УДК 004.04

Н.Г. ШИРМОВСЬКА, Т.О. ВАВРИК, Г.І. ЛЕВИЦЬКА, І.Р. МИХАЙЛЮК
Івано-Франківський національний технічний університет нафти і газу

ТЕОРЕТИЧНІ ЗАСАДИ ТА ОБҐРУНТУВАННЯ ІНФОРМАЦІЙНИХ ХАРАКТЕРИСТИК ДИСКРЕТИЗАЦІЇ ФОРМУВАННЯ ЦИФРОВИХ ПОВІДОМЛЕНЬ НА ОСНОВІ АВТОКОРЕЛЯЦІЙНИХ ФУНКЦІЙ У БАЗИСІ РАДЕМАХЕРА

У статті викладено теоретичні засади та проведено дослідження формування цифрових відліків на основі кореляційних функцій у теоретико-числовому базисі Радемахера, що дозволило обґрунтувати характеристики переваг та недоліків використання базису Радемахера в задачі ідентифікації станів об'єкта.

Ключові слова: теоретико-числовий базис, дискретизація, автокореляційна функція, інформаційний потік.

N.G. SHYRMOVSKA, T.A. VAVRYK, G.I. LEVYTSKA, I.R. MUHAJLUK
Ivano-Frankivsk National Technical University of Oil and Gas

THEORETICAL SUBSTANCES AND REASONING OF INFORMATION CHARACTERISTICS OF DISCRETION OF FORMATION OF DIGITAL MESSAGES BASED ON AUTOCORRELLATION FUNCTIONS IN THE RADEMAKER BASIS

Theoretical numerical bases are the fundamental theoretical foundations of numerical systems and data encoding methods. The choice of the basis function is performed depending on the systemic characteristics of the various communication channels and the operating conditions of the computer systems. In modern computer and telecommunication systems, widely used are theoretical numerical bases on the foundation of discrete piecewise linear functions, which provide a drastically simpler implementation of digital generators, as well as simplify algorithms for digital signal reception. This article describes the theoretical principles and conducts research on the formation of digital samples based on the correlation functions in the theoretic-numerical basis of Rademacher, which allowed to substantiate the characteristics of advantages and disadvantages for using the Rademacher basis in the problem of identifying the states of the object. Presented theoretical principles and research on the formation of digital samples, based on the correlation functions in the theoretical numerical basis of Rademacher, allowed to prove the characteristics of the advantages and disadvantages of using the Rademacher basis in the problem of identifying the states of the object, and they are an important reserve for the development of the theory and the improvement of diagnostic tools for the quasi-stationary objects.

Keywords: theoretical numerical basis, discretization, autocorrelation function, information flow.

Комп'ютерні системи та мережі широко впроваджуються у всі сфери суспільства. Важливим при цьому є створення проблемно-орієнтовних та спеціалізованих розподілених комп'ютерних систем для різних об'єктів і галузей промисловості. Актуальними задачами, що вирішують системи даного класу є оперативне діагностування технологічних об'єктів, які характеризуються нестаціонарністю, багатопараметричністю, екологічною небезпечністю, вибухонебезпечністю та ін. Особливо важливою задачею є прогнозування, передбачення та діагностування передаварійних та аварійних станів об'єктів управління (ОУ).

1. Теоретичні засади кодування інформаційних потоків у теоретико-числовому базисі Радемахера

В сучасних розподілених комп'ютерних системах реального часу найбільш широко використовуються методи формування та кодування інформаційних потоків у двійковій системі числення, тобто теоретико-числовому базисі (ТЧБ) Радемахера [1, 2]. При цьому, вхідні аналогові сигнали багатоканального аналого-цифрового перетворювача (БАЦП) перетворюються в інформаційні потоки даних. В залежності від використання різних типів БАЦП цифрові відліки X_{ij}, де i – часовий номер відліка, j – номер каналу, можуть формуватися одночасно або послідовно на деякому інтервалі дискретизації Δt (рис. 1).

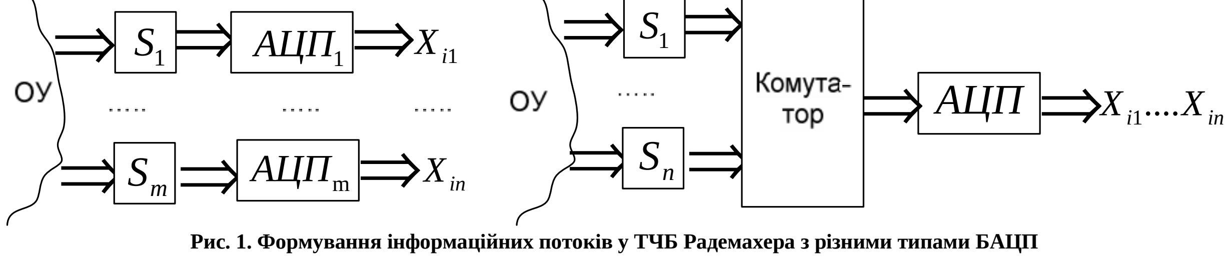

Рис. 1. Формування інформаційних потоків у ТЧБ Радемахера з різними типами БАЦП

При цьому формується інформаційна матриця цифрових відліків, які описують стан ОУ в деякий момент часу:

$$\begin{vmatrix} X_{11} & X_{21} & ... & X_{i1} & ... & X_{n1} \\ X_{12} & X_{22} & ... & X_{i2} & ... & X_{n2} \\ ... & ... & ... & ... & ... & ... \\ X_{1j} & X_{2j} & ... & X_{ij} & ... & X_{nj} \\ ... & ... & ... & ... & ... & ... \\ X_{1m} & X_{2m} & ... & X_{im} & ... & X_{nm} \end{vmatrix}.$$

На рис. 2 показаний приклад формування цифрових відліків.

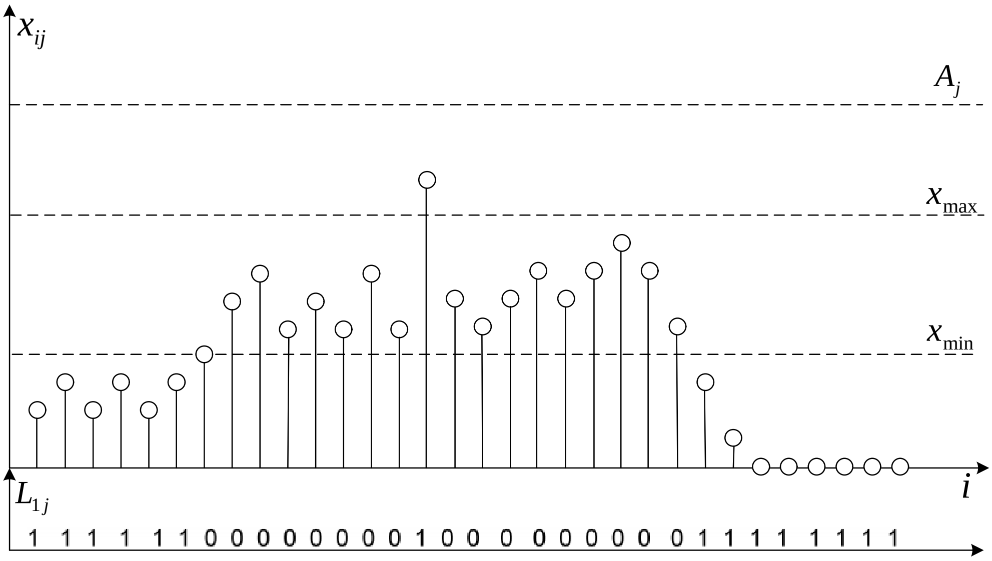

Рис. 2. Формування цифрових відліків

Операція ідентифікації стану ОУ на основі логіко-статистичних інформаційних моделей (ЛСІМ-1) у ТЧБ Радемахера виконується згідно виразів:

$$L_{1j} = 0, x_{\max} - x_{ij} > 0 \wedge x_{ij} - x_{\min} > 0;$$
$$L_{1j} = 1, x_{\max} - x_{ij} \le 0 \wedge x_{ij} - x_{\min} \le 0. \qquad (1)$$

В той же час в процесорах Радемахера, з метою спрощення його апаратної складності, операція замінена операцією додавання доповнюючих кодів, тобто вираз (1) отримає вигляд:

$$L_{1j} = 0, [x_{\max}]_{\partial} + [x_{ij}]_{\partial} > 1 \wedge [x_{ij}]_{\partial} + [x_{\min}]_{\partial} > 1;$$
$$L_{1j} = 1, [x_{\max}]_{\partial} + [x_{ij}]_{\partial} \le 1 \wedge [x_{ij}]_{\partial} + [x_{\min}]_{\partial} \le 1,$$

що потребує додаткового представлення двійкових кодів ТЧБ Радемахера в доповнюючому коді шляхом інвертування кодів від'ємних значень та додавання до їхнього розряду 1.

Крім того, великий рівень шумів та різнохарактерних впливів на мікроелектронні засоби показує, що кодові значення шумів можуть складати 20-30% істинних значень контрольованих параметрів буріння [3]. Цей недолік практично унеможливлює ефективність діагностування та виявлення передаварійних та аварійних ситуацій на об'єктах, що потребує використання інтегрованих характеристик вибіркового та ковзного математичних сподівань, а також розрахунку нормальних коефіцієнтів взаємокореляції між параметрами для побудови кластерних моделей.

2. Дослідження інформаційних характеристик дискретизації формування цифрових повідомлень на основі автокореляційних функцій у базисі Радемахера

Аналіз інформаційної технології кодування даних у ТЧБ Радемахера дозволяє сформулювати наступні функціональні обмеження та недоліки цього базису. Недоліками ТЧБ Радемахера є:

Великий об'єм інформації даних, які опрацьовуються системою з ентропією об'єму, яка обчислюється за формулою Хартлі:

$$I_{OY} = n \sum_{j=1}^{m} \widehat{E}[\log_2 A_j],$$
$$0 \le x_j \le A_j.$$

Наприклад, $0 \le x_j \le 1023, n = 128, I_{OY} = 128 \times 10 = 1028 б/с \approx 1 кбіт/с.$

Різна інформативність кожного інформаційного біта цифрового відліку x_{ij}. Тобто при виникненні помилки в одному з бітів відбувається різний рівень втрат інформації.

$$[x_i]_R = a_{k-1}2^{n-1} + a_{k-2}2^{n-2} + ... + a_1 2^1 + a_0.$$

Наприклад, при $K = 10$ втрата інформації в молодшому біті складає $0{,}001\%$, а втрата старшого біта $0{,}5\%$.

Абсолютна незахищеність кодів ТЧБ Радемахера від помилок, оскільки кожний біт відліку є інформативним.

Надлишковість кодування інформації, що приводить до зростання об'ємів інформаційних потоків, які опрацьовує комп'ютерна система. Це обумовлено тим, що діапазон квантування в БАЦП вибирається відповідно до максимального значення одного з контрольованих параметрів, в той час, коли багато параметрів мають менший діапазон квантування.

Наявність надлишкової інформації небажане, тому що канал зв'язку зайнятий на більш довгий час, ніж необхідно, або ж потрібні канали із більшою пропускною здатністю. При зберіганні надлишкової інформації збільшується необхідний об'єм пам'яті, зростає час пошуку та зчитування даних, а при обробці даних зростає час обробки.

Якість способу дискретизації, згідно якого відбираються відліки, слід оцінювати за похибкою, з якою вдається відновити вихідну функцію. Критеріями, в яких похибка відновлення сигналу наближується до нуля, є:

- частотний критерій В.А. Котельникова, в якому інтервали між відліками вибирають із врахуванням частотного спектра сигналу, що підлягає дискретизації;
- кореляційний критерій Н.А. Желєзнова, що встановлює зв'язок інтервалів між відліками з інтервалом кореляції сигналу;
- квантовий критерій Ф.Е. Темнікова, що вказує на залежність інтервалів між відліками за степенем квантування за рівнем та крутизною (першої похідної) сигналу.

Частіше за все при оцінці відхилення функції $Y(t)$, що відтворюється, від вихідного сигналу $X(t)$ на кожному інтервалі дискретизації $[0,T]$ використовують такі критерії: критерій найбільшого відхилення, середньоквадратичний, інтегральний, ймовірнісно-зональний. Як правило, неперервні функції не можуть мати обмежений спектр, тому, для сигналів обмеженої протяжності теорема Котельникова не може бути застосована. Критерій Желєзнова, що вводить поняття про середні характеристики сигналу, обмежує кінцеву точність відновлення функції структурою кореляційної функції та відношенням $t(оср)/t$. Квантовий критерій Темнікова зводить похибку відновлення вихідної неперервної функції з обмеженими похідними поліномами Лагранжа до залишкового члена.

У більшості практичних випадків оптимальна обробка сигналів є ускладненою, це потребує розробки методів оцінки точності представлення даних. Оптимальною будемо називати дискретизацію, в процесі якої вихідна функція із заданою точністю може бути представлена мінімальною кількістю відліків $X(t)$. В цьому випадку всі відліки є суттєвими для відновлення вихідної інформації. Розглянемо дискретизацію випадкових стаціонарних сигналів та синтез розкладу на основі узагальненого ряду Фур'є. Вважатимемо, що кореляційна функція випадкового сигналу $R(t)$ відома, а ефективний інтервал кореляції $\tau_0 << T$, де $[0,T]$ – час існування реалізації сигналу. Як критерій вірності використаємо функціонал [1]:

$$\nu = \frac{1}{T}\int_0^T M\left\{\Delta^2(t)\right\}dt\,, \tag{2}$$

де $\Delta(t) = f(t) - f_N(t)$ – похибка при заміні сигналу $f(t)$ функцією розкладу $f_N(t)$. Будемо вважати, що функція $f(t)$ інтегрується разом із своїм квадратом на інтервалі $[0,T]$. Мінімум середньо квадратичної похибки при обмеженому числі членів розкладу $N = T/\tau$ (τ – інтервал дискретизації) досягнутий у випадку представлення функції $f(t)$ відрізком ряду Фур'є в системі ортогональних функцій $\{\varphi_k(t)\}$:

$$f_N(t) = \sum_{k=0}^{N-1} a_k \varphi_k(t) \tag{3}$$

Без втрати сутності можна рахувати математичним очікуванням сигналу $M\{f(t)\} = 0$, а його нормою $|f(t)| = 1$. Підставляючи (2) в (1), отримаємо:

$$\nu = 1 - \frac{2}{\tau}\int_0^\tau \varphi(t)\int_0^\tau \varphi(x)R(t,x)dxdt + \int_0^\tau \varphi^2(t)\int_0^\tau \varphi(x)\int_0^\tau \varphi(y)R(x,y)dxdydt \tag{4}$$

Мінімум (3) з врахуванням показника якості $\lambda = \lambda_1$, де $\lambda_i < \lambda_{i+1}$, $i = \overline{1,m}$ на основі (4) буде таким:

$$\nu = 1 - \frac{1}{\tau \times \lambda_1}.$$

Тоді мінімальна похибка, що досягається при розкладі сигналу $f(t)$ на обмеженому інтервалі

$i = \overline{1,n}$:

$$\nu = 1 - \frac{1}{\tau}\sum_{i=1}^{n}\frac{1}{\lambda_i}.$$

Зрозуміло, що для будь-яких функцій розкладу випадкових сигналів з необмеженим спектром помилка розкладу залежить від виду функції кореляції $R(t)$ та інтервалу дискретизації. Порівнювати функції розкладу можна за коефіцієнтом надлишковості $\gamma = N / N_0$, де $N_0 = T / \tau_0$ – число некорельованих в середньому відліків, $N = T / \tau$ – число відліків, яке необхідне для досягнення заданої похибки. З врахуванням N та N_0 коефіцієнт надлишковості буде рівний $\gamma = \tau / \tau_0$. Найкращим слід вважати розклад, який при однаковій похибці вимагає меншого коефіцієнта надлишковості. Найбільш простим є визначення критерію вірності при:

$$R(t) = \begin{vmatrix} 1 - \frac{|t|}{\tau_0}, & |t| < \tau_0; \\ 0, & \textit{при} \;\; \textit{інших} \;\; \mathrm{t} \end{vmatrix}.$$

Для обміну даними в рамках інформаційної системи бурового виробництва (ІСБВ), а також організації діалогу з операторами термінальних пунктів управління технологічними процесами на рівні бурової, з'являється вимога видачі інформації планово-економічного та технологічного характеру в стандартному коді представлення алфавітно-цифрової інформації в реальному часі. На перешкоді стала величезна надлишковість інформації, що поступала переважно в аналоговому вигляді. З метою оптимізації відбору інформації, компактного представлення, передачі та зберігання, введений характеристичний параметр L, що являє собою зручну оцінку втрати повідомлень, зумовлених дискретизацією процесу. На основі оцифрованих значень параметрів технологічного процесу буріння, що підлягали контролю та запису на самопишучі прилади безпосередньо на буровій, були побудовані графіки нормованої автокореляційної функції процесів $R_{xx}(t)$. Побудова виконувалася з наперед надлишково заданих значень $Y(t)$, $t = 1{,}900$, а кількість $R_{xx}(t)$, $t = 1, k$, де k – число точок автокореляційної функції, причому $R_{xx}(k) < 0.01 R_{xx}(0)$. Тоді можна записати:

$$L = \frac{\sqrt{2}}{\Delta\xi}\sqrt{1 - R_{xx}(t)}\,, \qquad (5)$$

де $\Delta\zeta = \Delta y / \sigma$ – приведене (нормоване) значення запасу норми;

$\Delta y = y_{t+1} - y_t$ – приріст поточної ординати по відношенню до наступної на інтервалі дискретизації;

σ – середньоквадратичне відхилення параметра, що контролюється.

На підставі (5), коли відома автокореляційна функція $R_{xx}(t)$ або хоча б її початкова ділянка, з врахуванням мінімально допустимої похибки [2], знаходимо $y_{доп}$ – допустимий крок дискретизації, як абсцису перетину графіка $R_{xx}(t)$ з прямою, паралельною вісі абсцис, віддаленою від дотичної $R_{xx}(1)$ на значення 0.045Δy.

Оскільки всі етапи обчислень виконуються на комп'ютері, розглянемо аналітичний підхід визначення кроку дискретизації з графіка автокореляційної функції, показаного на рис. 3.

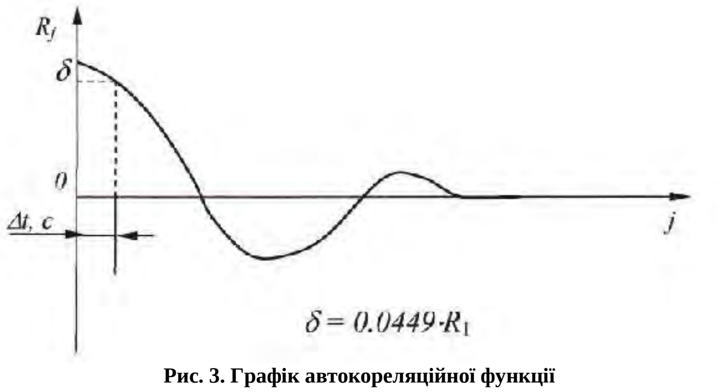

Рис. 3. Графік автокореляційної функції

Перш за все визначаємо, в межах якого інтервалу точок $R_{xx}(t)$, $R_{xx}(t+1)$ знаходиться величина $y = R_{xx}(1) - 0.045 R_{xx}(1)$. Запишемо рівняння прямої, яка проходить через точки визначеного інтервалу $R_{xx}(t)$, $R_{xx}(t+1)$:

$$\frac{t-t_i}{t_{i+1}-t_i}=\frac{R_{xx}-R_{xx}(i)}{R_{xx}(i+1)-R_{xx}(i)}.$$

Підставивши замість R_{xx} значення кроку дискретизації y, отримаємо:

$$\Delta t=\frac{\delta-R_{xx}(i)(t_{i+1}-t_i)}{|R_{i+1}-R_i|}+t_i=\frac{0.955\times R_{xx}(1)-R_{xx}(i)(t_{i+1}-t_i)}{|R_{i+1}-R_i|}+t_i.$$

При вхідних даних початок нумерації матриць, кількість точок дискретизації, ранжована змінна, відповідають значенням $ORIGIN=1$, $n=24$, $i=1,2....n$.

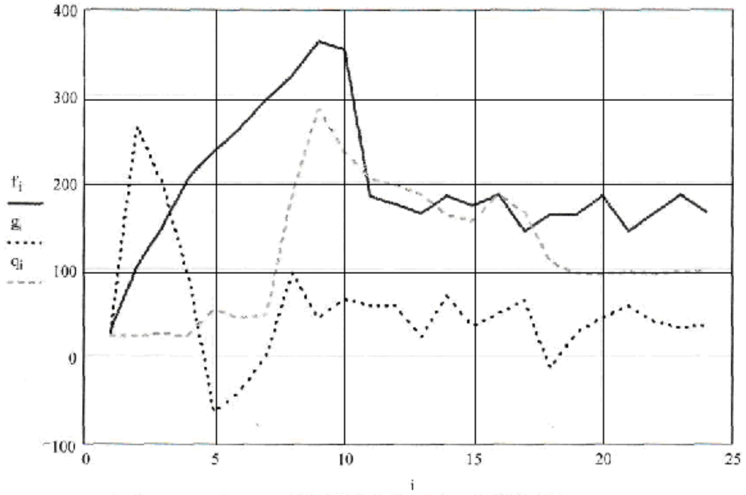

Рис. 4. Графіки зміни параметрів, що досліджуються в часі

Обчислюємо математичні сподівання цих функцій:

$$M(f)=\frac{1}{n}\times\sum_{i=1}^{n} f_i\ ;$$

$M(f)=196.625$; $M(g)=52.875$; $M(q)=121.042$.

Визначаємо центровані значення координат для цих функцій:

$F_i=f_i-M(f)$; $G_i=g_i-M(g)$; $Q_i=q_i-M(q)$.

Обчислюємо значення автокореляційних функцій: $j=1,2..n-1$

$$R_j=\frac{1}{n-j}\times\sum_{i=1}^{n-j}(F_i\times F_{i+j});$$

$$K_j=\frac{1}{n-j}\times\sum_{i=1}^{n-j}(G_i\times G_{i+j});$$

$$W_j=\frac{1}{n-j}\times\sum_{i=1}^{n-j}(Q_i\times Q_{i+j}).$$

В результаті розрахунків отримаємо:

$R_1=4.203\times10^3$; $K_1=1.735\times10^3$; $W_1=4.49\times10^3$.

Обчислюємо необхідну кількість значень автокореляційних функцій:

$$k(F)=\left|\begin{array}{l} i\leftarrow 1 \\ while\ \ F_i\ge 0.01\times F_1 \\ \quad i\leftarrow i+1 \\ i \end{array}\right.$$

$k(R)=4$

$k(K)=2$.

$k(W)=5$

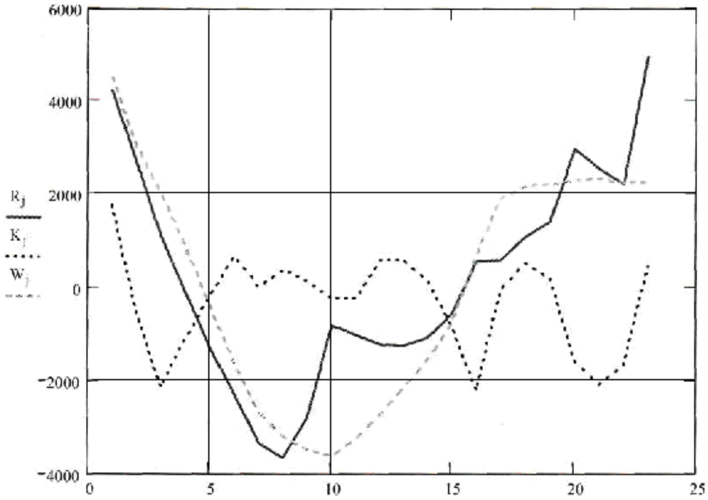

Рис. 5. Графіки автокореляційних функцій

Обчислюємо крок дискретизації для цих функцій:

$$\Delta t(F,k) = \left| \begin{array}{l} i \leftarrow 1 \\ t \leftarrow \dfrac{0.9551 \times F_1 - F_j}{\left|F_{j+1} - F_j\right|} + j \ \ if \ \ F_{j+1} < 0.9551 \times F_1 \\ t \end{array} \right. .$$

Результати розрахунку: $\Delta t(R, k(R)) = 7.519$; $\Delta t(K, k(K)) = 3.374$; $\Delta t(W, k(W)) = 8.8$.

В існуючих комп'ютерних системах оцифровування відбувається через рівні проміжки часу. Проте, в реальних аналогово-цифрових перетворювачах (АЦП) час моменту оцифровування схильний до відхилень через тремтінням фронту синхронізуючого сигналу (clock jitter). Вважаючи, що невизначеність моменту часу узяття відліку порядку A_t , одержуємо, що помилка, зумовлена цим явищем, може бути оцінена як

$$E_{a\delta} \le \left|\delta'(t)\Delta t\right| \le 2 \times A \times \pi \times f_0 \times \Delta t \ .$$

Легко побачити, що помилка відносно невелика на низьких частотах, проте на великих частотах вона може істотно зрости.

Ефект апертурної похибки може бути проігнорований, якщо її величина порівняно невелика в порівнянні з помилкою квантування. Таким чином, можна встановити наступні вимоги до тремтіння фронту сигналу синхронізації:

$$\Delta t < \frac{1}{2^n \times \pi \times f_0} \ ,$$

де π – розрядність АЦП; Δt – час тремтіння сигналу; f_0 – частота сигналу.

Таблиця 1

Вимоги до часу тремтіння фронту сигналу

Розрядність АЦП	Максимальна частота вхідного сигналу				
	44.1 кГц	192 кГц	1 Мгц	10 Мгц	100 Мгц
8	28.2 нс	6.48 нс	1.24 нс	124 пс	12.4 пс
10	7.05 нс	1.62 нс	311 пс	31.1 пс	3.11 пс
12	1.76 нс	405 пс	77.7 пс	7.77 пс	777 фс
14	441 пс	101 пс	19.4 пс	1.94 пс	194 фс
16	441 пс	101 пс	19.4 пс	1.94 пс	194 фс
18	27.5 пс	25.3 пс	4.86 пс	486 фс	48.6 фс
24	430 фс	98.8 фс	19.0 фс	1.9 фс	190 ас

З цієї таблиці можна зробити висновок про доцільність використання АЦП певної розрядності з урахуванням обмежень, що накладаються на час тремтіння фронту синхронізації (clock jitter).

Як правило, сигнали оцифровуються з мінімально необхідною частотою дискретизації з міркувань економії, при цьому шум квантування є білим, тобто його спектральна густина потужності рівномірно розподілена у всій смузі. Якщо ж оцифрувати сигнал з частотою дискретизації набагато більшої, ніж по теоремі Котельникова-Шеннона, а потім піддати цифровій фільтрації для придушення спектру зовні частотної смуги початкового сигналу, то відношення сигнал/шум буде кращим, ніж при використанні всієї смуги. Таким чином можна досягти ефективної роздільної здатності більшої, ніж розрядність АЦП.

Передискретизація також може бути використана для пом'якшення вимог до крутизни переходу від смуги пропускання до смуги придушення антиаліасінгового фільтру. Для цього сигнал оцифровують (наприклад, на вдвічі більшій частоті), потім проводять цифрову фільтрацію, пригнічуючи частотні компоненти зовнішньої смуги початкового сигналу, і, нарешті, знижують частоту дискретизації шляхом децимації.

Низький рівень інформативних кодів базису Радемахера, які включають істинні значення контрольованих параметрів плюс адитивні завади.

$$x_i = x_{i0} \pm \delta_i,$$

де x_{i0} – істинне значення контрольованих параметрів, δ_i – адитивна завада.

Спосіб реєстрації параметрів ОУ на основі БАЦП з часовим зміщенням вимірювання шляхом комутації параметрів призводить до появи ефекту старіння інформації, оскільки параметри не реєструються одночасно, що приводить до суттєвої втрати інформативної кореляції файлів та процесів обчислення матриць, коефіцієнту взаємної кореляцій, на основі яких будуються кластерні моделі виявлення та ідентифікації передавальних та аварійних ситуацій.

Оскільки двійкова система числення належить до позиційних систем числення, то при використанні операцій додавання чи віднімання в акумуляторі процесів виникають наскрізні переноси, що знижує швидкодію процесорів діагностування можливих передаварійних та аварійних ситуації на ОУ.

Висновки

Викладені теоретичні засади та дослідження формування цифрових відліків на основі кореляційних функцій у теоретико-числовому базисі Радемахера, дозволили обґрунтувати характеристики переваг та недоліків використання базису Радемахера в задачі ідентифікації станів об'єкта та є важливим резервом розвитку теорії та вдосконалення засобів діагностування квазістаціонарних об'єктів.

Література

1. Асташкин С. В. Функции Радемахера в симметричных пространствах / С. В. Асташкин // Функциональный анализ. – М., 2009. – С. 3–161.
2. Николайчук Я. М. Теорія джерел інформації : монографія / Николайчук Я. М. – Тернопіль : ТНЕУ, Економічна думка, 2008. – 396 с.
3. Ширмовська Н.Г. Оптимальна дискретизація заданих кореляційною функцією сигналів / Н.Г. Ширмовська // Методи та прилади контролю якості. – 2009. – № 22. – С. 107–111.

References

1. Astashkin S. V. Funkcii Rademahera v simmetrichnyh prostranstvah / S. V. Astashkin // Funkcional'nyj analiz. – M., 2009. – S. 3–161.
2. Nikolajchuk Ja. M. Teorija dzherel informacii : monografija / Nikolajchuk Ja. M. – Ternopil' : TNEU, Ekonomichna dumka, 2008. – 396 s.
3. Shirmovs'ka N.G. Optimal'na diskretizacija zadanih koreljacijnoju funkcieju signaliv / N.G. Shirmovs'ka // Metodi ta priladi kontrolju jakosti. – 2009. – № 22. – S. 107–111.

Рецензія/Peer review : 22.02.2018 р. Надрукована/Printed :28.03.2018 р.
Рецензент: д.т.н., проф. М.І. Горбійчук