

МАТЕМАТИЧНЕ МОДЕЛЮВАННЯ ТА ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ

УДК 621.395.7

Б.А. Мандзій, А. Бенч

Національний університет "Львівська політехніка"

ЕНЕРГЕТИЧНІ СПЕКТРИ СИГНАЛІВ, ЩО ПРЕДСТАВЛЯЮТЬ УКРАЇНОМОВНІ ТЕКСТИ

© Мандзій Б.А., Бенч А., 2009

Подано результати розрахунку енергетичних спектрів сигналів поширених типових кодів лінійних трактів цифрових систем пересилання інформації, що представляють українськомовні тексти, написані з використанням літер "кирилиці".

In the article the results of calculation of power spectrums of signals of widespread typical kodas of linear highways of the digital systems are given which present sending of information of, Ukrainian-language phototypographs written with the use of letters of cyrillic "alphabet".

Постановка завдання

Високі швидкості пересилання інформації в сучасних цифрових телекомунікаційних системах вимагають застосування відповідних кодів лінійних трактів, які задовольняли б вимогу узгодження спектра інформаційних сигналів з частотними характеристиками каналу зв'язку та уможлилювали б пересилання сигналів з мінімальними спотвореннями. Згадані коди повинні забезпечувати т. зв. компактність спектра, тобто відсутність постійної складової у спектрі та високочастотних складових за межами виділеної смуги пропускання каналу зв'язку, вони повинні мати властивість самосинхронізації, що особливо важливо під час пересилання довгих послідовностей "нулів" та "одиниць".

Сьогодні відомо багато різноманітних кодів лінійних трактів (цифрових форматів), які застосовуються у цифрових системах пересилання інформації [1]. Аналіз спектральних характеристик закодованих сигналів проводять, розглядаючи їх як випадкові послідовності імпульсів однакової форми (здебільшого прямокутної), які можуть відрізнятися тривалостями, амплітудами або полярностями, або одночасно амплітудами і полярностями. Кожній випадковій послідовності імпульсів відповідає своя дискретна послідовність чисел, яка залежить від використаного коду. Основною характеристикою таких випадкових послідовностей імпульсів є енергетичний спектр, який розраховують на підставі гіпотези про ергодичність дискретного випадкового процесу та про однакову апріорну ймовірність появи різних бінарних символів (нуля і одиниці) у кожному біті первинної бінарної послідовності, яка підлягає кодуванню [2, 3]. Такий підхід дає змогу достатньо просто визначити функцію автокореляції закодованої дискретної послідовності, а відтак на її підставі – енергетичний спектр. Розраховані в такий спосіб спектри характеризують достатньо тривалі послідовності імпульсних сигналів типових кодів, в яких імовірності появи бінарних символів в окремих бітах є взаємно незалежними.

Пересилання текстових повідомлень має ту особливість, що кожна конкретна мова характеризується певними імовірностями появи окремих літер в тексті, які являють собою стандартизовані набори бінарних символів. Практичний інтерес являє розрахунок енергетичних спектрів сигналів, які представляють українськомовні тексти, написані "кирилицею" і закодовані типовими

цифровими форматами. У цій роботі описано можливий спосіб розрахунку згаданих спектрів та наведено результати їх розрахунку для поширених кодів.

Математичні моделі спектральних густин та енергетичних спектрів сигналів типових кодів, які представляють літери "кирилиці"

Як вихідну інформацію прийнято кодову таблицю CP1251[4], в якій подано восьмибітові бінарні коди літер "кирилиці" та дані про частоту вживання літер українського алфавіту, наведені у [5].

У таблиці подано бінарні коди літер "кирилиці" та частоту їх вживання в українськомовних текстах.

Бінарні коди літер "кирилиці"

№ з/п	Літера (символ)	Частота вживання n	Бінарний код								Код 2BQ1			
			x ₀	x ₁	x ₂	x ₃	x ₄	x ₅	x ₆	x ₇	x ₀	x ₁	x ₂	x ₃
1	А	0,072	1	1	0	0	0	0	0	0	1	-3	-3	-3
2	Б	0,017	1	1	0	0	0	0	0	1	1	-3	-3	-1
3	В	0,052	1	1	0	0	0	0	1	0	1	-3	-3	3
4	Г	0,016	1	1	0	0	0	0	1	1	1	-3	-3	1
5	Д	0,035	1	1	0	0	0	1	0	0	1	-3	-1	-3
6	Е	0,017	1	1	0	0	0	1	0	1	1	-3	-1	-1
7	Є	0,008	1	1	0	0	1	0	1	0	3	3	3	3
8	Ж	0,009	1	1	0	0	0	1	1	0	1	-3	-1	3
9	З	0,023	1	1	0	0	0	1	1	1	1	-3	-1	1
10	И	0,061	1	1	0	0	1	0	0	0	1	-3	3	-3
11	Й	0,008	1	1	0	0	1	0	0	1	1	-3	3	-1
12	І	0,057	1	0	1	1	0	0	1	0	3	1	-3	3
13	Ї	0,006	1	0	1	1	1	1	1	1	3	3	1	1
14	К	0,035	1	1	0	0	1	0	1	0	1	-3	3	3
15	Л	0,036	1	1	0	0	1	0	1	1	1	-3	3	1
16	М	0,031	1	1	0	0	1	1	0	0	1	-3	1	-3
17	Н	0,065	1	1	0	0	1	1	0	1	1	-3	1	-1
18	О	0,094	1	1	0	0	1	1	1	0	1	-3	1	3
19	П	0,029	1	1	0	0	1	1	1	1	1	-3	1	1
20	Р	0,047	1	1	0	1	0	0	0	0	1	-1	-3	-3
21	С	0,041	1	1	0	1	0	0	0	1	1	-1	-3	-1
22	Т	0,055	1	1	0	1	0	0	1	0	1	-1	-3	3
23	У	0,040	1	1	0	1	0	0	1	1	1	-1	-3	+1
24	Ф	0,001	1	1	0	1	0	1	0	0	1	-1	-1	-3
25	Х	0,012	1	1	0	1	0	1	0	1	1	-1	-1	-1
26	Ц	0,006	1	1	0	1	0	1	1	0	1	-1	-1	3
27	Ч	0,018	1	1	0	1	0	1	1	1	1	-1	-1	1
28	Ш	0,012	1	1	0	1	1	0	0	0	1	-1	3	-3
29	Щ	0,001	1	1	0	1	1	0	0	1	1	-1	3	-1
30	Ь	0,029	1	1	0	1	1	1	0	0	1	-1	1	-3
31	Ю	0,004	1	1	0	1	1	1	1	0	1	-1	1	1
32	Я	0,029	1	1	0	1	1	1	1	1	1	-1	1	1
33	Пробіл –	0,017	1	0	1	1	0	1	-	-	3	3	3	x
34	Інші симв.	0,017	1	0	1	0	1	0	-	-	3	1	-1	x

Σ 1,00

Оскільки кожна літера представлена восьмибітовою послідовністю бінарних символів, то математичну модель комплексної спектральної густини сигналу, що представляє довільну літеру, можна подати у вигляді:

$$\dot{G}(\omega) = \dot{G}_1(\omega) \cdot \sum_{k=0}^7 x_k \cdot e^{-jk\omega\tau} + \dot{G}_0(\omega) \cdot \sum_{k=0}^7 (1-x_k) \cdot e^{jk\omega\tau}, \quad (1)$$

де $\dot{G}_1(\omega)$ та $\dot{G}_0(\omega)$ – комплексні спектральні густини елементарних (базових) сигналів, які представляють бінарні символи 1 та 0 відповідно; τ – тривалість одного біта; x_k – змінна, яка знаходиться на k -ій позиції восьмибітової послідовності і приймає значення бінарного символу, що знаходиться на k -ій позиції (див. таблицю).

Практично в усіх випадках елементарні (базові) сигнали, що представляють бінарні символи, є імпульсами прямокутної форми, які можуть змінювати амплітудні значення та полярність залежно від вибраного типового коду (цифрового формату).

Далі наведено формули для розрахунку амплітудних спектрів (модулів комплексної спектральної густини) поширених типових кодів, де значення амплітуди імпульсів позначено як A .

Уніполярний код:

$$\left| \dot{G}_{УП}(\omega) \right| = \frac{2A}{\omega} \cdot \left| \sin\left(\frac{\omega\tau}{2}\right) \right| \cdot \sqrt{\left[\sum_{k=0}^7 x_k \cos(k\omega\tau) \right]^2 + \left[\sum_{k=0}^7 x_k \sin(k\omega\tau) \right]^2}. \quad (2)$$

Код RZ з пасивною паузою:

$$\left| \dot{G}_{RZ}(\omega) \right| = \frac{2A}{\omega} \cdot \left| \sin\left(\frac{\omega\tau}{4}\right) \right| \cdot \sqrt{\left[\sum_{k=0}^7 x_k \cos(k\omega\tau) \right]^2 + \left[\sum_{k=0}^7 x_k \sin(k\omega\tau) \right]^2}. \quad (3)$$

Манчестерський код:

$$\left| \dot{G}_{МЧ}(\omega) \right| = \frac{A}{\omega} \cdot \sqrt{\left[1 - 2\cos\left(\frac{\omega\tau}{2}\right) + \cos(\omega\tau) \right]^2 + \left[2\sin\left(\frac{\omega\tau}{2}\right) - \sin(\omega\tau) \right]^2} \times \\ \times \sqrt{\left[\sum_{k=0}^7 (2x_k - 1) \cdot \cos(k\omega\tau) \right]^2 + \left[\sum_{k=0}^7 (2x_k - 1) \cdot \sin(k\omega\tau) \right]^2}. \quad (4)$$

Код АМІ:

$$\left| \dot{G}_{АМІ}(\omega) \right| = \frac{2A}{\omega} \cdot \left| \sin\left(\frac{\omega\tau}{2}\right) \right| \cdot \sqrt{\left[\sum_{k=0}^7 x_k \cos(k\omega\tau) \right]^2 + \left[\sum_{k=0}^7 x_k \sin(k\omega\tau) \right]^2}. \quad (5)$$

Біполярний код:

$$\left| \dot{G}_{БП}(\omega) \right| = \frac{2A}{\omega} \cdot \left| \sin\left(\frac{\omega\tau}{2}\right) \right| \cdot \sqrt{\left[\sum_{k=0}^7 (2x_k - 1) \cos(k\omega\tau) \right]^2 + \left[\sum_{k=0}^7 (2x_k - 1) \sin(k\omega\tau) \right]^2}. \quad (6)$$

Код 2BQ1:

$$\left| \dot{G}_{BQ}(\omega) \right| = \frac{2A}{\omega} \cdot \left| \sin(\omega\tau) \right| \cdot \sqrt{\left[\sum_{k=0}^3 x_k \cos(2k\omega\tau) \right]^2 + \left[\sum_{k=0}^3 x_k \sin(2k\omega\tau) \right]^2}. \quad (7)$$

У формулі (5) для коду АМІ змінні x_k , значення яких не дорівнюють нулеві, по черзі змінюють свій знак на протилежний (+, -, +, - тощо).

У випадку коду 2BQ1 первинна восьмибітова бінарна послідовність розбивається на чотири групи із двох символів, кожна з яких замінюється одним із чотирьох можливих рівнів амплітуди: -3A, -A, A, 3A, як подано в таблиці. Тому у цьому коді використано лише чотири змінні: x_0, x_1, x_2, x_3 .

Проте у цьому випадку тривалість одного біта подвоюється порівняно з тривалістю біта первинної бінарної послідовності.

Енергетичний спектр $W(\omega)$ пропорційний до квадрата амплітудного спектра:

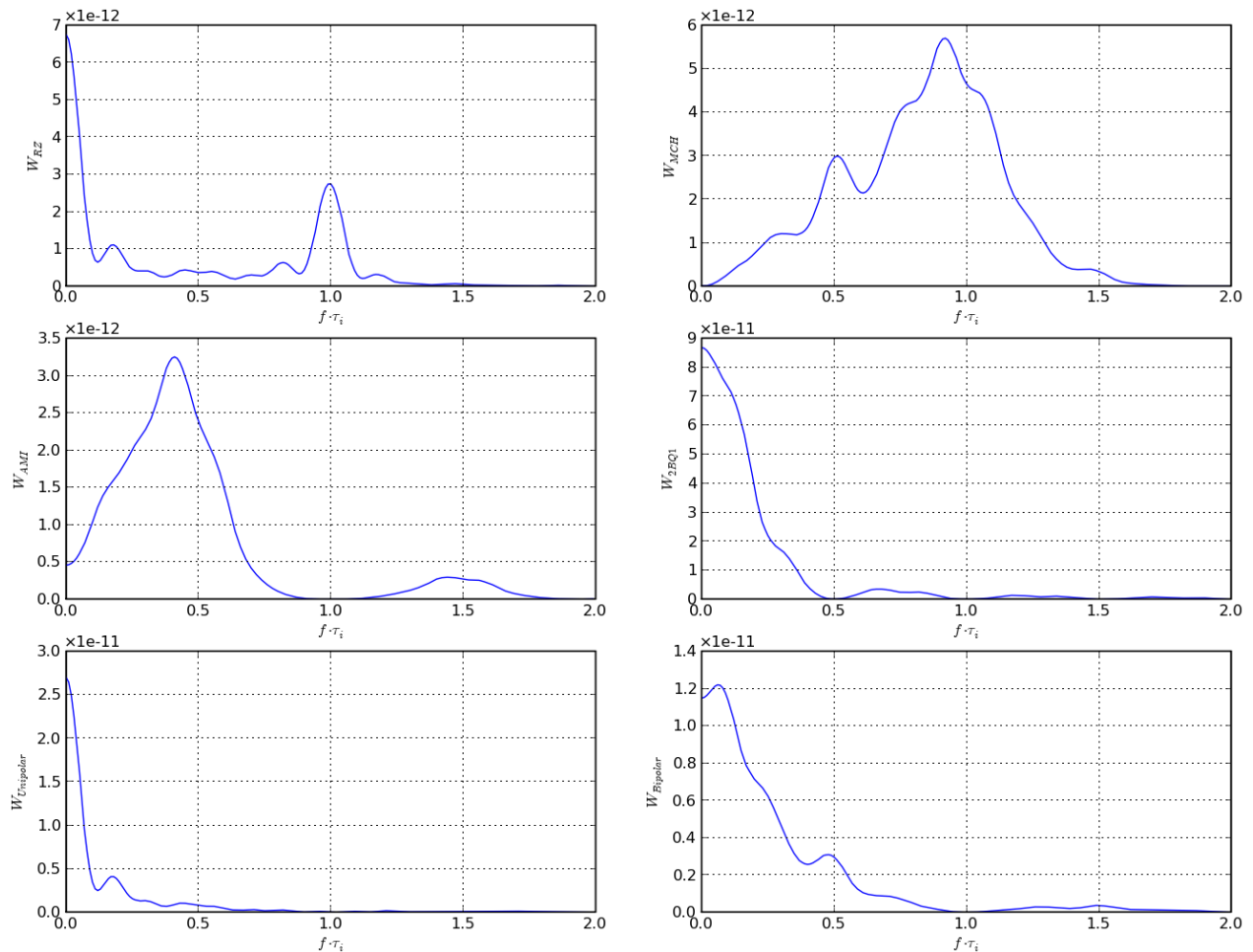
$$W(\omega) = |\dot{G}(\omega)|^2. \quad (8)$$

Під час розрахунку енергетичного спектра $W_{\Sigma}(\omega)$ сигналів, які відповідають у вибраному цифровому форматі україномовному текстові, написаному "кирилицею", потрібно підсумувати усі енергетичні спектри сигналів, які представляють літери (символи) у вибраному цифровому форматі з урахуванням частоти їх вживання:

$$W_{\Sigma}(\omega) = \sum_{s=1}^{34} n_i W_i(\omega), \quad (9)$$

де $W_i(\omega)$ – енергетичний спектр сигналу, що представляє i -ту літеру (символ) у вибраному цифровому форматі; n_i – частота вживання i -ї літери (символу).

На рисунку зображені енергетичні спектри сигналів типових кодів, розраховані на підставі формул (2)–(9).



Енергетичні спектри сигналів типових кодів

Енергетичні спектри сигналів, що представляють україномовні тексти з використанням типових кодів: а – RZ; б– Манчестерського; в – АМІ; г – 2ВQ1; д – уніполярного; е – біполярного.

Висновки

Аналіз отриманих результатів дає підставу зробити такі висновки:

1. Найвужчу смугу частот займають сигнали кодів уніполярного та 2ВQ1, проте в їхньому спектрі присутня значна постійна складова. Це пояснюється тим, що бінарні послідовності, які представляють часто вживані літери українського алфавіту, містять значну кількість одиниць.

2. Незначну постійну складову містять спектри сигналів кодів АМІ та Манчестерського, проте вони займають смугу частот, яка приблизно у 1,5...2 рази ширша від ширини спектра сигналів кодів уніполярного та 2BQ1.

3. Енергетичний спектр сигналів коду RZ має два явно виражені максимуми, які зосереджені біля постійної складової та на частоті $f=1/\tau_6$, і займає смугу частот, яка приблизно у 1,5 раза ширша, ніж у випадку кодів уніполярного та 2BQ1.

4. Енергетичний спектр сигналів біполярного коду містить значну постійну складову і зосереджений в діапазоні частот від 0 до частоти $f=1/\tau_6$.

Наведені результати можуть бути корисні під час аналізу процесів пересилання україномовних текстів з використанням типових кодів лінійних трактів сучасних телекомунікаційних систем.

1. Системи телекомунікацій: Підручник для ВНЗ / М.І. Мазурков, В.І. Правда, П.Ю. Баранов, І.М. Єремичой, В.Я. Чечельницький; За ред. М.І. Мазуркова і В.І. Правди. – Одеса: ТЕС, 2005. – 288 с. 2. Шварцман В.О., Ємельянов Г.А. Теорія передачі дискретної інформації: Учебник для вузів зв'язи. – М.: Связь, 1979. – 424 с. 3. Wesolowski K. Podstawy cyfrowych systemów telekomunikacyjnych. WKŁ. – Warszawa, 2003. – 408 s. 4. [http://msdn.Microsoft.com/uk-ua/global/cc305144\(en-us\).aspx](http://msdn.Microsoft.com/uk-ua/global/cc305144(en-us).aspx). 5. Конфорович А.Г. Добрий день, Архімеде: Цікаві задачі, ігри, головоломки. – К.: Молодь, 1988. – 152 с.

УДК 621.372.061

Ю.І. Шаповалов, С.В. Маньковський
Національний університет “Львівська політехніка”

ДВА МЕТОДИ ФОРМУВАННЯ ДИФЕРЕНЦІАЛЬНИХ РІВНЯНЬ ЛІНІЙНИХ ПАРАМЕТРИЧНИХ КІЛ У СИМВОЛЬНОМУ ВИГЛЯДІ

© Шаповалов Ю.І., Маньковський С.В., 2009

Для тестового прикладу розглянуто два методи формування символьних диференціальних рівнянь лінійних параметричних кіл. Проведено порівняння цих методів. Отримано миттєві значення вихідної напруги тестового кола символьним методом та за допомогою MicroCap.

This paper considers functioning of two methods of forming differential equations in symbolic form by the test linear parametric circuit. The results of performing these methods are compared. The momentary output voltage values are obtained by the symbolic method and by the MicroCap.

Вступ

Однією з важливих задач аналізу лінійних параметричних кіл є формування диференціального рівняння, яке пов'язує вхідний та вихідний сигнали заданого кола і має такий вигляд:

$$a_n(t)y^{(n)} + a_{n-1}(t)y^{(n-1)} + \dots + a_0(t)y = b_m(t)x^{(m)} + b_{m-1}(t)x^{(m-1)} + \dots + b_0(t)x, \quad (1)$$

де x, y – вхідний та вихідний сигнали відповідно; t – незалежна змінна (час); $a_i(t), b_j(t)$ – залежні від часу t коефіцієнти, які визначаються параметрами та структурою заданого параметричного кола.

У [1, 2] наведено два методи формування рівнянь (1) у символьному вигляді. Перший метод [1] ґрунтується на алгебраїчному виключенні, а другий [2] – на виключенні матричними чи топологічними методами [3] усіх змінних у рівняннях, що описують задане лінійне параметричне коло, крім y, x та їх похідних. Особливість цих двох методів полягає у тому, що вони застосовані