

НЕЛИНЕЙНАЯ ПУАССОНОВСКАЯ МОДЕЛЬ ФОТОЭЛЕКТРИЧЕСКОГО ДЕТЕКТИРОВАНИЯ

© 2008 г. Ю. Н. Раковский, канд. техн. наук

НПК “Государственный оптический институт им. С.И. Вавилова”, Санкт-Петербург

Выходной сигнал фотодетектора рассмотрен с позиции теории марковских процессов размножения и гибели. Получены выражения для первых двух моментов сигнала, порога чувствительности и динамического диапазона.

Коды OCIS: 040.5160, 160.2500.

Поступила в редакцию 22.09.2008.

Процесс фотоэлектрического детектирования можно представить в виде двух последовательных преобразований: преобразования излучения (потока фотонов) в поток носителей заряда (фотоэлектронов) и преобразования потока носителей заряда в выходной сигнал. В статье рассматривается в основном второе преобразование.

Полагая приращение выходного сигнала пропорциональным приращению числа носителей заряда и рассматривая процесс накопления заряда на интервале $[0, T]$, выходной сигнал можно представить в виде интеграла Стилтеса

$$u = \int_0^T G(t) dm(t), \quad (1)$$

где $G(t)$ – функция преобразования числа фотоэлектронов в выходной сигнал (отклик на один фотопоток), $m(t)$ – число носителей заряда в момент t .

Будем считать поток носителей заряда ординарным и независимым (пуассоновским). При этом математическое ожидание и дисперсия выходного сигнала в соответствии с теоремой Кэмпбелла [1] равны соответственно

$$U = \langle u \rangle = \int_0^T G(t) d\langle m(t) \rangle, \quad (2)$$

$$\sigma^2 = \int_0^T G^2(t) d\langle m(t) \rangle. \quad (3)$$

Здесь $\langle \rangle$ – символ математического ожидания (МО), σ – среднеквадратическое отклонение (СКО).

Не все появляющиеся носители заряда вносят свой вклад в выходной сигнал, некоторая их часть исчезает (рекомбинирует). Поэтому задачу нахождения характеристик потока $m(t)$ можно рассмотреть с позиции математической теории размножения и гибели.

Для ординарного и независимого потока дифференциальное уравнение размножения и гибели имеет вид [2]

$$d\langle m(t) \rangle / dt = \mu(t) - \gamma(t) \langle m(t) \rangle. \quad (4)$$

В выражении (4) $\mu(t)$ – интенсивность потока появления носителей заряда (МО появляющихся зарядов в единицу времени), $\gamma(t)$ – интенсивность потока разрежения (потока гибели). Уравнение (4) называется также “кинетическим уравнением рекомбинации”, а функция $1/\gamma(t)$ – интерпретируется как среднее время жизни носителя заряда [3].

Интенсивность потока появления зарядов полагается пропорциональной потоку фотонов:

$$\mu(t) = \int \eta(\lambda) v_\lambda(t, \lambda) d\lambda, \quad (5)$$

где $v_\lambda(t, \lambda)$ – спектральная плотность интенсивности потока фотонов, $\eta(\lambda)$ – вероятность появления носителя заряда под воздействием фотона, λ – длина волны. Пределы интегрирования определяются областью задания подынтегральных функций.

Введя относительную спектральную характеристику

$$\rho(\lambda) = \eta(\lambda) / \eta(\lambda_m), \quad \lambda_m: \eta(\lambda_m) = \eta_{\max} = \max \quad (6)$$

и эффективное значение интенсивности потока фотонов

$$v(t) = \int v_\lambda(t, \lambda) \rho(\lambda) d\lambda, \quad (7)$$

выражение (5) можно представить в виде

$$\mu(t) = \eta_{\max} v(t). \quad (8)$$

Будем считать, что помимо параметра (η), имеющего смысл вероятности появления (“рождения”) заряда, фотодетектор характеризуется также параметром (ξ), имеющим смысл вероятности его исчезновения (“гибели”). Тогда интенсивность потока разрежения (потока гибели) составит [2]

$$\gamma(t) = \xi \mu(t). \quad (9)$$

Рассмотрим детектирование в режиме накопления заряда при нулевых начальных условиях. При этом решение уравнения (4) с учетом (9) имеет вид

$$\langle m(t) \rangle = \int_0^t \mu(x) \exp\left(-\xi \int_x^t \mu(y) dy\right) dx, \quad (10)$$

а входящее в (2) и (3) приращение

$$d\langle m(t) \rangle = \mu(t) \left[1 - \xi \int_0^t \mu(x) \exp\left(-\xi \int_x^t \mu(y) dy\right) dx \right] dt. \quad (11)$$

Функция преобразования рассматривается в виде резистивно-емкостного интегратора (накопителя) с весовой функцией

$$G(t) = \exp(-t/\theta). \quad (12)$$

Параметр θ характеризует “неидеальность” накопителя: при $\theta = \infty$ накопитель идеальный (счетчик числа носителей заряда). Несущественный для дальнейшего анализа постоянный множитель в (12) опущен.

Расчет выходного сигнала при произвольной зависимости интенсивности μ от времени достаточно сложен. Здесь мы ограничимся случаем, когда интенсивность потока появления носителей заряда на интервале $[0, T]$ можно считать практически постоянной. В этом случае

$$d\langle m(t) \rangle = \mu \exp(-\xi \mu t) dt. \quad (13)$$

При этом математическое ожидание и дисперсия сигнала u равны соответственно

$$U(\mu, T, \theta, \xi) = \mu \theta [1 - \exp(-(1 + \xi \mu \theta) T / \theta)] / (1 + \xi \mu \theta), \quad (14)$$

$$\sigma^2(\mu, T, \theta, \xi) = U(\mu, T, \theta/2, \xi). \quad (15)$$

Корректность модели проверяется, в частности, возможностью сведения ее результатов к известным ранее соотношениям путем физически понятных предельных переходов при изменении соответствующих параметров модели.

При вероятности “гибели” носителя заряда, равной нулю, имеем линейную модель

$$U(\mu, T, \theta, 0) = \mu \theta [1 - \exp(-T/\theta)] \quad (16)$$

– “накопление на постоянной времени” [4]. Если, кроме того, накопитель идеальный, то

$$U(\mu, T, \infty, 0) = \mu T. \quad (17)$$

Очевидно, что выражение (16) можно использовать как приближение при условиях

$$\xi \mu T \ll 1 \text{ и } \xi \mu \theta \ll 1, \quad (18)$$

а выражение (17) – как приближение в случае достаточно малого времени накопления, т. е. при

$$\xi \mu T \ll 1 \text{ и } T/\theta \ll 1, \quad (19)$$

В случае идеального накопления

$$U(\mu, T, \infty, \xi) = \sigma^2(\mu, T, \infty, \xi) = [1 - \exp(-\xi \mu T)] / \xi. \quad (20)$$

Если формально ввести параметр (“время жизни”)

$$\tau = 1/\xi \mu, \quad (21)$$

то выражение (20) можно представить в виде

$$U(\mu, T, \tau) = \mu \tau [1 - \exp(-T/\tau)] \quad (22)$$

– “накопление на времени жизни” [4].

В общем случае зависимости МО и СКО сигнала как функций интенсивности потока появления носителей заряда (μ) существенно нелинейны и, более того, эти величины ограничены:

$$U(\infty, T, \theta, \xi) = \sigma^2(\infty, T, \theta, \xi) = 1/\xi \quad (23)$$

и, кроме того,

$$U(\mu, \infty, \theta, \xi) = \mu \theta / (1 + \xi \mu \theta). \quad (24)$$

Поток носителей заряда представляет собой сумму потока, вызванного излучением наблюдаемого источника (полезный сигнал), потока, вызванного излучением фона, и чисто темновой составляющей, обусловленной термогенерацией, так что полную интенсивность потока можно представить в виде

$$\mu = \mu_s + \mu_b + \mu_d, \quad (25)$$

где μ_s – составляющая полезного сигнала, μ_b и μ_d – фоновая и темновая составляющие. Последние два компонента объединим в одну (далее – “шумовую”) составляющую:

$$\mu_0 = \mu_b + \mu_d. \quad (26)$$

Ограничившись случаем, когда накопитель можно считать практически идеальным, т. е. используя выражение (20), посмотрим, как нелинейность фотодетектора влияет на такие его характеристики, как относительная спектральная характеристика (измеренная стандартным методом), пороговая характеристика (порог чувствительности) и динамический диапазон.

Спектральная характеристика

Будем считать, что устанавливаемая при измерении спектральной характеристики интенсивность потока фотонов много больше шумовой. Если измерения проводятся при постоянной интенсивности потока фотонов на заданной длине волны, т. е. при

спектральной плотности интенсивности потока фотонов вида

$$v_\lambda(\lambda') = v\delta(\lambda - \lambda'), \quad (27)$$

где $\delta(\cdot)$ – дельта-функция, то с учетом (5), (27) и (20) относительная спектральная характеристика имеет вид

$$s(\lambda) = [1 - \exp(-\xi\eta(\lambda)vT)]/[1 - \exp(-\xi\eta_{\max}vT)]. \quad (28)$$

Относительная спектральная характеристика (28) не совпадает с “истинной” характеристикой (6), и только при условии $\xi vT \ll 1$ имеем приближение

$$s(\lambda) \approx \eta(\lambda)/\eta_{\max} = \rho(\lambda). \quad (29)$$

Отметим, что при постоянном значении квантовой эффективности, т. е. при

$$\eta(\lambda) = \eta, \quad 0 \leq \lambda \leq \lambda_m \quad (30)$$

характеристика (28) без всяких условий совпадает с (29).

Иначе дело обстоит, если спектральная характеристика измеряется при постоянном лучистом потоке (Φ), т. е. при спектральной плотности лучистого потока вида

$$\Phi_\lambda(\lambda') = \Phi\delta(\lambda - \lambda'). \quad (31)$$

В этом случае с учетом связи между интенсивностью потока фотонов и лучистым потоком

$$v_\lambda(\lambda) = (\lambda/hc)\Phi_\lambda(\lambda) \quad (32)$$

(h – постоянная Планка, c – скорость света) “радиометрическую” относительную спектральную характеристику можно представить в виде

$$r(\lambda) = [1 - \exp(-(\lambda\eta(\lambda)/hc)\xi\Phi T)]/[1 - \exp(-(\lambda_m\eta_{\max}/hc)\xi\Phi T)]. \quad (33)$$

В линейном приближении, т. е. при условии

$$(\lambda_m\eta_{\max}/hc)\xi\Phi T \ll 1, \quad (34)$$

(33) переходит в известную форму

$$r(\lambda) = (\lambda/\lambda_m)(\eta(\lambda)/\eta_{\max}) = (\lambda/\lambda_m)\rho(\lambda). \quad (35)$$

Однако при том же условии (30) в отличие от предыдущего случая спектральная характеристика

$$r(\lambda) = [1 - \exp(-(\lambda/hc)\xi\eta\Phi T)]/[1 - \exp(-(\lambda_m/hc)\xi\eta\Phi T)] \quad (36)$$

не совпадает с линейным приближением (35). Введя обозначения

$$\lambda/\lambda_m = z, \quad (\lambda_m/hc)\xi\eta\Phi T = q, \quad (37)$$

характеристику (36) удобно представить в виде

$$r(z, q) = [1 - \exp(-qz)]/[1 - \exp(-q)]. \quad (38)$$

Функция (38) представлена на рис. 1. Сравнивая ее с измеренной спектральной характеристикой фотодетектора на основе арсенида индия из работы [5] (рис. 2), можно заключить, что рассматриваемая модель, по крайней мере качественно, соответствует реальности. Для получения спектральной характеристики (35) нужно либо обеспечить условие (34), либо проводить измерения при постоянном выходном сигнале.

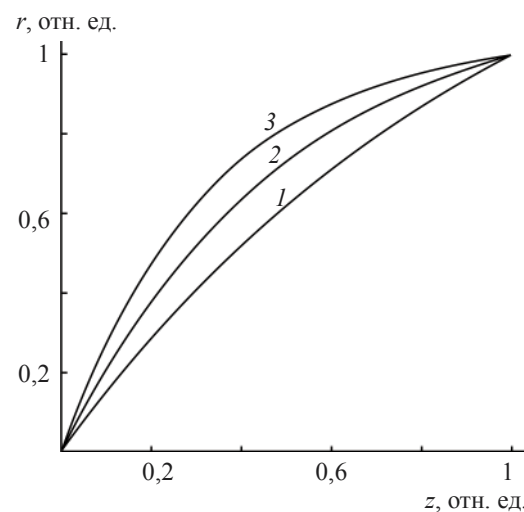


Рис. 1. Спектральные характеристики (38). 1 – $q = 1$, 2 – 2, 3 – 3.

Пороговая характеристика

Введем обозначения для МО числа появившихся за время T зарядов полезного сигнала и шума соответственно

$$M = \mu_s T, \quad N = \mu_0 T. \quad (39)$$

Значение величины M , удовлетворяющее условию

$$U_s(M) = U(M + N) - U(N) = \sigma(N), \quad (40)$$

считается пороговым. Соответствующее значение эффективного лучистого потока рассчитывается по формуле

$$\Phi = Mhc/T\lambda_m \eta_{\max}. \quad (41)$$

Заметим, что условие (40), соответствующее ГОСТ 17772-88, не является условием равенства отношения сигнал/шум единице, поскольку СКО шума зависит от полезного сигнала. “Порог чувст-

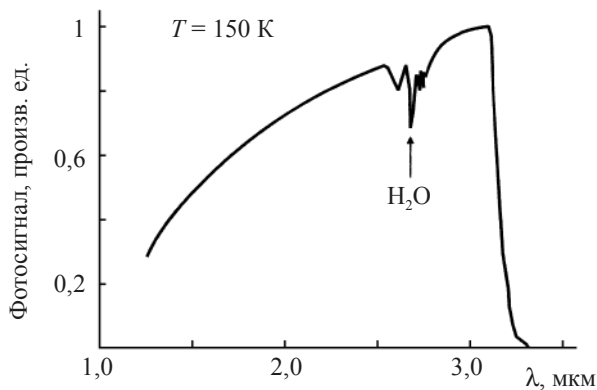


Рис. 2. Спектральная характеристика из работы [5].

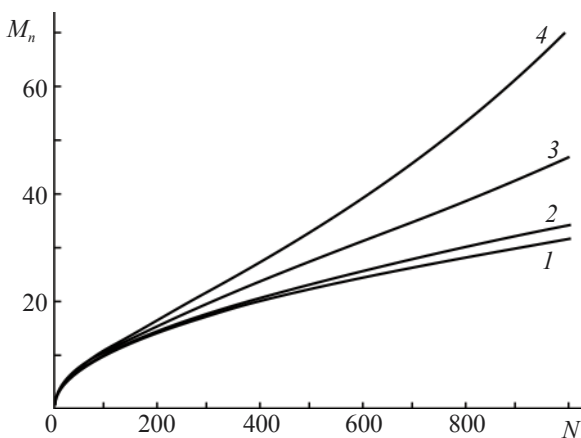


Рис. 3. Пороговые характеристики. 1 – $\xi = 10^{-6}$, 2 – 10^{-4} , 3 – 5×10^{-4} , 4 – 10^{-3} .

вительности” по определению (40) представляет собой уровень собственного шума (шума в отсутствие полезного сигнала) в единицах входного воздействия.

В рассматриваемом случае (идеальное накопление) пороговое значение величины M составит

$$M_n(N, \xi) = (1/\xi) \ln[1/(1 - \exp(\xi N)(\xi(1 - \exp(-\xi N)))^{1/2})]. \quad (42)$$

При $\xi = 0$ имеем известный результат линейной модели

$$M_n(N, 0) = N^{1/2}. \quad (43)$$

При $\xi \geq 10^{-4}$ отличие порога (42) от линейной модели уже достаточно ощутимо (рис. 3).

Динамический диапазон

В терминах и обозначениях, принятых в настоящей статье, динамический диапазон определяется как отношение

$$D = M_{\max}/M_{\min}. \quad (44)$$

ГОСТ 17772-88 определяет верхнюю границу диапазона (M_{\max}) из условия

$$U_s(M_k)/U_l(M_k) = k, \quad (k = 0,8) \quad (45)$$

где

$$U_l(M) = M \left. \frac{dU_s}{dM} \right|_{M=0}. \quad (46)$$

В качестве нижней границы ГОСТ определяет “порог чувствительности” из условия (40)

$$M_{\min} = M_n. \quad (47)$$

Для полезного сигнала вида

$$U_s(M) = [(1 - \exp(-\xi M))\exp(-\xi N)]/\xi \quad (48)$$

и $k = 0,8$ $M_{\max} = M_{0,8} = 0,465/\xi$. При таком значении M выходной сигнал составляет всего 37% своего максимального значения, что вряд ли оправдано. Для ограниченного по величине выходного сигнала предпочтительнее условие определения верхней границы диапазона вида

$$U_l(M_{\max}) = U_{s \max}, \quad U_{s \max} \cdot U_s(M) = \max. \quad (49)$$

Для сигнала (48) и условия (49) выходной сигнал на верхней границе диапазона составляет 63% своего максимума. При этом $M_{\max} = 1/\xi$, а динамический диапазон

$$D = M_{\max}/M_n = \left(\ln \left(1 / \left(1 - e^{\xi N} \sqrt{\xi (1 - e^{-\xi N})} \right) \right) \right)^{-1}. \quad (50)$$

Использование нижней границы диапазона (40) имеет смысл, если СКО шума не зависит от сигнала (или как приближение при $M \ll N$). Но при определении динамического диапазона нельзя считать полезный сигнал малым и необходимо учесть зависимость СКО шума от сигнала. Поэтому динамический диапазон должен зависеть от абсолютного значения полезного сигнала. При этом очевидно, что чем больше минимальный сигнал, тем меньше должен быть динамический диапазон.

Физически понятное и непротиворечивое определение динамического диапазона для ограниченного по величине сигнала может состоять в следующем.

Задается нижняя граница диапазона $M_{\min} = M$. Верхняя граница определяется как

$$M_{\max} = M(U_{s\max}/\sigma(M)). \quad (51)$$

При этом динамический диапазон

$$D(M) = M_{\max}/M = U_{s\max}/\sigma(M). \quad (52)$$

В рассматриваемом случае (идеальное накопление)

$$D(M) = [\xi(1 - \exp(-\xi M - \xi N))^{-1/2} \exp(-\xi N)]. \quad (53)$$

Функция $D(M, N, \xi)$ для некоторых значений переменных представлена на рис. 4. Как и следовало

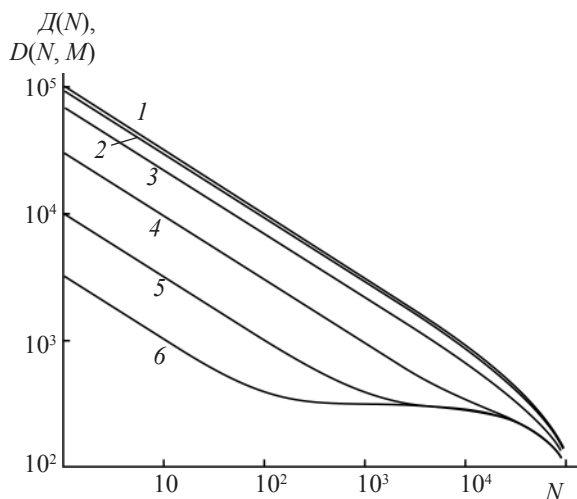


Рис. 4. Динамический диапазон $D(N) - 1$ и зависимости динамического диапазона $D(N, M) - 2-6$ ($\xi = 10^{-5}$). 2 - $M/N = 0,1$, 3 - 1, 4 - 10, 5 - 100, 6 - 1000.

ожидать, чем больше минимальное значение сигнала, тем меньше динамический диапазон.

При $M = 0$ имеем соотношения

$$D(0) = [1 - \exp(-\xi M_n)]^{-1} = [1 - \exp(-1/D)]^{-1}, \quad (54)$$

откуда при $D \gg 1$ получаем

$$D(0) \approx D = M_{\max}/M_n. \quad (55)$$

Выражения (54) и (55) можно использовать как приближения при $M \ll N$.

В статье не учтен внешний шум, т. е. шум усилителей, коммутаторов и т. п. В силу независимости внешнего шума его дисперсия (в сопоставимых единицах) должна быть добавлена к дисперсии, рассмотренной выше.

При выборе времени накопления (T) примерно равным длительности полезного сигнала (“согласованная фильтрация”) динамический диапазон может оказаться неприемлемо малым. В этом случае можно уменьшить время (аналогового) накопления и применить внешнее (дискретное) суммирование отсчетов. Эффективность такого метода существенно зависит от уровня внешнего шума.

Несложный, но громоздкий анализ влияния внешнего шума и неидеальности накопления не приводится в силу ограниченности объема статьи.

В заключение отметим, что в модели не учтено возможное в некоторых случаях ограничение числа накапливаемых зарядов. Учет этого фактора (математически это задача с ограниченным числом состояний) заслуживает специального исследования.

ЛИТЕРАТУРА

1. Кингман Дж. Пуассоновские процессы. М.: Изд. МЦНМО, 2007. 134 с.
2. Вентцель Е.С., Овчаров Л.А. Теория случайных процессов и ее инженерные приложения. М.: Наука, 1991. 382 с.
3. Рывкин С.М. Фотоэлектрические явления в полупроводниках. М.: Физматгиз, 1963. 496 с.
4. Трищенко М.А. Фотоприемные устройства и ПЗС. М.: Радио и связь, 1992. 400 с.
5. Ковчавцев А.П. Структуры металл-диэлектрик-полупроводник на основе арсенида индия // Автореферат докт. дис. Новосибирск: Изд. ФГУП “НИИСистем”, 2003. 46 с.