

МЕТОД ОЦЕНКИ ЧАСТНО-ВРЕМЕННОЙ ПЛОТНОСТИ ВЕРОЯТНОСТИ ЦИФРОВОГО СИГНАЛА С ИСПОЛЬЗОВАНИЕМ ЛИНЕЙНОЙ ИНТЕРПОЛЯЦИИ

DOI: 10.36724/2072-8735-2024-18-7-4-12

Manuscript received 14 June 2024;
Accepted 07 July 2024

Шевгунов Тимофей Яковлевич,
Московский авиационный институт (национальный
исследовательский университет), Москва, Россия,
shevgunov@gmail.com

Ключевые слова: циклоstationарность, частно-
временной анализ, частно-временная плотность,
линейная интерполяция, оценка плотности
вероятности

В работе представлена разработка нового инструмента частно-временного (fraction-of-time) подхода, в рамках которого случайный процесс описывается с использованием функциональных моделей, синтезируемых по его единственной наблюдаемой реализации, без необходимости построения абстрактных вероятностных моделей в условиях отсутствия достоверной априорной информации о проявлении процесса свойства эргодичности. На основе полученной ранее аналитической формулы, выражающей частно-временную плотность непрерывного сигнала в явной форме, был предложен метод оценки частно-временной плотности цифрового сигнала. В качестве аппроксимирующей модели, используемой для восстановления сигнала непрерывного времени из отсчётов цифрового, была выбрана кусочно-непрерывная линейная функция. Применение интерполяции моделью первого порядка позволило разработать быстродействующий алгоритм цифровой обработки сигналов, формирующий аналитическое выражение для оценки частно-временной плотности наблюдаемого сигнала. Показано, что синтезируемое выражение будет состоять из двух частей - регулярной, представленной суммой прямоугольных импульсов, и сингулярной, представленной суммой дельта-функций Дирака. В качестве демонстрации работы разработанного алгоритма оценивания представлены результаты моделирования для трёх случаев, включающих периодический сигнал, стационарный случайный процесс с известной корреляционной функцией и циклоstationарный процесс второго порядка. Для наиболее сложного случая циклоstationарного процесса предложенный алгоритм продемонстрировал выигрыш в оценке плотности вероятности по сравнению с известным методом ядерной оценки плотности вероятности при выборе гауссовской функции сглаживания. В качестве ожидаемого направления практического применения разработанных в рамках проведенного исследования метода и алгоритма на его основе рассматривается повышение качества решения задач анализа данных, для которых точность оценок плотности вероятности имеет существенное значение, например, задач обнаружения случайных сигналов с неизвестными параметрами.

Информация об авторе:

Шевгунов Тимофей Яковлевич, к.т.н., доцент кафедры "Теоретическая радиотехника", Московский авиационный институт (национальный исследовательский университет) "МАИ", Москва, Россия

Для цитирования:

Шевгунов Т.Я. Метод оценки частно-временной плотности вероятности цифрового сигнала с использованием линейной интерполяции // T-Comm: Телекоммуникации и транспорт. 2024. Том 18. №7. С. 4-12.

For citation:

Shevgunov T.Ya. (2024). A technique for the estimation of fraction-of-time probability density of a digital signal using linear interpolation. T-Comm, vol. 18, no.7, pp. 4-12. (in Russian)

Введение

В течение нескольких последних лет в научной литературе отмечается возрастание исследовательского интереса к частно-временному подходу (fraction-of-time, FOT) для анализа сигналов в задачах цифровой обработки сигналов [1]. Ключевая идея данного подхода, первые шаги по формализации которого были выполнены в конце 80-х годов [2], состоит в отказе от построения абстрактного вероятностного пространства при теоретическом синтезе моделей случайных процессов (СП) и замене этой аналитической процедуры построением функционального пространства, порождаемого сигналом как единственно наблюдаемой, бесконечно протяжённой во времени реализацией СП [3, 4].

При использовании частно-временного подхода в качестве эквивалента вероятностного закона распределения для сигнала непрерывного времени, наблюдаемого на интервале $[0, T]$, вводится частно-временное распределение (ЧВР):

$$F_{x[0;T]}(\xi) \triangleq \frac{1}{T} \int_0^T I(x(t) \leq \xi) dt = \frac{1}{T} \int_0^T u(\xi - x(t)) dt, \quad (1)$$

где $I(\bullet)$ – индикаторная функция, а $u(t)$ – функция Хэвисайда.

Частно-временная плотность (ЧВП), эквивалент плотности вероятности, определяется формально как производная ЧВР (1) в классическом или обобщённом смысле по его аргументу – переменной ξ :

$$f_{x[0;T]}(\xi) = \frac{dF_{x[0;T]}(\xi)}{d\xi} = \frac{1}{T} \int_0^T \delta(\xi - x(t)) dt, \quad (2)$$

в записи которой подынтегральным выражением является одномерная дельта-функция Дирака, зависящая от двух переменных: t и ξ . Справедливость такого перехода обоснована тем, что переменные t и ξ рассматриваются как независимые.

В работах [5, 6] была впервые предложена явная формула представления ЧВП для сигнала непрерывного времени $x(t)$, наблюдаемого на интервале $[0, T]$:

$$f_{x[0;T]}(\xi) = \frac{1}{T} \sum_m \frac{1}{|x'[t_{m[0;T]}(\xi)]|} + \frac{1}{T} \sum_k \mu[Ru_{x[0;T]}(x_{k[0;T]})] \delta(\xi - x_{k[0;T]}), \quad (3)$$

где $t_{m[0;T]}(\xi)$ – изолированные корни уравнения $x(t) = \xi$, принадлежащие интервалу $[0, T]$, $x'(t)$ – производная сигнала по времени, $x_{k[0;T]}$ – принимаемые сигналом значения, для которых множества неизоллированных $Ru_{x[0;T]}(x_{k[0;T]})$ суть непустые, $\mu(\bullet)$ – мера Лебега множества.

Практический интерес представляет получение оценок ЧВП на основе обработки цифрового сигнала $x[n]$, составленного из конечного числа наблюдений сигнала $x(t)$. Дискретизация $x(t)$ в самом общем случае состоит в получении отсчётов цифрового сигнала, наблюдаемого на интервале времени $[0, T]$. Конечная последовательность моментов времени

наблюдения $\mathbf{t} = \{t_n\}$ состоит из $N+1$ элемента, строго возрастает $t_n < t_{n+1}$ и принимает границы интервала в крайних значениях: $t_0 = 0$, $t_N = T$. Наблюдение сигнала в каждой точке последовательности $x_n = x(t_n)$ формирует последовательность $N+1$ вещественных чисел $\mathbf{x} = \{x_n\}$. В наиболее простом случае равномерной дискретизации: $t_n = nT_s$, где период дискретизации $T_s = T/N$. Однако, как будет показано в статье, требования к равномерности дискретизации не являются обязательными для предлагаемого метода оценивания ЧВП. Тем не менее, выбор постоянного шага дискретизации T_s позволит упростить аналитические выражения, и реализация алгоритма цифровой обработки будет выполнена в предположении о равномерной дискретизации.

Задача, решение которой представлено в настоящей статье, состоит в разработке пригодного для практической реализации метода оценивания ЧВП на основе отсчётов цифрового сигнала; в общей постановке она состоит в получении функции $\hat{f}_{x[0;T]}(\xi) = \hat{f}_{x[0;T]}(\xi; \mathbf{t}, \mathbf{x})$, которая в некотором формализуемом смысле хорошо приближает заданную аналитическим выражением (3) ЧВП $f_{x[0;T]}(\xi)$, полагаемую истинной для сигнала $x(t)$. При этом ожидается, что функция $\hat{f}_{x[0;T]}(\xi; \mathbf{t}, \mathbf{x})$ будет обладать свойствами, желательными для статистик (функций выборки), например, несмещённостью и состоятельностью.

1. Метод оценки частно-временной плотности

В основе предлагаемого метода оценки ЧВП лежит её аппроксимация с помощью функции $\hat{f}_{x[0;T]}(\xi)$, для которой оказывается сравнительно простым получить представление в форме конечных функциональных рядов. Выбор типа функций для слагаемых, формирующих функциональный ряд, возникает компромисс. С одной стороны, необходимо обеспечить высокую точность аппроксимации. С другой стороны, необходимо ограничить аналитическую и вычислительную сложность членов функционального ряда для того, чтобы сохранить возможность их представления посредством компактных, по возможности пригодных для как интуитивной интерпретации формальных выражений, так и построения алгоритмов, которые могут быть реализованы в современных вычислительных системах и устройствах цифровой обработки сигналов.

1.1. Разбиение интервала наблюдения

Рассмотрим множество $A \subset \mathbb{R}$, представляющее собой интервал времени наблюдения сигнала $x(t)$: $t \in A \subset \mathbb{R}$. Разбиением $\mathcal{B}$ множества A конечной меры Лебега: $0 < \mu(A) < \infty$, будем называть счётное семейство $\mathcal{B}$ (множество, элементами которого являются множества), сформированное из множеств $B_n \subseteq A \subset \mathbb{R}$ следующим образом:

$$\mathcal{B} = \left\{ B_n : \bigcup_n B_n = A \wedge (\forall n \neq m, \mu(B_n \cap B_m) = 0) \right\}. \quad (4)$$

В качестве величины, количественно характеризующей отношение между множеством A и его подмножеством B_n , будем использовать *условную относительную меру*, определяемую как отношение мер Лебега [7]:

$$\mu(B_n|A) = \frac{\mu(B_n \cap A)}{\mu(A)}, \quad (5)$$

где запись $\mu(B_n|A)$ также может быть кратко прочитана как «относительная мера B_n в A ».

Непосредственно из построения семейства $\mathcal{B}$ (4) следует, что условная относительная мера (5) может быть эквивалентно записана как отношение мер двух множеств:

$$\mu(B_n|A) = \frac{\mu(B_n)}{\mu(A)}, \quad (6)$$

а также следует свойство нормировки:

$$\sum_n \mu(B_n|A) = 1. \quad (7)$$

После построения разбиения $\mathcal{B}$ для множества A , ЧВП $f_{x|A}(\xi)$ сигнала $x(t)$, наблюдаемого в пределах заданного множеством A интервала времени, может быть найдена как:

$$f_{x|A}(\xi) = \sum_n \mu(B_n|A) f_{x|B_n}(\xi), \quad (8)$$

где условная ЧВП $f_{x|B_n}(\xi)$ определена для любого множества ненулевой меры B_n из (4) ($\mu(B_n) \neq 0$), в соответствии с формулой (3), если интервал $[0, T]$ заменить в ней на B_n .

Формула (8) допускает достаточно ясную интуитивную интерпретацию. В силу того, условная мера B_n в A есть число неотрицательное, то каждое слагаемое в правой части (8) внесёт парциальный вклад в полную ЧВП $f_{x|A}(\xi)$. В том случае, когда $\mu(B_n) = 0$, такой вклад обнуляется, и фактическая форма ЧВП $f_{x|B_n}(\xi)$ не имеет смысла.

Следует заметить, что выражение (8) представляет собой частно-временной аналог *формулы полной вероятности*, формулируемой в рамках классической теории вероятностей [8]. Так, используемые в (8) множества B_n , играют роль, эквивалентную несовместным, или взаимно-исключающим, случайным событиям в формуле полной вероятности [9].

1.2. Оценка парциальной ЧВП

В соответствии с определением (4) сформируем разбиение $\mathcal{B}$ полного интервала наблюдения $[0, T]$ на N интервалов B_n , границы которых определены парами последовательных элементов, взятых из множества отсчётов времени $\mathbf{t}$.

$$B_n = \begin{cases} [t_{n-1}, t_n), & 1 \leq n \leq N-1; \\ [t_{n-1}, t_n], & n = N. \end{cases} \quad (9)$$

В пределах n -го интервала B_n построим аппроксимацию $\hat{x}(t; n)$ сигнала $x(t)$ с помощью линейной модели, такой что её график в виде прямой линии будет проходить через две крайние точки: (t_{n-1}, x_{n-1}) и (t_n, x_n) .

Формальное выражение для функции, описывающей аппроксимирующий сигнал в пределах n -го интервала B_n , можно задать с использованием следующего уравнения:

$$\hat{x}(t; n) = b(n) + k(n)t, \quad t_{n-1} \leq t \leq t_n, \quad (10)$$

где коэффициенты $b(n)$ и $k(n)$ определяются в соответствии с формулами:

$$b(n) = \frac{x_n t_{n-1} - x_{n-1} t_n}{t_n - t_{n-1}}, \quad k(n) = \frac{x_n - x_{n-1}}{t_n - t_{n-1}}. \quad (11)$$

Объединение линейных аппроксимаций отдельных интервалов B_n (10) позволяет составить непрерывную кусочно-линейную аппроксимацию $\hat{x}(t)$ для сигнала $x(t)$ на всём интервале наблюдения $[0, T]$. Формальная запись $\hat{x}(t)$ может быть выражена суммой

$$\begin{aligned} \hat{x}(t) &= \sum_{n=1}^N \hat{x}(t; n) [u(t_n) - u(t_{n-1})] = \\ &= \sum_{n=1}^N \hat{x}(t; n) \operatorname{rect}\left(\frac{t - t_{n-1}}{t_n - t_{n-1}} - \frac{1}{2}\right), \quad 0 \leq t \leq T, \end{aligned} \quad (12)$$

где второй сомножитель каждого слагаемого введён в качестве интервального ограничителя: выражение отлично от нуля только в пределах интервала $[t_{n-1}, t_n]$. За пределами интервала наблюдения $[0, T]$ функция $\hat{x}(t)$ считается неопределённой.

Выбор линейной функции, аппроксимирующей сигнал на интервале B_n даёт возможность непосредственной подстановки (10) в (3). После аналитических преобразований в символической форме получается формальная запись, задающая парциальную ЧВП на интервале B_n :

$$\begin{aligned} \hat{f}_{x|B_n}(\xi) &= \hat{f}_x(\xi; n) = \\ &= \begin{cases} \frac{1}{|x_n - x_{n-1}|} \operatorname{rect}\left(\frac{\xi - 0.5(x_n + x_{n-1})}{|x_n - x_{n-1}|}\right), & x_n \neq x_{n-1}; \\ \delta(\xi - x_n), & x_n = x_{n-1}. \end{cases} \end{aligned} \quad (13)$$

Детальный анализ условий, используемых в аналитической записи ЧВП (13), позволяет отметить два различных случая в зависимости от коэффициента перед переменной t в линейной модели (10): один из этих случаев соответствует ненулевому коэффициенту, а другой – нулевому. Рассмотрим каждый из этих случаев подробнее и проиллюстрируем их графически.

График для случая, при котором выполняется первое условие в (13), представляет собой наклонную прямую линию, наклон которой без потери общности примем положительным. На рисунке 1 изображён комбинированный график, на котором ось x графика сигнала $x(t)$ совмещена с осью ξ графика оценки парциальной ЧВП $\hat{f}_x(\xi; n)$. При этом график оценки парциальной ЧВП повернут на 90 градусов против часовой стрелки для возможности такого совмещения. Площадь прямоугольного импульса, описывающего оценку ЧВП, должна быть равна единице, как достоверного события, что позволяет определить высоту прямоугольного импульса при

известной протяжённости его носителя, ограниченной значениями x_{n-1} и x_n – принимаемыми сигналом на границах интервала B_n .

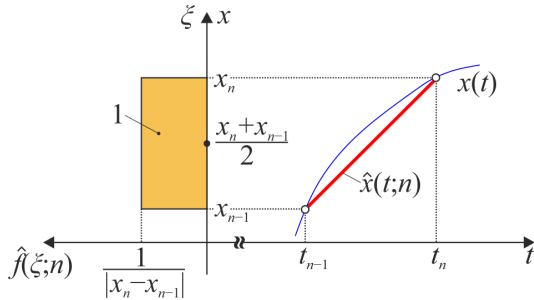


Рис. 1. Формирование оценки ЧВП на интервале с аппроксимирующей моделью в виде наклонной линии

Второе условие в (13) задаёт случай аппроксимирующей модели (10), заданной в пределах n -го интервала константой $\hat{x}(t; n) = x_{n-1} = x_n$, что проиллюстрировано горизонтальной прямой на графике рисунка 2. Постоянное значение x_n , занимающее весь интервал B_n , порождает дельта-функцию с единичной площадью. Здесь следует отдельно заметить, что кривые сигнала непрерывного времени $x(t)$ на рисунках 1 и 2 представлены исключительно в целях наглядности, поскольку после дискретизации достоверная информация о поведении сигнала между его отсчётами может быть потеряна.

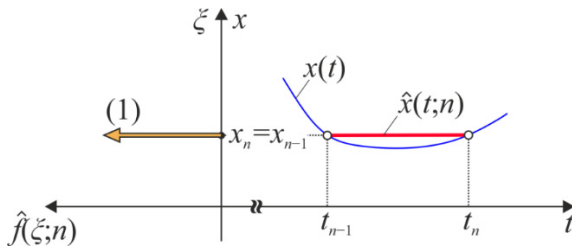


Рис. 2. Формирование оценки ЧВП на интервале с аппроксимирующей моделью в виде константы

1.3. Формирование оценки ЧВП

Необходимая для использования формулы (8) условная относительная мера B_n (9) в интервале $[0, T]$ определяется в соответствии с (6) как отношение длин временных интервалов:

$$\mu(B_n \parallel [0, T]) = \frac{t_n - t_{n-1}}{t_N - t_0} \quad (14)$$

что позволяет с использованием (8) записать выражение оценки ЧВП сигнала, наблюдаемого на всём интервале $[0, T]$, в виде суммы

$$\begin{aligned} \hat{f}_x(\xi) &= \sum_{n=1}^N \mu(B_n \parallel [0, T]) \hat{f}_x(\xi; n) = \\ &= \sum_{n=1}^N \frac{t_n - t_{n-1}}{t_N - t_0} \hat{f}_x(\xi; n). \end{aligned} \quad (15)$$

Здесь необходимо привести важное замечание о структуре получаемой оценки ЧВП. Функция $\hat{f}_x(\xi)$, как сумма (15) слагаемых вида (13), в общем случае может включать в себя

как регулярную, так и сингулярную часть. Можно показать, что если все последовательные пары (x_{n-1}, x_n) значений отсчётов сигнала различны: $x_{n-1} \neq x_n$, то оценка ЧВП будет состоять только из регулярной части. Если хотя бы в одной паре (x_{n-1}, x_n) элементы окажутся одинаковыми, в оценку ЧВП $\hat{f}_x(\xi)$ будет включена дельта-функция, расположенная в точке с координатой, соответствующем такому значению x_n .

С другой стороны, функция $\hat{f}_x(\xi)$ будет состоять только из сингулярной части, притом из единственной дельта-функции $\delta(\xi - x_0)$ с единичным весом, тогда и только тогда, когда все наблюдаемые значения x_n равны одному и тому же числу: $\forall n \in \mathbb{N}, n \leq N : x_n = x_0$. В противном случае $\hat{f}_x(\xi)$ будет непременно включать в себя регулярную часть, которая будет порождена по крайней мере одной парой (x_{n-1}, x_n) из различных элементов.

2. Алгоритм оценки частотно-временной плотности

В большинстве случаев дискретизация сигнала выполняется с равномерным интервалом T_s . При этом условные относительные меры всех интервалов (14) становятся одинаковыми $\mu(B_n \parallel [0, T]) = 1/N$, а выражение (15) упрощается до:

$$\hat{f}_x(\xi) = \frac{1}{N} \sum_{n=1}^N \hat{f}_x(\xi; n), \quad (16)$$

и вклады парциальных ЧВП (13) каждого из интервалов B_n становятся одинаковыми. При этом достаточными для оценки входными данными будет последовательность из N отсчётов цифрового сигнала $x = \{x_n\}$. Ожидаемым выходным результатом работы алгоритма будет функция $\hat{f}_x(\xi)$ (16), синтезируемая в форме функционального модуля или специального объекта языка программирования высокого уровня. В практическом случае выходными значениями алгоритма ЦОС, генерируемого такой программной функцией, могут служить:

1. вектор узлов $\xi = \{\xi_i\}$;

2. вектор значений оценок ЧВП $\mathbf{f} = \{\hat{f}_x(\xi_i)\}$ в узлах

сетки;

3. положения дельта-функций $\mathbf{DX} = \{x_m\}$ – вектор, содержащий значения, выделенные во втором условии (13);

4. веса дельта-функций $\mathbf{DW} = \{w_m\}$ – параллельный $\mathbf{DX}$ вектор, содержащий веса дельта-функций с учётом их вклада в (16).

Для генерации сетки узлов ξ требуется задать границы интервала наблюдения $[\xi_{min}, \xi_{max}]$ и величину шага дискретизации ξ_s . Выбор границ интервала может быть априорным, если имеется информация о носителе оценки ЧВП, например, на основе носителя плотности опорной теоретической модели. Иначе они могут быть определены на основе наблюдений сигнала.

В простейшем случае границы ξ_{min}, ξ_{max} совпадут соответственно с минимальным и максимальным отсчётами наблюдаемого сигнала.

Интервал дискретизации ξ_s определяет детализацию сетки значений ξ , в узлах которой

$$\xi_l = \xi_{\min} + l\xi_s, \quad 0 \leq l \leq L \quad (17)$$

проводится вычисление оценки ЧВП. Он может быть выбран априорно как $\xi_s = (\xi_{\max} - \xi_{\min}) / (L - 1)$, где L – желаемое количество узлов.

Дополнительно следует отметить, что выражение (13) предполагает проверку равенства двух последовательных значений x_n и x_{n-1} , которая может быть выполнена не вполне корректно для вещественных чисел, представленных в машинном формате конечной точности. В качестве возможного решения этой проблемы знак точного равенства может быть заменен на приближённое равенство $|x_n - x_{n-1}| < \varepsilon$, где величина ε определяет допуск и будет дополнительным параметром, выбираемым априорно, или пропорциональным константе машинной точности системы ЦОС, или примерно равным минимальному разрешению по амплитуде – типично, половине шага квантования АЦП.

3. Результаты

3.1. Гармонический сигнал

В качестве первого примера проведём оценку ЧВП гармонического сигнала, аналитическое выражение ЧВП которого было получено в [6]. При моделировании выберем следующие значения параметров: амплитуда $A_x = 4$ В, частота $\nu_x = 1$ МГц, $T_s = 1/(N\nu_x)$, где $N \in \mathbb{N}$ – число отсчётов, приходящихся на период. Выбрав $N = 12$, рассмотрим два характерных случая с начальными фазами равными $\varphi_1 = \pi/8$ и $\varphi_2 = \pi/12$, графики которых изображены соответственно на рисунках 3 а) и б). На каждом рисунке указаны отсчёты сигналов, используемые для оценки ЧВП, а также представлены графики линейной аппроксимирующей модели, построенной в соответствии с (12).

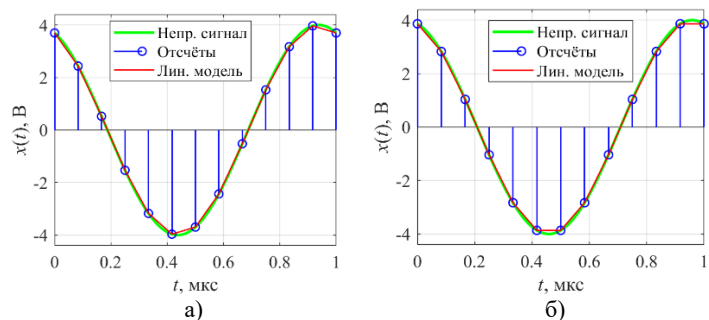


Рис. 3. Сигналы, отсчёты и их линейные аппроксимирующие модели для двух характерных случаев: а) без горизонтальных участков; б) при наличии горизонтальных участков

Для двух исследуемых случаев φ_1 и φ_2 на рисунках 4 а) и б) соответственно представлены оценки ЧВП, полученные с использованием разработанного алгоритма. Каждый график дополнен теоретической кривой ЧВП, построенной в соответствии с аналитической формулой из [6].

В графиках оценок ЧВП можно визуально проследить полигональную структуру, что объясняется относительно небольшим выбранным значением N . Следует также обратить

внимание на то, что оценка ЧВП на рисунке 4б содержит сингулярную часть, представленную двумя дельта-функциями. Эта часть была выделена из-за присутствия горизонтальных участков графика линейной аппроксимирующей модели, представленной на рисунке 3б.

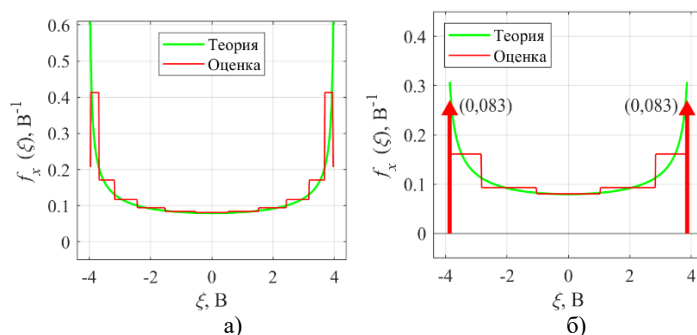


Рис. 4. Оценки частно-временной плотности для двух характерных случаев:

а) без сингулярной части; б) при наличии сингулярной части

На рисунках 5 а,б представлены ЧВР (1) – их графики всегда ограничены диапазоном $[0, 1]$ и не содержат сингулярных компонент. С учётом связи ЧВР и ЧВП (2), каждое ЧВР было получено численным интегрированием соответствующей ему ЧВП. Однако, на рисунке 5 б) необходимо обратить внимание на скачки величиной $1/12$, происходящие в точках, соответствующим положению дельта функций в ЧВП на рисунке 4б. Каждый график дополнен теоретической кривой ЧВР, построенной в соответствии с [6].

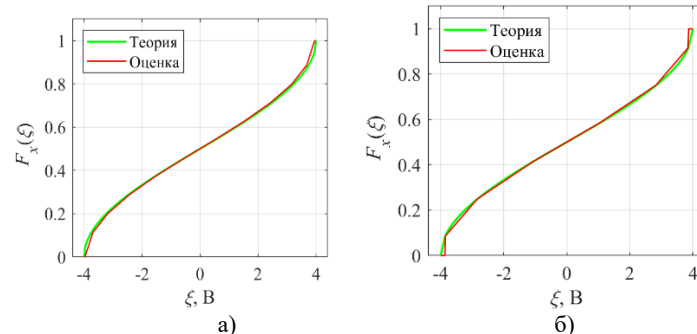


Рис. 5. Оценки частно-временного распределения для двух характерных случаев: а) без скачков, б) со скачками

Для того, чтобы количественно оценить различие в законах распределения случайных величин, предложено использовать две скалярные метрики – расстояние по норме L_1 [10] и статистику Колмогорова-Смирнова D_{KS} [11], формальные выражения которых представлены ниже

$$L_1(f, g) = \int_{\mathbb{R}} |f(\xi) - g(\xi)| d\xi, \quad (18)$$

$$D_{KS}(F, G) = \sup_{\xi \in \mathbb{R}} |F(\xi) - G(\xi)|, \quad (19)$$

Проведём анализ влияния числа отсчётов N , приходящихся на один период гармонического сигнала, на значение метрик L_1 и D_{KS} . Выберем интервал анализа $[\xi_{\min}, \xi_{\max}] = [-A, A]$, количество уровней разбиений доста-

точно большим $L = 10^5$, чтобы можно было пренебречь ошибками, вносимыми эффектом дискретизации по ξ . На рисунках 6 и 7 представлены графики зависимости метрик соответственно L_1 и D_{KS} , построенные в логарифмическом масштабе. Кривые демонстрируют убывающий характер, подтверждая этим состоятельность предложенного метода формирования оценки ЧВП. Действительно, при уменьшении периода дискретизации при сохранении времени наблюдения количество отсчётов возрастает, и линейная модель (12) всё лучше приближает исходный непрерывный сигнал.

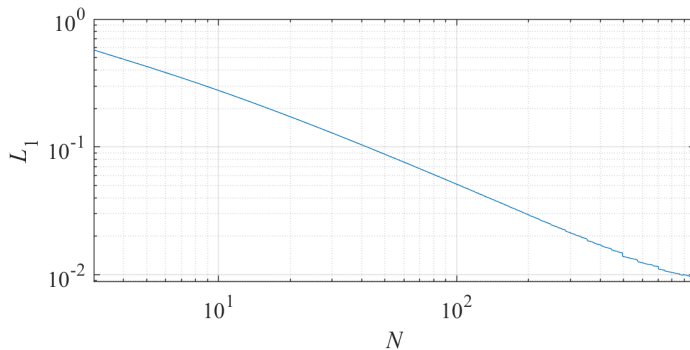


Рис. 6. Зависимость метрики L_1 от числа отсчётов, приходящихся на период

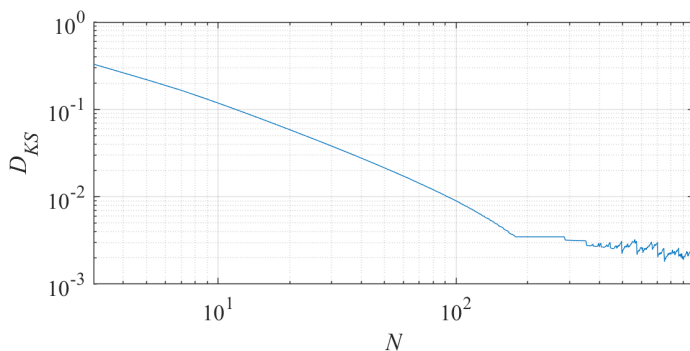


Рис. 7. Зависимость статистики Колмогорова-Смирнова от числа отсчётов, приходящихся на период

3.2. Стационарный случайный процесс

В качестве второго примера рассмотрим оценку ЧВП стационарного гауссовского случайного процесса $Y(t)$ с нулевым средним и известной одномерной корреляционной функцией:

$$R_y(\tau) = \mathcal{R}_y^{(0)}(\tau) = \sigma_y^2 \exp(-\alpha |\tau|), \quad (20)$$

где σ_y^2 – средняя мощность, а α – круговая частота пропускания формирующего ФНЧ 1-го порядка, на выходе которого наблюдается $y(t)$ при подаче на вход белого гауссовского шума.

На рисунке 8 представлены оценки ЧВП, полученные по реализациям разной длительности N , для значений параметров $T_s = 20$ нс, $\alpha = 11.2$ МГц, $\sigma_y^2 = 16$ В², числе уровней $L = 2000$, при ограничении интервала оценки $\pm 4\sigma_y$. Можно заметить, что в графиках оценок, найденным по реализациям короткой длительности, прослеживается полигональная структура оценщика (16), но по мере роста длительности реализации N оценки ЧВП визуально приближаются к теоретическому нормальному распределению.

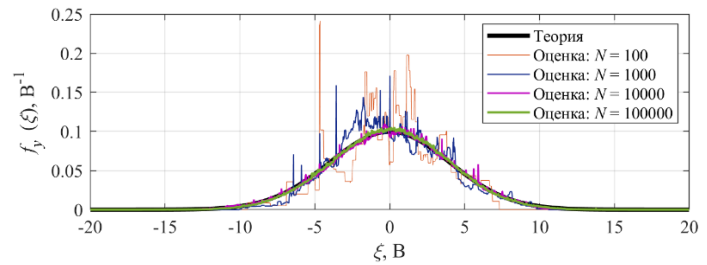


Рис. 8. Оценки ЧВП для гауссовского стационарного процесса

В силу того, что каждая реализация СП является последовательностью, выбираемой случайно, заключение о сходимости на основе визуального анализа не будет обоснованным. Чтобы убедиться в сходимости ЧВП к теоретической плотности вероятности СП, было проведено моделирование по методу Монте-Карло. Для длительности N сгенерируем $M = 100$ реализаций СП, по каждой из которых проведём оценку ЧВП с использованием предложенного алгоритма (16) и оценим интегральные меры различия: L_1 (18) и D_{KS} (19). Повторив это для нескольких значений длительности N , лежащих в диапазоне от 30 до 10^5 , сформируем параметрические выборки L_1 , D_{KS} , на основании которых можно сделать статистически обоснованный вывод.

На рисунке 9 в логарифмическом масштабе представлен график зависимости L_1 от длительности реализации N . Точки на графике соответствуют выборочному среднему, вокруг которого показан доверительный интервал в одно выборочное СКО. Непрерывная линия соединяет средние, давая представление о промежуточных значениях метрики.

На рисунке 10 представлен график зависимости метрики Колмогорова-Смирнова, построенный аналогичным образом

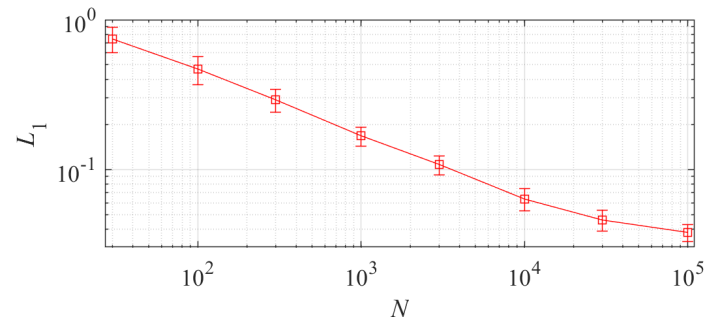


Рис. 9. Зависимость L_1 от длительности реализации СП

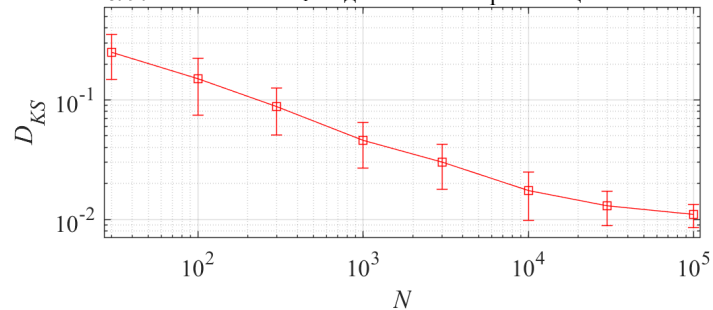


Рис. 10. Зависимость D_{KS} от длительности реализации СП

3.3 Циклостационарный случайный процесс

В качестве третьего примера рассмотрим циклостационарный СП, сформированный амплитудной модуляцией гармонического колебания $c(t)$ гауссовским стационарным СП:

$$Z(t) = Y(t)c(t) = Y(t)\cos(2\pi Ft + \varphi), \quad (21)$$

где $y(t)$ – СП, определяемой выражением (20), а F, φ – соответственно частота и начальная фаза колебания $c(t)$.

Для проведения сравнения оценка плотности вероятности (ПВ) СП также была получена с использованием метода ядерной оценки плотности вероятности (ЯОП) [12], также известного как метод Парзена-Розенблатта [13, 14]. При использовании этого метода оценка ПВ формируется на основе выборки СВ $\mathbf{z} = \{z_n\}$ конечного размера N как функция [15]:

$$\hat{f}_z^{(\text{ker})}(\xi) \triangleq \sum_{n=0}^{N-1} \frac{1}{Nb} W\left(\frac{\xi - z_n}{b}\right), \quad (22)$$

где $W(\psi)$ – функция сглаживающего ядра, которая была выбрана гауссовской:

$$W(\psi) = \frac{1}{\sqrt{2\pi}b} \exp\left(-\frac{\psi^2}{2b^2}\right), \quad (23)$$

где b – ширина окна, в общем случае зависящая от оценок размаха выборки и её размера N . Предложенное в работе [16] правило выбора ширины окна, также известное как правило Силвермана, предполагает значение

$$b = \left(\frac{4}{3N}\right)^{\frac{1}{5}} \hat{\sigma}_z, \quad (24)$$

где $\hat{\sigma}_z$ – выборочное СКО, полученное по выборке $\mathbf{z}$: квадратный корень из выборочной дисперсии. На практике возможно некоторое улучшение ПВ с помощью экспертного или эвристического подбора значения b , обычно за счёт небольшой его вариации в пределах от 0,5 до 2 раз от значения по правилу Силвермана (24).

При проведении моделирования были выбраны следующие значения параметров: $T_s = 20$ нс, $\alpha = 11.2$ МГц, $\sigma_y^2 = 4 \times 10^{-2} B^2$, $F = 1/(8T_s)$, $\varphi = 0$, интервал оценки ограничен $\pm 4\sigma_y$, число уровней $L = 2000$, длительность выборки $N = 2^{12} = 2048$ отсчётов, соответствующих времени наблюдения $T = 40,94$ мс.

На рисунке 11 представлены графики оценок ЧВП, полученных с помощью предложенного оценивателя (16) и метода ЯОП, с использованием которого были сформированы две оценки, соответствующие двум значениям ширины окна b : выбранному как (24) и выбранному эвристически в 1,3 раза больше. Для сравнения на графике показаны теоретическая кривая ЧВП, полученная численным интегрированием с использованием мультипликативной свёртки Мелина [5], а также кривая ПВ нормального распределения, построенная на основе выборочных оценок математического ожидания и дисперсии.

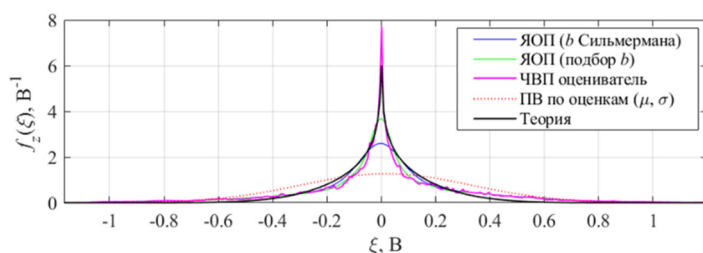


Рис. 11. Оценка ЧВП циклостационарного случайного процесса

Анализ графиков показывает, что отличительной особенностью оценки ЧВП, сформированной предлагаемым оценивателем, является то, что она хорошо приближается к теоретической кривой в области малых значений в окрестности нуля. В области больших значений кривые, полученные всеми использованными методами оценивания, ведут себя похожим образом, а именно: они формируют так называемые «хвосты» распределений, соответствующие гауссовскому закону, которые, в действительности, оказываются более «тяжелыми», чем те, которые имеет теоретическое распределение.

В таблице 1 представлены значения метрик сравнения ПВ: расстояния L_1 (18) и статистики Колмогорова-Смирнова D_{KS} (19). Полученные результаты показывают, что оценка ЧВП, сформированная с использованием предложенного метода оценивания, обладает меньшими интегральными мерами отличия от теоретической кривой по сравнению с оценками, сформированными методом ЯОП.

Таблица 1

Метрики различия законов распределения для оценок, полученных различными методами

Метод оценивания	Метрика	
	Статистика D_{KS}	Расстояние L_1
Предложенный метод	0.051	0.043
ЯОП (эвристическое значение b)	0.072	0.082
ЯОП (b по правилу Силвермана)	0.091	0.104

Заключение

В настоящем исследовании предложен способ оценивания частотно-временной плотности вероятности цифрового сигнала, основанный на восстановлении непрерывного сигнала из его отсчётов с использованием кусочно-непрерывной линейной интерполяции. В результате этого оказывается возможным сформировать оценку ЧВП в форме функционального ряда, членами которого будут функции, заданные на компактных носителях, а именно: прямоугольные импульсы и дельта-функции.

Одно из возможных объяснений выигрыша при анализе циклостационарного сигнала в точности оценок, получаемых с использованием предложенного метода (16), по сравнению с ЯОП (22) может состоять в следующем. Метод ЯОП формирует оценку как сумму функций одинаковой ширины, положение центров которых определяют элементы наблюдаемой выборки $\mathbf{z} = \{z_n\}$. При этом порядок следования элементов в выборке не оказывает какого-либо влияния на результирующую оценку ЧВП.

В предложенном новом методе слагаемые, входящие в сумму (16), представляют собой импульсы прямоугольной формы (13) переменной ширины. Так ширина n -го слагаемого в сумме (16) зависит от последовательной пары значений (z_{n-1}, z_n) выборки. Таким образом, предложенный метод можно считать, в некотором смысле, адаптивной версией метода ЯОП, в которой ширина окна подбирается с учётом порядка следования элементов в выборке $\mathbf{z}$.

Выигрыш, демонстрируемый предложенным методом, является экспериментальным подтверждением того, что при анализе цифровых сигналов, или временных рядов, порядок следования элементов в выборке имеет существенное значение.

Такие данные имеют структурное отличие от данных в неупорядоченных выборках. В результате обработки упорядоченных выборок следует ожидать получение более качественных оценок при использовании методов, подобных предложенному методу оценки ЧВП, т.е. таких, которые способны учитывать информацию о порядке следования отсчетов, составляющих цифровые сигналы.

Литература

1. Dehay D., Leškow J., Napolitano A., Shevgunov T. Cyclic Detectors in the Fraction-of-Time Probability Framework // *Inventions*. MDPI, 2023. Vol. 8, no. 6 – P. 152. DOI: 10.3390/inventions8060152.
2. Gardner W.A., Brown W.A. Fraction-of-time probability for time-series that exhibit cyclostationarity // *Signal Processing*. Vol. 23, no. 3. 1991, pp. 273-292. DOI: 10.1016/0165-1684(91)90005-4.
3. Napolitano A., Gardner W.A. Fraction-of-Time Probability: Advancing Beyond the Need for Stationarity and Ergodicity Assumptions // *IEEE Access*. Vol. 10. 2022, pp. 34591-34612. DOI: 10.1109/ACCESS.2022.3162620.
4. Gardner W.A. Transitioning away from stochastic process models // *Journal of Sound and Vibration*. Vol. 565. 2023, P. 117871. DOI: 10.1016/j.jsv.2023.117871.
5. Shevgunov T., Napolitano A. Fraction-of-Time Density Estimation Based on Linear Interpolation of Time Series // *2021 Systems of Signals Generating and Processing in the Field of on Board Communications*. Moscow, 2021. pp. 1-4, DOI: 10.1109/IEEECONF51389.2021.9415991.
6. Шевгунов Т.Я. Синтез явных аналитических формул вероятностных моделей сигналов при использовании частно-временного подхода // *T-Comm: Телекоммуникации и транспорт*. 2022. Том 16. №11. С. 21-29. DOI: 10.36724/2072-8735-2022-16-11-21-29.
7. Leškow J., Napolitano A. Foundations of the functional approach for signal analysis // *Signal Processing*. Vol. 86. 2006, pp. 3796-3825. DOI: 10.1016/j.sigpro.2006.03.028.
8. Taboga M. *Lectures on Probability Theory and Mathematical Statistics*, 3rd ed. USA, San Bernardino: CreateSpace, 2018. 670 p.
9. Pishro-Nik H. *Introduction to Probability, Statistics and Random Processes*. USA: Kappa Research, 2014. 744 p.
10. Deza M.M., Deza E. *Encyclopedia of Distances*, 4th ed. Springer, 2016. DOI: 10.1007/978-3-642-30958-8.
11. Stephens M.A. Tests of Fit for the Logistic Distribution Based on the Empirical Distribution Function // *Biometrika*. Vol. 66, no. 3. Dec. 1979, pp. 591-595. DOI: 10.1093/biomet/66.3.591.
12. Епанечников В.А. Непараметрическая оценка многомерной плотности вероятности // *Теория вероятностей и ее применения*. Том 14, № 1. 1969, С. 156-161.
13. Parzen E. On Estimation of a Probability Density Function and Mode // *The Annals of Mathematical Statistics*. Vol.33, no. 3. Sep. 1962, pp. 1065-1076. DOI: 10.1214/aoms/1177704472.
14. Rosenblatt M. Remarks on Some Nonparametric Estimates of a Density Function // *The Annals of Mathematical Statistics*. Vol. 27, no. 3. Sep. 1956, pp. 832-837. DOI:10.1214/aoms/1177728190.
15. Serfling R.J. *Approximation Theorems of Mathematical Statistics*. USA: John Wiley & Sons Inc., 1980. 400 p.
16. Silverman B.W. *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall/CRC, 1986. 45 p.

A TECHNIQUE FOR THE ESTIMATION OF FRACTION-OF-TIME PROBABILITY DENSITY OF A DIGITAL SIGNAL USING LINEAR INTERPOLATION

Timofey Ya. Shevgunov, Moscow Aviation Institute (National Research University), Moscow, Russia, shevgunov@gmail.com

Abstract

The paper presents the development of a new tool in fraction-of-time approach, where a random process is described by means of functional models synthesized based on its single realization, without the need to construct abstract probabilistic models as reliable a priori information on the process ergodicity property is lacking. Based on the analytical formula, expressing the fraction-of-time density of a continuous signal in explicit form, a method for estimating the fraction-of-time density of a digital signal was proposed. A piecewise continuous linear function was chosen as an approximating model used to restore a continuous-time signal from its samples. The first-order interpolating model made it possible to develop a light-weight digital signal processing algorithm that forms an analytical expression for the estimated fraction-of-time density of the observed signal. It is shown that the expression consists of two parts: a regular part represented by the sum of rectangle pulses, and a singular part represented by the sum of Dirac deltas. The simulation results are presented for three cases, including a periodic signal, a stationary random process with its known correlation function, and a second-order cyclostationary process. For the most complex case of a cyclostationary process, the proposed algorithm demonstrated a gain in estimating the probability density compared to the known method of kernel probability density estimation with a Gaussian smoothing function chosen. As an expected direction of practical application of the method and algorithm based on it developed within the framework of the conducted study, an improvement in the quality of solving data analysis problems for which the accuracy of probability density estimates is of significant importance is supposed, for example, problems of detecting random signals with unknown parameters.

Keywords: cyclostationarity, fraction-of-time, FOT density, linear interpolation, probability density estimator.

References

1. D. Dehay, J. Le?kow, A. Napolitano, T. Shevgunov (2023). Cyclic Detectors in the Fraction-of-Time Probability Framework. *Inventions*. Vol. 8, no. 6, P. 152. DOI: 10.3390/inventions8060152.
2. W.A. Gardner, W.A. Brown (1991). Fraction-of-time probability for time-series that exhibit cyclostationarity. *Signal Processing*. Vol. 23, no. 3, pp. 273-292. DOI: 10.1016/0165-1684(91)90005-4.
3. A. Napolitano, W.A. Gardner (2022). Fraction-of-Time Probability: Advancing Beyond the Need for Stationarity and Ergodicity Assumptions. *IEEE Access*. Vol. 10, pp. 34591-34612. DOI: 10.1109/ACCESS.2022.3162620.
4. W.A. Gardner (2023). Transitioning away from stochastic process models. *Journal of Sound and Vibration*. Vol. 565, P. 117871. DOI: 10.1016/j.jsv.2023.117871.
5. T. Shevgunov, A. Napolitano (2021). Fraction-of- Time Density Estimation Based on Linear Interpolation of Time Series. *2021 Systems of Signals Generating and Processing in the Field of on Board Communications*. Moscow. P. 1-4, DOI: 10.1109/IEEECONF51389.2021.9415991.
6. Shevgunov T.Ya. (2022). The synthesis of explicit analytical formulae for the probabilistic models of signals using fraction-of-time approach. *T-Comm*, vol. 16, no. 11, pp. 21-29. (in Russian) DOI: 10.36724/2072-8735-2022-16-11-21-29.
7. J. Le?kow, A. Napolitano (2006). Foundations of the functional approach for signal analysis. *Signal Processing*, vol. 86, pp. 3796-3825. DOI: 10.1016/j.sigpro.2006.03.028/
8. M. Taboga (2018). *Lectures on Probability Theory and Mathematical Statistics*, 3rd ed. USA, San Bernardino: CreateSpace. 670 p.
9. H. Pishro-Nik (2014). *Introduction to Probability, Statistics and Random Processes*. USA: Kappa Research, 744 p.
10. M. M. Deza, E. Deza (2016). *Encyclopedia of Distances*, 4th ed., Springer, 2016. DOI: 10.1007/978-3-642-30958-8.
11. M.A. Stephens (1979). Tests of Fit for the Logistic Distribution Based on the Empirical Distribution Function. *Biometrika*. vol. 66, no. 3, pp. 591-595. DOI: 10.1093/biomet/66.3.591.
12. V.A. Epanechnikov (1969). Nonparametric estimation of multidimensional probability density. *Theory of Probability and its Applications* 19. vol. 14, no. 1., pp. 156-161. (in Russian)
13. E. Parzen (1962). On Estimation of a Probability Density Function and Mode. *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065-1076. DOI: 10.1214/aoms/1177704472.
14. M. Rosenblatt (1956). Remarks on Some Nonparametric Estimates of a Density Function. *The Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 832-837. DOI:10.1214/aoms/1177728190.
15. R.J. Serfling (1980). *Approximation Theorems of Mathematical Statistics*. USA: John Wiley & Sons Inc., 400 p.
16. B.W. Silverman (1986). *Density Estimation for Statistics and Data Analysis*. London: Chapman & Hall/CRC, 45 p.

Information about author:

Timofey Ya. Shevgunov, Ph.D. (candidate of technical sciences), associate professor, Theoretical Radio Engineering department, Moscow Aviation Institute (National Research University) "MAI", Moscow, Russia