Воронежский государственный университет

# Анализ случайных векторов частот дискретных распределений отсчетов потоков излучения методом комплексных моментов

## Analysis of random vectors, frequencies of discrete distributions of reference streams, by the method of complex moments

Близняков Н. М., Вахтель В. М., Костомаха Д. Е., Работкин В. А.

При статистическом анализе данных если тип распределения дискретной случайной величены(СВ) – $k_j = 0; 1; \ldots$ и случайного вектора (СВР) $\nu(.) = \left(\nu_0, .. \nu_j \ldots, \nu_l\right)$;

$\nu_j = \nu_j(k = j)$ известен, то функцию и вектор параметров $\theta$ функции распределения $F(k|\theta)$ можно во многих случаях оценить, на основе известного метода моментов используя несколько первых целых эмпирических моментов распределения $\mu(k)$.

Последовательностям выборок малого объёма $n \leq 10$ чисел отсчётов  потоков частиц при среднем $0 \leq \bar{k} \leq 5$ соответствуют случайные векторы (СВР - $\nu(.)$ ) частот $\nu_j(k_j)$
значений

$$k_j : \nu(.) = (\nu_0, \nu_1, ..., \nu_l), \quad n = \sum_{j=0}^{l} \nu_j(k_j)$$

и  СВР относительных частот $\nu'_j(.) = \nu_j / n$.

Анализ однородности отдельных пар СВР в их больших $M >> 1$ последовательностях остаётся критической процедурой обработки данных.

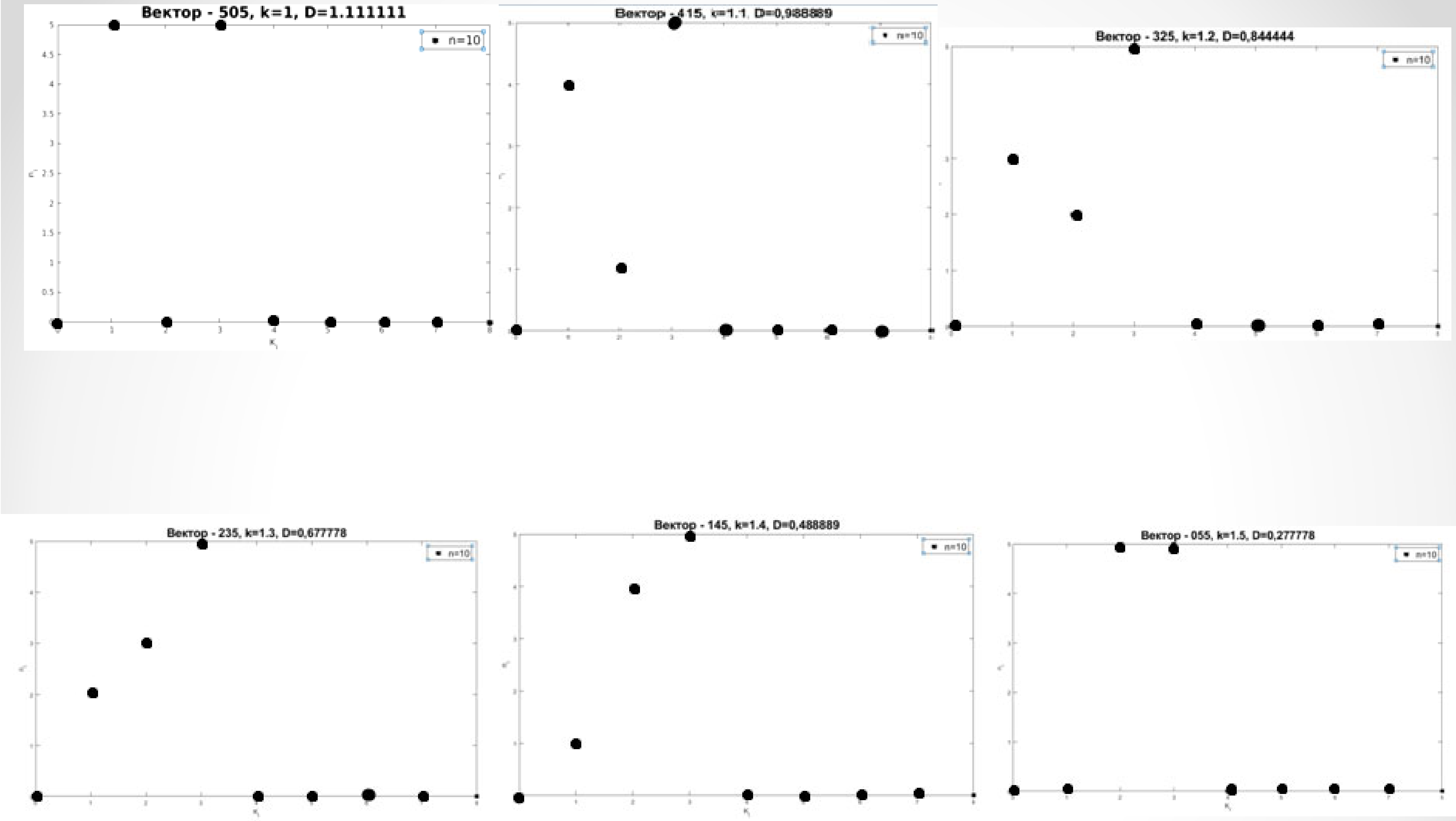


Рис. 1. Пример эмпирических векторов $\nu(.) = \nu_0(k=0), \nu_1(k=1), \nu_2(k=2)$: 505 – 055 «генеральной» выборки $\tilde{k} = 1{,}17623$, объёмом $\sim 10^7$.

В работе предложен метод оценки однородности СВР $\nu(.)$ и $\nu'(.)$ предварительно сгруппированных в пиках с несколькими фиксированными компонентами $\nu_j(.)$ многомодальных распределений функционалов – идентификаторов СВР $ID(\nu(.),a)$.

$$ID(\nu(.))_m = a_0\nu_{0.m} + a_1\nu_{1.m} + \ldots + a_l\nu_{l.m},$$

где $a = (a_0, \ldots, a_l)$- заданный проектирующий вектор, например, $a_0 < a_1 < \ldots < a_l$ или $a_0 > a_1 > \ldots > a_l$ с целочисленными компонентами.

Метод основан на анализе метрики - расстояния $\rho(\mu(\nu,\ldots,S)_m, \mu(\nu,\ldots,S))_q$ проекций фазовых траекторий комплексных функций эмпирических центральных моментов СВР **<u>дробных</u>** порядков $S > 1, S = r + \alpha, r = 0, 1, 2., 0 < \alpha < 1$

$$\mu(\nu(.), S) = \frac{1}{n-1}\sum_{j=1}^{n}(k_j - \bar{k})^S = \mathrm{Re}(\mu(\nu, S)) + i\,\mathrm{Im}(\mu(\nu, S)),$$

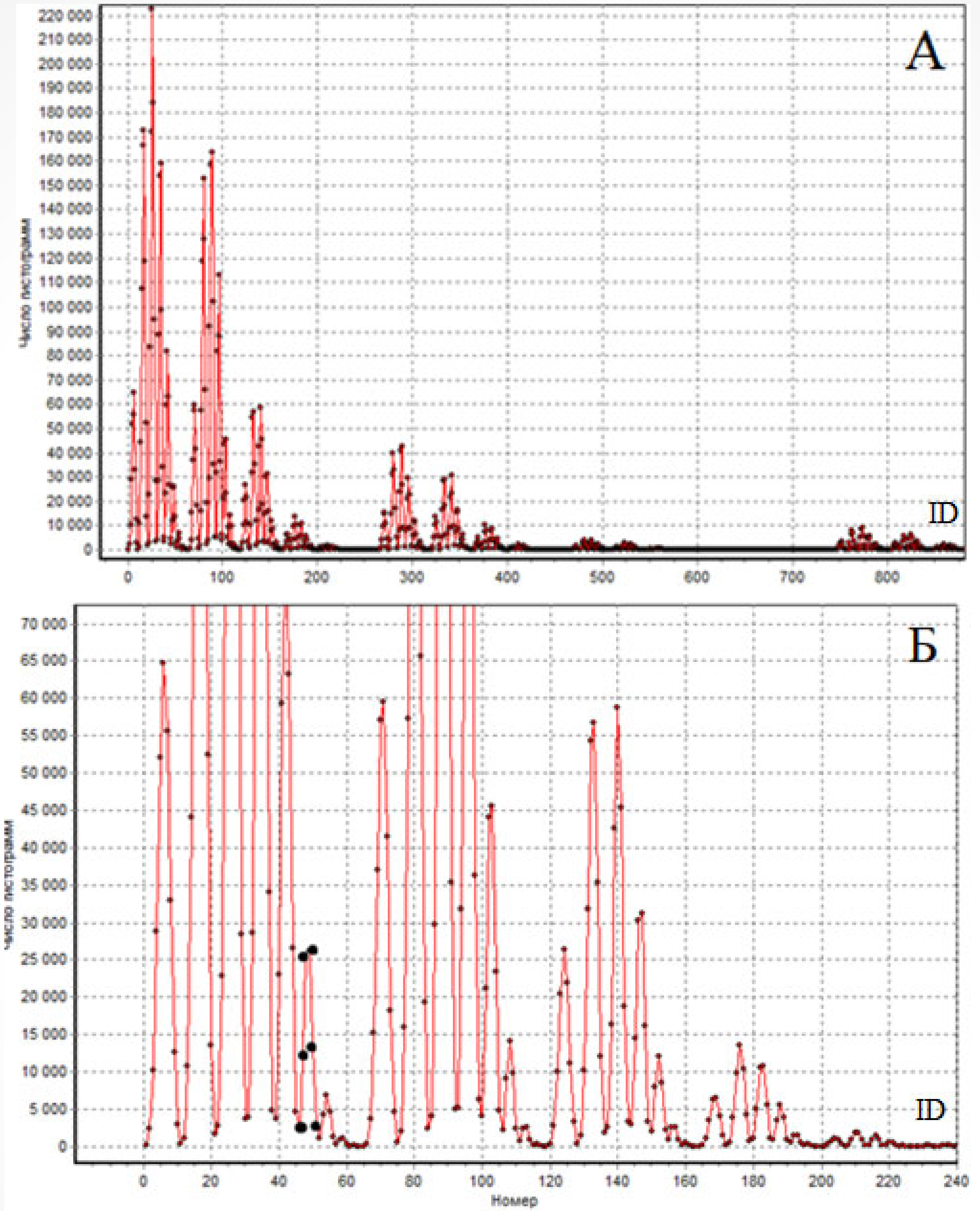


Рис. 2. А – Распределение - спектр СВР; Б – векторы [505-055](черные точки).

С увеличением порядка $r = 0, 1, 2$ , ...целых эмпирических моментов распределения СВ $k$, $\mu = \frac{1}{n-1}\sum_{j=0}^{n}(k_j - \tilde{k})^r$ значительно возрастает их разброс. Так как $\mu(r = 0) = 1, \mu(r = 1) = 0$, то в интервале порядка $r = 0, 1, \dots, 5$. Практически значимыми являются только эмпирические центральные моменты с $r = 2, 3, 4$,которые флуктуируют при $n \leq 10$ не менее чем на 20 – 30%. Поэтому при анализе эмпирических распределений и СВР, ограничиваются моментами не более четвертого порядка, что не позволяет при $n \leq 10$ определять $F(k|\theta)$ и проводить оценку однородности СВР особенно при значительных количества векторов M>>1.

Каждому СВР $\nu(.) = \nu_0(k_0 = 0); \nu_1(k_1 = 1)$ $\sum_{j=0}^{l} \nu_j = n$

однозначно соответствует реализация функции дробного $\mu(S \mid \bar{k})$ центрального комплексного момента порядка $S = r + \alpha$:

$$\bar{\mu}(S \mid \bar{k}) = \frac{1}{n}\sum_{j=1}^{n}(k_j - \tilde{k})^s = \mathrm{Re}(\mu(S \mid \bar{k})) + i\,\mathrm{Im}(\mu(S \mid \bar{k})) = P(\mu(S \mid \bar{k}))(\cos(\varphi(S)) + i\sin(\varphi(S))).$$

Проекция мнимой составляющей момента $\mu(S \mid \bar{k})$ на поверхность Im(μ(.)),S имеет вид

$$\mathrm{Im}(\mu(S \mid \tilde{k})) = \mathrm{Im}(\mu(S_0 \mid \tilde{k})e^{-\gamma S_0} \cdot e^{\gamma S} \cdot (-1)^r \sin(\pi\alpha))$$

где $\gamma = (\ln 10)\ln \bar{k}$

Фактор $Im(\mu(S_0|\tilde{k}))$ однозначно определен компонентами $\nu_i(.)$ СВР $\nu(.)$.

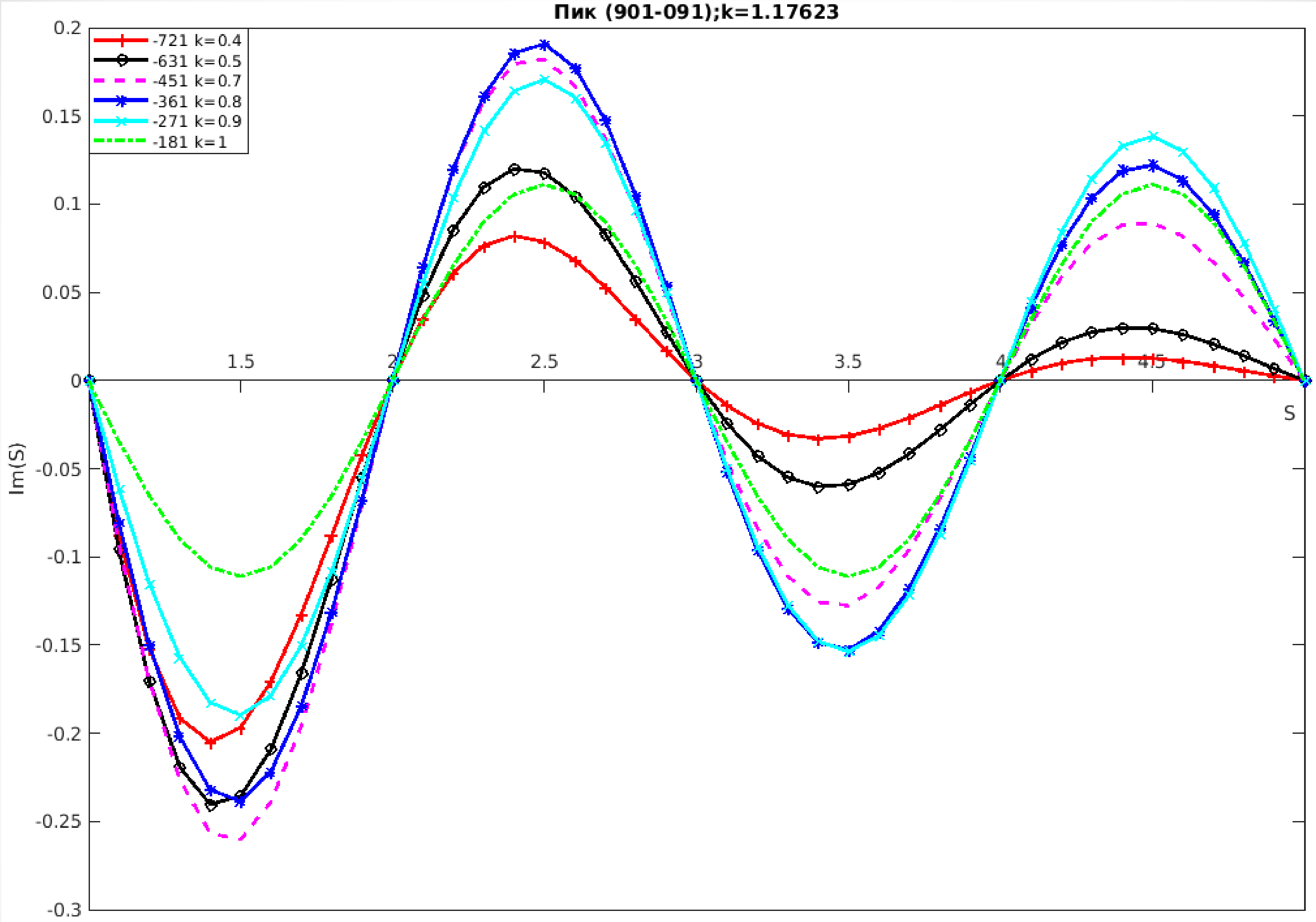


Рис. 3. Распределение мнимой составляющей Im(S) пика векторов 901-091 генеральной выборки $\bar{k} = 1{,}17623$.

$$\mathrm{Im} = \mathrm{Im}_0 \cdot 10^{\gamma \cdot \Delta S} \cdot \sin(2\pi \cdot \Delta S)$$

Для выборочного среднего значения вектора $\bar{k}(\nu) = \frac{1}{n}\sum_{j=0}^{l} \nu_j k_j;\ k_j = 0, 1, 2 \ldots$ справедливо $\tilde{k} < 1, \gamma < 0;\ \tilde{k} > 1, \gamma > 0;\ \tilde{k} \cong 1, \gamma = 0.$ При $\bar{k} < 1$ , значимая часть распределения $P_0(k)\left|k \div \bar{k}\right| < 1$ и центральные моменты $\mu\left(k - \bar{k}|S\right)$ с ростом порядка $S$ убывают. Периодические экстремумы $Im\left(\mu(s)\right)$ при $\mathrm{S} = r + 0.5$ позволяют оценить: соответствие СВР $\nu(.) = (\nu_0, \ldots, \nu_l)$имеющего полимиальную условную вероятность реализации и однородность совокупности векторов $\{\ \nu(.)\ \}$ образующих по комбинаторике компонент $\nu_j$, например $\nu_0, \nu_1$ подмножество в виде пика в эмпирическом распределении

$$M(ID(.)) = M(\nu_0 a_0 + \ldots + \nu_l a_l)$$

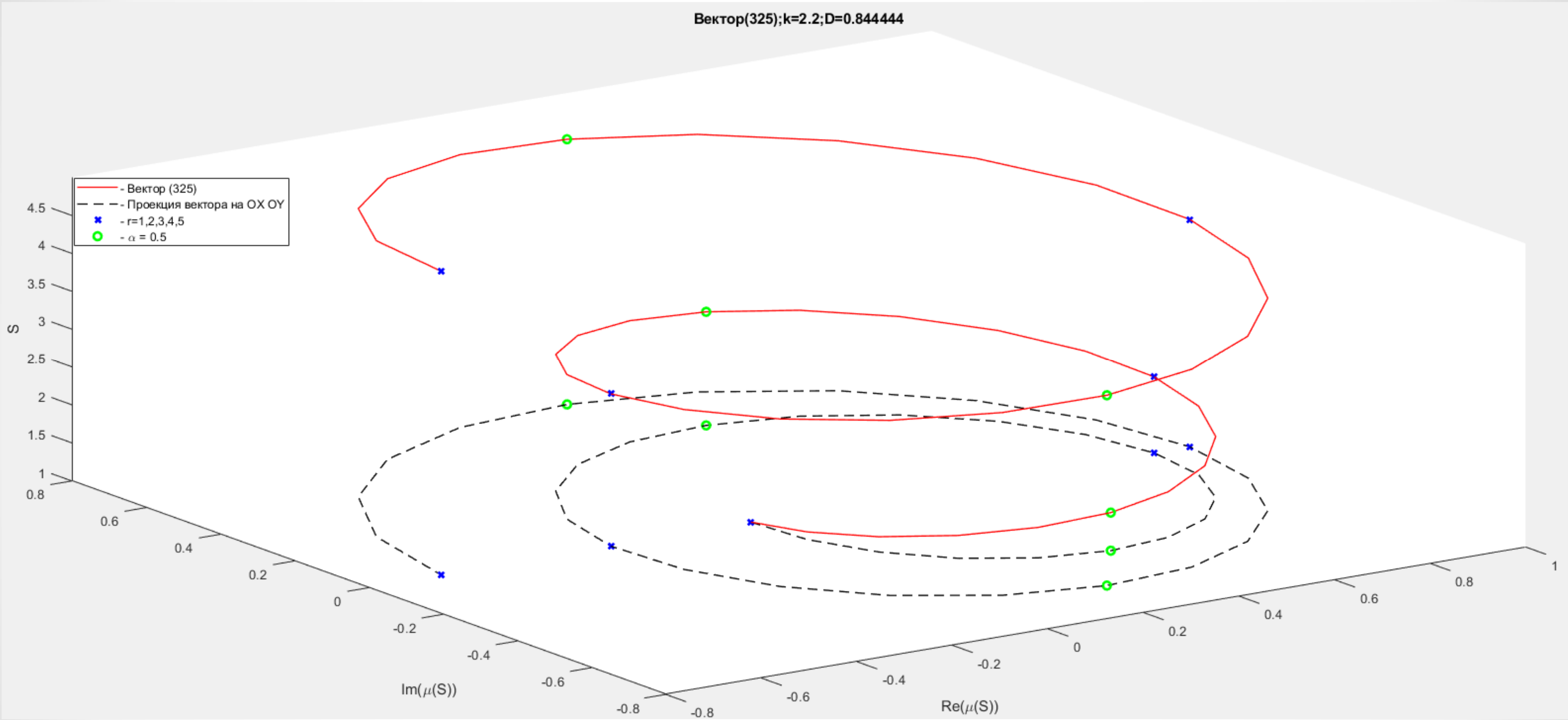


Рис. 5. Фазовая траектория вектора (325) в координатах Re(μ(S)), Im(μ(S)), S и проекция на плоскость Re(μ(S)), Im(μ(S)).

Предложенный метод учитывает кроме мнимой и реальную составляющую $\mathrm{Re}(\mu(\nu, S))$ функции моментов $\mu(\nu, S)$. В частности для СВР, образующих один из пиков в распределениями $M(ID(\nu(.)))$ при $\overline{k} = 1.176...$ ; $n = 10$ вероятности их реализации $P(\nu(.))$ и значения метрики $\rho(.)$ составляют при $S_0 = 1,\ S_{\max} = 4.9$

$\rho(m = 3, m)$ $\rho(m = 4, m)$

$$\rho = \frac{1}{N \cdot \Delta S}\left[\sum\left(\frac{\mathrm{Im}_1 - \mathrm{Im}_2}{\mathrm{Im}_1 + \mathrm{Im}_2}\right)^2 + \left(\frac{\mathrm{Re}_1 - \mathrm{Re}_2}{\mathrm{Re}_1 + \mathrm{Re}_2}\right)^2\right]^{1/2}$$

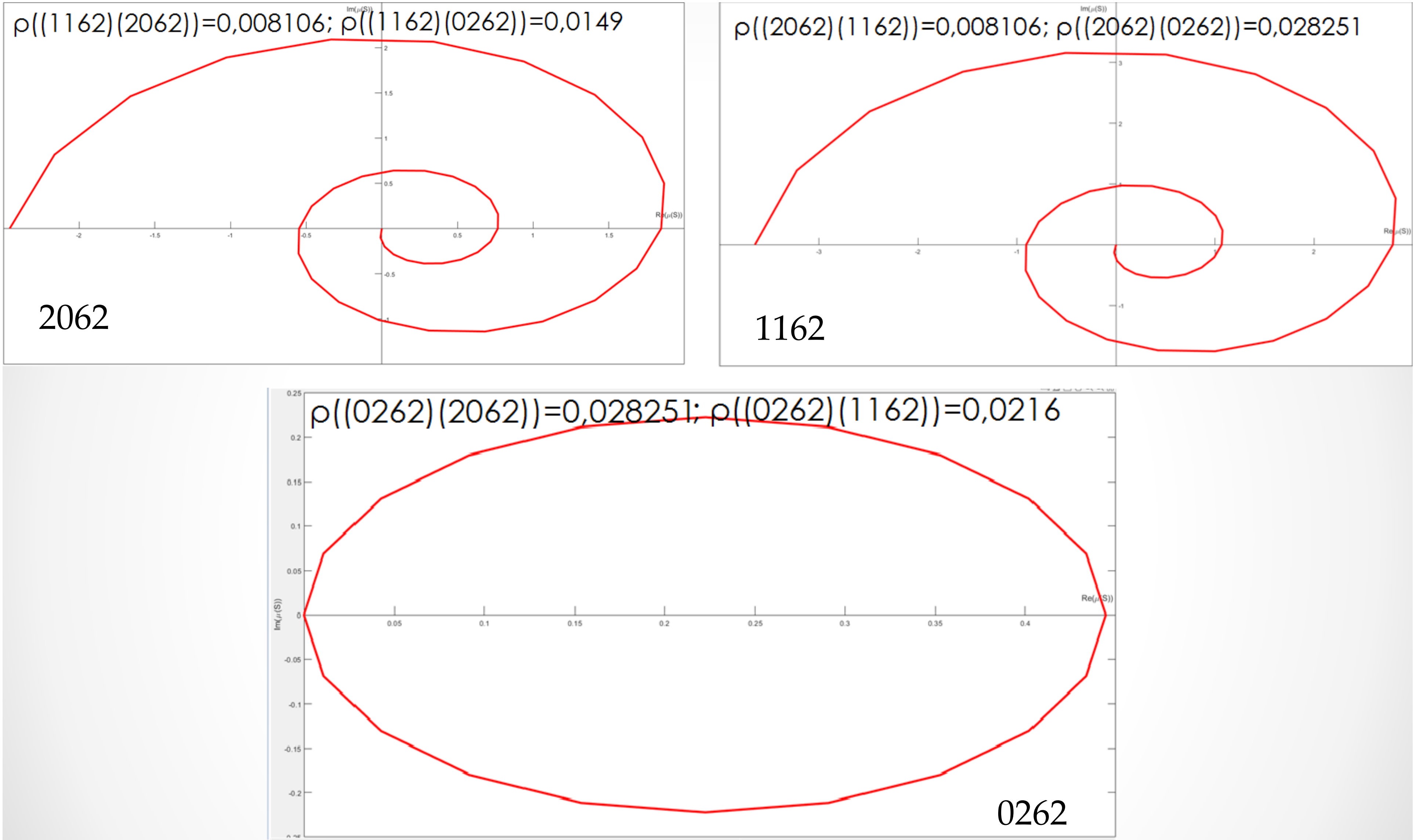


Рис. 6. Графики пика [2062-0262] Re(S), Im(S):

(2062): $\bar{k}$=1.8; $\bar{D}$=1.06;

(1162): $\bar{k}$=1.9; $\bar{D}$=0.76;

(0262): $\bar{k}$=2 ; $\bar{D}$=0.44.

| | 7021 | 6121 | 5221 | 4321 | 3421 | 2521 | 1621 | 0721 |
|---|---|---|---|---|---|---|---|---|
| 7021 | 0 | 0,003529 | 0,007207 | 0,010903 | 0,014423 | 0,017669 | 0,020647 | 0,023659 |
| 6121 | 0,00353 | 0 | 0,003772 | 0,007662 | 0,011449 | 0,014991 | 0,01827 | 0,021684 |
| 5221 | 0,00721 | 0,003772 | 0 | 0,004007 | 0,008013 | 0,011839 | 0,015453 | 0,019395 |
| 4321 | 0,01903 | 0,007662 | 0,004007 | 0 | 0,004138 | 0,008208 | 0,012196 | 0,016852 |
| 3421 | 0,01442 | 0,011449 | 0,008013 | 0,004138 | 0 | 0,004232 | 0,008627 | 0,014201 |
| 2521 | 0,01767 | 0,014991 | 0,011839 | 0,008208 | 0,004232 | 0 | 0,004737 | 0,011307 |
| 1621 | 0,02065 | 0,01827 | 0,015453 | 0,012196 | 0,008627 | 0,004737 | 0 | 0,007323 |
| 0721 | 0,02366 | 0,021684 | 0,019395 | 0,016852 | 0,014201 | 0,011307 | 0,007323 | 0 |

Таблица. 1. Матрица метрики ρ для пика векторов 7021-0721,

где $\rho = \frac{1}{N \cdot \Delta S}\left[\sum\left(\frac{\mathrm{Im}_1 - \mathrm{Im}_2}{\mathrm{Im}_1 + \mathrm{Im}_2}\right)^2 + \left(\frac{\mathrm{Re}_1 - \mathrm{Re}_2}{\mathrm{Re}_1 + \mathrm{Re}_2}\right)^2\right]^{1/2}$

# Заключение

Предложены и апробированы:

-методы анализа однородности флуктуации последовательностей из М выборок отсчетов зарегистрированных частиц К и соответствующих им случайных векторов СВР распределений $\nu(.)$ малого объёма $n \le 10$ на основе разработанной методики получения распределений идентификаторов $ID(\nu(.), a)$ векторов с различными проектирующими векторами $a$ , позволяющей группировать в пиках $M(ID(\nu(.)))$ однородные СВР с идентичными компонентами $\nu_j$. Метод особенно эффективен при больших последовательностях СВР $M >> 1$, когда существующие критерии и характеризации неприменимы;

-метод оценки однородности пар и групп выборок К и СВР при малой статистике в пиках распределений СВР, когда $M_j(\nu(.)) \le 10$ (в условиях: неполный пик) основанный на метрике расстояний $\rho(\mu((.)_j, \mu(.)_m) \ge 0$ между проекциями фазовых траекторий функций порядка S центральных дробных комплексных моментов СВР

$$\mu(\nu_j(.)S) = \mathrm{Re}(\mu(.)) + i \cdot \mathrm{Im}(\mu(.))$$

Совокупности однородных по заданным компонентам $\nu_j$ векторов с $\rho(.) < \rho_{крит}$ ,(например образующих пики в распределениях $ID(\nu(.), a)$ ) можно объединяя их в совокупности М>>1 обрабатывать и анализировать с помощью известных методов статистического анализа.

# Спасибо за внимание!