

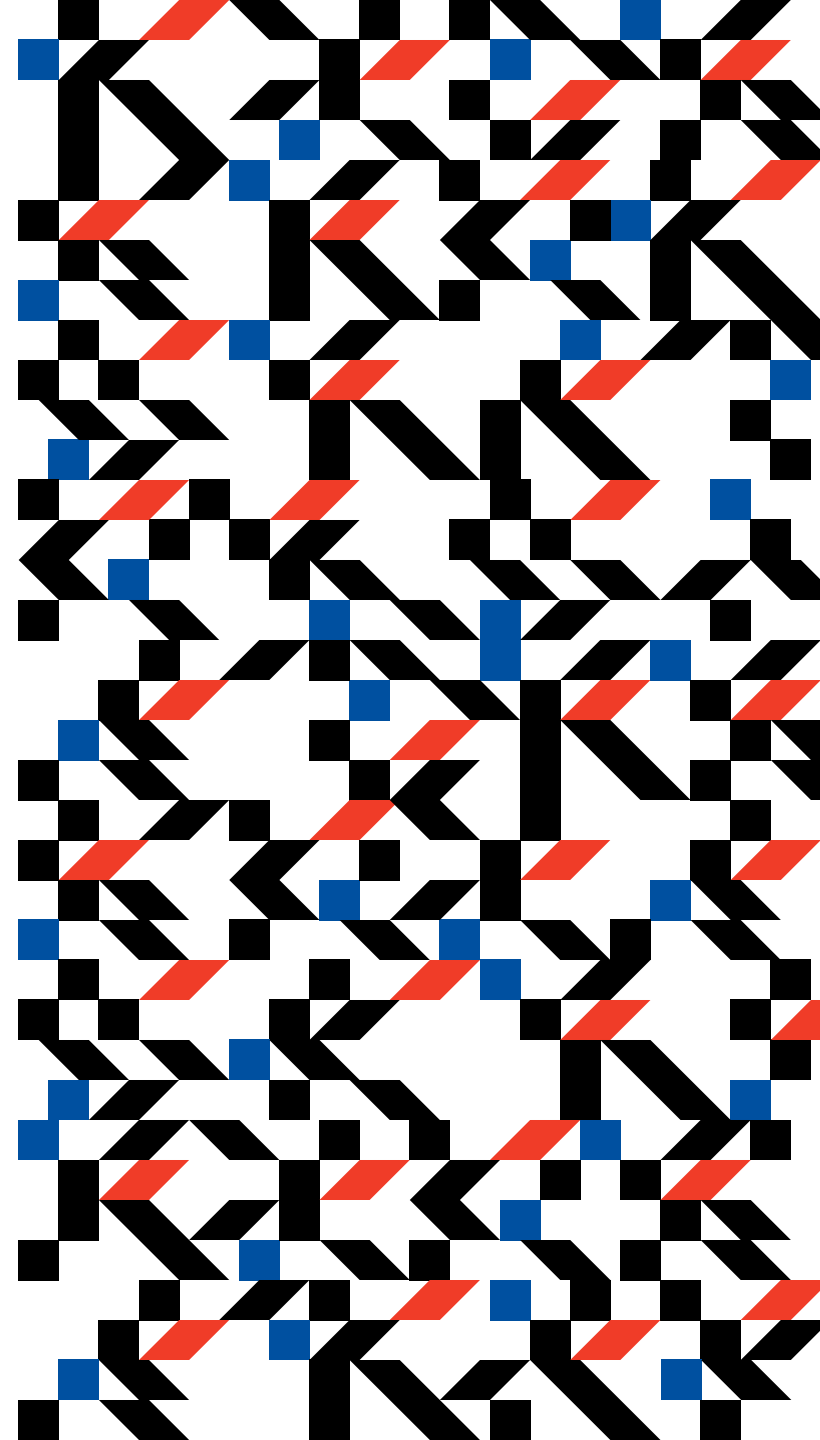


Теоретико-вероятностный анализ величины

преобладания
линейного статистического аналога для
несбалансированных сетей
Фейстеля с модульной редукцией

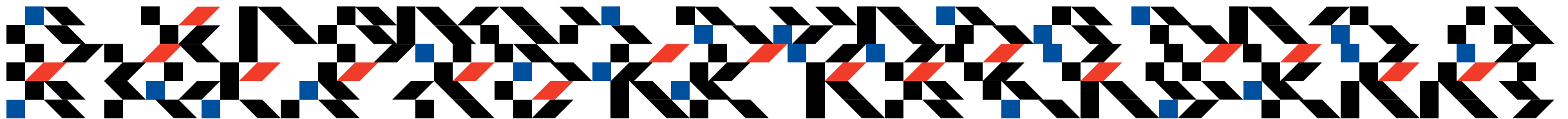
Каргин Александр, КриптоПро

Фомин Денис, Академия криптографии РФ



Что такое шифрование, сохраняющее формат (FPE – Format-preserving encryption)

Примеры, стандарты, уязвимости



Почему не получится использовать блочный шифр?



Почему, например, Кузнечик *не подходит*?

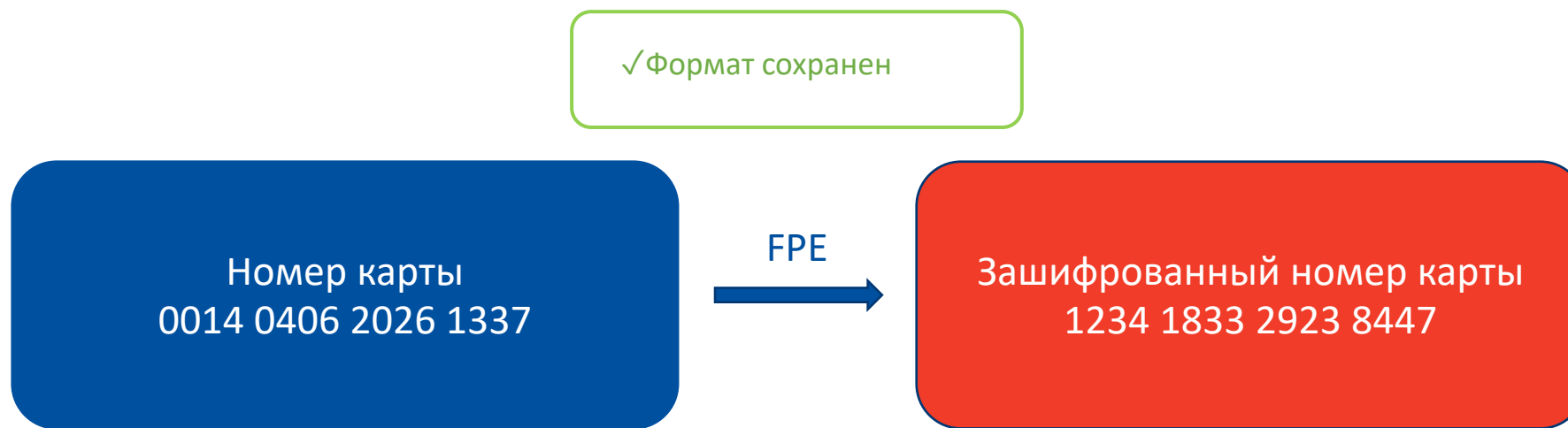
Кузнечик – это алгоритм шифрования, заданный отображением

$f: \{0, 1\}^{128} \rightarrow \{0, 1\}^{128}$ при фиксированном ключе, но:

- ▀ PAN – 16 десятичных цифр ≈ 53 бита, а не 128;
- ▀ Кузнечик отображает $\{0, 1\}^{128}$ в $\{0, 1\}^{128}$, а не в множество десятичных цифр;
- ▀ Нарушается схема БД, бизнес-логика.



Примеры применения FPE



Иные примеры в реальных задачах:

- ▀ $\{0, \dots, 9\}^{16}$ – номера банковских карт (PAN),
- ▀ $\{0, \dots, 9\}^{12}$ – ИНН,
- ▀ $\{0, \dots, 9\}^{11}$ – СНИЛС.

Data-Type Preserving Encryption – как пример FPE-алгоритма

▀ DTP – одна из первых практических FPE-схем

▀ Введен термин «*data-type preserving*» (позже заменен на *format-preserving*)

Упрощенный алгоритм DTP-2:

Для $X = x_1 x_2 \dots x_m \in \{0, \dots, d - 1\}^m$:

1. Разбить на блоки по 3 символа: (x_i, x_{i+1}, x_{i+2}) ,
2. Вычислить маску $z = E_K(x_i \| x_{i+1} \| x_{i+2})$,
3. Заменить $x_{i+2} \leftarrow (x_{i+2} + z) \bmod d$,
4. Повторить 3 раунда с циклическим сдвигом блоков.



Data-Type Preserving Encryption – как пример FPE-алгоритма

/// DTP – одна из первых практических FPE-схем

/// Введен термин «*data-type preserving*» (позже заменен на *format-preserving*)

Упрощенный алгоритм DTP-2:

Для $X = x_1 x_2 \dots x_m \in \{0, \dots, d - 1\}^m$:

1. Разбить на блоки по 3 символа: (x_i, x_{i+1}, x_{i+2}) ,
2. Вычислить маску $z = E_K(x_i \| x_{i+1} \| x_{i+2})$,
3. Заменить $x_{i+2} \leftarrow (x_{i+2} + z) \bmod d$,
4. Повторить 3 раунда с циклическим сдвигом блоков.

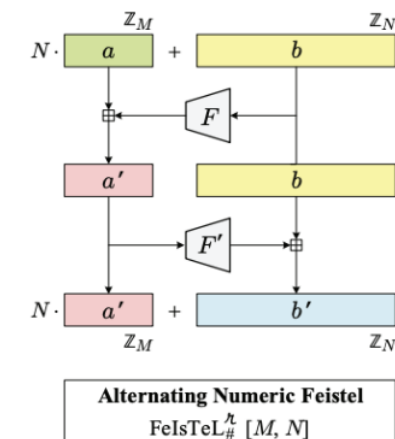
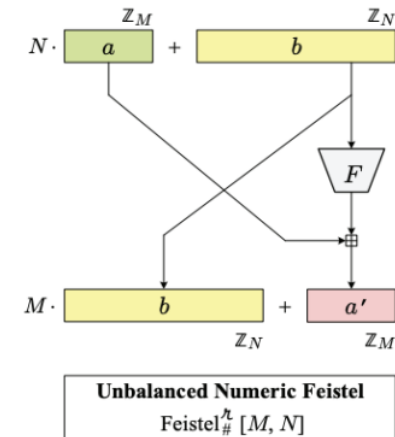
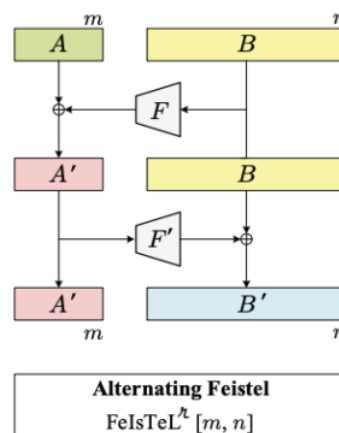
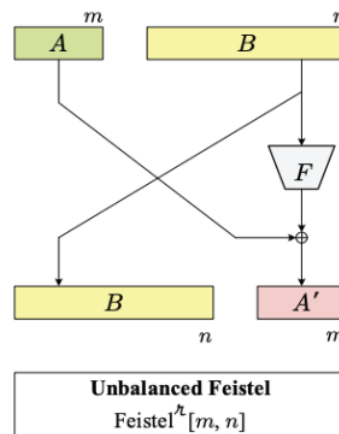
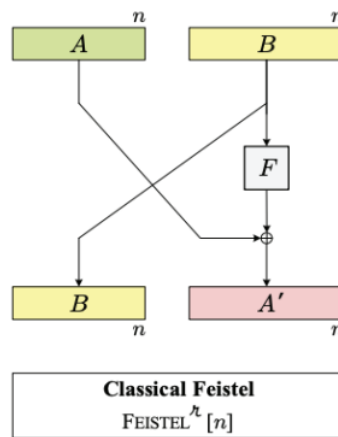
Data-Type Preserving Encryption – как пример FPE-алгоритма

Уязвимость к статистическим атакам:

- Если $d = 10$ (цифры), а $W = 256$ (1 байт), то символы 0 ... 5 появляются с вероятностью $26/256$, а 6 ... 9 – $25/256$.
- Имея порядка $q = 256$ шифртекстов одного PIN при разных T , можно восстановить его с вероятностью $> 90\%$;
- Для полного 16-значного PAN хватает $\sim 2^{41}$ шифртекстов.

Сети Фейстеля для FPE

Сеть Фейстеля – универсальная конструкция для построения FPE-алгоритмов.



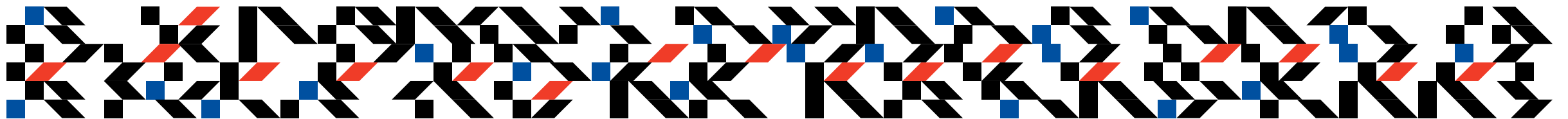
Стандарты

Схема	Когда стандартизировано	Уязвимости	Нынешнее время
FF1	Стандартизована NIST в 2016 году	Атаки 2016, 2019 годов — восстановление PIN за 256 запросов при малых областях значений	В актуальном NIST-стандарте
FF2	В черновике NIST 2013 года	Уязвимость 2015 года — уровень безопасности < 128 битов → не вошла в стандарт	Удалён в 2016: NSA указала, что он не обеспечивает 128-битную стойкость; анализ 2015 года подтвердил уязвимость
FF3	Стандартизована NIST в 2016 году	Атаки 2017, 2018, 2020 годов — восстановление CCN за 241 запросов; модифицирована в FF3-1 (2019)	Удалено в 2025 из-за ряда атак
FEA-1, FEA-2	Корейские стандарты	Аналогичны FF1/FF3 — уязвимы к атакам 2021 года	Нет информации



Теоретико-вероятностные характеристики распределения $\mathcal{G}(\mathbb{Z}_N)$

Определения, линейные соотношения,
характеристики



Определение распределения $\mathcal{G}(\mathbb{Z}_N)$

В данной работе используются следующие обозначения:

$$N = q^L, \ell = 2^{\lceil \log_2 N \rceil}, r = \ell - N.$$

Определение 1

Пусть $X \sim \mathcal{U}(\mathbb{F}_2^n)$, $Y = X \pmod N$. Тогда распределение Y на $\mathbb{Z}_N$ обозначается $\mathcal{G}(\mathbb{Z}_N)$ и задается функцией вероятности

$$\Pr\{Y = i\} = \begin{cases} 2^{1-n}, & 0 \leq i < 2^n - N, \\ 2^n, & 2^n - N \leq i < N. \end{cases}$$

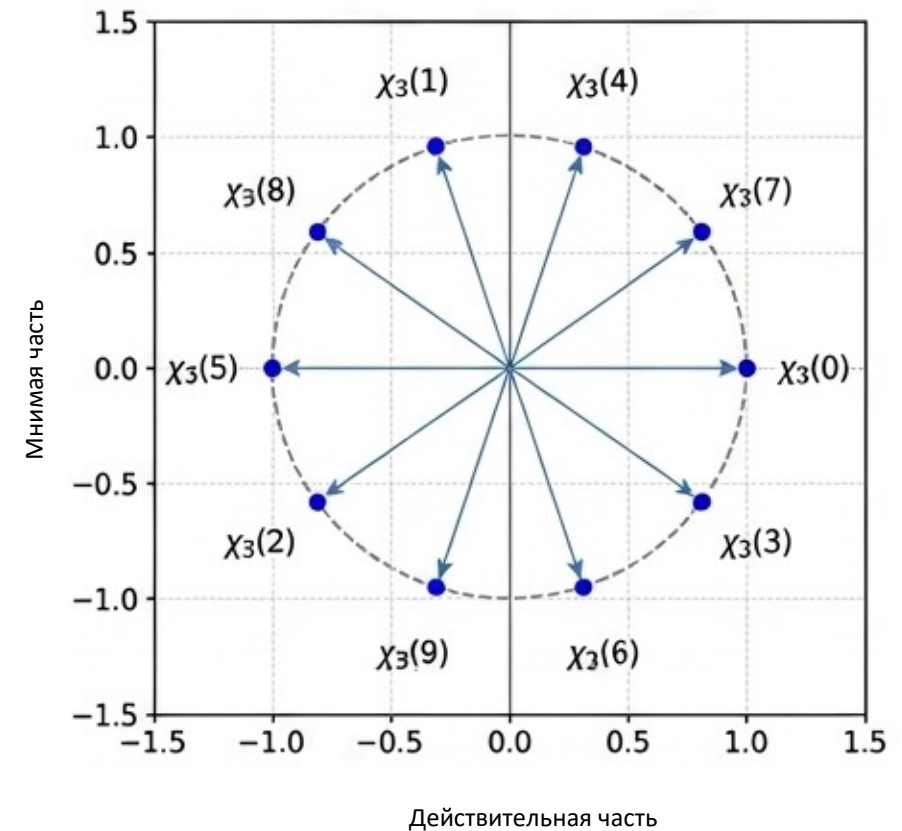
Характеры конечных абелевых групп и коэффициент преобразования линейного статистического аналога

/// Характером абелевой группы $\mathbf{X}$ называется любой гомоморфизм $\chi: \mathbf{X} \rightarrow \mathbb{C}^\times$ из группы $\mathbf{X}$ в мультипликативную группу поля комплексных чисел $\mathbb{C}^\times$.

/// Для циклической группы $\mathbb{Z}_n$ любой характер является комплексным корнем n -й степени из единицы. Все возможные характеры описывают n отображениями вида:

$$\chi_j(x) = \exp\left(\frac{2\pi i}{n} \cdot x \cdot j\right), j \in \mathbb{Z}_n.$$

Пример $\chi_3(x)$ в $\mathbb{Z}_{10}$

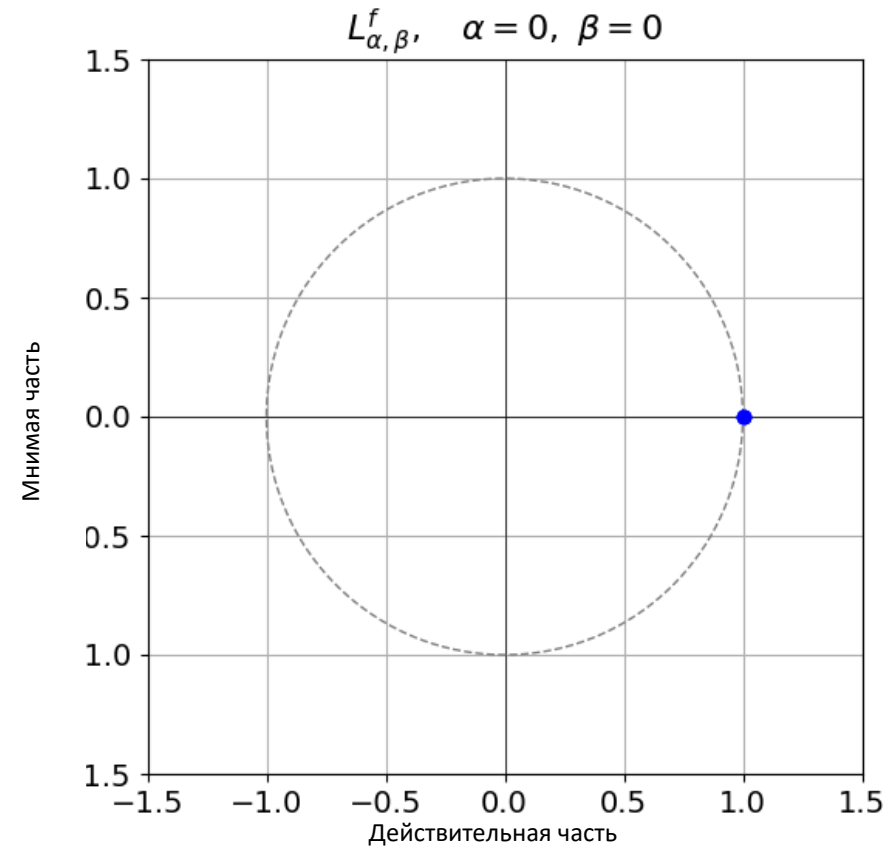


Характеры конечных абелевых групп и коэффициент преобладания линейного статистического аналога

Значение

$$L_{\beta, \alpha}^F = M[\psi_{\beta}(F(x)) \overline{x_{\alpha}(x)}]$$

будем называть преобладанием
линейного статистического аналога



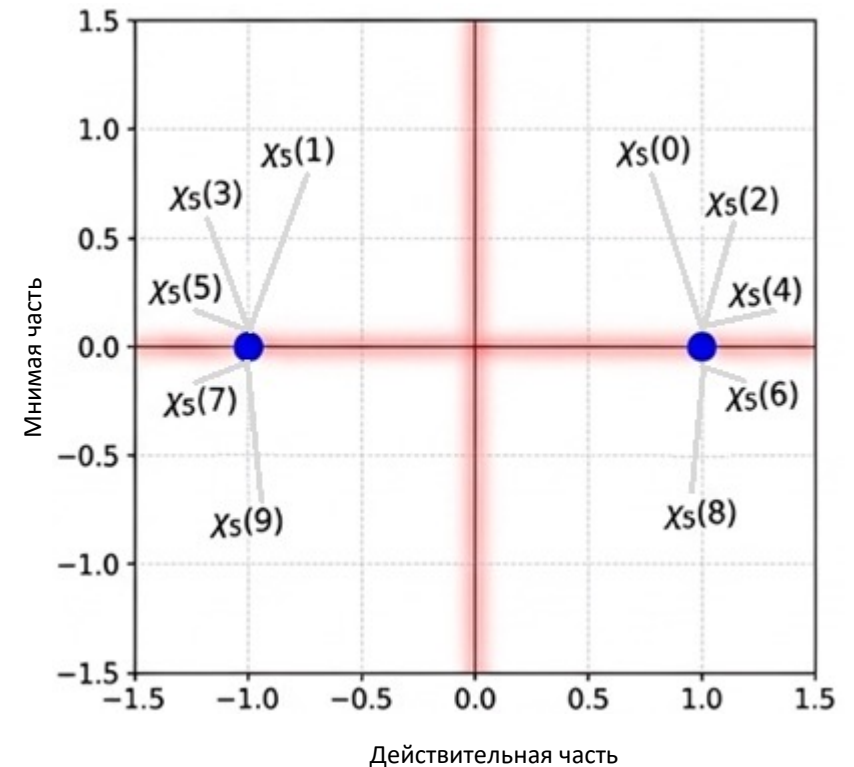
Характеры конечных абелевых групп и коэффициент преобразования линейного статистического аналога

Пусть $\mathbf{X}$ - произвольная абелева группа $|\mathbf{X}| \geq 2$, $\beta \in \mathbf{X} \setminus \{0\}$. Тогда характер ψ_β группы $\mathbf{X}$ будем называть *2-характером*, если для любого $x \in \mathbf{X}$ значение $\psi_\beta(x)$ можно представить в виде

$$\psi_\beta(x) = \exp(\pi i \cdot v(x)),$$

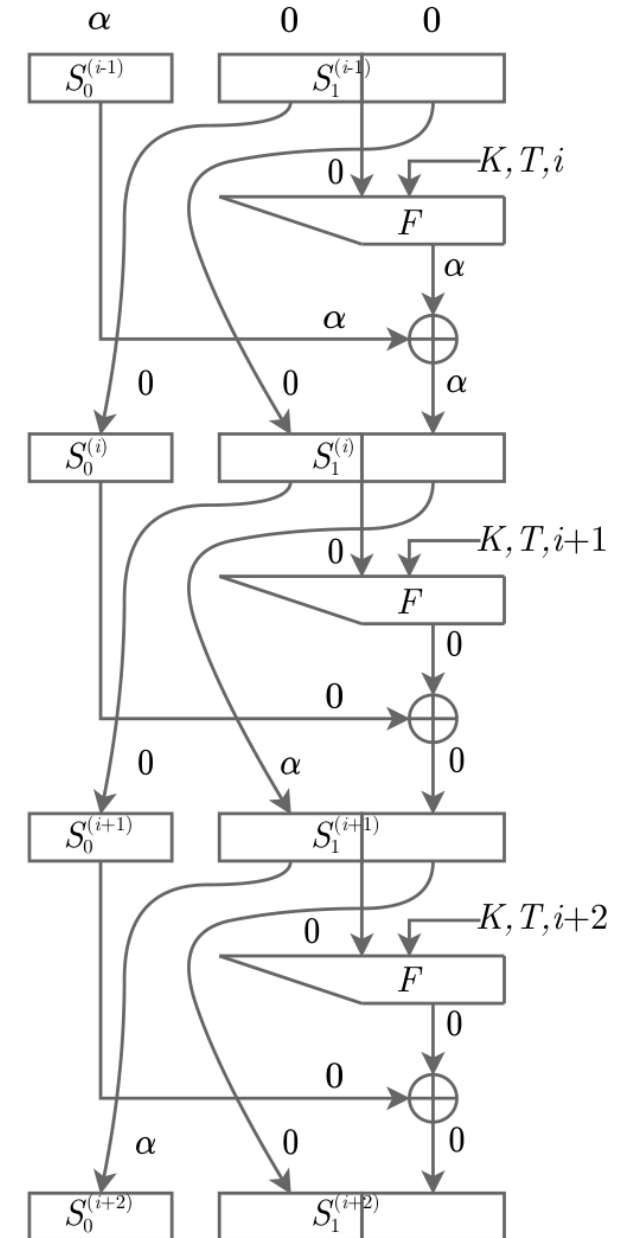
для некоторых $v(x) \in \mathbb{N}_0$.

Пример $\chi_5(x)$ в $\mathbb{Z}_{10}$



Согласованные линейные соотношения

Такой выбор текстов *максимизирует преобладание глобального линейного статистического аналога* — стандартный приём в линейном анализе сетей Фейстеля с небиективной раундовой функцией.



Математическое ожидание и дисперсия $L_{\beta,0}^F$

Утверждение 2

Для любого $\beta \in \mathbb{Z}_N$

$$ML_{\beta,0}^F = \begin{cases} 1 & , \beta = 0, \\ \frac{1}{l} \cdot \frac{1-e^{2\pi\beta r/N}}{1-e^{2\pi\beta/N}}, & \beta \not\equiv 0(\bmod N). \end{cases}$$

Утверждение 5

Для любого $\beta \in \mathbb{Z}_N$

$$DL_{\beta,0}^F = \begin{cases} 0 & , \beta = 0, \\ \frac{1}{q^R} \left(1 - \frac{1}{\ell^2} \cdot \frac{\sin^2(\pi\beta r/N)}{\sin^2(\pi\beta/N)} \right), & \beta \not\equiv 0(\bmod N). \end{cases}$$

Иные характеристики $L_{\beta,0}^F$ **Следствие 1**

Для любого $\beta \not\equiv 0 \pmod{N}$

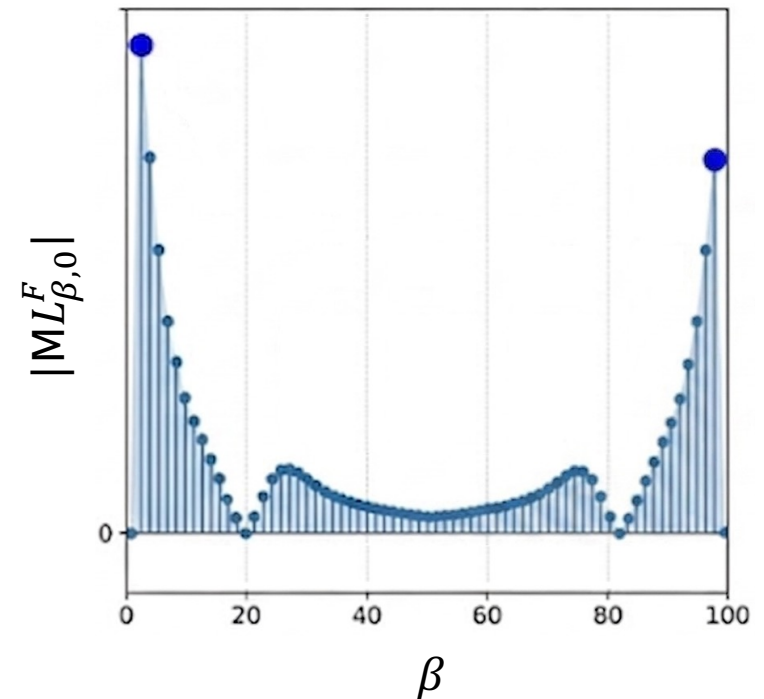
$$|ML_{\beta,0}^F| = \frac{1}{\ell} \cdot \left| \sin\left(\frac{\pi\beta r}{N}\right) / \sin\left(\frac{\pi\beta}{N}\right) \right|.$$

Следствие 2

Пусть $q, L \in \mathbb{N}$ и $1 \leq r < N$. Тогда

$$\max_{1 \leq \beta \leq N-1, \beta \in \mathbb{Z}} \left| \frac{\sin(\pi r \beta / N)}{\sin(\pi \beta / N)} \right| = \left| \frac{\sin(\pi r / N)}{\sin(\pi / N)} \right|$$

Пример при $q = 10, L = 2$



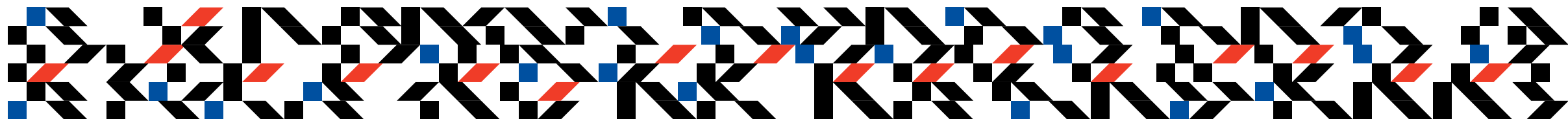
Иные характеристики $L_{\beta,0}^F$

Утверждение 3

Пусть $\beta \in \mathbb{Z}_N \setminus \{0\}$. Тогда $ML_{\beta,0}^F = 0$ тогда и только тогда, когда $N|\beta r$.

Асимптотическое распределение преобладания линейного статистического аналога

Основные результаты



Комплексное нормальное распределение

Комплексная случайная величина Z называется *комплексно нормальной* с параметрами $(\mu_0, \sigma^2, \gamma)$, что обозначается

$$Z \sim \mathcal{N}_{\mathbb{C}}(\mu_0, \sigma^2, \gamma),$$

если её можно представить как $Z = X + iY$, где вектор $(X, Y)^T$ имеет двумерное вещественное нормальное распределение $\mathcal{N}_2(\boldsymbol{\mu}, \Sigma)$ с

$$\boldsymbol{\mu} = \begin{pmatrix} \Re \mu_0 \\ \Im \mu_0 \end{pmatrix}, \Sigma = \frac{1}{2} \begin{pmatrix} \sigma^2 + \Re \gamma & -\Im \gamma \\ -\Im \gamma & \sigma^2 - \Re \gamma \end{pmatrix}.$$

Параметры имеют следующий смысл:

$$\sigma^2 = \mathbb{M}[|Z - \mu_0|^2] \text{ — дисперсия,}$$

$$\gamma = \mathbb{M}[(Z - \mu_0)^2] \text{ — псевдоковариация}$$

Дисперсия σ^2 всегда вещественна и неотрицательна; псевдоковариация $\gamma \in \mathbb{C}$ измеряет отклонение от *круговой симметрии*

В случае $\gamma = 0$ распределение называется *круговым*, $\mathcal{CN}(\mu_0, \sigma^2)$.

Основной результат: предельное распределение (сходимость к вещественно-нормальному закону)

Теорема 1

Пусть $Y \sim \mathcal{G}(\mathbb{Z}_N)$ и пусть $\beta \in \mathbb{Z}_N \setminus \{0\}$ удовлетворяет условиям:

1. χ_β является 2-характером, то есть $2\beta \equiv 0 \pmod{N}$;
2. $ML_{\beta,0}^F = 0$, то есть (см. утверждение 3) $N \mid \beta r$ и $N \nmid \beta$.

Тогда при $F(x) \xrightarrow{i.i.d.} \mathcal{G}(\mathbb{Z}_N)$, $x = 1, \dots, n = |\mathbb{Z}_N|$, нормированная величина

$$Z_n := \sqrt{n} \cdot L_{\beta,0}^F = \frac{1}{\sqrt{n}} \sum_{x=1}^n x_\beta(F(x))$$

сходится по распределению к стандартному вещественно-нормальному закону:

$$Z_n \xrightarrow{d} \mathcal{N}(0,1).$$

Теорема 3

Пусть $Y \sim \mathcal{G}(\mathbb{Z}_N)$ и пусть $\beta \in \mathbb{Z}_N \setminus \{0\}$ удовлетворяет условиям:

1. χ_β является 2-характером, то есть $2\beta \equiv 0 \pmod{N}$;
2. $ML_{\beta,0}^F = \mu \neq 0$, то есть (см. утверждение 3) $N \nmid \beta r$.

Тогда при $F(x) \xrightarrow{i.i.d.} \mathcal{G}(\mathbb{Z}_N)$, $x = 1, \dots, n = |\mathbb{Z}_N|$, центрированная и нормированная величина

$$\tilde{Z}_n := \sqrt{n} \cdot (L_{\beta,0}^F - \mu) = \frac{1}{\sqrt{n}} \sum_{x=1}^n (x_\beta(F(x)) - \mu)$$

сходится по распределению к вещественно-нормальному закону:

$$\tilde{Z}_n \xrightarrow{d} \mathcal{N}(0, 1 - \mu^2).$$

Основной результат: предельное распределение (сходимость к комплексно - нормальному закону)

Теорема 2

Пусть $Y \sim \mathcal{G}(\mathbb{Z}_N)$ и пусть $\beta \in \mathbb{Z}_N \setminus \{0\}$ удовлетворяет условиям:

1. χ_β не является 2-характером, то есть $2\beta \not\equiv 0 \pmod{N}$;
2. $ML_{\beta,0}^F = 0$, то есть (см. утверждение 3) $N \mid \beta r$ и $N \nmid \beta$;
3. $ML_{2\beta,0}^F = 0$, то есть $N \mid 2\beta r$ и $N \nmid 2\beta$.

Тогда при $F(x) \xrightarrow{i.i.d.} \mathcal{G}(\mathbb{Z}_N)$, $x = 1, \dots, n = |\mathbb{Z}_N|$, нормированная величина

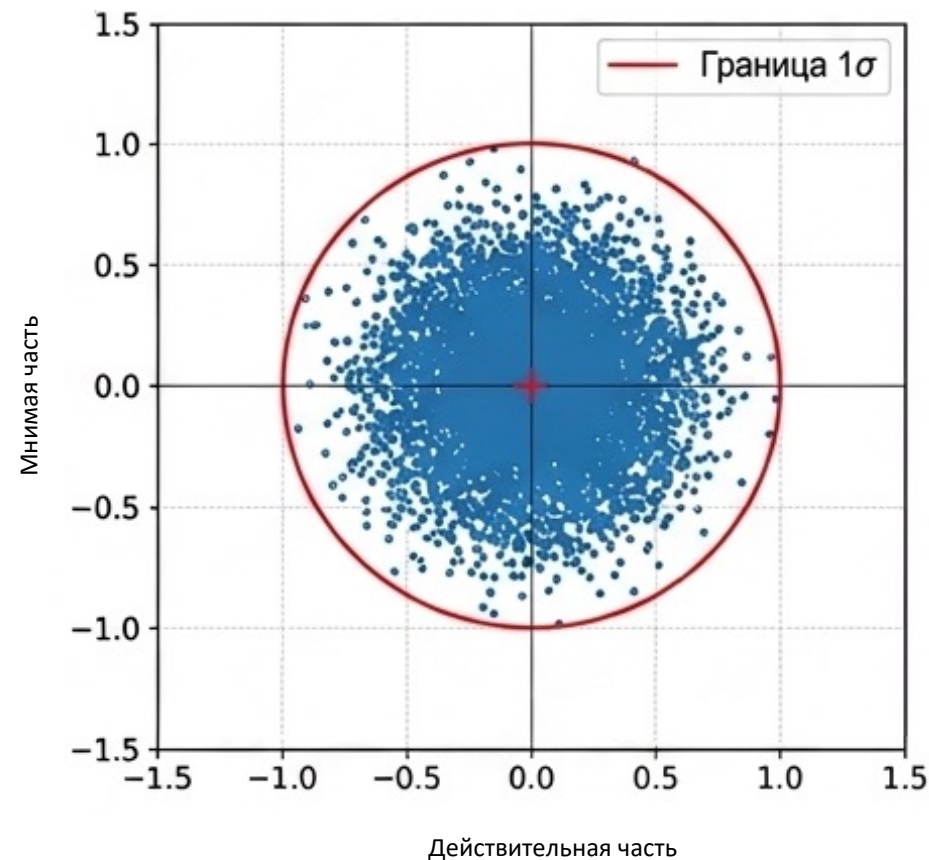
$$Z_n := \sqrt{n} \cdot L_{\beta,0}^F = \frac{1}{\sqrt{n}} \sum_{x=1}^n x_\beta(F(x))$$

сходится по распределению к стандартному комплексному нормальному закону:

$$Z_n \xrightarrow{d} \mathcal{CN}(0,1),$$

то есть $\Re Z_n, \Im Z_n \sim \mathcal{N}\left(0, \frac{1}{2}\right)$.

Пример при $q = 10, L = 2, \beta = 1$



Основной результат: предельное распределение (сходимость к комплексно - нормальному закону)

Теорема 4

Пусть $Y \sim \mathcal{G}(\mathbb{Z}_N)$ и пусть $\beta \in \mathbb{Z}_N \setminus \{0\}$ удовлетворяет условиям:

1. χ_β не является 2-характером, то есть $2\beta \not\equiv 0 \pmod{N}$;
2. $ML_{\beta,0}^F = \mu \neq 0$, то есть (см. утверждение 3) $N \nmid \beta$.

Обозначим $\mu_2 = ML_{2\beta,0}^F$. Тогда:

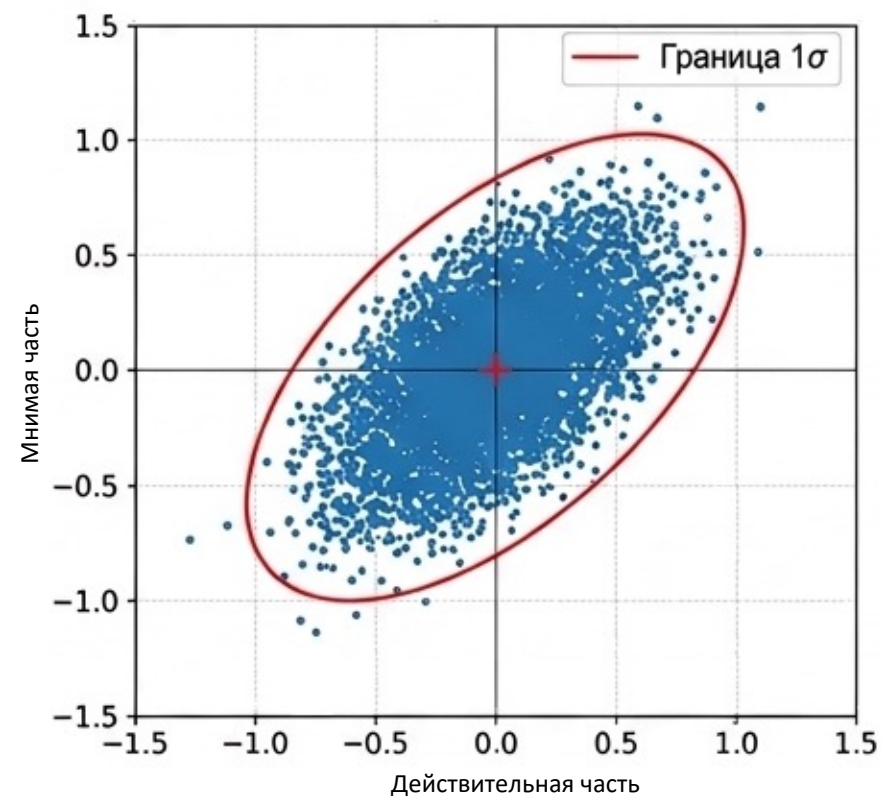
1. $\gamma := \mu_2 - \mu^2 \neq 0$;
2. при $F(x) \xrightarrow{i.i.d.} \mathcal{G}(\mathbb{Z}_N), x = 1, \dots, n = |\mathbb{Z}_N|$, центрированная и нормированная величина

$$\tilde{Z}_n := \sqrt{n} \cdot (L_{\beta,0}^F - \mu)$$

сходится по распределению к комплексному нормальному закону:

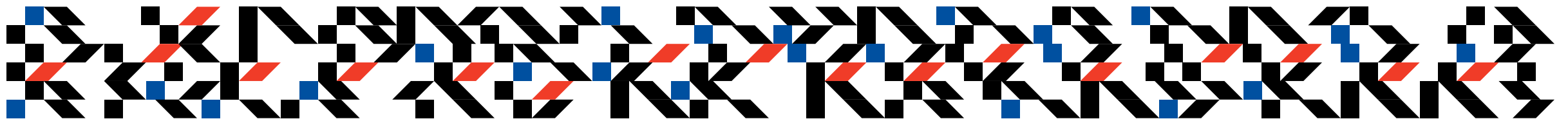
$$\tilde{Z}_n \xrightarrow{d} \mathcal{N}_{\mathbb{C}}(0, 1 - |\mu^2|, \gamma).$$

Пример при $q = 10, L = 2, \beta = 1$



Асимптотика модуля и квадрата модуля линейного статистического аналога

Основные результаты



Случаи классификации асимптотического поведения модуля и квадрата модуля линейного статистического аналога

$\mu = 0, \chi_\beta$ — 2-характер	$\mu = 0, \chi_\beta$ не является 2-характером
В этом случае $\chi_\beta(x) \in \{\pm 1\}, \forall x$, и из теоремы 1 следует, что $Z_n := \sqrt{n}L_{\beta,0}^F \xrightarrow{d} \mathcal{N}(0,1).$ Так как $L_{\beta,0}^F$ вещественнозначна, имеем $L_{\beta,0}^F = Z_n /\sqrt{n}$ $\Rightarrow$ $\sqrt{n} L_{\beta,0}^F \xrightarrow{d} \mathcal{N}(0,1) .$ Для квадрата модуля получаем $n L_{\beta,0}^F ^2 = Z_n^2 \xrightarrow{d} \chi_1^2,$ где $L_{\beta,0}^F$ — хи-квадрат распределение с одной степенью свободы	Здесь $\mu = 0, \mu_2 = 0$, и по теореме 2 $Z_n := \sqrt{n}L_{\beta,0}^F \xrightarrow{d} \mathcal{CN}(0,1),$ то есть $\Re Z_n, \Im Z_n \sim \mathcal{N}(0, 1/2)$ независимо. Тогда $\sqrt{n} L_{\beta,0}^F = Z_n \xrightarrow{d} \text{Rayleigh}\left(\frac{1}{\sqrt{2}}\right),$ с плотностью вероятности $f(r) = 2re^{-r^2}, r \geq 0$. Квадрат модуля имеет экспоненциальное распределение: $n L_{\beta,0}^F ^2 = Z_n^2 \xrightarrow{d} \text{Exp}(1),$ с плотностью вероятности $f(u) = e^{-u}, u \geq 0$.

Случаи классификации асимптотического поведения модуля и квадрата модуля линейного статистического аналога

$\mu \neq 0, \chi_\beta$ — 2-характер

В этом случае $\mu \in (-1, 1) \setminus \{0\}$, и по теореме 3

$$\tilde{Z}_n := \sqrt{n}(L_{\beta,0}^F - \mu) \xrightarrow{d} \mathcal{N}(0, 1 - \mu^2).$$

Для модуля, поскольку $\mu \neq 0$, функция $x \mapsto |x|$ дифференцируема в окрестности μ . По дельта-методу,

$$\sqrt{n}(|L_{\beta,0}^F| - |\mu|) \xrightarrow{d} \mathcal{N}(0, 1 - \mu^2),$$

поскольку производная $|\cdot|$ в точке $\mu \neq 0$ равна $\text{sign}(\mu)$, а её квадрат равен 1. Аналогично, для квадрата модуля, применяя дельта-метод к функции $g(x) = x^2$ с $g'(\mu) = 2\mu$, получаем

$$n(|L_{\beta,0}^F|^2 - \mu^2) \xrightarrow{d} \mathcal{N}(0, 4\mu^2(1 - \mu^2)).$$

$\mu \neq 0, \chi_\beta$ не является 2-характером

Здесь $\mu \in \mathbb{C}, |\mu| < 1$, и по теореме 4

$$\tilde{Z}_n := \sqrt{n}(L_{\beta,0}^F - \mu) \xrightarrow{d} \mathcal{N}_{\mathbb{C}}(0, \sigma^2, \gamma),$$

где $\sigma^2 = 1 - |\mu|^2, \gamma = \mu_2 - \mu^2 \neq 0$. Применим дельта-метод к функции $h(z) = |z|$. Её градиент в точке $\mu \neq 0$ равен $\nabla h(\mu) = (\Re(\bar{\mu}/|\mu|), \Im(\bar{\mu}/|\mu|))^T$. Тогда

$$\sqrt{n}(|L_{\beta,0}^F| - |\mu|) \xrightarrow{d} \mathcal{N}(0, \tau^2),$$

$$\text{где } \tau^2 = \frac{1}{2} \left(1 - \frac{|\mu|^2 + \Re(\gamma \frac{\bar{\mu}^2}{|\mu|^2})}{|\mu|^2} \right).$$

Для квадрата модуля, $g(z) = |z|^2 = z\bar{z}$, градиент равен $\nabla g(\mu) = (2\Re\mu, 2\Im\mu)^T$. Следовательно,

$$n(|L_{\beta,0}^F|^2 - \mu^2) \xrightarrow{d} \mathcal{N}(0, 4\tau_1^2),$$

где

$$\tau_1^2 = \frac{1}{2} (1 - |\mu|^2 + \Re\gamma).$$

Выводы:

Основные результаты

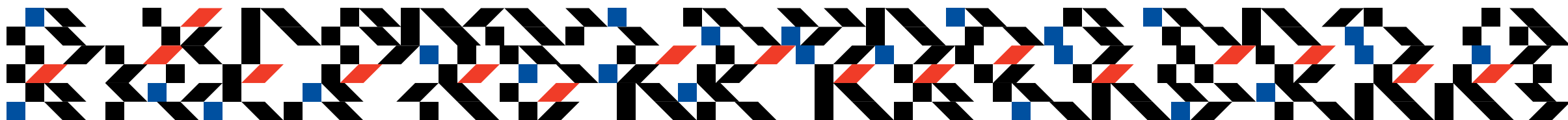
➤ Получено **компактное аналитическое выражение** для $ML_{\beta,0}^F$, устраняющее необходимость суммирования по $O(q^L)$ слагаемым.

➤ Установлено **точное условие обнуления смещения**: $ML_{\beta,0}^F = 0 \Leftrightarrow N|\beta r, N \nmid \beta$.

➤ Выведены **явные формулы** для модуля и дисперсии.

➤ Доказана **сходимость** централизованной и нормированной величины

$\tilde{Z}_n = \sqrt{n}(L_{\beta,0}^F - \mu)$ к одному из четырёх предельных распределений, полностью определяемых параметрами $\mu = ML_{\beta,0}^F, \mu_2 = ML_{2\beta,0}^F$ и типом характера χ_β .

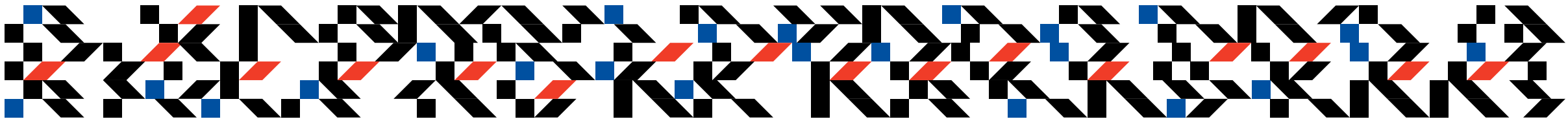


Выводы:

Предельные распределения $\tilde{Z}_n = \sqrt{n}(L^F_{\beta,0} - \mu)$ при $F(x) \sim \mathcal{G}(\mathbb{Z}_N)$

№	Условия на β	μ	Тип χ_β	Предел	Параметры
1	$N \beta r, N \nmid \beta$	0	2-характер	$\mathcal{N}(0,1)$	$\sigma^2 = 1$
2	$N \beta r, N \nmid \beta,$ $N \mid 2\beta r, N \nmid 2\beta$	0	Не 2-характер	$\mathcal{CN}(0,1,0)$	$\sigma^2 = 1, \gamma = 0$
3	$N \nmid \beta r$	$\neq 0$	2-характер	$\mathcal{N}(0,1 - \mu ^2)$	$\sigma^2 = 1 - \mu ^2$
4	$N \nmid \beta r$	$\neq 0$	Не 2-характер	$\mathcal{CN}(0,1 - \mu ^2, \mu_2 - \mu^2)$	$\sigma^2 = 1 - \mu ^2, \gamma \neq 0$

Здесь $N = q^L, r = 2^{\lceil \log_2 N \rceil} - N, \mu = ML^F_{\beta,0}, \mu_2 = ML^F_{2\beta,0}. \chi_\beta$ – 2 – характер $\Leftrightarrow 2\beta \equiv 0(\text{mod } N)$.

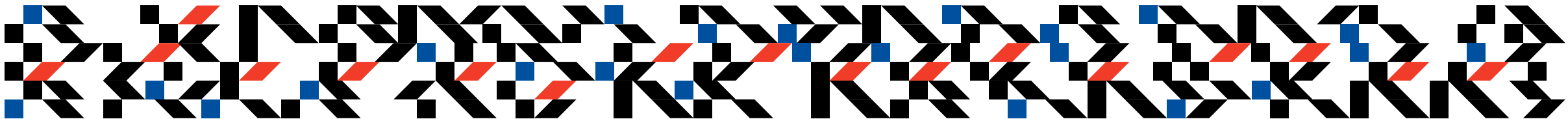


Выводы:

Асимптотика $|L_{\beta,0}^F|$ и $|L_{\beta,0}^F|^2$

№	$\sqrt{n}(L_{\beta,0}^F - \mu)$	$n(L_{\beta,0}^F ^2 - \mu ^2)$
1	$ \mathcal{N}(0,1) $	χ_1^2
2	$\text{Rayleigh}(\frac{1}{\sqrt{2}})$	$\text{Exp}(1)$
3	$\mathcal{N}(0,1 - \mu^2)$	$\mathcal{N}(0, 4\mu^2(1 - \mu^2))$
4	$\mathcal{N}(0, \tau^2)$	$\mathcal{N}(0, 4\tau_1^2)$

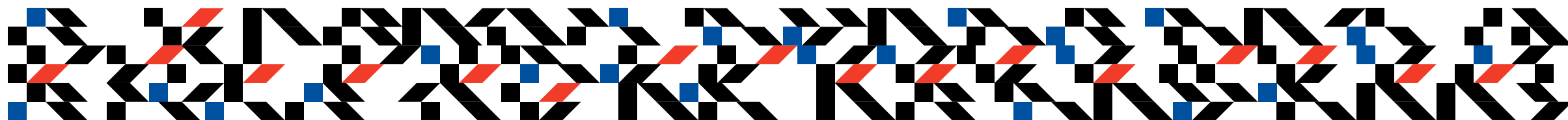
Здесь $\mu = ML_{\beta,0}^F$, $\gamma = ML_{2\beta,0}^F - \mu^2$, $\tau^2 = \frac{1}{2}(1 - |\mu|^2 + \Re(\gamma\bar{\mu}^2/|\mu|^2))$, $\tau_1^2 = \frac{1}{2}(1 - |\mu|^2 + \Re\gamma)$.



Выводы:

Перспективные направления

- /// Обобщение на разностный метод (через двойственность);
- /// Анализ композиции раундовых функций;
- /// Построение универсальных различителей на основе предельных распределений.





Спасибо за внимание