ПИСЬМА В РЕДАКЦИЮ

ПИСЬМА
В РЕДАКЦИЮ

УДК 537.874

ОБ АНАЛИТИЧЕСКОМ СИГНАЛЕ И СТАТЬЕ Ю. Н. ЗАЙКО «ИСТОРИЯ ОДНОГО АРТЕФАКТА»

М. В. Давидович

Саратовский государственный университет
E-mail: daviduvichmv@info.sgu.ru

Обсуждены ошибочные утверждения, допущенные в статье Ю. Н. Зайко [1] («Известия Саратовского университета. Нов. сер. Сер. Физика». 2012. Т. 12, вып. 1. С. 3–11).
Ключевые слова: аналитический сигнал, амплитуда, фаза, частота, распространение импульсов.

On the Analytical Signal and Yu. N. Zajko Article «A History of One "Artefact"»

M. V. Davidovich

The erroneous claims, made in the paper [1] (article of Yu. N. Zayko in «Izvestia of Saratov University. New Ser. Ser. Physics». 2012. Vol. 12, iss. 1. P. 3–11) have been discussed.
Key words: analytical signal, amplitude, phase, frequency, pulse propagation.

В журнале «Известия Саратовского университета» вышла статья [1] в основном обзорного характера по распространению импульсов, но в ней имеется ряд неверных и предвзятых утверждений автора (и даже передергиваний фактов), с которыми нельзя согласиться. Указанные ошибочные утверждения касаются критики теории аналитического сигнала (АС), который автор не приемлет, и конкретно – критике работ Л. А. Вайнштейна и Д. Е. Вакмана. Немного «досталось» и А. Ф. Голубенцеву.

Основные тезисы автора [1] сводятся к следующим ошибочным утверждениям. 1) АС отрицает быстрые (высокочастотные) осцилляции основных параметров сигнала – мгновенных (т.е. зависящих от текущего времени) амплитуды (огибающей), фазы и частоты, тогда как реально такие колебания имеются (соответствующая цитата: «С точки зрения аналитического сигнала никаких высокочастотных осцилляций волновых параметров сигнала (импульса) нет в принципе»). 2) Аналитический сигнал неоднозначен, так как «если добавить постоянную к действительной (мнимой) АС, то, поскольку результатом преобразования Гильберта постоянной является нуль, мы получим ту же мнимую (вещественную) часть, что и раньше, без добавления постоянной. Налицо неоднозначность». 3) АС в некотором роде эквивалентен методу усреднения (это утверждение фактически (неявно) содержится в п. 2 [1]). 4) АС ведет к нарушению принципа причинности (согласно п. 3, 4 в [1]).

При изложении автор противопоставляет АС комплексному сигналу (КС), считая, что он (КС) и есть правильное описание процессов: «В результате многие факты, полученные в рамках применения методов комплексного анализа, называемых Л. А. Вайнштейном и Д. Е. Вакманом теорией "комплексного сигнала", объявлялись артефактами и приписывались "недостаткам" последнего». Создается впечатление, что автор [1] либо не понял теорию АС и не прочитал внимательно работы [2–4], на которые он ссылается, либо занимается сознательным передергиванием результатов и вводит в заблуждение читателей журнала. Сразу надо отметить, что КС вводится лишь в [2, 3] как сигнал вида $f(t)=A(t)\exp(i\omega_0 t)$, соответствующий реальному сигналу (например, напряжению) $u(t)=A(t)\cos(\omega_0 t)$ (что, собственно, и отмечает автор [1]). В этом смысле КС есть приближенное обобщение метода комплексных амплитуд (символического метода), применяемого для гармонических процессов (которые, как известно, сигналами не являются), для которых $A(t)=\mathrm{const}$ (т.е. спектр есть $\delta(\omega)$). АС также есть обобщение метода (символического) комплексных амплитуд на произвольные нестационарные процессы (но уже точное [2, 3]). При этом требования на КС следующие: спектр $A(t)$ низкочастотный и лежит в полосе $(0,\Delta\omega)$, спектр же КС естественно переносится на частоту ω_0. Низкочастотность (медленность) $A(t)$ означает согласно [3], что $\Delta\omega/\omega_0 \ll 1$, т.е. сигнал квазимонохроматический. Известно, что сигнал, длящийся конечное время, имеет бесконечно большие частоты в спектре (реально они ограничены конечностью энергии сигнала E и квантовым условием $E<\hbar\omega$). Поэтому в [3] дается определение полосы частот $\Delta\omega$ так, чтобы в ней содержалась большая часть энергии сигнала (например, 99%). Определение указанной полосы можно сделать на основе временного и спектрального представлений энергии сигнала (теоремы Винера–Хинчина или равенства Парсеваля) [3]. Отсюда сразу имеем, что при $\Delta\omega \to 0$ АС переходит в КС монохроматического процесса, т.е. в КС. При $\Delta\omega/\omega_0 \ll 1$ КС отличается от АС быстрыми колебаниями с малой амплитудой (это совсем не означает, что в АС нет быстрых колебаний, а только то, что появившиеся в КС дополнительные колебания суть нефизические). На самом деле КС неоднозначен, и определить $A(t)$ можно несколькими путями [2, 3]. Несколько таких КС также отличаются на подобные быстрые колебания [2, 3]. Неоднозначность КС уже, помимо всего прочего, следует из произвольности выбора «несущей ω_0» для сигнала с полосой $\omega_1 \le \omega \le \omega_2$ при $\omega_2-\omega_1 < \omega_2+\omega_1$ (указанная неоднозначность в АС отсутствует). Спектр подобных КС сигналов слабо затекает в область отрицательных частот (энергия на отрицательных частотах мала), чем и обусловлено здесь слабое отличие КС от АС (проявляющееся в дополнительных осцилляциях с малой энергией). Как известно, АС есть сигнал с колебаниями на положительных частотах, спектральная амплитуда которых удваивается [1]. Таким образом, «нефизические» отрицательные частоты отбрасываются. Это совсем не значит, что спектральный интеграл со всеми частотами есть КС (как, по-видимому, трактует его автор [1], судя по смыслу его высказываний). Это обычный реальный сигнал, которому АС никак не противопоставляется [3]. Он реален в силу условия на его спектр: $U(\omega)=U^*(-\omega)$ (для АС $W(\omega)=2U(\omega)$, $\omega \ge 0$). Как известно, отрицательные частоты имеют определенный физический смысл, например, в квантовой электродинамике. Аналитический же сигнал $w(t)=u(t)+iv(t)$ имеет тот смысл, что $|w(t)|=a(t)$ есть амплитуда (огибающая) сигнала $u(t)$ (и соответственно $v(t)$), а $\arg(w(t))=\varphi(t)$ – его фаза (фаза сигнала $v(t)$ есть $\varphi(t)-\pi/2$ согласно преобразованию Гильберта). Величина же $A(t)$ не есть амплитуда, так же как и $\omega_0 t$ не есть мгновенная фаза. Правильно назвать $A(t)$ формфактором. Он может быть прямоугольным (функция $\Pi(t)$ в [3]), гауссовым, треугольным, трапецеидальным, в виде полукосинусоиды или нескольких полукосинусоид (скажем, квазигармоническая модуляция несущей) и т.п., но неправильно называть такие сигналы (импульсы). как сигналы с прямоугольной, гауссовой и т.п. огибающими или, например, гауссовым импульсом (что, кстати, делает автор [1]). Такие названия используются в литературе, но это жаргон. Указанные формфакторы никак огибающими не являются [2–4]. Обычно рассматривают конечные во времени формфакторы $A(t)$, за исключением гауссова. Огибающая прямоугольного видеоимпульса (без несущей или высокочастотного множителя) рассмотрена в [3]. Она имеет логарифмические сингулярности на фронтах, а также плавно спадающий фронт и «хвост». Собственно по поводу первого и идет дискуссия: фронт мол (сверхсветовой предвестник) опережает импульс и движется со сверхсветовой скоростью, нарушая принцип причинности! По этому поводу можно посмотреть разъяснения в [3]. Кстати, подобные эффекты (численное логарифмическое обострение фронтов «прямоугольного» радиоимпульса и сглаживание его бесконечно резких фронтов) получено и в работах Г. М. Стрелкова [5–7], на которые ссылается автор [1]. Сам Г. М. Стрелков также отрицает АС (это удалось установить в личной беседе с ним). Как ни странно, и он, и Ю. Н. Зайко при этом фактически используют АС при расчетах (вообще, весьма многие не принимали теорию АС в основном, как отмечено в [3], из-за заблуждений). В работе [8] Ю. Н. Зайко численно получил известные до этого (см. [1]) осцилляции мгновенной частоты и амплитуды на фронтах радиоимпульса в идеальном волноводе. При этом он использовал пространственно-временную функцию Грина (ФГ), вывод которой дан в работе Вайнштейна [4] и который основан на аналитических свойствах спектрального интеграла, т.е. фактически на тех же принципах, что и АС. Г. М. Стрелков использовал другой подход, позволивший ему учесть потери в плазме, но также сводящийся к АС (о подходах к моделированию распространения импульсов будет сказано далее). Заметим, что соотношения Крамерса–Кронига (т.е. квинтэссенция принципа причинности в электродинамике) есть следствие теории АС. Так в чем же причина «нарушения» принципа причинуости? На самом деле объяснение весьма простое: АС есть сигнал комплексный (в том смысле, что описывается комплексными числами как и КС). То есть сигналом (который движется, переносит информацию и энергию) он, по сути, не является. Тем более он относится к двум взаимно сопряженным по Гильберту реальным сигналам $u(t)$ и $v(t)=H(u(t))$ (которые и переносят информацию, не испытывая никаких нарушений причинности). АС есть математическая абстракция, описывающая амплитуду, фазу и частоту. Для разрывного формфактора (например, прямоугольного) он дает сглаженную амплитуду, которая обязана выходить за пределы его значений. Этот выход экспоненциально мал и не может быть обнаружен амплитудным детектором [3], который реагирует на реальный сигнал $u(t)$, но так, что $a(t)=|w(t)|$ (поэтому и можно говорить, что на вход «поступает сигнал» $w(t)$, а на выходе имеем $|w(t)|$). Реально на детектор поступают колебания, которые даже от предвестника приходят с запаздыванием, а в результате детектор «срабатывает» даже с некоторой задержкой, поэтому для коррекции времени прихода резкой части фронта импульса необходимо даже делать упреждение [3]. Видеоимпульсы, проходящие через реальную аппаратуру, всегда имеют сглаженные фронты (даже бесконечно дифференцируемые). Для таких формфакторов (пример – гауссов формфактор, но он нефизический) АС не дает опережения [9]. В любом случае видеоимпульс будет продетектирован не раньше, чем он поступит на вход. Как показано в [3], все приборы (функциональные устройства), выполняющие некоторые операторные преобразования над сигналами, воспринимают их так, как будто на их вход поступает АС. В частности, детектор (линейный) определяет $a(t)=|w(t)|$, а квадратичный – $|w(t)|^2$ (с той или иной степенью точности). Так, можно определять амплитуду с помощью двухполупериодного идеального линейного детектора (моста), поставив на выходе НЧ фильтр. Для компьютерного моделирования для этого можно взять модуль сигнала и выполнить усреднение методом скользящего среднего (что есть НЧ цифровой фильтр). Чем больше отсчетов (в том числе и на средний период осцилляций), тем более будет запаздывание, но амплитуда будет определяться тем точнее, и в пределе будет соответствовать сдвинутой (запаздывающей) амплитуде $|w(t)|$. Естественно реальный детектор осуществляет компромисс между временем срабатывания и точностью детектирования. АС абсолютно точно пригоден и в случаях $\Delta\omega \sim \omega_0$, $\Delta\omega > \omega_0$ (тогда определение ω_0 теряет смысл, и лучше писать так: $2(\omega_2-\omega_1)\geq\omega_2+\omega_1$), тогда как КС здесь теряет всякий смысл. В [2–3] АС определен для сигналов, заданных на всей временной оси. Такие сигналы соответствуют фильтрации 1-го рода [3, 10]. Можно определить сигналы, заданные на текущий момент времени *t*. Им соответствует фильтрация 2-го рода (физически реализуемая) [3, 10]. Если сигнал существует с конечного момента в прошлом – это фильтрация 3-го рода. Таким сигналам соответствуют мгновенные спектры:

$$U(\omega,t)=\int_{t-T}^{t}u(t')\exp(-i\omega t')dt',$$
$$u(t)=\frac{1}{2\pi}\int_{-\infty}^{\infty}U(\omega,t)\exp(i\omega t)d\omega. \qquad (1)$$

Здесь T – временное окно периодограммы сигнала на момент t. Мы как бы полагаем неизвестный нам в будущем сигнал равным нулю. Мгновенный спектр зависит от текущего времени как от параметра. Теория АС, построенная на соотношениях (1), имеет тот же вид с теми же свойствами АС.

Теперь ответим на основные пункты.

1) Теория АС утверждает, что амплитуда, фаза и частота связаны посредством АС. Это означает, что там, где сильно меняется амплитуда (например, на фронтах импульсов), сильно изменяется и мгновенная частота. Если на плоской вершине квазимонохроматического импульса мгновенная частота $\omega(t)$ приблизительно равна ω_0 и совершает небольшие осцилляции, то на фронте ее перепад изменяется от 0 до $2\omega_0$ и более. Тем более это справедливо для импульсов с широким спектром. Загвоздка заключается только в том, что строго АС для распространения импульсов почти никто не использовал (в отличие от сигналов) в силу сложности вычисления несобственных интегралов, поэтому применяют приближенные подходы. Если эти подходы (как в [5–7]) есть хорошее приближение к теории АС, результаты правильные. По этой же причине имеют место различные определения фазы в нелинейной динамике (поскольку использование АС и преобразования Гильберта приводит к большим вычислительным трудностям).

2) Действительно, АС не «воспринимает» постоянный (на всей временной оси) «сигнал», и это говорит в его пользу (было бы весьма плохо, если бы это было не так). Такой «сигнал», как известно, сигналом не является, как не является сигналом и монохроматический процесс (в частности, $\sin(\omega_0 t)$). Сигнал определяется как функция времени с конечной энергией, т.е. функция, интегрируемая с квадратом (в пространстве Гильберта L_2):

$$\int_{-\infty}^{\infty} [u(t)]^2 dt < \infty . \qquad (2)$$

Либо такая функция должна иметь конечные временные пределы и не более чем счетное число скачков первого рода, либо если временные границы бесконечные, она должна еще и убывать на бесконечностях так, чтобы интеграл (3) сходится. Условие (2) есть необходимое условие существования спектра (интеграла Фурье), т.е. и определения АС. Постоянный «сигнал» этому не удовлетворяет, поэтому возражение 2) не выдерживает критики. Заметим, что АС применяется во многих областях, в том числе и в теории случайных процессов (шумов). Там условие (2) может не выполняться, и стационарный случайный сигнал может иметь среднее значение, отличное от нуля. В таком случае образуют новый стационарный сигнал, вычитая это значение: $u'(t) = u(t) - \overline{u}(t)$, и такие сигналы считаются одинаковыми [3, 10]. Не важно, относительно какого среднего значения осуществляется колебание.

3) АС не есть некий метод усреднения. Наоборот, в [3] он противопоставлен методам усреднения, поскольку позволяет в любом приближении решать задачи, особенно если спектры разделяются. Нигде в теории АС нет усреднений (т.е. интегрирований по времени). Преобразование Гильберта – это фильтр, сдвигающий фазы на $-\pi/2$ (его можно написать в виде спектрального интеграла, который не есть усреднение).

4) По поводу нарушения принципа причинности все уже было разъяснено выше. Поэтому все претензии Ю. Н. Зайко к АС несостоятельны. Работа также изобилует рядом неточных или некорректных высказываний, большая часть из которых содержится в несоразмерном обилии сносок. В начале работы пространно обсуждаются не имеющие к ее дальнейшему содержанию вопросы гонки вооружений при холодной войне. Далее автор везде путает понятия спектральный интеграл (представление действительного сигнала или импульса в комплексной форме в виде спектрального интеграла) и КС, подменяя последний первым. Никто не называл такие интегралы КС (см. [2–4]). Автор утверждает, что в работах по распространению импульсов неопределенность в разбиении сигнала $u(t) = a(t)\cos(\varphi(t))$ на аплитудный и фазовый сомножители всегда снята уже из-за явного их задания. Но в спектральном интеграле (одномерный импульс вдоль оси z) задается величина $U(\omega)\exp(i\omega t - k_z(\omega)z)$, а в исходном сигнале – $u(t) = A(t)\cos(\omega_0 t)$. Мы уже выяснили, что $A(t) \neq a(t)$, $\varphi(t) \neq \omega_0 t$. И где тут это разделение? Тем более, если импульс задается не сигналом в точке $z = 0$, а некими источниками (током) в этой точке или в ее окрестности (точнее, плотностью тока). Не понятен смысл сноски 14: «Это равносильно не суммировать в ряду $\sin x + \sin 3x/3 + \sin 5x/5 + \ldots$, который сходится к ступенчатой функции, слишком много членов».

Во-первых, ряд sin x + sin $3x/3$ + sin$5x/5$ +... = = sin x + sin x +... расходится. Наверное, имелся в виду ряд sin(x) + sin(x)/3 + sin(x)/5 +..., но и в этом случае предложение безграмотно написано. В работах [3, 4] говорится, что колебания с удвоенной несущей частотой возникают для КС на всем протяжении его определения во времени. Причем введенный сигнал ([3, с. 8]) $u(t)=(1+m\cos(\Omega t))\cos(\omega_0 t)$ не имеет амплитуду $1+m\cos(\Omega t)$ – это некая усредненная амплитуда, которую определяет детектор. Формально $\Omega<\omega_0$, и согласно АС $\mathrm{v}(t)=(1+m\cos(\Omega t))\sin(\omega_0 t)$. На самом же деле следует записать $u(t)=\Pi(t)(1+m\cos(\Omega t))\cos(\omega_0 t)$, где $\Pi(t)$ – прямоугольное окно, в котором действует сигнал. Поэтому в центре и основной области этого окна действительно $a(t)\approx 1+m\cos(\Omega t)$, $\omega(t)\approx\omega_0$. На его краях (фронтах) это уже не так: амплитуда и частота меняются резко. Для этого надо строго вычислять $\mathrm{v}(t)$ с помощью преобразования Гильберта, что никто никогда обычно не делает. Отсюда, наверное, и возникают недоразумения у тех, кто получил подобные осцилляции. КС же дает нефизические осцилляции на всем временном интервале, что и отмечали авторы [2, 3]. Обычно такие осцилляции при вычислениях усредняют (естественно, они детектором не определяются).

Рассмотрим теперь вопрос о том, как осуществляется моделирование процесса распространения импульсов. Будем рассматривать только одномерные (плоские) импульсы. Возможна постановка задачи в виде задания «источника» импульса – сигнала $u(t,0)$ в точке $z=0$ с дальнейшим анализом, как он распространяется вправо:

$$u(t,z)=\frac{1}{2\pi}\int_{-\infty}^{\infty}U(\omega)\exp(i\omega t-ik_z(\omega)z)d\omega. \quad (3)$$

Здесь $k_z(\omega)$ определяет закон дисперсии и есть аналитическая функция в нижней полуплоскости комплексной частоты. Очевидно, АС для (2) есть

$$w(t,z)=\frac{1}{\pi}\int_{0}^{\infty}U(\omega)\exp(i\omega t-ik_z(\omega)z)d\omega, \quad (4)$$

а $|w(t,z)|$ есть искомая огибающая [3]. Этот же подход соответствует и постановке задачи, когда импульс заданной формы (например, с прямоугольным формфактором) и заданным спектром подошел слева к границе раздела вакуум-диспергирующая среда. Реальный источник в точке z создает два импульса, движущиеся в противоположных направлениях. Можно непосредственно вычислять интегралы (3), (4), выделяя полюса, явно вычисляя асимптотические интегралы и вычитая асимптотические члены. Этот процесс весьма сложен, поскольку $k_z(\omega)$ – функция комплексная, а экспонента не осциллирует периодически. Другой способ – введение пространственно-временной ФГ [3, 4] (мы назовем ее пропагаторной ФГ, поскольку она по смыслу аналогична пропагаторной ФГ в квантовой механике для волнового пакета, определяющего вероятность положения и скорости частицы [11]). Третий способ – решение дифференциальных и интегродифференциальных уравнений (ИДУ) электродинамики. ИДУ возникают, когда индукция связана с полями интегральным соотношением:

$$\mathbf{D}(\mathbf{r},t)=\varepsilon_0\int_{-\infty}^{t}\varepsilon(t-t')\mathbf{E}(\mathbf{r},t')\,dt'. \quad (5)$$

Здесь для простоты не учтены ни анизотропия (тензорный характер ε), ни бианизотропия (зависимость от поля $\mathbf{H}(t')$), ни пространственная дисперсия (интегрирование по координатам), ни неоднородность (зависимость ε от координаты). Можно решать непосредственно уравнения Максвелла (что обычно делают при использовании метода FDTD), либо перенося ток поляризации в их правую часть и решая их совместно с уравнением для тока поляризации. Часто при этом переходят от волнового уравнения к параболическому уравнению, пренебрегая второй производной по времени (именно так и моделировал Г. М. Стрелков). Оно тогда определяет «медленно меняющуюся амплитуду» по сравнению с быстрыми осцилляциями поля. Естественно это приближение, но оправданное упрощением вычислений. Другой строгий подход может быть основан на методе объемных интегральных уравнений (ОИУ). В этом подходе используются причинные пространственно-временные тензорные ФГ [12] и токи поляризации как вторичные источники поля. При этом естественно учитываются и первичные источники, т.е. сторонние токи, создавшие импульс. Естественно источники излучают во всех направлениях. Подобный подход удобен и при наличии в рассматриваемой среде структур (например, пластин с другими проницаемостями). Поскольку такие ОИУ, удовлетворяющие, кстати, явно принципу причинности, весьма громоздки, особенно если имеются зависимости типа (5) и среда неоднородна, мы не приводим их явного вида.

Список литературы

1. *Зайко Ю. Н.* История одного артефакта // Изв. Сарат. ун-та. Нов. сер. Сер. Физика. 2012. Т. 12, вып. 1. С. 3–11.
2. *Вайнштейн Л. А., Вакман Д. Е.* Амплитуда, фаза, частота – основные понятия теории колебаний // УФН. 1977. Т. 123, вып. 4. С. 657–682.
3. *Вайнштейн Л. А., Вакман Д. Е.* Разделение частот в теории колебаний и волн. М. : Наука, 1983. 288 с.
4. *Вайнштейн Л. А.* Распространение импульсов // УФН 1976. Т. 118, вып. 2. С. 339–367.
5. *Стрелков Г. М.* Распространение радиоимпульса в изотропной плазме // Радиотехника и электроника. 2006. Т. 51. С. 672–682.
6. *Стрелков Г. М.* Сложный радиосигнал в ионосферной плазме // Радиотехника и электроника. 2008. Т. 53, № 1. С. 1094–1103.
7. *Стрелков Г. М., Нарышкин В. И.* Распространение радиоимпульса с линейной частотной модуляцией в изотропной плазме // Радиотехника и электроника. 2008. Т. 53, № 1. С. 49–57.
8. *Зайко Ю. Н.* Частотная модуляция заполнения радиоимпульса, распространяющегося в диспергирующей среде // Изв. вузов. Радиофизика. 1989. Т. 32, № 12. С. 1558–1560.
9. *Давидович М. В.* Прохождение сигналов через фильтр с поглощением и отрицательное время задержки // ЖТФ. 2012. Т. 82, вып. 3. С. 15–22.
10. *Вайнштейн Л. А., Зубаков В. Д.* Выделение сигналов на фоне случайных помех. М. : Сов. радио, 1960. 447 с.
11. *Грибов В. Н.* Квантовая электродинамика. М. ; Ижевск, 2001. 288 с.
12. *Фелсен Л., Маркувиц Н.* Излучение и рассеяние волн : в 2 т. М. : Мир, 1978. Т. 1. 550 с.