TALLINNA POLÜTEHNILISE
INSTITUUDI TOIMETISED

ТРУДЫ ТАЛЛИНСКОГО
ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

СЕРИЯ А № 288

ТРУДЫ ПО ЭЛЕКТРОТЕХНИКЕ И АВТОМАТИКЕ

СБОРНИК СТАТЕЙ

VII

ТАЛЛИН 1970

Ф.6.7

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА
СЕРИЯ А № 288 1970

УДК 621

ТРУДЫ ПО ЭЛЕКТРОТЕХНИКЕ И АВТОМАТИКЕ

СБОРНИК СТАТЕЙ

VII

ТАЛЛИН 1970

Содержание

Стр.

I. Х.В. Силламаа. Систематика элементов в теории цепей 3

2. П.М. Плакк. Измерения при статическом двойном электрическом слое. I9

3. П.М. Плакк. Модель электрического двойного слоя. 33

4. О.М. Пикков. Математическая модель стационарного частотно-импульсного сигнала. 45

5. В.А. Крахт, Г.Э. Якобсон. К задаче последовательной декомпозиции конечных автоматов 55

6. В.А. Кукк. Случайное генерирование деревьев 63

7. В.А. Кукк. Рациональная аппроксимация передаточной функции 7I

8. Я.Н. Олль. Анализ точности оценивания весовой функции простейших линейных динамических объектов. 79

9. Я.М. Вырк. Применение методов теории информации при анализе измерительных устройств и преобразователей. 9I

I0. Я.М. Вырк, К.А. Хярмасте. Исследование влияния расположения полюсов на длительность переходной характеристики. I03

II. Я.М. Вырк. Последовательная электрическая коррекция динамических характеристик датчиков . . III

I2. А.А. Лаансоо. Расчет параметров измерительных цепей при наличии в их передаточных функциях диполей. I23

I3. А.А. Лаансоо. Расчет времени установления в УПТ с обратной связью при нелинейном режиме их работы. I33

Raamatukogu IV Tallinna Akadeemia

Ep. 9770

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

С Е Р И Я А № 288 1970

УДК 621.372.061.4

Х.В. Силламаа

СИСТЕМАТИКА ЭЛЕМЕНТОВ В ТЕОРИИ ЦЕПЕЙ

В классической теории цепей в качестве элементарных составных частей, из которых можно было построить любые цепи, применялись (кроме идеальных источников тока и напряжения) три типа двухполюсных элементов (дисторов) с параметрами - сопротивление R , индуктивность L и емкость C. Кроме них был введен еще многополюсный элемент - идеальный трансформатор (представляющий взаимную индуктивность). Система таких элементов действительно достаточна для описания свойств обратимых пассивных цепей [1,2]. Однако с развитием электроники стали широко применяться активные псевдолинейные и необратимые цепи [3], для описания которых необходимо было ввести новые элементы. Таковыми в первую очередь оказались управляемые источники,а вместе с ними и другие идеальные элементы, как например гираторы [3,4], трансакторы [5],преобразователи иммитанса [6], конверторы мощности [7], унисторы [8] и ряд других. Однако до настоящего времени отсутствует общая классификация таких элементов (возможных или же необходимых для описания любых цепей), а также не выяснены полностью их взаимные соотношения, представляющие несомненный интерес при развитии методов синтеза активных цепей.

В данной статье рассматривается задача анализа полной системы элементов применительно к узловой модели цепи (узловая концепция). Согласно этому любая цепь или ее часть может рассматриваться как n-полюсник, т.е. некоторая схема

(сочетание элементов) с n-узлами (полюсами), причем электрический режим каждого узла характеризуется потенциалом u_i (или напряжением относительно произвольного базиса) и током i_i [9,10]. Пользуясь законами Ома и Кирхгофа, можем для такого многополюсника составить n самостоятельных уравнений, которые эквивалентны матричному уравнению [11]

$$CU + DJ = 0, \qquad (1)$$

где $U = [u_1, \dots, u_n]^T$ - вектор-столбец узловых потенциалов,
$J = [i_1, \dots, i_n]^T$ - вектор-столбец узловых токов,
C и D - матрицы, определяющие свойства цепи.
При этом предполагается, что внутри многополюсника автономные источники отсутствуют (или же они разделены [9]).

Так как в любых физических цепях распределение токов не меняется при равном изменении потенциалов всех узлов, то уравнение (I) должно сохранить силу при замене U на новый вектор напряжений (изменение базиса) $U' = U + \mu E$, где $E - n \times 1$ матрица, все элементы которой равны I, а μ - произвольная скалярная постоянная. Отсюда вытекает условие

$$CE = 0, \qquad (2)$$

следовательно, матрица C имеет центрированные ряды (сумма элементов во всех рядах равняется нулю). Кроме того, из (2) следует $\det C = 0$, поэтому матрица C является ссобенной (ранг матрицы $r_c < n$). В то же время ранг блочной матрицы

$$[C, D] \qquad (3)$$

всегда n, что соответствует линейной независимости совокупности n уравнений, составляющих систему (I). Поэтому в матрице (3) всегда существует некоторый ненулевой минор порядка n. Если таковым является матрица D, то уравнение (I) может быть представлено в виде

$$J = -D^{-1}CU = YU, \qquad (4)$$

и мы получим представление n-полюсника через т.н. неопределенную матрицу узловых проводимостей Y [10]. При других ненулевых минорах порядка n в матрице (3) получается одно

из гибридных представлений многополюсника [I2] (по крайней мере одно из них всегда существует). В дальнейшем анализе пока предполагаем возможность представления цепи соотношением (4) (существование матрицы проводимостей). Как известно [I0], неопределенная матрица Y является двояко центрированной, и суммы элементов всех рядов и столбцов равны нулю. Если какой-либо ряд и соответствующий столбец матрицы Y состоят лишь из нулевых элементов, то n-полюсник содержит соответствующий изолированный узел. Прибавление к n-полюснику $m-n$ изолированных зажимов $(m>n)$, что означает дополнение $n\times n$ матрицы Y $m-n$ рядами и столбцами из нулевых элементов, позволяет превратить n-полюсник в m-полюсник без изменения режима исходных n полюсов.

При соединении всех полюсов двух n-полюсников с уравнениями $J_1 = Y_1 U_1$ и $J_2 = Y_2 U_2$ соответственно имеют место очевидные соотношения

$$J = J_1 + J_2; \quad U = U_1 = U_2, \tag{5}$$

где J и U представляют собой вектор токов и напряжений суммарного n-полюсника. Из формул (4) и (5) элементарным образом вытекает, что неопределенная матрица проводимости суммарного n-полюсника

$$Y = Y_1 + Y_2. \tag{6}$$

Формула (6), однако, пригодна для любых соединений узлов рассматриваемого класса многополюсников (обладающих матрицей Y). Если имеются m-полюсник (с матрицей узловых проводимостей $Y_{1(m)}$) и n-полюсник (с матрицей узловых проводимостей $Y_{2(n)}$) и хотим найти матрицу проводимостей многополюсника, получаемого соединением k узлов рассмотренных многополюсников и имеющего $m+n-k$ узлов, то в первую очередь проведем новую общую нумерацию узлов так, чтобы соединяемые k узлов были первыми, а остальным узлам ($m-k$ для первого многополюсника и $n-k$ для второго) присвоим последующие номера (до $m+n-k$). Теперь первому m-полюснику добавим $n-k$ изолированных узлов, номера

которых совпадают с несоединяемыми узлами второго многополюсника, а второму n-полюснику добавим аналогично $m-k$ изолированных узлов. В итоге получим матрицы проводимостей $Y_{1(m+n-k)}$ и $Y_{2(m+n-k)}$ одинакового порядка, допускающие суммирование согласно (6) и давая в результате искомую матрицу узловых проводимостей суммарного многополюсника.

Свойство суммируемости матриц проводимостей при соединении соответствующих многополюсников играет основную роль при выяснении системы элементов, посредством которых представимы любые цепи (рассматриваемого класса). Если ввести понятие эквимодульной матрицы, понимая под этой вещественную матрицу, все ненулевые элементы которой имеют равные абсолютные значения, то можно сформулировать следующую основную теорему:

Теорема I.[x] Любая двояко центрированная квадратная матрица (проводимостей) $Y_{(n)}$ n-го порядка может быть разложена на сумму не более чем $(n-1)^2$ двояко центрированных эквимодульных матриц (проводимостей) единичного ранга.

Доказательство может быть конструктивное. Эквимодульные двояко центрированные матрицы с наименьшим количеством ненулевых элементов должны иметь четыре ненулевых элемента в точках пересечения двух рядов и столбцов. Ранг такой матрицы равен единице. Образуем такую матрицу, принимая за основу один из элементов матрицы $Y_{(n)}$ среди первых $n-1$ рядов и столбцов, например y_{ij}, следующим образом (остальные ненулевые элементы расположены в n-ом ряду и столбце):

$$Y_{ij} = \begin{array}{c} \begin{matrix} 1 & & j & & n \end{matrix} \\ \begin{bmatrix} 0 & \cdots & 0 & \cdots & 0 \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & y_{ij} & \cdots & -y_{ij} \\ \vdots & & \vdots & & \vdots \\ 0 & \cdots & -y_{ij} & \cdots & y_{ij} \end{bmatrix} \begin{matrix} 1 \\ \\ i \\ \\ n \end{matrix} \end{array} \qquad (7)$$

[x] Теорема I, по существу, означает, что множество всех матриц проводимостей n-полюсника образует векторное пространство, размерность которого не превышает $(n-1)^2$.

Очевидно, на базе каждого ненулевого элемента матрицы $Y_{(n)}$ в пределах первых $n-1$ рядов и столбцов можно образовать одну матрицу типа Y_{ij} и поэтому их количество не превышает $(n-1)^2$. Остается еще показать, что

$$Y_{(n)} = \sum_{i,j=1}^{n-1} Y_{ij} . \qquad (8)$$

Так как в пределах первых $n-1$ рядов и столбцов лишь одна из суммируемых матриц имеет ненулевой элемент, то справедливость (8) в пределах $n-1$ рядов и столбцов очевидна. С другой стороны, сумма двояко центрированных матриц является также двояко центрированной матрицей. Поэтому условие (8) должно быть справедливым также для последнего n-го ряда и столбца соответственно, чем и утверждение теоремы доказано.

Таким образом, любая матрица проводимости может быть образована конечной суммой эквимодульных матриц типа (7). Если в такой матрице отбросить все пары рядов и столбцов, состоящие из только нулевых элементов и соответствующие изолированным узлам, то остается в общем случае матрица третьего порядка (с узлами i, j и n). Отсюда вытекает важное следствие.

Следствие I. Любые цепи, обладающие матрицей проводимостей, могут быть составлены из трехполюсных элементарных элементов.

Путем нормализации (вынесение общего скалярного множителя) эквимодульных матриц значения всех ее элементов можно привести к величинам +I, 0 и -I.

Если теперь рассматривать эквимодульные матрицы третьего порядка с такими (единичными) элементами, то очевидно, что из-за условия центрированности (равенства нулю суммы всех элементов) возможны лишь семь различных столбцовых векторов матриц

$$\begin{matrix} 0 & 0 & 0 & 1 & 1 & -1 & -1 \\ 0 & 1 & -1 & 0 & -1 & 0 & \cdot \\ 0 & -1 & 1 & -1 & 0 & 1 & 0 \end{matrix}$$

Из них можно составить 3I различную двояко центрированную матрицу, которые на основании следствия I и позволяют образовать любые цепи, обладающие матрицей проводимости. Задача теперь заключается в классификации этих элементов, так как среди полученного множества матриц некоторые соответствуют одному и тому же физическому элементу, отличаясь лишь различным порядком нумерации полюсов. Различная нумерация полюсов может быть учтена введением т.н. пермутационных матриц (матриц перестановок) P, которые позволяют заменить векторы токов J и напряжений U на пермутированные

$$J_1 = PJ, \qquad U_1 = PU, \tag{9}$$

отличающиеся порядком составляющих векторов. Количество возможных сочетаний составляющих для векторов третьего порядка равно $3!=6$, поэтому существуют шесть различных пермутационных матриц третьего порядка

$$\begin{bmatrix}1&0&0\\0&1&0\\0&0&1\end{bmatrix}, \begin{bmatrix}0&1&0\\0&0&1\\1&0&0\end{bmatrix}, \begin{bmatrix}0&0&1\\1&0&0\\0&1&0\end{bmatrix}, \begin{bmatrix}0&1&0\\1&0&0\\0&0&1\end{bmatrix}, \begin{bmatrix}1&0&0\\0&0&1\\0&1&0\end{bmatrix}, \begin{bmatrix}0&0&1\\0&1&0\\1&0&0\end{bmatrix}.$$

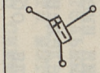
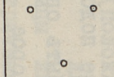
Эти пермутационные матрицы являются подклассом ортогональных матриц, все ненулевые элементы которых равны единице, и в алгебраическом смысле образуют известную шестиэлементную некоммутативную (мультипликативную) группу [I3].

Подстановка (9) в (4) дает нам соотношение

$$J_1 = P^{-1} Y P U_1, \tag{10}$$

откуда следует, что один физический трехполюсник при различном расположении относительно узлов (при различном нумерации узлов) генерирует класс ортогонально подобных матриц проводимостей ($Y = P^{-1} Y P$). Таким образом, отношение подобия матриц позволяет провести разложение интересующего нас множества матриц проводимостей трехполюсников на непересекающиеся классы (так как отношение подобия является

Таблица I

№	Название элементора	Кол-во различных элементоров	Пример матрицы элементора	Свойства матриц: собственные значения	Свойства матриц: характерист. полином	Свойства матриц: след S	Свойства матриц: алг. дополн. элемент D	Условное обозначение
I	Адмитор Admitor	6	$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	0; 0; 2	$\lambda^2(\lambda-2)$	2	0	
2	Транзор Transor	I2	$\begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$	0; 0; 1	$\lambda^2(\lambda-1)$	1	0	
3	Унатор Unator	4	$\begin{bmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \end{bmatrix}$	0; $\frac{1}{2}(3 \pm j\sqrt{3})$	$\lambda(\lambda^2-3\lambda+3)$	3	1	
4	Гиратор Gyrator	2	$\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix}$	0; $\pm j\sqrt{3}$	$\lambda(\lambda^2+3)$	0	1	
5	Диплор Diplor	6	$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix}$	0; $\pm\sqrt{3}$	$\lambda(\lambda^2-3)$	0	-1	
6	Изолятор Isolator	I	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$	0; 0; 0	λ^3	0	0	

отношением эквивалентности [I4]). Все матрицы в одном классе подобия имеют равные собственные значения. Они могут быть определены по характеристическому уравнению матриц, которое для общего случая матрицы проводимостей трехполюсника будет иметь вид

$$\lambda^3 - S\lambda^2 + 3D\lambda = 0, \qquad (II)$$

где $S = y_{11} + y_{22} + y_{33}$ - след матрицы (сумма элементов главной диагонали, $D = y_{11}y_{22} + y_{12}y_{21}$ - определитель соответствующей определенной матрицы $\mathcal{Y}$, получаемой из Y вычеркиванием одного (любого) ряда и соответствующего столбца [I0].

Анализ собственных значений и условий подобия матриц показал, что множество эквимодульных матриц третьего порядка распадается на шесть классов, каждый из которых соответствует определенному типу элементов. Следует подчеркнуть, что полученные классы исходят из свойств неопределенной матрицы проводимостей Y, которую можно рассматривать как матрицу проводимостей, если базисным узлом цепи [I0] избран произвольный изолированный узел.

Все классы трехполюсников и некоторые их характеристики представлены в таблице I. Большинство из полученных классов элементов (в отдельности) уже известны в литературе, хотя иногда под различными названиями.[x] Рассмотрим подробнее свойства этих классов:

I) Адмитор. Представляет собой трехполюсник, между двумя зажимами которого имеется обычная проводимость, а третий зажим изолирован. Таким образом, фактически имеем вырожденный трехполюсник - двухполюсник с дополнительным изолированным зажимом. В силу этого матрица проводимостей адмитора обладает двумя нулевыми собственными значениями. Харак-

---

[x] Ряд приведенных в таблице I названий элементов предложен автором.

терный признак адмитора - наличие на главной диагонали матрицы $Y_{(3)}$ двух ненулевых элементов одного знака (положительных для положительного адмитора и отрицательных для отрицательного адмитора). Ввиду симметричности матрицы любые схемные сочетания адмитора обладают также симметричными матрицами проводимостей (обратимые цепи).

2) Транзор. На самостоятельное значение такого элемента (с наименованием " transconductance ") обратил внимание С.И. Мэзон [8], хотя под названием "управляемый напряжением источник тока" он известен уже давно и широко применяется в схемах замещения в электронике.[x] Транзор является идеализацией пентода, полевого транзистора и т.д., являясь активным, необратимым элементом. Характерный признак его - наличие одного ненулевого элемента на главной диагонали матрицы проводимостей $Y_{(3)}$ (знак соответствует знаку транзора).

3) Унатор. Под таким названием элемент введен С.И. Мэзоном [8] как цикл из трех унисторов. Унатор является активным, необратимым элементом. Характерный признак его - наличие трех ненулевых элементов одного знака на главной диагонали матрицы проводимостей $Y_{(3)}$.

4) Гиратор. Элемент введен Б. Теллегеном [4] как неизбежный для описания необратимых цепей [3]. Является он пассивным элементом с кососимметричной матрицей проводимостей $Y_{(3)}$. Отсюда характерный признак - все элементы на главной диагонали матрицы нулевые.

5) Диплор. Описание данного элемента в литературе отсутствует. Является он активным неустойчивым элементом (матрица имеет разнозначные собственные значения). Его характерный признак: на главной диагонали матрицы проводимостей имеются два ненулевых разнозначных элемента.

6) Изолятор. Представляет он собой вырожденный трехполюсный элемент с нулевой матрицей проводимостей и может рассматриваться как совокупность трех изолированных узлов.

---

[x] Называть данный элемент "источником тока" неудачно хотя бы потому, что характерный параметр элемента имеет размерность не тока, а проводимости.

Нетрудно убедиться, что любая матрица проводимости может быть разложена на сумму эквимодульных матриц указанных классов. Например, для матрицы проводимостей с вещественными элементами получается:

$$\begin{bmatrix} -1 & -4 & 5 & 0 \\ 0 & 4 & -8 & 4 \\ 0 & 3 & 0 & -3 \\ 1 & -3 & 3 & -1 \end{bmatrix} = 5 \begin{bmatrix} 0 & -1 & 1 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} - 1 \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \end{bmatrix} + 3 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & 1 & 0 \end{bmatrix}. \quad (12)$$

Представленное разложение означает, что данную цепь можно представить как сумму транзора проводимостью 5 (на узлах I, 2 и 3), унатора проводимостью - I (на узлах I, 2 и 4) и гиратора проводимостью 3 (на узлах 2, 3 и 4).

Аналогично можно провести разложения матриц, элементами которых являются комплексные числа или даже многочлены и дробно-рациональные функции и т.д.

Легко убедиться, что такие разложения не единственны, а существует множество эквивалентных разложений. Причиной являются взаимные связи между элементами разных классов, допускающие взаимную их замену. Так, например, диплор может быть представлен как соединение положительного и отрицательного адмитора:

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & -1 & 1 \\ 0 & 1 & -1 \end{bmatrix}. \quad (13)$$

Гиратор можно представить как сумму положительного и отрицательного транзора

$$\begin{bmatrix} 0 & 1 & -1 \\ -1 & 0 & 1 \\ 1 & -1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}. \quad (14)$$

Унатор выражается через сумму адмитора и транзора

$$\begin{bmatrix} 1 & 0 & -1 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 & -1 \\ 0 & 0 & 0 \\ 0 & -1 & 1 \end{bmatrix}, \qquad (15)$$

а адмитор, в свою очередь, можно представить как сумму двух транзоров

$$\begin{bmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 0 \end{bmatrix}. \qquad (16)$$

При этом представленные разложения также не единственны и могут быть заменены другими, эквивалентными. Более подробный анализ взаимных связей показывает, что элементы всех классов могут быть представлены через положительные и отрицательные транзоры. Это непосредственно видно по разложениям (13) по (16), если в (13) и (15) адмиторы заменить разложением (16).

В итоге вытекает, что адмитор можно представить двумя транзорами одного знака, гиратор и диплор (а также изолятор) двумя транзорами разного знака, и унатор тремя транзорами одного знака. Поэтому транзор (управляемый источник тока) является истинным (и при этом трехполюсным) элементарным элементом всех цепей, обладающих матрицей проводимостей.

Для n -полюсников, согласно теореме 1, возможно разложение их матриц на конечную сумму матриц типа (7). Нетрудно убедиться, что эти матрицы Y_{ij} при $i \neq j$ соответствуют транзорам (дополненных $n-3$ изолированными узлами) при общем их количестве не более $(n-1)(n-2)$, а при $i=j$ будем иметь адмиторы (также дополненные изолированными узлами) с максимальным количеством $n-1$. Представляя теперь адмиторы также через транзоры (их будет не более $2(n-1)$), приходим к заключению, что любой n -полюсник можно разложить на

сумму не более $n(n-1)$ транзоров. Но, как следует из предыдущего изложения, такое разложение не единственное.

Дальнейшее исследование проблемы показало, что класс образующих элементов может быть еще сужен. Это основывается на взаимной связи положительных и отрицательных транзоров. Такая связь может быть представлена матричным уравнением

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & -1 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & -1 & 1 \\ -1 & 1 & 1 & -1 \\ 0 & 0 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \underset{\text{искл. } 4}{\Longrightarrow} \begin{bmatrix} -1 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 0 \end{bmatrix}. \quad (17)$$

Согласно этому уравнению суммирование трех положительных транзоров (на узлах I,2,4; I,3,4 и 2,3,4) дает в итоге четырехполюсник, откуда превращением четвертого узла во внутренний (путем исключения узла 4 [I0]) получается трехполюсник, соответствующий отрицательному транзору. Другими словами, свойства полученного четырехполюсника относительно узлов I, 2 и 3 эквивалентны отрицательному транзору при условии, что узловой ток четвертого узла идентично равняется нулю (условие внутренности узла), т.е. четвертый узел не должен соединяться никакими другими элементами. При этом примечательно, что такое представление отрицательного транзора через положительные единственное и минимальное (с наименьшим количеством транзоров). Требуется, однако, введение дополнительных, внутренних узлов в схему.

Использование описанного соотношения дает возможность построения реализаций любых цепей (и элементов) с применением лишь положительных транзоров. Например, для представления гиратора потребовались согласно (I4) один положительный и один отрицательный транзор. Заменяя теперь отрицательный транзор через положительные и вводя при этом внутренний узел, получим в итоге, например, следующее представление гиратора через положительные транзоры

$$\begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & -1 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -1 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ -2 & 2 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \qquad (I8)$$

$$\begin{bmatrix} 1 & -2 & 0 & 1 \\ 0 & 1 & 0 & -1 \\ -2 & 2 & 1 & -1 \\ 1 & -1 & -1 & 1 \end{bmatrix} \underset{\text{искл. } 4}{\Longrightarrow} \begin{bmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{bmatrix}.$$

Ввиду того, что исходное представление гиратора (I4) через транзоры было не единственное, не единственным окажется также представление гиратора через четыре положительных транзора.

Для представления отрицательного адмитора через положительные транзоры потребуется уже два внутренних узла, так как отрицательный адмитор может быть представлен двумя отрицательными транзорами (аналогично (I6)), а каждый отрицательный транзор может быть заменен в соответствии с (I7) тремя положительными одновременным введением своего внутреннего узла.

Аналогично нетрудно убедиться, что для представления отрицательного унатора потребуется в итоге девять положительных транзоров с тремя внутренними узлами.

Таким образом, увидим, что любой Π-полюсник, характеризуемый матрицей проводимостей, может быть представлен определенной схемой соединения простейших трехполюсников (в частности положительных транзоров). Целесообразно называть такие схемы "элементорными схемами замещения Π-полюсника" или же "элементорным представлением Π-полюсника", а самих элементарных трехполюсников элементорами. Из предыдущего изложения ясно, что, как правило, каждый Π-полюсник обладает множеством эквивалентных между собой элементорных представлений, раскрывающих внутреннюю структуру и свойства данного Π-полюсника.

Ограничение класса рассматриваемых Π-полюсников теми, которые обладают матрицей проводимостей, по существу может

быть снято. Известно [I5], что последовательное применение принципа введения внутренних узлов позволяет дать посредством матрицы проводимостей описание любых n -полюсников. Например, трехполюсник (идеальный усилитель напряжения), описываемый векторным уравнением (I) в виде

$$\begin{bmatrix} K & -1 & 1-K \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \end{bmatrix} = 0 , \qquad (I9)$$

ввиду сингулярности матрицы D не имеет матрицы проводимостей. Однако, если ввести четвертый, внутренний узел с нужным образом введенными дополнительными условиями, приходим, например, к уравнению

$$\begin{bmatrix} K & -1 & 1-K & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \\ i_4 \end{bmatrix} = 0 , \qquad (20)$$

которое, при условии внутренности четвертого узла ($i_4 \equiv 0$), полностью эквивалентно исходному уравнению (I9) относительно переменных узлов I, 2 и 3. Однако для уравнения(20) матрица уже регулярна и для такого четырехполюсника существует матрица проводимостей в виде

$$Y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ -K & 1 & K & -1 \\ K & -1 & 1-K & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ -K & 0 & K & 0 \\ K & 0 & -K & 0 \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 1 \\ 0 & 1 & 0 & -1 \\ 0 & -1 & 1 & 0 \end{bmatrix} , \qquad (2I)$$

которая может быть представлена как соединение транзора величиной k на узлах I, 3 и 4 и гиратора величиной I на узлах 2, 3 и 4. Внутренний узел может быть введен, вообще, различным образом [I5] и в результате получаются различные матрицы проводимостей. Например, для идеального усилителя напряжения могут быть получены кроме (2I) еще матрицы проводимости

$$Y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -K & 1 & K & -1 \\ 0 & 0 & -1 & 1 \\ K & -1 & 1-K & 0 \end{bmatrix} \quad \text{или} \quad Y = \begin{bmatrix} 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 1 \\ 1+K & -1 & 1-K & -1 \\ -K & 1 & K-1 & 0 \end{bmatrix} \qquad (22)$$

и многие другие. Однако все такие матрицы обеспечивают одинаковые свойства (эквивалентность) относительно внешних узлов (узлов исходного многополюсника). Таким образом, операция введения внутренних узлов является также не единственной.

Аналогично можно конструировать матрицу проводимостей для любой цепи, описываемой системой уравнений (I)и, следовательно, для всех многополюсников можно дать элементорную схему замещения.

Литература

I. W. Cauer. Theorie der linearen Wechselstromschaltungen, 2.Aufl. Akademie-Verlag, Berlin 1954.

2. H.J. Carlin, A.B. Giordano. Network theory. Prentice-Hall, N.Y.1964.

3. А.Н. Милях, А.К. Шидловский. Принцип взаимности и обратимость явлений в электротехнике. Наукова думка, Киев 1967.

4. B.D.H. Tellegen. The gyrator, a new electric network element. Philips Res.Rept.v.3. No.2.Apr.1948, p.81-101.

5. G.E. Sharpe. Axioms on transactors. Trans. IRE, CT-5. No.3.Sept. 1958, p.189-196.

6. J.G. Linvill. RC active filters.Proc.IRE, v.42. March 1954, p.555-564.

7. Э.В. Зелях. Идеальный преобразователь мощности - новый элемент электрической схемы. Электросвязь, 1957, № I.

8. S.J. Mason. Topological analysis of linear nonreciprocal networks. Proc.IRE,v.44, June,1957.

9. Э.В. З е л я х. Основы общей теории линейных электрических схем. Изд. АН СССР, М. 1951.

10. В.П. С и г о р с к и й. Методы анализа электрических схем с многополюсными элементами. Изд. АН УССР, Киев 1958.

11. R.W. N e w c o m b. Linear multiport synthesis. Mc-Graw Hill, N.Y.1966.

12. В.П. С и г о р с к и й. Гибридная матрица электронной схемы. Радиотехника, т.23 № 11, 1968, стр. 73.

13. А.Г. К у р о ш. Теория групп. Наука, М. 1967.

14. Ф.Р. Г а н т м а х е р. Теория матриц. Наука, М. 1967.

15. T. T o k a d. Some extention of the application of topologic formulas. Trans. IEEE, CT-13,Sept. 1966, p.332-335.

H. Sillamaa

NETWORK ELEMENT CLASSIFICATION

Summary

A classification of elementary network elements (elementors) is developed by using the properties of undefined admittance matrices. So every network, for which the admittance matrix description exists, may be presented by its elementor-circuit diagram. Such an elementor description may be obtained also for every network by using the principle of internal vertices (terminals). Is is shown that for admittance circuits elementor diagrams may be composed as a connection of only one type of elementors - transors. For every network there exists a set of equivalent elementor descriptions with the same terminal behaviour. Elementor descriptions may be used for solving network syntheses and realization problems, in particular for the active networks.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

С Е Р И Я А № 288 1970

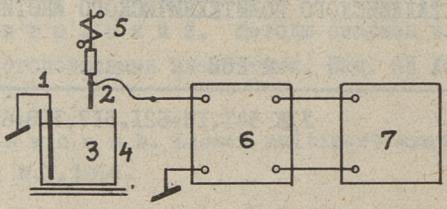
УДК 541.13+621.317.318+621.317.321

П.М. Плакк

ИЗМЕРЕНИЯ ПРИ СТАТИЧЕСКОМ ДВОЙНОМ ЭЛЕКТРИЧЕСКОМ СЛОЕ

При погружении электрода в раствор электролита на поверхности его образуется статический двойной электрический слой. Несмотря на широкое изучение последнего [1], в литературе отсутствуют количественные данные, а именно: о скорости образования двойного слоя, о величине образующегося при этом скачка потенциала между раствором и электродом и о величине заряда в двойном слое. Для получения ответа на эти вопросы использовались следующие методы измерения.

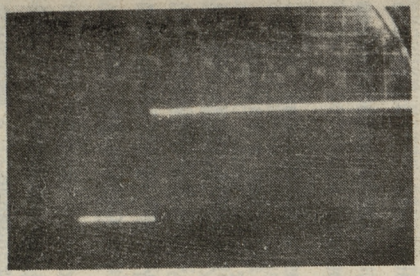
а) Метод потенциалов дает возможность следить за характером и динамикой образования статического двойного слоя и определяет скачок потенциала электрода или напряжение электрода относительно раствора. Измерительная схема представлена на фиг. 1. В водный раствор хлористого калия помещается пластинчатый платиновый электрод с большой поверхностью и острыми краями. Примерно через полчаса, когда на пластинчатом электроде образовался статический двойной слой, в тот же раствор погружается очень быстро (при помощи соответствующего пружинного механизма) чистый и сухой платиновый стержневой электрод. Последний прикреплен изолированно к соответствующему держателю. Пружинным механизмом синхронизирован осциллограф, работающий в режиме ждущей развертки. В соответствии с регулированием осциллографа фотографировались осциллограммы, которые показывают изменение потенциала стержневого электрода (в отношении заземленного пластинчатого элек-



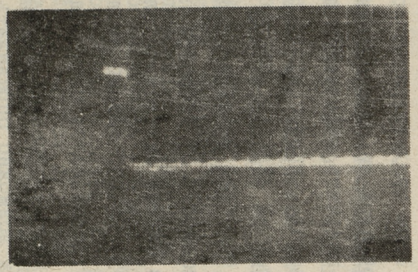
Фиг. 1. Схема метода потенциала:
1 – заземленный электрод с большой поверхностью из листовой платины, 2 – изолированный от земли зондовый электрод, 3 – водный раствор хлористого калия, 4 – изолированный от земли стеклянный сосуд, 5 – пружинный механизм, 6 – электрометрический усилитель, 7 – осциллограф

трода) в течение от нескольких тысячных долей до десяти секунд. Далее наблюдались изменения потенциала стержневого электрода по вольтметру электрометрического усилителя. Входное сопротивление усилителя вместе с сопротивлением изоляции держателя стержневого электрода находится в порядке величины 10^9 ом, и емкость двойного слоя стержневого электрода составляет примерно 10 мкф. Тем самым константа времени двойного слоя по измерительному прибору составляет приблизительно 3 часа и прибор не оказывает влияния на процесс, происходящий на стержневом электроде.

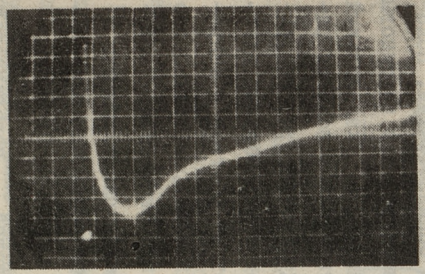
Изложенным выше способом была получена осциллограмма, приведенная на фиг. 2. На ней зарегистрирован скачок потенциала стержневого электрода в тот момент, когда последний вступает в контакт с раствором. Далее потенциал стержневого электрода уменьшается очень медленно и лишь в течение от получаса до одного часа он приходит к нулю. Так как чистый и сухой стержневой электрод до погружения в раствор не имеет двойного слоя, то скачок потенциала на осциллограмме показывает единственное имеющееся в измерительной цепи напряжение, т.е. напряжение U_0 между водным раствором хлористого калия и платиновым пластинчатым электродом. Следовательно, двойной слой образуется на изолированном от земли чистом платиновом электроде очень медленно и его успешно



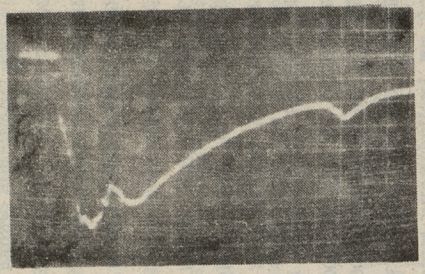
Фиг. 3. Осциллограмма насыщенного каломельного электрода, полученная методом потенциала



Фиг. 2. Осциллограмма платинового электрода, полученная методом потенциала



Фиг. 4. Осциллограмма, полученная в случае недостаточно чистого зонда



Фиг. 5. Осциллограмма, полученная в случае недостаточно чистого зонда

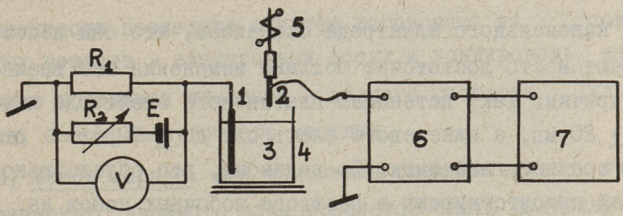
можно применять в качестве зонда при "нащупывании" потенциала жидкости. В дальнейшем называем платиновый стержень зондовым электродом.

Заменив платиновый пластинчатый электрод насыщенным каломельным электродом, мы получили вид осциллограммы, показанный на фиг. 3. Из осциллограммы явствует, что платиновый электрод является на 100 мв положительным, а каломельный электрод на 310 мв отрицательным в отношении к водному раствору хлористого калия. Для оценки достоверности результатов платиновый пластинчатый электрод и каломельный электрод были одновременно погружены в тот же раствор и напряжение между ними измерялось при помощи электрометрического усилителя. Оказалось, что соответственно ожидаемому, непосредственно измеренное падение напряжения равно разнице скачка потенциала, полученного из соответствующих осциллограмм.

В случае зондового электрода с нечистой поверхностью получаются осциллограммы весьма различного вида (например фиг. 4 и 5). Из последних видно, что адсорбировавшиеся добавки уже на воздухе образуют на зонде двойной слой, который при контакте с раствором начинает изменяться по времени. Если преднамеренно создать на зондовом электроде двойной слой из определенных добавок, то по осциллограммам можно получить информацию о его свойствах.

Когда повторили описанный выше опыт с никелевыми электродами, то оказалось, что никель не подходит в качестве зонда, так как двойной слой на нем образуется очень быстро. В зависимости от скорости погружения зонда в жидкость можно и в случае никелевого электрода получить известный скачок потенциала, однако последний в общем случае не совпадает с потенциалом электрода.

б) Компенсационным методом с точки зрения измерительной техники возможно определить скачок потенциала двойного слоя со сколько угодно большой точностью. Этот метод работает по схеме, представленной на фиг. 6. С помощью внешнего источника напряжения E и регулируемого сопротивления



Фиг. 6. Схема компенсационного метода. Числа с 1 до 7 относятся к тем же обозначениям, что и на фиг. 1

R_2 создается на малом сопротивлении R_1 падение напряжения, которое численно равно и противоположно напряжению U_0 электрода, полученному потенциальным методом. Таким способом электролитный раствор приводится к потенциалу земли. Зондовый электрод имеет, через входное сопротивление усилителя, также потенциал земли. При введении зонда в раствор в случае полной компенсации их потенциалы равны и скачка на осциллограмме не возникает. В случае же повышенной или пониженной компенсации на осциллограмме образуется скачок, по величине и направлению которого напряжение компенсации можно отрегулировать.

Измерялись скачки потенциала электродов относительно водного раствора хлористого калия. Полученные результаты приведены в таблице I. В качестве растворителя использовался бидистиллят, однако, к сожалению, степень чистоты хлористого калия была только "х.ч." и он содержал до 0,2% разных примесей. Максимальная ошибка, имевшая место при определении скачка потенциала каломельного электрода, составляла $\pm$ 10 мв, а у платинового электрода $\pm$ 5 мв. Достижение большей точности препятствовалось, с одной стороны, степенью чистоты поверхности зонда. С другой стороны, кажется, что также диффузионный потенциал каломельного электрода в соединительном мостике (в стеклянной мембране) не совсем константен, а в известной мере изменяется. При проверке постоянства скачка потенциала платинового и никелевого электродов в водном растворе хлористого калия при

помощи каломельного электрода оказалось, что они непостоянны. Имеют место достаточно большие изменения без внешне видимых причин. Так, потенциал платинового электрода изменялся до ± 20 мв, а никелевого электрода до ± 70 мв относительно среднего значения. По-видимому, это обусловлено адсорбцией присутствующих в растворе побочных ионов на поверхность электрода. Весьма вероятно, что такие изменения потенциала являются одной из причин, почему при измерениях двойного слоя следует говорить об "истории" электродов с мостом переменного тока.

Таблица I

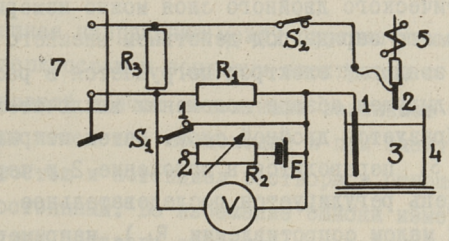
Скачок потенциала электрода или напряжение U_0 в мв относительно водного раствора хлористого калия. Представленные значения являются арифметическими средними из многих измерений

Концентрация раствора, моль/л	Насыщенный каломельный электрод	Платиновый электрод	Никелевый электрод
Насыщенный	-345	+20	от
I,0	-330	+50	-320
0,I	-3I8	+80	до
0,0I	-3I2	+I00	-460

Сравнивались между собой производимые промышленностью насыщенные каломельные электроды, снабженные стеклянной мембраной. Измерялась разность напряжения между их парами при помощи динамического конденсатора милливольт-электромера (входное сопротивление $I0^{I3}$ ом, измерительный ток приблизительно $I0^{-I5}$ а). Выяснилось, что в насыщенном водном растворе хлористого калия их напряжения были равными и составляли 0,5 мв в пределах погрешностей измерения. При разбавлении раствора различие в напряжениях повышалось и при концентрации 0,0I моль/л достигало до 20 мв. Электрод с мостиком из агар-агара дал еще большее расхождение в напряжениях. Данные таблицы I действительны для промышленно производимого каломельного электрода со стеклянной мембраной, в

мостике которого имел место наименьший диффузионный потенциал. Для точного определения скачка потенциала каломельного электрода последний следует изготовить из совершенно чистых веществ в электродный сосуд и зондировать непосредственно потенциал элемента - водного раствора хлористого калия, без промежуточного мостика.

в) Метод заряда. Для выяснения электрического строения статического двойного слоя требуется знать величину накапливающегося в нем электрического заряда. Последний можно измерить т.н. методом заряда по схеме, представленной на фиг. 7.

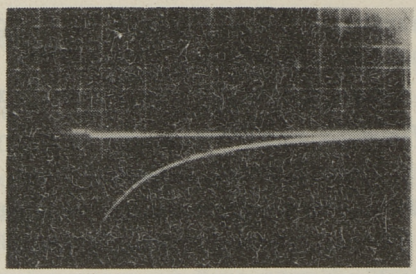


Фиг. 7. Схема метода заряда. Числа с 1 до 7 относятся к тем же обозначениям, что и на фиг. 1

Для измерения заряда, который накапливается в двойном слое при его образовании, в схеме имеется выключатель S_1 в положении I и выключатель S_2 в замкнутом положении. При погружении зондового электрода в раствор на нем очень быстро образуется двойной слой под действием напряжения U_0 платинового пластинчатого электрода. Требуемый для этого заряд проходит через сопротивление R_3 и на осциллограмме фиксируется протекание во времени тока заряда емкости двойного слоя (см. фиг. 8). Из осциллограммы находятся величина заряда Q_0 и напряжение U_0 . По последним вычисляется емкость статического двойного слоя

$$C_0 = Q_0/U_0 . \qquad (I)$$

Так как площадь пластинчатого электрода свыше ста раз превышает площадь зондового электрода, то снятый с него заряд не изменяет существенно скачка потенциала U_0 .

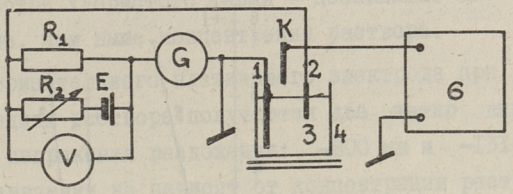


Фиг. 8. Осциллограмма, полученная по методу заряда

Емкость статического двойного слоя можно измерить также через изменение его заряда под действием внешнего напряжения. Для этого зондовый электрод погружается в раствор при описанном в предыдущем абзаце положении выключателей и на зонде быстро образуется двойной слой. Затем открывается выключатель S_2, S_1 переводится в положение 2 и через R_2 в измерительную цепь регулируется последовательное внешнее напряжение (при малом сопротивлении R_1), например 50 мв. После замыкания выключателя S_2 внешнее напряжение изменяет величину двойного слоя и необходимый для этого заряд фиксируется на осциллограмме. Эта процедура повторяется при противонаправленном внешнем напряжении. Полученные осциллограммы полностью совпадают с приведенной на фиг. 8. Вычисляя на основе осциллограмм емкость статического двойного слоя, получаем для нее одинаковое значение – 70 мкф/см^2, которое действительно для платинового электрода и при всех концентрациях водного раствора хлористого калия.

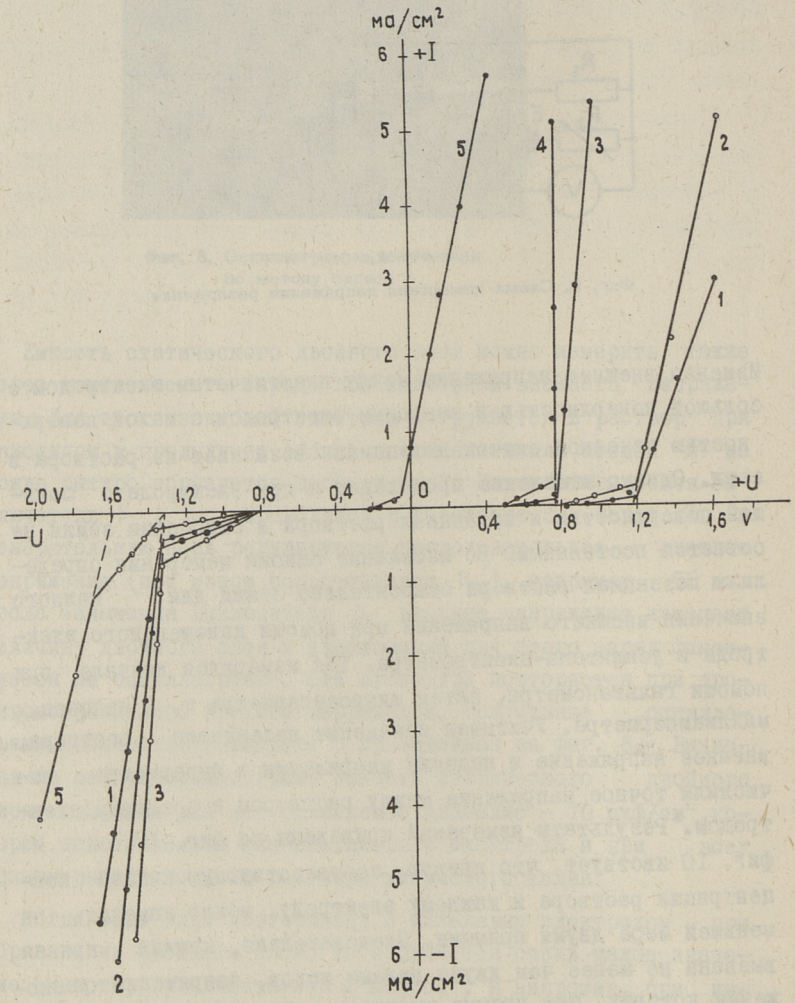
Когда этот опыт повторяется с никелевым электродом при образовании двойного слоя, то получается очень малое значение емкости, приблизительно 2 мкф/см^2. И напротив, при изменении двойного слоя, образовавшегося на никеле, получается значение емкости 55 мкф/см^2, которое не зависит от концентрации раствора. По-видимому, между никелевым электродом и водным раствором хлористого калия в случае потенциала нулевого заряда происходит весьма легко обмен зарядами. Однако у потенциала двойного слоя никеля, примерно – 400 мв, обмен заряда невозможен.

г) Измерение напряжения разложения. Для измерения напряжения разложения между электродом и раствором использовалась схема, показанная на фиг. 9.



Фиг. 9. Схема измерения напряжения разложения

Изменяя внешнее напряжение между пластинчатым электродом с большой поверхностью и зондовым электродом с малой поверхностью основной скачок потенциала возникает из раствора в зонд. Однако изменение происходит и при электроде с большой поверхностью и потенциал раствора в отношении земли не остается постоянным. Во избежание ошибки измерения определили потенциал раствора относительно земли для каждого значения внешнего напряжения при помощи каломельного электрода и усилителя-электрометра. Ток измерялся вначале при помощи гальванометра, затем микроамперметра и, наконец, миллиамперметра. Учитывая изменение потенциала раствора, внешнее напряжение и падение напряжения в амперметре вычислили точное напряжение между раствором и зондовым электродом. Результаты измерений приведены на фиг. 10. Из фиг. 10 явствует, что кривую, соответствующую каждой концентрации раствора и каждому электроду, можно заменить по меньшей мере двумя прямыми. Следовательно, каждая кривая вызвана не менее чем двумя видами ионов, напряжения разложения которых эти прямые именно и определяют. Кривые, полученные для положительного платинового электрода, определяют в случае насыщенного раствора и его концентрации I моль/л напряжения разложения 510 мв и 770 мв; при концентрации раствора 0,1 и 0,01 моль/л - напряжения 770 мв и



Фиг. 10. Кривая напряжения разложения. Платиновый электрод:
1 – конц. раствора 0,01 моль/л, 2 –0,1 моль/л,
3 – 1,0 моль/л, 4 – насыщенный раствор,
5 – никелевый электрод, 0,01 моль/л

1170 мв. Вероятно, напряжением разложения хлора является 1170 мв, каковое значение было получено при разбавленных растворах. Более низкие напряжения могут быть обусловлены нечистотой хлористого калия - добавками. Их влияние тем больше, чем выше концентрация раствора.

Для отрицательного платинового электрода при всех концентрациях раствора получаются два четко выраженных значения напряжения разложения: -800 мв и -1310 мв. Так как эти значения не зависят от концентрации раствора, то можно полагать, что -800 мв относится к водороду и -1310 мв к калию.

Измерения никелевыми электродами произвели при концентрации водного раствора хлористого калия 0,01 моль/л. При отрицательной полярности были получены такие же значения напряжения разложения, как и в случае платинового электрода. К ним добавилось еще значение, определенное большими плотностями тока, - 1590 мв.

При положительном никелевом электроде для напряжения разложения получены значения -80 мв и -240 мв. Потенциал нулевого разряда он проводит уже с плотностью тока 0,8 ма/см^2. Напротив, вблизи потенциала U_0 никелевого электрода его остаточный ток очень мал. Кривая напряжения разложения никелевого электрода показывает, почему никель является неподходящим в качестве зондового электрода при водном растворе хлористого калия, а также почему для емкости его статического слоя были получены двоякие значения.

Выводы

Используя известные электронные методы измерения и серийно выпускаемую промышленностью аппаратуру, можем получить дополнительную информацию о процессах, сопровождающих погружение в водный электролитический раствор чистого и сухого электрода с гладкой поверхностью из инертного металла. Автором применялись приведенные ниже методы и были получены следующие результаты.

I. Метод потенциалов позволяет наблюдать динамику образования двойного слоя и определить абсолютное значение скачка потенциала между раствором и электродом. Этот метод можно приспособить и для исследования процессов, сопровождаемых адсорбированными на электроде побочными добавками.

2. Компенсационный метод позволяет определить потенциальный скачок двойного слоя с большой точностью. При помощи этого метода установлено, что в каломельном электроде каломель является в отношении насыщенного водного раствора хлористого калия отрицательной 345 $\pm$ 10 мв. Платина в зависимости от концентрации раствора до 100 мв положительнее чем раствор, а никель до 460 мв отрицательнее его.

3. Метод заряда позволяет определить связанный в двойном слое заряд и на основе последнего вычислить емкость C_0 статического двойного слоя. Выяснилось, что в случае водного раствора хлористого калия емкость платинового электрода C_0 = = 70 мкф/см^2 и никелевого электрода - 55 мкф/см^2. Эти емкости не зависят от концентрации раствора.

4. Измерено напряжение разложения между раствором и электродом. Выяснилось, что его значение не зависит от концентрации раствора, но зависит только от вида иона. Так были получены значения напряжения разложения: хлор +1,17 в, калий -1,31 в и водород -0,80 в.

Литература

I. П. Делахей. Двойной слой и кинетика электродных процессов. Изд. "Мир" М. 1967.

P. Plakk

MESSUNGEN AN DER STATISCHEN ELEKTRISCHEN DOPPELSCHICHT

Zusammenfassung

Bei Anwendung der in der Elektronik heutzutage bekanten Meßmethoden und industriell serienweise hergestellten Apparate kann man Auskunft erhalten über Prozesse, die beim Eintauchen einer glatten, reinen und trockenen Metallelektrode in eine Elektrolytlösung stattfinden. Der Autor hat die nachfolgenden Meßmethoden angewandt und folgende Resultate erhalten:

1. Die Potentialmethode gestattet die Verfolgung der Dynamik der Doppelschichtbildung und die Bestimmung des Absolutwertes des zwischen Lösung und Elektrode entstehenden Potentialgefälles. Diese Methode kann modifiziert werden zur Untersuchung von Prozessen, welche durch die an die Elektrode aus der Lösung absorbierten Fremdstoffe hervorgerufen werden.

2. Mit Hilfe der Komponsationsmethode kann das Potentialgefälle der Doppelschicht mit großer Genauigkeit bestimmt werden. Mit dieser Methode wurde festgestellt, daß in der Kalomelelektrode das Kalomel um 345 ± 10 mV negativer ist als die wäßrige gesättigte Kaliumchloridlösung. In Abhängigkeit von der Konzentration ist das Platin bis 100 mV positiver und das Nickel bis 460 mV negativer als die Lösung.

3. Die Ladungsmethode erlaubt die Bestimmung der in der Doppelschicht angesammelten Ladung und daraus die Kapazität C_o der statischen Doppelschicht zu berechnen.

Es erwies sich, daß in der wäßrigen Lösung von Kaliumchlorid an der Platinelektrode C_o = 70 $\mu F/cm^2$ und an der Nickelelektrode 55 $\mu F/cm^2$. C_o ist unabhängig von der Konzentration der Lösung.

4. Die Zersetzungsspannung zwischen Lösung und Elektrode wurde gemessen. Dabei wurde festgestellt, daß der Wert der Zersetzungsspannung nicht von der Konzentration der Lösung, sondern nur von der Ionenart abhängt. Es wurden folgende Werte der Zersetzungsspannung erhalten: für das Chlor +1,17 V, das Kalium -1,31 V und den Wasserstoff -0,80 V.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

СЕРИЯ А № 288 1970

УДК 541.13

П.М. Плакк

МОДЕЛЬ ЭЛЕКТРИЧЕСКОГО ДВОЙНОГО СЛОЯ

Измерению диэлектрической проницаемости веществ посвящено много научно-исследовательских работ. Представленные в монографии Ф. Эме [1] приборы способны измерять диэлектрическую проницаемость только в случае веществ с малой проводимостью, до тангенса угла потерь $tg\delta \leq 1$. При этом не обращается внимания на электрический двойной слой.

Измерение электрических свойств электролитных растворов с большой проводимостью возможно при помощи классических электродов из инертного металла, если электрические величины двйного слоя известны с достаточной точностью [2]. В связи с кинетикой электродных процессов основательно изучен двойной слой и его строение [3, глава УШ, 4]. В этих работах основной электрической характеристикой его является зависимость дифференциальной емкости от потенциала. Эти работы не дают возможности учесть влияние двойного слоя при измерении электрических свойств веществ с большой электропроводностью. Для этого необходимо разработать модель двойного слоя, согласующуюся с электрическими экспериментальными данными.

Исходя из известного фактического материала и учитывая результаты измерений [5], оказывается возможным составить для двойного слоя платинового электрода и водного раствора хлористого калия упрощенную модель. При ее составлении исходили только из электрических точек зрения. Приемлема ли такая модель с химической точки зрения, - этот вопрос здесь не рассматривался.

При составлении модели исходили из следующих суждений.

а) Электрический двойной слой образуется по принципу минимальной энергии. О вероятности наличия той или другой его формации можно судить по энергетическим соображениям.

б) Переход заряда иона на электрод невозможен вследствие препятствующего действия гидратной оболочки. Если препятствующие условия отсутствуют, то гидратизированный ион может приблизиться к электроду и к другому иону до соприкасания.

в) В случае напряжения разложения гидратная оболочка иона под действием энергии электрического поля разрывается. Ион освобождается и может передать свой заряд электроду.

г) Расстояние гидратизированного иона двойного слоя от поверхности электрода по размерности сравнимо с радиусом расположения заряда в ионе. Поэтому электрическое поле, имеющее место в двойном слое, следует рассматривать как гомогенное и рассчитывать емкость двойного слоя по формуле пластинчатого конденсатора.

д) Электрическая емкость двойного слоя является постоянной величиной, которая не зависит от концентрации раствора и вещества инертного электрода. Она зависит только от того, какие ионы образуют двойной слой. В [5] двойные слои, измеренные методом заряда, образовались очень быстро и состоят, видимо, из основных ионов раствора.

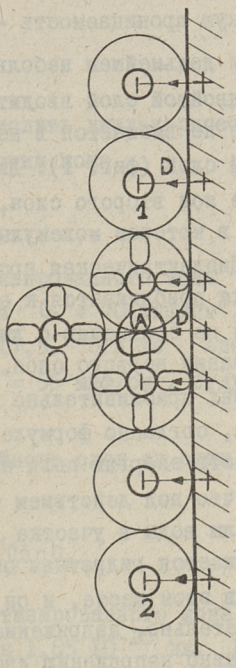
е) Аналогично электрической емкости конденсатора, статическая электроемкость двойного слоя C_0 вычисляется по накопившемуся в ней заряду Q_0 и напряжению электрода U_0 формулой $C_0 = Q_0/U_0$. Физически емкость не связана с уже накопившимся зарядом, она характеризует способность накопления заряда двойного слоя.

На металлические покрытия конденсатора можно приводить заряд в неограниченном количестве. Это относится и к металлическому покрытию двойного слоя - к инертному электроду. Напротив, находящееся в водном растворе покрытие образуется из гидратизированных ионов. Вместе с повышением концентрации ионов увеличивается также сила электрического поля в двойном слое. Начиная с известной концентрации ио-

нов, энергия электрического поля превышает энергию гидратации ионов, гидратная оболочка разрушается и начинается электролиз. В этом случае напряжение между раствором и электродом должно равняться напряжению разложения.

ж) Напряжение разложения в случае инертного электрода зависит только от вида иона [5].

Модель статического двойного слоя. Составляем ее на основе величин платинового электрода и водного раствора хлористого калия. В случае потенциала нулевого заряда двойной слой отсутствует. Если превращать платиновый электрод несколько более положительным, то к нему приближается известное число анионов до радиуса r гидратизированного иона. Эти ионы распределяются относительно поверхности электрода равномерно (см. фиг. I, ионы I и 2). При дальнейшем увеличении напряжения электрода число ионов возрастает до тех пор, пока вся поверхность электрода покроется одноионным слоем. Соответствующая картина представлена на фиг. I, где D обозначает электрическое смещение, связывающее электрические заряды разного знака. Сила электрического поля E определена выражением



Фиг. 1. В пояснение модели статического двойного слоя

$$E = \frac{D}{\varepsilon_0 \varepsilon_r}, \qquad (I)$$

причем электрическая постоянная $\varepsilon_0 = 8{,}85 \cdot 10^{-14}$ ф/см, а ε_r обозначает относительную диэлектрическую проницаемость данной среды. Электрическая емкость C_0 такого двойного слоя

является постоянной величиной, независимо от того, находимся ли мы в точке нулевого заряда (в двойном слое отсутствует заряд) или вся поверхность электрода покрыта ионами. Так как гидратная оболочка иона поляризована относительно центрального заряда, то она имеет сравнительно высокую диэлектрическую проницаемость - около I5.

При дальнейшем небольшом увеличении напряжения электрода в двойной слой вводится дополнительный заряд. Последний больше не вмещается в первый слой и начинает образовывать второй слой (фиг. I). Диэлектрическое смещение D , связывающее ион второго слоя, проходит через участок A первого слоя, в котором молекулы воды не поляризованы в отношении его. Диэлектрическая проницаемость ε_r первого слоя в этом участке приближается к единице. Она уменьшается не менее IO раз по сравнению с диэлектрической проницаемостью при связывании первого слоя. Так как значение смещения D в обоих слоях приблизительно равно, то сила электрического поля должна, согласно формуле (I), увеличиться свыше IO раз и плотность энергии поля не менее чем в IOO раз. Весьма вероятно, что под действием такого сильного электрического поля молекулы воды в участке A перевертываются в новое положение. Уничтожается гидратная оболочка иона первого слоя, находящаяся в этом месте, и он может отдать свой заряд электроду. Следовательно, напряжение разложения $U_к$ численно должно быть равно напряжению двойного слоя, которое имеет место в случае полностью выстроенного слоя толщиной в один ион.

Проверим возможность состоятельности этой модели количественными зависимостями. В случае полностью выстроенного однослойного двойного слоя ионы приблизились друг к другу до соприкасания, т.е. на расстояние $2r$. В таком случае на площадь двойного слоя в I $см^2$ помещается $1/4r^2$ иона. Тогда связанный в двойном слое заряд $Q_к$ выражается как

$$Q_к = \frac{q}{4r^2}, \qquad (2)$$

где элементарный заряд $q = I{,}6 \cdot IO^{-I9}$ К.

Напряжение двойного слоя в этом случае соответствует напряжению разложения $U_к$ и связано с емкостью двойного слоя формулой

$$U_к = \frac{Q_к}{C_o} = \frac{q}{4\pi r^2 C_o}. \qquad (3)$$

Так как $U_к$ и C_o определены проведенными в предыдущей статье [5]измерениями,то из формулы (3) можно найти радиус гидратизированного иона

$$r = 0{,}5\sqrt{\frac{q}{U_к C_o}}. \qquad (4)$$

Далее, при помощи r можно вычислить диэлектрическую проницаемость ε_r гидратной оболочки иона

$$\varepsilon_r = \frac{r C_o}{\varepsilon_o}. \qquad (5)$$

Произведя вычисления для гидратизированного иона хлора, при предпосылке, что $U_к$ = I,I7 в и C_o = 70 мкф/см^2, получаем r = 2,2 · I0^{-8} см и ε_r = I7,5. Эти же значения для калия, если $U_к$ = I,3I в и C_o = 55 мкф/см^2, будут r = = 2,4·I0^{-8}см и ε_r = I4,5.

Энергия электрического поля двойного слоя на один ион выражается формулой

$$W_{1\,ион} = \frac{C_o U_к^2}{2} : \frac{1}{4\pi r^2} = 0{,}5\, q U_к. \qquad (6)$$

Вычисляя эту величину для гидратизированного иона хлора, получаем 0,94 I0^{-I9} дж, а для калия I,05 I0^{-I9} дж. Принимая теплоту гидратации ионов хлора и калия приблизительно 80 ккал/г ион, получаем отсюда энергию гидратации на один ион 5,6 I0^{-I9} дж. Тем самым энергия гидратации одного иона приблизительно в 6 раз больше, чем энергия, приходящаяся на один ион в первом слое двойного слоя. Первый слой двойного слоя ионов может энергетически существовать. Присутствие ионов во втором слое двойного слоя невозможно, так как тогда энергия электрического поля в участке А превысит в несколько десятков раз гидратационную энергию иона.

Не размывается ли двойной слой толщиной в один ион под действием теплового движения молекул воды? Средняя энергия трансляции одной молекулы идеального газа выражается отношением $\frac{3}{2}kT$, что при комнатной температуре дает

значение $6,2\ 10^{-21}$ дж. Оно приблизительно в 15 раз меньше, чем энергия электрического поля, связывающая ион в двойном слое. Энергия теплового движения молекулы воды в растворе меньше, чем у идеального газа. Поэтому предположение, что первый слой ионов двойного слоя размывается молекулами воды, несостоятельно.

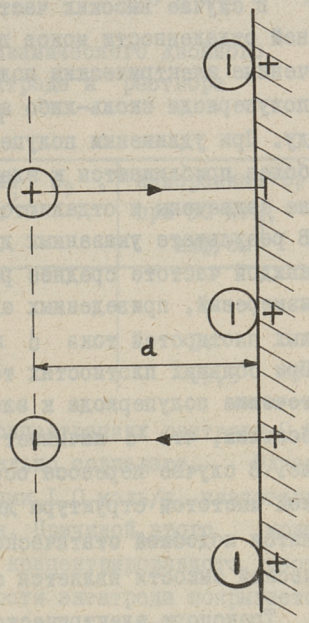
С электрической точки зрения динамическое равновесие двойного слоя не требуется, оно может быть вполне статического характера. Медленное образование двойного слоя на электрически изолированном платиновом электроде (до получаса) говорит даже против наличия динамического процесса.

С электрической точки зрения не требуется также и наличия диффузионной части двойного слоя при водных растворах сильных электролитов в случае, если скачок потенциала электрода не превышает напряжения разложения. Связанное в двойном слое число ионов весьма невелико по сравнению с числом ионов, присутствующих в растворе, также и в случае разбавленных растворов. Возьмем, например, раствор хлористого натрия концентрацией 0,01 моль/л и предположим, что на 1 $см^2$ поверхности электрода приходится около 1 $см^3$ раствора. В таком случае на каждый связанный в двойном слое ион приходится $1,4\ 10^5$ иона раствора и столь малое отсутствие равновесия не является причиной для существования диффузионного двойного слоя.

Модель динамического двойного слоя. При переменном напряжении двойной слой переводит заряд через динамическую или дифференциальную емкость C_p. Исходя из приведенной выше модели статического двойного слоя, модель динамического двойного слоя можно представить разным образом. Рассмотрим ниже некоторые варианты модели динамического двойного слоя и возможность их существования.

Вариант 1. Исходя из вышеприведенной модели статического двойного слоя, при напряжении разложения $U_к$ вся поверхность электрода покрыта слоем в один ион. При нормальном напряжении U_o электрода только часть поверхности электрода покрыта статическим двойным слоем, остальная же часть поверхности свободна (фиг. 2). Учитывая малое значение радиуса гидратизированного иона, можем предположить, что по-

крытая двойным слоем поверхность электрода относится ко всей поверхности так, как напряжение U_o электрода относится к напряжению разложения $U_к$. Пусть поверхность электрода равна I $см^2$, тогда часть поверхности, покрытая статическим двойным слоем, составляет $S_o = \frac{U_o}{U_к}$ $см^2$ и свободная от двойного слоя часть поверхности $S_1 = 1 - U_o / U_к$ $см^2$.



Фиг. 2. В пояснение модели динамического двойного слоя

Без внешнего напряжения ионы раствора распределены равномерно по всему объему раствора. Их среднее расстояние между собой и расстояние первого слоя от поверхности электрода значительно больше, чем толщина статического двойного слоя. Например, при концентрации раствора I моль/л среднее расстояние однозначных ионов составляет $I,I8 \cdot I0^{-7}$ см, а при концентрации 0,0I моль/л $5,57\ I0^{-7}$ см. Вследствие равномерного распределения ионов над всей поверхностью электрода, а также над свободной поверхностью его имеется равное число разнозначных ионов раствора. Включив на электрод переменное напряжение, к нему приводятся полупериодами "+" и "–" заряды. Смещение заряда на поверхности электрода происходит беспрепятственно (поверхность равного потенциала). Приводимые заряды располагаются на свободной поверхности электрода таким образом, чтобы можно было связать находящийся над поверхностью разнозначный ион (фиг. 2). Это связывание происходит через среду молекул воды, диэлектрическая проницаемость которой $\varepsilon_r = 80$. Образуется картина поля, позволяющая без большой ошибки вычислить также динамическую емкость C_p двойного слоя по формуле пластинчатого конденсатора. Активной поверхностью покрытия конденсатора остается половина свободной поверхности $S_1/2$ и

$$C_p = \frac{80 \cdot 8{,}85 \cdot 10^{-14} \cdot S_1}{2d} \quad \Phi/см^2, \qquad (7)$$

причем d обозначает среднее расстояние от электрода первого слоя ионов раствора.

В случае высоких частот d можно принять равным средней отдаленности ионов друг от друга, так как ионы, охваченные электрическим полем, не успевают в течение короткого полупериода сколь-либо значительно приблизиться к электроду. При удлинении полупериода захваченные ионы все более и более приближаются к электроду. Во втором полупериоде они не захвачены и отдаляются от электрода вследствие диффузии. В результате указанных движений образуется соответствующее каждой частоте среднее расстояние d первого слоя ионов. Из измерений, приведенных автором, выяснилось, что в случае малых плотностей тока d пропорционально логарифму частоты. При больших плотностях тока и низких частотах приводимый в течение полупериода к электроду заряд является настолько большим, что d начинает уменьшаться быстрее отмеченного выше. В случае переноса особенно больших зарядов с очень низкой частотой структура динамического двойного слоя становится подобной статическому и предельным значением динамической емкости является статическая емкость двойного слоя.

Транспорт электрического заряда происходит также вследствие движения иона, охваченного полем. Этот процесс происходит параллельно с переносом заряда емкостью C_p, имеет активный характер и на эквивалентной схеме двойного слоя представлен через параллельное сопротивление R_p.

В момент, когда сумма внешнего напряжения U и напряжения электрода U_o превышает напряжение разложения $U_к$, начинается сильный ток проводимости через двойной слой и R_p резко понижается. Согласно эквивалентной схеме начинается, с одной стороны, сильное разряжение C_p через R_p и, с другой стороны, искажение вида синусоидального тока. С измерительно-технической стороны оба эти эффекта могут дать кажущееся увеличение C_p, даже свыше значения C_o.

Сравним в случае платинового электрода и водного раствора хлористого калия вычисленные и измеренные автором по этому варианту значения C_p. Результаты представлены в таблице I.

Таблица I

Сравнение емкости C_p динамического двойного слоя при платиновом электроде и растворе хлористого калия

Концентрация раствора, моль/л	Вычисленная C_p, мкф/см^2	Измеренная C_p при 50 кгц, мкф/см^2
0,0I	5,8	7
0,I	I4,7	I4
I,0	28	I9,5

Из таблицы явствует, что при концентрациях раствора 0,0I и 0,I моль/л результаты приблизительно совпадают. Однако емкость, измеренная при концентрации I,0 моль/л, значительно меньше ее вычисленного значения. Причиной этого может быть обстоятельство, что в случае концентрированного раствора одна часть свободной поверхности электрода покрывается ионами добавок адсорбированного хлористого калия и вместе со свободной поверхностью емкость уменьшается.

Экстраполируя значение емкости, измеренное автором при малой плотности тока (0,09 ма/см) на частоту I гц, получаем C_p = 42 мкф/см^2. Так как $C_p < C_o$, то и при этой частоте еще существует статический двойной слой.

Если взять за основу приведенную выше модель двойного слоя, то утверждение о том, что минимум динамической емкости имеет место в точке нулевого заряда электрода, несостоятельно. Сомнения о правильности этого утверждения высказаны также в [4, стр. I40 последний абзац].

Рассматривались вариант 2 (что перенос переменного тока происходит путем увеличения и уменьшения заряда статического двойного слоя, соответственно включенному на электрод внешнему напряжению) и вариант 3 (заряд переменного тока

переносится ионами, которые находятся за статическим двойным слоем во втором слое). Энергетические расчеты показали, что эти варианты не достоверны.

В целях выяснения достоверности модели динамического двойного слоя был проведен ряд непосредственных измерений при помощи переменного тока четырехугольной и синусоидальной формы. Методы и результаты измерений будут опубликованы в следующем сообщении. Последние подтверждают, что с электрической точки зрения модель динамического двойного слоя, представленная в первом варианте, правдоподобна. Эта модель поясняет также зависимость емкости динамического двойного слоя от плотности тока, на что обращалось внимание в [2].

Выводы

I. Исходя из общепринятых положений и полученных при измерении результатов была составлена модель статического двойного слоя. Достоверность модели проверялась энергетическими вычислениями. При помощи модели вычислили радиус r гидратизированного иона и диэлектрическую проницаемость ε_r. Для хлора получены $r = 2{,}2\ 10^{-8}$ см и $\varepsilon_r = 17{,}5$; для калия $r = 2{,}4\ 10^{-8}$ см и $\varepsilon_r = 14{,}5$.

2. Составлена модель динамического двойного слоя и ее достоверность проверена путем соответствующих непосредственных измерений, проведенных автором при помощи переменного тока. Эти методы измерений и полученные результаты будут опубликованы в следующем сообщении. Выяснилось, что динамический двойной слой существует параллельно со статическим, на свободной от последнего поверхности. Если емкость C_o статического двойного слоя вычисляется из диэлектрической проницаемости гидратизированного иона, которая составляет около 15, то динамическая емкость C_p вычисляется диэлектрической проницаемостью воды. Определены предельные значения емкости динамического двойного слоя в зависимости от частоты и плотности тока.

Литература

I. Ф. Эме. Диэлектрические измерения. Изд. "Химия", М. 1967.

2. П.М. Плакк. Об измерении электрических свойств электролитов с большой проводимостью. Труды Таллинского политехнического института, серия А, № 268, 1968, 73.

3. В.В. Скорчеллетти. Теоретическая электрохимия. Госхимиздат, Ленинград 1963.

4. П. Делахей. Двойной слой и кинетика электродных процессов. Изд. "Мир" М. 1967.

5. П.М. Плакк. Измерения при статическом вдойном электрическом слое. См. наст.сборник, стр. 19.

P. Plakk

MODELL DER ELEKTRISCHEN DOPPELSCHICHT

Zusammenfassung

1. Ausgehend von allgemein bekannten Standpunkten und den erhaltenen Meßresultaten wird ein Modell der statischen Doppelschicht gegeben. Die Glaubwürdigkeit des Modells wurde mit energetischen Berechnungen kontrolliert. Nach dem Modell wurden der Radius des hydratisierten Ions r und die Dielektrizitätskonstante berechnet. Für das Chlorion erhielt man $r=2{,}2\cdot10^{-8}$ cm und $\varepsilon_r = 17{,}5$; für das Kaliumion $r=2{,}4\cdot10^{-8}$ cm und $\varepsilon_r = 14{,}5$.

2. Es wurde auch ein Modell der dynamischen Doppelschicht gegeben und dessen Glaubwürdigkeit durch entsprechende, vom Autor durchgeführte direkte Messungen mit Wechselstrom kontrolliert. Diese Meßmethoden und Meßresultate werden in einer folgenden Arbeit veröffentlicht. Es erwies sich, daß die dynamische Doppelschicht parallel mit der statischen an der von der letzteren freien Oberfläche existiert. Wenn die Kapazität der statischen Doppelschicht C_o mit der Dielektrizitätskonstante der hydratisierten

Ions berechnet wird (ca 15). so wird die Kapazität der dynamischen Schicht C_p mit der Dielektrizitätskonstante des Wassers berechnet (ε_r =80). Es wurden die Grenzwerte der dynamischen Kapazität in Abhängigkeit von der Frequenz und der Stromdichte bestimmt.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

С Е Р И Я А № 288 1970

УДК 681.2:519.272

О.М. Пикков

МАТЕМАТИЧЕСКАЯ МОДЕЛЬ СТАЦИОНАРНОГО ЧАСТОТНО-ИМПУЛЬСНОГО СИГНАЛА

Введение

Частотно-импульсный сигнал в принципе представляет собой дискретные приращения фазы $\Delta\varphi$ некоторого гармонического частотного сигнала, частота $F(t)$ которого пропорциональна контролируемому параметру

$$F(t) = kx(t), \qquad (1)$$

где k – коэффициент пропорциональности.

Появляющиеся в результате этого процесса импульсы отмечают моменты окончания отдельных циклов приращения $\Delta\varphi$ [1]. Дискретные циклы фазы частотного сигнала связаны с временным интервалом $\vartheta(i)$ между двумя последующими импульсами

$$\vartheta(i) = t_i - t_{i-1} = \frac{\Delta\varphi}{2\pi F(t)_i}, \qquad (2)$$

где $F(t)_i$ – средняя частота сигнала в i-ном интервале импульсов.

Наиболее естественное представление о частотно-импульсном сигнале получаем, предполагая, что циклы приращения

$$\Delta\varphi = 2\pi. \qquad (3)$$

Тогда частота появления импульсов совпадает с частотой этого фиктивного гармонического сигнала $F(t)$.

Интервалы между соседними импульсами образуют своеобразный время-импульсный сигнал, в котором каждый предыдущий импульс t_{i-1} является тактовым в отношении текущего,

сигнального импульса t_i.

Информация содержится в длительности интервала времени $\vartheta(i)$ между "тактовым" t_{i-1} и сигнальным t_i импульсами.

Контролируемый параметр $x(t)$ в (I) является непрерывной случайной величиной. Согласно (I) и (2) непрерывными случайными величинами окажутся также $F(t)$ и $\vartheta(i)$.

Длительность интервалов $\vartheta(i)$ между импульсами функционально связана с частотой импульсов $F(t)$ и контролируемым параметром $x(t)$, согласно (2) и (3).

$$\vartheta(i) = \frac{1}{F(t)_i} = \frac{1}{k\overline{x(t)}_i}, \qquad (4)$$

где $F(t)_i$ и $\overline{x(t)}_i$ – средние значения частоты импульсов и контролируемого параметра в интервале $\vartheta(i)$ соответственно.

Внутренняя структура самого частотно-импульсного сигнала

а) Флуктуация момента появления импульса. Последовательность импульсов частотно-импульсного сигнала даже при постоянной величине контролируемого параметра $x(t) = C = \text{const}$ не является строго регулярной.

Возникновение каждого импульса в последовательности сопровождается некоторыми случайными воздействиями, которые могут вызвать смещение моментов возникновения отдельных импульсов. Время корреляции этих воздействий обычно небольшое по сравнению с периодом следования импульсов и случайные смещения (флуктуации) отдельных импульсов $\xi(t_i)$ можно считать независимыми. Поэтому значения корреляционных моментов этой случайной величины $\xi(t_i)$

$$K_{\xi ij} = M[\xi_i \xi_j] = \sigma_\xi^2 \delta(j-i), \qquad (5)$$

а закон распределения ее имеет симметричный характер

$$m_\xi = M[\xi(t_i)] = 0. \qquad (6)$$

Поскольку эта флуктуация вызвана многими малыми случайными воздействиями, то можно предполагать предельное, приближающееся к нормальному закону распределение случайной величины $\xi(t_i)$ с плотностью

$$w_1(\xi)=\frac{1}{\sqrt{2\pi}\,\sigma_\xi}\exp\left[-\frac{\xi(t_i)^2}{2\sigma_\xi^2}\right] \tag{7}$$

и с дисперсией, согласно (5),

$$D_\xi=\sigma_\xi^2. \tag{8}$$

Такими являются вероятностные характеристики момента появления импульса сигнала. Вся последовательность импульсов с учетом флуктуации $\xi(t_i)$ образует периодический нестационарный случайный процесс [2] с n-мерной плотностью распределения

$$\begin{aligned} w_n &= (\xi_1, t_1; \xi_2, t_2; \ldots; \xi_n, t_n) = \\ &= w_n(\xi_1, t_1+T; \xi_2, t_2+T; \ldots; \xi_n, t_n+T). \end{aligned} \tag{9}$$

б) Образующийся время-импульсный сигнал. Длительность интервала между импульсами при $x(t)=C=\text{const}$, учитывая флуктуацию $\xi(t_i)$ моментов появления импульсов,

$$\vartheta(i)=(t_i+\xi_i)-(t_{i-1}+\xi_{i-1})=T+\xi_i-\xi_{i-1}=T+\eta(i) \tag{10}$$

имеет в своем составе случайную величину $\eta(i)$, являющуюся разностью двух независимых случайных величин $\xi(t_i)$

$$\eta(i)=\xi_i-\xi_{i-1}, \tag{11}$$

которые обе распределены нормально (7).

Плотность вероятности длительности интервала

$$w_1(\vartheta)=w_1(\eta)=\frac{1}{\sqrt{2\pi}\,\sigma_\eta}\exp\left[-\frac{\eta(i)^2}{2\sigma_\eta^2}\right], \tag{12}$$

где $\sigma_\eta=\sigma_\vartheta=\sqrt{\sigma_{\xi i}^2+\sigma_{\xi i-1}^2}$ – среднеквадратическое отклонение длительности интервала между двумя последующими импульсами.

Поскольку у нестационарного периодического случайного процесса $\xi(t_i)$ все его характеристики периодически повторяются (9), то

$$\sigma_{\xi i}=\sigma_{\xi i-1}=\sigma_\xi$$

и согласно (6)

$$m_{\xi i} = m_{\xi i-1} = m_{\xi} = 0.$$

Поэтому

$$\sigma_{\eta} = \sqrt{2}\sigma_{\xi}; \quad \sigma_{\eta}^2 = 2\sigma_{\xi}^2; \quad D_{\eta} = 2D_{\xi} \tag{13}$$

и

$$m_{\eta} = m_{\xi i} + m_{\xi i-1} = 0. \tag{14}$$

Следовательно, длительность интервала $\vartheta(i)$ оказывается случайной величиной, даже если контролируемый параметр $x(t) = C = \text{const}$ и составляющая T, определяющая информационное содержание сигнала, является детерминированной величиной (10). Дисперсия длительности интервала между двумя последующими импульсами равна двойной дисперсии момента появления импульса (13).

Корреляционные моменты длительности интервалов, учитывая (5) и (10)

$$K_{\vartheta ij} = K_{\eta ij} = M\left[(\xi_i - \xi_{i-1})(\xi_j - \xi_{j-1})\right] =$$

$$= M(\xi_i\xi_j - \xi_i\xi_{j-1} - \xi_{i-1}\xi_j + \xi_{i-1}\xi_{j-1}) =$$

$$= \begin{cases} 2\sigma_{\xi}^2 & \text{при } j = i; \\ -\sigma_{\xi}^2 & \text{при } |j-i| = 1; \\ 0 & \text{при } |j-i| \geq 2. \end{cases} \tag{15}$$

Полученная в (15) отрицательная корреляция указывает на существование сильной линейной связи между соседними интервалами $\vartheta(i)$ и $\vartheta(j)$.

Интервалы $\vartheta(i)$ между импульсами составляют стационарную случайную последовательность, потому что многомерная плотность вероятности

$$w_n\left[\vartheta_1, iT; \vartheta_2, (i+1)T; \ldots; \vartheta_n, (i+n-1)T\right] =$$

$$w_n\left[\vartheta_1, (i+1)T; \vartheta_2, (i+2)T; \ldots; \vartheta_n, (i+n)T\right]$$

является периодической функцией с периодом T от дискретного аргумента i.

Корреляционная матрица этой n-мерной стационарной периодической случайной последовательности при $x(t)=C=\text{const}$ и $T=\text{const}$ симметрична в отношении главной диагонали и имеет кроме нее отличные от нуля элементы только в смежных с ней позициях (I5).

$$\| K_{\eta ij} \| = D_\eta \| r_{\eta ij} \|, \qquad (I6)$$

где $\| r_{\eta ij} \|$ – нормированная корреляционная матрица случайной последовательности $\vartheta(i)$,

$$\| r_{\eta ij} \| = \begin{Vmatrix} 1 & -0{,}5 & 0 & \dots & 0 \\ & 1 & -0{,}5 & \dots & 0 \\ & & 1 & \dots & 0 \\ & & & \dots & \\ & & & & 1 \end{Vmatrix} \qquad (I7)$$

Последовательность интервалов $\vartheta(i)$ характеризуется n-мерной нормальной функцией распределения [2] с плотностью

$$w_n(\vartheta_1 \vartheta_2 \dots \vartheta_n) = \frac{1}{(2\pi D_\eta)^{n/2}\sqrt{\Delta_\eta}} \exp\left(-\frac{1}{2\Delta_\eta}\sum_{i=1}^{n}\sum_{j=1}^{n}\Delta_{\eta ij}\frac{\eta_i \eta_j}{D_\eta}\right), \qquad (I8)$$

где Δ_η и $\Delta_{\eta ij}$ – детерминант и алгебраические дополнения матрицы коэффициентов корреляции $\| r_{\eta ij} \|$.

Установившийся режим контролируемого объекта

Характер изменения контролируемого параметра $x(t)$ зависит от состояния объекта управления. Установившемуся режиму работы объекта соответствует стационарный характер контролируемого параметра $x(t)$.

В этом режиме отклонение параметра $x(t)$ от заданного значения невелико и носит случайный характер. Математическое ожидание значения параметра $m_x=\text{const}$ и дисперсия $D_x=\text{const}$.

Случайная функция $x(t)$ в этом случае может быть представлена как сумма двух составляющих

$$x(t) = C + \nu(t), \qquad (19)$$

где $C = m_x$ – неслучайная величина,

$\nu(t)$ – стационарная случайная функция нормального распределения с математическим ожиданием $m_\nu = 0$ и дисперсией $D_\nu = \sigma_\nu^2 = D_x$.

Находим длительность интервалов $\vartheta(i)$, соответствующую значению функции $x(t)$.

Функциональная зависимость $\vartheta(i) = f(x)$ (4) является нелинейной. В силу малости случайных возмущений параметра $x(t)$, $\vartheta(i) = f(x)$ может быть линеаризована.

После замены нелинейной функциональной зависимости участком касательной в точке, соответствующей математическому ожиданию аргумента, случайные величины ϑ и x окажутся связанными линейно.

$$\vartheta(i) = f(m_x) + f'(m_x)\cdot(x - m_x) =$$

$$= \frac{1}{k m_x} + \frac{\partial}{\partial x}\left(\frac{1}{kx}\right)_{m_x}\cdot\left(x - m_x\right) = \frac{1}{k m_x} - \frac{x - m_x}{k m_x^2}. \qquad (20)$$

Стационарная случайная величина ϑi (20) может быть разложена на две составляющие

$$\vartheta(i) = T + \varepsilon(i), \qquad (21)$$

где T – неслучайная величина,

ε – случайная величина, имеющая нормальную плотность распределения с математическим ожиданием $m_\varepsilon = 0$ и дисперсией $D_\varepsilon = \sigma_\varepsilon^2$.

Математическое ожидание этой линейной функции можно найти, если подставить в ее выражение (20) значение $x = m_x = C$. Получим:

$$m_\vartheta = \frac{1}{k m_x} = \frac{1}{kC} = T. \qquad (22)$$

Дисперсия величины ϑ определяется как

$$D_\vartheta = D_\varepsilon = \left[f'(m_x)\right]^2 D_x = \frac{D_x}{k^2 m_x^4} = \frac{D_\nu}{k^2 C^4} = D_\nu k^2 T^4. \qquad (23)$$

Корреляционная функция исходного нормального процесса (19) обычно имеет экспоненциальный характер

$$K_x(\tau) = K_y(\tau) = K_y(0)\, e^{-\alpha|\tau|} = D_y e^{-\alpha|\tau|} = D_y \rho_x(\tau), \qquad (24)$$

где $\rho_x(\tau) = \dfrac{K_x(\tau)}{D_y}$ – нормированная корреляционная функция процесса $x(t)$.

Корреляционные моменты K_{xij} дискретов исходного процесса $x(t)$ можно найти, если подставим в (24)

$$\tau = t_j - t_i = (j-i)T.$$

В результате находим корреляционную матрицу

$$\| K_{xij} \| = \| K_{yij} \| = D_y \| r_{yij} \|, \qquad (25)$$

где $\| r_{yij} \|$ – нормированная корреляционная матрица дискретов процесса $x(t)$

$$\| r_{yij} \| = \begin{Vmatrix} 1 & e^{-\alpha|T|} & e^{-\alpha|2T|} & e^{-\alpha|3T|} & \dots \\ & 1 & e^{-\alpha|T|} & e^{-\alpha|2T|} & \dots \\ & & 1 & e^{-\alpha|T|} & \dots \\ & & & \dots\dots\dots\dots & \\ & & & \dots\dots\dots\dots & \\ & & & & 1 \end{Vmatrix} \qquad (26)$$

Элементы этой матрицы

$$r_{yij} = \rho_x(\tau) \text{ при } \tau = (j-i)T.$$

Корреляционная матрица случайной последовательности интервалов $\vartheta(i)$ (2I)

$$\| K_{\vartheta ij} \| = D_\varepsilon \| r_{\varepsilon ij} \| = \frac{D}{k^2 C^4} \| r_{yij} \| = D_y k^2 T^4 \| r_{yij} \|, \qquad (27)$$

где $\| r_{\varepsilon ij} \| = \| r_{yij} \|$ – нормированная корреляционная матрица последовательности $\vartheta(i)$.

Зная корреляционную матрицу (27) стационарной случайной последовательности $\vartheta(i)$ и предполагая закон распределения вероятности его нормальным (2I), можем выразить плотность распределения этого процесса в виде многомерной плотности нормального закона аналогично (I8)

$$w_n(\vartheta_1 \vartheta_2 \dots \vartheta_n) = \frac{1}{(2\pi D_\varepsilon)^{n/2}\sqrt{\Delta_\varepsilon}} \exp\left(-\frac{1}{2\Delta_\varepsilon} \sum_{i=1}^{n} \sum_{j=1}^{n} \Delta_{\varepsilon ij} \frac{\varepsilon_i \varepsilon_j}{D_\varepsilon}\right), \qquad (28$$

где Δ_ε и $\Delta_{\varepsilon ij}$ – детерминант и алгебраические дополнения матрицы коэффициентов корреляции $\| r_{\varepsilon ij} \|$ (26).

Полное математическое описание образуемого время-импульсного сигнала контроля в стационарном режиме контролируемого параметра $x(t)$ определяется как сумма двух независимых процессов (последовательностей) – (II) и (2I). Эти процессы между собой совершенно независимы, оба распределены нормально: (I8) и (28). Поэтому математическое ожидание суммарного процесса из (I4) и (22)

$$m_{\eta\varepsilon} = m_\eta + m_\varepsilon = T = \frac{1}{kC}. \qquad (29)$$

Корреляционная матрица композиции этих двух случайных процессов (длительности результирующих интервалов) из (I6) и (27)

$$\| K_{\eta\varepsilon ij} \| = \| K_{\eta ij} \| + \| K_{\varepsilon ij} \| = D_\eta \| r_{\eta ij} \| + D_\varepsilon \| r_{\varepsilon ij} \| = $$
$$= D_\eta \| r_{\eta ij} \| + D_\nu k^2 T^4 \| r_{\nu ij} \|. \qquad (30)$$

Плотность n-мерного распределения результирующего процесса при установившемся режиме контролируемого объекта

$$w_n(\vartheta_1, \vartheta_2 \dots \vartheta_n) = \frac{1}{(2\pi D_{\eta\varepsilon})^{n/2}\sqrt{\Delta_{\eta\varepsilon}}} \exp\left[-\frac{1}{2\Delta_{\eta\varepsilon}} \sum_{i=1}^{n}\sum_{j=1}^{n} \Delta_{\eta\varepsilon ij} \frac{(\eta_i+\varepsilon_i)(\eta_j+\varepsilon_j)}{D_{\eta\varepsilon}}\right], \qquad (3I)$$

где $\Delta_{\eta\varepsilon}$ и $\Delta_{\eta\varepsilon ij}$ – детерминант и алгебраические дополнения матрицы коэффициентов корреляции $\| r_{\eta\varepsilon ij} \|$ с элементами

$$r_{\eta\varepsilon ij} = \frac{K_{\eta\varepsilon ij}}{\sqrt{D_\eta D_\varepsilon}}, \qquad (32)$$

$D_{\eta\varepsilon} = D_\eta + D_\varepsilon$ – дисперсия совместного случайного процесса $(\eta + \varepsilon)$.

З а к л ю ч е н и е

Образующийся из частотно-импульсного время-импульсный сигнал контроля при стационарном режиме контролируемого параметра $x(t)$ имеет в своем составе один регулярный и два случайных компонента.

Регулярный компонент T, который является постоянной величиной, служит носителем полезной информации.

Первый случайный компонент η возникает в самом процессе преобразования. Величина его определяется шумом преобразования (конструкцией преобразователя).

Второй случайный компонент содержится в контролируемом параметре $x(t)$ и переносится преобразователем в образуемый время-импульсный сигнал контроля. Величина второго компонента ε функционально связана со случайной составляющей процесса $x(t)$.

Результирующая последовательность интервалов $\vartheta(i)$ между соседними импульсами частотно-импульсного сигнала может быть характеризована многомерным нормальным распределением.

Литература

I. В.Г. Кнорринг. О механизме работы измерительных преобразователей с частотным выходом. Измерительная техника, № 8, 1966, 51-54.

2. Б.Р. Левин. Теоретические основы статистической радиотехники. Книга вторая. Советское Радио, Москва 1968, 198-200.

O. Pikkov

A MATHEMATICAL MODEL OF THE STATIONARY PULSE-FREQUENCY SIGNAL

Summary

This article describes the inherent structure of ideal pulse-frequency signal on its stationary condition. A Gaussian distribution of the random fluctuations in controlled parameter and noise in converter is assumed.

The statistical approach to the description of pulse-frequency signal allows to split it into three statistically independent components. One of these is a carrier of the useful information, the two others represent its inherent and converted noise.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

СЕРИЯ А № 288 1970

УДК 62 507

В.А.Крахт, Г.Э.Якобсон

К ЗАДАЧЕ ПОСЛЕДОВАТЕЛЬНОЙ ДЕКОМПОЗИЦИИ КОНЕЧНЫХ АВТОМАТОВ

I. Введение

Одни из первых работ по последовательной декомпозиции автоматов принадлежат Ю. Хартманису [1] и А. Гиллу [2]. В работе [1] предложен метод декомпозиции, основанный на нахождении разбиения с подстановочным свойством (СП-разбиения) на множестве состояний автомата. В работе [2] использован матричный метод декомпозиции. На тесную связь, существующую между этими двумя методами, указано в работе [3]. В настоящей статье применено несколько иное по сравнению с работой [3] понятие последовательной декомпозиции, дающее возможность значительно расширить класс автоматов, для которых такая декомпозиция существует. Усовершенствован также способ построения автоматов, образующих последовательную декомпозицию.

Мы воспользуемся автоматом Мили, но полученные результаты легко могут быть распространены и на автоматы Мура.

Как обычно, под автоматом Мили будем понимать пятерку $A=(X,S,Y,\delta,\lambda)$ (X - входной алфавит, S -множество состояний, Y - выходной алфавит, $\delta: S\times X \to S$ - функция переходов, $\lambda: S\times X \to Y$ - функция выходов).

Определение I. $A'=(X',S',Y',\delta',\lambda')$ называется подавтоматом автомата $A=(X,S,Y,\delta,\lambda)$ ($A'\subseteq A$) если и только если

$$X'\subseteq X,\ S'\subseteq S,\ \delta'(S',X')\subseteq S'$$

и для любого $x \in X'$ и $s \in S'$

$$\delta'(s,x) = \delta(s,x), \lambda'(s,x) = \lambda(s,x).$$

Определение 2. Автоматы $A = (X, S, Y, \delta, \lambda)$ и $A' = (X', S', Y', \delta', \lambda')$ называются изоморфными, если и только если существуют три взаимно-однозначных отображения $\Psi_X: X \to X'$, $\Psi_S: S \to S'$, $\Psi_Y: Y \to Y'$, таких что для любого $x \in X$ и $s \in S$:

$$\Psi_S(\delta(s,x)) = \delta'(\Psi_S(s), \Psi_X(x)),$$

$$\Psi_Y(\lambda(s,x)) = \lambda'(\Psi_S(s), \Psi_X(x)).$$

Тройка $\Psi = (\Psi_X, \Psi_S, \Psi_Y)$ называется изоморфизмом между автоматами A и A', и пишется $A \overset{\Psi}{\sim} A'$.

Определение 3. Разбиение π на множестве состояний S автомата $A = (X, S, Y, \delta, \lambda)$ называется разбиением с подстановочным свойством (СП-разбиением), если и только если для любого $P \in \pi$ и $x \in X$ существует $P' \in \pi$ такой, что

$$\delta(P, x) \subseteq P'.$$

Очевидно, на множестве S всегда существуют два тривиальных СП-разбиения - поэлементное разбиение 0 и целое разбиение I. Пусть π некоторое разбиение на множестве S. Обозначим через $\bar{\pi}$ наибольшее разбиение на том же множестве S такое, что $\pi \cdot \bar{\pi} = 0$.

2. Теорема последовательной декомпозиции

Определение 4. Последовательной композицией автоматов $A_1 = (X_1, S_1, Y_1, \delta_1, \lambda_1)$ и $A_2 = (X_2, S_2, Y_2, \delta_2, \lambda_2)$ для которых $X_1 = X_2$, называется автомат $A = A_1 \circ A_2 = (X, S, Y, \delta, \lambda)$, где $X = X_1$, $S = S_1 \times S_2$, $Y = Y_2$ и для всех $(s, r) \in S_1 \times S_2$ и $x \in X$

$$\delta((s,r),x) = (\delta_1(s,x), \delta_2(r, \lambda_1(s,x))),$$

$$\lambda((s,r)x) = \lambda_2(r, \delta_1(s,x)).$$

Из определения 4 непосредственно следует ассоциативность операции последовательной композиции, т.е. $A_1 \circ (A_2 \circ A_3) = (A_1 \circ A_2) \circ A_3$.

Определение 5. Автомат $A_1 \circ A_2$ называется последовательной декомпозицией автомата A, если и только если ав -

томат A изоморфен некоторому подавтомату $A' \subseteq A_1 \circ A_2$.

Последовательную декомпозицию $A_1 \circ A_2$ считают нетривиальной, если и только если $|S_1| < |S|$ и $|S_2| < |S|$ ($|S|$ -мощность множества S).

В [I] и [3] различными друг от друга способами показана справедливость следующей теоремы.

Теорема. Автомат $A=(X,S,Y,\delta,\lambda)$ разложим на нетривиальную последовательную декомпозицию $A_1 \circ A_2$,если и только если существует нетривиальное СП-разбиение на множестве состояний S .

Автоматы A_1 и A_2 , образующие последовательную декомпозицию $A_1 \circ A_2$, строятся на базе СП-разбиения π и разбиения $\bar{\pi}$ следующим образом:

I. $A_1=(X_1,S_1,Y_1,\delta_1,\lambda_1)$, $X_1=X$, $S_1=\pi$, $Y_1=\pi\times X$,

$$\delta_1(P,x)=P' \Longleftrightarrow \delta(P,x)\subseteq P',$$

$$\lambda_1(P,x)=(P,x),$$

где $P,P'\in\pi$, $x\in X$.

2. $A_2=(X_2,S_2,Y_2,\delta_2,\lambda_2)$, $X_2=\pi\times X$, $S_2=\bar{\pi}$, $Y_2=Y$,

$$\delta_2(Q(P,x))=Q' \Longleftrightarrow \delta(Q\cap P,x)\subseteq Q',$$

$$\lambda_2(Q(P,x))=\lambda(Q\cap P,x),$$

где $Q,Q'\in\bar{\pi}$, $P\in\pi$, $x\in X$.

Из способа построения автоматов A_1 и A_2 следует, что состоянию $s\in P\cap Q$ ($P\in\pi, Q\in\bar{\pi}$) автомата A соответствует состояние P автомата A_1 и состояние Q автомата A_2 .

Обратим внимание на то, что автомат A_2 может оказаться неполностью определенным, даже при полной определенности разложимого актомата A. Действительно, если для некоторых $P\in\pi$ и $Q\in\bar{\pi}$ $P\cap Q=\emptyset$, то неопределенность функции δ и λ на паре $(P\cap Q,x)$ влечет неопределенность функций δ_2 и λ_2 на паре $(Q,(P,x))$. Нетрудно убедиться, что только при условии равномощных классов разбиения π автомат A_2 будет полностью определен. Необходимым для этого является условие, что число состояний в множестве S автомата A должно быть четным числом.

3. Минимизация промежуточного алфавита

Назовем входной (выходной) алфавит автомата $A_2(A_1)$ промежуточным алфавитом. Часто промежуточный алфавит удается сократить - минимизировать - путем разбиения его на классы неразличимых входных букв и замены каждого такого класса одной буквой, принадлежащей этому классу.

Определение 5. Две буквы $x,w \in X$ называются неразличимыми в автомате $A(X,S,Y,\delta,\lambda)$ если и только если для любого $s \in S$

$$\delta(s,x)=\delta(s,w),\quad \lambda(s,x)=\lambda(s,w).$$

Отношение неразличимости букв входного алфавита рефлексивно, симметрично и транзитивно, т.е. является отношением эквивалентности. Отношение неразличимости букв входного алфавита определяет разбиение ε на алфавите X такое, что для любого класса $E \in \varepsilon$ и любой буквы $s \in S$

$$\delta(s,E)=s',\ \lambda(s,E)=y,$$

где

$$s' \in S,\ y \in Y.$$

Разбиение ε может быть легко найдено по таблицам δ и λ (таблицам переходов и выходов) автомата. Из определения 5 следует, что в один класс E разбиения ε входят только такие буквы входного алфавита, соответствующие которым столбцы как в таблице δ, так и в таблице λ имеют одинаковые заполнения.

Существующие в автомате A_2 неопределенности целесообразно доопределить так, чтобы получить наибольшее разбиение ε.

Пример. Найдем последовательную декомпозицию $A_1 \circ A_2$ для автомата A, заданного своими таблицами δ и λ.

	x_1	x_2	x_3
s_1	s_1	s_5	s_3
s_2	s_5	s_7	s_4
s_3	s_6	s_3	s_7
s_4	s_7	s_2	s_3
s_5	s_5	s_4	s_1
s_6	s_2	s_1	s_4
s_7	s_7	s_6	s_3

δ

	x_1	x_2	x_3
s_1	y_1	y_2	y_3
s_2	y_3	y_3	y_2
s_3	y_2	y_1	y_2
s_4	y_3	y_2	y_2
s_5	y_1	y_3	y_3
s_6	y_2	y_1	y_1
s_7	y_1	y_2	y_1

λ

Методом Хартманиса на множестве состояний S найдено СП-разбиение $\pi = \{\{s_1, s_4, s_7\}, \{s_2, s_5, s_6\}, \{s_3\}\}$. Выберем $\bar{\pi} = \{\{s_1, s_2, s_3\}, \{s_4, s_5\}, \{s_6, s_7\}\}$.

Строим автоматы A_1 и A_2.

I. $X_1 = \{x_1, x_2, x_3\}$, $S_1 = \{P_1, P_2, P_3\}$, $(P_1 = \{s_1, s_4, s_7\}$, $P_2 = \{s_2, s_5, s_6\}$, $P_3 = \{x_3\})$, $Y_1 = \{(P_1 x_1), (P_1 x_2), \ldots, (P_3 x_3)\}$.

	x_1	x_2	x_3
P_1	P_1	P_2	P_3
P_2	P_2	P_1	P_1
P_3	P_2	P_3	P_1

δ_1

	x_1	x_2	x_3
P_1	P_1x_1	P_1x_2	P_1x_3
P_2	P_2x_1	P_2x_2	P_2x_3
P_3	P_3x_1	P_3x_2	P_3x_3

λ_1

2. $X_2 = Y_1$, $S_2 = \{Q_1, Q_2, Q_3\}$ $(Q_1 = \{s_1, s_2, s_3\}$, $Q_2 = \{s_4, s_5\}$, $Q_3 = \{s_6, s_7\})$, $Y_2 = \{y_1, y_2, y_3\}$.

	P_1x_1	P_1x_2	P_1x_3	P_2x_1	P_2x_2	P_2x_3	P_3x_1	P_3x_2	P_3x_3
Q_1	Q_1	Q_2	Q_1	Q_2	Q_3	Q_2	Q_3	Q_1	Q_3
Q_2	Q_3	Q_1	Q_1	Q_2	Q_2	Q_1	(Q_2)	(Q_3)	(Q_2)
Q_3	Q_3	Q_3	Q_1	Q_1	Q_1	Q_2	(Q_1)	(Q_3)	(Q_1)

δ_2

	P_1X_1	P_1X_2	P_1X_3	P_2X_1	P_2X_2	P_2X_3	P_3X_1	P_3X_2	P_3X_3
Q_1	y_1	y_2	y_3	y_3	y_3	y_2	y_2	y_1	y_2
Q_2	y_3	y_2	y_2	y_1	y_3	y_3	(y_3)	(y_3)	(y_3)
Q_3	y_1	y_2	y_1	y_2	y_1	y_1	(y_1)	(y_1)	(y_1)

λ_2

Доопределим функции δ_2 и λ_2, как это показано в скобках таблиц δ_2 и λ_2. На промежуточном алфавите X_2 найдено следующее наибольшее разбиение $\varepsilon = \{z_1, z_2, z_3, z_4, z_5, z_6\}$, где

$$z_1=\{(P_1X_1),(P_3X_2)\},\ z_2=\{(P_1X_1)\},\ z_3=\{(P_1X_3)\},\ z_4=\{(P_2X_1)\},$$
$$z_5=\{(P_2X_2),(P_2X_1),(P_3X_3)\},\ z_6=\{(P_2X_3)\}.$$

После минимизации промежуточного алфавита функции δ_1, λ_1 δ_2, λ_2 примут вид:

	X_1	X_2	X_3
P_1	P_1	P_2	P_3
P_2	P_2	P_1	P_1
P_3	P_2	P_3	P_1

δ_1

	X_1	X_2	X_3
P_1	z_1	z_2	z_3
P_2	z_4	z_5	z_1
P_3	z_5	z_6	z_5

λ_1

	z_1	z_2	z_3	z_4	z_5	z_6
Q_1	Q_1	Q_2	Q_1	Q_2	Q_3	Q_2
Q_2	Q_3	Q_1	Q_1	Q_2	Q_2	Q_1
Q_3	Q_3	Q_3	Q_1	Q_1	Q_1	Q_2

δ_2

	z_1	z_2	z_3	z_4	z_5	z_6
Q_1	y_1	y_2	y_3	y_3	y_3	y_2
Q_2	y_3	y_2	y_2	y_1	y_3	y_3
Q_3	y_1	y_2	y_1	y_2	y_1	y_1

λ_2

Литература

I. J. Hartmanis, R.E. Stearns. Algebraic Structure Theory of Sequential Machines. Prentice-Hall, Inc., N.Y. 1966.

2. A. Gill. Cascaded Finite-State Machines, IRE Trans., EC-10, No.3, 1961.

3. Г.Э. Якобсон. Последовательная декомпозиция конечных автоматов. Автоматика и вычислительная техника, 1970 (в печати).

W.Kracht, G.Jakobson

ON THE PROBLEM OF THE SERIAL DECOMPOSITION OF FINITE AUTOMATA

Summary

A definition of the serial decomposition based on the concept of the isomorfic subautomaton is given. The definition is linked with Hartmanis' theorem of the serial decomposition. The method of the minimization of the intermediate alphabet is formalized.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

СЕРИЯ А № 288 1970

УДК 621.372.061

В.А. Кукк

СЛУЧАЙНОЕ ГЕНЕРИРОВАНИЕ ДЕРЕВЬЕВ

Методам генерирования всех деревьев графа посвящено немало работ [1]. Это связано с тем, что вычисление разных определителей может быть приведено к вычислению суммы величин, связанных с деревьями некоторого графа [2]. В частности, согласно формуле Максвелла, детерминант электрической цепи равен сумме величин всех деревьев цепи [3]. Ряд важных свойств цепи в свою очередь определяется детерминантом (т.н. внутренние свойства).

Следует, однако, отметить, что получение простого и надежного алгоритма генерирования всех деревьев графа не решает задачу нахождения внутренних характеристик цепи. Дело в том, что число деревьев графа, вообще говоря, может быть огромным. Так, например, полный граф с n вершинами содержит n^{n-2} деревьев, что при $n = 10$ составляет уже 10^8.

Это указывает на возможность (и, видимо, целесообразность) применения для оценки внутренних характеристик цепи методов Монте-Карло. Последние в свою очередь требуют случайного генератора деревьев. В настоящей статье приводятся два алгоритма и рассматриваются возможности их реализации на ЭВМ. Оба генератора действуют на ненаправленном (реберном) графе (мультиграфе). В первом из них деревья строят независимо друг от друга, во втором реализуется случайное блуждание на графе деревьев.

Будем обозначать [2, 4] через n число вершин графа G, через m число его ребер. Предположим, что G связен, тогда цикломатическое число $\nu = m-n+1$ и коцикломатическое

число [5] $\varsigma = m-1$. Деревьями t называем частичные деревья графа G. Лесом τ называем частичный лес графа G.

Отношение связности на множестве вершин X графа G, определяемое лесом τ, обозначим через $Con(\tau)$: $(x,y) \in Con(\tau)$, если в τ существует цепь, соединяющая x и y.

Кроме того, определяем на произведении $U \times T$ множеств ребер U и лесов T отношение включаемости $\|$. Ребро $u \in U$ включаемо в лес τ $(u \| \tau)$, если присоединение ребра u к лесу τ не приводит к образованию циклов (т.е. снова получается лес). Если x и y – концевые вершины ребра u, то $u \| \tau$ эквивалентно $(x,y) \notin Con(\tau)$.

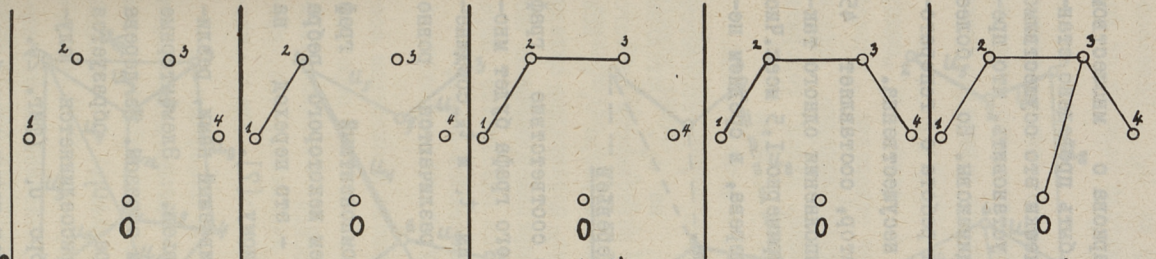
Случайный генератор деревьев

Алгоритм генерирования состоит в следующем: строится последовательность лесов (0–пустой лес)

$$0 \subset \tau_1 \subset \tau_2 \subset \cdots \subset \tau_\varsigma = t,$$

где τ_{i+1} получается из τ_i присоединением к последнему случайно выбранного ребра $u \| \tau_i$. Случайный генератор ребер действует на множестве U и имеет равномерное распределение. Выбранное ребро испытуется на $u \| \tau_i$; при $u \not\| \tau_i$ оно бракуется и снова обращаются к генератору ребер.

Описанный генератор легко осуществляется на ЭВМ. Пусть в массиве A (m ячеек) описаны ребра графа с помощью пар концевых вершин. В массиве B (n ячеек) отображаем отношение связности $Con(\tau)$ следующим образом. Каждой вершине отведена ячейка, которая заполняется командой (безусловного) перехода таким образом, что ячейки вершин каждого класса эквивалентности X_i $(X_i \in X/Con(\tau))$ образуют замкнутый цикл. Если требуется проверить $u \| \tau$, то управление дается по содержимому ячейки x, записав в ячейку x выход „$u \| \tau$" и в ячейку y выход „$u \not\| \tau$". При положительном исходе $(u \| \tau)$ циклы, соответствующие x и y, объединяются в один цикл обменом содержимых ячеек x и y. На фиг. I приведен пример генерирования дерева и соответственно содержания массива B. (Под 4–ым столбцом представлен массив B при испытании ребра (2,I) на включаемость).



№		u = (2,1)	u = (2,3)	u = (3,4)	u = (3,0)
0	→ 0	→ 0	→ 0	→ 0	→ 4
1	→ 1	→ 2	→ 2	→ 2	→ 2
2	→ 2	→ 1	→ 3	→ 3	→ 3
3	→ 3	→ 3	→ 1	→ 4	→ 0
4	→ 4	→ 4	→ 4	→ 1	→ 1

	u = (2,1)
	→ 0
y	→ u//τ
x	→ u\\τ
x : (→3) →	→ 4
	→ 1

Фиг. 1

Наконец, в массиве C собираются коды ребер дерева для дальнейшего использования.

Описанный процесс является цепью Маркова с множеством состояний T . Исследование ее может быть проведено стандартными методами [6, 8], хотя практически это осуществимо лишь для простейших графов. Нетрудно установить, что предельные вероятности для деревьев не одинаковы. Но более подробный анализ сложен. Практическая работа с генератором показала, что эта неравномерность несущественна.

Длина программы, реализующей генератор, составляет 45 команд в коде ЭВМ "Минск-22". Время выполнения одного такта (выбор ребра u и проверки $u \| \tau$) примерно I,5 мсек.Число тактов, необходимое для получения дерева, в среднем несколько раз превышает число ребер графа.

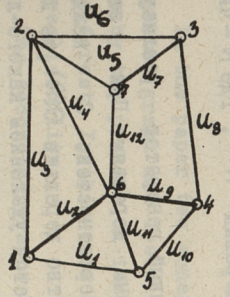
Блуждание по графу деревьев

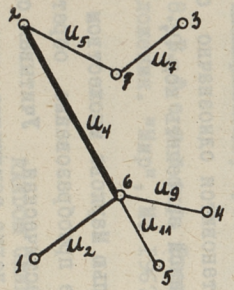
Связному графу G можно ставить в соответствие граф деревьев Γ . Множеством вершин Y этого графа будет множество деревьев графа G , а две вершины t_i и t_k соединены ребром, если сами деревья t_i и t_k различаются ровно одним ребром графа [7].

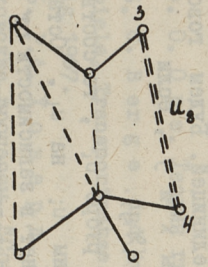
Двойное дерево $\mathfrak{D}=(t,u)$ - это двухкомпонентный граф (лес), полученный из дерева t удалением некоторого ребра u . Элементарное преобразование дерева - это переход на графе деревьев от одной вершины к смежному [7].

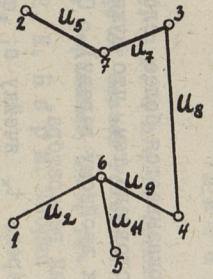
Второй генератор деревьев, рассматриваемый нами, реализует случайное блуждание по графу деревьев. Элементарные преобразования осуществляются следующим образом. Из дерева t_k удалением случайно выбранного ребра u_j образуется двойное дерево $\mathfrak{D}_{kj}$. Затем к $\mathfrak{D}_{kj}$ присоединяется случайно выбранное (включаемое в $\mathfrak{D}_{kj}$) ребро u_i (фиг. 2).

Выбор ребра u_i может быть произведен двумя способами. Во-первых, можно использовать генератор ребер графа G .Во-вторых, можно образовать множество U_{kj} ребер, включаемых









1	2	3	4	5	6	7	
1	0	0	0	0	0	0	u_2
1	0	0	1	1	1	0	u_4
0	0	1	0	0	0	1	u_5
0	0	1	0	0	0	0	u_7
0	0	0	1	0	0	0	u_9
0	0	0	0	1	0	0	u_{11}

	34		
	00		u_2
10	01	110	u_3
	10		u_5
	10		u_7
	01		u_9
	00		u_{11}

1	2	3	4	5	6	7	
1	0	0	0	0	0	0	u_2
1	0	0	1	1	1	0	u_3
1	0	1	1	1	1	1	u_5
1	0	1	1	1	1	0	u_7
1	0	0	0	1	1	0	u_9
0	0	0	0	1	0	0	u_{11}

Фиг. 2

в $\mathfrak{R}_{kj}$)регулярным способом, например, перебором ребер графа) и затем произвести случайный выбор из U_{kj} . Целесообразность применения того или другого метода зависит от конкретной реализации генератора.

Кодирование этого генератора оказывается более сложным, чем в предыдущем случае. Это обусловлено тем, что нужно завершить обратный шаг (от дерева к двойному дереву). Опишем кратко одну возможность реализации генератора.

Отведем каждому ребру u_j дерева t_k ячейку a_j ,содержащую n битов. В ячейку a_j закодируем отношение эквивалентности $Con(\mathfrak{R}_{kj})$, т.е. вершины, принадлежащие одному классу, отметим нулем, остальные единицей. Будем говорить, что ребро u_j разделяет вершины x и y , если

$$(x,y) \notin Con(\mathfrak{R}_{kj}).$$

Если из дерева t_k было удалено рубро u_i , то соответствующий код не меняется после замены u_i на u_l . Остальные коды подлежат преобразованию или нет в зависимости от того, разделяет ли соответствующее ребро вершины ребра u_l . Если разделение имеет место, то коду ребра u_j поразрядно складывается код ребра $u_i(u_l)$. На фиг. 2 приведен вариант кодирования. (Код каждого дерева становится однозначно определенным, если за некоторой вершиной закрепить код 0; в примере выбрана вершина 2).

Описанный процесс является цепью Маркова множеством состояний $\mathcal{Y}$. Так как элементарное преобразование обратимо, то матрица переходов дважды стохастическая. Учитывая связность графа деревьев Γ , можем легко заключить, что предельные вероятности всех состояний равны N^{-1} , где N – число деревьев [8].

Наконец отметим, что рассмотренный генератор реализует также случайное блуждание по двойным деревьям. При этом сохраняется равенство предельных вероятностей. С каждым двойным деревом $\mathfrak{R}$ связано множество деревьев $\Theta(\mathfrak{R})$,из которых данное двойное дерево получается удалением какого-либо ребра: $\Theta(\mathfrak{R}) = \{t : \exists u\,[(t,u) = \mathfrak{R}]$. Учитывая на каждом шагу все деревья множества $\Theta(\mathfrak{R})$, можем увеличить скорость сходимости при оценке характеристик цепей.

Литература

I. G.P. Char. Generation of Trees, Two-Trees, and Storage of Master Forests. IEEE Transactions, vol. CT-15, No. 3, September 1969, pp. 228-238.

2. К. Берж. Теория графов и ее приложения. ИЛ., М. 1962.

3. С. Сешу, Н. Балабанян. Анализ линейных цепей. Госэнергоиздат, М. 1963.

4. О. Оре. Теория графов. "Наука", М. 1968.

5. C. Berge und A. Ghouila-Houri. Programme, Spiele, Transportnetze. G.B. Teubner Verlagsgesellschaft, Leipzig 1967.

6. Р.А. Ховард. Динамическое программирование и марковские процессы. "Советское Радио", М. 1964.

7. R.L. Cummins. Hamilton Circuits in Tree Graphs. IEEE Transactions, vol. CT-13, No. 1, March 1966, pp.82-90.

8. В. Феллер. Введение в теорию вероятностей и ее приложения. "Мир", М. 1967.

V. Kukk

RANDOM GENERATION OF TREES

Summary

In the paper two generators of trees of an undirected, connected graph are described. The first one builds up a tree by the process of inserting stochastically chosen edges in a forest. In this case trees have different probabilities of realizability.

The second generator realizes random walk on the tree graph of the given graph.

A method of coding is given for both generators.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

С Е Р И Я А № 288 1970

УДК 621.372.061

В.А. Кукк

РАЦИОНАЛЬНАЯ АППРОКСИМАЦИЯ ПЕРЕДАТОЧНОЙ ФУНКЦИИ

В настоящей статье приводится алгоритм построения дробно-рациональной передаточной функции $R(s)$, исходя из разложения соответствующей импульсной функции по функциям Лагерра. Обсуждаются основы и вычислительные аспекты алгоритма.

Исходное разложение запишем следующим образом

I) для импульсной функции

$$h(t) = \sum_{k=0}^{\infty} a_k l_k(t), \qquad (1)$$

2) для передаточной функции

$$F(s) = \sum_{k=0}^{\infty} a_k \lambda_k(s). \qquad (2)$$

Здесь $F(s)$ - лапласово изображение от $h(t)$, а $l_k(t)$- функции Лагерра [1]

$$l_k(t) = Y(t)\frac{e^{\frac{t}{2}}}{k!}\frac{d^k}{dt^k}(e^{-t}t^k),$$

$Y(t)$- функция Хэвисайда и

$$\lambda_k(s) = \frac{(s-\frac{1}{2})^k}{(s+\frac{1}{2})^{k+1}}.$$

Вместе с рядами (I) и (2) будем рассматривать ряд

$$H(z) = \sum_{k=0}^{\infty} a_k z^k \qquad (3)$$

и последовательность

$$a_0,\ a_1,\ a_2,\ \dots,$$

которую обозначим через A. Разложения (2) и (3) связаны соотношением

$$F(s) = (1-z)H(z)\Big|_{z=\frac{s-\frac{1}{2}}{s+\frac{1}{2}}}. \qquad (4)$$

Предполагается, что известны первые $N+1$ членов последовательности A. Построим дробь

$$G_n(z) = \frac{N(z)}{D(z)} = \frac{c_0 + c_1 z + \dots + c_m z^m}{1 + d_1 z + \dots + d_n z^n}$$

такую, что в её тейлоровском разложении

$$G_n(z) = g_0 + g_1 z + g_2 z^2 + \dots$$

коэффициенты $g_0, g_1, \dots, g_N$ как можно меньше отличались бы соответственно от $a_0,\ a_1,\ \dots,\ a_N$. Без ограничения общности можно принимать $n = m+1$. (Ниже будет видно, что должно быть $1 \le 2n-1 \le N$). Применяя к $G_n(z)$ преобразование (4), получим рациональную дробь $R(s)$, степень которой (большая из степеней числителя и знаменателя) не превышает n.

При построении $G_n(z)$ мы сначала найдем $D(z)$, стараясь получить $g_n = a_n,\ g_{n+1} = a_{n+1}, \dots, g_N = a_N$. Затем определяем $N(z)$ из условия $g_0 = a_0,\ g_1 = a_1, \dots, g_{n-1} = a_{n-1}$. Для этого (и с учетом (4)) вычислим произведение многочленов

$$1-z,\ 1 + d_1 z + \dots + d_n z^n = D(z),\ a_0 + a_1 z + \dots + a_N z^N,$$

сохраняя в результате лишь члены до z^n.

Определение $D(z)$ является основной частью алгоритма и имеет самостоятельное значение. Заметив, что [2]

$$g_r = -\sum_{i=1}^{n} d_i\, g_{r-i}$$

при $r \ge n$; отсюда при условии $g_i = a_i$ $(i = n, n+1, \dots, N)$, получим систему

$$\begin{pmatrix} a_0 & a_1 & \cdots & a_{n-1} \\ a_1 & a_2 & \cdots & a_n \\ \cdots & \cdots & \cdots & \cdots \\ a_{N-n} & a_{N-n+1} & \cdots & a_{N-1} \end{pmatrix} \begin{pmatrix} d_n \\ d_{n-1} \\ \cdots \\ d_1 \end{pmatrix} = \begin{pmatrix} -a_n \\ -a_{n+1} \\ \cdots \\ -a_N \end{pmatrix}. \qquad (5)$$

Система эта, вообще говоря, несовместна и поэтому ее решение следует искать методом наименьших квадратов.

Очевидно, если последовательность A соответствует рациональной передаточной функции, степень которой не выше n, то наш алгоритм (при $N \geq 2n-1$) восстанавливает эту функцию. Итак, метод имеет в этом смысле наивысшую степень точности.

С другой стороны, $g_i = a_i$ при $i = 0,1,\cdots,n-1$, а $\sum_{i=n}^{N} (g_i - a_i)^2$ имеет наименьшее значение. Нетрудно показать, что функционал

$$I(X_n) = \sum_{i=0}^{N} (x_i - a_i)^2$$

при условиях нашей задачи достигает минимума на $X_{N_0} = g_0, g_1, \ldots, g_N$. Кроме того, если $h(t) \in L_2(0,\infty)$, то

$$\lim_{N\to\infty} I(X_{N_0}) = \int_0^{\infty} [h(t) - g(t)]^2 dt,$$

т.е. при достаточно большом N мы решаем задачу минимизации среднеквадратичной погрешности. Поэтому можно сказать, что рассмотренный алгоритм дает также наилучшую среднеквадратичную аппроксимацию.

Перейдем к вычислительной процедуре. Нормальную систему для (5) запишем в виде матрицы

$$\begin{pmatrix} \alpha_{00} & \alpha_{01} & \cdots & \alpha_{0,n-1} & \alpha_{0,n} \\ \alpha_{10} & \alpha_{11} & \cdots & \alpha_{1,n-1} & \alpha_{1,n} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ \alpha_{n-1,0} & \alpha_{n-1,1} & \cdots & \alpha_{n-1,n-1} & \alpha_{n-1,n} \end{pmatrix}, \qquad (6)$$

где $\alpha_{ij} = \sum_{l=0}^{N-n} a_{i+l} a_{j+l}$. Эта матрица содержит также матрицы нормальных систем более низкого порядка, что позволяет попутно получить приближения низшего порядка (однако при этом не используются все коэффициенты a_i). Кроме того,

ранг матрицы (6) в общем случае неизвестен. Это приводит к необходимости использования метода исключения при решении системы. Мы подробнее опишем алгоритм, слегка отличающийся от стандартной схемы Гаусса.

Обозначим элемент в i-ой строке и j-ом столбце матрицы после k-го шага через $b_{ij}^{(k)}$. На k-ом шагу в преобразованиях участвуют лишь столбцы с номерами $\geq k$; притом элементы k-го столбца (кроме $b_{kk}^{(k)}$) не меняются.

I. На k-ом шагу делим элементы k-ой строки на $-b_{kk}^{(k-1)}$; если это невозможно, процесс заканчивается; $b_{kk}^{(k-1)}$ заменяем нулем.

2. Умножим k-ую строку на элементы k-го столбца и прибавляем к соответствующим строкам.

3. Переходим к следующему шагу.

После окончания процесса в каждом столбце имеем решение. Именно, над элементом $b_{kk}^{(p)}$ стоит $d_1^{(k)}$, над последним $d_2^{(k)}$ и т.д. Решения в столбцах $k \geq p$ эквивалентны. В остальных столбцах сохранились решения более низкого порядка.

Если процесс заканчивался после $p < n$ шагов, то рекомендуется решение повторять, взяв $n = p$, чтобы полностью использовать исходную информацию.

Алгоритм малочувствителен к погрешностям вычислений. Пусть ранг матрицы $\varrho < n$. Вполне возможно, что из-за неточностей вычислений на главной диагонали после ϱ-го шага не получается нуля и процесс продолжается. Но благодаря тому, что в последующих столбцах также имеются решения, результат будет правильным.

Пример. Рассмотрим последовательность

$$A = 1, 3, 5, 7, \ldots,$$

полученную из разложения функции

$$h(t) = 2\frac{d}{dt}\delta(t) = 2\delta'(t) = \sum_0^\infty (2k+1)\, l_k(t).$$

Пусть $N = 20$ и $n = 4$ (сильная переопределенность, для восстановления $h(t)$ достаточно взять $N = 3$ и $n = 2$). Вычисления приведены в таблице I.

Таблица I

6545	7I23	770I	8279	8857
7I23	7769	84I5	906I	9707
770I	84I5	9I29	9843	I0557
8279	906I	9843	I0625	II407
0	-I.0883II	-I.I76623	-I.264935	-I.353246
7I23	I6.95584	33.9II68	50.86755	67.82336
770I	22.9II68	67.82342	I0I.735I	I35.6467
8279	50.86755	I0I.735I	I52.6026	203.4700
0	-I.0883II	I.000000	2.00000I	3,000000
7I23	0	-2.000000	-3.000000	-4.000000
770I	33.9II68	2^{-I4}	0.000000	0.000000
8279	50.86755	0.000000	-5.2^{-I5}	-2^{-I3}
I	-I.0883II	I.000000	2.00000I	3.000000
7I23	0	-2.000000	-3.000000	-4.000000
770I	33.9II68	0	0.000000	0.000000
8279	50.86755	0.000000	-5.2^{-I5}	-2^{-I3}
0	-I.0883II	I.000000	2.00000I	I.400000
7I23	0	-2.000000	-3.000000	-I.599998
770I	33.9II68	0-	0.000000	0.000000
8279	50.86755	0.000000	0	-0.800000

К о м м е н т а р и и: После второго шага в третьем столбце получается 2^{-14} (при точном расчете 0). В этом столбце содержится простейшее решение $(1-2z+z^2)$, в следующих столбцах: $1-3z^2+2z^3=(1-2z+z^2)(1+2z)$ и $1-4z^3+3z^4=$ $=(1-2z+z^3)(1+2z+3z^2)$.

Во втором столбце первое приближение $1-1.0883112$.

При продолжении вычислений получим решения (округленно) $1-3z^2+2z^3$ и $1-0.8z-1.6z^3+1.4z^4$. Из всех этих решений, начиная со второй степени, получается одинаковый результат. Так для $1-3z^2+2z^3$ имеем

$$(1-z)(1-3z^2+2z^3)(1+3z+5z^2+7z^3+\cdots)=1+2z-z^2-2z^3+\cdots$$

и

$$(1-z)G_3(z)=\frac{1+2z-z^2-2z^3}{1-3z^2+2z^3}. \qquad (7)$$

После сокращения на $(1-z)(1+2z)$ и подстановки $z=\frac{s-\frac{1}{2}}{s+\frac{1}{2}}$, получим $R(s)=2s$

или $h(t)=2\delta'(t)$. Впрочем, непосредственная подстановка в (7) даёт

$$R(s)=\frac{6s^2-s}{3s-\frac{1}{2}}=2s.$$

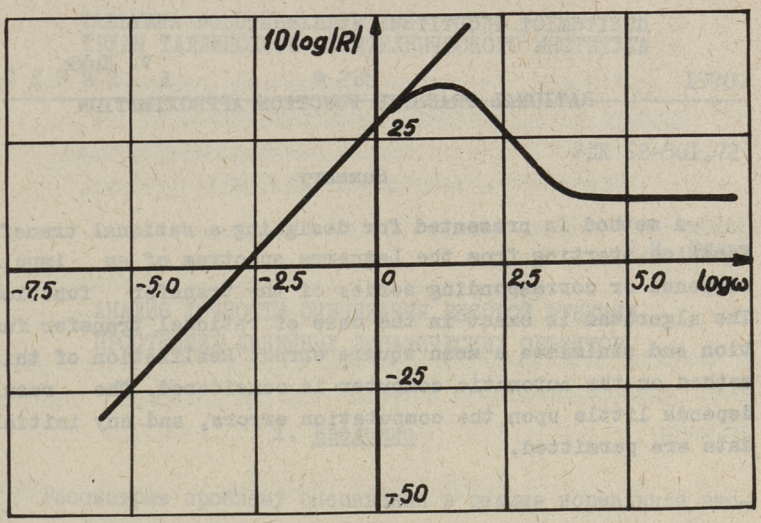
Окончательный результат, полученный на ЭВМ:

$$R(s)=2s\times\frac{0.2999999-0.3999990s+3.599998s^2+0.1370907\cdot10^{-5}s^3}{0.2999998-0.3999989s+3.599998s^2+0.1698732\cdot10^{-5}s^3-0.1192093\cdot10^{-6}s^4}. \qquad (8)$$

На фиг. I представлены графики функции 2ω и $|R(j\omega)|$ в зависимости от ω.

Замечания.

I) Второй множитель в (8), появившийся в результате неточностей в вычислениях, неустойчив (полюсы в правой полуплоскости). Этот множитель, разумеется, следует просто отбросить. Появление этого множителя еще раз подтверждает тот факт, что любое представление конечной точности имеет локальный характер и не несет ответственности за "чужие области".



Фиг. 1.

2) Аналогичный алгоритм был получен Стейглитцем [3] при более жестких требованиях ($h(t) \in L_2$, хотя приводится пример $h(t) = \delta(t-\tau) \notin L_2$).

3) Мы не сделали никаких предположений о свойствах последовательности A. Алгоритм, конечно, работает при любых начальных данных (так как N конечно).

Литература

1. Е. Янке, Ф. Эмде. Ф. Лёш. Специальные функции. "Наука", М. 1968.

2. Ф.Р. Гантмахер. Теория матриц. "Наука", М. 1966.

3. K. Steiglitz. Rational Transform Approximation via the Laguerre Spectrum. J. Franklin Inst., vol. 280 No.5, pp. 387-394; November 1965.

V. Kukk

RATIONAL TRANSFER FUNCTION APPROXIMATION

Summary

A method is presented for designing a rational transfer function starting from the Laguerre spectrum of an impulse response or corresponding series of the transfer function. The algorithm is exact in the case of rational transfer function and minimizes a mean square error. Realization of this method on the automatic computer is considered. The result depends little upon the computation errors, and any initial data are permitted.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА
СЕРИЯ А № 288 1970

УДК 62-501.72

Я.Н. Олль

АНАЛИЗ ТОЧНОСТИ ОЦЕНИВАНИЯ ВЕСОВОЙ ФУНКЦИИ ПРОСТЕЙШИХ ЛИНЕЙНЫХ ДИНАМИЧЕСКИХ ОБЪЕКТОВ

I. Введение

Рассмотрим проблему оценивания в режиме нормальной эксплуатации весовой функции (ВФ) линейного динамического объекта, описываемого уравнением

$$y(t)=\int_0^\infty h(\tau)\,x(t-\tau)\,d\tau\,, \qquad (1)$$

где $x(t)$ - входной сигнал объекта,
$y(t)$ - выходной сигнал объекта,
$h(\tau)$ - ВФ.

При этом через интервал времени Δ наблюдается сигнал

$$v(i\Delta)=y(i\Delta)+\varphi(i\Delta),\quad i=1,2,\dots,N\,, \qquad (2)$$

где $\varphi(t)$ - помеха на выходе объекта.

Впервые в такой постановке эта задача была рассмотрена в работе [1], а [2 - 6] посвящены исследованию ошибок определения ВФ, возникающих из-за шума на выходе $\varphi(t)$, квантования по времени ВФ объекта и непрерывных сигналов $x(t)$ и $y(t)$ и отбрасывания $h(\tau)$ при больших значениях τ.

В [6] впервые были получены выражения, позволяющие вычислить точность оценки ВФ, полученной методом наименьших квадратов при наличии следующих ограничений относительно вероятностных характеристик случайных процессов $x(t)$ и $\varphi(t)$:

$$M\{x(t)\}=0,\quad M\{\varphi(t)\}=0,\quad R_x(\tau)=D_x e^{-b|\tau|},\qquad (3)$$
$$R_\varphi(\tau)=D_\varphi e^{-a|\tau|},\quad R_{x\varphi}(\tau)=0,$$

где $R_x(\tau), R_\varphi(\tau)$ – автокорреляционные функции $x(t)$ и $\varphi(t)$ соответственно,

$R_{x\varphi}(\tau)$ – их взаимокорреляционная функция,

$M\{\cdot\}$ – знак математического ожидания.

Дискретная модель объекта (I) в [6] выражена в виде

$$y(j\Delta)=\Delta\sum_{i=0}^{J} h\left(i\Delta+\frac{\Delta}{2}\right)x\left[(j-i)\Delta-\frac{\Delta}{2}\right],\qquad (4)$$

а качество восстановленной модели объекта (4) характеризовано среднеквадратичной погрешностью предсказания выхода объекта:

$$D_\eta=M\left\{\left[y(t)-\tilde{y}(t)\right]^2\right\},\qquad (5)$$

где

$$\tilde{y}(t)=\Delta\sum_{i=0}^{J}\tilde{h}\left(i\Delta+\frac{\Delta}{2}\right)x\left[(j-i)\Delta-\frac{\Delta}{2}\right]\qquad (6)$$

и $\tilde{h}(\cdot)$ оценка ординаты ВФ.

В [6] было получено общее выражение для D_η при ограничениях (3), однако вопросы выбора основных параметров модели (4) – периода квантования Δ, числа ординат ВФ J и времени наблюдения $T_н=N\Delta$, необходимого для получения заданной точности прогнозирования, остались до сих пор не решенными. Решению этих вопросов препятствует то обстоятельство, что выражение для D_η весьма громоздкое даже для самых простых видов ВФ. В настоящей работе выводится это выражение для объекта первого порядка и даются графики для выбора J, $T_н$ и оптимального времени памяти модели (4) $T_o=(J+1)\Delta$. Для выяснения особенностей оценивания объектов более высоких порядков было проведено моделирование оценивания объекта второго порядка, так как уже в этом случае аналитическое выражение для D_η становится слишком сложным. Применимость асимптотических формул

[6] при практических значениях времени наблюдения $T_н$, превышающих несколько десятков раз время памяти модели T_o, проверяется также моделированием.

2. Точность оценивания объекта первого порядка

Пусть задан объект первого порядка, имеющий ВФ:

$$h(\tau) = \frac{K}{T} e^{-\frac{\tau}{T}}, \qquad (7)$$

дискретная модель которого выражена в виде (4). Находим выражения для математического ожидания оценок ВФ:

$$\underset{0 \le i < J}{A(i)} = M\left\{\tilde{\Theta}(i)\right\} = \Delta \cdot h\left(i\Delta + \frac{\Delta}{2}\right) + \delta\Theta_\varphi(i), \qquad (8)$$

$$A(J) = M\left\{\tilde{\Theta}(J)\right\} = \Delta \cdot h\left(J\Delta + \frac{\Delta}{2}\right) + \delta\tilde{\Theta}_\varphi(J) + \delta\Theta_\varepsilon(J),$$

где $\delta\Theta_\varphi(i)$ – ошибка от квантования, $\tilde{\Theta}(i) = \Delta \cdot \tilde{h}\left(i\Delta + \frac{\Delta}{2}\right)$

и $\delta\Theta_\varepsilon(J)$ – ошибка от отбрасывания $h(\tau)$ при $\tau > T_o$.

После преобразований на основе формул [6] имеем:

$$A(0) = \frac{K}{1-Tb}\left[\beta^{1/2} - \frac{2Tb\,\alpha^{1/2}(1-\alpha\beta)}{(1+Tb)(1-\beta^2)}\right]$$

$$\underset{1 \le i \le J-1}{A(i)} = \frac{2KTb(\beta-\alpha)(1-\alpha\beta)}{(1-\beta^2)\left[1-(Tb)^2\right]\alpha^{1/2}} \cdot \alpha^i = B\alpha^i \qquad (9)$$

$$A(J) = \frac{2KTb\,\alpha^J(\beta-\alpha)}{(1-\beta^2)\left[1-(Tb)^2\right]\alpha^{1/2}},$$

где

$$\alpha = \exp\left(-\frac{\Delta}{T}\right)$$

и $\beta = \exp(-b\Delta)$.

Так как предсказанный выходной сигнал $\tilde{y}(t)$ выражается в виде (6) и

$$\Delta\tilde{h}\left(i\Delta + \frac{\Delta}{2}\right) = \Delta \cdot h\left(i\Delta + \frac{\Delta}{2}\right) + \delta\Theta_\varphi(i) +$$

$$+ \delta\Theta_\varepsilon(i) + \delta\Theta_\varphi(i) = A(i) + \delta\Theta_\varphi(i),$$

где $\delta\Theta_\varphi(i)$ – ошибка от шума $\varphi(t)$ на выходе, то ошибка предсказания выхода равна:

$$\eta(t) = y(t) - \tilde{y}(t) = y(t) - \sum_{i=0}^{J} A(i)\, x\left[(j-i)\Delta - \right. \tag{10}$$
$$\left. - \frac{\Delta}{2}\right] - \sum_{i=0}^{J} \delta\Theta_{\varphi}(i)\, x\left[(j-i)\Delta - \frac{\Delta}{2}\right] = \eta_1 - \eta_2 - \eta_3 .$$

Из-за центрированности и статистической независимости $\delta\Theta_{\varphi}(i)$ от $x(t)$ при предсказании (это соблюдается, если между реализациями $x(t)$ при оценивании и предсказании находится достаточный временный сдвиг) имеем:

$$M\left\{\eta_1 \cdot \eta_3\right\} = 0, \quad M\left\{\eta_2 \cdot \eta_3\right\} = 0.$$

Поэтому среднеквадратичная ошибка предсказания выхода объекта D_{η} равна:

$$D_{\eta} = M\left\{\left[\eta_1 - \eta_2 - \eta_3\right]^2\right\} = D_{\eta_1} + D_{\eta_2} + D_{\eta_3} - 2K_{\eta_1\eta_2}, \tag{11}$$

где $$D_{\eta\kappa} = M\left\{\eta^2_{\kappa}\right\}, \quad \kappa = 1, 2, 3$$

и $$K_{\eta_1\eta_2} = M\left\{\eta_1 \cdot \eta_2\right\}.$$

После преобразований получим следующие выражения для компонентов D_{η}:

$$D_{\eta_1} = \frac{D_x \cdot K^2}{1 + Tb}, \tag{12}$$

$$D_{\eta_2} = D_x\left\{A^2(0) + B^2\frac{\alpha^2 - \alpha^{2J}}{1-\alpha^2} + A^2(J) + 2\left\{A(0)\cdot\right.\right.$$

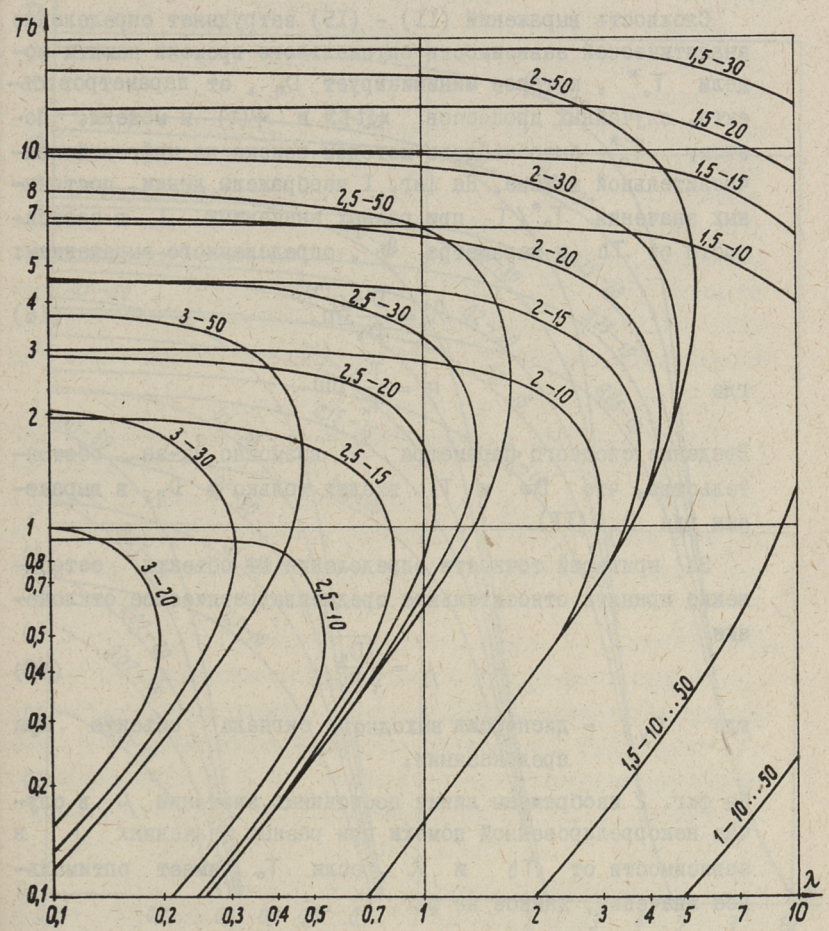
$$\left[B\frac{\alpha\beta - (\alpha\beta)^J}{1-\alpha\beta} + A(J)\cdot\beta^J\right] + B\left[\frac{B}{1-\alpha\beta}\left(\alpha\beta\frac{\alpha^2 - \alpha^{2(J-1)}}{1-\alpha^2} - \right.\right. \tag{13}$$

$$\left.\left.\left. - \frac{\alpha^{J+1}\beta^{J-1} - \alpha^{2J-1}\beta}{1-\alpha\beta^{-1}}\right) + A(J)\frac{\alpha\beta^{J-1} - \alpha^J}{1-\alpha\beta^{-1}}\right]\right\}\right\},$$

$$K_{\eta_1\eta_2} = \frac{D_x \cdot K}{1-Tb}\left\{\left[A(0) + B\frac{\alpha\beta - (\alpha\beta)^J}{1-\alpha\beta} + A(J)\cdot\beta^J\right]\beta^{1/2} - \right.$$

$$\left. - \frac{2b\alpha^{1/2}}{1/T + b}\left[A(0) + B\frac{\alpha^2 - \alpha^{2J}}{1-\alpha^2} + A(J)\cdot\alpha^J\right]\right\}. \tag{14}$$

Компонент D_{η_3}, независящий от типа объекта, согласно [6] выражается в виде:



Фиг. 1. Линии постоянных значений оптимальной памяти дискретной модели на плоскости Tb и λ при разных значениях J. Первое число на кривой означает T_0^*/T, второе J.

$$D_{\eta_3} = \frac{D_\varphi}{N}\left[J + \frac{1+\delta\beta}{1-\delta\beta}\right], \qquad (I5)$$

где $$\delta = \exp(-a\Delta).$$

Сложность выражений (II) - (I5) затрудняет определение аналитической зависимости оптимального времени памяти модели T_o^*, которое минимизирует D_η, от параметров объекта, случайных процессов $x(t)$ и $\varphi(t)$ и модели. Поэтому T_o^* было найдено методом поиска на цифровой вычислительной машине. На фиг. I изображены линии постоянных значений T_o^*/T при разных значениях J в зависимости от Tb и параметра λ, определенного выражением:

$$\lambda = \frac{D_\varphi}{D_y} \cdot m, \qquad (I6)$$

где $$m = \frac{T}{T_н} \cdot 200.$$

Введение сводного параметра λ возможно из-за обстоятельства, что D_φ и $T_н$ входят только в D_{η_3} в выражении для D_η (II).

За критерий точности определения ВФ объекта естественно принять относительное среднеквадратическое отклонение

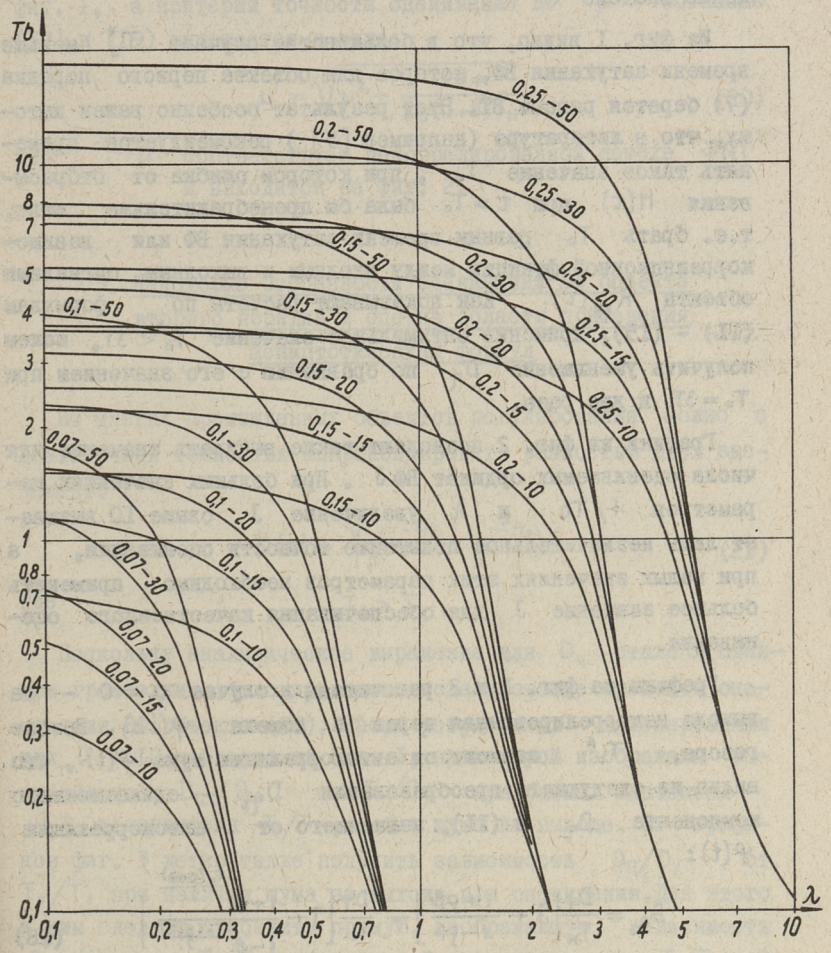
$$\mu = \sqrt{\frac{D_\eta}{D_y}}, \qquad (I7)$$

где D_y - дисперсия выходного сигнала объекта при предсказании.

На фиг. 2 изображены линии постоянных значений μ в случае некоррелированной помехи при разных значениях J в зависимости от Tb и λ, если T_o имеет оптимальное значение, данное на фиг. I.

Для пользования графиками на фиг. I и 2 надо предварительно оценить T, b и D_φ/D_y вычислить λ по (I6) и интерполировать между соседними линиями при заданном J.

Пример: Пусть $D_\varphi/D_y = 0{,}5$; $T = 100$ сек; $T_н/T = 50$; $J = 30$, $b = 0{,}04$ сек$^{-1}$.



Фиг. 2. Линии постоянных значений показателя точности модели μ при оптимальной памяти модели на плоскости Tb и λ при разных значениях J. Первое число на кривой означает μ, а второе J. Помеха φ не коррелирована

вычислим: $\lambda = 2$. Из фиг. I находим, что $T_o^*/T = 2{,}25$, а из фиг. 2 $\mu = 0{,}2$. Задавая показателем точности модели μ, можем по фиг. 2 найти требуемое время наблюдения, определив вначале λ.

Из фиг. I видно, что в большинстве случаев T_o^* меньше времени затухания ВФ, которое для объекта первого порядка (7) берется равным $3T$. Этот результат особенно важен потому, что в литературе (например [5]) рекомендуется применять такое значение T_o, при котором ошибка от отбрасывания $h(\tau)$ при $\tau > T_o$ была бы пренебрежительно мала, т.е. брать T_o равным времени затухания ВФ или взаимокорреляционной функции между входным и выходным сигналами объекта $R_{xy}(\tau)$. Как показывают расчеты по формулам (II) - (I5), применив оптимальное значение $T_o^* < 3T$, можем получить уменьшение D_η по сравнению с его значением при $T_o = 3T$ в два раза.

Графики на фиг. 2 позволяют также выбирать значение для числа оцениваемых ординат ВФ J. При больших значениях параметров $1/Tb$ и λ увеличение J свыше I0 вызывает лишь незначительное повышение точности оценивания, а при малых значениях этих параметров необходимо применять большое значение J для обеспечивания качественного оценивания.

Графики на фиг. I и 2 рассчитаны в случае $\rho = 0$ - на выходе некоррелированная через Δ помеха $\varphi(i\Delta)$. Вообще говоря, T_o^* зависит от автокорреляции шума $\varphi(t)$, что видно из следующего преобразования D_{η_3} - единственного компонента D_η в (II), зависящего от автокорреляции $\varphi(t)$:

$$D_{\eta_3} = \frac{D_\varphi}{T_H}\left[J + \frac{1+\rho\beta}{1-\rho\beta}\right] = \frac{D_\varphi}{N}\left[J + \frac{1+e^{-\frac{T_o(a+b)}{J}}}{1-e^{-\frac{T_o(a+b)}{J}}}\right]. \qquad (I8)$$

Однако дробь в скобках в (I8) существенно отличается от J только при степенях e вблизи I. Тогда можно степень e разложить в ряд по T_o/J и получим:

$$D_{\eta_3} \approx \frac{D_\varphi}{T_H}\left[T_o + \frac{2}{a+b}\right]. \qquad (I9)$$

Здесь видно, что T_o^* практически не зависит от автокорреляции шума, которая, однако, увеличивает минимум D_η. Поэтому в случае коррелированной помехи T_o^* можно найти на фиг. I, а критерий точности оценивания ВФ μ' согласно (I7) и (I9) выражается в виде:

$$\mu' = \sqrt{\mu^2 + \frac{2D\varphi}{D_y(a+b)T_H}}, \qquad (20)$$

где μ соответствует некоррелированной помехе $\varphi(t)$ и находится на фиг. 2.

3. Некоторые особенности оценивания ВФ объекта второго порядка и оценка области применения асимптотических формул

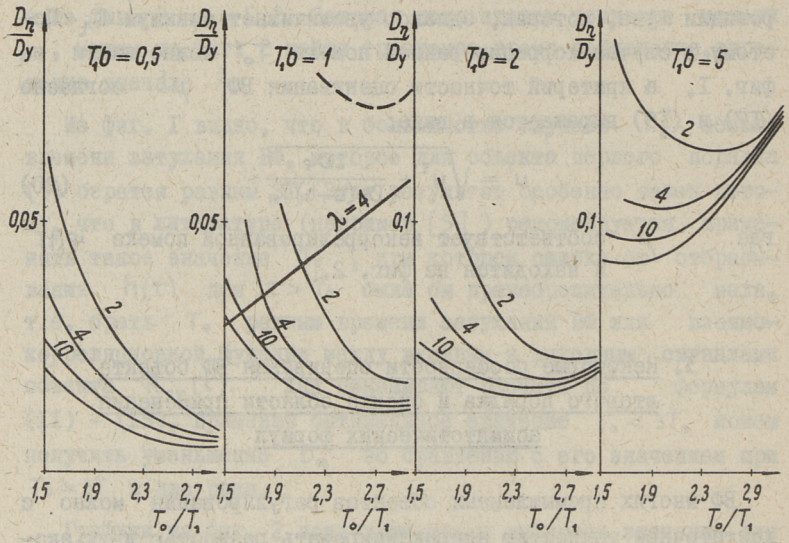
ВФ многих промышленных объектов регулирования можно с достаточной точностью аппроксимировать разностью двух экспонент:

$$h(\tau) = \frac{K}{T_1}\left(e^{-\tau/T_1} - e^{-\tau/T_2}\right), \qquad (2I)$$

где $T_1 > T_2$.

Поскольку аналитическое выражение для D_η стало бы слишком громоздким, было проведено моделирование процесса оценивания ВФ объекта (2I). Часть результатов моделирования при J = I0 показана на фиг. 3, на которой изображены зависимости D_η / D_y от T_o/T_1 при разных значениях $T_1 b$ и отношения T_1/T_2, без шума на выходе. Из графиков фиг. 3 легко также получить зависимость D_η/D_y от T_o/T_1 при наличии шума на выходе при оценивании. Для этого к ним следует прибавить прямую, изображающую зависимость D_{η_3}/D_y от T_o/T_1 (на фиг. 3 это сделано при $T_1/T_2 = 2$, $T_1 b = I$, $\lambda = 4$ и белом шуме, суммарная кривая начерчена пунктиром).

На фиг. 3 видно, что при уменьшении T_1/T_2 T_o^* увеличивается. Например при $T_1/T_2 = 2$, $T_1 b = I$ и $\lambda = 4$ имеем $T_o^* = 2{,}7\,T_1$, а для объекта первого порядка при $T = T_1$, $Tb = 1$,



Фиг. 3. Число на кривой означает T_1/T_2 самая нижняя кривая действительна для объекта первого порядка

$\lambda = 4$ и $J = 10$ получим $T_o^* = 1,9\,T$. Если принять за критерий затухания ВФ ее уменьшение на 20 раз по сравнению с ее максимальным значением, то у объекта второго порядка (2I) время затухания будет больше, чем у объекта первого порядка, получающегося из (2I) при пределе $T_2 \to 0$, так как у объекта (2I) максимум ВФ будет меньше, чем у соответствующего объекта первого порядка, а при $\tau \to \infty$ их ВФ $h(\tau)$ асимптотически совпадают. Например при $T_1/T_2 = 2$ время затухания ВФ (2I) будет $4,25\,T_1$. Благодаря этому отношение T_o^* к времени затухания $h(\tau)$ будет у объектов (2I) и (7) приблизительно то же самое:

$$\frac{1,9\,T}{3\,T} \approx \frac{2,7\,T_1}{4,25\,T_1} = 0,635. \qquad (22)$$

На основании результатов моделирования было вычислено T_o^* для объекта (2I) при разных комбинациях параметров случайных процессов, объекта и модели (T_1b изменялось от 0,I до I0, T_1/T_2 от I0 до 2, λ от 0,I до I0). При всех этих случаях с несущественной погрешностью подтверж-

далось правило, что для объекта второго порядка (2I) можно оптимальное значение памяти модели T_o^* вычислить по графикам на фиг. I, если положить $T = T_1$, а результат преобразовать по формуле

$$T_{oII}^* = \frac{T_o^*}{3T_1} \cdot T_з, \qquad (23)$$

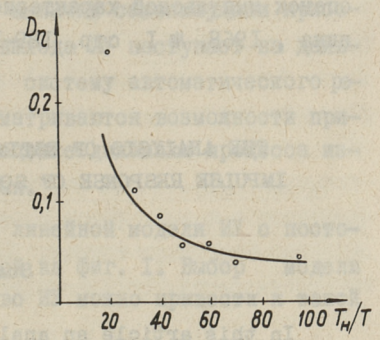
где T_o^* найдено по фиг. I,

T_{oI}^* - оптимальное T_o для объекта второго порядка,

$T_з$ - время затухания ВФ объекта второго порядка.

Точность оценивания с уменьшением T_1/T_2 уменьшается, как видно из графиков на фиг. 3.

Полученное аналитическое выражение для D_η составлено на основе асимтотических формул, в которых оценки корреляционных функций, полученных на интервале наблюдения $T_н$, заменены точными значениями этих функций, что справедливо только при $T_н \to \infty$. Для выяснения возможности применения полученных соотношений при достаточно малых $T_н$ было проведено статистическое моделирование процесса оценивания ВФ объекта первого порядка (7) на цифровой вычислительной машине. Моделирование показало, что аналитической формулой (II) можно успешно пользоваться уже при времени наблюдения, равном $30T \ldots 40T$. При этом истинное значение среднеквадратической погрешности предсказания выхода объекта немного больше расчетного, так как в последнем не были учтены некоторые случайные компоненты ошибок определения ВФ. На фиг. 4 приведены вычисленные на основе асимптотической формулы D_η (сплошная кривая) и полученные моделированием оценки $\tilde{D}_\eta$ (точки) при $J = 10$, $D_y/D_\varphi = 1$, $a = \infty$ и $Tb = 1$.



Фиг.4. Результаты моделирования.

Литература

1. T.P. Goodman, I.B. Reswick. Determination of system characteristic from normal operating records". Transactions of ASME, 1956, Feb. ,pp.259-271.

2. R.B. Kerr, W. Surber. Precision of impulse response identification based on short normal operating record. IRE Transactions on Automatic Control, 1961, May, pp.173-181.

3. H. Kwaakernaak. Some observation on the errors in the estimation of impulse response. IEEE Transactions on Circuit Theory, 1963, No.3, pp.152-159.

4. M.L. Levin. Optimum estimation of impulse response in the presence of noise. IRE Transactions on Circuit Theory, 1960, March, pp.50-56.

5. В.М. Бабурин, В.Л. Ленский, П.С. Матвеев, Ю.Б. Рождественский. О погрешностях при использовании статистического метода для исследования объектов управления . Сб. Аналитические самонастраивающиеся системы автоматического управления , М. 1965.

6. И.И. Перельман. Асимптотические свойства оценок импульсной характеристики . Автоматика и телемеханика , 1968, № I, стр. 189-199.

J. Oll

THE ANALYSIS OF ESTIMATING ACCURACY OF THE IMPULSE RESPONSE OF SOME SIMPLEST LINEAR PLANTS

Summary

In this article an analytical expression is derived, that characterizes estimating accuracy of the impulse response of a first order plant. Graphics have been calculated to find the optimal memory time and appropriate accuracy of the discrete model of the plant. The region of application of asympthotical expressions of 6 have been established by Monte-Carlo method. Some aspects of estimating secondary order plants have also been concerned.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

С Е Р И Я А № 288 1970

УДК 621.317.7.083.94:621.391.23

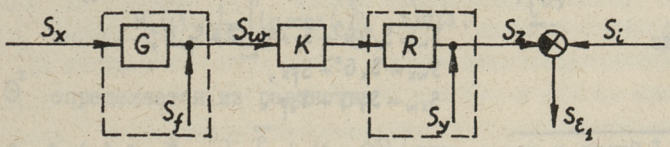
Я.М. Вырк

ПРИМЕНЕНИЕ МЕТОДОВ ТЕОРИИ ИНФОРМАЦИИ ПРИ АНАЛИЗЕ ИЗМЕРИТЕЛЬНЫХ УСТРОЙСТВ И ПРЕОБРАЗОВАТЕЛЕЙ

В последние годы для оценки средств и процессов измерения предложен ряд критериев, основанных на понятии количества информации, введенном Шенноном [1,2]. Разработанные критерии удобны при анализе цифровых и циклических способов измерения.

Во многих случаях можно считать, что как на входе, так и на выходе измерительного устройства (ИУ) имеется непрерывный стационарный случайный процесс [3] (например регистрация непрерывных измеряемых величин самопишущими приборами; случаи, когда сигнал от выхода ИУ поступает на дальнейшую обработку в непрерывную систему автоматического регулирования и т.д.). Ниже рассматриваются возможности применения теории информации при оценке динамики процесса измерения для стационарного случая.

Исходная модель. Исходим из линейной модели ИУ с постоянными параметрами, изображенной на фиг. 1. Выбор модели объясняется тем, что большинство ИУ можно привести к такой схеме.



Фиг. 1

Здесь

S_x – спектральная плотность входного (измеряемого)сигнала $x(t)$;

G – передаточная функция элементов ИУ до корректирующего фильтра (включая датчик);

S_f – спектральная плотность аддитивной погрешности $f(t)$, приведенной к выходу части ИУ, описываемой G;

K – передаточная функция корректирующего фильтра;

R – передаточная функция элементов ИУ, последующих за корректирующим фильтром (включая вторичный прибор);

S_y – спектральная плотность аддитивной погрешности $y(t)$, вызванная элементами ИУ, соответствующих R;

S_z – спектральная плотность выходного сигнала ИУ $z(t)$;

S_i – спектральная плотность желаемого выходного сигнала $i(t)$ (сигнал без погрешности);

S_{ε_1} – спектральная плотность погрешности ИУ $\varepsilon_1(t)$.*

Согласно вышесказанному считаем, что $x(t)$, $f(t)$ и $y(t)$ – стационарные случайные процессы с нормальным распределением. Предположим, что они имеют нулевые математические ожидания и что $y(t)$ некоррелирован с $x(t)$ и $f(t)$, а $f(t)$ может быть коррелирован с $x(t)$. Известны S_x, S_f и S_y, которые могут быть представлены дробно-рациональными функциями от квадрата частоты, а также взаимные спектральные плотности S_{xf} и S_{fx}.

Среднеквадратичная ошибка и передаточная функция корректирующего фильтра. Для оценки точности измерения используем среднеквадратичную ошибку. На основе известных спектральных соотношений, согласно фиг. I, получим

$$\overline{\varepsilon_1^2} = \frac{1}{2\pi}\int_{-\infty}^{\infty}\left(S_x L L^* + S_w K K^* R R^* - S_{wx} K^* R^* L - S_{xw} K R L^* + S_y\right) d\omega, \qquad (I)$$

где

$$S_w = S_x G G^* + S_f + S_{xf} G^* + S_{fx} G,$$
$$S_{wx} = S_x G^* + S_{fx},$$
$$S_{xw} = S_x G + S_{xf},$$

* Здесь и далее для краткости записи $S_x = S_x(\omega)$, $G = G(j\omega)$, $G^* = G(-j\omega)$ и т.д.

а L есть желаемая передаточная функция ИУ, т.е.

$$S_i = S_x L L^*. \qquad (2)$$

Нас интересует минимальная среднеквадратичная ошибка. Поэтому целесообразно синтезировать корректирующий фильтр, обеспечивающий минимальное значение ошибки. Физически реализуемая передаточная функция такого фильтра (вывод ее аналогичен приведенному в [4])

$$K = \frac{1}{R\Psi}\left[\frac{S_{wx} L}{\Psi^*}\right]_+, \qquad (3)$$

где $\Psi\Psi^* = S_w$.

Количество информации и точность. Основой для информационной характеристики ИУ является количество информации на единицу времени $I(i,z)$, содержащееся в выходном сигнале $z(t)$ относительно желаемого выходного сигнала $i(t)$. По [5]

$$I(i,z) = \frac{1}{2\pi}\int_0^\infty \ln \frac{S_i S_z}{S_i S_z - |S_{iz}|^2}\, d\omega,$$

откуда получим окончательно

$$I(i,z) = \frac{1}{2\pi}\int_0^\infty \ln \frac{S_w + \dfrac{S_y}{KK^* RR^*}}{S_f + \dfrac{S_y}{KK^* RR^*} - \dfrac{S_{fx} S_{xf}}{S_x}}\, d\omega. \qquad (4)$$

Для установления связей между количеством информации и точностью используем понятие ε-энтропии. ε-энтропия $H_\varepsilon(i)$ или скорость создания информации - это минимальное количество информации, которое необходимо в единицу времени для того, чтобы с заданной среднеквадратичной ошибкой $\bar{\varepsilon}^2$ воспроизвести данный случайный процесс [6]. Согласно теореме 5.2 из [7]

$$H_\varepsilon(i) = \frac{1}{4\pi}\int_{-\infty}^{\infty} \ln \max\left(\frac{S_i}{\Theta^2}, 1\right) d\omega, \qquad (5)$$

где Θ^2 определяется из равенства

$$\bar{\varepsilon}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \min(S_i, \Theta^2)\, d\omega. \qquad (6)$$

Заметим, что $H_\varepsilon(i)$ не зависит от свойств ИУ, а определяется только измеряемым процессом и желаемой передаточной функцией.

Информационные и точностные характеристики. Введем понятие информационного к.п.д. для ИУ без корректирующего фильтра, обозначая его через $\eta_и$ и определяя как

$$\eta_и = \frac{H_\varepsilon(i)}{I_0(i,z)}, \qquad (7)$$

где $\overline{\varepsilon^2}$ в (6) определяется уравнением (I) и $I_0(i,z)$ уравнением (4) при $K = K^* = 1$. Информационный к.п.д. для ИУ с корректирующим фильтром с передаточной функцией по (3) будет аналогично

$$\eta_{и1} = \frac{H_{\varepsilon_1}(i)}{I_0(i,z)}, \qquad (8)$$

где H_{ε_1} получается из (5) и (6) при замене $\overline{\varepsilon^2}$ на $\overline{\varepsilon_1^2}$.

Чтобы определить соответствующие точностные характеристики, введем понятие информационно-теоретически минимальной среднеквадратичной ошибки $\overline{\varepsilon_0^2}$. $\overline{\varepsilon_0^2}$ определяется уравнениями (5) и (6), если $H_\varepsilon(i) = I_0(i,z)$, т.е. предполагается, что при количестве информации $I_0(i,z)$ ошибка зависит только от свойств желаемого выходного сигнала $i(t)$. Точностной к.п.д. без корректирующего фильтра

$$\eta_т = \frac{\overline{\varepsilon_0^2}}{\overline{\varepsilon^2}} \qquad (9)$$

и с корректирующим фильтром (3)

$$\eta_{т1} = \frac{\overline{\varepsilon_0^2}}{\overline{\varepsilon_1^2}}. \qquad (10)$$

Так как минимальное количество информации $H_\varepsilon(i)$ и определяемое при его помощи информационно-теоретически минимальная среднеквадратическая ошибка $\overline{\varepsilon_0^2}$ зависят от желаемого выходного сигнала $i(t)$, а количество информации $I_0(i,z)$ и среднеквадратичная ошибка кроме того от свойств ИУ (как видно из (I) и (4) при $K = K^* = 1$), то $\eta_и$ и $\eta_т$ характеризуют согласованность динамических свойств ИУ и измеряемого процесса. Такие же характеристики есть и $\eta_{и1}$ и $\eta_{т1}$. Кроме

того, они являются максимально достижимыми значениями к.п.д. системы по фиг. I, ибо для рассматриваемого случая (стационарные сигналы с нормальным распределением) линейная фильтрация является и оптимальной.

При измерениях нас интересует прежде всего получаемая точность. Поэтому точностные характеристики $\overline{\varepsilon_0^2}$, $\eta_т$ и $\eta_{т1}$ окажутся более наглядными характеристиками ИУ, нежели $I_0(i,z)$, $\eta_и$, $\eta_{и1}$.

Возможности применения. I) Предложенные характеристики применимы при анализе ИУ. Они могут быть критериями при сравнении приборов различного типа, при оценке согласованности динамических свойств ИУ и измеряемого процесса, при выяснении целесообразности введения корректирующих фильтров в схему ИУ.

2) Аналогичные характеристики можно определить для отдельных измерительных преобразователей. Для этого придется во всех формулах поставить $R=1$ и $S_y=0$, а под G понимается тогда передаточная функция измерительного преобразователя, для которой определяются информационные и точностные характеристики. Подобно анализу ИУ эти характеристики можно использовать при сравнении различных преобразователей, а также при оценке согласованности их динамических свойств с измеряемым процессом. Особенно полезным может оказаться информационный анализ для датчиков, поскольку во многих случаях динамические свойства (передаточная функция, влияние помех) ИУ определяются именно датчиком.

3) Возможности информационного анализа не исчерпываются вышеописанной моделью. Представляет интерес информационный анализ ИУ, в котором корректирующие фильтры можно включить на выходны нескольких измерительных преобразователей одного ИУ при последовательном их соединении. При этом могут оказаться полезными результаты, полученные в теории информации [8, 9]. Интересно также исследование влияния динамических свойств отдельных преобразователей на информационные и точностные характеристики всего ИУ в зависимости от порядка и способа (последовательное, параллельное, в цепи обратной связи) их включения.

4) Так как количество информации $I_0(i,z)$ зависит от свойств ИУ (см.(6) при $K = K^* = 1$), а тем самым от использованных в нем элементов, то целесообразно, очевидно, уже в стадии разработки ИУ использовать информационные или точностные характеристики с целью их максимизации и ввести в схему ИУ корректирующие фильтры. При этом придется учитывать и некоторые статистические свойства (спектральная плотность, закон распределения) процессов, для которых прибор в основном предназначен.

Пример. В заключение рассмотрим простой пример расчета. Исследуем влияние передаточной функции измерительного преобразователя (датчика) на его точностные характеристики. Предположим, что датчик описывается динамически апериодическим звеном первого порядка с константом времени T, т.е.

$$G = \frac{1}{1+j\omega T}.$$

Пусть, кроме того,

$$S_x = \frac{1}{1+\omega^2};\quad S_f = \frac{0{,}01}{0{,}01+\omega^2};\quad R = 1;\quad S_y = 0;\quad L = 1;\quad S_{fx} = S_{xf} = 0.$$

Уравнение (4) дает при $K = K^* = 1$

$$I_0(i,z) = \frac{1}{2\pi}\int_0^\infty \ln \frac{\omega^4 + \omega^2\left(\frac{101}{T^2}+1\right)+\frac{2}{T^2}}{(\omega^2+1)\left(\omega^2+\frac{1}{T^2}\right)}\,d\omega, \qquad (II)$$

откуда получим (см. приложение)

$$I_0(i,z) = \frac{1}{2}\left(a+b-1-\frac{1}{T}\right),$$

где

$$a^2 = \frac{\frac{101}{T^2}+1}{2} + \sqrt{\left(\frac{\frac{101}{T^2}+1}{2}\right)^2 + \frac{2}{T^2}},$$

$$b^2 = \frac{\frac{101}{T^2}+1}{2} - \sqrt{\left(\frac{\frac{101}{T^2}+1}{2}\right)^2 - \frac{2}{T^2}}.$$

По (3) получим для передаточной функции фильтра

$$K = \frac{A(1+j\omega T)(0{,}1+j\omega)}{0{,}01\,T^2(a+j\omega)(b+j\omega)}.$$

Значения A, a и b для различных T даны в таблице I.

Таблица I

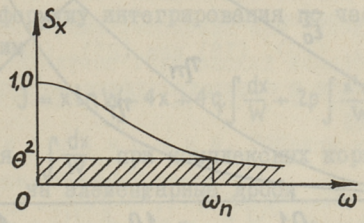
T	50	I0	5	I	0,5	0,I	0,05	0,0I
A	0,530	0,4I5	0,30I	0,0869	0,0459	$9{,}5\ I0^{-3}$	$4{,}77\ I0^{-3}$	$0{,}96\ I0^{-3}$
a	I,02	I,4I	2,24	I0,I	20,I	I00,5	20I	I005
b	0,028	0,I00	0,I26	0,I40	0,I4I	0,I4I	0,I4I	0,I4I

Среднеквадратичные ошибки $\overline{\varepsilon^2}$ и $\overline{\varepsilon_i^2}$ можно вычислить по (I), используя таблицы интегралов [4].

Для определения $\overline{\varepsilon_o^2}$ найдем сперва из (5) Θ^2. Поставим $H_\xi(i) = I_o(i,z)$, где $I_o(i,z)$ определяется равенством (II). Для данного случая (см. фиг. 2, а также [6])

$$I_o(i,z) = \frac{1}{2\pi}\int_0^{\omega_n} \ln \frac{1}{(1+\omega^2)\Theta^2}\, d\omega, \qquad (I2)$$

$$\Theta^2 = \frac{1}{1+\omega_n^2}. \qquad (I3)$$



Фиг. 2.

Представим (I3) в (I2) и вычислим интеграл (I2) при помощи формулы 48.I из [I0]. Получим

$$I_o(i,z) = \frac{\omega_n - \operatorname{arctg}\omega_n}{\pi},$$

откуда для каждого конкретного $I_o(i,z)$ можно вычислить соответствующее ω_n. $\overline{\varepsilon_o^2}$ определяется площадью заштрихованной части на фиг. 2, т.е.

$$\overline{\varepsilon_o^2} = \frac{1}{\pi}\int_0^{\omega_n} \Theta^2 d\omega + \frac{1}{\pi}\int_{\omega_n}^{\infty} \frac{1}{1+\omega^2} d\omega.$$

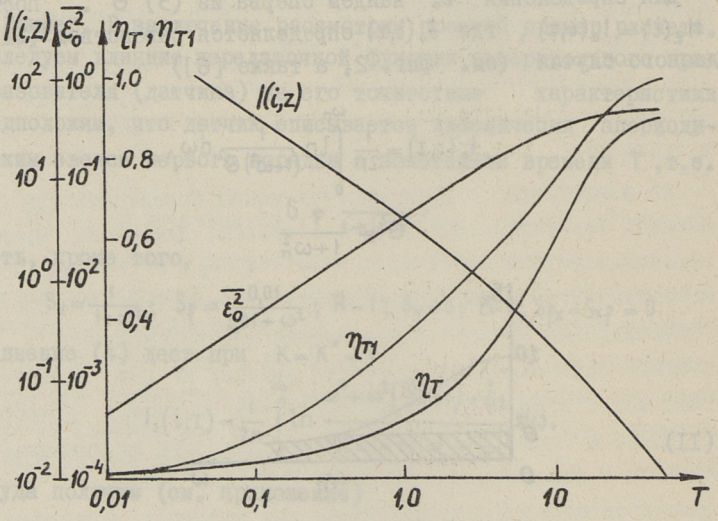
Применяя формулу 5.I из I0 и уравнение (I3), получим окончательно

$$\overline{\varepsilon_0^2} = \frac{\omega_n}{\pi(1+\omega_n^2)} + \frac{1}{2} - \frac{1}{\pi}\operatorname{arctg}\omega_n ,$$

откуда для больших ω_n ($\omega_n >> 1$)

$$\overline{\varepsilon_0^2} \approx \frac{2}{\pi^2[I_0(i,z)+0{,}5]} .$$

Выполненные на основе приведенных формул вычисления изображены на фиг. 3.



Фиг. 3

Как видно из фигуры, малым значениям $I_0(i,z)$ соответствуют сравнительно высокие значения точностных к.п.д. ($T>5$). Линейная фильтрация эффективна при $0{,}1 < T < 2$. При малых T ($T<0{,}1$) точностные к.п.д. малы, что указывает на несогласованность свойств датчика с измеряемым процессом.

Приложение. При дробно-рациональных спектрах случайных процессов определение количества информации на единицу времени ведет к вычислению интеграла вида

$$I = \frac{1}{2\pi}\int_0^\infty \ln\frac{A(\omega)}{B(\omega)}\,d\omega = \frac{1}{2\pi}\int_0^\infty \ln\frac{\sum_{i=0}^{n} a_i\omega^{2i}}{\sum_{i=0}^{n} b_i\omega^{2i}}\,d\omega, \quad b_n = a_n = 1. \qquad (14)$$

Для вычисления интеграла находим корни многочленов $A(\omega)$ и $B(\omega)$ относительно ω^2. Получим

$$I = I_1 + I_2 = \sum \frac{1}{2\pi i} \int_0^\infty \ln \frac{\omega^2 + z_i}{\omega^2 + \omega_i} d\omega + \sum \frac{1}{2\pi i} \int_0^\infty \ln \frac{\omega^4 + c_i\omega^2 + d_i}{\omega^4 + e_i\omega^2 + f_i} d\omega, \quad (15)$$

где в интеграл I_2 выделены трехчлены, имеющие комплексные корни относительно ω^2 или в числителе, или в знаменателе, или в обоих местах.

Интегралы типа I_1 найдем по формуле 4.222.1 из [11]

$$\frac{1}{2\pi i} \int_0^\infty \ln \frac{\omega^2 + z_i}{\omega^2 + \omega_1} d\omega = \frac{1}{2}\left(\sqrt{z_i} - \sqrt{\omega_i}\right). \quad (16)$$

Для определения интегралов в I_2 найдем выражение для неопределенного интеграла

$$J = \int \ln W dx,$$

где

$$W = x^4 + px^2 + q.$$

Используя формулу интегрирования по частям и формулу 14.6 из [10], получим

$$J = x \ln W - 4x + 4q \int \frac{dx}{W} + 2p \int \frac{x^2 dx}{W}. \quad (17)$$

Для нахождения $\int \frac{dx}{W}$ при комплексных корнях (т.е. $q > \frac{p^2}{4}$)* разложим $\frac{1}{W}$ на элементарные дроби

$$\frac{1}{W} = \frac{1}{2\sqrt{q}} \left[\frac{-\frac{x}{2\lambda} + 1}{x^2 - 2\lambda x + \sqrt{q}} + \frac{\frac{x}{2\lambda} + 1}{x^2 + 2\lambda x + \sqrt{q}} \right],$$

где

$$\lambda = \sqrt{\frac{1}{2}\sqrt{q} - \frac{p}{4}}. \quad (18)$$

Формула 13.1 из [10] дает

$$\int \frac{dx}{x^4 + px^2 + q} = \frac{1}{8\lambda\sqrt{q}} \ln \left| \frac{x^2 + 2\lambda x + \sqrt{q}}{x^2 - 2\lambda x + \sqrt{q}} \right| + \frac{1}{4\mu\sqrt{q}} \left(\operatorname{arctg} \frac{x - \lambda}{\mu} + \operatorname{arctg} \frac{x + \lambda}{\mu} \right), \quad (19)$$

где

$$\mu = \sqrt{\frac{1}{2}\sqrt{q} + \frac{p}{4}}. \quad (20)$$

* Формула 14.1 из [10] при комплексных корнях неверна.

При помощи (I8), (I9), (20) и формул I4.4 и II.5 [I0] получим из (I7) при $q_i > \frac{p^2}{4}$ после несложных преобразований

$$J = \int \ln(x^4 + px^2 + q_i)\,dx = x\ln(x^4 + px^2 + q_i) - 4x + \\ + \lambda \ln\left|\frac{x^2 + 2\lambda x + \sqrt{q_i}}{x^2 - 2\lambda x + \sqrt{q_i}}\right| + 2\mu\left(\operatorname{arctg}\frac{x-\lambda}{\mu} + \operatorname{arctg}\frac{x+\lambda}{\mu}\right). \qquad (2I)$$

При $\frac{p^2}{4} > q_i$ на основе формул I4.4, II.5 и I4.I [I0] получим для J

$$J = \int \ln(x^4 + px^2 + q_i)\,dx = x\ln(x^4 + px^2 + q_i) - 4x + \\ + \sqrt{2(p-\sqrt{\delta})}\operatorname{arctg}\frac{\sqrt{2}x}{\sqrt{p-\sqrt{\delta}}} + \sqrt{2(p+\sqrt{\delta}}\operatorname{arctg}\frac{\sqrt{2}x}{\sqrt{p+\sqrt{\delta}}}, \qquad (22)$$

где

$$\delta = p^2 - 4q_i. \qquad (23)$$

Найдем теперь выражение для несобственных интегралов в I_2. Используя (2I), (20), (I8) или (23), (22) и подставляя пределы интегрирования в соответствующие формулы, получим после преобразований

$$\frac{1}{2\pi}\int_0^\infty \ln\frac{\omega^4 + c_i\omega^2 + d_i}{\omega^4 + e_i\omega^2 + f_i}\,d\omega = \frac{1}{2}\left(\sqrt{2\sqrt{d_i} + c_i} - \sqrt{2\sqrt{f_i} + e_i}\right). \qquad (24)$$

Конечный результат (24) не зависит от того, являются ли корни трехчленов относительно ω^2 комплексными или реальными.

Формулы (I5), (I6) и (24) дают возможность вычислить количество информации на единицу времени, определяемое (I4).

Литература

I. П.В. Новицкий. Основы информационной теории измерительных устройств. Энергия, Л. I968.

2. В.И. Рабинович. Информационные методы в теории измерений. Автометрия, 5, I967, I9-26.

3. А.С. Немировский. Вероятностные методы в измерительной технике. Изд. стандартов, М. 1964.

4. Дж. К. Ньютон, Л.А. Гулд, Дж. Ф. Кайзер. Теория линейных следящих систем. ФМЛ, М.1961.

5. М.С. Пинскер. Информация и информационная устойчивость случайных величин и процессов. Изд. АН СССР, М. 1960.

6. Ф.П. Тарасенко. Введение в курс теории информации. Изд. Томского университета. Томск 1963.

7. М.С. Пинскер. Гауссовские источники. Проблемы передачи информации, вып. 14, 1963, 59-100.

8. Р.Л. Добрушин, Б.С. Цыбаков. Передача информации с дополнительным шумом. Проблемы передачи информации, вып. 14, 1963, 21-42.

9. Я.Г. Синай. Наименьшая ошибка и наилучший способ передачи стационарных сообщений при линейном кодировании и декодировании в случае гауссовских каналов связи. Проблемы передачи информации, вып. 2, 1959, 40-48.

10. М.Л. Смолянский. Таблицы неопределенных интегралов. Наука. М. 1967.

11. И.С. Градштейн, И.М. Рыжик. Таблицы интегралов, сумм, рядов и произведений. Физматгиз, М. 1962.

J. Võrk

INFORMATION METHODS IN ANALYSIS OF MEASURING INSTRUMENTS AND TRANSDUCERS

Summary

Using notions of information transmission rate and $\mathcal{E}$ - entropy the information and precision efficiencies are determined. Possibilities of its utilization are observed. In the case of rational spectral densities of normal stationary signals formulae are given for the calculation of information transmission rate.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

СЕРИЯ А № 288 1970

УДК 62-50I.433

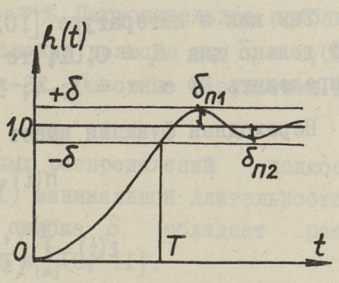
Я.М. Вырк, К.А. Хярмасте

ИССЛЕДОВАНИЕ ВЛИЯНИЯ РАСПОЛОЖЕНИЯ ПОЛЮСОВ НА ДЛИТЕЛЬНОСТЬ ПЕРЕХОДНОЙ ХАРАКТЕРИСТИКИ

В данной статье исследуются переходные характеристики, соответствующие передаточным функциям (ПФ) со всеми нулями в бесконечности, с нормированным свободным членом и коэффициентом при старшем члене в знаменателе

$$\Phi(s) = \frac{1}{s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + 1}. \qquad (I)$$

Задачей является определить для конкретного n такое расположение полюсов нормированной ПФ, для которой длительность переходной функции T (определенная допустимым отклонением δ от установившегося значения, см. фиг.I) была бы минимальна. При этом δ = 0,0I; 0,02; 0,03; 0,05 и n = 3; 4; 5; 6.



Фиг. 1

Очевидно, что целесообразно выбрать ПФ так, что перерегулирование не превосходит допустимой ошибки. Поэтому сразу исключаются из рассмотрения распределения полюсов по Баттерворсу и Чебышева, обладающие большим перерегулированием [I]. Среди распределений полюсов, основанных на минимуме различного вида интегральных критериев, с точки зрения T и величины перерегулирования лучшими являются полюсы, вытекающие из минимума [2]

$$I = \int_0^{\infty} |t(1-h(t))|\, dt.$$

От ПФ, предназначенных для формирования импульсов, обычно требуются быстрое время нарастания (от 0,1 до 0,9 от установившегося значения) и малые значения перерегулирования. Это сопровождается и малыми значениями T.

Принимая во внимание вышесказанное, под рассмотрение были взяты законы расположения полюсов, которые приведены в таблице I. Соответствующие ПФ были надлежащим нормированием приведены к форме (I) и затем вычислены переходные функции. Полученные результаты даны в таблице I.

Сравнивая данные таблицы I, видим, что наилучшими показателями при $n = 5$, $n = 6$ обладает расположение полюсов с равноволновой переходной характеристикой. Расположение характеризуется тем, что для $h(t)$ $n-1$ первых отклонений от конечного значения $(\delta_m \cdots \delta_{n(n-1)})$ равны. В [12] выяснено, что среди ПФ с $n = 3$ минимальное T при заданном $\delta = \delta_t$ имеют ПФ, у которых $\delta_t = \delta_{n1} = \delta_{n2}$, т.е. они соответствуют ПФ с равноволновой переходной характеристикой. Поэтому признаем это расположение среди рассмотренных расположений полюсов наилучшим.

Так как в литературе [10,11] имеются данные о выбранных ПФ только для $\delta_t = 0{,}01$, то поставим дальнейшей задачей определить ПФ с $n = 3; 4; 5; 6$ и $\delta_t = 0{,}02; 0{,}03; 0{,}05$.

Переходная функция имеет вид

$$h(t) = 1 + \varepsilon(t),$$

$$\varepsilon(t) = \sum_{i=1}^{n} \left(\frac{1}{s_i} \prod_{\substack{k=1 \\ k \neq i}}^{n} \frac{1}{s_i - s_k} \right) e^{s_i t},$$

где s_i и s_k – полюса искомых ПФ.

Придется определить для конкретных n все $s_i = \sigma_i + j\omega_i$ при соблюдении условий нормировки (I) и требования, что

$$\delta_t = \delta_{n1} = \cdots = \delta_{n(n-1)}.$$

Используем линеаризированный метод Ремеза, предложенный в [13]. Шаги итерационного процесса следующие:

I) Найдем первые $n-1$ экстремальные точки $\varepsilon(t)$. Соответствующие значения времени обозначим через t_1, t_2, ..., t_{n-1}.

2) Даем всем σ_i и ω_i по одному маленькие приращения (в данной работе использовано приращение 2%) и найдем соответствующие чувствительности

$$\vartheta_{\sigma_i} = \frac{\Delta\varepsilon(t)}{\Delta\sigma_i}, \quad \vartheta_{\omega_i} = \frac{\Delta\varepsilon(t)}{\Delta\omega_i}$$

для моментов времени $t_0 = 0$, $t_1, t_2 \cdots t_{n-1}$.

3) Составим систему линейных уравнений

$$\sum_i \left(\vartheta_{\sigma_i}\Delta\sigma_i + \vartheta_{\omega_i}\cdot\Delta\omega_i\right)_{t=t_0} = 1 + \varepsilon(t_0),$$

$$\sum_i \left(\vartheta_{\sigma_i}\cdot\Delta\sigma_i + \vartheta_{\omega_i}\cdot\Delta\omega_i\right)_{t=t_j} = \delta_t(-1)^j + \varepsilon(t_j),$$

$$j = 1, 2 \cdots n-1$$

и разрешим относительно n приращения $\Delta\sigma_i$ и $\Delta\omega_i$.

4) Изменим все параметры: $\sigma_i = \sigma_i + \Delta\sigma_i$, $\omega_i = \omega_i + \Delta\omega_i$ и повторяем процесс.

Первоначальным приближением для δ_t = 0,02 в данной работе выбрано распределение полюсов при δ_t = 0,01 [10]. Точность определения $\delta_{n1} \cdots \delta_{n(n-1)}$ было ±1%. Первоначальное приближение для δ_t = 0,03 – распределение полюсов для δ_t = 0,02. Задача была решена на ЭЦВМ Минск-22. Основные результаты даны в таблице 2.

Заключение. Среди рассмотренных распределений полюсов [I-II] нормированных ПФ вида (I) минимальной длительностью переходной функции при заданной ошибке δ обладает распределение полюсов, предложенное в [10, II].

Нормированные ПФ являются базой для определения желаемой ПФ системы автоматического регулирования и ПФ различных корректирующих цепей, их применяют при расчетах цепей, формирующих импульсы и т.д. Поэтому полученные результаты можно использовать во всех случаях, когда при заданной ошибке требуется минимальная длительность переходной характеристики.

Таблица I

Расположение полюсов \ n	3			4			5			6			Примечания
	$T_{0,05}$	$T_{0,02}$	δ_{n1} %	$T_{0,05}$	$T_{0,02}$	δ_{n1} %	$T_{0,05}$	$T_{0,02}$	δ_{n1} %	$T_{0,05}$	$T_{0,02}$	δ_{n1} %	
Минимальное время регулирования [3]	4,4	5,I	-	4,6	-	5,I	5,7	6,55	-	6,25	-	4,6	
Критическое затухание переходного процесса [3]	4,4	4,8	2,0	5,I	-	3,0	6,05	6,6	2,0	6,6	-	3,5	Данные из графиков в [3]
Минимальное $\int\|t(t-h(t))\|dt$ [2]	3,65	-	2,0	4,3	4,55	2,0	5,I5	5,4	2,0	5,65	-	5,0	при n = 3, δ_{n2} =2,5%
Плоская аппроксимация К.Г.В.З. [4]	4,6	5,0	0,8	5,4	5,85	0,8	6,25	6,75	0,8	7,I	7,6	0,7	
Чебышевская аппроксимация К.Г.В.З. [5]	5,0	6,7	-	4,65	4,95	I,4	5,25	5,6	0,9	6,0	6,3	0,7	n =3 ε =0,025, n =4 ε =0,0I25, n =5, n =6 ε = =0,0I. См. [5].
Промежуточное между Баттерворсом и плоской аппроксимацией К.Г.В.З. [6]	4,2	-	2,5	4,8	-	3,I	5,7	-	3,5	-	-	-	m =0,6 См. [6] Данные из графиков в [6]
	4,35	4,8	I,4	5,I5	5,5	I,8	5,9	6,25	I,8	-	-	-	m =0,8 См. [6]
Параболическое [7]	3,8	4,I	I,3	5,0	6,8	0,4	6,4	7,8	0,4	-	-	-	b =I,25 см. [7]
Эллиптическое [8]	3,7	-	2,2	4,55	4,9	0,7	5,4	5,7	I,2	6,25	6,6	0,7	a = 0,5 см. [8]
Гиперболическое [9]	3,75	-	2,I	4,7	5,I	-	5,6	6,2	-	-	-	-	b = 5,0 см. [9]
Равноволновая переходная характеристика [I0, II]	3,92	4,24	I,0	4,38	4,64	I,0	4,89	5,I3	I,0	5,42	5,65	I,0	δ_t =0,0I см. [I0]

$T_{0,05}$; $T_{0,02}$ – длительности переходной характеристики при δ = 0,05 и δ = 0,02.

К.Г.В.З. – константное групповое время запаздывания.

Таблица 2

n	δ_t	полюсы			$T_{0,01}$	$T_{0,02}$	$T_{0,03}$	$T_{0,05}$
	0,01	-0,754	-0,613 ± j 0,975	-	4,39	4,24	4,12	3,92
	0,02	-0,722	-0,537 ± j 1,047	-	-	3,87	3,78	3,64
3	0,03	-0,703	-0,484 ± j 1,090	-	-	-	3,58	3,46
	0,05	-0,681	-0,408 ± j 1,141	-	-	-	-	3,22
	0,01	-0,648 ± j 0,392	-0,433 ± j 1,248	-	4,76	4,64	4,54	4,38
	0,02	-0,617 ± j 0,407	-0,349 ± j 1,307	-	-	4,31	4,23	4,10
4	0,03	-0,600 ± j 0,416	-0,295 ± j 1,337	-	-	-	4,04	3,93
	0,05	-0,584 ± j 0,426	-0,224 ± j 1,365	-	-	-	-	3,73
	0,01	-0,628	-0,539 ± j 0,679	-0,295 ± j 1,426	5,25	5,13	5,04	4,89
	0,02	-0,611	-0,511 ± j 0,698	-0,219 ± j 1,462	-	4,82	4,75	4,63
5	0,03	-0,605	-0,497 ± j 0,708	-0,171 ± j 1,476	-	-	4,58	4,48
	0,05	-0,605	-0,489 ± j 0,717	-0,116 ± j 1,477	-	-	-	4,30
	0,01	-0,575 ± j 0,290	-0,449 ± j 0,895	-0,196 ± j 1,539	5,77	5,65	5,57	5,42
	0,02	-0,565 ± j 0,296	-0,427 ± j 0,911	-0,130 ± j 1,553	-	5,39	5,32	5,20
6	0,03	-0,566 ± j 0,299	-0,421 ± j 0,916	-0,095 ± j 1,547	-	-	5,17	5,07
	0,05	-0,576 ± j 0,301	-0,425 ± j 0,917	-0,055 ± j 1.521	-	-	-	4,92

Литература

1. K.W. Henderson, W.H. Kautz. Transient Responses of Conventional Filters. IRE Transactions on Circuit Theory,CT-5,No.4,1958,pp.333-347.

2. F. Csâki, A short survey of deterministic optimization techniques based on integral criteria. Budapest. Periodica Polytechnica. Electrical Engineering - Electrotechnic. 12, No1, 1968,1-15.

3. Н.И. Соколов. Аналитический метод синтеза линеаризированных систем автоматического регулирования. Машиностроение, М. 1966.

4. H.J. Orchard. The Roots of the Maximally Flat-Delay Polynomials. IEEE Transactions on Circuit Theory, CT-12,No 3, 1965,pp.452-454.

5. A.B. Macnee. Chebyshev Approximation of a Constant Group Delay. IEEE Transactions on Circuit Theory, CT-10, No 2, 1963,pp.284-285.

6. Y. Peless, T. Murakami. Analysis and synthesis of transitional Butterworth-Thomson filters and bandpass amplifiers. RCA Review,vol.XVIII, No 1, 1957, pp. 60-94.

7. S.K. Mullick. Pulse Networks with Parabolic Distribution of Poles. IRE Transactions on Circuit Theory,CT-9, No 3, 1961, pp.302-305.

8. J.O. Scanlan. Transfer Functions with Elliptic Distribution of Poles at Equal Frequency Spacings. IEEE Transactions on Circuit Theory, CT-12,No 2, 1965,pp.260-266.

9. M.S. Ghausi, M. Adamowicz. A new class of filters for pulse applications. Journal of Franklin Institute, 282, No 2, 1966, 20-30.

10. J. Jess, H.W. Schuessler. A Class of Pulse-Forming Networks. IEEE Transactions on Circuit Theory, CT-12, No 2, 1965, pp.296-299.

11. J. Jess. Über Impulsfilter mit Tschebyscheffschem Verhalten im Zeit- und Frequenzbereich. A.E.Ü., Nr. 8, 1963, pp.391-401.

12. R. A n r e e s c u. Analiza si sinteza uneî functii de transfer de ordinul trei din punctul de vedere al răspunsului tranzitoriu la impulsuri. Telecomunicatii, Anul 12, №.9, 1968, pp. 345-354.

13. G.C. T e m e s, M. G y i. Design of Filters with Arbitrary Passband and Chebyschev Stopband Attenuation. IEEE, International Convention Record, №.4, 1967, pp. 2-12.

J. Vŏrk, K. Härmaste

THE EFFECT OF POLE LOCATION ON THE CONTROL TIME

Summary

Various transfer functions given in literature /1-11/ are observed. It is found that the transfer functions with equiripple error /10,11/ have minimal duration of step response. Numerical solutions are obtained for these transfer functions with different values of error.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

С Е Р И Я А № 288 1970

УДК 62-501.455

Я.М. Вырк

ПОСЛЕДОВАТЕЛЬНАЯ ЭЛЕКТРИЧЕСКАЯ КОРРЕКЦИЯ ДИНАМИЧЕСКИХ ХАРАКТЕРИСТИК ДАТЧИКОВ

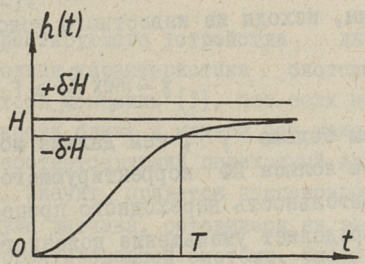
Последовательной электрической коррекции инерционных датчиков первого порядка, т.е. имеющих передаточную функцию (ПФ) вида

$$W_g(s) = \frac{K_g}{T_1 s + 1},$$

посвящено много статей (см., например, перечень литературы в [1]). Гораздо меньшее внимание уделено коррекции инерционных датчиков высокого порядка, имеющих в общем виде ПФ

$$W_g(s) = \frac{B(s)}{A(s)} = \frac{\sum_{i=0}^{m} b_i s^i}{\sum_{i=0}^{n} a_i s^i}, \quad b_0 = a_0 = 1, \quad n > m, \tag{I}$$

где статический коэффициент передачи датчика $W_g(0) = 1$, что не влияет на общность полученных результатов.



Фиг. 1.

В настоящей статье будет рассматриваться вопрос выбора ПФ корректирующего устройства (КУ) для такого датчика при условии, что динамическая ошибка определяется длительностью переходной характеристики (определяется допустимым относительным отклонением δ от установившегося значения, см. фиг. I). Предполагается, что все корни $B(s)$ и $A(s)$ (нули и полюсы) с отрицательными вещественными частями.

Основные понятия. При последовательной коррекции выходной сигнал датчика преобразуется в КУ. ПФ системы датчик-КУ

$$W_c(s) = W_g(s)\,W_к(s).$$

При идеальной коррекции

$$W_c(s) = 1 \quad \text{и} \quad W_к(s) = \frac{1}{W_g(s)},$$

т.е. система воспроизводит без искажений входной сигнал (в том числе и ступенчатое возмущение). Идеальные КУ нереализуемые, так как полином в числителе ПФ имеет порядок выше, чем полином в знаменателе. Поэтому имеет смысл говорить только о приближенной коррекции, где ПФ корректирующего устройства l-ого порядка

$$W_к(s) = K_к \frac{C(s)}{D(s)} = K_к \frac{\sum_{i=0}^{l} c_i s^i}{\sum_{i=0}^{r} d_i s^i}, \quad c_0 = d_0 = 1, \; r \geq l. \qquad (2)$$

Идея последовательной коррекции заключается в компенсации полюсов ПФ датчика нулями ПФ корректирующего устройства и введении новых полюсов (полюсы КУ), которые находятся на комплексной плоскости дальше от мнимой оси, чем первоначальные полюсы датчика. При этом T уменьшается.

Основной характеристикой КУ считаем коэффициент глубины коррекции (или просто глубина коррекции), который определим, исходя из известного условия Фиалкова [2], как

$$\gamma = \max_i \frac{c_i}{d_i} \geq 1 \qquad i = 1 \ldots l. \qquad (3)$$

Чем больше γ, тем дальше можно переместить от мнимой оси полюсы ПФ корректирующего устройства и тем меньше длительность переходного процесса. С другой стороны, γ определяет уменьшение полезного (измеряемого) сигнала, если КУ реализуется пассивными RLC-цепями. Согласно условию Фиалкова, для того, чтобы передаточная функция $W_к(s)$ была реализуема пассивной RLC-цепью, необходимо, чтобы коэффициенты полинома числителя $W_к(s)$ не превышали соответствующих коэффициентов знаменателя. Значит, пассивными RLC-цепями реализуема ПФ корректирующего устройства со статическим коэффициентом передачи

$$W_к(0) = K_к \leq \frac{1}{\gamma}.$$

Если коррекция не должна сопровождаться изменением амплитуды выходного сигнала, то γ определяет минимально нужное усиление, т.е. нужны активные корректирующие цепи. При этом увеличивается влияние помех, действующих на выходе датчика, на скорректированный полезный сигнал по сравнению с нескорректированным сигналом.

Принимая во внимание вышесказанное, поставим дальнейшей задачей определить ПФ корректирующего устройства при известной ПФ датчика и заданном γ так, чтоб динамическая ошибка (в рассматриваемом случае длительность переходного процесса) была бы минимальная. Рассмотрим более детально случай, где в (2)

$$l = r = n,$$

т.е. для коррекции используются форсирующие звенья, а γ определяется условием

$$\gamma = \frac{c_n}{d_n}. \qquad (4)$$

При достаточно больших γ это так и бывает.

ПФ датчика и КУ имеют реальные нули и полюсы. Такой случай встречается часто. Например, такие ПФ имеют многие термометры сопротивления и термопары. Для их коррекции удобно применить КУ на RC-элементах, позволяющие просто реализовать реальные нули и полюсы [2].

а) Определим сперва ПФ корректирующего устройства для случая, когда требуемая переходная характеристика системы должна быть монотонной. Для этого напомним [3], что если на реальной оси комплексной плоскости ближе к мнимой оси находится нуль (или нули) ПФ, то соответствующий переходный процесс имеет перерегулирование. Значит, придется компенсировать КУ-ом кроме полюсов и нули датчика, оставшиеся на реальной оси правее полюсов скорректированной системы. Сначала предполагаем, что все нули датчика остаются правее полюсов скорректированной системы (обычно так и бывает). В этом случае на основе вышесказанного выберем по (I)

$$W_к(s) = K_к \frac{A(s)}{B(s)\,G(s)}, \qquad (5)$$

где

$$G(s) = \sum_{i=0}^{n-m} g_i s^i, \quad g_0 = 1.$$

Сравнивая (5) с (I), (2) и (4), находим

$$D(s) = B(s) \cdot G(s)$$

$$C(s) = A(s)$$

$$\gamma = \frac{a_n}{b_m \cdot g_{n-m}}, \tag{6}$$

откуда

$$g_{n-m} = \frac{a_n}{b_m \cdot \gamma}. \tag{7}$$

ПФ всей системы

$$W_c(s) = K_к \frac{1}{G(s)}.$$

Для данной ПФ определено его порядок $n-m$ и произведение значений полюсов $\frac{1}{g_{n-m}}$. Известно [4], что при реальных полюсах и этих условиях получается минимальная длительность переходного процесса, когда имеется $n-m$-кратный полюс в точке

$$s = -\frac{1}{g_к}, \quad g_к = \sqrt[n-m]{g_{n-m}}. \tag{8}$$

Итак

$$W_к(s) = K_к \frac{A(s)}{B(s)(g_к s+1)^{n-m}}, \tag{9}$$

$$W_c(s) = K_к \frac{1}{(g_к s+1)^{n-m}}. \tag{10}$$

Покажем, используя полученные результаты, что при условии (4) нет смысла использовать для коррекции КУ, где $l < n$ (в частности в [5] предложено использовать КУ с $l = 1$).

При $l < n$ часть из полюсов датчика остается нескомпенсированными. Пусть

$$A(s) = A'(s) \cdot A''(s),$$

где

$$A'(s) = \sum_{i=0}^{q} a'_i s^i, \quad a'_0 = 1 \tag{11}$$

$$A''(s) = \sum_{i=0}^{n-q} a''_i s^i, \quad a''_0 = 1,$$

причем полюсы $A'(s)$ компенсируются. Тогда

$$W'_к(s) = K_к \frac{A'(s)}{B(s)\, G'(s)},$$

где

$$G'(s)=\sum_{i=0}^{q-m} g'_i s^i.$$

Согласно (4)

$$\gamma=\frac{a'_q}{g'_{q-m}\,b_m}, \qquad (12)$$

$$W'_c(s)=K_K\frac{1}{A''(s)\,G'(s)}. \qquad (13)$$

Здесь произведение значений полюсов есть $\frac{1}{g'_{q-m}\,a''_{n-q}}$.
Из (6) и (11) получим при $l=n$

$$\gamma=\frac{a'_q\,a''_{n-q}}{b_m\,g_{n-m}}. \qquad (14)$$

(12) и (14) дают

$$\frac{1}{g_{n-m}}=\frac{1}{a''_{n-q}\,g_{q-m}}. \qquad (15)$$

Так как согласно (15) произведения значений полюсов $W'_c(s)$ из (13) и $W_c(s)$ из (10) равны, $W_c(s)$ имеет согласно (10) кратные полюсы, а $W'_c(s)$ согласно (13) нет (даже не в том случае, если $G'(s)$ имеет кратные полюсы), то длительность переходного процесса, соответствующая $W_c(s)$, меньше [3]. Значит, целесообразно выбрать $l=n$.

Может оказаться, что после выбора передаточной функции КУ по (9) некоторые нули ПФ датчика являются по абсолютной величине больше чем $\frac{1}{g_K}$ из (8). Известно [3], что нули, оставшиеся на реальной оси левее полюсов (по крайней мере для данного случая), не вызывают перерегулирования в переходном процессе. Значит, в таком случае некоторые нули ПФ датчика можно оставить нескомпенсированными. Перепишем $B(s)$ из (1) в виде

$$B(s)=\prod_{i=1}^{m}(T_i s+1),$$

где

$$T_1<T_2<\cdots<T_m.$$

Предположим сперва, что только

$$\frac{1}{T_1}>\frac{1}{g_K} \quad \text{или} \quad g_K>T_1,$$

$$T_1=\alpha_1 g_K \qquad \alpha_1<1. \qquad (16)$$

Оставим нескомпенсированным нуль ПФ датчика в точке $\frac{1}{T_1}$, т.е. выберем ПФ корректирующего устройства

$$W_{к1}(s) = K_к \frac{A(s)\cdot(T_1 s+1)}{B(s)\cdot(g_{к1}s+1)^{n-m+1}},$$

$$\gamma = \frac{T_1 a_n}{b_m g_{к1}^{n-m+1}}. \qquad (I7)$$

При равных γ получим из (I7) с учетом (6) и (I6)

$$g_{к1} = \sqrt[n-m+1]{\alpha_1}\cdot g_к, \quad g_{к1} < g_к. \qquad (I8)$$

ПФ всей системы

$$W_{c1}(s) = K_к \frac{T_1 s+1}{(g_{к1}s+1)^{n-m+1}}.$$

Хотя $W_{c1}(s)$ имеет кратность полюса на один выше чем $W_c(s)$ из (I0), но все-таки можно ожидать, что длительность переходного процесса T, соответствующая $W_{c1}(s)$, меньше, так как по (I8) $\frac{1}{g_{к1}} > \frac{1}{g_к}$ и на быстроту протекания переходного процесса оказывает положительное влияние нескомпенсированный нуль в точке $\frac{1}{T_1}$. Уменьшение T явно видно при $\alpha_1 << 1$.

В общем случае, если

$$T_1 < \cdots < T_i < \cdots < T_j < g_к < T_{j+1}, \qquad (I9)$$

$$\alpha_1 < \cdots < \alpha_i < \cdots < \alpha_j < 1, \qquad \alpha_i = \frac{T_i}{g_к},$$

то можно выбрать

$$W_{кi}(s) = K_к \frac{A(s)}{B_i(s)\cdot(g_{кi}s+1)^{n-m+i}}, \qquad (20)$$

где

$$B_i(s) = \frac{B(s)}{\prod_{q=1}^{i}(T_q s+1)},$$

$$g_{кi} = g_к \sqrt[n-m+i]{\prod_{q=1}^{i}\alpha_q}. \qquad (2I)$$

i определяется из (2I) методом проб при условии что

$$T_i < g_{кi} < T_{i+1}.$$

ПФ всей системы

$$W_{ci}(s) = K_к \frac{\prod_{q=1}^{i}(T_q s+1)}{(q_{ki} s+1)^{n-m+1}}. \qquad (22)$$

Проведя итоги, получим следующую последовательность выбора ПФ корректирующего устройства, если требуемый переходный процесс монотонный. Сперва определим $W_к(s)$ по (9), которая является и окончательной, если $q_к < T_1$. Если же $q_к$ соответствует неравенству $q_к > T_1$, то определим ПФ корректирующего устройства из (20).

б) Если переходный процесс может иметь перерегулирование, то введением диполя [6] можно уменьшить длительность переходного процесса T. Диполь состоит из близко расположенных нуля и полюса ПФ. В данном случае можно диполь получить так, что один полюс (нуль) ПФ датчика не совпадает точно с нулем (полюсом) ПФ корректирующего устройства. Если ближе к мнимой оси остается нуль ПФ скорректированной системы, то переходный процесс имеет перерегулирование. Выбором длины диполя можно добиться, что перерегулирование не превышает допустимой ошибки, в результате чего длительность переходного процесса T уменьшается. Исследуем влияние диполя на переходный процесс, соответствующий ПФ всей системы. После введения диполя (9) и (10) преобразуются в

$$W_{kd}(s) = K_к \frac{A_d(s)\cdot(T_d s+1)}{B(s)\cdot(q_{kd} s+1)^{n-m}}, \qquad (23)$$

$$W_{cd}(s) = K_к \frac{T_d s+1}{(q_d s+1)(q_{kd} s+1)^{n-m}}, \qquad (24)$$

где

$$A_d(s) = \frac{A(s)}{(q_d s+1)},$$

т.е. $\frac{1}{q_d}$ есть нескомпенсированный полюс ПФ датчика.

$$\frac{1}{T_d} < \frac{1}{q_d} \quad \text{или} \quad T_d > q_d.$$

Длина диполя

$$\lambda = \frac{1}{q_d} - \frac{1}{T_d}.$$

Сравнивая (9) и (23), получим при равных γ

$$g_{kd} = g_k \sqrt[n-m]{\frac{T_d}{g_d}} . \qquad (25)$$

При заданном γ точное аналитическое определение влияния диполя на переходный процесс в общем виде сложно. Согласно (25) при введении диполя изменяется положение полюсов КУ. Кроме того, это влияние зависит от длины диполя λ и от его положения на реальной оси.

Если $\frac{1}{g_k} >> \frac{1}{g_d}$ и $\lambda << \frac{1}{g_d}$, то можно приближенно оценить влияние диполя. В таком случае $\frac{T_d}{g_d} \approx 1$ и из (25)

$$g_{kd} \approx g_k .$$

Легко убедиться, исследуя выражение для переходного процесса $h_d(t)$, соответствующего (24), что при этих условиях в окресности T

$$h_d(t) \approx h(t) + g_d \lambda \cdot W_c(0) ,$$

где $h(t)$ – переходный процесс, соответствующий $W_c(s)$ из (10),

$W_c(0)$ – есть статический коэффициент передачи скорректированной системы.

Аналогично определяется влияние диполя на переходный процесс, соответствующий (22).

ПФ датчика и КУ могут иметь комплексные нули и полюсы.

Все сказанное относительно компенсации реальных нулей и полюсов ПФ датчика форсирующими звеньями остается в силу и здесь. ПФ корректирующего устройства можно определить по (9), (20) или (23).

Но как показывают исследования нормированных ПФ [7], при комплексных полюсах можно получить уменьшение длительности переходного процесса T (при тех же γ и порядке КУ). Для определения расположения комплексных полюсов КУ исходим из уравнения (5), а (20) перепишем в виде

$$W_{ki}(s) = K_k \frac{A(s)}{B_i(s) G_i(s)} .$$

Придется определить полюсы $G(s)$ или $G_i(s)$, для которых известны их порядок и коэффициенты члена с наивысшей степенью s: g_{n-m} по (7) или g_{n-m+i}, которое можно определить из (8), (19) и (21) как

$$g_{n-m+i} = g_{ki}^{n-m+i} = \prod_{q=1}^{i} \alpha_q g_{n-m}^{1+\frac{i}{n-m}}.$$

Распределение полюсов $G(s)$ или $G_i(s)$ второго порядка можно определить по [8], а с третьего по шестой порядок по [7]. Для этого выбираем нормированное распределение полюсов соответствующей степени так, что перерегулирование равняется допустимому δ для конкретного случая. Затем произведем денормирование, чтоб получить требуемый $G(s)$ или $G_i(s)$. Денормированием рассматриваемого случая изменится только масштаб времени. Для $W_{ci}(s)$ из (22) придется произвести дополнительно контроль на перерегулирование, так как при выборе $G_i(s)$ мы не учитываем нулей ПФ.

Некоторые соотношения общего вида.

Если КУ должно обладать фильтрующими свойствами (например относительно высокочастотных помех, действующих на выходе датчика), то в (2) целесообразно выбрать $r > l$, $l = n$.

Рассмотрим случай, где

$$W_g(s) = \frac{1}{\sum_{i=0}^{n} a_i s^i} \tag{26}$$

и все полюсы реальные. КУ имеет реальные нули и полюсы (реализуется RC -цепями). Выберем с точки зрения минимальности переходного процесса полюсы КУ кратными

$$W_K(s) = K_K \frac{\sum_{i=0}^{n} a_i s^i}{\left(\frac{ds}{r}+1\right)^r}, \tag{27}$$

т.е. в (2) $c_i = a_i$.

Придется определить d при известных r и γ.

На основе [9]

$$\left(\frac{ds}{r}+1\right)^r = \sum_{i=0}^{r} \frac{r!}{i!(r-i)!} \frac{d^i s^i}{r^i}, \tag{28}$$

где обозначаем

$$d_i = \frac{r!}{i!(r-i)!} \frac{d^i}{r^i},$$

так что

$$W_K(s) = K_K \frac{\sum_{i=0}^{n} a_i s^i}{\sum_{i=0}^{r} d_i s^i}. \qquad (29)$$

На основе (3), (28) и (29)

$$\gamma = \max_i \frac{a_i}{d_i} = \max_i \frac{a_i i!(r-i)!\, r^i}{r!\, d^i}, \quad i = 1 \dots n,$$

откуда

$$d = \max_i \sqrt[i]{\frac{a_i i!(r-i)!\, r^i}{r!\, \gamma}}, \quad i = 1 \dots n. \qquad (30)$$

При реальных корнях ПФ датчика и КУ формулы (30) и (27) определяют ПФ корректирующего устройства, если в (2) $r > l$, $l = n$. ПФ датчика определяется из (26) и γ из (3). Эти формулы действительны и при $r = l$, т.е. для коррекции используются форсирующие звенья.

Представляет интерес поведение переходного процесса при $r \to \infty$. Тогда, так как [9]

$$\lim_{r \to \infty} \frac{1}{\left(\frac{ds}{r} + 1\right)^r} = e^{-ds}$$

$$W_c(s) = K_K e^{-ds},$$

т.е. с увеличением r при константном γ ПФ скорректированной системы приближается к запаздывающему звену. При этом

$$d = \max_i \sqrt[i]{\frac{a_i i!}{\gamma}}, \quad i = 1 \dots n,$$

так как

$$\lim_{r \to \infty} \frac{(r-i)!\, r^i}{r!} = 1.$$

Если в (26) имеются дополнительно реальные нули, то соответствующие формулы для d и при $r \to \infty$ получаются слишком громоздкие. Если полюсы КУ могут быть комплексными, то не оказывается возможным дать аналогичные формулы. При этих случаях при выборе ПФ корректирующего устройства можно пользоваться рассуждениями, приведенными для КУ на форсирующих звеньях, с учетом того, что γ определяется формулой (3).

Заключение. Для уменьшения динамической ошибки датчиков высокого порядка, если ошибка определяется длительностью переходной характеристики, можно использовать включенные

последовательно с датчиком корректирующие устройства (КУ). Если КУ имеет передаточную функцию (ПФ) форсирующего звена, то введением понятия коэффициента глубины коррекции, исходя из известного условия Фиалкова, удалось доказать, что порядок ПФ корректирующего устройства должен равняться порядку ПФ датчика (при условии, что коэффициент глубины коррекции определяется отношением коэффициентов старших членов ПФ корректирующего устройства, как это обычно и бывает). При этом условии дана методика расчета ПФ корректирующего устройства для датчика с ПФ, имеющими все нули и полюсы в левой полуплоскости.

Если КУ должно обладать фильтрующими свойствами, то целесообразно выбрать порядок ПФ корректирующего устройства выше порядка ПФ датчика. При этом выведены формулы расчета ПФ корректирующего устройства, исходя из введенного понятия коэффициента глубины коррекции для случая, когда датчик имеет только реальные полюсы, а КУ может иметь действительные отрицательные нули и полюсы.

Литература

I. D. Hofmann. Zur elektrischen Korrektur des dynamischen Verhaltens von trägen Messwandlern. MSR, 1. 1967, S. 20-27.

2. Н. Балабанян. Синтез электрических цепей. Госэнергоиздат, 1961.

3. З.Ш. Блох. Переходные процессы в линейных системах автоматического регулирования. ФМЛ, М. 1961.

4. А.А. Красовский, Г.С. Поспелов. Основы автоматики и технической кибернетики. Госэнергоиздат, М.-Л. 1962.

5. Е.М. Душин, Д.Ф. Хофман. Последовательная коррекция промышленных электрических термометров. Изв. ВУЗ "Приборостроение", I, 1966, 141-145.

6. Д. Т р а к с е л. Синтез систем автоматического регулирования. Изд. Машиностроительной литературы, М. 1959.

7. Я.М. В ы р к, К.А. Х я р м а с т е. Исследование влияния расположения полюсов на длительность переходной характеристики. См. настоящий сборник, стр. 103-109.

8. С.Д. Т р о п с к и й, Э.М. И б р а г и м о в. Исследование переходных характеристик в комплексной плоскости. Сб. Трудов ЛМИ № 62 "Техническая кибернетика", 3, 1967, 88-99.

9. СМБ. Математический анализ. Функции, пределы, ряды, цепные дроби, ФМЛ, М. 1961.

J. Vörk

SERIAL ELECTRICAL CORRECTION OF THE DYNAMIC CHARACTERISTICS OF TRANSDUCERS

Summary

The coefficient of correction is defined. If the dynamic error is determined by stabilization time, the rules for choice of rational transfer functions of serial correction circuits will be given.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

СЕРИЯ А № 288 1970

УДК 621.317.745.024.001.24

А.А. Лаансоо

РАСЧЕТ ПАРАМЕТРОВ ИЗМЕРИТЕЛЬНЫХ ЦЕПЕЙ ПРИ НАЛИЧИИ В ИХ ПЕРЕДАТОЧНЫХ ФУНКЦИЯХ ДИПОЛЕЙ

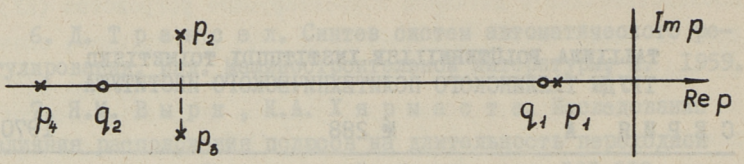
Многие задачи измерительной техники требуют исследования хода переходных характеристик при точности установления порядка 0,1...0,01% и выше, вплоть до уровня шумов в элементах схемы. Нередко после затухания основной, быстрой части переходного процесса происходит медленное "подползание" выходного сигнала к его установившемуся значению. Одной из причин такого явления является наличие в передаточной функции схемы близко расположенных пар "нуль-полюс". В теории автоматического управления [1, 2] такие пары принято называть диполями. Если диполь p_1, q_1 (см. фиг. 1) находится на комплексной плоскости значительно ближе к мнимой оси, чем остальные полюсы, то ход переходной характеристики при больших временах ($t \to \infty$) определяется выражением

$$x_{вых}(t) \approx x_{вых}(\infty) + x_{1вых}(t), \qquad (1)$$

где $x_{вых}(\infty)$ – установившееся значение выходного сигнала,

$x_{1вых}(t)$ – составляющая переходной характеристики, обусловленная полюсом p_1.

Затухание составляющей $x_{1вых}(t)$ происходит значительно медленнее, чем затухание остальной части переходного процесса. Это приводит к уменьшению быстродействия цепи (приходится ожидать окончания "подползания"). Поэтому во многих случаях, где это возможно, целесообразно рассчитывать



Фиг. 1. К влиянию диполя

схему так, чтобы выполнялось неравенство:

$$\left|\frac{x_{1вых}(0)}{X_{вых\,н}}\right| < \delta_1 , \qquad (2)$$

где $X_{вых\,н}$ – номинальное значение выходного сигнала,

δ_1 – часть допустимой погрешности, отведенная для $x_{1вых}(t)$.

При соблюдении неравенства (2) влиянием диполя на переходную характеристику цепи можно пренебречь.

Ниже рассматриваются условия, обеспечивающие выполнение неравенства типа (2) для разомкнутой цепи и для измерительного усилителя постоянного тока (УПТ) с обратной связью.

Разомкнутая цепь

Для анализа представим передаточную функцию (ПФ) цепи в виде

$$W(p)=W(0)\,\frac{\left(1-\frac{p}{q_1}\right)\prod_{i=2}^{m}\left(1-\frac{p}{q_i}\right)}{\left(1-\frac{p}{p_1}\right)\prod_{i=2}^{n}\left(1-\frac{p}{p_i}\right)} , \qquad (3)$$

где $W(0)$ – коэффициент передачи для постоянной составляющей сигнала,

p_i, q_i – полюсы и нули ПФ цепи,

p_1, q_1 – полюс и нуль, образующие диполь (см. фиг. I).

Из формулы (3) получаем для составляющей $x_{1вых}(t)$ следующее выражение:

$$x_{1вых}(t) = x_{вх} \cdot W(0) \cdot \left(1 - \frac{p_1}{q_1}\right) \cdot \frac{\prod_{i=2}^{m} \left(1 - \frac{p_1}{q_{i1}}\right)}{\prod_{i=2}^{n} \left(1 - \frac{p_1}{p_i}\right)} e^{p_1 t}, \qquad (4)$$

где $x_{вх}$ – величина входного скачка.

Для модуля $|x_{1вых}(t)|$ получаем из (4):

$$|x_{1вых}(t)| \leq |x_{вх} \cdot W(0)| \cdot \left|1 - \frac{p_1}{q_1}\right| \cdot \frac{\prod_{i=2}^{m} \left|1 - \frac{p_1}{q_i}\right|}{\prod_{i=2}^{n} \left|1 - \frac{p_1}{p_i}\right|} e^{p_1 t}, \qquad (5)$$

Если остальные полюсы находятся от мнимой оси значительно дальше, чем диполь, то можно считать, что для $i > 1$

$$\begin{aligned} 1 - \frac{p_1}{q_i} &\approx 1, \\ 1 - \frac{p_1}{p_i} &\approx 1. \end{aligned} \qquad (6)$$

Учитывая (6), выражение (5) приобретает вид:

$$|x_{1вых}(t)| \leq |x_{вх} \cdot W(0)| \cdot \left|1 - \frac{p_1}{q_1}\right| e^{p_1 t}. \qquad (7)$$

Неравенство (2) будет выполнено, если

$$\frac{|p_1 - q_1|}{|q_1|} < \delta_1 \gamma_1, \qquad (8)$$

где

$$\gamma_1 = \left| \frac{x_{вых\,н}}{x_{вх} \cdot W(0)} \right|. \qquad (9)$$

Коэффициент γ_1 учитывает величину входного воздействия. Обычно $\gamma_1 \approx 0{,}5 \ldots 1$.

Неравенство (8) определяет максимально допустимую длину диполя (длиной диполя принято называть модуль разности $|p_1 - q_1|$), при которой данный диполь можно еще не учитывать при анализе динамики цепи.

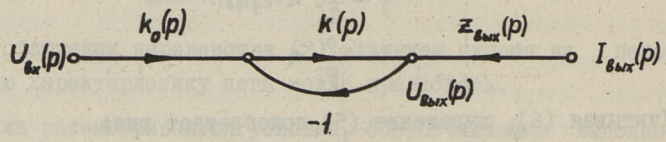
Пример I. Для $\gamma_1 = 1$, $\delta_1 = 10^{-4}$ получаем из формулы (8):

$$\frac{|p_1 - q_1|}{|q_1|} < 10^{-4},$$

т.е. допускается разница между нулем и полюсом в $\pm$ 0,0I%. Выполнять такое требование весьма затруднительно, поэтому в разомкнутой цепи желательно избегать появления диполей.

Измерительный УПТ

Рассмотрим общую структурную схему УПТ, представленную в виде сигнального графа на фиг. 2. Исходными параметрами для анализа являются коэффициент передачи по петле обратной связи $k(p)$, выходное сопротивление разомкнутого усилителя $z_{вых}(p)$ и масштабный коэффициент усиления УПТ $k_0(p)$.



Фиг. 2. Сигнальный граф УПТ.

Для анализа представим $k(p)$ в виде:

$$k(p) = k\frac{A(p)}{B(p)}, \qquad (10)$$

где $k = k(0)$ – петлевое усиление,

$A(p), B(p)$ – многочлены от p.

Выходное напряжение $U_{вых}(p)$ на основе фиг. 2 с учетом формулы (10) выражается:

$$U_{вых}(p) = k_з(p)U_{вх}(p) + z_{вых з}(p)\cdot I_{вых}(p), \qquad (11)$$

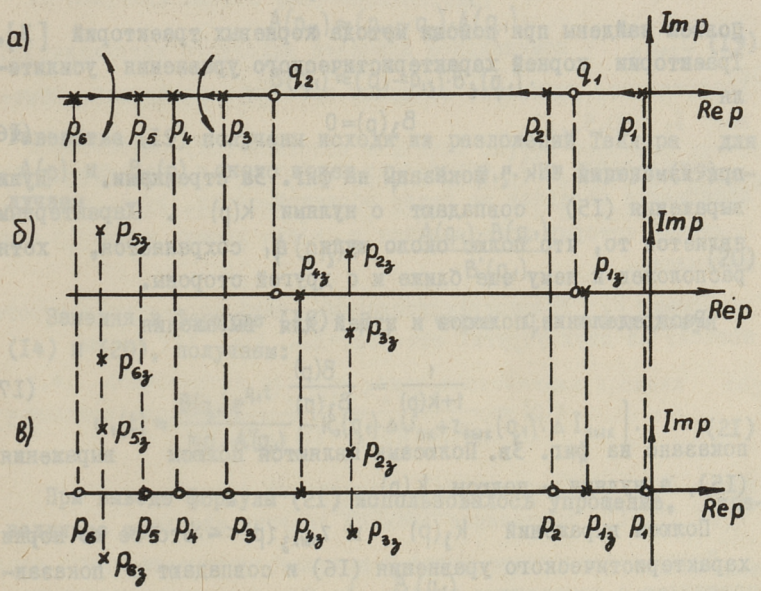
где $k_з(p)$, $z_{вых з}(p)$ – коэффициент усиления и выходное сопротивление замкнутого УПТ:

$$k_з(p) = k_о(p)\,\frac{kA(p)}{B_з(p)}, \qquad (I2)$$

$$z_{вых\,з}(p) = z_{вых}(p)\,\frac{B(p)}{B_з(p)}, \qquad (I3)$$

$$B_з(p) = B(p) + kA(p). \qquad (I4)$$

Параметры $k_з(p)$ и $z_{вых\,з}(p)$ определяют ход переходных характеристик при скачкообразных изменениях $U_{вх}(p)$ и $I_{вых}(p)$.



Фиг. 3. Расположения полюсов и нулей:
а) для $k(p)$, б) для $\frac{k(p)}{1+k(p)}$,
в) для $\frac{1}{1+k(p)}$

Рассмотрим расположения их полюсов и нулей.

Типичное расположение нулей и полюсов для $k(p)$ показано на фиг. 3а. Полюсы $p_3 \ldots p_6$ обусловлены паразитными высокочастотными параметрами элементов схемы, а полюс p_1 и нуль q_2 – параметрами корректирующей цепи. При больших ко-

эффициентах петлевого усиления цепь коррекции обычно составляется из двух или большего числа упругих звеньев, которые вводятся в различные усилительные каскады. Из-за разброса параметров элементов отдельных упругих звеньев имеет место их неточное сопряжение, которое выражается в появлении диполей (диполь p_2 , q_1 на фиг. 3а).

На фиг. 3б изображено распределение нулей и полюсов для выражения

$$\frac{k(p)}{1+k(p)} = k\,\frac{A(p)}{B_3(p)}. \qquad (15)$$

Полюсы найдены при помощи метода корневых траекторий [3]. Траектории корней характеристического уравнения усилителя

$$B_3(p)=0 \qquad (16)$$

при изменении " k " показаны на фиг. 3а стрелками. Нули выражения (15) совпадают с нулями $k(p)$. Характерным является то, что полюс около нуля q_1 сохраняется, хотя расположен к нему еще ближе и с другой стороны.

Распределение полюсов и нулей для выражения

$$\frac{1}{1+k(p)} = \frac{B(p)}{B_3(p)} \qquad (17)$$

показано на фиг. 3в. Полюсами являются полюсы выражения (15), а нулями - полюсы $k(p)$.

Полюсы выражений $k_3(p)$ и $z_{вых\,3}(p)$ - это те же корни характеристического уравнения (16) и совпадают с показанными на фиг. 3б и 3в. Расположения нулей для $k_3(p)$ и $z_{вых\,3}(p)$ в общем случае могут не совпадать с показанными на фиг. 3б и 3в. В дальнейшем предполагаем, что нуль q_1 выражения (15) является нулем и для $k_3(p)$. Это справедливо, если q_1 не является полюсом для $k_0(p)$.

Учитывая изложенное выше, получаем из формул (11)...(14) для составляющей $u_1(t)$ переходной характеристики УПТ, обусловленной полюсом p_{13} (фиг. 3б и 3в), следующее выражение:

$$u_1(t) = \frac{k \cdot k_0(p_{13}) \cdot A(p_{13})}{p_{13} \cdot B'(p_{13})} e^{p_{13}t} \cdot \Delta U_{вх} + \frac{z_{вых}(p_{13}) B(p_{13})}{p_{13} \cdot B'(p_{13})} e^{p_{13}t} \Delta I_{вых}, \quad (18)$$

где $\Delta U_{вх}$, $\Delta I_{вых}$ – приращения входного напряжения и выходного тока – независимые внешние воздействия.

Формулу (18) можно упростить, заменяя в ней p_{13} через q_1. Такая замена не оправдывает себя только для $A(p_{13})$, так как $A(q_1) = 0$. Приближенное выражение для $A(p_{13})$ можно получить, используя равенства:

$$A(p_{13}) \approx (p_{13} - q_1) \cdot A'(q_1),$$
$$B(q_1) \approx (q_1 - p_{13}) \cdot B'_3(q_1). \quad (19)$$

Равенства (19) получены исходя из разложений Тейлора для $A(p)$ и $B_3(p)$ около точек q_1 и p_{13}. Из формул (19) получаем

$$A(p_{13}) \approx - \frac{A'(q_1) \cdot B(q_1)}{B'(q_1)}. \quad (20)$$

Заменяя в формуле (18) p_{13} через q_1 и используя (14) и (20), получаем:

$$u_1(t) \approx \frac{B(q_1) e^{q_1 t}}{k q_1 \cdot A'(q_1)} \left[-k_0(q_1) \cdot \Delta U_{вх} + z_{вых}(q_1) \cdot \Delta I_{вых} \right]. \quad (21)$$

При выводе формулы (21) использовалось упрощение, справедливое при $k >> 1$:

$$1 + \frac{1}{k} \cdot \frac{B'(q_1)}{A'(q_1)} \approx 1. \quad (22)$$

Формула (21) позволяет оценить величину "подползания", возникающего из-за наличия диполя, используя только параметры разомкнутого УПТ. Погрешность определения величины $u_1(t)$ с помощью формулы (21) будет невелика (порядка нескольких процентов), если: 1) q_1 не является полюсом для $k_0(p)$ и 2) расстояние между полюсом p_{13} и ближайшим к нему нулем выражения $z_{вых3}(p)$ значительно (более, чем в 15 ...20 раз) превышает длину диполя p_{13}, q_1. Случаи, когда

эти требования не выполняются, в данной статье не рассматриваются.

Условие (2), обеспечивающее малость составляющей $U_1(t)$ переходной характеристики, в данном случае имеет вид:

$$\left| \frac{B(q_1)}{k \cdot q_1 \cdot A'(q_1)} \right| < \gamma_2 \delta_1 , \qquad (23)$$

где

$$\gamma_2 = \left| \frac{U_{вых\,н}}{-k_0(q_1) \cdot \Delta U_{вх} + z_{вых}(q_1) \cdot \Delta I_{вых}} \right| , \qquad (24)$$

$U_{вых\,н}$ – номинальное значение выходного напряжения УПТ.

Коэффициент γ_2 учитывает величину внешних воздействий. Обычно $\gamma_2 \approx 0{,}5 \ldots 1$.

Неравенство (23) определяет требования к параметрам коэффициента передачи по петле обратной связи $k(p)$, при выполнении которых содержащийся в $k(p)$ диполь можно не учитывать.

Пример 2. Рассмотрим усилитель, цепь коррекции которого состоит из двух упругих звеньев. Для $k(p)$ УПТ имеем:

$$k(p) = \frac{k(pT^*_{к1}+1)(pT^*_{к2}+1)}{(pT_{к1}+1)(pT_{к2}+1)\prod_{i=3}^{n}(pT_i+1)} , \qquad (25)$$

где T_i – "паразитные" постоянные времени,
$T_{к1}, T_{к2}, T^*_{к1}, T^*_{к2}$ – постоянные времени упругих звеньев.

Пусть расположение нулей и полюсов для выражения (25) совпадает с показанным на фиг. 3а (p_2, q_1 образуют диполь):

$$p_1 = -\frac{1}{T_{к1}} ; \quad p_2 = -\frac{1}{T_{к2}} ; \quad p_i = -\frac{1}{T_i} ; \qquad (26)$$

$$q_1 = -\frac{1}{T^*_{к1}} ; \quad q_2 = -\frac{1}{T^*_{к2}} .$$

Предполагаем еще, что

$$k_0(q_1) = k_0(0); \; z_{вых}(q_1) = z_{вых}(0); \; \gamma_2 = 1. \qquad (27)$$

Тогда неравенство (23) приобретает вид:

$$\left|\frac{\left(1-\frac{T_{к1}}{T^*_{к1}}\right)\left(1-\frac{T_{к2}}{T^*_{к1}}\right)\prod_{i=3}^{n}\left(1-\frac{T_i}{T^*_{к1}}\right)}{k\left(-\frac{1}{T^*_{к1}}\right)\left(1-\frac{T^*_{к2}}{T^*_{к1}}\right)T^*_{к1}}\right| < \delta_1 . \qquad (28)$$

Так как обычно

$$\frac{T_{к1}}{T^*_{к1}} >> 1; \quad \frac{T_i}{T^*_{к1}} << 1; \quad \frac{T^*_{к2}}{T^*_{к1}} << 1, \qquad (29)$$

то получаем вместо (28) неравенство:

$$\frac{T_{к1}}{kT^*_{к1}}\left|1-\frac{T_{к2}}{T^*_{к1}}\right| < \delta_1 . \qquad (30)$$

Неравенство (30) является в данном случае условием для сопряжения параметров (постоянных времени) отдельных корректирующих звеньев. Например, если $T_{к1}$ = 10 мсек, $T^*_{к1}$ = = 0,5 мсек, $k = 10^4$, $\delta_1 = 10^{-4}$, то

$$\left|1-\frac{T_{к2}}{T^*_{к1}}\right| < \frac{kT^*_{к1}}{T_{к1}}\delta_1 = 0{,}05,$$

т.е. допускается разница между постоянными времени корректирующих звеньев в ± 5%.

Заключение

Одной из причин "подползания" выходного сигнала к своему установившемуся значению является наличие в передаточной функции цепи диполей, расположенных ближе к мнимой оси, чем остальные полюсы. Обеспечение малости составляющих переходного процесса, обусловленных диполями, в случае разомкнутой цепи весьма трудно.

В УПТ (цепь с обратной связью) такие диполи могут возникать из-за неточного сопряжения постоянных времени корректирующих звеньев. Из-за влияния обратной связи условия сопряжения постоянных времени, обеспечивающие малость "подползания", становятся менее жесткими и могут быть легко выполнены.

В статье получены расчетные формулы, позволяющие оценить величину погрешности, позникающей за счет "подползания", для случая разомкнутой цепи и для УПТ с обратной связью.

Литература

I. Н.Т. Кузовков. Динамика систем автоматического управления. "Машиностроение", Москва 1968.

2. Дж. Тракселл. Синтез систем автоматического регулирования. Машгиз, М. 1959.

3. Г.А. Бедриков, К.Ф. Теодорчик. Траектории корней линейных автоматических систем. Наука, Москва 1964.

A. Laansoo

BESTIMMUNG DER PARAMETER DER MESSUNGSKETTEN, DEREN ÜBERTRAGUNGSFUNKTIONEN DIE DIPOLE BEHALTEN

Zusammenfassung

Die Dipole in der Übertragungsfunktion, die der imaginären Achse näher stehen, als die anderen Pole und Nullstellen, ergeben im Übertragungsprozeß die Komponenten, die kleine Absolutbeträge und ein langsames Exponentialabklingen haben. Es werden Formeln gegeben, mit deren Hilfe man die Parameter der Kette so berechnen kann, daß die Absolutbeträge dieser Komponenten kleiner werden, als der zulässige Fehler der Kette.

TALLINNA POLÜTEHNILISE INSTITUUDI TOIMETISED
ТРУДЫ ТАЛЛИНСКОГО ПОЛИТЕХНИЧЕСКОГО ИНСТИТУТА

СЕРИЯ А № 288 1970

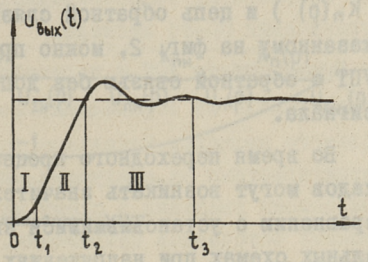
УДК 621.317.745.024.001.24

А.А. Лаансоо

РАСЧЕТ ВРЕМЕНИ УСТАНОВЛЕНИЯ В УПТ С ОБРАТНОЙ СВЯЗЬЮ ПРИ НЕЛИНЕЙНОМ РЕЖИМЕ ИХ РАБОТЫ

Многокаскадные УПТ с глубокой отрицательной обратной связью широко применяются в технике. Анализ их динамики и расчет цепей коррекции обычно ведутся исходя из линейной модели УПТ. Такой подход во многих случаях является односторонним. Например, при помощи линейной модели можно определить параметры корректирующих цепей, но их расстановка по каналу усиления (первые, последние каскады) линейной моделью не учитывается. Рассмотрение нелинейной модели УПТ позволяет получить критерии для определения правильной расстановки корректирующих цепей.

При достаточно большом скачкообразном входном воздействии характерный ход выходного напряжения УПТ показан на фиг. 1.



Фиг. 1. Переходная характеристика УПТ

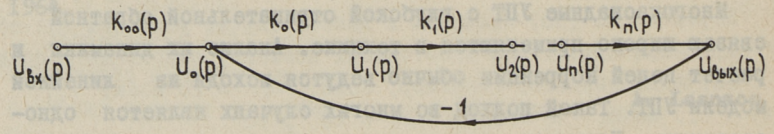
Переключение на выходе происходит в момент времени $t = 0$. При $t = t_1$ часть усилительных каскадов входит в ограничение. В течение времени от t_1 до t_2 усилитель работает в нелинейном режиме. Далее ($t > t_2$) происходит окончательное установление выходного напряжения в линейном режиме. Время установления t_y определяется как сумма трех слагаемых:

$$t_y = t_1 + t_н + t_л, \qquad (I)$$

где t_1, $t_н$, $t_л$ – длительности работы усилителя на участках I, II, III (см. фиг. I).

При невысокой точности установления (порядка нескольких процентов) определяющим для t_y обычно является $t_н$. При точностях установления порядка 0,I...0,0I% и выше слагаемые $t_н$ и $t_л$ обычно сравнимы. Слагаемое t_1 во всех случаях является малым по сравнению с t_y и при расчетах им можно пренебречь. Изложенные положения служат исходными для дальнейшего анализа.

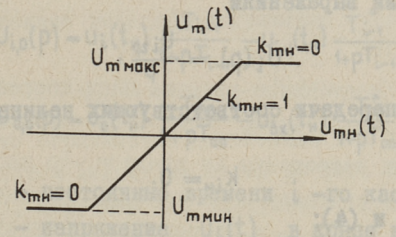
Рассмотрим УПТ, структура которого показана на фиг. 2 в виде сигнального графа.



Фиг. 2. Структура УПТ

Усилитель содержит n усилительных каскадов ($k_1(p)$ .., $k_n(p)$) и цепь обратной связи ($k_{oc}(p)$, $k_o(p)$). К виду, показанному на фиг. 2, можно привести сигнальный граф любого УПТ с обратной связью без дополнительных каналов передачи сигнала.

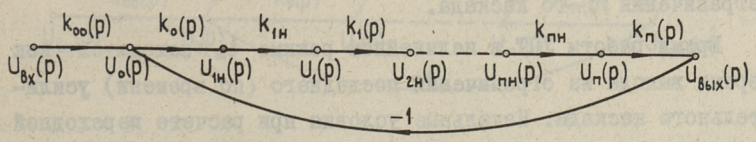
Во время переходного процесса на входах отдельных каскадов могут возникать значительные перенапряжения (по сравнению с установившимися значениями напряжений). В реальных схемах при напряжениях, превышающих определенное значение, возникает режим ограничения. Ограничение на входе m-го каскада можно учесть при помощи введения в схему на фиг. 2 перед m-ым каскадом специального нелинейного звена, передаточная функция которого зависит от уровня сигнала. Для дальнейшего анализа предполагаем, что: I) нелинейное звено является безынерционным и 2) его характеристику можно аппроксимировать прямыми линиями (фиг. 3).



Фиг. 3. Характеристика нелинейного звена

На фиг. 3 применены обозначения $U_{mн}(t)$ - сигнал на входе нелинейного звена; $U_{mмакс}, U_{mмин}$ - граничные значения сигнала на входе m -го каскада; $k_{mн}$ - коэффициент передачи нелинейного звена. Таким образом, при $U_{mмин} < U_{mн}(t) < U_{mмакс}$ получаем $k_{mн}=1$ (линейный режим); в остальных случаях $k_{mн}=0$. В реальных схемах предположения I и 2 обычно выполняются при применении специальных схем ограничения на диодах.

В ограничении одновременно могут находиться и несколько (2,3...) усилительных каскадов. Общая структура усилителя с учетом эффекта ограничения показана в виде сигнального графа на фиг. 4. В схему включены дополнительные нелинейные звенья $k_{1н}, \ldots, k_{nн}$.



Фиг. 4. Структура нелинейного УПТ

Для упрощения анализа схемы на фиг. 4 предполагаем, что режим ограничения наступает одновременно с подачей входного сигнала. Это позволяет пренебречь участком I (фиг. I) и дает следующую методику расчета. В момент времени $t=0$ к схеме на фиг. 4 подключается напряжение

$$U_{вх}(p) = \frac{U_{вх}}{p}; \qquad (2)$$

входные напряжения каскадов, попадающих в ограничение, принимаются равными выражениям

$$U_i(p) = \frac{U_{izp}}{p}, \qquad (3)$$

а коэффициент передачи соответствующих нелинейных звеньев равными нулю

$$k_{iн} = 0. \qquad (4)$$

В формулах (3) и (4):

i – номера каскадов, попадающих в насыщение,

U_{izp} – граничное значение входного напряжения i-го каскада (максимум или минимум).

Затем определяются оригиналы выражений для входных напряжений нелинейных звеньев $L^{-1}[U_{iн}(p)]$. Длительность работы m-го каскада в режиме ограничения определяется исходя из неравенства

$$\left|L^{-1}[U_{mн}(p)]\right| > \left|U_{mzp}\right|. \qquad (5)$$

После выхода из ограничения m-го каскада ($L^{-1}[U_{mн}(p)]$ становится меньше своего граничного значения) дальнейший ход переходного процесса определяется при помощи видоизмененной схемы ($k_{mн}$ принимает значение 1) с учетом величин напряжений, установившихся в схеме к моменту выхода из ограничения m-го каскада.

Время работы УПТ в нелинейном режиме $t_н$ находится как время выхода из ограничения последнего (по времени) усилительного каскада. Начальные условия при расчете переходной характеристики на следующем, линейном участке можно учесть при помощи введения в схему на фиг. 2 дополнительных независимых источников напряжения. Соответствующий граф показан на фиг. 5.

Если i-ый усилительный каскад является инерционным или упругим звеном первого порядка [x]

$$k_i(p) = k_i \frac{1+pT_i^*}{1+pT_i}, \qquad (6)$$

x) Здесь и в дальнейшем применены обозначения $k_i(0) \equiv k_i$.

то напряжения $U_{i,0}(p)$ на фиг. 5 выражаются:

$$U_{i,0}(p) = u_i(t_н)\frac{T_{i-1}}{1+pT_{i-1}} - u_{i-1}(t_н)\frac{T^*_{i-1}}{1+pT_{i-1}},$$

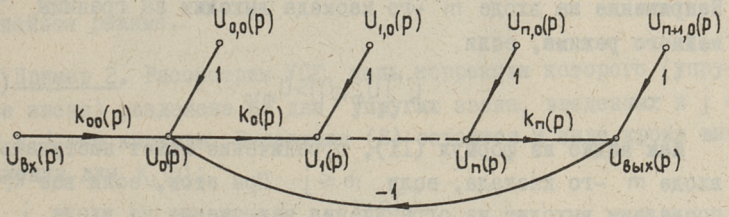
$$U_{0,0}(p) = u_0(t_н)\frac{T_{00}}{1+pT_{00}} - u_{вх}(t_н)\frac{T^*_{00}}{1+pT_{00}}, \qquad (7)$$

где T_i, T_i^* – постоянные времени i -го каскада,

$u_i(t_н)$ – напряжение $u_i(t)$ в конце нелинейного участка (при $t = t_н$).

Метод учета начальных условий при помощи схемы на фиг.5 аналогичен приведенному в [I] методу сведения расчета переходных процессов к нулевым начальным условиям.

Указанная методика дает хорошие результаты при практических расчетах, если время нахождения в ограничении определяется постоянными времени, большими по сравнению с "паразитными" постоянными времени каскадов. Такими обычно являются постоянные времени цепей коррекции при больших коэффициентах усиления по разомкнутой цепи.



Фиг. 5. К учету начальных условий

Пример I. Рассмотрим УПТ, цепь коррекции которого состоит из одного упругого звена, введенного в j -ый каскад:

$$k_{00}(p) = k_{00},$$

$$k_i(p) = k_i\frac{1}{1+pT_i} \quad \text{для } i = 0,\ldots,n;\ i \neq j,$$

$$k_j(p) = k_j \frac{1 + pT^*_{k1}}{(1+pT_{k1})(1+pT_j)}, \qquad (8)$$

где T_{k1}, T^*_{k1} – постоянные времени цепи коррекции,
T_i – "паразитные" постоянные времени.

Выражение для напряжения на входе m -го каскада, обусловленного выходным сигналом и сигналом ограничения на входе l -го каскада ($l \geq m$), в данном случае имеет вид:

$$U_m(p) = \frac{U_{вх}}{p} k_{oo} \prod_{i=0}^{m-1} k_i(p) - \frac{U_{i2p}}{p} \prod_{i=l+1}^{n} k_i(p) \prod_{i=0}^{m-1} k_i(p). \qquad (9)$$

Учитывая, что

$$T_{k1} >> T_i ; \quad T_{k1} >> T^*_{k1}, \qquad (10)$$

получаем из выражения (9):

$$L^{-1}[U_m(p)] \approx \begin{cases} \left[k_{oo} \cdot U_{вх} - U_{i2p} \prod_{i=m}^{n} k_i\right] \left(1 - e^{-\frac{t}{T_{k1}}}\right) \prod_{i=0}^{m-1} k_i & \text{при } m > j \\ \left[k_{oo} U_{вх} - U_{i2p} \left(1 - e^{-\frac{t}{T_{k1}}}\right) \prod_{i=m}^{n} k_i\right] \prod_{i=0}^{m-1} k_i & \text{при } m \leq j. \end{cases} \qquad (11)$$

Напряжение на входе m -го каскада выходит за границы линейного режима, если

$$L^{-1}[U_m(p)] > U_{m2p}. \qquad (12)$$

Как видно из формул (II), ограничение может наступать на входе m -го каскада, если $m \leq j$. При этом, если все $k_i > 1$, последним выходит из ограничения напряжение на входе j -го каскада. Время нахождения в ограничении $t_н$ определяется из уравнения

$$\left[k_{oo} U_{вх} - U_{j2p} \left(1 - e^{-\frac{t_н}{T_{k1}}}\right) \prod_{i=j}^{n} k_i\right] \prod_{i=0}^{j-1} k_i = U_{j2p} \qquad (13)$$

в виде

$$t_н \approx T_{k1} \frac{k_{oo} U_{вх} \prod_{i=0}^{j-1} k_i - U_{j2p}}{U_{j2p} \prod_{i=0}^{n} k_i}. \qquad (14)$$

При выводе формулы (I4) использовалось приближенное равенство, справедливое при $x << 1$:

$$\ln(1+x) \approx x. \qquad (15)$$

В нелинейном режиме выходное напряжение определяется выражением:

$$U_{вых}(t) \approx U_{j\,гр}\left(1-e^{-\frac{t}{T_{k1}}}\right)\prod_{i=j}^{n} k_i. \qquad (16)$$

После выхода из ограничения дальнейший ход переходного процесса определяется исходя из схемы на фиг. 5. Учитывая неравенства (10), получаем для следующего, линейного участка переходного процесса (отсчет времени начинается с конца нелинейного участка):

$$U_{вых}(p) = U_{вх}(p)\,k_{00}\frac{k(p)}{1+k(p)} + \frac{T_{k1}\,U_{j+1}(t_н)\prod_{i=j+1}^{n} k_i(p)}{(1+pT_{k1})\left[1+k(p)\right]}. \qquad (17)$$

Ход переходного процесса будет определяться корнями характеристического уравнения УПТ

$$1+k(p)=0, \qquad (18)$$

так как множитель $(1+pT_{k1})$ в знаменателе выражения (17) сокращается.

Как показывает анализ выражения (17), весь дальнейший ход переходного процесса в данном случае будет происходить в линейном режиме.

Пример 2. Рассмотрим УПТ, цепь коррекции которого (упругое звено) разделена на два упругих звена, введенных в j-ый и l-ый каскады. Выражения (8) остаются в силе, кроме выражения для $k_l(p)$:

$$k_l(p) = k_l\frac{1+pT^*_{k2}}{(1+pT_{k2})(1+pT_l)}. \qquad (19)$$

Постоянные времени выбраны так, что

$$T_{k1} >> T^*_{k1} = T_{k2} >> T^*_{k2}. \qquad (20)$$

Возможны следующие случаи: 1) $l<j$, 2) $l>j$. Как показывает анализ, при $l<j$ (случай 1) последним выходит из ограничения напряжение на входе j-го каскада. При $l>j$ последним может выходить из ограничения или напряжение на входе j-го (случай 2а) или l-го (случай 2б) каскада. Формулы для расчета времени $t_н$ в случаях 1 и 2а совпадают с приведенными выше формулами (13), (14), а поведение УПТ на

последующем линейном участке определяется корнями характеристического уравнения (I8).

В случае 2б время нахождения в ограничении определяется из уравнения

$$U_{вх}k_{оо}\prod_{i=0}^{l-1}k_i - U_{вх}k_{оо}\left(1-\frac{T_{к1}^*}{T_{к1}}\right)e^{-\frac{t_н}{T_{к1}}}\prod_{i=0}^{l-1}k_i -$$

$$-U_{lгр}\prod_{i=0}^{n}k_i + U_{lгр}e^{-\frac{t_н}{T_{к1}}}\prod_{i=0}^{n}k_i = U_{lгр}. \qquad (2I)$$

Используя равенство (I5), после некоторых упрощений получаем из уравнения (2I):

$$t_н \approx T_{к1}^* \frac{U_{вх}k_{оо}}{U_{lгр}\prod_{i=l}^{n}k_i - U_{вх}k_{оо}}. \qquad (22)$$

Как показывает анализ схемы на фиг. 5, в выходном напряжении УПТ на линейном участке имеется компонент с постоянной времени $T_{к2}$:

$$U_{вых}^*(t) = ce^{-\frac{t}{T_{к2}}}, \qquad (23)$$

где

$$c \approx U_{вых}(t_н) - U_{вх}k_{оо},$$

$$U_{вых}(t_н) \approx U_{lгр}\left(1-e^{-\frac{t_н}{T_{к2}}}\right)\prod_{i=l}^{n}k_i. \qquad (24)$$

Из-за наличия экспоненты (23) в выходном напряжении время установления в линейном режиме $t_л$ (см. формулу (I)) может сильно увеличиваться. Поэтому расстановку цепей коррекции, при которой получается случай 2б, следует считать неправильной. Для правильной расстановки следует добиться выполнения неравенства

$$U_{jгр}\frac{T_{к1}^*}{T_{к1}}\prod_{i=j}^{l-1}k_i < U_{lгр}, \qquad (25)$$

при котором получаем случай 2а.

Экспериментальная часть. Для операционного УПТ с параметрами $k_{оо} \approx I,2$; $k_о \approx 0,5$; $k_1 \approx I5$; $k_2 \approx k_3 \approx k_4 \approx I2$; $k_5 \approx 20$;

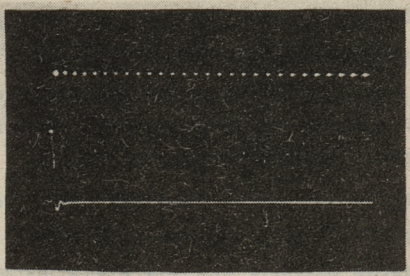
$T_{к1} \approx 45$ мксек; $T^{*}_{к1} \approx T_{к2} \approx 10$ мксек; $U_{отр} \approx 1$в; $U_{5гр} \approx 0{,}7$в получаем при $U_{вх} \approx 4$в и $j = 1$, $l = 4$ случай 2а. Используя формулу (14), получаем для данного случая:

$$t_н \approx 0{,}25 \text{ мксек}.$$

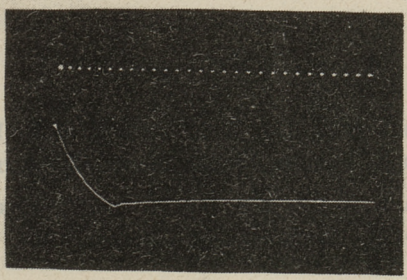
Для этого же усилителя при $j = 1$, $l = 5$ получаем случай 2б. Расчет по формулам (22)...(24) дает в данном случае:

$$t_н \approx 5 \text{ мксек},$$
$$U_{вых}(t_н) \approx 5{,}6 \text{ в},$$
$$U^{*}_{вых}(t) \approx 0{,}8\, e^{-\frac{t}{T_{к2}}},$$

т.е. на линейном участке имеется экспонента величиной в 0,8 в, имеющая постоянную времени 10 мксек.



а



б

Фиг. 6. Осциллограммы переходного процесса:
а) правильная коррекция, б) неправильная коррекция

На фиг. 6а и 6б показаны осциллограммы входного (верхние кривые с отметками времени в I мксек) и выходного (нижние кривые без отметок времени) напряжения для УПТ с вышеприведенными параметрами. Осциллограммы на фиг. 6а получены для случая $j = I$, $l = 4$, осциллограммы на фиг. 6б - для случая $j = I$, $l = 5$.

Сравнение расчетных (см. выше) и определенных с осциллограмм (фиг. 6а и 6б) параметров показывает хорошее совпадение.

Заключение. При достаточно больших скачкообразных входных сигналах переходный процесс в УПТ может иметь нелинейный участок, обусловленный ограничениями в каскадах усиления. На длительность нелинейного, а также и последующего линейного участков большое влияние оказывает расстановка цепей коррекции по каналу усиления.

Приводится методика расчета УПТ с обратной связью в нелинейном режиме его работы. Получены расчетные формулы для определения длительности нелинейного участка и критерии для правильной расстановки корректирующих цепей.

Следует отметить, что из-за введения цепей коррекции в первые каскады усиления уровень шумов на выходе УПТ будет определяться главным образом уровнем шумов последних (ближайших к выходу) усилительных каскадов.

Литература

I. Г.В. Зевеке, П.А. Ионкин, А.В. Нетушил, С.В. Страхов. Основы теории цепей. Энергия, Москва 1965.

2. Под ред. А.В. Нетушила. Теория автоматического управления. Часть I, Высшая школа, Москва 1968.

A. Laansoo

BERECHNUNG DER ZEIT DES ÜBERGANGSPROZESSES IN RÜCKGEKOPPELTEN GLEICHSTROMVERSTÄRKERN IN NICHTLINEAREM REGIME

Zusammenfassung

Bei großen sprungartigen Eingangssignalen gerät ein Teil der Verstärkerstufen zu Beginn des Übergangsprozesses in Besättigung. Die Dauer der Arbeit im nichtlinearen Regime (in Besättigung) ist von der Stelle der Einführung des Korrekturnetzwerkes (die ersten, die letzten Stufen) abhängig. Es werden die Berechnungsmethode und die Formeln gegeben, die die Dauer der Arbeit im nichtlinearen Regime zu berechnen und die richtigen Stellen der Einführung des korrekturnetzwerkes zu bestimmen erlauben.

Eesti NSV Teaduslik Raamatukogu IV Tehnika Akadeemia

ТРУДЫ ПО ЭЛЕКТРОТЕХНИКЕ
И
АВТОМАТИКЕ

С б о р н и к с т а т е й
VII

Таллинский политехнический институт

Редактор Г. В я л ь я м я э
Технический редактор Л. Л о о п е р

Сборник утвержден коллегией Трудов ТПИ 27 мая 1969 года.

Сдано в набор 9 окт. 1969 года. Подписано к печати 16 апр. 1970 года. Бумага 60x90/16. Печ. л. 9 + приложение. Уч.-изд. л. 6,81. Тираж 400. МВ-04333. Зак. №219. Ротапринт ТПИ, Таллин, ул. Коскла, 2/9. Цена 69 коп.



EESTI AKADEEMILINE RAAMATUKOGU
1 0200 00066741 4

Цена 69 коп.