1}{k}(a + b - ab) = \frac{1}{k}(1 - (1-a)(1-b)).\end{aligned}$$

Отсюда вытекает, что $k\min\limits_i(\boldsymbol{p}*\boldsymbol{q}) \geqslant 1 - (1 - k\min\limits_i p_i)(1 - k\min\limits_i q_i)$ или эквивалентно:

$$1 - k\min_i(\boldsymbol{p}*\boldsymbol{q}) \leqslant (1 - k\min_i p_i)(1 - k\min_i q_i).$$

Это и есть неравенство $d_0(\boldsymbol{p}*\boldsymbol{q}) \leqslant d_0(\boldsymbol{p})d_0(\boldsymbol{q})$. Теорема доказана.

Из теорем 5 и 4 вытекает, что для распределений $\boldsymbol{p}$ с неравномерностью $d_0(\boldsymbol{p}) = 1 - k \min\limits_i p_i < 1$ будет иметь место экспоненциальная сходимость квазигрупповых сверток к равномерному распределению. Условие $d_0(\boldsymbol{p}) < 1$, как несложно видеть, равносильно $\mu(\boldsymbol{p}) = Q$. Отметим, что это выполняется для распределений, которые *не* удовлетворяют неравенству $d_1(\boldsymbol{p}) < 1$. Таким образом, теорема 5 влечет экспоненциальную сходимость в некоторых случаях, для которых недостаточно неравенства (2).

§ 5. Оценки для конечных групп

Оценки скорости сходимости, полученные выше для квазигрупповых сверток, могут быть использованы и в частном случае, когда $\circ$ — групповая операция. В этой ситуации экспоненциальная оценка $C\lambda^n$ (для любой неравномерности) вытекает из общей теории цепей Маркова. Значение λ может быть оценено сверху через собственные значения матрицы переходов $P = (p_{j^{-1} \circ i})_{i,j}$ случайного блуждания на группе $\langle Q, \circ \rangle$, порожденного распределением $\boldsymbol{p}$.

Матрица P имеет k собственных значений, среди которых обязательно присутствует 1 — наибольшее по модулю собственное значение. Остальные собственные значения матрицы P, $\beta_1, \ldots, \beta_{k-1}$ будем считать расставленными в порядке убывания модулей: $1 = \beta_0 \geqslant |\beta_1| \geqslant |\beta_2| \geqslant \ldots \geqslant |\beta_{k-1}| \geqslant 0$.

Для эргодической цепи Маркова выполнено $|\beta_1| < 1$ и основание экспоненты λ в оценке скорости сходимости $C\lambda^n$ может быть оценено сверху значением $|\beta_1|$. Согласно теореме 4 это же значение λ оценивается сверху значениями неравномерностей $d_0(\boldsymbol{p})$, $d_1(\boldsymbol{p})$, $d_\Delta(\boldsymbol{p})$.

Оценки величины λ через неравномерности можно соотнести с утверждениями из теории случайных блужданий на группах, где собственное значение β_1 оценивается функцией от распределения $\boldsymbol{p}$. В частности, имеет место следующая теорема, приведенная в [21, теорема 6.2, с. 297]. В качестве ее первоисточника указываются работы Д. Алдоуса [10] и совместная работа П. Диакониса и Л. Салофф-Коста [12].

Т е о р е м а 6 [21]. *Пусть группа $\langle Q, \circ \rangle$ и распределение $\boldsymbol{p}$ на Q таковы, что $p_i = p_{i^{-1}}$ для всех $i \in Q$ и $\mu(\boldsymbol{p})$ порождает относительно операции $\circ$ все множество Q. Положим*

$$D = \max_{x \in Q} \min\{l\colon x = s_1 \circ \ldots \circ s_l, \ \text{где } s_1, \ldots, s_l \in \mu(\boldsymbol{p})\}.$$

Тогда для собственного значения β_1 матрицы переходов $P = (p_{j^{-1} \circ i})_{i,j}$ случайного блуждания на группе $\langle Q, \circ \rangle$, порождаемого распределением $\boldsymbol{p}$, выполнено неравенство $\beta_1 \leqslant 1 - \dfrac{1}{D^2} \min\limits_{i \in \mu(\boldsymbol{p})} p_i$.

Из теоремы 6 вытекает оценка $\lambda \leqslant 1 - \dfrac{1}{D^2} \min\limits_{i \in \mu(\boldsymbol{p})} p_i$. Полученные ранее оценки для λ будут нетривиальны (отличны от $\lambda \leqslant 1$) лишь для распределений $\boldsymbol{p}$ с достаточно большим носителем. Таким образом, область применимости теоремы 6 шире, чем у полученных нами оценок через неравномерности. Однако для распределений $\boldsymbol{p}$, носители которых обеспечивают нетривиальные оценки, можно сопоставить наши результаты с оценкой из теоремы 6.

Для нетривиальности оценки через $d_0(\boldsymbol{p})$ требуется $|\mu(\boldsymbol{p})| = k$. В этом случае имеем

$$\lambda \leqslant d_0(\boldsymbol{p}) = 1 - k \min_i p_i < 1 - \min_i p_i < 1 - \frac{1}{D^2} \min_i p_i.$$

Для нетривиальности оценки через $d_\triangle(\boldsymbol{p})$ достаточно выполнения неравенства $|\mu(\boldsymbol{p})| > k/2$. При этом получаем

$$\lambda \leqslant d_\triangle(\boldsymbol{p}) = \sum_{i=1}^{\lfloor k/2 \rfloor} p_{[i]} - \sum_{\lceil k/2 \rceil + 1}^{k} p_{[i]} \leqslant 1 - \sum_{\lfloor k/2 \rfloor + 1}^{k} p_{[i]} \leqslant 1 - \frac{1}{D^2} \min_{i \in \mu(\boldsymbol{p})} p_i.$$

Итак, при условии нетривиальности оценки $\lambda \leqslant d_0(\boldsymbol{p})$ и $\lambda \leqslant d_\triangle(\boldsymbol{p})$ оказываются сильнее оценки из теоремы 6.

§ 6. Возможные обобщения

Естественное обобщение бинарных квазигрупповых операций — n-арные. Функция $f(x_1, \ldots, x_n)\colon Q^n \to Q$ называется *n-арной квазигрупповой операцией*, если для всех $i = 1, \ldots, n$ и любых $c_1, \ldots, c_{n-1} \in Q$ функция $g(x) = f(c_1, \ldots, c_{i-1}, x, c_i, \ldots, c_{n-1})$ осуществляет перестановку на множестве Q. Для преобразований независимых случайных величин, составленных из квазигрупповых операций произвольной арности, сходимость распределений к равномерному доказана в работе автора [9]. Для этого доказательства так же, как и в [6], используется неравномерность $d_\delta(\boldsymbol{p})$. Вместе с тем, по-видимому, требуются лишь технические изменения для того, чтобы перенести доказательства для неравномерностей $d_0(\boldsymbol{p})$ и $d_\triangle(\boldsymbol{p})$ (а вместе с ними и экспоненциальные оценки сходимости) с бинарного на n-арный случай.

Следующее возможное обобщение — переход к операциям, введенным в [7, 9], обладающих квазигрупповыми свойствами лишь на некотором подмножестве $Q' \subseteq Q$. Операцию $f(x_1, \ldots, x_n)$ будем называть *n-арной Q'-поглощающе-квазигрупповой**), если для всех $i = 1, \ldots, n$ и любых $c_1, \ldots, c_{n-1} \in Q'$ функция $g(x) = f(c_1, \ldots, c_{i-1}, x, c_i, \ldots, c_{n-1})$ осуществляет перестановку на множестве Q'. Это обобщение естественно с той точки зрения, что представляет собой класс операций, которые приводят к возникновению предельного вероятностного закона при использовании их для построения преобразований независимых одинаково распределенных случайных величин на конечном множестве Q (см. [7, 9]).

Для Q'-поглощающе-квазигрупповых операций в работе [9] также доказана сходимость к равномерному на Q' распределению, однако доказательство вновь использует меру неравномерности, подобную d_δ. Для этого класса операций переход к другим мерам неравномерности, возможно, потребует бóльших усилий.

Еще одно возможное обобщение — рассмотрение вероятностных свойств бинарных операций, у которых требуется обратимость только справа или слева, а не обе обратимости сразу. Их свойства в качестве криптографических преобразователей рассматривались в работах И. В. Чередника (например, [5]).

---

*) Ранее использовался менее точный термин «Q'-поглощающая квазигрупповая».

СПИСОК ЛИТЕРАТУРЫ

1. Артамонов В. А. Квазигруппы и их приложения // Чебышевский сборник. — 2018. — Т. 19, № 2. — С. 111–122.
2. Воробьев Н. Н. Сложение независимых случайных величин на конечных абелевых группах // Математический сборник. — 1954. — Т. 34(76), № 1. — С. 89–126.
3. Глухов М. М. О применениях квазигрупп в криптографии // Прикладная дискретная математика. — 2008. — № 2. — С. 28–32.
4. Маршалл А., Олкин И. Неравенства: теория мажоризации и ее приложения. — М.: Мир, 1983. 576 с.
5. Чередник И. В. Развитие одного подхода к построению множества блочных биективных преобразований // Математические вопросы криптографии. — 2021. — Т. 12, № 3. — С. 49–66.
6. Яшунский А. Д. О преобразованиях распределений вероятностей бесповторными квазигрупповыми формулами // Дискретная математика. — 2013. — Т. 25, № 2. — С. 149–159.
7. Яшунский А. Д. О необходимых условиях предельных вероятностных теорем в конечных алгебрах // Докл. РАН. Математика, информатика, процессы управления. — 2020. — Т. 493. — С. 47–50.
8. Яшунский А. Д. О скорости сходимости квазигрупповых сверток вероятностных распределений // Дискретная математика. — 2022. — Т. 34, № 3. — С. 160–171.
9. Яшунский А. Д. О конечных алгебрах с предельным вероятностным законом // Алгебра и анализ. — 2022. — Т. 34, № 5. — С. 211–234.
10. Aldous D. On the Markov-chain simulation method for uniform combinatorial simulation and simulated annealing // Probability in the Engineering and Informational Sciences. — 1987. — V. 1, N1. — P. 33–46.
11. Chauhan D., Gupta I., Verma R. Quasigroups and their applications in cryptography // Cryptologia. — 2021. — V. 45, N3. — P. 227–265.
12. Diaconis P., Saloff-Coste L. Comparison techniques for reversible Markov chains // The Annals of Applied Probability. — 1993. — V. 3, N3. — P. 696–730.
13. Markovski S. Design of crypto primitives based on quasigroups // Quasigroups and Related Systems. — 2015. — V. 23. — P. 41–90.
14. Markovski S., Gligoroski D., Bakeva V. Quasigroup String Processing: Part 1 // Contrib. MANU, Sec. Math. Tech. Sci. — 1999. — V. XX, N1–2. — P. 13–28.
15. Markovski S., Bakeva V. Quasigroup String Processing: Part 4 // Contrib. MANU, Sec. Math. Tech. Sci. — 2006–2007. — V. XXVII–XXVIII, N1–2. — P. 41–53.
16. Markovski S., Kusakatov V. Quasigroup String Processing: Part 2 // Contrib. MANU, Sec. Math. Tech. Sci. — 2000. — V. XXI, N1–2. — P. 15–32.
17. Markovski S., Kusakatov V. Quasigroup String Processing: Part 3 // Contrib. MANU, Sec. Math. Tech. Sci. — 2002–2003. — V. XXIII–XXIV, N1–2. — P. 7–27.
18. Markovski S., Gligoroski D., Kocarev L. Unbiased Random Sequences from Quasigroup String Transformations // Fast Software Encryption. FSE 2005. LNCS, vol 3557/ Eds. H. Gilbert, H. Handschuh. — Springer, Berlin, Heidelberg, 2005. — Pp. 163–180.
19. Maurer U., Pietrzak K., Renner R. Indistinguishability Amplification // Advances in Cryptology — CRYPTO 2007, LNCS, vol. 4622. — Springer-Verlag, 2007. — Pp. 130–149.
20. Rioul O. What Is Randomness? The Interplay between Alpha Entropies, Total Variation and Guessing // Physical Sciences Forum. — 2022. — V. 5, N30. — P. 1–9.
21. Saloff-Coste L. Random walks on finite groups // Probability on discrete structures. Encyclopaedia Math. Sci., vol. 110. — Springer, Berlin, 2004. — Pp. 263–346.
22. Shaltiel R. An introduction to randomness extractors // Automata, Languages and Programming. ICALP 2011. LNCS, vol 6756/ Eds. L. Aceto, M. Henzinger, J. Sgall. — Springer, Berlin, Heidelberg, 2011. Pp. 21–41.
23. Scherbakov V. A. Quasigroups in cryptology // Computer Science Journal of Moldova. — 2009. — V. 17, N2(50). — P. 193–228.

Поступило в редакцию 30 X 2023