316.920
3
1974

VOL. / TOM 3 * NUMBER / НОМЕР 1

ACADEMY OF SCIENCES OF THE USSR
HUNGARIAN ACADEMY OF SCIENCES

PROBLEMS OF CONTROL AND INFORMATION THEORY

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

АКАДЕМИЯ НАУК СССР
АКАДЕМИЯ НАУК ВЕНГРИИ

1974

AKADÉMIAI KIADÓ, BUDAPEST

PROBLEMS OF CONTROL AND INFORMATION THEORY

is an international quarterly sponsored jointly by the Presidium of the Academy of Sciences of the USSR and of the Hungarian Academy of Sciences. It offers publicity for original papers and short communications of the following topics:

— general theory of control systems and system theory

— theory of automata

— information theory

— operation research; theory of complex systems

— theory of economic control; system modelling

— theory and methods of adaptation, learning, identification and pattern recognition

— methods of information processing; application of digital computers in control and communication systems

— new physical principles in constructing technical devices for automation, control and information processing

The four issues published per year make up a volume of some 320 pp.

While this quarterly is mainly a publication forum of the research results achieved within the USSR and Hungary, also papers of international interest from other countries are welcome.

Distributor:

KULTURA

Hungarian Trading Co. for Books and Newspapers

H-1389 Budapest 62, P.O. Box 149, Hungary

The quarterly is published by

AKADÉMIAI KIADÓ

Publishing House of the Hungarian Academy of Sciences

H-1363 Budapest, P.O. Box 24, Hungary

Президиумами Академии Наук СССР и Академии Наук ВНР было принято решение о совместном издании журнала

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

Журнал создан для обеспечения быстрейшей публикации материалов о новейших результатах научных исследований, разработок и кратких сообщений о достижениях в следующих областях науки:

— общая теория процессов управления и теория систем
— теория автоматов
— теория информации
— исследование операций; теория сложных систем
— теория управления экономическими системами; модели систем
— теория и методы адаптации, обучения, идентификации и распознавания образов
— методы обработки информации; применение вычислительных машин в системах управления и передачи информации
— новые физические принципы создания технических средств автоматики, управления и передачи информации

Журнал издается четыре раза в год общим объёмом около 320 печатных страниц.

Журнал в первую очередь будет публиковать научные достижения обеих стран, а также статьи из других стран, представляющие международный интерес.

Распространитель:

KULTURA

Венгерское Общество по распространению книг и журналов

H-1389 Будапешт 62, п. о. 149, Венгрия

Издатель журнала:

AKADÉMIAI KIADÓ

Издательство Академии Наук Венгрии

H-1363 Будапешт, п. о. 24, Венгрия

PROBLEMS OF CONTROL AND INFORMATION THEORY

Volume 3, 1974

PAPERS

Beck, J.: On the entropy of the sum of IID random variables. **3,** 2, pp. 167—170 167
Belavkin, V. P.: Optimal linear randomized filtration of quantum boson signals. **3,** 1, pp. 47—62 47
Bliznakov, J.: Main principles of the formation and management of organizational-productive systems. **3,** 3, pp. 227—230 227
Csáki, F.: Transformation methods from phase-variable form to canonical forms. **3,** 3, pp. 209—225 209
Goppa, V. D.: Universal algebraic codes for symmetric channels. **3,** 4, pp. 241—246 241
Győrfi, L.: On some estimation problems in pattern recognition. **3,** 1, pp. 11—17.. 11
Haber, R.—Keviczky, L.: Adaptive dual extremum control by finite order Volterra model. **3,** 4, pp. 247—260 247
Kartsatos, A. G.: Global controllability of perturbed quasi-linear systems. **3,** 2, pp. 137—145 137
Khazen, E. M.: Risk increase informational estimates under observed data reduction in sequential analysis. **3,** 1, pp. 35—45 35
Kolmanovsky, V. B.: On the approximation of linear control systems with time lag. **3,** 1, pp. 63—76 63
Komáromi, É.: A finite primal method for the assignment problem. **3,** 2, 157—166 157
Krasovskii, N. N.: Game-theoretic control and problems of stability. **3,** 3, pp. 171—182 171
Kwiatek, B.—Zieba, R.: Homomorphisms of stochastic and nondeterminate automata. **3,** 4, pp. 277—284 277
Levin, V. I.: Calculation of the transient processes in the combinational finite automata. **3,** 4, pp. 261—276 261
Milenin, N. K.: Linear coding in the retransmission of the signal. **3,** 2, pp. 127—136 127
Nemetz, T.: Finite data reduction in the case of a long sequence of observations. **3,** 1, pp. 19—34 19
Nemetz, T.: On the α-divergence rate for Markov-dependent hypotheses. **3,** 2, pp. 147—155 147
Oganesyan. S. Sh.—Tairyan, V. I.—Yagdzyan, V. G.: Decomposition of cyclic codes into equal-weight classes. **3,** 2, pp. 117—125 117
Pupkov, K. A.—Kapalin, V. I.: The use of Volterra series in nonlinear system designing. **3,** 2, pp. 105—115 105
Rizhikov, Yu. I.: Analysis of mass service system with switching on threshold. **3,** 1 pp. 85—97 85
Siforov, V. I.: The development of science in the Soviet Union. **3,** 1, pp. 7—9.... 7
Tarján, T. G.: On complexity of switching circuits. **3,** 3, pp. 183—195.... 183
Tolmacheva, A. Yu.: Choice of base elements in structural synthesis algorithms. **3,** 1, pp. 77—83 77
Voronov, A. A.—Skaletskii, V. V.—Chistyakov, Yu. V.: Determination of the input distribution in a one-channel queue system. **3,** 3, pp. 197—207 197

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

AUTHORS

Beck, J. **3**, 2, pp. 167—170
Belavkin, V. P. **3**, 1, pp. 47—62
Bliznakov, J. **3**, 3, pp. 227—230
Chistyakov, Yu. V. **3**, 3. pp. 197—207
Csáki, F. **3**, 3, pp. 209—225
Goppa, V. D. **3**, 4, pp. 241—246
Győrfi, L. **3**, 1, pp. 11—17
Haber, R. **3**, 4, pp. 247—260
Kapalin, V. I. **3**, 2, pp. 105—115
Kartsatos, A. G. **3**, 2, pp. 137—145
Keviczky, L. **3**, 4, pp. 247—260
Khazen, E. M. **3**, 1, pp. 35—45
Kolmanovsky, V. B. **3**, 1, pp. 63—76
Komáromi, É. **3**, 2, 157—166
Krasovskii, N. N. **3**, 3, pp. 171—182
Kwiatek, B. **3**, 4, pp. 277—284
Levin, V. I. **3**, 4, pp. 261—276
Milenin, N. K. **3**, 2, pp. 127—136
Nemetz, T. **3**, 1, pp. 19—34; **3**, 2, pp. 147—155
Oganesyan, S. Sh. **3**, 2, pp. 117—125
Pupkov, K. A. **3**, 2, pp. 105—115
Rizhikov, Yu. I. **3**, 1, pp. 85—07
Siforov, V. I. **3**, 1, pp. 7—9
Skaletskii, V. V. **3**, 3, pp. 197—207
Tairyan, V. I. **3**, 2, pp. 117—125
Tarján, T. G. **3**, 3, pp. 183—195
Tolmacheva, A. Yu. **3**, 1, pp. 77—83
Voronov, A. A. **3**, 3, pp. 197—207
Yagdzyan, V. G. **3**, 2, pp. 117—125
Zieba, R. **3**, 4, pp. 277—284

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

Том 3, 1974

СТАТЬИ

Бек Й: Об энтропии одинаково распределенных независимых случайных величин. **3,** 2, стр. 167—170 167

Белавкин В. П. Оптимальная линейная рандомизированная фильтрация квантовых бозонных сигналов. **3,** 1, стр. 47—62 47

Близнаков Й. Основные принципы формирования и управления организационно-производственными системами. **3,** 3, стр. 227—230 227

Воронов А. А., Скалецкий В. В., Чистяков Ю. В. Определение функции распределения входящего потока в однолинейной системе массового обслуживания. **3,** 3, стр. 197—207 197

Гоппа В. Д. Универсальные алгебраические коды для симметричных каналов. **3,** 4, стр. 241—246 241

Дерфи Л. Некоторые проблемы оценок в теории распознавания образов. **3,** 1, стр. 11—17 11

Картсатос А. Г. Управляемость в целом квазилинейных систем в условиях возмущений. **3,** 2, стр. 137—145 137

Квиатек Б., Зиеба Р. Гомоморфизм стохастических и недерминированных автоматов. **3,** 4, стр. 277—284 277

Колмановский В. Б. Об аппроксимации линейных управляемых систем с последействием. **3,** 1, стр. 63—76 63

Комароми Е. Конечный примальный метод для решения транспортной задачи. **3,** 2, стр. 157—166 157

Красовский Н. Н. Игровое управление и задачи устойчивости. **3,** 3, стр. 171—182 ....

Левин В. И. Вычисление переходных процессов в комбинационных конечных автоматах. **3,** 4, стр. 261—276 261

Миленин Н. К. Линейное кодирование при ретрансляции сигнала. **3,** 2, стр. 127—136 127

Немец Т. Коррекция конечного числа данных в случае продолжительной последовательности наблюдений. **3,** 1, стр. 19—34 19

Немец Т. Об асимптотическом исследовании α-расходимости мер, соответствующих двум однородным марковским процессам. **3,** 2, стр. 147—155 147

Оганесян С. Ш., Таирян В. И., Ягджян В. Г. Разбиение циклических кодов на равновесные классы. **3,** 2, стр. 117—125 117

Пупков К. А., Капалин В. И. Применение рядов Вольтерра для исследования нелинейных систем. **3,** 2, стр. 105—115 105

Рыжиков Ю. И. Расчет системы массового обслуживания с порогом включения. **3,** 1, стр. 85—97 85

Сифоров В. И. Развитие науки в СССР (к 250-летию Академии наук СССР). **3,** 1, стр. 3—5 3

Тарян Т. Г. О сложности релейно-контактных схем. **3,** 3, стр. 183—195 183

Толмачева А. Ю. Выбор базовых элементов в алгоритмах структурного синтеза. **3,** 1, стр. 77—83 77

Хабер Р., Кевицки Л. Адаптивное дуальное экстремальное управление при помощи конечной модели Вольтерра. **3,** 4, стр. 247—260 247

Хазен Э. М. Информационные оценки для возрастания риска при редукции наблюдаемых данных в задачах последовательного анализа. **3,** 1, стр. 35—45 35

Чаки Ф. Методы преобразования уравнений состояния к каноническому виду. **3,** 3, стр. 209—225 209

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

АВТОРЫ

Бек Й. **3,** 2, стр. 167—170
Белавкин В. П. **3,** 1, стр. 47—62
Близнаков Й. **3,** 3, стр. 227—230
Воронов А. А. **3,** 3, стр. 197—207
Гоппа В. Д. **3,** 4, стр. 241—246
Дерфи Л. **3,** 1, стр. 11—17
Зиеба Р. **3,** 4, стр. 277—284
Капалин В. И. **3,** 2, стр. 105—115
Картсатос А. Г. **3,** 2, стр. 137—145
Квиатек Б. **3,** 4, стр. 277—284
Кевицки Л. **3,** 4, стр. 247—260
Колмановский В. Б. **3,** 1, стр. 63—76
Комароми Е. **3,** 3, стр. 157—166
Красосвкий Н. Н. **3,** 3. стр. 171—182
Левин В. И. **3,** 4, стр. 261—276
Миленин Н. К. **3,** 2, 127—136
Немец Т. **3,** 1, стр. 19—34; **3,** 2, стр. 147—155
Оганесян С. Ш. **3,** 2, стр. 117—125
Пупков К. А. **3,** 2, стр. 105—115
Рыжиков Ю. И., **3,** 1, стр. 85—97
Сифоров В. И. **3,** 1, стр. 3—5
Скалецкий В. В. **3,** 3, стр. 197—207
Таирян В. И. **3,** 2, стр. 117—125
Тарян Т. Г. **3,** 3, стр. 183—195
Толмачева А. Ю. **3,** 1, стр. 77—83
Хабер Р. **3,** 4, стр. 247—260
Хазен Э. М. **3,** 1, стр. 35—45
Чаки Ф. **3**[7] 3, стр. 209—225
Чистяков Ю. В. **3,** 3, стр. 197—207
Ягджян В. Г. **3,** 2, стр. 117—125

PROBLEMS OF CONTROL AND INFORMATION THEORY

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

EDITORS

B. N. PETROV (Moscow)
F. CSÁKI (Budapest)

DEPUTY EDITORS

V. S. PUGACHEV (Moscow)
V. I. SIFOROV (Moscow)
S. CSIBI (Budapest)

CO-ORDINATING EDITORS

S. V. EMELIANOV (Moscow)
L. KALMÁR (Budapest)

M. A. GAVRILOV (Moscow)
I. CSISZÁR (Budapest)

A. M. LETOV (Moscow)
A. PRÉKOPA (Budapest)

L. VARGA (Budapest)

E. D. TERYAEV (Moscow)
J. KOCSIS (Budapest)

РЕДАКТОРЫ ЖУРНАЛА

Б. Н. ПЕТРОВ (Москва)
Ф. ЧАКИ (Будапешт)

ЗАМЕСТИТЕЛИ РЕДАКТОРОВ

В. С. ПУГАЧЕВ (Москва)
В. И. СИФОРОВ (Москва)
Ш. ЧИБИ (Будапешт)

ЧЛЕНЫ РЕДАКЦИОННОЙ КОЛЛЕГИИ

С. В. ЕМЕЛЬЯНОВ (Москва)
Л. КАЛМАР (Будапешт)

М. А. ГАВРИЛОВ (Москва)
И. ЧИСАР (Будапешт)

А. М. ЛЕТОВ (Москва)
А. ПРЕКОПА (Будапешт)

Л. ВАРГА (Будапешт)

Е. Д. ТЕРЯЕВ (Москва)
Я. КОЧИШ (Будапешт)

AKADÉMIAI KIADÓ

PUBLISHING HOUSE OF THE HUNGARIAN ACADEMY OF SCIENCES

BUDAPEST

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

Printed in Hungary

AKADÉMIAI NYOMDA, BUDAPEST

Problems of Control and Information Theory, Vol. 3 (1), pp. 3—5 (1974)

РАЗВИТИЕ НАУКИ В СССР

(к 250-летию Академии наук СССР)

В 1974 году Академии наук СССР исполняется 250 лет. За два с половиной столетия своей деятельности ученые Академия наук обогатили многие разделы науки. Имена таких ученых как М. В. Ломоносов, Н. И. Лобачевский, П. Л. Чебышев, Д. И. Менделеев, А. М. Бутлеров, К. А. Тимирязев, А. М. Ляпунов и многие другие широко известны во всем мире.

Она сыграла большую роль в изучении и освоении природных богатств СССР, в развитии промышленности, сельского хозяйства, здравоохранения, образования и культуры.

На современном этапе исторического развития особенно важное значение приобрели фундаментальные исследования, создающие предпосылки прогресса всей науки, техники, производства.

В настоящее время в Академии наук СССР имеется 250 хорошо оснащенных научных учреждений. Организованы и ведут плодотворную работу академии наук во всех союзных республиках, Сибирское отделение АН СССР, Уральский и Дальневосточный научные центры.

В наши дни в Советском Союзе насчитывается более одного миллиона научных работников. Достижения в области изучения космоса, развития космонавтики, ядерной физики, математики, теории информации и управления, квантовой электроники, наук о земле и во многих других областях естествознания и обществоведения выдвинули советских ученых на передовые рубежи мировой науки.

Большое внимание в Советском Союзе уделяется развитию фундаментальных исследований в области теории управления и теории информации, а также применению результатов этих исследований и, в частности, для создания и совершенствования Автоматизированных систем управления различных уровней, Автоматизированных систем сбора и обработки информации, Единой автоматизированной системы связи страны, Государственной сети вычислительных центров и многих других.

Созданная академиком В. А. Котельниковым теория потенциальной устойчивости и развитая академиком А. Н. Колмогоровым теория фильтрации и экстраполяции случайных процессов вместе с работами американских ученых К. Шеннона и Н. Винера привели к созданию теории информации и кибернетики.

Важное значение имеют работы академиков Б. Н. Петрова, В. А. Трапезникова и других по теории автоматического управления.

Советские ученые имеют большие заслуги в создании современной теории оптимального управления, а также в решении ряда важных конкретных задач управления (академик Л. С. Понтрягин и его школа, академик Н. Н. Красовский и др.).

В последние годы в Советском Союзе получили существенное развитие исследования в области систем управления с переменной структурой, в теории терминального управления, в теории координатно-параметрического управления, в методике решения многокритериальных задач. Дальнейшее развитие получили также классические направления в теории автоматического управления, в частности, теория оптимального управления, теория адаптивных систем, теория беспоисковых самонастраивающихся систем, теория инвариантности, теория многосвязных систем, статистические проблемы технической кибернетики.

Была развита теория адаптивных и обучающихся систем, позволившая систематизировать и обобщить некоторые из полученных ранее результатов в области идентификации, распознавания и фильтрации и указать пути регулярного получения новых алгоритмов решения задач, связанных с принятием решения в условиях неопределенности. Получили также развитие работы в области автоматизации проектирования различных технических объектов с использованием электронных вычислительных машин.

Большое внимание было уделено в последние годы фундаментальным исследованиям в области передачи, распределения и обработки информации с физических, технических и биологических системах. Были проведены исследования, направленные на разрешение коренных проблем техники связи — повышения эффективности и надежности, а также разнообразных проблем, связанных с построением Единой автоматизированной сети связи. К числу таких исследований относятся: поиски эффективных видов кодирования и процедур декодирования, обеспечивающих высокую надежность передаваемой информации; изучение количественных соотношений, связывающих скорость передачи информации, вероятность появления ошибок, уровень помех с видом используемого кода и алгоритмами декодирования; исследование информационных свойств различных каналов и, в частности, каналов с обратной связью; асимптотическое изучение свойств сложных сетей;

исследование различных систем распределения информации и, в частности, систем, учитывающих приоритеты; изучение различных систем коммутации в сетях связи и многие другие.

Во всех этих направлениях получены интересные результаты, имеющие большое теоретическое и прикладное значение.

Решение ряда актуальных проблем современной науки требует совместных усилий ученых многих стран. Результаты их исследований, обогащая друг друга, позволяют науке быстрее продвигаться вперед. Поэтому все более широкий размах приобретает международное научное сотрудничество, в котором ученые Академии наук СССР принимают активное участие.

В. И. Сифоров

член-корреспондент АН СССР

Problems of Control and Information Theory, Vol. 3 (1), pp. 7—9 (1974)

THE DEVELOPMENT OF SCIENCE IN THE SOVIET UNION

(The 250th anniversary of the foundation of the Academy of Sciences of the USSR)

In 1974 the Academy of Sciences of the USSR celebrates the 250th anniversary of its existence. During the past two and a half centuries the scientists of the Academy have enriched several branches of science. The name of M. V. Lomonosov, N. I. Lobatchevsky, P. L. Chebyshev, D. I. Mendeleev, A. M. Butlerov, K. A. Timiryazev, A. M. Lyapunov and of many other scientists has become known all over the world.

The Academy of Sciences has played an important role in the discovery of the natural resources of the country, in the development of industry, agriculture, public health, education and culture.

At the present state of historic development fundamental research creating the preconditions of the advancement of science, technology and production has obtained great significance.

At present the Academy of Sciences of the USSR has 250 well-equipped research institutes. In every federal republic of the Soviet Union an academy of science has been organized and is conducting successful research; branches of the Academy are working in many autonom republics. The Siberian, Uralian and Far East Departments of the Academy of Science have been established.

Today more than one million research scientists are working in the Soviet Union. With their results in the fields of space research, aeronautics, atomic physics, mathematics, information and control theory, quantum electronics, earth science as well as several other branches of natural and social science Soviet scientists have reached leading ranks in world science.

In the Soviet Union great attention is being paid to the fundamental research in control and information theory, to the usage of the results of this research among others in the creation and development of automated control systems of different levels, automated information collecting and processing systems, the uniform, automated telecommunication system of the country and the national network of computer centers etc.

The theory of potentional stationarity founded by academician V. A. Kotelnikov and the theory of filtration and extrapolation of stochastic proces-

ses developed by academician A. N. Kolmogorov together with the works of the American scientists C. E. Shannon and N. Wiener led to the foundation of information theory and cybernetics.

The work of academicians B. N. Petrov and V. A. Trapeznikov and others in the theory of automated control has great importance.

Soviet scientists have great merits in the foundation of the modern theory of optimal control and in the solution of a series of important, concrete control problems (academician L. S. Pontryagin and his school, academician N. N. Krasovsky and others).

In the past years significant development could be experienced in the Soviet Union in the fields of control systems with changing structure, the theory of terminal control, the theory of coordinate-parametric control, the method of solving multi-criterion tasks. The classical direction of the theory of automatic control has been advanced, specifically the theory of optimal control, the theory of adaptive systems, the theory of disturbed self-tuning systems, the theory of invariance, the theory of multi-connected systems and the statistical problems of technical cybernetics.

The theory of adaptive and self-teaching systems has been developed, with the help of which it becomes possible to systematize and generalize results obtained earlier in the fields of identification, recognition and filtration, to systematically develop new algorithms necessary for the solution of tasks connected with decision making among undetermined conditions. The field of automatization of technical design with the help of electronic computers has advanced.

Great attention has been given in the past years to fundamental research in the field of information transfer, spreading and processing in physical, technical and biological systems. Research has been directed in the solution of the basic problems of telecommunication — to increase efficiency and reliability, and many other problems connected with the establishment of a unified automated telecommunication network. The following topics belong to this research:

the search of effective coding and decoding procedures which assure a high degree of reliability of the information transmitted;

the study of the quantitative connections connecting the speed of information transfer, error probability, noise level with the applied code and the decoding algorithm;

research into the information characteristics of channels, among them channels having a feedback connection;

the asymptotic examination of the characteristics of complex networks;

the study of information spreading systems, among them systems considering priorities;

the investigation of switching methods in communication systems, etc.

In all these areas interesting results were obtained, having great theoretical and practical significance.

The solution of the problems of modern science requires the joint effort of the scientists of many countries. The results of their research enrich the science of the individual countries and they also make a faster advancement of science possible. Just because of this, international scientific cooperation is becoming more and more widespread and the scientists of the Academy of Sciences of the USSR play an important role in it.

V. I. Siforov

Corrseponding Member of the Academy of Sciences of the USSR

Problems of Control and Information Theory, Vol. 3 (1), pp. 11–17 (1974)

ON SOME ESTIMATION PROBLEMS IN PATTERN RECOGNITION

L. GYŐRFI

(Budapest)

(Received October 9, 1973)

Two estimations are investigated for which the results of the algorithms are very useful in pattern recognition. The targets of estimations are the L_2 norm squares of a probability density and an optimal Bayes decision function. This paper presents also two examples for the application of these quantities in pattern recognition.

Notations

Let $(\Omega, \mathcal{A}, P)$ be a probability space, ξ a random variable taking the values in the N dimensional Euclidean space, H_+, H_- two disjoint events (alternative hypotheses) for which $H_+ \cup H_- = \Omega$. The decision problem is to decide for the observation of ξ which of the hypotheses holds.

Let us assume that the distribution of ξ (denoting by Q) is absolutely continuous and f denotes the probability density of ξ. Put

$$D(\xi) = P(H_+/\xi) - P(H_-/\xi) \tag{1}$$

and

$$d(\xi) = f(\xi) \cdot D(\xi). \tag{2}$$

For knowing the functions d or D it can be realized the Bayes decision minimizing the probability of misdecision. If $D(\xi) \geq 0$ (or $d(\xi) \geq 0$) then the Bayes decision is H_+ and H_- elsewhere. We show algorithms to estimate the $\| d \|^2 = \int d^2(x)dx$ and $\| f \|^2 = \int f^2(x)dx$.

The algorithms

Let us assume that it is given a sequence of independent and identically distributed random variables $(x_1, \varrho_1), (x_2, \varrho_2) \ldots$ such that x_n have the common probability density f and ϱ_n take the values $+1$ or -1 and

$$E(\varrho_n/x_n) = D(x_n) \qquad (n = 1, 2 \ldots). \tag{3}$$

Let us adopt

$$\xi_n = \frac{\sum_{k=1}^{n} \varrho_{2k-1} K_n(x_{2k-1}, x_{2k}) \varrho_{2k}}{n} \tag{4}$$

for the estimation of $\| d \|^2$ and estimate $\| f \|^2$ by

$$\eta_n = \frac{\sum_{k=1}^{n} K_n(x_{2k-1}, x_{2k})}{n} . \tag{5}$$

Here

$$K_n(x, y) = \frac{1}{a_n^N} h\left(\frac{x-y}{a_n}\right) \tag{6}$$

where

$$\int h(z)dz = 1 \tag{7}$$

and $\{a_n\}_{n=1}^{\infty}$ is a sequence of positive numbers.

Theorem 1. For definitions (4) and (5)

$$E(\xi_n) = \iint h(z)d(x)d(x - a_n z)dxdz \tag{8}$$

and

$$E(\eta_n) = \iint h(z)f(x)f(x - a_n z)dxdz. \tag{9}$$

Proof. Because of the independence of $(x_1, \varrho_1), (x_2, \varrho_2) \ldots$

$$\begin{aligned} E(\xi_n) &= E(\varrho_1 K_n(x_1, x_2)\varrho_2) = E(E(\varrho_1 K_n(x_1, x_2)\varrho_2/x_1, x_2)) = \\ &= E(K_n(x_1, x_2)\, E(\varrho_1\, \varrho_2/x_1, x_2)) = E(K_n(x_1, x_2)\, E(\varrho_1/x_1)\, E(\varrho_2/x_2)) = \\ &= E(K_n(x_1, x_2)\, D(x_1)\, D(x_2)) \end{aligned} \tag{10}$$

and

$$E(\eta_n) = E(K_n(x_1, x_2)). \tag{11}$$

(10), (11) and (2) imply that

$$E(\xi_n) = \iint K_n(x, y)\, d(x)\, d(y)\, dx\, dy \tag{12}$$

and

$$E(\eta_n) = \iint K_n(x, y)\, f(x)\, f(y)\, dx\, dy\,. \tag{13}$$

Using definition (6) and the notation $z = \dfrac{x-y}{a_n}$ from (12) and (13) we have (8) and (9).

Corollary. Theorem 1 implies that the sequences $E(\xi_n)$ and $E(\eta_n)$ are bounded and

$$E(\xi_n) \leq \| d \|^2 \int | h(z) | dz \tag{14}$$

and

$$E(\eta_n) \leq \| f \|^2 \int | h(z) | dz. \tag{15}$$

For each a_n and z

$$|\int d(x)\, d(x - a_n z)\, dx| \leq \sqrt{\int d^2(x)\, dx \cdot \int d^2(x - a_n z)\, dx} = ||d||^2 \tag{16}$$

and

$$|\int f(x)\, f(x - a_n z)\, dx| \leq \sqrt{\int f^2(x)\, dx \cdot \int f^2(x - a_n z)\, dx} = ||f||^2. \tag{17}$$

Then (16), (17), (8), (9) give the statement. If for each $z \in R^N$ $h(z) \geq 0$ (h is Parzen kernel) then because of (7), (14), (15) the estimates "bias down".

Theorem 2. If $\lim a_n = 0$ then estimations (5) and (4) are asymptotically unbiased

$$\lim_{n\to\infty} E(\xi_n) = ||d||^2 \tag{18}$$

and

$$\lim_{n\to\infty} E(\eta_n) = ||f||^2. \tag{19}$$

Proof. Using the result of Theorem 1 we have to prove that for each square integrable function g

$$\lim_{n\to\infty} \iint h(z)\, g(x)\, g(x - a_n z)\, dz\, dx = ||g||^2 \tag{20}$$

if $\lim\limits_{n\to\infty} a_n = 0$. Let us define the operators $U_n : L_2 \to L_2$ by the equation

$$g_n(x) = U_n g(x) = \int h(z) g(x - a_n z) dz. \tag{21}$$

Thus U_n is a sequence of bounded linear operators and

$$\begin{aligned}
&||U_n g||^2 = \int (\int h(z)\, g(x - a_n z) dz)^2\, dx = \\
&= \int (\iint h(u)\, h(v)\, g(x - a_n u)\, g(x - a_n v)\, du\, dv)\, dx \leq \\
&\leq ||g||^2 \iint |h(u)\, h(v)|\, du\, dv = ||g||^2 \left|\int |h(z)|\, dz\right|^2
\end{aligned}$$

so for each n

$$||U_n|| \leq \int |h(z)|\, dz. \tag{22}$$

We prove at first that for continuous $g \in L_2$

$$\lim_{n\to\infty} ||U_n g - g||^2 = 0$$

$$\begin{aligned}
&||U_n g - g||^2 = ||U_n g||^2 - 2\,(U_n g, g) + ||g||^2 = \\
&= \int (\iint h(u)\, h(v)\, g(x - a_n u)\, g(x - a_n v)\, du\, dv)\, dx - \\
&- 2 \iint h(z)\, g(x)\, g(x - a_n z)\, dz\, dx + ||g||^2 = \\
&= \iint h(u)\, h(v)\, (\int g(x - a_n u)\, g(x - a_n v)\, dx)\, du\, dv - \\
&- 2 \int h(z)\, (\int g(x)\, g(x - a_n z)\, dx)\, dz + ||g||^2.
\end{aligned} \tag{23}$$

If g is continuous then the functions $\int g(x - a_n u) g(x - a_n v) dx$ as the functions of (u, v) are at $(0, 0)$ continuous and have uniform bound $(\| g \|^2)$ and the functions $\int g(x)\, g(x - a_n z)\, dx$ as the functions of z are at 0 continuous and have also uniform bound $(\| g \|^2)$. So for (23) the Lebesgue integral theorem can be applied and

$$\lim_{n\to\infty} \| U_n g - g \|^2 = \iint h(u)\, h(v) \left(\lim_{n\to\infty} \int g(x - a_n u)\, g(x - a_n v)\, dx\right) du\, dv -$$

$$- 2 \int h(z) \left(\lim_{n\to\infty} \int g(x)\, g(x - a_n z)\, dx\right) dz + \| g \|^2 = 0. \tag{24}$$

Let us use the Banach–Steinhaus theorem for the sequence U_n. The continuous functions in L_2 are everywhere dense, for continuous square integrable g (24) holds and $\| U_n \|$ are bounded. So the Banach–Steinhaus theorem [8] implies that for each $g \in L_2$

$$\lim_{n\to\infty} \| U_n g - g \| = 0 \tag{25}$$

which gives (20).

Remark. If $\int d(x)\, d(x - z)\, dx$ has bounded second order derivatives in a neighbourhood of 0 then h can be chosen such that

$$| E(\xi_n) - \| d \|^2 | \leq C\, a_n^2 . \tag{26}$$

The proof of this statement is the same as in [1] for the investigation of probability density estimation.

Lemma. If h is square integrable then

$$\mathrm{Var}(\xi_n) \leq \frac{\| h \|^2 \| f \|^2}{n a_n^N} \tag{27}$$

and

$$\mathrm{Var}(\eta_n) \leq \frac{\| h \|^2 \| f \|^2}{n a_n^N} . \tag{28}$$

Proof. From the definition of ξ_n and η_n

$$\mathrm{Var}(\xi_n) \leq \frac{E\big((\varrho_1 K_n(x_1, x_2)\, \varrho_2)^2\big)}{n} = \frac{E\big(K_n^2(x_1, x_2)\big)}{n} \tag{29}$$

and

$$\mathrm{Var}(\eta_n) \leq \frac{E\big(K_n^2(x_1, x_2)\big)}{n} . \tag{30}$$

Since

$$E\big(K_n^2(x_1, x_2)\big) = \iint \frac{1}{a_n^{2N}} h^2\left(\frac{x-y}{a_n}\right) f(x) f(y)\, dx\, dy =$$

$$= \frac{1}{a_n^N} \int h^2(z) \left(\int f(x) f(x - a_n z)\, dx\right) dz \leq \tag{31}$$

$$\leq \frac{1}{a_n^N} \|f\|^2 \int h^2(z)\, dz = \frac{\|f\|^2 \|h\|^2}{a_n^N}.$$

Thus (29), (30) and (31) imply the statement of the lemma.

Remark. The kernel function $K_n(x, y)$ can be defined more generally, for example the kernel function in [2] can be used which is investigated for the estimation of probability density function.

Applications

1. In many actual pattern recognition problems one has to estimate decision function minimizing the probability of misclassification. For solving this problem it would be useful to know the best result, the Bayes error probability. But it is an unsolved problem how to estimate in a simple way the Bayes error probability using the sample (x_1, ϱ_1), $(x_2, \varrho_2)\ldots$

However there are many information type measures associated with the Bayes error. This application gives an example for the estimation of such measure in a special case.

Igor Vajda [6] presented a measure for finite or countable infinite number of hypotheses. For two hypotheses this measure is

$$h = \frac{1}{2}\big(1 - \int |D(x)|^2 Q(dx)\big). \tag{32}$$

The Bayes error probability

$$P_c(d) = \frac{1}{2}\big(1 - \int |D(x)|\, Q(dx)\big). \tag{33}$$

Vajda proved generally that

$$\frac{h}{1 + \sqrt{1-h}} \leq P_e(d) \leq h. \tag{34}$$

But for two hypotheses h is the asymptotic error probability of the nearest neighbour decision rule and Cover and Hart proved [3] that

$$\frac{1}{2}(1 - \sqrt{1 - 2h}) \leq P_e(d) \leq h. \tag{35}$$

The definition of h implies

$$h = \frac{1}{2}\left(1 - \int\limits_{\{f \neq 0\}} \frac{d^2(x)}{f(x)} dx\right). \tag{36}$$

Let us assume that Q is uniform distribution on an unknown Borel set A with finite Lebesgue measure i.e.

$$f(x) = \begin{cases} c & \text{if} \quad x \in A \\ 0 & \text{if} \quad x \notin A. \end{cases} \tag{37}$$

Thus

$$\int f^2(x) dx = c \tag{38}$$

and in this special case

$$h = \frac{1}{2}\left(1 - \frac{\|d\|^2}{\|f\|^2}\right). \tag{39}$$

In one dimensional case for known absolutely continuous distribution Q the $Q(\xi)$ has uniform distribution in $[0, 1]$ ($\|f\|^2 = 1$) and the Bayes errors using the observations ξ and $Q(\xi)$ are equal.

2. The second example is associated with the estimation of d. In [4] there is an algorithm tending to the projection d^* of d. In this case for the error probability using the decision function d^* (denoting by $P_e(d^*)$)

$$|P_e(d^*) - P_e(d)|^2 \leq \mu(A) \|d - d^*\|^2 \tag{40}$$

where A is a Borel set such that $Q(A) = 1$ and μ is the Lebesgue measure. Since d^* is the projection of d so

$$|P_e(d^*) = P_e(d)|^2 \leq \mu(A) (\|d\|^2 - \|d^*\|^2). \tag{41}$$

If we know $\mu(A)$ and $\|d^*\|^2$ then using the estimation (4) we have the upper bound (41).

References

1. *Rejtő, L., Révész, P.*, "Density estimation and pattern classification", Problems of Control and Information Theory, **2** (1973) 67–80.
2. *Földes, A., Révész, P.*, "A general method for density estimation" Submitted to the Annals of Statistics, 1973.
3. *Cover, T., Hart, P.*, "The nearest neighbour pattern classification", IEEE Transactions on Information Theory, **13** (1967).
4. *Győrfi, L.*, "On the estimation of probability density and optimal decision function in RKHS" Submitted to Proceedings of European Meeting of Statisticians, Budapest, 1972.

5. *Parzen, B.*, "On estimation of probability density function and mode", Ann. Math. Stat. **33** (1962).
6. *Вайда И.* Оценка минимальной вероятности ошибки при проверке конечного или счетного числа гипотез. Проблемы передачи информации (1968) 1.
7. *Halmos, P. R.*, "Measure Theory" Van Nostrand, New York, 1950.
8. *Kantorovich, L. V.*, *Akilov, G. P.*, "Functional Analysis in Normed Spaces" Pergamon Press, Macmillan, New York, 1964.

Некоторые проблемы оценок в теории распознавания образов

Л. ДЕРФИ

(Будапешт)

Предложены два рекуррентных алгоритма для оценки квадратичного интеграла функции плотности вероятностей и функции решений. Алгоритмы являются шаговыми: коррекция осуществляется на основе оценки следующих друг за другом проб.

В теореме 2 представлены условия воспроизведения оценки без искажений, а приложенной леммой устанавливается верхняя граница вариации при любом количестве шагов. Результаты иллюстрируются на примерах (1) оценки информационного измерительного числа (включающего в себя вероятность ошибки) и (2) непосредственного исследования вопроса приближения по вероятности ошибки.

L. Győrfi
Research Institute for Telecommunication
1026 Budapest, II. Gábor Áron u. 65
Hungary

Problems of Control and Information Theory, Vol. 3 (1), pp. 19–34 (1973

FINITE DATA REDUCTION IN THE CASE OF A LONG SEQUENCE OF OBSERVATIONS*

T. NEMETZ

(Budapest)

(Received March 20, 1973)

Let us be given a finite number of simple hypotheses and a sequence of observation $\xi_1, \xi_2, \ldots, \xi_N$ such that under all possible hypotheses the ξ_i's are independent and identically distributed random variables, and N is a very large integer. However, we may use only m symbols (say the numbers $1, 2, \ldots, m$) for describing the value of each individual random variable. (E.g. we want to use a k-bit storage in a computer for each random variable, where $m = 2^k$.) The problem is to find an m-valued statistic T such that the sequence $T\xi_1, T\xi_2, \ldots, T\xi_N$ should contain as much information as possible, at least in an asymptotic sense. We introduce an optimality criterion based on the Chernoff information-number. For an optimal m-valued statistic, the rate of convergence to the prior uncertainty of the information contained in the sequence $T\xi_1, T\xi_2, \ldots$ is the maximum. As it is well known this maximal rate equals the Bayesian error-exponent (see e.g. Rényi [4]). We prove a theorem concerning the form of the "best" statistic. Also we give an iterative procedure providing, step by step, a better statistic, and terminating in the optimal one in many cases. Our results are illustrated in the exponential case.

1. Introduction

Let us recall some definitions and basic results. Let $\Theta = \{\vartheta_1, \ldots, \vartheta_s\}$ be a finite parameter-space, let ϑ be a random variable, called parameter, with an a priori probability distribution $W = \{w_1, \ldots, w_s\}$ where $w_i = \text{Prob}\{\vartheta = \vartheta_i\}$. Let ξ be a random variable defined on a certain probability space and taking its values in the sample space $\{X, \mathfrak{X}\}$. Denote the probability measure generated by ξ on $\{X, \mathfrak{X}\}$ under the hypothesis $\vartheta = \vartheta_i$ by P_i, and the corresponding density functions with respect to a given σ-finite measure P by $p_i(x)$. (P_i's are assumed to be absolutely continuous with respect to P.) Let $\{Y, \mathfrak{Y}\}$ denote the $[0, 1]$ interval with the σ-algebra of its Borel-sets, and consider a random variable ζ having uniform distribution on Y and totally independent of $\{\xi, \vartheta\}$. Given any measurable mapping $d(x, y)$ of $X * Y$ into Θ we can perform a decision $d(\xi, \zeta)$ on the basis of the observations of ξ and ζ. (Here the sign $*$ means the Cartesian product.)

* This paper was done while the author was at Carleton University, Dept. of Math., Ottawa.

The probability of error of a decision is defined as the probability of wrong decision

$$e(d) = \sum_{i=1}^{s} w_i P_i(d \neq \vartheta_i) = \sum_{i=1}^{s} w_i \sum_{\substack{j=1 \\ j \neq i}}^{s} P_i(d = \vartheta_j). \tag{1.1}$$

The error probability e^* for the decision problem is the minimal value of $e(d)$, i.e. $e^* = \inf_{(d)} e(d)$. Following Lindley [2], the amount of information provided by the observation ξ is defined as

$$I(\vartheta, \xi) = H(\vartheta) - E[H(\vartheta \mid \xi)] \tag{1.2}$$

where $H(\cdot)$ denotes the Shannon entropy.* Here the average conditional entropy

$$R(\vartheta \mid \xi) = E[H(\vartheta \mid \xi)] \tag{1.3}$$

is called the missing information.

Between the error-probability and the missing information the following inequality holds

$$2e^* \leq R(\vartheta, \xi). \tag{1.4}$$

The missing information (and so the error-probability, too) may be bounded from above by α-divergences. Let P_1 and P_2 be arbitrary probability measures. Their α-divergence is defined as follows:

$$H_\alpha(P_1, P_2) = \int [p_1(x)]^\alpha [p_2(x)]^{1-\alpha}, \qquad 0 < \alpha < 1. \tag{1.5}$$

We would like to remind that Rényi's original definition of the divergence of order α was

$$\frac{1}{\alpha - 1} \log_2 H_\alpha (P_1, P_2).$$

Setting

$$D(P_1, P_2) \overset{\text{def}}{=} \inf_{0<\alpha<1} H(P_1, P_2) \tag{1.6}$$

we get another useful measure, called the D-divergence of P_1 and P_2.

Consider now a sequence of independent identically distributed observations $\xi_1, \xi_2, \ldots$, where all ξ_i are associated with the same probability law described above for ξ. It is known that the following limit relation is valid

$$\lim_{n\to\infty} \frac{1}{n} \log e_n^* = \lim \frac{1}{n} \log R(\vartheta, \xi^n) = \log \max_{i \neq j} D(P_i, P_j) \tag{1.7}$$

* I.e. $H(\vartheta) = -\Sigma\, w_i \log_2 w_i$, where $0 \cdot \log 0$ is understood to be 0.

for the error-probability e_n^* and for the missing information $R(\vartheta, \xi^n)$ in the case of the observation vector $\xi^n = \{\xi_1, \ldots, \xi_n\}$, see Chernoff [1] and Vajda[5].

This limit relation gives us key to a possible solution of the problem outlined in the Abstract. In fact, as a consequence of (1.7) the m-valued statistic having minimal D-divergence should be considered the best one in the case of a long sequence of observations. On the basis of this idea we can now start investigating our problem. Through this paper we will suppose that there are only two possible hypotheses.

2. The form of the m-valued statistic yielding asymptotically most information

Let us consider a sequence of observation $\xi_1, \xi_2, \ldots$ where the ξ_i's are independent and identically distributed under both the hypotheses H_0 and H_1. Let us denote the probability measures generated by ξ_1 on the sample space $\{X, \mathfrak{X}\}$ by P_0 and P_1, respectively, and their density functions with respect to a certain dominating σ-finite measure P by $f(x)$ and $g(x)$. We assume that P_0 and P_1 are equivalent. Let ξ and $\{Y, \mathfrak{Y}\}$ be the same as in the introduction. Denoting a measurable mapping of $X * Y$ into the set $M = \{1, 2, \ldots, m\}$ by T we define a randomized m-valued statistic of ξ by setting $\eta = T(\xi, \zeta)$. An infinite sequence $\zeta_1, \zeta_2, \ldots$ of independent repetitions of ζ gives us a new sequence of observations $\eta_1, \eta_2, \ldots$ where $\eta_i = T(\xi_i, \zeta_i)$. We seek that mapping T for which the sequence $\eta_1, \eta_2, \ldots$ is asymptotically most informative. (Taking into account that the information provided by the sequence of m-valued statistics is independent of the range of the statistics, we can, without loss of generality, restrict our investigation to the statistics having values in M.) In order to do this we have to find a mapping T such that the D-divergence of $P_{0,\eta}$ and $P_{1,\eta}$ is minimal, as stated in the introduction. Here $P_{i,\eta}$ denote the probability measures generated on M by η under the hypothesis H_i. Now we are going to prove the existence of such a mapping T and, as a result of the proof, we shall also get the form of T. In the last part of this section we prove that the m-valued statistics yielding asymptotically most information are never randomized, except for pathological situations. Let us begin with the following simple lemma:

Lemma 1. Define $H_0(P_0, P_1)$ and $H_1(P_0, P_1)$ as follows:

$$H_0(P_0, P_1) = \lim_{\alpha \to 0} H_\alpha(P_0, P_1) \quad \text{and} \quad H_1(P_0, P_1) = \lim_{\alpha \to 1} H_\alpha(P_0, P_1).$$

If the measures P_0 and P_1 are equivalent and their densities $f(x)$ and $g(x)$ are not identical, then their α-divergence as a function of α attains its minimum at an inner point α_0 of $(0, 1)$.

Proof. It can be seen that

$$H_0(P_0, P_1) = \int_{\{x:f(x)\neq 0\}} g(x)P(dx).$$

Furthermore, $H_\alpha(P_0, P_1)$ is a continuous function of α, thus it attains its minimum on the closed interval $0 \leq \alpha \leq 1$. Using the Hölder-inequality we obtain immediately that $H_\alpha(P_0, P_1) \leq 1$ for $0 < \alpha < 1$ with equality iff (i.e. if and only if) $P_0 \equiv P_1$. Let us assume now that $\alpha_0 = 0$ yields minimal α-divergence. Then

$$H_0(P_0, P_1) = \int_{\{x:f(x)\neq 0\}} g(x)P(dx) = 1$$

since P_0 and P_1 are equivalent, which is obviously a contradiction. Because of the symmetry, $\alpha_0 = 1$ cannot be the minimum point either.

Remark 1. In this case $P_{0,\eta}$ and $P_{1,\eta}$ are also equivalent, so their α-divergence takes on its minimal value at an inner point, too.

Remark 2. It is known that the α-divergence is a strictly $\cup$-convex function of α, independently of whether P_0 and P_1 are equivalent or not provided, they are neither identical nor orthogonal. This means that the minimizing α is uniquely determined except for the last two cases.

Example 1. Let P be the Lebesgue measure on the real line and define P_0 and P_1 by their densities

$$f(x) \equiv 1 \text{ if } 0 \leq x \leq 1 \text{ and } g(x) \equiv \frac{1}{2}, \text{ if } 0 \leq x \leq 2.$$

For these measures we have

$$H_\alpha(f, g) \geq \frac{1}{2} = H_0(f, g),\ 0 \leq \alpha \leq 1$$

with equality iff $\alpha = 0$.

Remark 3. The example just mentioned shows it may happen that $\alpha_0 = 0$. In such a case a sequence of 2-valued statistics may provide all information asymptotically, namely the mapping

$$\eta^*(x) = \begin{cases} 0 & \text{if } f(x) = 0 \\ 1 & \text{else.} \end{cases}$$

In this case we have $D(P_0, P_1) = H_0(P_0, P_1) = H_0(P_{1,\eta^*}, P_{0,\eta^*} = D(P_{0,\eta^*}, P_{1,\eta^*})$. Of course, a similar statement holds for the case $\alpha_0 = 1$. In both cases our problem is trivially solved. We refer to these as trivial cases.

Remark 4. Another trivial case is the following. Let $l(x)$ denote the likelihood ratio of P_0 and P_1

$$l(x) = \frac{f(x)}{g(x)}$$

and let us assume that there exists a function $l^*(x)$ such that

$$P_i(l(x) \neq l^*(x)) = 0, \qquad i = 0, 1$$

where $l^*(x)$ has at most m values. Since $l^*(\xi)$ is a sufficient statistic, the sequence $l^*(\xi_1), l^*(\xi_2), \ldots$ contains the same information as the sequence $\xi_1, \xi_2, \ldots$ does, so our problem is trivially solved again.

Let us consider now the general form of the α-divergence of a 2-valued random variable χ, $\chi \in \{0; 1\}$, using the notation $P_0\{\chi = 0\} = s = 1 - P_0\{\chi = 1\}$, $P_1\{\chi = 0\} = t = 1 - P_1\{\chi = 1\}$

$$\psi_\alpha(s, t) = s^\alpha t^{1-\alpha} + (1 - s)^\alpha (1-t)^{1-\alpha}, \qquad 0 < \alpha < 1. \tag{2.1}$$

This function is well defined on the unit-square $0 \leq s, t \leq 1$ and as we have already stated

$$0 \leq \psi_\alpha(s, t) \leq 1$$

where $\psi_\alpha(s, t) = 1$ iff $s = t$.

In addition, we have

Lemma 2. $\psi_\alpha(s, t)$, $0 < \alpha < 1$, is a strictly $\cap$-convex function along every straight line in $0 \leq s, t \leq 1$ except for the diagonal $s = t$, where it is identically one.

Proof. We have to prove that for $0 < \tau < 1$

$$\tau\psi_\alpha(s_1, t_1) + (1 - \tau)\psi_\alpha(s_2, t_2) < \psi_\alpha[\tau s_1 + (1 - \tau) s_2, \tau t_1 + (1 - \tau) t_2] \tag{2.2}$$

except for the cases $(s_1, t_1) = (s_2, t_2)$ or $s_1 = t_1$ and $s_2 = t_2$. Using the fact that $f(u) = u^\alpha$, $0 < \alpha < 1$ is strictly $\cap$-convex on the whole positive half-line, the proof goes as follows:

$$\tau s_1^\alpha t_1^{\alpha-\alpha} + (1 - \tau) s_2^\alpha t_2^{1-\alpha} < [\tau s_1 + (1 - \tau) s_2]^\alpha \cdot [\tau t_1 + (1 - \tau) t_2]^{1-\alpha}$$

since

$$\frac{\tau t_1}{\tau t_1 + (1 - \tau) t_2} \left(\frac{s_1}{t_1}\right)^\alpha + \frac{(1 + \tau) t_2}{\tau t_1 + (1 - \tau) t_2} \left(\frac{s_2}{t_2}\right)^\alpha < \left[\frac{\tau s_1 + (1 - \tau) s_2}{\tau t_1 + (1 - \tau) t_2}\right]^\alpha$$

with equality iff $s_1/t_1 = s_2/t_2$.

Similarly

$$\tau(1-s_1)^\alpha(1-t_1)^{1-\alpha}+(1-\tau)(1-s_2)^\alpha(1-t_2)^{1-\alpha}\leq$$
$$\leq[\tau(1-s_1)+(1-\tau)(1-s_2)]^\alpha[\tau(1-t_1)+(1-\tau)(1-t_2)]^{1-\alpha}$$

with equality iff $\dfrac{1-s_1}{1-t_1}=\dfrac{1-s_2}{1-t_2}$.

These two conditions lead to the two exceptional cases. It has to be mentioned that the boundary points of the unit-square should need some further routine discussion, which we omit. Lemma 2 enables us to get the result for $m=2$, namely it means that minimum can be reached at some point of the boundary of the proper domain of (s, t). I.e. we have to look for the optimum among those (s, t) pairs only, for which there is a set $A\in\mathcal{X}*\mathcal{Y}$ such that

$$s=P_0(A)\quad\text{and}\quad t=P_1(A).$$

Let us denote the set of all these pairs (s, t) by $\tilde{G}=\tilde{G}(P_0, P_1)$. Concerning $\tilde{G}$ the following statements hold (see Nemetz [3]):

(i) Given any s there exists at least one $t(s)$ such that $(s, t(s))\in\tilde{G}$. In fact, $t(s)=s$ will do.

(ii) If $(s_1, t_1)\in\tilde{G}$ and $(s_2, t_2)\in\tilde{G}$ then for any $0\leq\tau\leq 1$ $(\lambda s_1+(1-\tau)s_2, \tau t_1+(1-\tau)t_2)\in\tilde{G}$ i.e. $\tilde{G}$ is a convex set.

(iii) $\tilde{G}$ is a closed set.

(iv) $\tilde{G}$ is symmetric with respect to the point $(1/2, 1/2)$. Let $t^*(s)=\min\{t:(s,t)\in\tilde{G}\}$, i.e. $t^*(s)$ denotes the "lower" boundary of $\tilde{G}$. Defining

$$A(c,\gamma)=\{(x,y):\frac{g(x)}{f(x)}<c\}\cup\{(x,y):\frac{g(x)}{f(x)}=c,\ y\leq\gamma\}\tag{2.3}$$

we can state

(v) There are numbers c and γ such that

$$s=P_0(A(c,\gamma))\quad\text{and}\quad t^*(s)=P_1(A(c,\gamma)).$$

Summarizing these results we have

Lemma 3. There exists a mapping T, yielding minimal α-divergence for $m=2$, given α, which has the form:

$$T(x,y)=\begin{cases}0 & \text{if}\quad (x,y)\in A(c,\gamma)\\ 1 & \text{else,}\end{cases}\tag{2.4}$$

where the numbers c, γ are not necessarily uniquely determined.

Remark 5. Because of Lemma 1, the same result holds for the D-divergence, too.

In (2.4) the constant γ plays the role of randomization parameter. Since more information cannot be gained by randomization anyone considers it immediately unnecessary. Indeed, it plays only an auxiliary role in this derivation. Let $t_g(s) \leq s$ be defined by the equation

$$g = \psi_\alpha(s, t_g(s))$$

provided there is such a $t_g(s)$. It may be proved that $t_g(s)$ is a strictly $\cup$-convex function of s on its domain. Since a straight line corresponds to the randomization in the $t^*(s)$-curve, it is obvious that no inner point of this line, i.e. no $0 < \gamma < 1$ can provide an optimal point, in general. Indeed, it follows immediately from Lemma 2, that $0 < \gamma < 1$ may yield optimal mapping iff $\tilde{G}(P_0, P_1)$ is the diagonal $s = t$, i.e. iff P_0 and P_1 are identical.

This means that in (2.4) $\gamma = \gamma(c)$ is equal to either 0 or 1, provided P_0 and P_1 are not identical.

In order to extend our results to the case $m > 2$, let us observe that any m-valued statistic η defines a measurable partition $A(\eta) = \{A_1, A_2, \ldots \ldots, A_m\}$ of $X * Y$ where

$$A_i = \{(x, y) : \eta(x, y) = i\}.$$

Therefore in this case we have to investigate the set

$$\tilde{G}(P_0, P_1, m) = \{s_i = P_0(A_i), t_i = P_1(A_i), i = 1, 2, \ldots m\}$$

instead of $\tilde{G}(P_0, P_1)$, where $\{A_1, A_2, \ldots, A_m\}$ denotes measurable partitions.

It is easy to deduce from a theorem of Blackwell [6] that $\tilde{G}(P_0, P_1, m)$ is a closed, convex set.

Futhermore instead of $\psi_\alpha(s, t)$ now we have to consider the function

$$\psi_\alpha(s, t) = \sum_{\alpha=1}^{m} s_i^\alpha \, t_i^{1-\alpha}. \tag{2.5}$$

A straightforward generalization of Lemma 2 shows, that $\psi_\alpha(\underline{s}, \underline{t})$ is an $\cap$-convex function of $(\underline{s}, \underline{t})$, $0 \leq s_i \leq 1$, $0 \leq t_i \leq 1$, $\sum_{i=1}^{m} s_i = \sum_{i=1}^{m} t_i = 1$.

This means, that in this case also exists optimal m-valued statistic. Excluding the trivial cases, let us consider now a measurable partition $\mathcal{A} = \{A_1, A_2, \ldots, A_m\}$ such that the corresponding m-valued statistic

$$\eta(\xi, \zeta) = i, \quad \text{if} \quad (\xi, \zeta) \in A_i$$

is an optimal one, and investigate the connection between A_i and A_j, where i and j, $i \neq j$ are fixed. Let $B = A_i \cup A_i$ and define another statistic $\eta^*(\xi, \zeta)$ for a given $A \subset B$, $A \in \mathfrak{X} \otimes \mathfrak{Y}$ by setting

$$\eta^*(\xi, \zeta) = \begin{cases} \eta(\xi, \zeta), & \text{if} \quad (\xi, \zeta) \notin B \\ i, & \text{if} \quad (\xi, \zeta) \in A \\ j, & \text{if} \quad (\xi, \zeta) \in B - A. \end{cases}$$

Here $H_\alpha(P_{0,\eta}, P_{1,\eta})$ and $H_\alpha(P_{0,\eta^*}, P_{1,\eta^*})$ differ only in terms corresponding to the pairs of sets (A_i, A_j) and $(A, B - A)$. Defining the probability measures

$$P_k^*(\ ) = \frac{1}{P_k(B)} P_k(\ .\ ), k = 0, 1, \tag{2.6}$$

we can apply Lemma 3. (In fact, it is easy to see that $P_k(A_i)$ and $P_k(A_j)$ cannot be equal to zero at the same time, $k = 0, 1$, except for the trivial cases, therefore (2.6) really defines probability measures.) Obviously, here the densities are connected in the same manner:

$$p_k^*(x) = \frac{1}{P_k(B)} p_k(\ .\ ), k = 0, 1.$$

Considering the expression

$$H_\alpha(P_{0,\eta^*}, P_{1,\eta^*}) = \sum_{\substack{k \neq i \\ k \neq j}} P_0(A_k)^\alpha P_1(A_k)^{1-\alpha} + [P_0(A)^\alpha P_1(A)^{1-\alpha} +$$

$$+ P_0(B - A)^\alpha P_1(B - A)^{1-\alpha}] \sum_{\substack{k \neq i \\ k \neq j}} P_0(A_k)^\alpha P_1(A_k)^{1-\alpha} +$$

$$+ P_0(B)^\alpha P_1(B)^{1-\alpha} \cdot [P_0^*(A)^\alpha P_1^*(A)^{1-\alpha} + P_0^*(B - A)^\alpha P_1^*(B - A)^{1-\alpha}]$$

we see that (A_i, A_j) must define one of the best 2-valued statistics, in the above sense, for the pair of probability distributions P_0^* and P_1^*. According to Lemma 3 this means that one of the sets A_i and A_j is of the form

$$\{(x, y) \in B\} \cap [\{(x, y): \frac{g(x)}{f(x)} < c \cdot \frac{P_0(B)}{P_A(B)} \overset{\text{def}}{=} c^*\} \cup \{(x, y): \frac{g(x)}{f(x)} = c^*, y \leq \gamma(c^*)\}].$$

If here $0 < \gamma(c^*) < 1$, then $\dfrac{g(x)}{f(x)}$ is constant on B with probability one. This, by simple induction, yields that we have a trivial case, which is contradiction. Therefore here $\gamma(c^*) = 0$ or 1. Furthermore, our derivation is valid for every pair (i, j).

Let

$$c_i^- = \inf_{(x,y)\in A_i} \frac{g(x)}{f(x)} \qquad \text{and} \qquad c_i^+ = \sup_{(x,y)\in A_i} \frac{g(x)}{f(x)} .$$

Rearranging the sets, if necessary, we have $c_1^+ \leq c_2^+ \leq \ldots \leq c_m^+$ and, if $c_i^+ = c_{i+1}^+$ then $c_1^- \leq c_{i+1}^-$. Obviously it holds for c^* in (2.7) for $i = 1$ and $j = 2$ that

$$c_1^+ \leq c^* \leq c_2^+$$

and therefore the set given in (2.7) is A_1. In addition, $c_2^- \geq c_1^+$. Similarly, we obtain $c_1^- \leq c_1^+ \leq c_2^- \leq \ldots \leq c_{m-1}^+ \leq {}_m^-c \leq c_m^+$. Thus we can conclude in the following

Theorem. Let P_0 and P_1 equivalent probability measures with densities f and g with respect to a certain σ-finite measure. Let ξ be a random variable associated on the sample space with the probability measures P_0 and P_1 under the hypotheses H_0 and H_1, respectively. Then there are real numbers $0 \leq c_1 \leq c_2 \leq \ldots \leq c_m \leq c_{m+1} = +\infty$ and a sequence of zeros and ones $\gamma_1 = 0, \gamma_2, \ldots, \gamma_m, \gamma_{m+1} = 1$ such that the m-valued statistic

$$\eta(\xi) = \mathrm{i} \quad \text{iff} \quad \xi \in A(c_{i+1}, \gamma_{i+1}) - A(c_i, \gamma_i) \tag{2.8}$$

yields minimal α-divergence among all measurable m-valued statistics. Here the numbers $\{c_i\}$ may depend on α. In addition, there exists m-valued statistic of the same form which provides minimal D-divergence.

Remark 6. In our derivation we assumed the equivalence of P_0 and P_1 only to assure that the minimal D-divergence for the m-valued statistics is attained at a point $0 < \alpha_0 < 1$. Nothing else was used. So, if P_0 and P_1 are not equivalent, but α_0 satisfies the condition $0 < \alpha_0 < 1$ the proof is correct, and a mapping of the form (2.8) provides minimal D-divergences. If $\alpha_0 = 0$ (or $\alpha_0 = 1$) there exists a trivial 2-valued statistic yielding minimal D-divergence, as it is stated in Remark 3, and this statistic has the same form. Thus we can conclude, that our theorem holds for every pair of probability measures. We are not going into the details, whether situations, where

$$\min_{(\eta)} D(P_0, P_1) = \min_{(\eta)} H_0(P_{0\eta}, P_{1\eta})$$

and at the same time

$$D(P_0, P_1) \neq H_0(P_0, P_1)$$

can exist or not.

The following example shows that the m-valued statistic yielding minimal α-divergence and having the form (2.8) is not necessarily uniquely determined.

Example 2. Let ξ have uniform distribution on $[0, 1]$ under the hypothesis H_0 and distribution function

$$P_1(\xi < x) = 1/2\,[3/2 - x - \sqrt{3x(1-x)}] \text{ if } 3/4 \leq x \leq 1$$
$$P_1(\xi = 1) = 3/4$$

under the hypothesis H_1.

It is easy to see that for all the 2-valued statistics

$$\eta_c(\xi) = \begin{cases} 0 & \text{if} \quad \xi < c \\ 1 & \text{if} \quad c \leq \xi \end{cases} \tag{2.9}$$

where $3/4 < c < 1$, the 1/2-divergence is equal to 1/2 i.e.

$$H_{1/2}(P_{0,\eta_c}, P_{1,\eta_c}) = 1/2.$$

Here, obviously,

$$P_k(\eta_{c_1} \neq \eta_{c_2}) \neq 0, \qquad k = 0, 1$$

for every $3/4 < c_1 < c_2 < 1$.

3. An iterative procedure for obtaining the optimal statistic

The method of proof suggests a possible way for obtaining the optimal m-valued statistic.

Let us choose an initial value (C_1, Γ_1) for the vector

$$(C, \Gamma) = \{c_1 = 0 \leq c_2 \leq \ldots \leq c_m \leq c_{m+1} = +\infty,$$
$$\gamma_1 = 0, \gamma_2, \ldots, \gamma_m, \gamma_{m+1} = 1\}$$

so that the restriction of P_0 and P_1 on the sets $A(c_{i+2}, \gamma_{i+2}) = A(c_i, \gamma_i)$ would not be identical, $i = 1, 2, \ldots, m-1$. Were this impossible, the likelihood function would have at most m different values with P-probability one, and so we would have a trivial problem. Let α_1 be that value of α for which

$$H_{\alpha_1}(P_{0,\eta_1}, P_{1,\eta_1}) = \inf_{0<\alpha<1} H_\alpha(P_{0,\eta_1}, P_{1,\eta_1}) \overset{\text{def}}{=} D_1$$

or a good approximation thereof, where η_1 is the mapping defined by (C_1, Γ_1) and (2.8). Now let $(C_{1,1}, \Gamma_{1,1}) = (C_1, \Gamma_1)$ and having defined $(C_{1,n}, \Gamma_{1,n})$ determine $(C_{1,n+1}, \Gamma_{1,n+1})$ as it follows:

$$c_{i,n+1} = c_{i,n};\ \gamma_{i,n+1} = \gamma_{i,n} \quad \text{if} \quad i = 1, i = m+1$$

and if

$$2 \leq i \leq m,\ i \not\equiv n+1 \pmod{m-1}$$

$$c_{i,n+1} = c^*,\ \gamma_{i,n+1} = \gamma^*_i,\ 2 \leq i \leq m,\ i \equiv n+1 \pmod{m-1}$$

where $c_{i-1,n} \leq c^*_i \leq c_{i+1,n}$, and γ^* denote such values for which $\eta_{1,n+1}$ defined by (2.8) and $(C_{1,n+1}, \Gamma_{1,n+1})$ yield minimal α_1-divergence, given all the others c_i, γ_i. Doing this we have a sequence of statistics $\eta_{1,n}$, each of them providing α_1-divergence $H_1(n)$ less than (or equal to) D_1. So the sequence $H_1(n)$ converges to a number $H_1 \leq D_1$. Let (C_2, Γ_2) be one of the limit vectors of the sequence $\{(C_{1,n}, \Gamma_{1,n})\}$ or a good approximation it. Define α_2, D_2 in the same manner as in the case of α_1, D_1 and repeat the whole procedure. So we have $D_1 \geq H_1 \geq D_2 \geq \ldots$ The limit mappings defined by the sequence $\{(C_n, \Gamma_n)\}$ are called suboptimal mappings. It is clear that the D-divergence provided by any of them equals $\lim_{(n)} D_n \overset{\text{def}}{=} D_0$. The convexity properties characterizing our problem are good enough reasons to hope that D_0 is the minimal D-divergence attainable by m-valued statistics in many cases. Indeed, computer-studies show that in some important cases our procedure terminates in optimal m-valued statistics. But there are examples when it does not do so. As to such a counter example we should like to refer to Example 2. Let $m = 2$. It is obvious that in this case one of the statistics defined by (2.9) yields minimal D-divergence. If there exist a, b, $3/4 < a < 1$ and $3/4 < b < 1$ such that

$$D(P_{0,\eta_a}, P_{1,\eta_a}) = H_{1/2}(P_{0,\eta_a}, P_{1,\eta_a}) = 1/2 \tag{3.1}$$

and

$$D(P_{0,\eta_b}, P_{1,\eta_b}) < H_{1/2}(P_{0,\eta_b}, P_{1,\eta_b}) = 1/2 \tag{3.2}$$

then we have a counter example. Indeed, our procedure with initial vector $(C_1, \Gamma_1) = \{c_1 = 0, c_2 = a, c_3 = 1, \gamma_1 = 0, \gamma_2, \gamma_3 = 1\}$ repeatedly may result in the statistic η_a, which is obviously not optimal, and, in fact, such an a and b do exist. Let us consider the following expression

$$\left.\frac{dH_\alpha(s(c), t(c))}{d\alpha}\right|_{\alpha=1/2} = \left(\log \frac{s(c)}{t(c)}\right)\sqrt{s(c)\,t(c)} + \left(\log \frac{1-s(c)}{1-t(c)}\right)\sqrt{(1-s(c))(1-t(c))} \tag{3.3}$$

where

$$s(c) = c \quad \text{if} \quad 3/4 < c < 1$$

$$t(c) = 1/2\left[3/2 - c - \sqrt{3c(1-c)}\right] \text{ if } 3/4 < c < 1.$$

Here the right hand side of (3.3) is continuous, and

$$\lim_{c\to 3/4} \frac{dH_\alpha}{d\alpha}\bigg|_{\alpha=1/2} = -\log 2, \qquad \lim_{c\to 1} \frac{dH_\alpha}{d\alpha}\bigg|_{\alpha=1/2} = +\log 2.$$

This means the existence of an a and b required in (3.1) and (3.2).

4. The exponential case

In order to illustrate the ideas above, let us consider, as an example, the exponential case.

Let $X = [0, +\infty)$ with its Borel σ-algebra $\mathfrak{X}$, and let the measures P_0 and P_1 be given by the distribution functions

$$F(x) = P_0(\xi < x) = 1 - e^{-\lambda x} \quad \text{if} \quad 0 \leq x < +\infty$$

and

$$G(x) = P_1(\xi < x) = 1 - e^{-\mu x} \quad \text{if} \quad 0 \leq x < +\infty,$$

respectively. Without loss of generality we can assume that $\lambda < \mu$. In this case the statistic

$$\xi^* = e^{-\lambda\xi}$$

is obviously a sufficient statistic, so we may restrict ourselves to the mapping ξ^*. It is easy to see that

$$P_0(\xi^* < x) = x \quad \text{if} \quad 0 \leq x \leq 1$$

$$P_1(\xi^* < x) = x^{\mu/\lambda} \overset{\text{def}}{=} x^\varrho, \quad \text{if} \quad 0 \leq x \leq 1$$

and

$$H_\alpha(P_1, P_0) = \int_0^1 \varrho^\alpha x^{(\varrho-1)\alpha}\, dx = \frac{1}{(\varrho-1)\alpha + 1}\, \varrho^\alpha$$

$$D(P_0, P_1) = \min_{(\alpha)} H_\alpha(P_1, P_0) = H_{\alpha_0}(P_1, P_0) = \frac{\log \varrho}{\varrho - 1} \cdot e^{1-\frac{\log\varrho}{\varrho-1}} \tag{4.1}$$

where

$$\alpha_0 = \frac{1}{\log \varrho} - \frac{1}{\varrho - 1}\,. \tag{4.2}$$

Our theorem tells us that there exists an optimal m-valued statistic of the form

$$\eta(\xi) = i \quad \text{iff} \quad x_i \leq \xi^* < x_{i+1}, \quad i = 1, 2, \ldots, m \tag{4.3}$$

where $0 = x_1 < x_2 < \ldots < x_m < x_{m+1} = 1$. The corresponding probabilities are given by

$$s_i = P_0(\eta = i) = x_{i+1} - x_i,$$

$$t_i = P_1(\eta = i) = x_{i+1}^{\varrho} - x_i^{\varrho}, \qquad i = 1, 2, \ldots, m$$

and so

$$H_\alpha(P_{1,\eta}, P_{0,\eta}) = \sum_{i=1}^{m} (x_{i+1} - x_i)^{1-\alpha} (x_{i+1}^{\varrho} - x_i^{\varrho})^{\alpha}. \tag{4.4}$$

Rearranging, we have

$$H_\alpha(P_{1,\eta}, P_{0,\eta}) = \left[\left(\frac{x_2}{x_3}\right)^{1+(\varrho-1)\alpha} + \left\{1 - \left(\frac{x_2}{x_3}\right)^{\varrho}\right\}^{\alpha} \left(1 - \frac{x_2}{x_3}\right)^{1-\alpha}\right] \cdot x_3^{1+(\varrho-1)\alpha} +$$

$$+ \sum_{i=3}^{m} (x_{i+1} - x_i)^{1-\alpha} (x_{i+1}^{\varrho} - x_i^{\varrho})^{\alpha}.$$

Given $x_3, \ldots, x_{m+1}$ and α we have to choose that value or x_2 for which $[\ldots]$ takes on its minimal value or, what is the same, the minimum place $0 \leq t_2 \leq 1$ of the function

$$\varphi_2(t) = t^{1+(\varrho-1)\alpha} + (1 - t^{\varrho})^{\alpha}(1 - t)^{1-\alpha}.$$

Let us observe, that $\varphi_2(t)$ is the α-divergence of the measures generated by the 2-valued statistic

$$\eta_i^*(\xi) = \begin{cases} 1 & \text{if} \quad 0 \leq \xi^* < t \\ 2 & \text{if} \quad t \leq \xi^* \leq 1. \end{cases}$$

Now let $A_2 = \varphi_2(t_2)$ and let us look for the optimal value x_3, where $x_2 = t_2 x_3$ and $x_4, \ldots, x_{m+1}$ are given numbers. We have

$$H_\alpha(P_{1,\eta}, P_{0,\eta}) = \left[A_2 \left(\frac{x_3}{x_4}\right)^{1+(\varrho-1)\alpha} + \left\{1 - \left(\frac{x_3}{x_4}\right)^{\varrho}\right\}^{\alpha} \left(1 - \frac{x_3}{x_4}\right)^{1-\alpha}\right] \cdot x_4^{1+(\varrho-1)\alpha} +$$

$$+ \sum_{i=4}^{m} (x_{i+1} - x_i)^{1-\alpha} (x_{i+1}^{\varrho} - x_i^{\varrho})^{\alpha}.$$

Continuing this procedure in $(m - 1)$ steps it is obvious that we have to find the minimum places $0 \leq t_i \leq 1$ of the functions

$$\varphi_i(t) = A_{i-1} t^{1+(\varrho-1)\alpha} + (1 - t^{\varrho})^{\alpha}(1 - t)^{1-\alpha}, \quad i = 2, \ldots, m. \tag{4.5}$$

Here A_i denotes the minimal α-divergence attainable by i-valued statistics, and so by the Hölder inequality

$$1 \geq A_i > H_\alpha(P_1, P_0) > 0. \tag{4.6}$$

It is obvious that $\varphi_i(t)$ is a $\cup$-convex function, so the numbers t_i are uniquely defined. Furthermore $0 < t_i < 1$. Once we had determined the numbers $t_2, t_3, \ldots, t_m$ we have that the η defined by (4.3) with $x_m = t_m$, $x_i = t_i x_{i+1}$, $i = m - 1, \ldots, 2$ results in the m-valued statistic yielding minimal α-divergence.

The procedure suggested in paragraph 3 works as it follows: Choose an initial value $\alpha = \alpha_1$, e.g. the one given in (4.2). Determine the numbers $t_i = t_i(\alpha_1)$ and by them the numbers $x_i = x_i(\alpha_1)$. Compute the new value of α, $\alpha = \alpha_2$ minimizing the expression (4.4) given x_i's, and continue. As numerical computations show, only a few iterations (in α, $m \leq 5$) are needed to approximate the optimal statistic well enough).

As to the computational work, the numbers t_i can be easily obtained by succesive approximation. In fact, t_i is the unique root in $0 \leq t \leq 1$ of the equation

$$\frac{d\varphi_i(t)}{dt} = A_{i-1}[1 + (\varrho - 1)\alpha]\, t^{(\varrho-1)\alpha} - \alpha\varrho(1 - t^\varrho)^{\alpha-1} t^{\varrho-1}(1 - t)^{1-\alpha} - \tag{4.7}$$
$$- (1 - \alpha)(1 - t^\varrho)^\alpha (1 - t)^{-\alpha} = 0\,,$$

or, in a rearranged form

$$\left[\frac{t^\varrho}{1 - t^\varrho} \cdot \frac{1 - t}{t}\right]^\alpha = \frac{1}{A_{i-1}[\varrho\alpha + (1 - \alpha)]} \cdot \left[\alpha\varrho \frac{t^\varrho}{1 - t^\varrho} \cdot \frac{1 - t}{t} + 1 - \alpha\right].$$

Introducing the new variable

$$u = \frac{t^\varrho}{1 - t^\varrho} \cdot \frac{1 - t}{t} \tag{4.8}$$

we have

$$u = B_{i-1}\left(u + \frac{1 - \alpha}{\alpha\varrho}\right)^{1/\alpha},$$

where

$$B_{i-1} = \left(\frac{\alpha\varrho}{A_{i-1}[\varrho\alpha + 1 - \alpha]}\right)^{1/\alpha}.$$

It is easy to see, that $u = u(t)$ is a monotone increasing function of t in $0 < t < 1$ and $\lim_{t\to 0} u(t) = 0$, $\lim_{t\to 1} u(t) = 1/\varrho$. So (4.9) must have a unique root in $(0, 1/\varrho)$. Indeed, this can be seen immediately from the convexity of the function $(u + a)^\beta$ with $\beta < 1$ and from the inequality

$$B_{i-1}\left(\frac{1}{\varrho} + \frac{1 - \alpha}{\alpha\varrho}\right)^{1/\alpha} \leq 1/\varrho$$

obtained from (4.6) and (4.1). This means that the sequence $\{u_n\}$

$$u_{n+1} = B_{i-1}\left(u_n + \frac{1-\alpha}{\alpha\varrho}\right)^{1/\alpha}, \; 0 \leq u_0 < 1/\varrho$$

tends to the root $u(i)$ of (4.9). Similarly the root t_i of (4.7) may be obtained from $u(i)$ by successive approximation based on the form

$$t = (t^\varrho(1-u) + u)^{1/\varrho - 1}.$$

We mention the special case $\varrho = 2$, when $t = \dfrac{u}{1-u}$

Acknowledgement

The author is indebted to professors I. Csiszár, M. Hellman and A. Perez for their valuable remarks.

References

1. *Chernoff, H.*, A measure of asymptotic efficiency for tests of a hypothesis based on the sum of observations, Ann. Math. Stat., **23** (1952) 493–507.
2. *Lindley, D. V.*, On a measure of the information provided by an experiment, Ann. Math. Stat. **27** (1956) 986–1005.
3. *Nemetz, T.*, Information theory and testing hypotheses, Matematikai Lapok **20** (1969) 241–275 (In Hungarian), also Proc. Coll. on Inf. Theory, Debrecen (Hungary) 1967, 283–293.
4. *Rényi, A.*, On some problems of statistics from the point of view of information theory, Proc. Coll. on Inf. Theory, Debrecen (Hungary) 1967, 343–357.
5. *Vajda, I.*, On the convergence of information contained in a sequence of observations, Proc. Coll. on Inf. Theory, Debrecen (Hungary), 1967; or: Limit theorems for total variation of Cartesian Product Measures, Studia Sci. Math. Hung., **6** (1971) 317—334.
6. *Blackwell, D.*, The range of certain vector integrals, Proc. Amer. Math. Soc., **2** (1951) 390–395.

Коррекция конечного числа данных в случае продолжительной последовательности наблюдений

Т. НЕМЕЦ

(Будапешт)

Пусть заданы конечное число простых гипотез и последовательность наблюдений $\xi_1, \xi_2, \ldots, \xi_N$ таких, что при всех возможных гипотезах эти значения являются независимыми, равномерно распределенными случайными переменными и N — большое целое число. Обычно ограничиваются учетом только m символов (скажем: 1, 2, . . ., m) для описания значений всех отдельных случайных переменных. Проблема заключается в нахождении такой статистики T с m-значениями, при которой последовательность $T\xi_1$, $T\xi_2, \ldots, T\xi_N$ содержит в себе максимально возможную информацию в асимптотическом

смысле. Такую статистику назовем «лучшей». Введен критерий оптимальности, основанный на применении информационного числа Чернова. Для оптимальной статистики с m-значениями скорость сходимости максимальна и равняется бейесовой экспоненте ошибки.

Представлена теорема в соответствии с «лучшей» статистикой. Разработан интеративный процесс для определения «лучшей» статистики, являющийся оптимальным во многих случаях. Результаты иллюстрируются для экспоненциального случая.

T. Nemetz
Mathematical Research Institute
Hungarian Academy of Sciences
1053 Budapest, V. Reáltanoda u. 13—15.
Hungary

Problems of Control and Information Theory, Vol. 3 (1), pp. 35–45 (1974)

RISK INCREASE INFORMATIONAL ESTIMATES UNDER OBSERVED DATA REDUCTION IN SEQUENTIAL ANALYSIS

E. M. KHAZEN

(Moscow)

(Received June 21, 1972)

For observed data reduction risk increase estimates are derived in the term-of losses of Shannon's amount of information, contained in the observed sample, about estimated parameter θ and results of further observations (the latter inforl mation is used for choosing the final solution and the time to stop observations)s Construction of optimal and suboptimal decision rules is discussed. The result-obtained enable one to estimate, in a variety of cases, closeness of risks of sub. optimal and optimal decision rules, and limiting potential accuracy of statisticas decisions as well.

1. Introduction

Limited computer memory and requirements for simple device implementation stipulate data reduction in constructing optimal statistical decision rules. If a decision is to be taken on the basis of a fixed number N of observations $x_1, \ldots, x_N$, reduction to a sufficient statistics $Z(x_1, \ldots, x_N)$ in Bayes problems is known not to lead to risk growth. Sufficiency here is understood in Fisher's sense: a function or a set of functions $Z(x_1, \ldots, x_N)$ is a sufficient statistics for estimation of a parameter θ (about which the final decision should be made) if under given values of $x_1, \ldots, x_N$ conditional probability distribution for θ depends on $x_1, \ldots, x_N$ only through $Z(x_1, \ldots, x_N)$. A. Perez [1, 2] has found a number of risk increase informational estimates for reduction where sufficiency is not observed.

In the cases of sequential analysis, observed data reduction without risk growth takes place if statistics $Z(x_1, \ldots, x_m)$ (on which both decision to stop observations and the final decision are based) is transitive apart from being Fisher-sufficient [3]. It means that, provided the values of $x_1, \ldots, x_m$ are given, conditional probability distribution for the value of $Z_{m+1} = Z(x_1, \ldots, x_{m+1})$ depends on $x_1, \ldots, x_m$ only through $Z(x_1, \ldots, x_m) = Z_m$.

In [4] some risk change estimates for sequential analysis were obtained caused by a probability distribution modification and observed data reduction. The present paper aims at further extension of the results obtained in [1, 2, 4, 6] to the sequential analysis and optimal control.

2. Risk increase informational estimates under data reduction

Let $x_1, \ldots, x_m$ be the results of sequentially made observations. A sequential decision rule is given by "stop sets" Q_n (where Q_n is the set of all samples $(x_1, \ldots, x_n)$ for which decision is taken at nth step, $n = 1, 2 \ldots$), and decision functions $b(x_1, \ldots, x_n)$ for the choice of final decision after the stop of observations.

The loss function $W(b(x_1, \ldots, x_n); \theta; n)$ (or "fine" for a taken decision $b(x_1, \ldots, x_n)$ under the true value θ of parameter) in Bayes sequential analysis problems is assumed to be given. The loss function W may allow for the cost of n observations. Risk is defined as the mathematical expectation of losses. Introduce the following notations:

$R(\theta)$ — risk if parameter takes value θ,
R — mean risk, and
ϱ — Bayes risk, i.e. risk corresponding to the optimal decision rule.

For the case where the sets Q_n are fixed, and the probability density $p(x_1, \ldots, x_n/\theta)$ of the results of $x_1, \ldots, x_n$ observations under parameter value θ is substituted by $\tilde{p}(x_1, \ldots, x_n/\theta)$, in [4] the following risk change estimates have been obtained similarly to [1, 2]:

$$-\sqrt{2R(W^2/\theta)\overline{I(\tilde{p}:p/\theta)}} \leq \tilde{R}(\theta) - R(\theta) \leq \sqrt{2\tilde{R}(W^2/\theta)\overline{I(p:\tilde{p}/\theta)}} \tag{2.1}$$

$$-\sqrt{2W_0 R(\theta)\overline{I(\tilde{p}:p/\theta)}} \leq \tilde{R}(\theta) - R(\theta) \leq \sqrt{2W_0\tilde{R}(\theta)\overline{I(p:\tilde{p}/\theta)}} \tag{2.2}$$

$$\overline{I(p:\tilde{p}/\theta)} \geq \frac{1}{W_0}\left[R(\theta)\ \ln\frac{R(\theta)}{\tilde{R}(\theta)} + \big(W_0 - R(\theta)\big)\ln\frac{W_0 - R(\theta)}{W_0 - \tilde{R}(\theta)}\right], \tag{2.3}$$

where $R(W^2/\theta)$ is the mathematical expectation of squared losses under the true parameter value θ. It is assumed in (2.2) and (2.3) that $W(b(x_1, \ldots, x_n); \theta; n) \leq$ $\leq W_0$. $\overline{I(p:\tilde{p}/\theta)}$ is the mean value of the informational divergence between $p(x_1, \ldots, x_n/\theta)$ and $\tilde{p}(x_1, \ldots, x_n/\theta)$ defined by the following equality:

$$\begin{aligned}\overline{I(p:\tilde{p}/\theta)} &= \sum_{n=1}^{\infty}\int\ldots\int_{Q_n} p(x_1,\ldots,x_n/\theta)\ln\frac{p(x_1,\ldots,x_n/\theta)}{\tilde{p}(x_1,\ldots,x_n/\theta)}\,dx_1,\ldots,dx_n,\\ \overline{I(\tilde{p}:p/\theta)} &= \sum_{n=1}^{\infty}\int\ldots\int_{Q_n} \tilde{p}(x_1,\ldots,x_n/\theta)\ln\frac{\tilde{p}(x_1,\ldots,x_n/\theta)}{p(x_1,\ldots,x_n/\theta)}\,dx,\ldots,dx_n.\end{aligned} \tag{2.4}$$

Similar estimates may be done for the mean risks R and $\tilde{R}$, $\overline{I(p:\tilde{p}/\theta)}$ being substituted by

$$\overline{I(p:\tilde{p})} = \int \overline{I(p:\tilde{p}/\theta)}\ p(\theta)\,d\theta. \tag{2.5}$$

If for the distributions $\tilde{p}$ and p, under any fixed value of θ the observation results are independent, the following holds:

$$\overline{I(p:\tilde{p}/\theta)} = \overline{n}(\theta)\, I_1(p:\tilde{p}/(\theta);\ \overline{I(\tilde{p}:p/\theta)} = \bar{\tilde{n}}(\theta)\, I_1(\tilde{p}:p/\theta)\,,$$
$$\overline{I(p:\tilde{p})} = \overline{n}(\theta^*)\, I_1(p:\tilde{p}/\theta^*);\ \overline{I(\tilde{p}:p)} = \bar{\tilde{n}}(\theta^*)\, I_1(\tilde{p}:p/\theta^*)\,, \tag{2.6}$$

where $\overline{n}(\theta)$ is the mathematical expectation of the number of observations required by a given sequential decision rule under the distribution density $p(x_1, \ldots, x_n/\theta)$; $\bar{\tilde{n}}(\theta)$ is defined similarly for the density $\tilde{p}(x_1, \ldots, x_m/\theta)$; $m = 1, 2, 3, \ldots$; and

$$I_1(p:\tilde{p}/\theta) = \int_{-\infty}^{+\infty} p(x/\theta)\, \ln \frac{p(x/\theta)}{\tilde{p}(x/\theta)}\, dx$$

is the informational divergence between distributions $p(x/\theta)$ and $\tilde{p}(x/\theta)$ for the result x of an observation. It is assumed that $p(x/\theta) = 0$ for the points x where $\tilde{p}(x/\theta) = 0$, and *vice versa*, integrals (2.4) and (2.6) thus being finite.

Estimates (2.1), (2.2), and (2.3) hold also for Bayes risks ϱ and $\tilde{\varrho}$, the only difference being that in this case $\overline{I(p : \tilde{p}/\theta)}$ and $\overline{I(\tilde{p} : p/\theta)}$ are defined as follows:

$$\overline{I(p:\tilde{p}/\theta)} = \sum_{n=1}^{\infty} \int \underset{Q_n^*}{\ldots} \int p(x_1, \ldots, x_n/\theta)\ \ln \frac{p(x_1, \ldots, x_n/\theta)}{\tilde{p}(x_1, \ldots, x_n/\theta)}\, dx_1 \ldots dx_n$$

$$\overline{I(p:\tilde{p}/\theta)} = \sum_{n=1}^{\infty} \int \underset{\tilde{Q}_n^*}{\ldots} \int \tilde{p}(x_1, \ldots, x_n/\theta)\ \ln \frac{\tilde{p}(x_1, \ldots, x_n/\theta)}{p(x_1, \ldots, x_n/\theta)}\, dx_1 \ldots dx_n\,, \tag{2.7}$$

where Q_n^* and $\tilde{Q}_n^*$ are optimal stop sets for the densities $p(x_1, \ldots, x_m, \theta)$ and $p(x_1, \ldots, x_m, \theta)$, respectively.

Consider reduction of observed data $x_1, \ldots, x_m$. We pass from the values $x_1, \ldots, x_m$ to the values $x_1', x_2', \ldots, x_m'$, $x_i' = \varphi_i(x_i', \ldots, x_{i-1}', x_i)$ having joint probability density $p'(x_1', \ldots, x_m', \theta)$ induced by the original density $p(x_1, \ldots, x_m, \theta)$.

In the original space θ and $x_1, x_2, \ldots, x_m, \ldots$ consider the following density:

$$\tilde{\tilde{p}}(x_1, \ldots, x_m, \theta) = p'(\theta/x_1', \ldots, x_m')\, p(x_m/x_1', \ldots, x_{m-1}') \times$$
$$\times p(x_{m-1}/x_1', \ldots, x_{m-2}') \ldots p(x_2/x_1') p(x_1/x_1') p(x_1') \tag{2.8}$$

where $x_i' = \varphi_i(x_i', \ldots, x_{i-1}', x_i)$; $m = 1, 2, \ldots$

Under distribution $\tilde{\tilde{p}}(x_1, \ldots, x_m, \theta)$ statistics $T_m = (x'_1, \ldots, x'_m)$ is sufficient and transitive.

Let Q_n^* be stop sets for the optimal sequential decision rule corresponding to the probability distribution $p(x_1, \ldots, x_m, \theta)$. If Q_n^* is preserved, and $\tilde{\tilde{p}}(x_1, \ldots, x_m, \theta)$ (2.8) is substituted for $p(x_1, \ldots, x_m, \theta)$, risk $\tilde{\tilde{R}}$ satisfies inequalities (2.1), (2.2) and (2.3) for mean risks. In this case:

$$\overline{I(p:\tilde{\tilde{p}})} = \sum_{m=1}^{\infty} \int \underset{Q_m^*}{\ldots} \int p(x_1, \ldots, x_m \theta) \left[\ln \frac{p(\theta/x_1, \ldots, x_m)}{p'(\theta/x'_1, \ldots, x'_m)} + \right.$$

$$\left. + \ln \frac{p(x_m/x_1, \ldots, x_{m-1})}{p(x_m/x'_1, \ldots, x'_{m-1})} + \ldots + \ln \frac{p(x_2/x_1)}{p(x_2/x'_1)} \right] dx_1 \ldots dx_m \, d\theta = \tag{2.9.}$$

$$= J(\theta, (x_1, \ldots, x_n)) - J(\theta, (x'_i, \ldots, x'_n)) +$$

$$+ \overline{J(x_m, (x_1, \ldots, x_{m-1}))} - \overline{J(x_m, (x'_1, \ldots, x'_{m-1}))}.$$

The two former terms $-J(\theta, (x_1, \ldots, x_n))$ and $J(\theta, (x'_1, \ldots, x_n))$ — represent Shannon's amount of information about θ in the original sequential sample $(x_1, \ldots, x_n)$, and that in the reduced sample $(x'_1, \ldots, x'_n)$, where n is the time when observations are stopped. The latter two terms in (2.9) give the mean increase of Shannon's amount of information about x_m contained in $(x_1, \ldots, x_{m-1})$ as compared with that contained in $(x'_1, \ldots, x'_{m-1})$, taken over all the samples $(x_1, \ldots, x_m)$ where decision is made not earlier than on mth step.

It follows from (2.9) that

$$\overline{J(x_m, (x_1, \ldots, x_{m-1}))} - \overline{J(x_m, (x'_1, \ldots, x'_{m-1}))} = \bar{n} \Delta J_*$$

where $\bar{n}$ is the mathematical expectation of the number of observations, and ΔJ_* is the value of $\ln [p(x_m \mid x_1, \ldots, x_{m-1})/p(x_m \mid x'_1, \ldots, x'_{m-1})]$ for a mean sample $(x_1, \ldots, x_m)$ having $(x_1, \ldots, x_{m-1})$ in the domain of continued observations. If divergencies

$$\int_{-\infty}^{+\infty} p(x_m/x_1, \ldots, x_{m-1}) \ln \frac{p(x_m/x_1, \ldots, x_{m-1})}{p(x_m/x'_1, \ldots, x'_{m-1})} dx_m \leq \Delta J = \text{const}.$$

then

$$\overline{J(x_m, (x_1, \ldots, x_{m-1})} - \overline{J(x_m, (x'_1, \ldots, x'_{m-1}))} \leq \bar{n} \, \Delta J.$$

These terms characterize losses of information concerning further observations that is necessary for the optimal choice of observation stop time.

Under $\tilde{\tilde{p}}(x_1, \ldots, x_m, \theta)$ stop sets Q_n^* are not optimal, thus the Bayes risk $\tilde{\tilde{\varrho}} \leq \tilde{\tilde{R}}$. Since, under distribution (2.8), statistics $T_m = (x_1', \ldots, x_m')$ is sufficient and transitive, the optimal rule for density $\tilde{\tilde{p}}(x_1, \ldots, x_m, \theta)$ (2.8) is constructed only through $x_1', \ldots, x_m'$ without recourse to more complete observation results $x_1, \ldots, x_m, \ldots$ For a sequential rule optimal with respect to a reduced problem, the risk $\varrho' = \tilde{\tilde{\varrho}} \leq \tilde{\tilde{R}}$. (We have also $\varrho = R \leq \varrho' \leq \tilde{\tilde{R}}$.) Therefore for the increase of ϱ' as compared with ϱ the following estimates hold:

$$\varrho' - \varrho \leq \sqrt{2\,\varrho'(W^2)\,(\overline{\Delta J}_1 + \overline{\Delta J}_2)}\,; \tag{2.10}$$

$$\overline{\Delta J}_1 + \overline{\Delta J}_2 \leq \frac{1}{W_0}\left[\varrho \ln \frac{\varrho}{\varrho'} + (W_0 - \varrho) \ln \frac{W_0 - \varrho}{W_0 - \varrho'}\right]; \tag{2.11}$$

$$\varrho' - \varrho \leq \sqrt{2\varrho W_0(\overline{\Delta J}_1 + \overline{\Delta J}_2)}\,, \tag{2.12}$$

where

$$\overline{\Delta J}_1 = J\big(\theta,(x_1,\ldots,x_n)\big) - J\big(\theta,(x_1',\ldots,x_n')\big);$$

$$\overline{\Delta J}_2 = \overline{J\big(x_m,(x_1,\ldots,x_{m-1})\big)} - \overline{J\big(x_m,(x_1',\ldots,x_{m-1}')\big)};\ m \leq n\,.$$

n is the time of observation stop (optimal for probability densities $(p(x_1,\ldots,x_m,\theta))$.

3. Construction of suboptimal sequential decision rules and estimation of their closeness to the optimal ones

A conditional risk function $S(Z; m)$ may be introduced into Bayes sequential analysis problems. It is the mathematical expectation of losses for a sequential decision rule, providing m observations were made and a value $Z(x_1, \ldots, x_m) = Z$ was obtained. As it was mentioned above, if statistics $Z(x_1, \ldots, x_m) = Z_m$ is sufficient and transitive, the optimal sequential decision rule uses only the current value of Z_m. In this case, the optimal decision rule is to stop observations when $Z(x_1, \ldots, x_m)$ gets into the domain R_m defined by $S(Z; m) = W(Z; m)$. The domain of continued observations C is defined by $S(Z; m) < W(Z; m)$; $m = 1, 2, 3, \ldots$ where $W(Z; m)$ is the conditional mathematical expectation of losses upon stopping observations and taking the final decision at time m, provided that $Z(x_1, \ldots, x_m) = Z$. Function $S(Z; m)$ of the optimal decision rule satisfies the following recurrent equation [5, 6]:

$$S(Z; m) = \min\left[W(Z; m) \int_{-\infty}^{\infty} S(Z_{m+1};\ m+1)\, P(Z_{m+1}/Z; m)\, d\, Z_{m+1}\right]. \tag{3.1}$$

Equation (3.1) makes it possible to find effectively the optimal sequential decision rule for the case where dimensionality of statistics $Z(x_1, \ldots, x_m)$ is

comparatively small and does not increase with the number of observations. If the number of observations is bounded by a value of T, then $S(Z; T) \equiv \equiv W(Z; T)$, and Eq. (3.1) enables one to find $S(Z; T-1), \ldots, S(Z; 1)$. For a wide variety of problems, a T-truncated rule turns under $T \to \infty$ into the optimal non-truncated one [5, 6].

Consider now a reduction consisting in transition from a multidimensional sufficient and transitive statistics $Z(x_1, \ldots, x_m)$ to a statistics of lesser dimensionality $T_m = \varphi_m(Z_m)$ which may be neither sufficient, nor transitive. Consider the class of sequential decision rules where decision is made on the basis of values of T_m. If T_m is a Markov process ($m = 1, 2, \ldots$), the sequential rule (optimal for the given class) may be obtained through the recurrent equation (3.1) taking the following form:

$$\widetilde{S}(T;m) = \min\left[\widetilde{W}(T; m); \int_{-\infty}^{\infty} \widetilde{S}(T_{m+1}; m+1)\, P\,(T_{m+1}/T; m)\, d\,T_{m+1}\right], \tag{3.2}$$

where $\widetilde{S}$ and $\widetilde{W}$ are conditional mathematical expectations under continuation and stop of observations, respectively, for a decision rule using value of $(T_m; m)$.

One can estimate the Bayes risk increase caused by transition from multidimensional and transitive statistics Z_m to Markov statistics T_m by inequalities (2.10), (2.11) and (2.12) where

$$\overline{\Delta J_1} = J(\theta; Z_n) - J(\theta; T_n); \overline{\Delta J_2} = \overline{J(Z_m, Z_{m-1})} - \overline{J\big(Z_m, (T_1, \ldots, T_{m-1})\big)}. \tag{3.3}$$

Consider also the case where T_m is a close-to-Markov stochastic process, i.e. the process of weak aftereffect. If a decision rule obtained through Eq. (3.1) is applied to a non-Markov process T_m, the change of mean risk can be estimated by inequalities (2.1), (2.2) and (2.3). In this case the following holds:

$$\begin{aligned}
\overline{I(p{:}\tilde{p}\,|\,\theta)} &= \sum_{n=1}^{\infty} \int \underset{Q_n^*}{\ldots} \int \prod_{i=1}^{n-1} p(T_{i+1}|T_i; \theta)\, p(T|\theta) \times \\
&\times \ln \frac{\prod\limits_{i+1}^{n-1} p(T_{i+1}|T_i; \theta)\, p(T_1|\theta)}{p(T_1, \ldots, T_n|\theta)}\, dT_i \ldots dT_n, \\
\overline{I(\tilde{p}{:}p|\theta)} &= \sum_{h=1}^{\infty} \int \underset{Q_n^*}{\ldots} \int p(T_1, \ldots, T_n|\theta) \times \\
&\times \ln \frac{p(T_1, \ldots, T_n|\theta)}{\prod\limits_{i=1}^{n-1} p(T_{i+1}|T_i; \theta)\, p(T_1|\theta)}\, d\,T_1 \ldots d\,T_n.
\end{aligned} \tag{3.4}$$

Here $p(T_1, \ldots, T_m \mid \theta)$, the true joint probability density for $T_1, \ldots, T_m$, corresponds to $\tilde{p}$, and p is substituted by $\prod_{i=1}^{m-1} p(T_{i+1} \mid T_i, \theta) P(T_1 \mid \theta)$ — the joint probability density for $T_1, \ldots, T_m$ which should have taken place, had these random variables constituted a Markov chain with probabilities $p(T_i \mid \theta)$ and $p(T_{i+1} \mid T_i; \theta)$, $i = 1, 2, \ldots$ correlated with the true probability $p(T_1, \ldots, T_m \mid \theta)$. Denote by δ and $\tilde{\delta}$ the optimal rules for densities p and $\tilde{p}$, respectively. Assume that the informational divergencies are bounded as follows:

$$\int_{-\infty}^{\infty} p(T_i|T_1, \ldots, T_{i-1}; \theta) \ln \frac{p(T_i|T_1, \ldots, T_{i-1}; \theta)}{P(T_i|T_{i-1}; \theta)} dT_i \leq \Delta I_1 ,$$

under $(T_1, \ldots, T_{i-1}) \in B_{i-1}(\delta)$, $B_{i-1}(\delta)$ being the set of samples $(T_1, \ldots, T_{i-1})$ where decision is taken by rule δ not before the ith step. Then in (3.4):

$$\overline{I(\tilde{p}:p|\theta)} \leq \bar{\tilde{n}}(\theta) \Delta I_1 , \tag{3.5}$$

where $\bar{\tilde{n}}(\theta)$ and $\bar{n}(\theta)$ are the mathematical expectations of the number of observations required by decision rule δ for densities $\tilde{p} = p(T_1, \ldots, T_n \mid \theta)$ and $p = \prod_{i=1}^{m-1} p(T_{i+1} \mid T_i, \theta) p(T_1 \mid \theta)$, respectively. Similarly find

$$\Delta I_2 \geq \int_{-\infty}^{\infty} p(T_i|T_{i-1}; \theta) \ln \frac{p(T_i|T_{i-1}; \theta)}{p(T_i|T_1, \ldots, T_{i-1}; \theta)} d T_i$$

for $(T_1, \ldots, T_{i-1}) \in B_i(\delta)$. Then $\overline{I(p : \tilde{p} \mid \theta)} \leq \bar{n}(\theta) \Delta I_2$.

The difference of Bayes risks $\varrho' = R(\delta)$ and $\varrho = \tilde{R}(\tilde{\delta})$ is estimated through (3.4) in a similar way.

For process T_m with small aftereffect, the values of ΔI_{i_1} and ΔI_2 are small. If one regards approximately T_m as a Markov process rejecting thus its "history" $T_1, \ldots, T_{m-1}$, the increase of risk due to such a sub-optimal rule is bounded by the inequalities above:

$$\text{for } \Delta I_1, \Delta I_2 \to 0 \quad \text{obtain} \quad | \varrho' - \varrho | \to 0.$$

4. Application examples

The informational estimates obtained may be used to find potential accuracy of statistical decision rules. They enable one also to judge whether one or another of data reductions is possible under a given restriction on the allowable risk increase. In doing so, losses of information (2.9) and (3.3) may

be evaluated by computer or, sometimes, manually. Thus, accuracies of sub-optimal and optimal sequential decision rules may be compared without construction of optimal decision rules through solution of recurrent risk equation (3.1), the latter, in the case of many variables, being cumbersome even for the computer.

Show, as an example, signal detection in a multichannel system. Let H_k be a hypothesis about presence of a signal in channel k, and H_0 be a hypothesis about signal absence in all channels. Signal may be only in one channel at a time. An *a priori* signal absence probability is 1/2; and presence of signal is equally probable for all channels. At the output of channel k at time m a variable $x_k(m)$ is observed having gaussian distribution with variance σ^2 and mathematical expectation equal to zero and a, respectively, under the absence and presence of signal in the channel. The random variables $x_k(m), k = 1, \ldots, N$, are independent. After m observations, the logarithm of likelihood ratio between H_k and its negation $\overline{H}_k$ is as follows:

$$Z_k(m) = \frac{a}{\sigma^2} \sum_{j=1}^{m} x_k(j) - \frac{ma^2}{2\,\sigma^2}.$$

Statistics $(Z_1, \ldots, Z_N)$ is transitive and sufficient for the sequential discrimination of hypotheses $H_0(\theta = \theta_0)$ and $\overline{H}_0 = \bigcup_{k=1}^{N} H_k(\theta = \theta_1)$.

Consider reduction of multidimensional statistics $(Z_1, \ldots, Z_N)$ to a one-dimensional one

$$Z(m) = \max_{1 \leq k \leq N} Z_k(m). \tag{4.1}$$

Shannon's amounts of information about parameter $\theta(\theta = \theta_0, \theta_1)$ contained in $(Z_1(m), \ldots, Z_N(m))$ and $Z(m)$ may be readily computed. Denote these amounts by $J(m)$ and $J'(m)$, respectively. Their graphs are plotted in Fig. 1. (Integrals representing $J(m)$ and $J'(m)$ were defined by the computer M—20 using the method of Monte Carlo.) Figure 2 shows graphs of informational divergences between hypotheses H_0 and $\overline{H}_0$:

$$I_1 = I(H_0 : \overline{H}_0) = \int_{-\infty}^{\infty} p(Z \mid H_0) \ln [p(Z \mid H_0)/p(Z \mid \overline{H}_0)]\, dZ;$$

$$I_2 = I(\overline{H}_0 : H_0),$$

when using multidimensional variable $(Z_1, \ldots, Z_N)$ as statistics Z, and one-dimensional statistics $Z(m) = \max_{1 \leq k \leq N} Z_k(m)$. These results show that $Z(m)$, though not being sufficient, contains nearly complete information about θ.

Inequality (2.3) enables one to establish the lower bounds of error probabilities or of mean number of observations for discrimination of many simple and complex hypotheses.

While non-Markov, process $Z(m) = \max_{1 \leq k \leq N} Z_k(m)$ has low aftereffect. For the sequential decision rule based on $Z(m)$ of (4.1), the continuation and stop

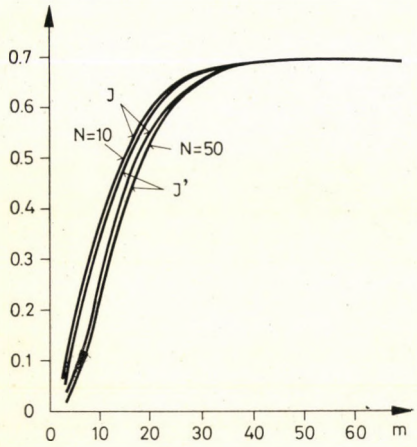


Fig. 1

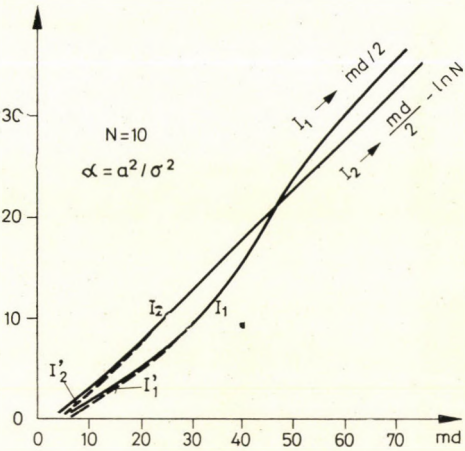


Fig. 2

domains were defined approximately by means of Eq. (3.1) as is indicated in Section 3. The decision rule is to compare $Z(m)$ with variable thresholds $A(m)$ and $B(m)$ (computed in advance on the basis of Eq. (3.1) solution): if $B(m) < Z(m) < A(m)$ observations go on; if $Z(m) \geq A(m)$ hypothesis $\overline{H}_0$ is accepted, if $Z(m) \leq B(m)$, H_0 is accepted.

Error probabilities for the obtained sub-optimal rule based on statistics Z_m of (4.1) were compared with the lower bounds of error probability in the absence of reduction (see also examples in [4, 6, 7, 8]). The comparison has born out that the constructed decision rule is close to the optimal one.

Transition from continuous observation of stochastic processes to discrete one in problems of statistical decisions and optimal control is another example of reduction.

In the case of gaussian or asymptotically gaussian statistics, evaluation (or finding of upper bounds) of (2.9) and (3.3) is simpler. In this case, $\Delta \bar{J}_2$ is written in the terms of variance of next value prediction error for a sufficient and transitive statistics Z_{m+1} based on observations of its preceeding value Z_m and reproduced data $Z'_1, \ldots, Z'_m$:

$$\overline{\Delta J_2} = \overline{J(Z_{m+1}, Z_m)} - \overline{J(Z_{m+1}, (Z'_i, \ldots, Z'_m))} = \\ = \frac{k}{2} \overline{\ln \frac{\det \| B_{ij}(Z'_1, \ldots, Z'_m \|}{\det \| B_{ij}(Z_m) \|}} \tag{4.2}$$

where k is statistics' Z_{m+1} dimensionality; det $\| B_{ij}(Z_m) \|$ and det $\| B_{ij}(Z'_1, \ldots \ldots, Z'_m \|$ are determinants of error variance matrix for the best prediction of Z_{m+1}, based on observations, respectively, of Z_m and $Z'_1, \ldots, Z'_m$. In (4.2), as above, mathematical expectation is taken over all $x_1, \ldots, x_m$ where decision is taken not before step m.

To compute Shannon's amount of information (with respect to a continuous parameter θ) contained in a sequential sample $x_1, \ldots, x_n$, one can make use of asymptotical sufficiency and asymptotical normality (under $n \to \infty$) of maximum likelihood estimate $\hat{\theta}(x_1, \ldots, x_n)$. For independent observations obtain asymptotically in this case:

$$J((x_1, \ldots, x_n), \theta) = \frac{1}{2} \int p(\theta) \ln \left[(\overline{n}(\theta))^r \det \| I_{ij}(\theta) \| \right] d\theta + \text{const}\left[p(\theta)\right],$$

where r is the dimensionality of parameter θ; $\overline{n}(\theta)$ is the mathematical expectation of number of observations in the function of θ. $J((x'_1, \ldots, x'_m), \theta)$ is written similarly with $\| I'_{ij}(\theta) \|$ being substituted for $\| I_{ij}(\theta) \|$, where $\| I_{ij}(\theta) \|$ and $\| I'_{ij}(\theta) \|$ are Fisher's information matrices corresponding to observation of original variable x and reduced one x'. The difference $J - J'$ is expressed through the logarithm of determinants' $\| I_{ij} \|$ and $\| I'_{ij} \|$ ratio.

In papers by A. Rényi [9] and I. Vajda [10] asymptotic estimates were obtained for Shannon's amount of information in $x_1, \ldots, x_n$ with respect to a discrete parameter θ having a finite number of values $\theta_1, \ldots, \theta_k$ under fixed number of independent observations n:

$$J((x_1, \ldots, x_n), \theta) = H - \frac{\lambda^n}{\sqrt{n}} \text{const};$$

$$\lambda = \max_{i \neq j} \min_{0 \leq a \leq 1} \int_{-\infty}^{\infty} (p(x \mid \theta_i))^a \, (p(x \mid \theta_j))^{1-a} \, dx;$$

$$H = - \sum_{q=1}^{k} P(\theta = \theta_q) ln \, P(\theta = \theta_q).$$

References

1. *Perez, A.*, Information, ε-Sufficiency and Data Reduction Problems. Kybernetika (1965) 4; (1967) 1.
2. *Perez, A.*, Information theory methods in reducing complex decision problems. In: Trans. of the Fourth Prague Conference on Information Theory, Random Processes and Statistics. Prague, 1965. Academia, Prague, 1967, pp. 55—87.

3. *Bahadur, R. R.*, Sufficiency and Statistical Decision Functions. Ann. of Math. Stat. **25**, *3* (1954) 423—462.
4. *Хазен Э. М.* Информационные оценки для риска в задачах последовательного анализа. Изв. АН СССР, Техническая кибернетика, **6** (1969) 126—140.
5. *Wald, A.*, Statistical Decision Functions. N. I., 1950.
6. *Хазен Э. М.* Методы оптимальных статистических решений и задачи оптимального управления. М., Сов. радио, 1968.
7. *Бененсон З. М., Хазен Э. М.* Методы статистического последовательного анализа в задачах распознавания многих гипотез. Изв. АН СССР, Техническая кибернетика, (1966) *4*.
8. *Хазен Э. М.* Определение потенциальной точности решения некоторых задач распознавания и оценивания. Изв. АН СССР, Техническая кибернетика (1971) *3*, 184—192.
9. *Rényi, A.*, On some problems of statistics from the point of view of information theory. In: Proceedings of the Colloquium on Information Theory. Debrecen (Hungary), 1967, 343—357.
10. *Vajda, I.*, On the Amount of Information Contained in a Sequence of Observations. Kybernetika, (1970) 5.

Информационные оценки для возрастания риска при редукции наблюдаемых данных в задачах последовательного анализа

Э. М. ХАЗЕН

(Москва)

Пусть $x_1, \ldots, x_m, \ldots$ — результаты последовательно выполняемых наблюдений; редуцированные величины $x_1', \ldots, x_m', \ldots$ определяются по ним, $x_i' = \varphi_i(x_1', \ldots, x_{i-1}', x_i)$, где φ_i — некоторые функции. Выводятся оценки для возрастания бейесовского риска ϱ' для статистического последовательного решающего правила, использующего только редуцированные значения $x_i', \ldots, x_m', \ldots$, по сравнению с бейесовским риском ϱ для последовательного правила, использующего полные результаты наблюдений $x_1, \ldots, x_m \ldots$ Возрастание риска ограничивается неравенствами (2.10)—(2.12), связывающими изменение риска и изменение шенноновского количества информации. При этом $\overline{\Delta J_1} = J(\theta, (x_1, \ldots, x_n)) - J(\theta, (x_1', \ldots, x_n'))$ представляет собой величину потерь (вследствие редукции) шенноновского количества информации, содержащейся в наблюдаемой выборке относительно оцениваемого параметра θ, а $\overline{\Delta J_2} = \overline{J(x_m, (x_1, \ldots, x_{m-1}))} - \overline{J(x_m, (x_1', \ldots, x_{m-1}'))}$ характеризует среднее значение потерь шенноновского количества информации, содержащейся в полученных результатах наблюдений относительно результата следующего наблюдения. (θ — параметр, о котором должно приниматься окончательное решение; W — весовая функция потерь; неравенство (2.11) выполняется, если $W \leq W_0 \equiv \text{const}$).

Полученные оценки представляют собой развитие результатов работ А. Переза [1, 2], который рассмотрел задачи оптимальных статистических решений с фиксированным заранее числом наблюдений $n \equiv \text{const}$. В отличие от случая $n \equiv \text{const}$ в задачах статистического последовательного анализа риск может возрастать не только из-за нарушения условия достаточности статистики $(x_1', \ldots, x_m')$ для оценки параметра θ, но также из-за нарушения транзитивности, связанного с потерей информации о будущем течении процесса, которая должна использоваться для выбора момента остановки наблюдений.

Рассматривается построение оптимальных и подоптимальных последовательных решающих правил. Результаты работы позволяют в ряде случаев оценивать близость рисков подоптимальных последовательных решающих правил и оптимальных, а также находить оценки предельной потенциальной точности статистических решений.

Э. М. Хазен

Московский Государственный университет им. М. В. Ломоносова, механико-математический факультет, кафедра теории вероятности

СССР, Москва В-234

Ленинские горы

Problems of Control and Information Theory, Vol. 3 (1), pp. 47—62 (1974)

OPTIMAL LINEAR RANDOMIZED FILTRATION OF QUANTUM BOSON SIGNALS

V. P. BELAVKIN

(Moscow)

(Received June 16, 1972)

An equation defining the optimal linear randomized estimation $z = Ay + \eta$ of non-observables x based on indirect non-ideal measurement of the received quantum (boson, non-Gaussian) signal y is derived. Its solutions for several important cases are presented. The linear filtration of the stationary boson signals is discussed, and an example is given illustrating the stationary signal filtration at the output of the non-ideal wave communication line.

The progress of communications on super-high and optical frequencies where the quantum nature of the electromagnetic signal becomes essential, makes it necessary to develop the quantum theory of optimal signal processing [1]. Within the framework of the theory, the received signal y is described by a set of non-commuting observables $\{y_t, t \in T\}$ which are considered below as real: $y_t^* = y_t$. The fact that the observables are non-commuting with respect to the multiplication ($y_t y_{t'} \neq y_{t'} y_t$) is an adequate mathematical interpretation of their physical incompatibility (i.e. impossibility of their arbitrarily precise measurement). In the quantum physics, the non-commuting real observables are represented by linear Hermitian operators in the Hilbert space of quantum-mechanical states. Such a specific representation, however, is not required in this paper.

The non-commutating generalization of communication theory problems is naturally proceeded with the simplest optimization problem of boson[1] signal linear processing based on the linearity of the optimal estimation in the Gaussian case, proved in [2].

The problem of finding the optimal set of commuting observables $z^0 = \gamma^0(y)$ whose precise measurement gives estimates $\{z_1^0, \ldots, z_m^0\} \equiv z^0$ minimizing the mean square losses (1.3) in the class of linear transforms $\gamma(y) =$

[1] The quantum signal y is referred to as boson if it is described by secondary commuting observables $\{y_t\}$, i.e. observables commuting with their own commutators $y_t y_{t'} - y_{t'} y_t$. In the quantum field theory, particles and fields adhering to the Bose—Einstein statistics are described by secondary commuting operators of generation and destruction. The electric and magnetic field strengths constitute an important, for the communication theory, example of secondary commuting real observables.

$= Ay + \eta$ of the received boson non-Gaussian signal y was formulated and studied in [3]. In general, however, the solution does not meet the correspondence principle, that is to say it does not turn into the solution of the corresponding classical (i.e. commutative) problem under approximation to the weak non-commutativity $\hbar \to 0$. This difficulty may be overcome by means of estimation algorithm randomization corresponding to the expansion the class of measurements under consideration from direct to indirect ones. The indirect measurements are done through the direct measurements of a set of commuting observables of an enlarged system comprising the initial system (accessible to observations) and an additional, usually independent one (measuring apparatus). Owing to such a randomization it becomes possible to consider joint quasi-measurements of non-commuting observables of the initial system [4, 5] realized through an indirect measurements. Making use of this concept, derive an equation for the optimal randomized algorithm of quantum signal linear filtration meeting the correspondence principle.

1. Derivation of the optimal linear randomized filtration equation

For the sake of simplicity, confine ourselves at first to the finite-dimensional signal $y = \{y_j, j = 1, \ldots, n\}$ describing observable sequence of non-commuting physical values $y_1, \ldots, y_n$. We have to find among transformations $z = \gamma(y)$ of linear kind

$$z = Ay + \eta \equiv \left\{ \sum_{j=1}^{n} a_{ij} y_j + \eta_i, \quad i = 1, \ldots, m \right\} \tag{1.1}$$

satisfying the commutativity condition

$$[z, z^T] = \| z_i z_j - z_j z_i \| = 0 \tag{1.2}$$

the optimal one $z^0 = A^0 y + \eta^0$ in the terms of the minimal mean square risk criterion customary to the linear theory

$$R = \frac{1}{2} M(z - x)^T G(z - x) \equiv \frac{1}{2} M \sum_{i,j} (z_i - x_i) g_{ij} (z_j - x_j). \tag{1.3}$$

Hereafter the following matrix notation will be used: z, z^T are column vector and row vector, respectively, consisting of commuting observables $\{z_1, \ldots, z_m\}$ giving, as the result of joint precise measurement, numerical estimates $\{z_1, \ldots, z_m\}$ of estimated $x^T = (x_1, \ldots, x_m)$; $G = \| g_{ij} \|$ is a real symmetric positively defined matrix $m \times m$, and M is expectation.

For the quantum randomization of the linear estimation algorithm (1.1) given by the $m \times n$ matrix $A = \| a_{ij} \|$ and the m-dimensional vector $\eta = \{\eta_i, i = 1, \ldots, m\}$, the components of the latter $\eta_1, \ldots, \eta_m$ should be regarded as non-commuting y-independent stochastic variables. Substituting (1.1) into (1.2) and taking into account that the independence condition for y and η implies their commutativity $[y, \eta^T] = 0$, write the commutativity condition for the observed variables z in the following matrix form:

$$A[y, y^T] A^T + [\eta, \eta^T] = 0. \tag{1.4}$$

One can easily see that this condition may be met only if the commutators $[y_i, y_j]$ and $[\eta_i, \eta_j]$ commute with y and η i.e. are c-numbers

$$[y, y^T] = \| c_{ij} \| \equiv C_y, \quad [\eta, \eta^T] = \| c_{ij} \| = C_\eta. \tag{1.5}$$

The commutation matrix C_η of quantum observables η for each A is defined quite uniquely as

$$C_\eta = -AC_yA^T. \tag{1.6}$$

It follows from the fact that operators describing observables y and η act in different Hilbert state spaces of the main and an additional systems, and that equality of the commutators $[\eta, \eta^T] = -A[y, y^T]A^T$ may take place only if they are multiples of the unit operator. Further we shall confine ourselves only to the secondary commuting observables y having known commutation (imaginary, anti-symmetric) matrix C_y giving with (1.6) quantization of the vector η for each A. This quantum randomization of the linear estimation (1.1) enables us to look through arbitrary matrices A rather than only through those satisfying the condition $AC_yA^T = 0$ as it was done in [3] for non-randomized linear estimation.

Substituting $z = Ay + \eta$ into (1.3) and averaging it, write the mean square risk allowing for independence of the observables η of x and y as the following sum:

$$R = \frac{1}{2} \operatorname{Tr} GM(Ay + \eta - x)(Ay + \eta - x)^T = R_0(A) + R_\eta(A), \tag{1.7}$$

where

$$R_0(A) = \frac{1}{2} \operatorname{Tr} G(K_x, - AK_{xy}^T - K_{xy}A^T + AK_yA^T), \tag{1.8}$$

$$R_\eta(A) = \frac{1}{2} \operatorname{Tr} G\big(K_\eta - (m_\eta - m_x + Am_y)(m_\eta - m_x + Am_y)^T\big). \tag{1.9}$$

Here m_x, m_y, m_η are expectations of x, y, η, respectively, and K_x, K_{xy}, K_y, K_η are respective symmetrized correlation matrices of the variables x, y, and η:

$$K_x = M\frac{1}{2}[(x-m_x),(x-m_x)]_+,\ K_{xy} = M\frac{1}{2}[(x-m_x),(y-m_y)^T]_+,$$

$$K_y = M\frac{1}{2}[y-m_y),(y-m_y)^T]_+,\ K_\eta = M\frac{1}{2}[(\eta-m_\eta),(\eta-m_\eta)^T]_+,$$

(the brackets $[\,.\,,\,.\,]_+$ define the anti-commutator $[y, y^T]_+ = ||\, y_i y_j + y_j y_i ||$).

Risk (1.7) minimization in the class of linear transformations $Ay + \eta$ may be done in two steps: first, for a fixed A, the minimum of losses (1.9) is found having the sense of indirect measurement error of non-commuting $u = Ay$, and second, minimization with respect to A is done. Minimization of losses (1.9) in the terms of statistical states of the boson vector η having bounded below quantum entropy $S_\eta \geq S$ was carried out in [5]. The optimal state thus defined is Gaussian and characterized by the expectation vector m_η and symmetrized correlation matrix K_η

$$m_\eta = m_x - Am_y,\ K_\eta = \frac{1}{2}C_\eta \operatorname{cth}\frac{1}{2\theta}GC_\eta. \tag{1.10}$$

Parameter $\theta \geq 0$ is defined from equation $S_\eta(\theta) = S$, where

$$S_\eta(\theta) = \frac{1}{2}\operatorname{Tr}\left(\frac{1}{2\theta}GC_\eta \operatorname{cth} GC_\eta - \ln\left|2\operatorname{sh}\frac{1}{2\theta}GC_\eta\right|\right)$$

is the entropy of boson Gaussian vector η corresponding to the temperature θ. The function $S_\eta(\theta)$ may be easily seen to be monotonically increasing, and the zero temperature $\theta = 0$ corresponds to the zero entropy $S = 0$. Below, for the sake of convenience, the temperature θ, rather than the entropy S, will be regarded as the independent parameter characterizing degree of linear estimation randomization. The randomized linear estimation (1.1) having the temperature θ is interpreted in [5] as indirect measurement of secondary commuting Ay by the optimal linear measuring apparatus with phase coordinates η and non-ideality degree given by the temperature θ.

The minimal indirect measurement error for fixed θ and with due regard to (1.10), (1.6) is as follows

$$R_\eta^\theta(A) = \frac{1}{2}\operatorname{Tr} GAC_yA^T \operatorname{cth}\frac{1}{2\theta}GAC_yA^T \equiv \frac{1}{2}\operatorname{Tr} f_\theta\left(\frac{1}{2}GAC_yA^T\right) \tag{1.11}$$

and increases monotonically with θ. The absolute error minimum (1.9) attained under absence of limitations on entropy and corresponding to the ideal indi-

rect measurement of the observables $u = Ay$ is obtained from (1.11) through the passage to the limit $\theta \to 0$:

$$R_\eta^0(A) = \lim_{\theta \to 0} \frac{1}{2} \operatorname{Tr} f_\theta \left(\frac{1}{2} GAC_y A^T \right) = \frac{1}{2} \operatorname{Tr} \left| \frac{1}{2} GAC_y A^T \right| \tag{1.12}$$

where the fact that the limit of $f_\theta(\alpha) = \alpha \operatorname{cth} \frac{1}{\theta} \alpha$ with respect to the parameter $\theta \to 0$ equals to the module $|\alpha|$ is allowed for. In the classical (i.e. commutative $C_y = 0$) case, the losses (1.12) turn into zero and the linear estimation (1.1) becomes for $\theta \to 0$ non-randomized, that is the vector η is identified with its expectation. For the quantum case, the correlation matrix K_η^θ defining the error (1.11) is not equal to zero even for $\theta = 0$, and in this sense the linear estimation is still randomized even under the ideal indirect measurement.[2]

Derive an equation defining the optimal matrix A^θ minimizing the risk $R^\theta(A) \equiv R(A) + R_\eta^\theta(A)$ under the fixed θ. The functional

$$R^\theta(A) = \frac{1}{2} \operatorname{Tr} \left[G(K_x - AK_{xy}^T - K_{xy}A^T + AK_yA^T) + f_\theta \left(\frac{1}{2} GAC_yA^T \right) \right] \tag{1.13}$$

is convex and by virtue of its analiticity allows differentiation by A^T. Equating variation

$$\delta R^\theta(A) = \operatorname{Tr} \left[G(AK_y - K_{xy}) + f'_\theta \left(\frac{1}{2} GAC_yA^T \right) \frac{1}{2} GAC_y \right] \delta A^T$$

to zero and taking into consideration independence of variations δA^T, obtain the following equation

$$G(AK_y - K_{xy}) + \frac{1}{2} f'_\theta \left(\frac{1}{2} GAC_yA^T \right) GAC_y = 0 \tag{1.14}$$

where $f'_\theta(\alpha) \equiv \frac{d}{d\alpha} f_\theta(\alpha)$ is a monotonic increasing odd function with values $|f'(\alpha)| \leq 1$:

$$f'_\theta(\alpha) = \operatorname{cth} \frac{\alpha}{\theta} \operatorname{cosech}^2 \frac{\alpha}{\theta} = \frac{\operatorname{sh}(2\alpha|\theta) - 2\alpha/\theta}{\operatorname{ch}(2\alpha/\theta) - 1}. \tag{1.15}$$

The matrix non-linear equation (1.14) is that desired. Its solution A_θ for each θ defines the optimal randomized linear estimation $u^\theta = Ay^\theta + \eta^\theta$ characterized by the temperature θ, the optimal Gaussian boson vector η^θ

[2] The ideal indirect measurement realized by measurement of linear superposition (1.1) of signal y and vacuum noise η is referred to as coherent, i.e. it may be described by a set of coherent measurement state vectors [1].

being described in the terms of the commutation matrix (1.6), the expectation vector and the correlation matrix (1.10) for $A = A^\theta$. The minimal risk defining the quality of the optimal linear filtration under indirect measurement is as follows:

$$R^\theta = \frac{1}{2}\operatorname{Tr}\left[G(K_x - A^\theta K_{xy}) + g_\theta\left(\frac{1}{2}\, GA^\theta C_y A^{\theta T}\right)\right] \tag{1.16}$$

$\big(g_\theta(\alpha) = f_\theta(\alpha) - \alpha f'_\theta(\alpha)\big)$, and is readily obtainable by multiplying (1.14) by A^T from the right and substituting the obtained identity into (1.13). The second term in (1.16) is due to the estimation randomization and under $\theta \to \infty$ increases as $m\theta/2$ (where m is the dimensionality of the vectors x and z). Under $\theta \to 0$ it vanishes, and for the Gaussian case risk (1.16) reaches the lower bound obtained in [2] through minimization in the class of *arbitrary* indirect measurements. This proves equivalence of linear and Gaussian approximations in optimization of boson signal processing.

2. Study of the optimal linear filtration equation

By applying the matrix identity $f(DA^T)D = Df(A^TD)$ to (1.15) and to matrix D equal to $\frac{1}{2}\, GAC_y$, write solution of (1.14) as follows:

$$A^\theta = K_{xy}(K_y + B_\theta)^{-1} \tag{2.1}$$

where B_θ standing for the matrix $\frac{1}{2}\, C_y f'_\theta\left(\frac{1}{2}\, A^{\theta T} G A^\theta C_y\right)$ is, for fixed θ, the solution of

$$B = \frac{1}{2}\, C_y f'_\theta\left(\frac{1}{2}\,(K_y + B)^{-1} K_{yx} G K_{xy}(K_y + B)^{-1} C_y\right) \tag{2.2}$$

with respect to B. By introducing $\alpha = \varphi_\theta(\,\cdot\,)$ which is the inverse function of (1.15), rewrite this matrix equation in a more suitable form:

$$\varphi_\theta(2\, C_y^{-1}B)\, 2\, C_y^{-1} K_y = (1 + K_y^{-1}B)^{-1} K_y^{-1} K_{yx} G K_{xy}(1 + K_y^{-1}B)^{-1}. \tag{2.3}$$

In its domain of definition $-1 \leq \beta \leq 1$, $\varphi_\theta(\beta)$ is odd analytical function increasing monotonically from $-\infty$ to ∞. The right part of (2.3) is monotonically decreasing non-negative bounded function of argument $2\, K_y^{-1}B$. Under such conditions, the right and left parts of (2.3) as independent functions of matrix B always have for $C_y \neq 0$ a single common point $B = B_0$ such that dependence of B_0^θ on $\theta > 0$ is monotonically non-increasing: $B_\theta \leq B'_\theta$ for $\theta \geq \theta'$ with the product $2C_y^{-1}B_\theta$ lying within the domain

$$-1 \leq 2C_y^{-1}B_\theta \leq 1. \tag{2.4}$$

Solution of (2.2) for $\theta \neq 0$ may be looked for in the form of a series $B_\theta = B^{(0)} + \frac{1}{\theta} B^{(1)} + \frac{1}{\theta^2} B^{(2)} + \ldots$ converging rapidly for $\theta \gg 1$ (strong estimation randomization). Taking into consideration that $f'(\alpha) = \frac{1}{3}\left(\frac{2\alpha}{\theta}\right) + 0\left(\frac{2\alpha}{\theta}\right)^\alpha$ find readily that

$$B^{(0)} = 0, B^{(1)} = \frac{1}{6} C_y A_0^T G A_0 C_y,$$

where $A_0 = K_{xy} K_y^{-1}$ is a matrix defining estimation (1.1) for the commutative case of $C_y = 0$. Under weak non-commutativity $C_y = \varepsilon S_y$ (where ε is a small parameter), $B^{(1)} = o(\varepsilon^2)$ and matrix A^θ has the classical form $A^0 = A_0$ not only in the zero approximation, but in the first quasi-classical one for $\varepsilon \ll 1 (\theta \neq 0)$ as well. Thus, the randomized linear extimation, as opposed to the non-randomized one of [3], satisfies the correspondence principle. Quantum corrections in (1.16) to the losses of the classical randomized estimation $R_0^\theta = \frac{1}{2} \mathrm{Tr}\left(G(K_x - AK_{xy}^T) + \theta\right)$ are for $\theta \neq 0$ on the order of ε^2.

In the case of the ideal indirect measurement $\theta = 0$, the matrix $B_0 \equiv \lim_{\theta \to 0} B_\theta$ realizes, evidently, the boundaries of the inequality (2.4). Consider solution of (2.2) at $\theta = 0$ for the following important case. Let the signal y having non-degenerate commutation matrix of rank $n = 2s$ consist of two groups of physically different observables $\{p_1, \ldots, p_s\} \equiv p$ and $\{q_1, \ldots, q_s\} \equiv p$ described, respectively, by matrices C_y and K_y having the following form

$$C_y = \begin{pmatrix} 0 & -iC \\ iC & 0 \end{pmatrix} \equiv C \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad K_y = \begin{pmatrix} K & 0 \\ 0 & K \end{pmatrix} = K \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \tag{2.5}$$

where the $s \times s$ matrices C and K are real, symmetric and positive definite. (Otherwise, one would have to obtain such a signal $\tilde{y} = (p, q)$ by means of non-degenerate real transform $\tilde{y} = Ty$ which always may be done [3].) If in doing so the matrix $K_{yx} G K_{xy}$ also breaks down into direct product

$$K_{yx} G K_{xy} = \begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix} \equiv D \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{2.6}$$

solution B_0 of (2.2) for $\theta \to 0$ is as follows:

$$B_0 = \frac{1}{2} \begin{pmatrix} C & 0 \\ 0 & C \end{pmatrix} \equiv \frac{1}{2} C \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}. \tag{2.7}$$

Indeed, substitute (2.5) and (2.6) into (2.2) and search its solution in the form of $B = H \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$. Taking into account the following identity:

$$f'_\theta \left(\frac{1}{2} (K+H)^{-1} D(K+H) C \otimes \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \right) =$$

$$= \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \otimes f'_\theta \left(\frac{1}{2} (K+H)^{-1} D(K+H)^{-1} C \right)$$

which holds because function $f_\theta(\cdot)$ is odd, obtain the following equation defining matrix H:

$$H = \frac{1}{2} C f'_\theta \left(\frac{1}{2} (K+H)^{-1} D(K+H)^{-1} C \right).$$

For $\theta \to 0$ function f'_θ turns into the signature of the product of two positive definite matrices that is equal to the identity matrix. Thus $H = 1/2\, C$.

Condition (2.6) is met if matrices K_{xy} and G have the following form:

$$K_{xy} = k \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad G = g \otimes \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

In this case the estimated observables x and their estimates z fall into two groups $x = \{x_p, x_q\}$, $z = \{z_p, z_q\}$ conveniently joined with complexification:

$$a = \frac{1}{\sqrt{2}} (x_p + i x_q), \quad c = \frac{1}{\sqrt{2}} (z_p + i z_p)$$

by halving the dimensionality of the linear space. The commuting optimal complex estimates $c = \{c_j\}$ under the ideal indirect measurement $\theta = 0$ are as follows:

$$c = k \left(K + \frac{1}{2} C \right)^{-1} b + \beta$$

where $b = \frac{1}{\sqrt{2}} (p + iq)$ is a complex boson s-dimensional vector described by commutators $[b, b^+]$ and correlators $\frac{1}{2} M\, [b, b^+]$ equal to positive definite matrices C and K; and $\beta = \frac{1}{\sqrt{2}} (\eta_p + i\eta_q)$ is an independent coherent noise with the following commutation and correlation matrices:

$$C_\beta = - \left(K + \frac{1}{2} C \right)^{-1} k^T C k \left(K + \frac{1}{2} C \right),$$

$$K_\beta = \frac{1}{2} \left(K + \frac{1}{2} C \right)^{-1} k^T C k \left(K + \frac{1}{2} C \right)^{-1}.$$

The appendix gives solution of (2.3) for the two-dimensional signal in the general case where condition (2.6) is not met. A critical point exists in such cases where the ideal indirect measurement degenerates into a direct measurement of optimal linear combination $(b_1 p + b_2 q)$.

3. Optimal linear filtration of stationary boson signals

The vectors x, y were assumed above to be finite-dimensional. Since all results were written in the matrix form, their generalization to the infinite-dimensional and continual cases is evident. Consider such a generalization for the stationary case.

Let the received boson signal $y = \{y_t\}$ and the estimated useful signal $x = \{x_t\}$ be stationary and stationarily correlated signals (either discrete, or continuous). Confine ourselves to the case where y_t is one-dimensional (x_t may be multi-dimensional: $x_t = \{x_j(t)\}$), and assume that expectations m_x and m_y are equal to zero.

By means of the following spectral representation

$$y_t = \int_0^{N/2} \sqrt{2} \cos 2\pi\nu t \, dP(\nu) + \int_0^{N/2} \sqrt{2} \sin 2\pi\nu t \, dQ(\nu) = \int_{-N/2}^{N/2} e^{-2\pi i\nu t} \, dY(\nu) \tag{3.1}$$

($N = 1$ if t has integral values $-\infty, \ldots, -1, 0, 1, \ldots, \infty$ and $N = \leq$ if t is continuous: $t \in (-\infty, \infty)$) come to the pairs $(dP(\nu), dQ(\nu))$ which may be easily complexified as follows:

$$dY(\nu) = \begin{cases} \frac{1}{\sqrt{2}}\big(dP(\nu) + i\,dQ(\nu)\big), & \nu \geq 0 \\ \frac{1}{\sqrt{2}}\big(dP(-\nu) + i\,dQ(-\nu)\big), & \nu < 0. \end{cases}$$

The spectral components $dY(\nu)$ under different ν are orthogonal

$$[dY(\nu), dY(\nu')^*]_- = 0, \; M[dY(\nu), dY(\nu')^*]_+ = 0 \qquad \nu \neq \nu',$$

and under similar ν are characterized by the following commutation and correlation measures:

$$dC_y(\nu) = [dY(\nu), dY(\nu)^*]_-, \; dK_y(\nu) = \frac{1}{2} M[dY(\nu), dY(\nu)^*]_+, \tag{3.2}$$

defining spectral expansions

$$C_y(\tau) = \int_{-N/2}^{N/2} e^{2\pi i\nu\tau} \, dC_y(\nu), \; K_y(\tau) = \int_{-N/2}^{N/2} e^{2\pi i\nu\tau} \, dK_y(\nu)$$

of the commutation $C_y(t-t') = [y_t, y_{t'}]_-$ and correlation $K_y(t-t') = \frac{1}{2} M[y_t, y_{t'}]_+$ functions of the stationary (in a wide sense) boson signal y_t.

Since all the matrices involved into (2.3) are assumed to be dependent only on the difference between indices t, t', its solution in the spectral representation boils down to the solution of independent algebraic equations with respect to one unknown (for each ν) that may be found graphically. Indeed, by introducing the derivatives

$$\varepsilon(\nu) = \frac{1}{2} dC_y(\nu)/dK_y(\nu), \quad \gamma(\nu) = dG_y(\nu)/dK_y(\nu), \tag{3.3}$$

$$\beta(\nu) = dB(\nu)/dK_y(\nu),$$

where $dG_y(\nu)$ is the spectral measure of matrix $K_{yx}GK_{xy} = \| g_y(t-t') \|$, and $dB(\nu)$ is the spectral measure of the desired $B = \| b(t-t') \|$, obtain

$$\varphi_\theta\left(\frac{\beta(\nu)}{\varepsilon(\nu)}\right) = \varepsilon(\nu)\,\gamma(\nu)/\big(1+\beta(\nu)\big)^2 \tag{3.4}$$

with respect to the new unknowns $\beta(\nu)$, $|\nu| \leq N/2$. The solution $\beta_\theta(\nu)$ of (3.4) for each ν, θ lies within $0 \leq \beta_\theta(\nu) \leq |\varepsilon(\nu)|$ and increases monotonically from zero to $|\varepsilon(\nu)| \leq 1$ for $\theta \to 0$.

The spectral density $a^\theta(\nu)$ of the matrix $A^\theta = \| a^\theta(t-t') \|$ defining the optimal randomized linear estimation (1.1) is as below:

$$a^\theta(\nu) = \frac{1}{1+\beta_\theta(\nu)} dK_{xy}(\nu)/dK_y(\nu) \equiv \frac{1}{1+\beta_\theta(\nu)} a_0(\nu) \tag{3.5}$$

where $dK_{xy}(\nu)$ is the spectral measure of correlation matrix $K_{xy} = \left\| \frac{1}{2} M[x_t, y_{t'}]_+ \right\|$. The frequency characteristic of the optimal linear filter (without regard to the realizability in the physical causal sense) is defined by (1.5). This formula enables one also to write readily the spectral measures $dC_\eta(\nu)$, $dK_\eta(\nu)$ of the "stationary self-noise $\eta(t)$ of the optimal measuring apparatus" having temperature θ:

$$dC_\eta(\nu) = \frac{1}{\big(1+\beta_\theta(\nu)\big)^2} a_0(\nu)\, a_0(\nu)^+\, dC_y(\nu);$$

$$dK_\eta(\nu) = \frac{\operatorname{cth} \varphi\big(\beta_\theta(\nu)/\varepsilon(\nu)\big)}{2\big(1+\beta_\theta(\nu)\big)^2} a_0(\nu)\, a_0(\nu)^+ + dC_y(\nu) \tag{3.6}$$

where $\varphi(\cdot)$ is the inverse of the derivative $\frac{d}{d\alpha}(\alpha \operatorname{cth} \alpha)$. By substituting (3.5) into (1.16) and taking into consideration (3.4), obtain the optimal density

of optimal filtration losses:

$$R^{\theta}(\nu) = R_0(\nu) + \frac{1}{2}\left[\frac{\gamma(\nu)\,\beta_\theta(\nu)}{1+\beta_\theta(\nu)} + \varphi\left(\frac{\beta_\theta(\nu)}{\varepsilon(\nu)}\right)\left(\operatorname{cth}\varphi\left(\frac{\beta_\theta(\nu)}{\varepsilon(\nu)}\right) - \frac{\beta_\theta(\nu)}{\varepsilon(\nu)}\right)\right], \tag{3.7}$$

where $R_0(\nu)$ is the spectral density of losses of the non-randomized classic estimation $A_0 y$ without allowance for the losses introduced by the indirect measurement. In this formula (as, incidentally, in (3.5) and (3.6)), parameter $\beta(\nu) = \beta_\theta(\nu)$ may be regarded as independent because of the one-to-one monotonic correspondence between θ and β. For the ideal indirect measurement one should assume that $\beta = |\,\varepsilon(\nu)\,|$, and the last term in (3.7) vanishes. In this case, losses (3.7) are greater than losses $R_0(\nu)$ for the "energy" of vacuum fluctuations $\frac{1}{2}\gamma(\nu)\,\varepsilon(\nu)/(1+\varepsilon(\nu))$. The zero value of parameter β corresponds to the strongly non-ideal optimal indirect measurement.

Example. Let us apply the results obtained to the case of stationary signal filtration at the output of a wave transmission line with strong attenuation and temperature T. The wave y_t received at the output of the line (e.g. travelling wave of voltage) is represented by superposition $x_t + \xi_t$ of the stationary classical signal x_t and the boson equilibrium noise ξ_t having the following [6] spectral commutation $k_\xi(\nu)\,dK_\xi(\nu)/d\nu$ and correlation $k_\xi(\nu) = dK_\xi(\nu)\,d\nu$ densities:

$$c_\xi(\nu) = 2r(\nu)\,h\nu,\ \ k_\xi(\nu) = r(\nu)\,h\nu\,\operatorname{cth}\frac{1}{2}\,\frac{h\nu}{kT}$$

where h is the Plank constant, k is the Boltzmann constant, and $r(\nu)$ is the wave resistance at frequency ν. Taking into consideration that $dK_{xy}(\nu) = k_x(\nu)d\nu$, $dK_y(\nu) = (k_x(\nu) + k_\xi(\nu))d\nu$ ($k_x(\nu)$ being the spectral intensity of a process x_t), obtain the following form of the frequency response (3.5) of the optimal filter under the ideal indirect measurement:

$$a^0(\nu) = k_x(\nu)/\left(k_x(\nu) + 2r(\nu)\,h\nu(1 - e^{-h\nu/kT})^{-1}\right).$$

The spectral density of filtration losses is as follows:

$$R^0(\nu) = \frac{1}{2}\,k_x(\nu)\Big/\left(1 + \frac{k_x(\nu)}{2r(\nu)\,h\nu}\,(1 - e^{-h\nu/kT})\right),$$

for weak signals $k_x(\nu) \lesssim 2r(\nu)\,h\nu$ at high frequencies $h\nu \gtrsim kT$ it differs significantly from the corresponding classical limit

$$R_0(\nu) = \frac{1}{2}\,k_x(\nu)/\left(1 + k_x(\nu)/2r(\nu)\,kT\right).$$

In the case where the spectral density $k_x(\nu)$ of the useful signal becomes equal to zero at $|\nu| \geq N/2$, $N < \infty$, the stationary estimation $z^0 = A^0 y + \eta$ at the output of the optimal filter has the following spectral intensity

$$|k_x(\nu)|^2 \Big/ \left(k_y(\nu) + \frac{1}{2} Cy(\nu)\right)$$

equal to zero at $|\nu| \geq N/2$. In this case, precise measurement of optimal values realizing the ideal indirect measurement of observables $u^0 = A^0 y$ may be done at discreet times $t_j = j\Delta$ ($\Delta \leq 1/N$), and the optimal estimates may be reestablished through the usual Kotelnikov interpolation formula.

Appendix

Optimal linear measurement of a pair of secondary commuting observables

Let y be a pair of conjugate observables p and q whose commutation matrix is characterized only by one parameter $\hbar$: $C_y = \hbar \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \equiv \hbar\sigma_2$.

Decompose the arithmetic roots of positively defined 2×2 matrices $\sqrt{K_y^{-1}}\, K_{yx}\, GK_{xy} \sqrt{K_y^{-1}}$, $\sqrt{K_y^{-1}}\, B \sqrt{K_y^{-1}}$ with respect to the Pauli matrices:

$$1 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Obtain

$$\sqrt{K_y^{-1}}\, K_{yx}\, GK_{xy} \sqrt{K_y^{-1}} = (\varkappa + k)^2, \quad \sqrt{K_y^{-1}}\, B \sqrt{K_y^{-1}} = (\alpha - a)^2 \tag{A.1}$$

where $\varkappa + k \equiv \varkappa 1 + \sum_{i=1}^{3} k_i \sigma_i$; $\alpha + a \equiv \alpha 1 + \sum_{i=1}^{3} a_i \sigma_i$ ($k_2 = 0$, $a_2 = 0$ because the matrices of (A.1) are symmetrical). By multiplying the left part of (2.3) by $\sqrt{K_y}(1 + K^{-1}B)$, and the right one by $(1 + K_y^{-1}B)\sqrt{K_y^{-1}}$, and allowing for decompositions (A.1), we can write it as follows:

$$\left(1 + (\alpha + a)^2\right) \varphi_\theta \left(\frac{1}{\varepsilon}\sigma_2(\alpha + a)^2\right) \frac{1}{\varepsilon}\sigma_2\left(1 + (\alpha + a)^2\right) = (\varkappa + k)^2. \tag{A.2}$$

Here formula $A\sigma_2 A^T = \det A\sigma_2$ is used which leads to $\sqrt{K_y}\, 2C_y^{-1} \sqrt{K_y} = \frac{2}{\hbar}\sqrt{\det K_y}\,\sigma_2 \equiv \frac{1}{\varepsilon}\sigma_2$. Taking into consideration the matrix identity $A\varphi(\sigma_2 A^2) = \varphi(A\sigma_2 A)A$ and the fact that for any symmetric matrix A and any odd function φ the equality $\varphi(A\sigma_2 A) = \varphi(\det A\sigma_2) = \varphi(\det A)\sigma_2$ holds, rewrite (A.2) as follows:

$$\varphi_0\left(\frac{\beta}{\varepsilon}\right) \frac{\beta}{\varepsilon} \left(\frac{\alpha - a}{\beta} + \alpha + a\right)^2 = (\varkappa + k)^2$$

where $\beta = \alpha^2 - a^2 = \sqrt{\det K_y^{-1} B}$. By solving it with respect to α and a, obtain:

$$\alpha = \frac{1}{\beta + 1} \sqrt{\frac{\beta\varepsilon}{\varphi_\theta(\beta/\varepsilon)}}\, \varkappa, \quad a = \frac{1}{\beta - 1} \sqrt{\frac{\beta\varepsilon}{\varphi_\theta(\beta/\varepsilon)}}\, k \tag{A.3}$$

where parameter β is defined for each θ through the condition $\alpha^2 - a^2 = \beta$, thus leading to the following algebraic equation:

$$\frac{1}{\varepsilon} \varphi_\theta\left(\frac{\beta}{\varepsilon}\right) = \frac{\varkappa^2}{(1+\beta)^2} - \frac{k^2}{(1-\beta)^2} = \frac{1}{(1-\beta^2)^2}\big((1+\beta^2)\, d - 2\beta S\big). \tag{A.4}$$

The following notation is introduced here:

$$d = \varkappa^2 - k^2 = \sqrt{\det K_y^{-1} K_{yx} G K_{xy}},$$

$$S = \varkappa^2 + k^2 = \frac{1}{2} \operatorname{Tr} K_y^{-1} K_{xy} G K_{xy}. \tag{A.5}$$

The obtained formulas (A.3) together with (A.4) having one unknown β define the matrix $B_\theta = \sqrt{K_y}(\alpha + a)^2 \sqrt{K_y}$ and lead to the solution of (1.14) which in the case under consideration may be represented as:

$$A^\theta = \frac{1}{1-\beta^2} K_{xy}\big(K_y^{-1} - \beta d (K_{yx} G K_{xy})^{-1}\big). \tag{A.6}$$

Thus, in the two-dimensional case, the matrix B_θ defining the optimal linear estimation is defined by a single parameter determined through solution of non-linear algebraic equation (A.4). Under $S = d = \gamma$, the right part of the equation becomes simpler, and it takes the form of $\varphi_\theta(\beta/\varepsilon) = \varepsilon\gamma(1+\beta)^2$ similar to that of (3.4) defining the optimal measurement of one-dimensional complex amplitude of frequency ν. Under $S \neq d$, the right-hand part of (A.4) decreases monotonically within the interval $0 \leq \beta < 1$ and has there a positive root:

$$\beta^0 = \frac{1}{d}\,(S - \sqrt{S^2 - d^2}).$$

It will be recalled that the function $\varphi_\theta(\beta/\varepsilon)$ increasing monotonically from $-\infty$ to ∞ is odd and has the domain of definition $|\beta/\varepsilon| < 1$. Thus (A.4) for each $\theta \geq 0$ has a single monotonically non-increasing solution $\beta = \beta_\theta$ lying within the domain:

$$0 \leq \beta_\theta \leq \min(|\varepsilon|, \beta^0) = \begin{cases} |\varepsilon| & \text{for} \quad |\varepsilon| < \beta^0, \\ \beta^0 & \text{for} \quad |\varepsilon| \geq \beta^0. \end{cases}$$

For $\theta \to \infty$, the lower bound ($\beta_\infty = 0$) of the inequality is reached, and for $\theta \to 0$, $\beta_0 = \min(|\varepsilon|, \beta^0)$ is reached.

By substituting (A.6) into (1.16), write the risk of optimal estimation for the case under consideration:

$$R^0 = \frac{1}{2} \operatorname{Tr} GK_x - \frac{1}{1-\beta^2}(S - \beta d) = R_0 + \frac{\beta}{1-\beta^2}(d - \beta S) \qquad \text{(A.7)}$$

where R_0 is the estimation risk without allowance for measurement losses. Comparison of the equality (A.7) of the optimal estimation allowing randomized solutions with that of non-randomized estimation

$$R^* = R_0 + \frac{\beta^0}{1-\beta^{\theta 2}}(d - \beta^0 S) = R_0 - \frac{1}{2}(S - \sqrt{S^2 - d^2})$$

based on measurement of the optimal combination $b_1^0 p + b_2^0 q$ [3] shows that within the domain $|\varepsilon| < \beta^0$ the randomization with temperature $\theta = 0$ gives essential advantage (see Fig. 1). For $|\varepsilon| \geq \beta^0$ (essentially quantum domain) losses (A.7) become under ideal measurement ε-independent and coincide with the non-randomized estimation losses. It means that at the point $|\varepsilon| = \beta^0$ the randomized estimation based on the ideal measurement undergoes the "phase transition of the second kind" and degenerates into the non-randomized

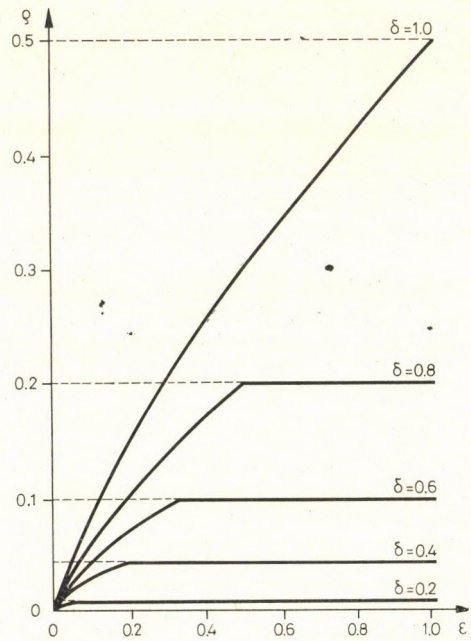


Fig. 1. Dependence of optimal ideal indirect measurement losses $\varrho^0 = (R^0 - R_0)/S$ on ε under various values of $d/S \equiv \delta$. Corresponding levels of ε-independent losses $\varrho^* = (R^* - R_0)/S$ of the optimal direct measurement are shown by dashed lines

estimation. Indeed, (A.6) may be represented as

$$A^\theta = K_{xy}\left(\frac{1-\beta\sqrt{\gamma_2/\gamma_1}}{1-\beta^2}\, b_{\gamma_1} b_{\gamma_1}^T + \frac{1-\beta\sqrt{\gamma_1/\gamma_2}}{1-\beta^2}\, b_{\gamma_2} b_{\gamma_2}^T\right), \tag{A.8}$$

where $b_{\gamma_1}, b_{\gamma_2}$ are eigenvectors of matrix $K_y K_{yx} G K_{xy}$ corresponding to its eigenvalues γ_1, γ_2 and normalized with respect to weight K_y: $b_{\gamma_i} K_y b_{\gamma_i}^T = \delta_{ij}$. Taking into consideration that $\beta = \min(\sqrt{\gamma_1/\gamma_2}, \sqrt{\gamma_2/\gamma_1})$, obtain that matrix (A.8) within the domain $|\varepsilon| \geq \beta^0$ under ideal measurement $\theta = 0$ degenerates and has the rank $r(A^0) = 1$:

$$A^0 = K_{xy} b^0 b^{0T}$$

where b^0 is that of vectors $b\gamma_1$, $b\gamma_2$ which corresponds to the greater eigenvalue $\gamma^0 = \max(\gamma_1, \gamma_2)$. Observables $A_y^0 = K_{xy} b^0 b^{0T} y$ are linearly dependent on the only observable $b^{0T} y = b_1^0 p + b_2^0 q$ which may be measured without randomization because a single observable may be measured directly.

For $S = d$, $\beta_0 = 1$ and measurement randomization is advantegeous along all the domain of the parameter $\varepsilon = \hbar/2\sqrt{\det K_y} \leq 1$ values. This takes place if condition (2.6) is met and, specifically, in the stationary case allowing complexification.

References

1. *Helstrom, C. W.* et al., "Quantum-Mechanical Communication Theory", Proc. of IRE, **58** (1970) 10, 186–207.
2. *Belavkin V. P., Grishanin B. A.*, "Issledovanie zadachi optimal'nogo otsenivaniya v kvantovykh kanalakh svyazi metodom obobschennogo neravenstva Geizenberga" (Study of Optimal Estimation in Quantum Communication Channels by Generalized Geizenberg Inequality) in "Problemy Peredachi Informatsii" (Problems in Information Transmission), **9** (1973) 3 (in Russian).
3. *Belavkin V. P., Grishanin B. A.*, "Optimal' noe izmerenie kvantovykh peremennykh" (Optimal Measurement of Quantum Variables) in "Problemy Peredachi Informatsii" (Problems in Information Transmission), **8** (1972) 3 (in Russian).
4. *Gordon, J. P., Louisell, W. H.*, "Simultaneous Measurement of Noncommuting Observables", in "Physics of Quantum Electronics", P. Z. Kelley et al. (Eds.), New York, McGraw-Hill, 1966, 833–840.
5. *Belavkin V. O.*, "Lineinoe otsenivanie nekommutiruyushchikh kvantovykh peremennykh pri kosvennom ikh izmerenii" (Linear Estimation of Noncommuting Quantum Variables under Indirect Measurement), in "Radiotekhnika i Elektronika" (Radio and Electronic Engineering), **17** (1972) 12, 2533 (in Russian).
6. *Haus, A.*, "Steady-State Quantum Analysis of Linear Systems", Proc. of IRE, **58** (1970) 10.

Оптимальная линейная рандомизированная фильтрация квантовых бозонных сигналов

В. П. БЕЛАВКИН

(Москва)

Линейная рандомизированная фильтрация бозонного действительного сигнала $y = \{y_t, t \in T\}$ описывается матрицей $\| A_{ts} \|$ и независимым от y бозонным шумом $\eta = \{\eta_s, s \in S\}$, которые определяют линейную оценку $z = Ay + \eta$ оцениваемого сигнала

$x = \{x_s, s \in S\}$. Поскольку измеряемые $z = \{z_s, s \in S\}$ должны коммутировать: $[z_s, z_{s'}] = 0$ c–числовая матрица C_η коммутаторов $[\eta_s, \eta_{s'}]$ равна матрице AC_yA^T с противоположным знаком ($C_y = \|[y_t, y_{t'}]\|$ — коммутационная матрица принимаемого сигнала y). Оптимальное состояние бозонной системы, описываемой наблюдаемыми η, минимизирующее при фиксированной энтропии среднеквадратичные потери косвенного измерения $\frac{1}{2}\langle (z-x)^T G(z-x)\rangle$, есть гауссово.

Выводимое уравнение, определяющее оптимальное A, отличается от уравнения $AK_y - K_{xy} = 0$ в классическом (коммутативном) случае наличием нелинейного члена:

$$AK_y - K_{xy} + \frac{1}{2} f'_\theta\left(\frac{1}{2} AC_yA^TG\right) AC_y = 0 \tag{1}$$

(K_y, K_{xy} — матрицы корреляторов: $K_y = \frac{1}{2}\|\langle y'_t y_{t'} + y_{t'} y_t\rangle\|$, $K_{xy} = \|\langle x_s y_t\rangle\|$, а функция $f'_\theta(.)$ есть производная от $f_\theta(\alpha) = \alpha \mathrm{cth}(\alpha/\theta)$ по α). Параметр $\theta \geq 0$ есть «температура» шума η, и характеризует степень рандомизации, т. е. степень неидеальности квазиизмерения некоммутирующих Ay.

Решение A^θ уравнения (1) удовлетворяет принципу соответствия: $A^\theta \to A_0$ при $C_y \to 0$ ($A_0 = K_{xy}K^{-1}$ есть решение уравнения (1) в классическом случае $C_y = 0$), и может быть найдено при $\theta \neq 0$ методом последовательных приближений: $A^\theta = A_0 + \frac{1}{\theta} A_1 + \frac{1}{\theta^2} A_2 + \ldots$.

В идеальном случае $\theta = 0$ оптимальное квазиизмерение становится когерентным; и в существенно квантовых областях может вырождаться в прямое измерение. Последнее не имеет места, если

$$K_{xy} = \begin{pmatrix} K & 0 \\ 0 & K \end{pmatrix}, \quad K_y = \begin{pmatrix} R & 0 \\ 0 & R \end{pmatrix}, \quad C_y = \begin{pmatrix} 0 & -iC \\ iC & 0 \end{pmatrix} \tag{2}$$

(R, C — положительно определенные матрицы). В этом случае сигналы $x = \{u_s, v_s\}$, $y = \{p_t, q_t\}$ допускают простое комплексное представление: $a = \frac{1}{\sqrt{2}}(u + iv)$, $b = \frac{1}{\sqrt{2}}(p + iq)$, и оптимальная комплексная оценка $c = \{c_s\}$ оцениваемого комплексного $a = \{a_s\}$ имеет вид:

$$c = K\left(R + \frac{i}{2} C\right)^{-1} b + \beta,$$

где $\beta = \{\beta_s\}$ есть комплексный шум, пропорциональный операторам рождения бозонов оптимальной вспомогательной системы, находящейся в когерентном состоянии.

В частности, (2) имеет место, если y_t есть стационарный случайный бозонный процесс, стационарно коррелированный с x_s. В идеальном случае $\theta = 0$ оптимальная спектральная характеристика линейного фильтра есть $A^0(\nu) = A_0(\nu)/(1 + \varepsilon(\nu))$, где $0 \leq \varepsilon(\nu) \leq 1$, а $A_0(\nu)$ есть оптимальная спектральная характеристика в классическом случае $\varepsilon(\nu) = 0$. Если y_t есть принимаемая бегущая волна напряжения в линии передачи с затуханием и температурой T, то $\varepsilon(\nu) = \mathrm{th}(h\nu/2kT)$, где h, k есть постоянные Планка и Больцмана.

В. П. Белавкин
Московский Государственный Университет им. М. В. Ломоносова,
физический факультет
СССР, Москва В-234,
Ленинские горы

Problems of Control and Information Theory, Vol. 3 (1), pp. 63–76 (1974)

ОБ АППРОКСИМАЦИИ ЛИНЕЙНЫХ УПРАВЛЯЕМЫХ СИСТЕМ С ПОСЛЕДЕЙСТВИЕМ

В. Б. КОЛМАНОВСКИЙ

(Москва)

(Поступила в редакцию 3 мая 1972 г.)

Рассматривается задача об оптимальном управлении, минимизирующем квадратичный функционал в линейной системе с последействием. Предложены два способа аппроксимации оптимального управления и минимизируемого функционала. Получены оценки разности между приближенным и точным решением задачи, формулируемые в терминах коэффициентов уравнений.

1. Пусть вектор фазовых координат $x(t)$ описывается системой уравнений

$$x^{\cdot}(t) = A(t)\,x(t) + B(t)\,x(t-h) + D(t)\,u(t)\,, \quad 0 \leq t \leq T, \tag{1.1}$$

где $x(t) \in E_n$ (E_n — евклидово пространство размерности n). Матрицы $A(t)$, $B(t)$ размерности $n \times n$ и матрица $D(t)$ размерности $n \times m$ с кусочно-непрерывными элементами заданы. Наконец, постоянная $h \in [0, h_0]$, а управление $u(t) \in E_m$. Решение системы (1.1) при $t \geq 0$ определяется начальными условиями

$$x(\tau) = \varphi(\tau), \qquad -h \leq \tau \leq 0, \tag{1.2}$$

где $\varphi(\tau) \in E_n$ — заданная, кусочно-непрерывная функция.

Задача 1. Выбором управления $u(t)$ минимизировать функционал

$$x'(T)\,Hx\,(T) + \int_0^T [x'(t)\,M(t)\,x(t) + u'(t)\,N(t)\,u(t)]\,dt\,. \tag{1.3}$$

Здесь постоянная матрица $H \geq 0$, кусочно-непрерывные матрицы $M(t) \geq 0$, $N(t) > 0$, $0 \leq t \leq T$ (штрих — знак транспонирования, символ $H \geq 0$ означает неотрицательную определенность матрицы H, а символ $N(t) > 0$ — положительную определенность матрицы $N(t)$).

Вопросу об аппроксимации уравнений (1.1) с последействием управляемой системой обыкновенных дифференциальных уравнений посвящены работы [1, 2]. Обоснование возможности такой аппроксимации дано в [3]; здесь предполагается, что коэффициенты в (1.1), (1.3) постоянны и управление $u(t)$ — скалярная функция. Далее в [3] рассматривается вспомогатель-

ная задача минимизации квадратичного функционала вида (1.3) на траекториях системы обыкновенных дифференциальных уравнений

$$y^{(i)}(t) + ky^{(i)}(t) = ky^{(i-1)}(t)\,,$$
$$y^{\cdot(0)}(t) = Ay^{(0)}(t) + By^{(k)}(t) + Du\,, \qquad 0 \leq t \leq T\,, \tag{1.4}$$

где k — целое число, $i = 1, 2, \ldots, k$, а $y^{(i)}(t)$ — есть i-ая производная функции $y(t)$. В [3] установлена сходимость при $k \to \infty$ оптимальных управлений, траекторий и значений оптимизируемого функционала для системы (1.4) к соответствующим величинам в исходной задаче 1.

Задаче оптимального управления линейными системами с последействием и квадратическим функционалом качества (1.3) посвящены также работы [4, 5]. В [4] для детерминированных систем вида (1.1) и функционала (1.3) при $H = 0$ с помощью метода Беллмана установлен вид последовательных приближений к оптимальному управлению и оптимальному значению функционала (1.3). Наконец, в [5] для общих стохастических систем с последействием и произвольной матрице $H \geq 0$ в (1.3) найдены формулы последовательных приближений к оптимальным управлению и функционалу, установлена сходимость последовательных приближений, предельные уравнения в частных производных для коэффициентов оптимального управления, а также условия в терминах параметров системы (1.1) существования решения у этих предельных уравнений. Первая цель настоящей работы — оценить скорость сходимости указанного метода последовательных приближений. Этому посвящен раздел 2.

На основании результатов работ [3—6] оптимальное управление в линейных системах с последействием и квадратическим функционалом качества есть линейный функционал от траектории процесса на длины времени запаздывания, то-есть при формировании управления необходимо хранить в памяти всю траекторию на длине времени запаздывания. Последнее, однако, не всегда возможно.

Вторая цель настоящей работы установить связь задачи 1 и следующей задачи 2.

Задача 2. Определить управление $v(t)$, минимизирующее функционал

$$y'(T)\,Hy(T) + \int_0^T [y'(t)\,M(t)\,y(t) + v'(t)\,N(t)v\,(t)]\,dt, \tag{1.5}$$

где $y(t)$ — решение системы обыкновенных дифференциальных уравнений

$$y^{\cdot}(t) = [A(t) + B(t)]\,y(t) + D(t)\,v(t), \quad 0 \leq t \leq T \qquad y(0) = \varphi(0)\,. \tag{1.6}$$

В разделе 3 получены оценки разности оптимальных величин в задачах 1 и 2. Отметим, что рассмотрением систем вида (1.1) мы ограничиваемся лишь для простоты изложения. Незначительно измененные результаты разделов 2 и 3 справедливы также и для линейных стохастических систем с последействием, содержащих как несколько дискретных, так и распределенное запаздывание.

Для более точной постановки задачи об оценке скорости сходимости последовательных приближений приведем некоторые необходимые результаты работ [4, 5]. Определим квадратные матрицы $P(t) \geq 0, Q(t, \tau), R(t, \tau, \varrho) = R'(t, \varrho, \tau)$ $(0 \leq t \leq T, -h \leq \tau, \varrho \leq 0)$ размерности $n \times n$ как решение системы уравнений

$$D_1(t) = D(t) N^{-1}(t) D'(t), \; t - h < \tau_1, \; \varrho \leq t$$

$$P^{\cdot}(t) + A'(t) P(t) + P(t) A(t) + Q(t, 0) + Q'(t, 0) = P(t) D_1(t) P(t) - M(t),$$

$$A'(t) Q(t, \tau) + R(t, 0, \tau) + \frac{\partial Q}{\partial t}(t, \tau_1 - t) = P(t) D_1(t) Q(t, \tau), \tag{1.7}$$

$$\frac{\partial R(t, \tau_1 - t, \varrho_1 - t)}{\partial t} = Q'(t, \tau) D_1(t) Q(t, \varrho).$$

Граничными условиями для этой системы уравнений служат

$$P(T) = H, \; Q(T, \tau) = R(T, \tau, \varrho) \equiv 0, \; -h < \tau, \; \varrho \leq 0$$

$$B'(t) P(t) - Q'(t_1 - h) = 0, \; 0 \leq t \leq T \tag{1.8}$$

$$B'(t) Q(t_1, \tau) - R(t_1, -h, \tau) + Q'(t_1, \tau) B(t) - R'(t, \tau, -h) = 0.$$

Обозначим через $P_0(t) \geq 0$, $Q_0(t, \tau)$, $R_0(t, \tau, \varrho)$ $(R_0'(t, \tau, \varrho) = R_0(t, \varrho, \tau))$ произвольные непрерывно-дифференцируемые по t, τ, ϱ квадратные матрицы $n \times n$, удовлетворяющие требованиям (1.8). Определим далее последовательность матриц $P_i(t)$, $Q_i(t, \tau)$, $R_i(t, \tau, \varrho)$ как решения задачи (1.7), (1.8), в которой вместо P, Q, R стоят P_i, Q_i, R_i, а вместо слагаемых $P(t) D_1(t) P(t)$, $P(t) D_1(t) Q(t, \tau)$, $Q'(t, \tau) D_1(t) Q(t, \varrho)$ стоят соответственно

$$P_{i-1}(t) D_1(t) P_i(t) + P_i(t) D_1(t) P_{i-1}(t) - P_{i-1}(t) D_1(t) P_{i-1}(t),$$

$$P_{i-1}(t) D_1(t) Q_i(t, \tau) + P_i(t) D_1(t) Q_{i-1}(t, \tau) - P_{i-1}(t) D_1(t) Q_{i-1}(t, \tau), \tag{1.9}$$

$$Q_{i-1}'(t, \tau) D_1(t) Q_i(t, \varrho) + Q_i'(t, \tau) D_1(t) Q_{i-1}(t, \varrho) - Q_{1-i}'(t, \tau) D_1(t) Q_{i-1}(t, \varrho).$$

Тогда на основании [6] последовательности $P_i(t)$, $Q_i(t, \tau)$, $R_i(t, \tau, \varrho)$ сходятся поточечно к решению задачи (1.7), (1.8). Оптимальные управление $u_0(t)$ и зна-

чение J_0 функционала (1.3) в задаче 1 определяются формулами

$$u_0(t) = -N^{-1}D^1(t)\left[P(t)\,x(t) + \int_{-h}^{0} Q(t,\tau)\,x(t+\tau)\,d\tau\right],$$

$$J_0 = \varphi'(0)\,P(0)\,\varphi(0) + \varphi'(0)\int_{-h}^{0} Q(0,\tau)\,\varphi(\tau)\,d\tau + \tag{1.10}$$

$$+\int_{-h}^{0} \varphi'(\tau)\,Q'(0,\tau)\,d\tau\,\varphi(0) + \int_{-h}^{0}\int_{-h}^{0} \varphi'(\tau)\,R(0,\tau,\varrho)\,\varphi(\varrho)\,d\tau\,d\varrho,$$

где $\varphi(\tau)$ — начальное условие (1.2) для системы (1.1). Пусть $A^{ij}(t)$ — элемент матрицы $A(t)$. Положим

$$\|A(t)\| = \max_{i,j} |A^{ij}(t)| \qquad (i, j = 1, \ldots, n). \tag{1.11}$$

Оценить скорость сходимости — значит оценить нормы разностей

$$\|P(t) - P_i(t)\|, \quad \|Q(t,\tau) - Q_i(t,\tau)\|_1 = \max_\tau \|Q(t,\tau) - Q_i(t,\tau)\|$$

$$\|R(t,\tau,\varrho) - R_i(t,\tau,\varrho)\|_2 = \max_{\tau,\varrho} \|R(t,\tau,\varrho) - R_i(t,\tau,\varrho)\|. \tag{1.12}$$

Ввиду некоторой громоздкости доказательств вынесем их в приложение.

2. Обозначим

$$\alpha_i^P = \|P(t) - P_i(t)\|, \quad \alpha_i^Q = \|Q(t,\tau) - Q_i(t,\tau)\|_1, \tag{2.1}$$

$$\alpha_i^R = \|R(t,\tau,\varrho) - R_i(t,\tau,\varrho)\|_2. \tag{2.2}$$

Оценки выражений (2.1), (2.2) будут получены в два этапа.

1) Оценим вначале нормы матриц $P_i(t)$, $Q_i(t,\tau)$, $R_i(t,\tau,\varrho)$. Воспользуемся для этого формулами, определяющими P_i, Q_i, R_i через $P_{i-1}, Q_{i-1}, R_{i-1}$ ([5, 6]):

$$P_1(t) = \int_t^T \Big[K_i'(s,t)\big(M(s) + P_{i-1}(s)\,D_1(s)\,P_{i-1}(s)\big)\,K_i(s,t) +$$

$$+\int_{s-h}^{s} K_i'(s,t)\,P_{i-1}(s)\,D_1(s)\,Q_{i-1}(s,\tau-s)\,K(\tau,t)\,d\tau + \tag{2.3}$$

$$+\int_{s-h}^{s}\int_{s-h}^{s} K_i'(\tau,t)\,Q_{i-1}'(s,\tau-s)\,D_1(s)\,Q_{i-1}(s,\varrho-s)\,K_i(\varrho,t)\,d\tau\,d\varrho\Big]\,ds +$$

$$+ K_i'(T,t)\,HK_i(T,t),$$

где $K_i(s,t)$ — фундаментальное решение системы уравнений (1.1) при управлении $u(t)$, равном

$$u_i(t) = -N^{-1}(t)\,D'(t)\left[P_{i-1}(t)\,x(t) + \int_{-h}^{0} Q_{i-1}(t,\tau)\,x(t+\tau)\,d\tau\right]. \tag{2.4}$$

Матрица $Q_i(t, \tau)$ определяется из соотношений

$$
\begin{aligned}
Q_i(t, \tau) = \int_t^T & [K_i'(s, t)\,(M(s) + P_{i-1}(s)\,D_1(s)\,P_{i-1}(s))\,K_{ii}(s, t, \tau) + \\
& + \int_{s-h}^{s} (K_i'(s, t)\,P_{i-1}(s)\,D_1(s)\,Q_{i-1}(s, s_1 - s))\,K_{ii}(s_1, t, \tau) + \\
& + K_s'(s_1, t)\,Q_{i-1}(s, s_1 - s)\,D_1(s)\,P_{i-1}(s)\,K_{ii}(s, t, \tau)\,ds_1 + \qquad (2.5) \\
& + 2 \int_{s-h}^{s} \int_{s-h}^{s} ds_1\,ds_2\,K_i'(s_1, t)\,Q_{i-1}'(s, s_1 - s)\,D_1(s)\,Q_{i-1}(s, s_2 - s) \times \\
& \times K_{ii}(s_2, t, \tau)]\,ds + \int_t^{\min(T,\, t+\tau+h)} [K_i'(s, t)\,P_{i-1}(s)\,D_1(s)\,Q_{i-1}(s, t + \tau - s) + \\
& + 2 \int_t^{s} ds_1\,K_i'(s_1, t)\,Q_{i-1}(s, s_1 - s)\,D_1(s)\,Q_{i-1}(s, t + \tau - s)]\,ds + \\
& + K_i'(T, t)\,HK_{ii}(T, \tau, t)\,,
\end{aligned}
$$

где

$$
\begin{aligned}
K_{ii}(s, t, \tau) = {} & K_i(s, t + \tau + h)\,B(t + \tau + h) - \\
& - \int_{-h}^{0} K_i(s, t + \tau - \varrho)\,D_1(t + \tau - \varrho)\,Q_{i-1}(t + \tau - \varrho, \varrho)\,d\varrho\,. \qquad (2.6)
\end{aligned}
$$

Наконец, $R_i(t, \tau, \varrho)$ задается соотношением

$$
\begin{aligned}
R_i(t, \tau, \varrho) = \int_t^T & [K_{ii}'(s, t, \tau)\,(M(s) + P_{i-1}(s)\,D_1(s)\,P_{i-1}(s))\,K_{ii}'(s, t, \varrho) + \\
& + \int_{s-h}^{s} ds_1\,K_{ii}'(s, t, \tau)\,P_{i-1}(s)\,D_1(s)\,Q_{i-1}(s, s_1 - s)\,K_{ii}(s_1, t, \varrho) + \\
& + \int_{s-h}^{s} \int_{s-h}^{s} ds_1\,ds_2\,K_{ii}'(s_1, t, \tau)\,Q_{i-1}'(s, s_1 - s)\,D_1(s)\,Q_{i-1}(s, s_2 - s)\,K_{ii}(s_2, t, \varrho)]\,ds + \\
& + \int_t^{\min(T,\, t+\tau+h)} Q_{i-1}'(s, t + \tau - s)\,D_1(s)\,P_{i-1}(s)\,K_{ii}(s, t, \varrho)\,ds + \qquad (2.7) \\
& + 2 \int_t^{\min(T,\, t+\tau+h)} ds \int_t^{s} ds_1\,K_{ii}'(s_1, t, \tau)\,Q_{i-1}(s, s_1 - s) D_1(s) Q_{i-1}((s, t + \varrho - s) + \\
& + \int_t^{\min T,\, +\tau+h,\, t+\tau+h)} \\
& + K_{ii}'(T, t, \tau)\,HK_{ii}(T, t, \varrho)\,.
\end{aligned}
$$

Замечание. Отметим, что соотношения (2.3)—(2.7) представляют собой формулы, задающие решение уравнений в частных производных для P_i, Q_i, R_i с граничными условиями (1.8). Это обстоятельство будет использовано в дальнейшем.

На основании результатов [4—6] построенные таким образом матрицы удовлетворяют при любых начальных функциях $\varphi(\tau)$ неравенству

$$J_i = \varphi'(0)\, P_i(t)\, \varphi(0) + \varphi'(0) \int_{-h}^{0} Q_i(t, \tau)\, \varphi(\tau)\, d\tau +$$

$$+ \int_{-h}^{0} \varphi'(\tau)\, Q_i'(t, \tau)\, \varphi'(0)\, d\tau + \int_{-h}^{0} \int_{-h}^{0} \varphi'(\tau)\, R_i(t, \tau\, \varrho)\, \varphi(\varrho)\, d\tau\, d\varrho \leq$$

$$\leq J_{i-1} = \varphi'(0)\, P_{i-1}(t)\, \varphi(0) + \varphi'(0) \int_{-h}^{0} Q_{i-1}(t, \tau)\, \varphi(\tau)\, d\tau + \tag{2.8}$$

$$+ \int_{-h}^{0} \varphi'(\tau)\, Q_{i-1}'(t, \tau)\, \varphi'(0)\, d\tau + \int_{-h}^{0} \int_{-h}^{0} \varphi'(\tau)\, R_{i-1}(t, \tau, \varrho)\, \varphi(\varrho)\, d\tau\, d\varrho\,.$$

$$0 \leq t \leq T, \quad i = 1, 2, \ldots,$$

причем J_i при $t = 0$ есть значение функционала (1.3) при управлении $u_i(t)$, задаваемом (2.4).

Будем далее считать, что $P_0(t) = Q_0(t, \tau) = R_0(t, \tau, \varrho) \equiv 0$. Тогда матрица $K_1(s, t)$ есть фундаментальное решение системы (1.1) при $u(t) \equiv 0$ и легко может быть найдена в явном аналитическом виде (например, интегрированием с помощью метода шагов [8]). Из определения матриц $P_i(t)$ следует, что $P_i(t) \geq 0$, $0 \leq t \leq T$, $i = 1, 2, \ldots$ Отсюда и из (2.8) вытекает, что

$$\| P_i(t) \| \leq \| P_1(t) \|, \qquad i = 1, 2, \ldots, \tag{2.9}$$

где $P_1(t)$ ввиду (2.3) и определения P_0, Q_0 есть

$$P_1(t) = \int_{t}^{T} K_1'(s, t)\, M(s)\, K_1(s, t)\, ds + K_1'(T, t)\, H K_1(T, t)\,, \tag{2.10}$$

т. е. $\| P_1(t) \|$, а вместе с нею и все $\| P_i(t) \|$ оценены через параметры исходной системы (1.1) и функционала качества (1.3).

Обратимся теперь к оценкам нормы $\| R_i(t, \tau, \varrho) \|$. Прежде всего на основании (2.8) при всех $0 \leq t \leq T$, $-h \leq \tau \leq 0$

$$0 \leq R_k^{ii}(t, \tau, \tau) \leq R_1^{ii}(t, \tau, \tau)\,, \qquad k = 1, 2, \ldots \tag{2.11}$$

Кроме того, ввиду (2.8) и определения матриц $R_i(t, \tau, \varrho)$ имеем для всех $k \geq 1$, любых пар (τ, ϱ), (i, j) и любых чисел q_1, q_2

$$0 \leq q_1^2 R_k^{ii}(t, \tau, \tau) + R_k^{jj}(t, \varrho, \varrho) q_2^2 + \left(R_k^{ij}(t, \tau, \varrho)\right) + R_k^{ji}(t, \varrho, \tau)\right) q_1 q_2;$$

$$R_k^{ij}(t, \tau, \varrho) = R_k^{ji}(t, \varrho, \tau). \tag{2.12}$$

Отсюда заключаем, что для произвольных фиксированных t, τ, ϱ, i, j матрица r с элементами

$$r^{11} = R_k^{ii}(t, \tau, \tau), \qquad r^{22} = R_k^{jj}(t, \varrho, \varrho),$$

$$r^{12} = r^{21} = R_k^{ij}(t, \tau, \varrho)$$

неотрицательно определена. Поэтому

$$| R_k^{ij}(t, \varrho, \tau) | \leq \max [R_k^{ii}(t, \tau, \tau), R_k^{jj}(t, \varrho, \varrho)] .$$

Значит, с учетом (2.11)

$$|| R_k(t, \tau, \varrho) ||_2 \leq || R_1(t, \tau, \varrho) ||_2 , \tag{2.13}$$

где матрица $R_1(t, \tau, \varrho)$ в силу (2.7) равна

$$R_1(t, \tau, \varrho) = \int_t^T K_{11}'(s, t, \tau) M(s) K_{11}'(s, t, \varrho) ds + K_{11}'(T, t, \tau) H K_{11}(T, t, \varrho). \tag{2.14}$$

Рассмотрим, наконец, функции $Q_k(t, \tau)$. Аналогично (2.12) нетрудно получить, используя (2.8), что при любых фиксированных i, j, t, τ и любых числах q_1, q_2

$$0 \leq q_1^2 P_k^{ii}(t) + 2 q_1 q_2 Q_k^{ij}(t, \tau) + R_k^{jj}(t, \tau) .$$

Поэтому матрица r_1 с элементами

$$r_1^{11} = P_k^{ii}, \quad r_1^{22} = R_k^{jj}(t, \tau, \tau), \quad r_1^{12} = r_1^{21} = Q_k^{ij}(t, \tau)$$

неотрицательно определена.

Следовательно,

$$| Q_k^{ii}(t,\tau) | \leq \max [P_k^{ii}(t), R_k^{jj}(t, \tau, \tau)] .$$

Отсюда и из (2.9), (2.13) заключаем, что при всех k

$$|| Q_k(t, \tau) ||_1 \leq \max [|| P_1(t) ||, || R_1(t, \tau, \tau) ||_2] . \tag{2.15}$$

2. Введем в рассмотрение функции $\Delta_k^1(t)$, $\Delta_k^2(t)$, $\Delta_k^3(t, \tau, \varrho)$ с помощью соотношений

$$\Delta_{k+1}^1(t) = - P_{k+1}(t) + P_k(t), \quad \Delta_{k+1}^2(t, \tau) = - Q_{k+1}(t, \tau) + Q_k(t, \tau),$$

$$\Delta_{k+1}^3(t, \tau, \varrho) = R_k(t, \tau, \varrho) - R_{k+1}(t, \tau, \varrho).$$

Покажем, что введенные таким образом функции Δ^1_{k+1} удовлетворяют тем же уравнениям, что и P_{k+1} при $M(t) \equiv 0$ с однородными граничными условиями (1.8). Однородность граничных условий очевидна. Далее, вычитая из уравнения для P_k уравнение для P_{k+1}, имеем

$$\begin{aligned}\dot{\Delta}^1_{k+1}(t) + A'(t)\,\Delta^1_{k+1}(t) + \Delta^1_{k+1}(t)\,A(t) + \Delta^2_{k+1}(t, 0) + \Delta^2_{k+1}(t, 0)' = \\ = \beta(t) = P_k(t)\,D_1(t)\,P_{k-1}(t) + P_{k-1}(t)\,D_1(t)\,P_k(t) - \\ - P_{k-1}(t)\,D_1(t)\,P_{k-1}(t) - P_{k+1}(t)\,D_1(t)\,P_k(t) - \\ - P_k(t)\,D_1(t)\,P_{k+1}(t) + P_k(t)\,D_1(t)\,P_k(t)\,.\end{aligned} \tag{2.16}$$

Преобразуем правую часть (2.16). Нетрудно видеть, что

$$P_{k-1}(t)\,D_1(t)\,P_{k-1}(t) = \Delta^1_k(t)\,D_1(t)\,\Delta^1_k(t) + P_k(t)\,D_1(t)\,\Delta^1_k(t) + \Delta^1_k(t)\,D_1(t)\,P_k(t). \tag{2.17}$$

Кроме того

$$P_k(t)\,D_1(t)P_{k-1}(t) - P_k(t)\,D_1(t)\,P_{k+1}(t) = P_k(t)\,D_1(t)\,(\Delta^1_{k+1}(t) + \Delta^1_k(t))\,.$$

Отсюда и из (2.17) получаем, что

$$\beta(t) = P_k(t)\,D_1(t)\,\Delta^1_{k+1}(t) + \Delta^1_{k+1}(t)\,D_1(t)\,P_k(t) - \Delta^1_k(t)\,D_1(t)\,\Delta^1_k(t)\,.$$

Поэтому $\Delta^1_{k+1}(t)$ с учетом (2.16) удовлетворяют тем же уравнениям, что и $P_{k+1}(t)$. Совершенно аналогично убеждаемся, что $\Delta^2_{k+1}(t, \tau)$ и $\Delta^3_{k+1}(t, \tau, \varrho)$ удовлетворяют уравнениям, определяющим $Q_{k+1}(t, \tau)$ и $R_{k+1}(t, \tau, \varrho)$. Отсюда и из сделанного в п. 1) замечания о представлении решений системы уравнений для P_k, Q_k, R_k в виде (2.3), (2.5), (2.7) следует, что для функций Δ^j_i справедливо интегральное представление вида (2.3), (2.5)—(2.7), в котором следует положить $M(t) \equiv 0$, $H = 0$. Таким образом, имеем, например, из (2.3), что

$$\begin{aligned}\Delta^1_i(t) = \int_t^T [K^1_i(s, t)\,\Delta^1_{i-1}(s)\,D_1(s)\,\Delta^1_{i-1}(s)\,K_i(s, t) + \\ + \int_{s-h}^{s} K'_i(s, t)\,\Delta^1_{i-1}(s)\,D_1(s)\,\Delta^2_{i-1}(s, \tau - s)\,K_i(\tau, t)\,d\tau + \\ + \int_{s-h}^{s}\int_{s-h}^{s} K'_i(\tau, t)\,\Delta^2_{i-1}(s, \tau - s)'\,D_1(s)\,\Delta^2_{i-1}(s, \varrho - s)\,K_i(\varrho, t)\,d\tau\,d\varrho]\,ds\,.\end{aligned} \tag{2.18}$$

Подобным образом записываются с учетом (2.5)—(2.7) и выражения для Δ^2_i и Δ^3_i. Применяя к последовательности Δ^i_j в точности те же рассуждения, которые были использованы выше при получении оценок (2.9), (2.13), (2.15),

убеждаемся, что

$$\| \Delta_i^1(t) \| \leq \| P_1(t) \|, \ \| \Delta_i^2(t, \tau) \|_1 \leq \max [\| P_1(t) \|, \ \| R_1(t, \tau, \tau) \|_2] \tag{2.19}$$

$$\| \Delta_i^3(t, \varrho, \tau) \|_2 \leq \| R_1(t, \tau, \varrho) \|_2, \qquad i = 1, 2, \ldots$$

Функции $K_i(s, t)$, будучи фундаментальными решениями системы линейных уравнений с равномерно ограниченными коэффициентами (ввиду (2.9), (2.13), (2.15), (2.3)—(2.6)), легко оцениваются известным образом. Именно (см. [9], стр. 153)

$$\| K_i(s, t) \| \leq \exp \int_t^s [\| B(\tau) \| + \| A(\tau) \| + \| N^{-1}(\tau) D'(\tau) P_1(\tau) \| + \\ + h \| Q_1(\tau, \varrho) \|_2] \, d\tau \, . \tag{2.20}$$

Обозначим через $\beta_1(t, s)$ функцию, определенную соотношением

$$\beta_1(t, s) = \max [\| K_i'(s, t) \| (\| D_1(s) \Delta_{i-1}^1(s) K_i(s, t) + \\ + \int_{s-h}^s D_1(s) \Delta_{i-1}^2(s, \tau - s) K_i(t, \tau) \, d\tau \|, \tag{2.21}$$

$$\| \int_{s-h}^s \int_{s-h}^s \| K_i'(\tau, t) \| D_1(s) \Delta_{i-1}^2(s, \varrho - s) K_i(\varrho, t) \, d\tau \, d\varrho \|] \, , \quad i = 2, 3, \ldots$$

и положим

$$C_1 = \max_{t,s,i} \beta_1(t, s), \qquad 0 \leq t, s \leq T, \qquad i = 2, 3, \ldots \tag{2.22}$$

Отметим, что ввиду (2.19)—(2.22) постоянная C_1 выражается через параметры системы (1.1), (1.3).

Из (2.18), (2.21), (2.22) следует, что

$$\| \Delta_i^1(t) \| \leq C_1 \int_t^T [\| \Delta_{i-1}(s) \| + \| \Delta_{i-1}^2(s, \tau) \|_1] \, ds \, . \tag{2.23}$$

Совершенно аналогично, используя соотношения для $\Delta_i^2(t, \tau)$, которые получаются из (2.5), убеждаемся в существовании такой постоянной $C_2 \geq 0$, что

$$\| \Delta_i^2(t, \tau) \|_1 \leq C_2 \int_t^T [\| \Delta_{i-1}^1(s) \| + \| \Delta_{i-1}(s, \tau) \|_2] \, ds \, . \tag{2.24}$$

Оценку для C_2, подобную (2.22), для краткости опустим. Отметим лишь, что она получается из соотношений для $\Delta_i^2(t, \tau)$ точно так же, как и (2.22) получена из (2.18). Из (2.23), (2.24) имеем

$$\| \Delta_i^1(t) \| + \| \Delta_i^2(t, \tau) \|_1 \leq (C_1 + C_2) \int_t^T [\| \Delta_{i-1}^1(s) + \| \Delta_{i-1}^2(s, \tau) \|_2] \, ds \, .$$

Значит, итерируя, получаем

$$\|\Delta^1_{i+1}(t)\| + \|\Delta^2_{i+1}(t,\tau)\|_1 \leq \frac{(C_1+C_2)^i C_3(T-t)^i}{i!}, \qquad i=0,1,\ldots, \tag{2.25}$$

где постоянная C_3 есть ввиду (2.19)

$$C_3 = \max\,[\max_t \|P_1(t)\|, \max_t \|R_1(t,\tau,\tau)\|_2], \qquad 0 \leq t \leq T.$$

Далее ясно, в силу поточечной сходимости $P_i(t)$ к $P(t)$, что

$$P(t) - P_i(t) = \sum_{j=i}^{\infty}\big(P_{j+1}(t) - P_j(t)\big) = -\sum_{j=1}^{\infty}\Delta^1_{j+1}(t).$$

Таким образом

$$\|P(t) - P_i(t)\| \leq \sum_{j=1}^{\infty}\frac{(C_1+C_2)^j C_3(T-t)^j}{j!}, \quad i=1,2,\ldots. \tag{2.26}$$

Точно так же устанавливаем, что

$$\|Q(t,\tau) - Q_i(t,\tau)\|_1 \leq \sum_{j=1}^{\infty}\frac{(C_1+C_2)^j C_3(T-t)^j}{j!}. \tag{2.27}$$

Наконец, из уравнения для $\Delta^3_i(t,\tau,\varrho)$ следует, что для $i \geq 2$

$$\|\Delta^3_i(t,\tau,\varrho)\|_2 \leq C_4 \int_t^T [\|\Delta^1_{i-1}(s)\| + \|\Delta^2_{i-1}(s,\tau)\|_1]\,ds,$$

где постоянная C_4 легко оценивается через параметры системы (1.1) подобно тому, как была получена оценка для C_1. Отсюда и из (2.25) аналогично выводу (2.26) следует, что

$$\|R(t,\tau,\varrho) - R_i(t,\tau,\varrho)\|_2 \leq \sum_{j=1}^{\infty}\frac{(C_1+C)^{j+1} C_3 C_4 (T-t)^{j-1}}{(j-1)}. \tag{2.28}$$
$$i=1,2,\ldots$$

Формулы (2.26)—(2.28) и оценивают скорость сходимости метода последовательных приближений в задаче оптимального управления (1.1)—(1.3). Итак, доказана

Теорема. Пусть коэффициенты в (1.1), (1.3) и начальная функция (1.2) — кусочно-непрерывны. Тогда нормы разностей между значениями коэффициентов оптимального управления и критерия качества и i-ого последовательного приближения к ним определяются формулами (2.26)—(2.28) с постоянными $C_i \geq 0$, оцениваемыми через параметры исходных соотношений (1.1), (1.3).

Входящие в (2.26)—(2.28) суммы представляют собой остаток разложения в ряд Тейлора функции $C_3 \exp (C_1 + C_2)(T - t)$. Используя это замечание и формулу остаточного члена ряда Тейлора в форме Лагранжа, делаем вывод, что

$$\sum_{j=1}^{\infty} \frac{(C_1 + C_2)^j C_3 (T - t)^j}{j!} \leq C_3 \frac{(C_1 + C_2)^i T^i}{i!} \exp (C_1 + C_2) T .$$

3. Рассмотрим теперь некоторые соотношения между оптимальными значениями управления и критерия качества в задачах 1 и 2. Известно, что оптимальное управление в задаче 2 есть ([10])

$$v_0(t) = - N^{-1}(t) D'(t) \overline{P}_1(t) y(t) , \tag{3.1}$$

а оптимальное значение функционала (1.5) равно $\varphi'(0) \overline{P}(0) \varphi(0)$, где функция $\overline{P}(t)$ — решение матричного уравнения Риккати.

Оценим вначале

$$\| u_0(t) - v_0(t) \| = \| N^{-1}(t) D'(t) \big(P(t) \varphi(0) + \int_{-h}^{0} Q(t, \tau)\big) \varphi(\tau) d\tau - \overline{P}(t) \varphi(0) \|,$$

где $\varphi(\tau)$ — произвольная кусочно-непрерывная функция на отрезке $-h \leq \leq \tau \leq 0$. Эта формула представляет собой оценку разности между оптимальными управлениями в системах (1.1) и (1.6). Ясно, что

$$\| u_0(t) - v_0(t) \| \leq \| N^{-1}(t) D'(t) \int_{-h}^{0} Q(t, \tau) \varphi(\tau) d\tau \| +$$
$$+ \| N^{-1}(t) D'(t) \big(P(t) - \overline{P}(t)\big) \varphi(0) \| . \tag{3.2}$$

Поэтому достаточно оценить $\| P(t) - \overline{P}(t) \|$, ибо оценка $\| Q(t, \tau) \|_1$ уже получена в (2.15), а $\varphi(\tau)$ считается заданной.

Из (1.7) и (1.8) заключаем, что

$$P^{\cdot}(t) + \big(A'(t) + B'(t)\big) P(t) + P(t) \big(A(t) + B(t)\big) + \tag{3.3}$$
$$+ \left(\frac{\partial Q(t, -h\varepsilon)}{\partial \tau} + \frac{\partial Q'(t, -h\varepsilon)}{\partial \tau}\right) h = P'(t) D_1(t) P(t) - M(t), \qquad P(T) = H,$$

где $\dfrac{\partial Q(t, -h\varepsilon)}{\partial \tau}$ есть производная $\dfrac{\partial Q(t, \tau)}{\partial \tau}$ в точке $\tau = -h\varepsilon$, $0 \leq \varepsilon \leq 1$. Напомним ([10]), что матрица $\overline{P}(t)$ есть решение задачи (3.3) при $h = 0$.

Оценим вначале $\left\| \dfrac{\partial Q(t, \tau)}{\partial \tau} \right\|_2$, не выписывая явных выражений мажорирующих функций через параметры систем ввиду их громоздкости, а ограничиваясь лишь описанием алгоритма их получения.

Из (2.5) по точечной сходимости P_i, Q_i к P, Q и теоремы Лебега о предельном переходе под знаком интеграла следует, что матрица $Q(t, \tau)$ удовлетворяет интегральному тождеству (2.5), всюду в котором следует опустить индексы i, $i-1$. Продифференцируем полученное тождество по τ. Тогда, используя еще неравенства (2.9), (2.15), (2.20), получаем, что при всех $0 \leq h \leq h_0$

$$\sup_t \left\| \frac{\partial Q(t, \tau)}{\partial \tau} \right\|_2 \leq C_1, \qquad 0 \leq t \leq T, \tag{3.4}$$

причем указанным способом постоянную C_1 нетрудно оценить и через параметры исходных соотношений (1.1), (1.3). Положим $\Delta(t) = P(t) - \overline{P}(t)$. Ввиду (2.10) и аналогичного соотношения для $\overline{P}(t)$ имеем

$$\| \Delta(t) \| \leq C_2(t), \tag{3.5}$$

где $C_2(t)$ также выражается через (2.10) и оценку $\| \overline{P}(t) \|$, для которой справедливо соотношение $(0 \leq t \leq T)$

$$\| \overline{P}(t) \| \leq \left\| \int_t^T K'(s, t) M(s) K(s, t)\, ds + K'(T, t) H K(T, t) \right\|,$$

где $K(s, t)$ — фундаментальная матрица системы $x^{\cdot}(t) = A(t) x(t)$. Из уравнений для $\overline{P}(t)$ и (3.3) вытекает, что

$$\Delta^{\cdot}(t) + \big(A'(t) + B'(t)\big) \Delta(t) + \Delta(t) \big(A(t) + B(t)\big) + \tag{3.6}$$

$$+ h \left(\frac{\partial Q(t, -h\varepsilon)}{\partial \tau} + \frac{\partial Q'(t, -h\varepsilon)}{\partial \tau} \right) = P(t) D_1(t) P(t) - \overline{P}(t) D_1(t) \overline{P}(t), \quad \Delta(T) = 0.$$

Но нетрудно проверить, что

$$P(t) D_1(t) P(t) = \Delta(t) D_1(t) \Delta(t) + P(t) D_1(t) \Delta(t) + \Delta(t) D_1(t) \overline{P}(t).$$

Отсюда и из (3.6) следует, что

$$\Delta(t) = h \int_t^T z'(t, s) \left[\frac{\partial Q(s, -h\varepsilon)}{\partial \tau} + \frac{\partial Q'(s, -h\varepsilon)}{\partial \tau} \right] z(t, s)\, ds, \tag{3.7}$$

где

$$Z(t, s) = \exp \int_s^t \left[-A'(\tau) - B'(\tau) + \overline{P}(\tau) D_1(\tau) + \frac{1}{2} \Delta(\tau) D_1(\tau) \right] d\tau.$$

Из последнего равенства и соотношений (3.4)—(3.7) заключаем, что

$$\sup_t \| \Delta(t) \| \leq h C_3, \qquad 0 \leq t \leq T, \tag{3.8}$$

где постоянная C_3 выражается через параметры системы с помощью (3.4), (3.5), (3.7).

Итак, если коэффициенты (1.1), (1.3) удовлетворяют требованиям теоремы, а элементы матрицы $B(t)$ имеют ограниченную измеримую производную, то справедлива оценка (3.8).

Используя (3.2), (3.8), убеждаемся в существовании такой постоянной $\overline{C}_3$, что

$$\| u_0(t) - v_0(t) \| \leq \overline{C}_3 h \sup_\tau \| \varphi(\tau) \| , \qquad -h \leq \tau \leq 0 .$$

Незначительно модифицируя рассуждения, использованные при доказательстве оценки (3.8), можно установить следующие утверждения:

1) Если J_0 и J_1 — оптимальные значения функционалов (1.3), (1.5) в задачах 1 и 2, то при некоторой постоянной δ

$$| J_0 - J_1 | \leq \delta h \, [\sup_\tau \| \varphi(\tau) \|]^2 .$$

2) Если $x_0(t)$ и $y_0(t)$ — оптимальные траектории в системах (1.1), (1.6), то

$$\| x_0(t) - y_0(t) \| \leq C_7 h .$$

3) Пусть J_0 — оптимальное значение функционала (1.3) в задаче 1, а $\bar{J}$ — значение функционала (1.3) при управлении

$$u(t) = - N^{-1}(t) D'(t) \overline{P}(t) x(t) ,$$

где $\overline{P}(t)$ — решение задачи (3.3) с $h = 0$. Тогда

$$| J_0 - \bar{J} | \leq C_1 h \, [\sup_\tau \| \varphi(\tau) \|]^2 .$$

Литература

1. *Салуквадзе М. Е.* К задаче синтеза оптимального регулятора в линейных системах с запаздыванием, подверженных постоянно действующим возмущениям. Автоматика и телемеханика, **23** (1962), *2*, 721—731.
2. *Репин Ю. М., Третьяков В. Е.* Решение задачи об аналитическом конструировании регулятора на моделирующих устройствах. Автоматика и телемеханика, **24** (1963), *6*, 738—743.
3. *Красовский Н. Н.* Об аппроксимации одной задачи аналитического конструирования регуляторов в системе с запаздыванием. ПММ, **28** (1964) вып. 4, 319—326.
4. *Kushner, H. J., Barnea, D. I.* On the control of a linear functional-differential equation with quadratic cost. Siam journ. of control, **8** (1970), *2*, 156—174.
5. *Колмановский В. Б., Майзенберг Т. Л.* Синтез оптимальных процессов с последействием. Материалы первой конференции по оптимальному планированию и управлению народным хозяйством. Секция математических проблем. Изд-во ЦЭМИ АН СССР, 1971, 112—118.
6. *Колмановский В. Б., Майзенберг Т. Л.* Оптимальное управление стохастическими системами с последействием. Автоматика и телемеханика, *1* (1973), 47—62.
7. *Красовский Н. Н.* Об аналитическом конструировании оптимального регулятора в системе с запаздываниями времени. ПММ, **26** (1962), I, 39—51.

8. *Эльсгольц Л. Э.* Введение в теорию дифференциальных уравнений с запаздывающим аргументом. «Наука», М., 1964.
9. *Красовский Н. Н.* Некоторые задачи теории устойчивости движения. Физматгиз, М., 1959.
10. *Летов А. М.* Аналитическое конструирование регулятора. Метод динамического программирования. Автоматика и телемеханика, 22 (1961), *4*, 387—398.

On the approximation of linear control systems with time lag

V. B. KOLMANOVSKY

(Moscow)

Given linear control system with time lag of the form (1.1) where A, B, D, are matrices, x — vector, $h \geq 0$ — time lag und u — control. These control u must be found so that functional (1.3) will achieve his minimal value.

Some papers (see e.g. [1—7]) are devoted to the exact solution of this problem. In these papers [1—7] are shown, particularly, that exact definition of solutions is difficult.

The present paper is devoted to some approximation problems of optimal control for systems (1.1) with functional (1.3). It is found two different forms of approximative control (see formulae (2.4), (3.17)). After this the difference norm between approximative control, trajectories and functional value and optimal one are estimated (see formulae (2.26)—(2.28), (3.8), (3.16), (3.18). The received estimates are expressed through coefficients of initial control systems (1.1) and functional (1.3).

В. Б. Колмановский
Институт проблем механики
СССР, Москва А—40
Ленинградский просп., д. 7

Problems of Control and Information Theory, Vol. 3 (1), pp. 77–83 (1974)

CHOICE OF BASE ELEMENTS IN STRUCTURAL SYNTHESIS ALGORITHMS

A. YU. TOLMACHEVA

(Moscow)

(Received november 24, 1972)

For a given multi-output device (system of Boolean functions) and a particular structural synthesis algorithm, a technique choosing between two functionally complete sets of elements that resulting in simpler implementation of the given multi-output device is suggested. Decision rule is constructed for Boolean functions given in the tabular form. This technique is an expansion of [1]. Experimental verification of the Gavrilov–Kopylenko algorithm [2] is described as well.

1. Problem formulation

A method [3] suggested in the paper enables one to choose for a given state table (system of Boolean functions) and a particular structural synthesis algorithm that of two functionally complete sets of elements which leads to simpler physical implementation of the given multi-output device.

Consider a multi-output device with r inputs and k outputs given by a system of Boolean functions $F(n, k)$. The general formulation of the problem is as follows. There are a synthesis algorithm for multi-output logic networks and two functionally complete sets of elements:

$$f_1^{(1)}, f_2^{(1)}, \ldots, f_n^{(1)} \tag{1.1}$$

$$f_1^{(2)}, f_2^{(2)}, \ldots, f_n^{(2)}. \tag{1.2}$$

For an arbitrary multi-output device (system of Boolean functions $F(n, k)$), without synthesizing its structure one has to determine which of the two bases is more appropriate to an algorithm $\mathfrak{A}$ in the terms of structural simplicity.

Let $\mathfrak{F}(n, k)$ be a set of Boolean function systems $F(n, k)$. For a system of Boolean functions $F(n, k)$ belonging to $\mathfrak{F}(n, k)$ and algorithm $\mathfrak{A}$ denote by $L_{\mathfrak{A}}^{(1)}(F(n, k))$ $(L_{\mathfrak{A}}^{(2)}(F(n, k)))$ implementation complexity in bases (1.1) and (1.2), respectively.

The set $\mathfrak{F}(n, k)$ is broken into two subclasses $\mathfrak{F}^{(1)}(n, k)$ and $\mathfrak{F}^{(2)}(n, k)$:

$$F(n, k) \in \mathfrak{F}^{(1)}(n, k) \quad \text{if} \quad L_{\mathfrak{A}}^{(1)}(F(n, k)) \leq L_{\mathfrak{A}}^{(2)}(F(n, k))$$

$$F(n, k) \in \mathfrak{F}^{(2)}(n, k) \quad \text{if} \quad L_{\mathfrak{A}}^{(1)}(F(n, k)) > L_{\mathfrak{A}}^{(2)}(F(n, k)). \tag{1.3}$$

Thus, the original problem is that of classification (division) of the set $\mathfrak{F}(n, k)$

$$\mathfrak{F}(n, k) = \mathfrak{F}^{(1)}(n, k) \cup \mathfrak{F}^{(2)}(n, k), \tag{1.4}$$

$$\mathfrak{F}^{(1)}(n, k) \cap \mathfrak{F}^{(2)}(n, k) = \varnothing .$$

We are going to construct a program [4] which would, by means of a teaching sequence, work out a classification rule based on (1.3) for all presented sets of Boolean functions (generally speaking, different from those of the teaching sequence).

The quality of the decision rule will be verified through classification of a randomly and independently selected set of function systems called "examining sequence". This set enables us to find percentage of disagreement between decision rule classification and the true one. This percentage characterizes the decision rule quality, as well as the frequency computed through a finite sample characterizes the probability.

One of the basic questions is how to choose the elements of the teaching sample. To ensure the high quality of the decision rule, we have to envision the environment where the device will operate after completion of the teaching. That is why it pays to choose among practical systems the teaching set for division of Boolean function systems $F(n, k)$. Assume that Boolean function systems $F(n, k)$ are given in tabular form.

Our task is to classify the partially definite Boolean $m \times k$ matrices (where m is the number of inputs for which function systems of the set $\mathfrak{F}(n, k)$ are defined and k is the number of functions) so that the corresponding function systems are separated by (1.3) into two classes.

2. Set $\mathfrak{F}(n, k)$ classification algorithm

Let us proceed with the set $\mathfrak{F}(n, 1)$. For a given structural synthesis algorithm $\mathfrak{A}$ and functionally complete sets of elements (1.1) and (1.2), an algorithm is to be worked out in accordance with (1.3) classifying the set of columns of length m (corresponding to the set of values of functions $F(n, 1) \in \mathfrak{F}(n, 1)$ and denoted by $\mathfrak{F}_m(n, 1)$ into two classes $\mathfrak{F}_m^{(1)}(n, 1)$ and $\mathfrak{F}_m^{(2)}(n, 1)$ where $\mathfrak{F}_m^{(1)}(n, 1)$ $(\mathfrak{F}_m^{(2)}(n, 1))$ is the set of columns corresponding to the set $\mathfrak{F}^{(1)}(n, 1)$ $(\mathfrak{F}^{(2)}(n, 1))$. As was shown in [5], the number N of columns in the teaching sample for the algorithm below should exceed or, at least, be equal to $2\,m$, because the probability of classification error $P(N, m)$ for an arbitrary column under $N = [m\beta]$ $\left(\text{where } \beta = \dfrac{N}{m}\right)$ has the following property:

$$\lim_{N\to\infty} P(N, m) = \begin{cases} \dfrac{1}{2} & 0 \leq \beta \leq 2 \\ \dfrac{1}{2}\left(\dfrac{1}{\beta - 1}\right) & \beta > 2 \end{cases} \tag{2.1}$$

Assume we are given a set $Q = \{q^s\} = \begin{Bmatrix} q_1^s \\ q_2^s \\ \vdots \\ q_m^s \end{Bmatrix}$, $q_i \in \{0, 1, -\}$ of columns of length m belonging to $\overline{\mathfrak{F}}_m(n, 1)$ and are told which class they belong to, or, to put it differently, assume that functions from Q were synthesized in bases (1.1) and (1.2) by means of $\mathfrak{A}$ and complexities of the resulted networks were compared:

$$Q^1 \in \overline{\mathfrak{F}}_m^{(1)}(n, 1), \qquad Q^2 \in \overline{\mathfrak{F}}_m^{(2)}(n, 1) \tag{2.2}$$

where

$$Q = Q^1 \cup Q^2 \quad |Q| = |Q^1| + |Q^2| = l + r \geq 2m$$

$$Q^1 = \{q^s\}\ s = 1, 2, \ldots, l \qquad Q^2\{q^s\}\ s = l + 1, l + 2, \ldots, l + r.$$

Now, let an arbitrary function $F(n, 1)$ be given together with corresponding Boolean column φ of length m

$$\varphi = \begin{pmatrix} \varphi_1 \\ \varphi_2 \\ \vdots \\ \varphi_m \end{pmatrix} \qquad \begin{matrix} \varphi_i \in \{0, 1, -\} \\ i = 1, 2, \ldots, m. \end{matrix} \tag{2.3}$$

We have to decide in which of the bases (1.1) or (1.2) algorithm $\mathfrak{A}$ will give more advantageous implementation of $F(n, 1)$. Make use of the procedure (algorithm $A(b)$ suggested in [1]).

The set Q of columns of length m may be represented as a matrix

$$\begin{Bmatrix} q_1^1 & q_1^2 & \cdots & q_1^l & q_1^{l+1} & \cdots & q_1^{l+r} \\ q_2^1 & q_2^2 & \cdots & q_2^l & q_2^{l+1} & \cdots & q_2^{l+r} \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ q_m^1 & q_m^2 & \cdots & q_m^l & q_m^{l+1} & \cdots & q^{l+r} \end{Bmatrix}. \tag{2.4}$$

Separate among columns of length m in the matrix all sets of length b as follows. Take first the set consisting of the rows numbered $1, 2, \ldots b$. In a presented column φ which is to be classified and in columns $q^1, q^2, \ldots, \ldots, q^l, \ldots, q^{l+r}$ separate only the first b rows. Denote the resulting columns

by $\bar{q}^1, \bar{q}^2, \ldots, \bar{q}^l, \ldots, \bar{q}^{l+r}, \ldots, \bar{\varphi}$. Denote by $\Gamma^{(1)}_{1,2,\ldots b}$, the number of columns $\bar{q}^{1+l}, \ldots, \bar{q}^l$ coinciding with $\bar{\varphi}$, and by $\Gamma^{(2)}_{1,2,\ldots b}$, the number of columns $\bar{q}^{l+1}, \ldots, \bar{q}^{l+r}$ coinciding with $\bar{\varphi}$ as well. Compute $\Gamma^{(1)}_{i_1,i_2,\ldots,i_b}$ and $\Gamma^{(2)}_{i_1,i_2,\ldots,i_b}$ for all sets $i_1, i_2, \ldots, i_b$ of length b.

Define sums of $\Gamma^{(1)}_{i_1,i_2,\ldots,i_b}$ and $\Gamma^{(2)}_{i_1,i_2,\ldots,i_b}$ over all possible combinations C^b_m and denote the corresponding values by $\Gamma^{(1)}_{(b)}$ and $\Gamma^{(2)}_{(b)}$:

$$\Gamma^{(1)}_{(b)} = \sum_{\langle i_1,i_2,\ldots,i_b\rangle} \Gamma^{(1)}_{i_1,i_2,\ldots,i_b}, \tag{2.5}$$

$$\Gamma^{(2)}_{(b)} = \sum_{\langle i_1,i_2,\ldots,i_b\rangle} \Gamma^{(2)}_{i_1,i_2,\ldots,i_b}. \tag{2.6}$$

Values $\dfrac{\Gamma^{(1)}(b)}{l}$ and $\dfrac{\Gamma^{(2)}(b)}{r}$ will be referred to as "specific weights" of column φ in the first and second classes, respectively.[1] The specific weights being equal, classification is not done.

The following equality [1] holds:

$$\Gamma^{(1)}(b) = \sum_{s=1}^{l} C^b_{m-\varrho(\varphi,q^s)} \tag{2.7}$$

$$\Gamma^{(2)}(b) = \sum_{s=l+1}^{l+r} C^b_{m-\varrho(\varphi,q^s)} \tag{2.8}$$

where $\varrho(\varphi, q^s)$ is the number of non-coincidences in separate digits between columns φ and q^s. (If in any of the summandes $b > m - \varrho(\varphi, q^s)$, it is assumed to be equal to zero.)

The number of operations for classification of a vector-column φ is, evidently, on order $(l + r) \cdot m$.

Turn now to discussion of the set of systems of Boolean functions $\mathfrak{F}(n, k)$. In this case, for a given algorithm $\mathfrak{A}$ and two chosen functionally complete sets of elements (1.1) and (1.2), an algorithm should be worked out, in accordance with (1.3), for classification of a set of partially definite Boolean $m \times k$ matrices corresponding to Boolean function systems $F(n, k)$ from $\mathfrak{F}(n, k)$ (denote it by $\bar{\mathfrak{F}}_m(n, k)$) into clases $\bar{\mathfrak{F}}^{(1)}_m(n, k)$ and $\bar{\mathfrak{F}}^{(2)}_m(n, k)$.

Let $Q = Q^1 \cup Q^2$ be a teaching sequence of such matrices and $Q^1 \in \bar{\mathfrak{F}}^{(1)}_m(n, k)$, $Q^2 \in \bar{\mathfrak{F}}^{(2)}_m(n, k)$:

$$Q = \left\{ \begin{array}{l} \begin{pmatrix} q^s_{11} & q^s_{21} & \cdots & q^s_{k1} \\ q^s_{21} & q^s_{22} & \cdots & q^s_{k2} \\ \vdots & \vdots & & \vdots \\ q^s_{1m} & q^s_{2m} & \cdots & q^s_{km} \end{pmatrix} \end{array} \right\} \begin{array}{l} i = 1, \quad s = 1, 2, \ldots, l \\ \\ i = 2, \quad s = l + 1, l + 2, \ldots, l + r \end{array} \tag{2.9}$$

[1] This corresponds to voting for both classes in [1].

and let an arbtirary matrix φ be given

$$\varphi = \begin{pmatrix} \varphi_{11} & \varphi_{21} & \cdots & \varphi_{k1} \\ \varphi_{12} & \varphi_{22} & \cdots & \varphi_{k2} \\ \vdots & \vdots & & \vdots \\ \varphi_{1m} & \varphi_{2m} & \cdots & \varphi_{km} \end{pmatrix}. \tag{2.10}$$

We have to determine which of the bases (1.1) and (1.2) is more appropriate to synthesis of multi-output device corresponding to the Boolean function system $F(n, k)$ (matrix φ). First, reorder the output variables of the function systems $F(n, k)$ involved into the teaching sequence Q and those of the function system being classified (i.e. all columns of matrices of set Q and matrix φ) by Sholomov formula [6] estimating implementation complexity for each output

$$I_p = N_p \log N_p - n_{1p} \log n_{1p} - n_{0p} \log n_{0p} \qquad p = 1, 2, \ldots, k \tag{2.11}$$

where n_{1p} is the number of working states of pth function,

n_{0p} is the number of non-working states of pth function,

N_p is total number of states of pth function; so that $I_{p1} \leq I_{p2}$ at $p_1 < p_2$.[2]

Next, to each output Z_p of system $F(n, k)$ (column p of matrix φ) the above voting algorithm $A(b)$ is applied, and if the matrix φ as a whole is tested to belong either to $\mathfrak{F}^{(1)}_{(nk)}$ of to $\mathfrak{F}^{(1)}_{(n,k)}$, its specific weight in respective classes is determined as follows:

$$\Gamma^{(1)}(b) = \sum_{p=1}^{k} \sum_{s=1}^{l} C^{b}_{m-\varrho(\varphi_{p_i}, q^s_{p_i})}, \tag{2.12}$$

$$\Gamma^{(2)}(b) = \sum_{p=s}^{k} \sum_{s=1}^{l} C^{b}_{m-\varrho(\varphi_{p_i}, q^s_{p_i})}, \tag{2.13}$$

where $\varrho(\varphi_{p_i}, q^s_{p_i})$ is the number of non-coincidences in separate digits of pth of column of φ and pth column of $\varrho \in Q$.

Matrix φ is assigned to the class $\mathfrak{F}^{(1)}(n, k)$ if $\frac{\Gamma^{(1)}(b)}{l} > \frac{\Gamma^{(2)}(b)}{r}$. The number classification operations is of order $(l + r) \cdot m \cdot k$.

[2] It is possible to reorder the output variables of systems of functions involved into the teaching sequence Q so that output pth variable has minimal Hamming distance from the output pth variable of the function corresponding to matrix φ, but it would take more operations.

3. Practical verification of the classification algorithm

For structural synthesis the Gavrilov–Kopylenko arbitrary element base algorithm [2] for logic networks was used, and 4-input invert plus NOT and 5-input majority element plus NOT were taken as bases (1.1) and (1.2), respectively. Verification of the above algorithm $A(1)$ was carried out for a set of completely definite functions of 5 variables $\mathfrak{F}(5, 1)(n = S, k = 1)$. Computations were carried out on the BESM–4M computer, programs were written in the Automation Institute of the Kirghiz Academy of Sciences. The teaching sequence Q included 64 Boolean functions obtained through the randomizer as follows: 32 top digit in the binary representation of a random number were regarded as values of completely definite Boolean function of 5 variables. For the Boolean functions of Q, logic networks in bases (1.1) and (1.2) were synthesized, and in accordance with (1.3), Q was broken into Q_1, and Q_2. Next, for a testing set of Boolean functions K consisting of 20 functions obtained through the randomizer as well, percentage of disagreement between decision rule classification (see section 2) and that of (1.3) was determined. The error was 15%, but since the teaching sample Q was rather small, results of the algorithm $A(1)$ may be regarded as fair enough.

Results of the paper are readily generalized to the case of more than two functionally complete sets of elements. Moreover, the method above may be used for effectiveness comparison of two structural synthesis algorithms.

References

1. *Zhuravlev Yu. I., Tulyaganov Sh. B.*, "Izmerenie vazhnosti priznaka" (Attribute importance measurement), in "Voprosi kibernetiki" (Subjects of the cybernetics), Tashkent, 1970 (in Russian), vyp. 38, 29–42.
2. *Gavrilov M. A., Kopylenko V. M.*, "Transition tables method of synthesis of multi-output combinational structures with arbitrary components", in "Abstraktnaya i strukturnaya teoriya releinykh ustroistv i konechnykh avtomatov" (Abstract and structural theory of switching circuits and finite automata), Moscow "Nauka", 1972 (in Russian), 7–32.
3. *Tolmacheva A. Yu.*, "Sravnitel'naya otsenka effectivnosti algoritmov sinteza konechnykh avtomatov i raspoznavanie ob'ektov" (Comparative estimates of effectiveness of finite automata synthesis algorithms and object recognition), in "Avtomatika i vychislitel'naya tekhnika" (Automation and computing technique), **2** (1972) 22 (in Russian).
4. *Vapnik V. N.*, "Zadacha obucheniya raspoznavaniyu obrazov" (The problem of pattern recognition teaching), Moscow, "Znanie", 1971 (in Russian).
5. *Cover, T. M.*, "Geometrical and statistical properties of systems of linear inequalities with applications in pattern recognition", IEEE Trans. on Electronic Computers, **EC–14** (1965) 326–334.
6. *Sholomov L. A.*, "Kriterii slozhnosti bulevykh funktsii" (Boolean function complexity criteria), in "Problemi kibernetiki" (Problems in the cybernetics), Moscow, **17** (1966) 91–127 (in Russian).

Выбор базовых элементов в алгоритмах структурного синтеза

А. Ю. ТОЛМАЧЕВА

(Москва)

В работе предлагается метод, позволяющий для данного многовыходного устройства (системы булевых функций), конкретного алгоритма структурного синтеза выбрать из двух функционально полных наборов элементов набор, который дает более простую реализацию заданного многовыходного устройства.

Метод построения решающего правила рассматривается для случая табличного задания системы булевых функций и является развитием алгоритма работы [1].

Качество решающего правила проверяется путем случайного и независимого отбора некоторого множества систем функцией, задающих многовыходные устройства, относительно которых выясняется, к какому классу относить их решающее правило. Такое множество примеров принято называть экзаменационной последовательностью.

На экзаменационной последовательности проверяется процент несовпадений в классификациях решающего правила и истинной. Найденный процент характеризует качество решающего правила так же, как вычисленная по конечной выборке частота характеризует вероятность.

Проверка метода была осуществлена экспериментально для алгоритма М. А. Гаврилова и В. М. Копыленко [2], ошибка метода около 15%.

А. Ю. Толмачева
Институт проблем управления
СССР, Москва В-485,
Профсоюзная ул., 81

Problems of Control and Information Theory, Vol. 3 (1), pp. 85—97 (1974)

РАСЧЕТ СИСТЕМЫ МАССОВОГО ОБСЛУЖИВАНИЯ С ПОРОГОМ ВКЛЮЧЕНИЯ

Ю. И. РЫЖИКОВ

(Москва)

(Поступила в редакцию 12 июня 1972 г.)

Предлагается алгоритм расчета стационарных вероятностей состояний системы типа $M/G/1$, в которой обслуживание начинается при накоплении в системе $q > 1$ требований и прекращается при полном рассасывании очереди. Получены производящая функция вероятностей и факториальные моменты длины очереди. Найдено преобразование Лапласа распределения времени ожидания и начальные моменты этого распределения. Указан способ выбора оптимального q.

В ряде систем массового обслуживания, особенно при относительно малой загрузке, оказывается целесообразным введение «порога включения», когда обслуживание начинается при скоплении в системе $q > 1$ заявок и заканчивается при полном рассасывании очереди. В подобных системах обеспечиваются значительно более благоприятные условия работы обслуживающего органа и его персонала. С другой стороны, при этом несколько ухудшается обслуживание потребителей, в частности увеличивается среднее число заявок в системе. В задачах о восстанавливаемом ЗИП [1] введение порога включения приводит к необходимости увеличения оборотного запаса агрегатов. Количественная оценка работы «пороговых» систем невозможна без распределения числа требований в них. Ниже предлагается алгоритм построения такого распределения для однолинейной системы с простейшим потоком заявок, произвольной функцией $B(\tau)$ распределения времени обслуживания и неограниченной очередью.

Алгоритм основан на введении линейчатого марковского процесса.

Будем характеризовать состояние системы числом k находящихся в ней требований ($k = 0, 1, \ldots$), и для включенного обслуживания — временем τ истекшего обслуживания очередного требования. Переходы между состояниями могут быть описаны с помощью рис. 1. (Условимся отмечать все характеристики, относящиеся к состояниям без восстановления, звездочкой.) Обозначим через λ интенсивность поступления заявок и положим $\mathcal{B}(\tau) = 1 - B(\tau)$. Будем считать, что $\mathcal{B}(+0) = 1$ (это эквивалентно допущению о невозможности мгновенного обслуживания). Повторяя проведенные в [1] рассуждения применительно к указанной выше организации обслуживания

(рис. 1), получим для стационарных условных плотностей

$$\tilde{\varrho}_k(\tau) = \frac{\varrho(\tau)}{\mathfrak{B}(\tau)}, \qquad k = 1, 2, \ldots \tag{1}$$

и вероятностей состояний систему интегро-дифференциальных уравнений

$$\left.\begin{aligned} & p_0^* = \frac{1}{\lambda} \int\limits_0^\infty \tilde{\varrho}_1(\tau)\, dB(\tau), \\ & p_k^* = p_{k-1}^*, \quad k = 1, 2, \ldots, q-1, \\ & \frac{d\tilde{\varrho}_1(\tau)}{d\tau} = -\lambda\tilde{\varrho}(\tau), \\ & \frac{d\tilde{\varrho}_k(\tau)}{d\tau} = -\lambda\tilde{\varrho}(\tau) + \lambda\tilde{\varrho}_{k-1}(\tau), \quad k = 2, 3, \ldots \end{aligned}\right\} \tag{2}$$

с начальными условиями

$$\left.\begin{aligned} & \tilde{\varrho}_k(0) = \int\limits_0^\infty \tilde{\varrho}_{k+1}(\tau)\, dB(\tau), \quad k = 1, 2, \ldots; \ k \neq q \\ & \tilde{\varrho}_q(0) = \int\limits_0^\infty \tilde{\varrho}_{q+1}(\tau)\, dB(\tau) + \lambda\, p_{q-1}^* \end{aligned}\right\}. \tag{3}$$

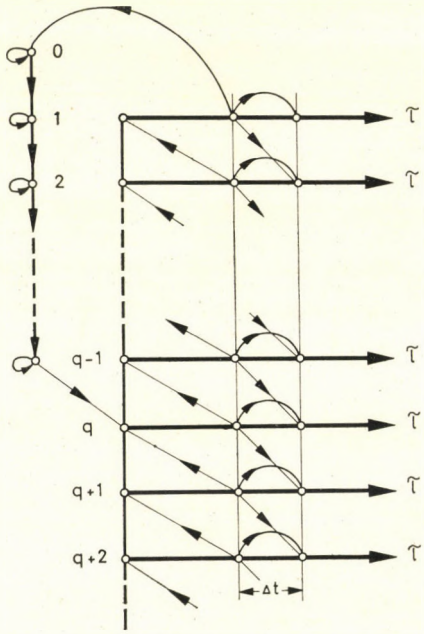


Рис. 1

Решая систему (2), начиная с уравнения для $\tilde{\varrho}_1(\tau)$ и применяя последовательные подстановки, находим

$$\tilde{\varrho}_k(\tau) = e^{-\lambda\tau} \sum_{j=1}^{k} C_j \frac{(\lambda\tau)^{k-j}}{(k-j)!}, \quad k = 1, 2, \ldots . \tag{4}$$

Очевидно, для любого k $\varrho_k(0) = C_k$. Положим

$$A_j = \int_0^\infty \frac{(\lambda\tau)^j}{j!} e^{-\lambda\tau} dB(\tau), \quad j = 0, 1, \ldots \tag{5}$$

— это вероятность прихода ровно j новых заявок за время обслуживания одной старой. Тогда

$$\int_0^\infty \tilde{\varrho}_{k+1}(\tau)\, dB(\tau) = \int_0^\infty \sum_{j=1}^{k+1} \frac{(\lambda\tau)^{k+1-j} e^{-\lambda\tau}}{(k+1-j)!} dB(\tau) = \sum_{j=1}^{k+1} C_j A_{k+1-j},$$

и система (3) может быть представлена в форме

$$\left.\begin{aligned} C_k &= \sum_{j=1}^{k+1} C_j A_{k+1-j}, \quad k = 1, 2, \ldots, q-1, q+1, q+2, \ldots, \\ C_q &= \sum_{j=1}^{q+1} C_j A_{q+1-j} + \lambda p_0^* \end{aligned}\right\}. \tag{6}$$

Стационарные вероятности подсчитываются согласно

$$p_k = \int_0^\infty \varrho_k(\tau)\, d\tau = \int_0^\infty \tilde{\varrho}_k(\tau)\, \mathcal{B}(\tau)\, d\tau .$$

Можно показать, что

$$\int_0^\infty \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau} \mathcal{B}(\tau)\, d\tau = \frac{1}{\lambda}\left(1 - \sum_{z=0}^{k} A_z\right),$$

откуда с учетом выражений (4) следует

$$p_k = \frac{1}{\lambda} \sum_{j=1}^{k} C_j \left(1 - \sum_{z=0}^{k-j} A_z\right), \quad k = 1, 2, \ldots . \tag{7}$$

Получим для p_k более простое представление. Прежде всего,

$$p_k = \frac{1}{\lambda}\left[\sum_{j=1}^{k-1} C_j \left(1 - \sum_{z=0}^{k-1-j} A_z - A_{k-j}\right) + C_k(1 - A_0)\right] =$$

$$= \frac{1}{\lambda}\left[\sum_{j=1}^{k-1} C_j \left(1 - \sum_{z=0}^{k-1-j} A_z\right) - \sum_{j=1}^{k} C_j A_{k-j} + C_k\right] =$$

$$= p_{k-1} + \frac{1}{\lambda}\left(C_k - \sum_{j=1}^{k} C_f A_{k-j}\right). \tag{8}$$

Согласно (6), стоящая в скобках сумма при всех $k \neq q$ есть C_{k-1}. Следовательно,

$$p_k = p_{k-1} + \frac{1}{\lambda}(C_k - C_{k-1}), \quad k = 2, 3, \ldots ; k \neq q + 1 . \tag{9}$$

На основании (8) и уравнения системы (6) для C_q имеем

$$p_{q+1} = p_q + \frac{1}{\lambda}\left(C_{q+1} - \sum_{j=1}^{q+1} C_j A_{q+1-j}\right) = p_q + \frac{1}{\lambda}(C_{q+1} - C_q + \lambda p_0^*) =$$

$$= p_q + p_0^* + \frac{1}{\lambda}(C_{q+1} - C_q). \tag{10}$$

Прямой подстановкой в (7) $k = 1$ можно найти выражение для p_1. Объединив результат с (9) и (10), получаем формулы для вероятностей

$$\left.\begin{aligned} p_1 &= \frac{1}{\lambda} C_1 (1 - A_0); \\ p_k &= p_{k-1} + \frac{1}{\lambda}(C_k - C_{k-1}), \ k = 2, 3, \ldots ; \ k \neq q + 1 \\ p_{q+1} &= p_q + p_0^* + \frac{1}{\lambda}(C_{q+1} - C_q). \end{aligned}\right\} \tag{11}$$

Для расчета входящих в (11) постоянных просуммируем уравнения системы (6)

$$\sum_{k=1}^{\infty} C_k = \lambda p_0^* + \sum_{k=1}^{\infty} \sum_{j=1}^{\infty} C_j A_{k+1-j} = \lambda p_0^* + \sum_{j=1}^{\infty} C_j \sum_{i=0}^{\infty} A_i - C_1 A_0 =$$

$$= \lambda p_0^* + \sum_{j=1}^{\infty} C_j - C_1 A_0 .$$

Отсюда находим $C_1 = \frac{\lambda p_0^*}{A_0}$. Последующие постоянные легко вычисляются с помощью выводимых из той же системы (6) рекуррентных формул

$$\left.\begin{aligned}
C_1 &= \frac{\lambda p_0^*}{A_0};\\
C_2 &= \frac{1}{A_0}[C_1(1-A_1)];\\
C_k &= \frac{1}{A_0}\left[C_{k-1}(1-A_1) - \sum_{j=1}^{k-2} C_j A_{k-j}\right],\ k = 3, 4, \ldots; k \neq q+1\\
C_{q+1} &= \frac{1}{A_0}\left[C_q(1-A_1) - \lambda p_0^* - \sum_{j=1}^{q-1} C_j A_{q+1-j}\right] =\\
&= \frac{1}{A_0}\left[C_q(1-A_1) - \sum_{j=1}^{q-1} C_j A_{q+1-j}\right] - C_1.
\end{aligned}\right\} \quad (12)$$

Для завершения методики необходимо указать способ расчета p_0^*. Эта вероятность может быть найдена из уравнения баланса требований. За время T в систему в среднем поступает λT требований, и в среднем столько же должно быть обслужено. Доля времени включения обслуживания есть $1 - \sum_{k=0}^{q-1} p_k^* = 1 - qp_0^*$. Считая среднее время обслуживания равным $\bar{\tau}$, получаем уравнение баланса

$$\lambda T = \frac{T(1 - qp_0^*)}{\bar{\tau}},$$

откуда следует

$$p_0^* = \frac{1 - \lambda\bar{\tau}}{q}. \quad (13)$$

Полная вероятность наличия в системе k требований ($k = 1, 2, \ldots, q-1$) определяется по формуле

$$P_k = p_0^* + p_k. \quad (14)$$

Техника вычисления постоянных коэффициентов $\{A_j\}$ излагается, например, в [2].

Для нахождения временных показателей работы системы обслуживания необходимо получить производящую функцию стационарных вероятностей. Введем интенсивность обслуживания

$$r(\tau) = \frac{B'(\tau)}{\mathcal{B}(\tau)}.$$

Теперь система дифференциальных уравнений (2) для плотностей стационарных вероятностей может быть переписана в виде

$$\left.\begin{aligned} &\frac{d\varrho_1(\tau)}{d\tau} + [\lambda + r(\tau)]\,\varrho_1(\tau) = 0\,, \\ &\frac{d\varrho_k(\tau)}{d\tau} + [\lambda + r(\tau)]\,\varrho_k(\tau) = \lambda\varrho_{k-1}(\tau), \quad k = 2, 3, \ldots \end{aligned}\right\}. \tag{15}$$

Умножим уравнение для $\varrho_k(\tau)$ на u^k и просуммируем по всем k. В результате получим

$$\sum_{k=1}^{\infty} u^k \varrho_k'(\tau) + [\lambda + r(\tau)] \sum_{k=1}^{\infty} u^k \varrho_k(\tau) = \lambda u \sum_{k=1}^{\infty} u^k \varrho_k(\tau)\,. \tag{16}$$

Но $\mathcal{S}(u, \tau) = \sum_{k=1}^{\infty} u^k \varrho_k(\tau)$ по определению есть производящая функция. Отсюда следует, что равенство (16) эквивалентно дифференциальному уравнению

$$\frac{\partial \mathcal{S}(u, \tau)}{\partial \tau} = [\lambda(u - 1) - r(\tau)]\,\mathcal{S}(u, \tau)\,,$$

а потому

$$\mathcal{S}(u, \tau) = \mathcal{S}(u, 0) \exp\left[\lambda(u - 1)\tau - \int_0^{\tau} r(\tau) d\tau\right].$$

Поскольку $\exp\left(-\int_0^{\tau} r(\tau) d\tau\right) = \exp\left[\ln \mathcal{B}(\tau)\right] = \mathcal{B}(\tau)$, последнее равенство сводится к

$$\mathcal{S}(u, \tau) = \mathcal{S}(u, 0)\, \mathcal{B}(\tau) \exp\left[\lambda\tau(u - 1)\right]. \tag{17}$$

Определим начальное значение $\mathcal{S}(u, 0)$. С помощью (3) и первого из уравнений (2) находим, что

$$\left.\begin{aligned} \lambda p_0^* &= \int_0^{\infty} \varrho_1(\tau)\, r(\tau)\, d\tau\,, \\ \varrho_k(0) &= \int_0^{\infty} \varrho_{k+1}(\tau)\, r(\tau)\, d\tau, \quad k = 1, 2, \ldots, k \neq q, \\ \varrho_q(0) &= \int_0^{\infty} \varrho_{q+1}(\tau)\, r(\tau)\, d\tau + \lambda p_{q-1}^*\,. \end{aligned}\right\}.$$

Применив к этой системе те же преобразования, что и к (15), с учетом доказанного выше равенства p_0^* и p_{p-1}^* получаем

$$u\,\mathcal{S}(u, 0) = \lambda u(u^q - 1)\, p_0^* + \int_0^{\infty} \mathcal{S}(u, \tau)\, r(\tau)\, d\tau\,.$$

Подставив в интеграл правую часть (17), найдем

$$\int_0^\infty \mathcal{S}(u, \tau) r(\tau)\, d\tau = \mathcal{S}(u, 0)\, \beta(-\lambda u + \lambda),$$

где $\beta(s) = \int_0^\infty e^{-s\tau} dB(\tau)$ — преобразование Лапласа распределения $B(\tau)$ времени обслуживания. Значит $u\mathcal{S}(u, 0) = \lambda u(u^q - 1)\, p_0^* + \mathcal{S}(u, 0)\, \beta(\lambda - \lambda u)$, откуда

$$\mathcal{S}(u, 0) = \frac{\lambda u(u^q - 1)\, p_0^*}{u - \beta(\lambda - \lambda u)}. \tag{18}$$

С учетом (17) и (18) производящая функция стационарных вероятностей будет

$$Q(u) = \sum_{k=0}^{q-1} p_k^* u^k + \int_0^\infty \mathcal{S}(u, \tau)\, d\tau = \frac{1 - \lambda\bar{\tau}}{q} \cdot \frac{u^q - 1}{\dfrac{u}{\beta(\lambda - \lambda u)} - 1}. \tag{19}$$

При $q = 1$ эта формула совпадает с приведенной в [3] формулой для обычной однолинейной системы. С ее помощью можно найти моменты распределения числа требований. Известно, что k-й факториальный момент дискретного распределения равен

$$\nu_{[k]} = \frac{d^k}{du^k} Q(u) \bigg|_{u=1}, \tag{20}$$

причем между двумя первыми факториальными и начальными моментами существуют соотношения

$$\nu_1 = \nu_{[1]}, \quad \nu_2 = \nu_{[2]} + \nu_{[1]}. \tag{21}$$

Непосредственное дифференцирование правой части (19) с последующей подстановкой $u = 1$ приводит к громоздким выкладкам и трудно устранимой неопределенности. Поэтому преобразуем (19) в степенной ряд. Введем новую переменную $z = 1 - u$. Тогда при $u \to 1$ $z \to 0$ и (19) принимает вид

$$Q(1 - z) = p_0^* \frac{[(1 - z)^q - 1]\, \beta(\lambda z)}{(1 - z) - \beta(\lambda z)}.$$

Далее необходимо разложить числитель и знаменатель в ряды с удержанием членов, содержащих степени z не выше третьей (при отыскании моментов k-го порядка — не выше $k + 1$), сократить числитель и знаменатель на z, обратить знаменатель и результат умножить на числитель. Опуская эти

вычисления, приведем окончательный результат

$$Q(1-z) = p_0^* \frac{q}{\lambda b_1 - 1}\left\{-1 + \left[\frac{q-1}{2} + \lambda b_1 + \frac{\lambda^2 b_2}{2!\,(\lambda b_1 - 1)}\right] z + \right.$$

$$+\left[-\frac{(q-1)(q-2)}{3!} - \frac{\lambda b_1}{2!}(q-1) - \frac{\lambda^2 b_2}{2!} + \frac{\lambda^3 b_3}{3!\,(\lambda b_1 - 1)} - \frac{\lambda^4 b_2^2}{4(\lambda b_1 - 1)^2} + \right.$$

$$\left.\left. + \frac{\lambda^2 b_2}{2!\,(\lambda b_1 - 1)}\left(\frac{q-1}{2!} + \lambda b_1\right)\right] z^2\right\} + o(z^2).$$

В этой формуле и далее через $\{b_k\}$ обозначены начальные моменты распределения времени обслуживания $B(\tau)$. С учетом (20), выполненной выше замены переменных, выражения для p_0^* и равенства $b_1 = \bar{\tau}$ находим первые два момента распределения числа требований в системе

$$\left.\begin{aligned} \nu_{[1]} &= \frac{q-1}{2} + \lambda b_1 + \frac{\lambda^2 b_2}{2!\,(1-\lambda b_1)}, \\ \nu_{[2]} &= (q-1)\left(\frac{q-2}{3} + \lambda b_1\right) + \frac{\lambda^2}{1-\lambda b_1}\left[\frac{(q+1)\,b_2}{2} + \frac{(\lambda b_2)^2}{2(1-\lambda b_1)} + \frac{\lambda b_3}{3}\right] \end{aligned}\right\}. \tag{22}$$

Полученные результаты позволяют найти распределение времени ожидания начала обслуживания заявки. Если требование застало систему в состоянии $E_k(\tau)$, то оно будет принято к обслуживанию после завершения уже начатого обслуживания и полного обслуживания $(k-1)$ требований, поступивших в систему ранее данного $(k = 1, 2, \ldots)$. Если же требование застало систему в состоянии E_{k^*}, т. е. в режиме с отключенным обслуживанием, то предварительно нужно дождаться поступления еще $(q - k^* - 1)$ заявок, а затем обслужить $k^*(k^* = 0, 1, \ldots, q-1)$ заявок. Таким образом, для преобразование Лапласа-Стилтьеса стационарного распределения времени ожидания имеем

$$\omega(s) = \int_0^\infty e^{-st}\, dW(t) =$$

$$= \sum_{k=0}^{q-1} p_k^* \left(\frac{\lambda}{\lambda + s}\right)^{q-k-1} \beta^k(s) + \int_0^\infty d\tau \sum_{k=1}^\infty \varrho_k(\tau)\, \beta^{k-1}(s) \int_0^\infty e^{-sv} \frac{B'(\tau + v)}{\mathfrak{B}(\tau)}\, dv. \tag{23}$$

Здесь $\dfrac{\lambda}{\lambda + s}$ — преобразование Лапласа распределения интервалов между

требованиями для простейшего потока. Первая из входящих в (23) сумм равна

$$\sum_{k=0}^{q-1} p_k^* \left(\frac{\lambda}{\lambda+s}\right)^{q-k-1} \beta^k(s) = p_0^* \left(\frac{\lambda}{\lambda+s}\right) \sum_{k=0}^{q-1} \left[\frac{(\lambda+s)\,\beta(s)}{\lambda}\right]^k =$$

$$= \frac{\beta^q(s) - \left(\dfrac{\lambda}{\lambda+s}\right)^q}{\beta(s) - \dfrac{\lambda}{\lambda+s}}\, p_0^* . \tag{24}$$

Вторая составляющая может быть переписана в виде $\dfrac{R[\beta(s)]}{\beta(s)}$, где

$$R(u) = \int_0^\infty \frac{d\tau}{\mathfrak{B}(\tau)} \sum_{k=1}^\infty u^k\, \varrho_k(\tau) \int_0^\infty e^{-sv} B'(\tau+v)\, dv .$$

Но $\sum\limits_{k=1}^\infty u^k \varrho_k(\tau)$ есть производящая функция $\mathfrak{F}(u, \tau)$. Подставляя выражение последней из (17), убеждаемся, что

$$R(u) = \mathfrak{F}(u, 0) \int_0^\infty e^{-\lambda\tau(u-1)} \left[\int_0^\infty e^{-sv} B'(\tau+v)\, dv\right] d\tau = \mathfrak{F}(u, 0)\, \frac{\beta(s) - \beta(\lambda - \lambda u)}{\lambda - \lambda u - s} .$$

Подставив найденное выше выражение для $\mathfrak{F}(u, 0)$, получаем

$$R(u) = \frac{\lambda u(u^q - 1)\, p_0^*}{u - \beta(\lambda - \lambda u)} \cdot \frac{\beta(s) - \beta(\lambda - \lambda u)}{\lambda - \lambda u - s} ,$$

откуда

$$\frac{R[\beta(s)]}{\beta(s)} = \frac{\lambda[\beta^q(s) - 1]\, p_0^*}{\lambda - \lambda\beta(s) - s} .$$

Итак, преобразование Лапласа для распределения времени ожидания $W(t)$

$$\omega(s) = \frac{1 - \lambda\bar{\tau}}{q} \left\{ \frac{\beta^q(s) - \left(\dfrac{\lambda}{\lambda+s}\right)^q}{\beta(s) - \dfrac{\lambda}{\lambda+s}} + \frac{\lambda[\beta^q(s) - 1]}{\lambda - \lambda\beta(s) - s} \right\}. \tag{25}$$

При $q = 1$ из (25) следует известная формула Полячека—Хинчина [3].

Найдем моменты $W(t)$. На основании (23)

$$\omega(s) = 1 - \frac{w_1}{1!}s + \frac{w_2}{2!}s^2 - \frac{w_3}{3!}s^3 + \ldots,$$

где $\{w_k\}$ суть начальные моменты распределения $W(t)$. Значит w_k есть умноженный на $(-1)^k k!$ коэффициент при s^k в разложении правой части (25) по степеням s. Выполним это разложение, имея в виду получение w_1 и w_2. Положим

$$\left.\begin{aligned}
A &= \frac{1}{\lambda} - b_1, \\
B &= \frac{b_2}{2!} + \frac{q-1}{2!}b_1^2 - \frac{q+1}{2!\,\lambda^2}, \\
C &= \frac{(q+1)(q+2)}{3!\,\lambda^3} - \frac{b_3}{3!} - \frac{q-1}{2!}b_1 b_2 - \frac{(q-1)(q-2)}{3!}b_1^3, \\
D &= \frac{\dfrac{b_2}{2!} - \dfrac{1}{\lambda^2}}{\dfrac{1}{\lambda} - b_1}, \\
E &= \frac{\dfrac{1}{\lambda^3} - \dfrac{b_3}{3!}}{\dfrac{1}{\lambda} - b_1}.
\end{aligned}\right\} \tag{26}$$

В этих обозначениях

$$\frac{\beta^q(s) - \left(\dfrac{\lambda}{\lambda+s}\right)^q}{\beta(s) - \dfrac{\lambda}{\lambda+s}} = \frac{q\lambda}{1-\lambda b_1}\cdot\frac{A + Bs + Cs^2 + \ldots}{1 + Ds + Es^2 + \ldots} =$$

$$= \frac{q\lambda}{1-\lambda b_1}\{A + (B - AD)s + [C - BD + A(D^2 - E)]s^2 + \ldots\}. \tag{27}$$

Аналогичным методом выводим разложение

$$\frac{\beta^q(s) - 1}{1 - \beta(s) - \dfrac{s}{\lambda}} =$$

$$= \frac{q\lambda}{1-\lambda b_1}\{A_1 - (B_1 + A_1 D_1)s + [C_1 + B_1 D_1 + A_1(D_1^2 + E)]s^2 + \ldots\}, \tag{28}$$

где

$$\left.\begin{aligned}
A_1 &= b_1, \\
B_1 &= \frac{b_2}{2!} + \frac{q-1}{2!} b_1^2, \\
C_1 &= \frac{b_3}{3!} + \frac{q-1}{2!} b_1 b_2 + \frac{(q-1)(q-2)}{3!} b_1^3, \\
D_1 &= \frac{\dfrac{b_2}{2!}}{\dfrac{1}{\lambda} - b_1}, \\
E_1 &= \frac{b_3}{3!\left(\dfrac{1}{\lambda} - b_1\right)}.
\end{aligned}\right\} \tag{29}$$

Подставляя (27) и (28) в (25) получаем искомое разложение $\omega(s)$ по степеням s, из которого выводим интересующие нас начальные моменты распределения времени ожидания начала обслуживания

$$w_1 = \lambda(B_1 + A_1 D_1 - B + AD), \tag{30}$$

$$w_2 = 2\lambda[C - BD + A(D^2 - E) + C_1 + B_1 D_1 + A_1(D_1^2 + E_1)]. \tag{31}$$

В частности, с учетом (26) и (29) из формулы (30) можно получить

$$w_1 = \frac{q-1}{\lambda} + \frac{\lambda b_2}{2(1 - \lambda b_1)},$$

которая при $q = 1$ дает известный для традиционной системы $M|G|1$ результат ([3], стр.80).

Таким образом, основные характеристики продвижения требований в системе получены. На их основе можно решить вопрос об оптимальном значении q. Пусть, например, комплексный показатель оценки системы имеет вид

$$L(q) = g \cdot n(q) + h \cdot w_1(q),$$

где $n(q)$ — число включений системы в единицу времени, а g и h — соответствующие весовые коэффициенты. Очевидно,

$$n(q) = \lambda p_{q-1}^* = \frac{\lambda}{q}(1 - \lambda b_1),$$

так что

$$L(q) = g\frac{\lambda}{q}(1 - \lambda b_1) + h\left[\frac{q-1}{\lambda} + \frac{\lambda b_2}{2(1-\lambda b_1)}\right]. \tag{32}$$

Дифференцируя найденное соотношение по q и приравняв производную нулю, получаем оптимальный порог включения

$$\tilde{q} = \lambda\sqrt{\frac{g}{h}(1 - \lambda b_1)}. \tag{33}$$

Практически выбирается то из целых чисел $[\tilde{q}]$ и $[\tilde{q}] + 1$, которое доставляет меньшее значение правой части (32). Разумеется, выбор q можно основывать и на других вероятностных характеристиках системы.

В заключение еще раз подчеркнем достоинства исследованного класса систем массового обслуживания. «Пороговое» включение при $q > 1$ приводит к более редкой смене режимов работы системы, что повышает ее надежность и уменьшает загрузку оператора в частично автоматизированных комплексах, яысвобождая время последнего для решения других задач. Таким образом, применение пороговых систем массового обслуживания (и развитых в настовщей статье методов их расчета) следует признать весьма перспективным.

Литература

1. *Рыжиков Ю. И.* Многономенклатурная задача о восстанавливаемом ЗИП. Известия АН СССР, «Техническая кибернетика» (1969) *6*, 65—75.
2. *Рыжиков Ю. И.* Управление запасами. «Наука», М., 1969.
3. *Риордан Дж.* Вероятностные системы обслуживания. «Связь», 1966.

Analysis of mass service system with switching on threshold

Yu. I. RIZHIKOV

(Moscow)

In a number of mass service systems, especially at a relatively low load, it is desirable to include "switching threshold", then the service begins at the accumulation in the system $q > 1$ demands and stops at the full queue vanishing. The more favourable conditions for serving techniques and personnel are given in the similar systems. At the same time the serving of users become worse.

This article gives the quantitative characteristics of $M/G/1$ mass service system with an infinite queue at the q servicing threshold. The way to choose an optimal threshold value q is indicated. On the basis of diagram of the associated linear Markov process (Fig. 1) the difference equations are obtained for the state probabilities. From these equations the system of the integro-differential equations is derived for the stationary probability densities and corresponding system connecting original conditions. The evident expressions (4) for the densities through recurrently defined by (12) constants are found. Formulae (11) for the stationary state probabilities are obtained by the density integration.

The probability p_0 — formula (13) — required for the beginning of the recurrent count from the balance condition of the input and output demands in the stationary regime is obtained.

The system time characteristics are analyzed by use of the generating function apparatus. The density generating functions and state probabilities were derived on the basis of the earlier obtained differential equations for densities. The factorial moments (22) of demand number in the system, the Laplace transform of the waiting time distribution for the arbitrary demand and the first two moments of this distribution — formulae (30)—(31) — are found.

The check of the obtained formulae by $q = 1$ substitution gives the earlier known results: the probability generating function for the ordinary $M/G/1$ system and Pollaczek—Khinchin's formulae follow from (19) and (25) respectively.

In this article the question on the optimal design of threshold q is also discussed on the basis of founded characteristics of the demand advancement. For its calculating formula (33) was obtained supposing that the system efficiency is defined by the sum of two parts from which the first is proportional to the mean number of switching in the time unit and the second — to the mean time of the demand waiting.

Ю. И. Рыжиков
Институт проблем управления
СССР, Москва В-485
Профсоюзная ул., 81

6th CONGRESS OF THE INTERNATIONAL MEASUREMENT CONFEDERATION

After Budapest, Stockholm, Warsaw and Paris with the largest participation in IMEKO's 15-year history the 6th Congress was held under the headline

Acquisition and Processing of Measurement Data for Automation

in Dresden (GDR) from 17 to 23 June. In front of the Cultural Palace of the once magnificant residence of the Saxon prince-electors, the ancient city of Dresden, the flags of 30 nations announced the beginning of the Congress to the 1500 registered participants and the invited guests. After the welcoming address of Prof. Harry Trumpold, head of the National Organizing Committee, Dr. Herbert Weiz, Deputy Prime Minister of the GDR, patron of the event, opened the official program.

167 high-level papers (selected from more than 300 submitted manuscripts by an International Paper Selection Committee) were presented in 7 Sections.

The Sections were introduced by Survey lectures delivered by best-known scientists. The interlacing between measuring technology, information processing and automation, particularly of measurement and production processes require close cooperation among scientists and engineers working in these fields. This cooperation is furthered by the Five International Associations Coordinating Committee (FIACC) consisting of IMEKO and the following four non-governmental organizations:

AICA — Association Internationale pour le Calcul Analogique
IFAC — International Federation of Automatic Control
IFIP — International Federation for Information Processing
IFORS — International Federation of Operational Research Societies.

In view of these aims Prof. Dr. T. J. Williams, USA was invited to deliver a survey lecture on behalf of IFIP to Section 3 and W. E. Miller, USA on behalf of IFAC to introduce Section 4. *Section 1* — "Theory of Measurement Systems" dealt with the fundamentals of measurements and the theory of measuring methods, systems and instruments, the theory and practice of error analysis and the relevant economic aspects. The starting point hereto was the survey lecture of Prof. L. Finkelstein, from The City University of London entitled Fundamental Concepts of Measurement. It was to be seen from the contributions that there is a hard work going on concerning the theory of measuring systems in several countries, though with different main points and development tendencies. Consequently an exchange of experience would significantly extend the scope of knowledge available. The Automation of Instrument Design, Manufacture and Testing — this was the subject of *Section 2*. The survey paper of Prof G. M. Aristov, USSR bore the title "Current Trends in the Automation of Processes in Design, Construction and Evaluation of Measuring Instruments". Topics of lectures were the computer-aided design of elements and measuring systems, the automatic assembly as well as the automatic testing of instruments and the error location.

The survey lecture of Prof. T. J. Williams, USA "Means and Methods for Measurement and Computer Interfacing" introduced the similar-headed *Section 3*. The section papers gave some examples on complex measuring systems applied in data acquisition

and processing and provided an overall picture of the present situation regarding the interfacing of measurement systems and transducers with computers of partial pre-processing abilities and the organizing of programming for data acquisition and processing — bringing the generally arising errors clearly to the surface. They also treated the problem of an optimal division of tasks between the measurement system and the computer.

The papers of *Section 4* "On-like Measurement in Continuous Processes" handled mechanical, thermal, electrical and chemical process, resp. system parameters and the questions of humidity measurement. The survey lecture by W. E. Miller and W. G. Wright, USA, treated "Sensors for On-line Process Control". New, in practice not yet adopted measuring methods were described to improve and expand the measuring principles of proved value, e.g. using microwave, laser, ultrasonic and X-ray technique in production engineering. Great emphasize was put on the principles and methods in connection with the quickly developing oil- and gas-industry, first of all as to the transporting facilities through the pipelines. The most interesting lectures were those introducing devices for the measurement of process parameters, which have absolutely no mechanically working parts and are very reliable and of high durability. They are particularly experimented with in the chemical industry. An especially valuable contribution from Japan should perhaps be still mentioned here about the development of a flowmeter for fluid metals. *Section 5* had the title of "Measurement in Discontinuous Processes" which meant geometrical quantities, stochastic spatial and time functions as well as multidimensional measurements. Prof. H. Trumpold and Prof. W. Lotze, GDR in their survey lecture spoke about the importance of data processing in the machine engineering. Investigations of a working group led to the conclusion that by combining the usual mechanical and electronical measuring methods with the modern, optoelectronic ones optimal results can be achieved in solving the difficulties of surface measurement. *Section 6* dealt with an up-to-date problem causing a lot of troubles all over the world and crying therefore for a possibly rapid solution, namely the *measurement of air and water pollution*. In the complex measurement systems for the control of air pollution the following processes have been underlined:

- optical radiance absorption and reflection
- lightmodulation
- aerosol and flame ionization
- coulometric and colorimetric processes
- utilization of the piezo-effect.

In some lectures automatic systems and instruments were described for the measurement of water pollution and the quality control of water. The most significant questions are here the parameters: pH-value, conductivity, turbidity, the radiants, further the dissolved oxygen content. On the occasion of a round-table discussion the attendees could learn interesting data about the design and engineering of water pollution monitoring stations. "Testing, Calibration and Parameter Estimation of Industrial Sensors and Measuring Systems" — this was the *title* of *Section 7*. It was introduced by the survey paper of S. S. Carlisle, United Kingdom, entitled "Instrument Evaluation, Type-Approval, Standards — a Review of International Activity". The steadily growing requirements toward the accuracy of the measurement of length in the machine engineering and precision-mechanics make it inevitable to apply measuring methods in practice which until now have only been used in metrologic laboratories. Suggestions have been given for the rationalization of high-precision measurements of length. One of the interesting contributions described a laser-interferometer in metric length-units, others up-to-date automatic measuring methods and instruments. Further one dealt with the measurement of mechanical vibrations and impulses. An American scientist spoke about a new equipment consisting of an electron-computer and some specially developed peripheral instruments which is able to measure short-wave radiation at large protuberances with better results, since the computer makes an accurate classification of the tuberances possible, by the comparison of nature, intensity and duration.

Every section was followed by a round-table discussion of the most up-to-date and exciting subjects under the leadership of excellent invited experts. Numerous useful and significant proposals turned up which must be considered regarding future international cooperation. Despite of the fact that 3 different sections ran in parallel, all of

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

them were well-attended and participation cards for the round-table discussions had to be obtained well ahead.

Tired participants had the opportunity to get refreshed — and acquainted with each other — on the exceptionally pleasant social gatherings, in the romantic Albrechtsburg of Meissen built in the 15th century, the modern restaurant "Am Zwinger", the wonderful evening boat-trip on the river Elbe by the famous "Weisse Flotte", or during the successful banquet at "Weisser Hirsch", one of the most beautiful spots in Dresden's surroundings.

By the closing speech of IMEKO's newly elected President, Dr. S. S. Carlisle, Director of the SIRA Institute in England the five days' work of the Congress came to the end.

Experts of the measurement science and technology will meet again under the auspices of the International Measurement Confederation in London, 1976, at IMEKO 7.

Participants on various occasions expressed their great satisfaction and appreciation concerning the excellent organization and cordial hospitality of the Society for Measurement and Automation, our Member Organization in the GDR. The Dresden Congress is an important milestone, not only because it closed the first 15 years' period in the history of the International Measurement Confederation, but also showed the way how to go on in the future.

Papers presented at the Congress will appear in ACTA IMEKO 1973 published jointly by Akadémiai Kiadó (the Publishing House of the Hungarian Academy of Sciences) and the North-Holland Publishing Co., Amsterdam. It is due to come out at the beginning of 1974.

Concurrently with the 6th Congress representatives of the Member Organizations met again to talk about organizational questions. The usual yearly Session was a really plenary one since delegates from all of the IMEKO countries were attending. (We have at present 20 Member Organizations.)

On the basis of a recent exchange in the Confederation's Statutes a President was elected in the person of Mr. S. S. Carlisle, Director of the world-famous SIRA Institute, a research and development centre for measurement and control in the United Kingdom, who has been associated with the British delegation to IMEKO for well over ten years.

Dr. O. Winkler, President of the VDI/VDE Gesellschaft für Mess- und Regelungstechnik, our Member Organization in the Federal Republic of Germany became Treasurer.

The Secretary General was elected again for another 3 year period.

The 16th Plenary Session of the General Council will be held in London, May, 1974 where the first steps will be taken to shape the scientific program of IMEKO 7.

Assoc. Prof. Dr. GY. STRICKER
Secretary General of IMEKO

BOOK REVIEW

J. A. Cadzow: Discrete-time Systems. An Introduction with Interdisciplinary Applications. Prentice Hall. Inc. Englewood Cliff, New Jersey, 1973 (440 pages)

With the advent of the digital computers the discrete-data systems become more important. At the same time the theory of such systems was considered rather complex and complicated. This book develops in a simple, clear and straightforward manner the fundamental conceptions of the discrete-data systems based on the Kronecker delta response (instead of the Dirac delta distribution).

This book may be considered as an introductory analytical treatment developing in an easy going but precise manner the topics of discrete-time, discrete-data signals and systems. It shows applications of linear discrete systems from various disciplines such as physical, engineering, industrial, biological, social, sociological, management and economic systems. It gives the rules of convolution summation those of the z-transform and the inverse z-transform. Introduces the notion of system transfer function and applies it to the computation of system response. Surveys the methods of digital simulation and digital filters. Summarizes the basic rules of state-space representation.

Despite its introductory character, this book gives a sufficiently deep insight into modern system concepts. Its advantage is the completely selfcontained treatment requiring a relatively low level of mathematical preparation for comprehension. The text is particularly well adapted for self study purposes by introducing numerous examples and selected answers to problems. It shows up also the possibilities of computer programming in some cases.

F. CS

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

PROBLEMS OF CONTROL AND INFORMATION THEORY Vol. 3 (1)
ENGLISH TRANSLATION OF THE COMMUNICATIONS IN RUSSIAN

ON THE APPROXIMATION OF LINEAR CONTROL SYSTEMS WITH TIME LAY

V. B. KOLMANOVSKY

(Moscow)

Such a task of optimal control is dealt with which minimizes a quadratic functional in a linear system with delay. Two approximative methods are proposed in order to solve the task of the optimal control and the functional minimization. It succeeded to determine the difference between the approximative and exact solutions by the terms of the coefficients of the equations.

1. Let the vector $x(t)$ of the phase-coordinates be described by the following system of equations:

$$x(t) = A(t)x(t) + B(t)\,x(t-h) + D(t)\,u(t),\ 0 \leq t \leq T, \tag{1.1}$$

where $x(t) \in E_n$ (E_n is the n-dimensional Euclidean space). The $n \times n$ matrices $A(t)$, $B(t)$ and the $n \times m$ matrix $D(t)$ are given by their periodically continuous elements. Finally, $h \in [O, h_0]$ is a constant, the control is $u(t) \in E_n$. If $t \geq 0$, the solution of the system (1.1) can be obtained by the initial conditions:

$$x(\tau) = \varphi(\tau), \qquad -h \leq \tau \leq 0, \tag{1.2}$$

where $\varphi(\tau) \in E_n$ is a given periodically continuous function.

Problem 1. Minimize the following functional by selection of the control $u(t)$:

$$x'(T)\,Hx(T) + \int_0^T [x'(t)\,M(t)\,x(t) + u'(t)\,N(t)\,u(t)]\,dt. \tag{1.3}$$

There is a constant matrix $H \geq O$, and the periodically continuous matrice) $M(t) \geq 0$, $N(t) \geq 0$, $(0 \leq t \leq T)$ (prime means transpose, $H \geq 0$ indicatess a positive semidefinite matrix, $N(t) > 0$ means a positive definite one).

The approximation of Eq. (1.1) with delay was treated in [1,2], where the delay is controlled by ordinary differential equations. The foundation of the possibility of such an approximation can be found in [3], where the coeffi-

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

cients are assumed to be constant in Eqs. (1.1), (1.3) and the control $u(t)$ is a scalar function. Furthermore [3] considers the auxiliary task of minimizing the (1.3) type quadratic functional on the following trajectories of the system of ordinary differential equations:

$$\begin{aligned} &y^{(i)}(t) + ky^{(i)}(t) = ky^{(i-1)}(t), \\ &\dot{y}^{(0)}(t) = Ay^{(0)}(t) + By^{(k)}(t) + Du, \quad 0 \leq t \leq T, \end{aligned} \tag{1.4}$$

where k is an integer, $i = 1,2, \ldots k$; $y^{(i)}(t)$ is the derivative of the function. For the system (1.4) the optimal solutions, trajectories and values of the functional to be optimized are proved to tend to the corresponding values of the original Problem 1, if $k \to \infty$. The optimal control of the linear systems with delay and that of the quadratic functional (1.3) are considered in [4,5]. The successive approximation for the optimal control and the optimal value of the functional (1.3) is given in [4] by Bellman's method for the (1.1) type deterministic system and for the functional (1.3) in case of $H = 0$. Finally, formulae are given in [5] for the successive approximation to the optimal control and functional in the case of generalized stochastic systems with delay and arbitrary matrix $H \geq 0$ in (1.3). The convergence of the successive approximation, the boundary equations of partial derivatives have been proved. Further some conditions were found for the existence of the solutions of the boundary equations for determining the parameters of the (1.1) type system. The main purpose of this paper is to appreciate the convergence rate of the above-mentioned method of the successive approximations. This is treated in Section 2.

On the base of results of [3—6] in linear systems with delay with quadratic quality-criteria there exists a linear functional derived from the trajectories of the process, which is related to the length of the delay time. It means that for the duration of the delay time the whole trajectory must be stored in the memory. This latter cannot, however, always be realized.

The second purpose of this paper is to find a relation between the first problem and the following second one.

Problem 2. Determine control $v(t)$, which minimizes the functional

$$y'(T)\, Hy(T) + \int_0^T [y'(t)\, M(t)\, y(t) + v'(t)\, N(t)\, v(t)]\, dt, \tag{1.5}$$

where $y(t)$ is the solution of the following system of ordinary differential equations

$$\begin{aligned} &\dot{y}(t) = [A(t) + B(t)]\, y(t) + D(t)\, v(t), \\ &y(0) = \varphi(0), \qquad 0 \leq t \leq T. \end{aligned} \tag{1.6}$$

The difference of the optimal quantities in the first task and the second one is given in Section 3.

It should be noticed that the limitation at the (1.1)-type system is done only to simplify the explanations. The results in Sections 2 and 3 are valid with little alterations at linear stochastic systems with delay too, which are known to have discrete and distributed delays.

Some necessary results from [4, 5] are introduced in order to formulate more precisely the convergence rate of the successive approximations.

Let the $n \times m$ $P(t) > 0$, $Q(t, \tau)$, $R(t, \tau, \varrho) = R'(t, \varrho, \tau)$ $(0 \leq t \leq T$, $-h \leq \tau, \varrho \leq 0)$ quadratic matrices be the solutions of the following system of equations:

$$
\begin{aligned}
&D_1(t) = D(t)\, N^{-1}(t)\, D'(t), \quad t - h < \tau_1,\ \varrho_1 \leq t,\\
&\dot{p}(t) + A'(t)\, p(t) + p(t)\, A(t) + Q(t, 0) + Q'(t, 0) = p(t)\, D_1(t)\, P(t) - M(t),\\
&A'(t)\, Q(t, \tau) + R(t, 0, \tau) + \frac{\partial Q}{\partial t}(t, \tau_1 - t) = P(t)\, D_1(t)\, Q(t, \tau),\\
&\frac{\partial R(t, \tau_1 - t, \varrho_1 - t)}{\partial t} = Q'(t, \tau)\, D_1(t)\, Q(t, \varrho).
\end{aligned}
\tag{1.7}
$$

The boundary conditions for this system of equations are:

$$
\begin{aligned}
&P(t) = H, Q(T, \tau) = R(T, \tau, \varrho) \equiv 0, \quad -h < \tau, \varrho \leq 0,\\
&B'(t)\, P(t) - Q'(t_1 - h) = 0, \quad 0 \leq t \leq T,\\
&B'(t)\, Q(t_1 \tau) - R(t_1 - h, \tau) + Q'(t_1 \tau)\, B(t) - R'(t, \tau, -h) = 0.
\end{aligned}
\tag{1.8}
$$

Let $P_0(t) \geq 0$, $Q_0(t, \tau)$, $R_0(t, \tau, \varrho)(R_0'(t, \tau, \varrho) = R_0(t, \varrho, \tau))$ be arbitrary $n \times m$ quadratic matrices, continuously differentiable in t, τ, ϱ satisfying the conditions of (1.8). Moreover the sequence of matrices $(P_i(t), Q_i(t, \tau), R_i(t, \tau, \varrho))$ is determined as the solution of (1.7) and (1.8). Instead of $P, Q, R; P_i, Q_i, R_i$ are in their solutions and instead of $P(t)D_1(t)P(t)$, $P(t)D_1(t)Q(t)$, $Q'(t, \tau)D_1(t)Q(t, \varrho)$ there are respectively.

$$
\begin{aligned}
&P_{i-1}(t)\, D_1(t)\, P_i(t) + P_i(t)\, D_1(t)\, P_{i-1}(t) - P_{i-1}(t)\, D_1(t)\, P_{i-1}(t)\\
&P_{i-1}(t)\, D_1(t)\, Q_i(t, \tau) + P_i(t)\, D_1(t)\, Q_{i-1}(t, \tau) - P_{i-1}(t)\, D_1(t)\, Q_{i-1}(t, \tau)\\
&Q_{1-1}'(t, \tau)\, D_1(t)\, Q_i(t, \varrho) + Q_i'(t, \tau)\, D_1(t)\, Q_{i-1}(t, \varrho) - Q_{i-1}'(t, \tau)\, D_1(t)\, Q_{i-1}(t, \varrho).
\end{aligned}
\tag{1.9}
$$

In accordance with [6] sequences $P_i(t)$, $Q_i(t, \tau)$, $R_i(t, \tau, \varrho)$ are pointwise convergent to the solution of the tasks (1.7) and (1.8). The following formulae deter-

mine the optimal control $u(t)$ and the value I_0 of the functional (1.3):

$$u_0(t) = -N^{-1} D'(t) [P(t) x(t) + \int_{-h}^{0} Q(t, \tau) x(t + \tau) d\tau],$$

$$I_0 = \varphi'(0) P(0) \varphi(0) + \varphi'(0) \int_{-h}^{0} Q(0, \tau) \varphi(\tau) d\tau + \tag{1.10}$$

$$+ \int_{-h}^{0} \varphi'(\tau) Q'(0, \tau) d\tau \varphi(0) + \int_{-h}^{0} \int_{-h}^{0} \varphi'(\tau) R(0, \tau, \varrho) \varphi(\varrho) d\tau d\varrho,$$

where $\varphi(\tau)$ is the (1.2) initial condition of the system (1.1). Let $A^{ij}(t)$ the entry of the matrix $A(t)$. Let

$$\| A(t) \| = \max_{i,j} | A^{ij}(t) |, \ (i, j = 1, \ldots, n). \tag{1.11}$$

To determine the convergence rate is just to determine the norm of the following differences

$$\| P(t) - P_i(t) \|, \| Q(t, \tau) - Q_i(t, \tau) \|_1 = \max_\tau \| Q(t, \tau) - Q_i(t, \tau) \|,$$

$$\| R(t, \tau, \varrho) - R_i(t, \tau, \varrho) \|_2 = \max_{\tau, \varrho} \| R(t, \tau, \varrho) - R_i(t, \tau, \varrho) \|. \tag{1.12}$$

This proof is given in the Appendix, due to some technical difficulties.

2. Let

$$\alpha_i^P = \| P(t) - P_i(t) \|, \quad \alpha_i^Q = \| Q(t, \tau) - Q_i(t, \tau) \|_1, \tag{2.1}$$

$$\alpha_i^R = \| R(t, \tau, \varrho) - R_i(t, \tau, \varrho) \|_2 . \tag{2.2}$$

The estimation of formulae (2.1), (2.2) is in two steps.

1) The norm of the matrices $P_i(t), Q_i(t, \tau), R_i(t, \tau, \varrho)$ are estimated at first. In order to do this formulae are used where P_i, Q_i, R_i are determined in terms of $P_{i-1}, Q_{i-1}, R_{i-1}$ ([5—6]):

$$P_i(t) = \int_t^T [K_i'(s, t) (M(s) + P_{i-1}(s) D_1(s) P_{i-1}(s)) K_i(s, t) +$$

$$+ \int_{s-h}^{s} K_i'(s, t) P_{i-1}(s) D_1(s) Q_{i-1}(s, \tau - s) K_i(\tau, t) d\tau + \tag{2.3}$$

$$+ \int_{s-h}^{s} \int_{s-h}^{s} K_i'(\tau, t) Q_{i-1}'(s, \tau - s) D_1(s) Q_{i-1}(s, s - s) K_i(\varrho, t) d\tau d\varrho] ds +$$

$$+ K_i'(T, t) H K_i(T, t),$$

where $K_i(s, t)$ is the fundamental solution of the system (1.1) of the equations with control $u(t)$:

$$u_i(t) = -N^{-1}(t) D'(t) [P_{i-1}(t) x(t) + \int_{-h}^{0} Q_{i-1}(t, \tau) x(t + \tau) d\tau. \tag{2.4}$$

The matrix $Q_i(t, \tau)$ can be determined by:

$$
\begin{aligned}
Q_i(t, \tau) = & \int_t^T [K_i'(s, t)\,(M(s) + P_{i-1}(s)\,D_1(s)\,P_{i-1}(s))\,K_{ii}(s, t, \tau) + \\
& + \int_{s-h}^{s} (K_i'(s, t)\,P_{i-1}(s)\,D_1(s)\,Q_{i-1}(s, s_1 - s)\,K_{ii}(s_1, t, \tau) + \\
& + K_1'(s_1, t)\,Q_{i-1}'(s, s_1 - s)\,D_1(s)\,P_{i-1}(s)\,K_{ii}(s, t, \tau))\,ds_1 + \qquad (2.5) \\
& + 2\int_{s-h}^{s}\int_{s-h}^{s} ds_1 ds_2\,K_i'(s_1, t)\;Q_{i-1}'(s, s_1 - s)\;D_1(s)\;Q_{i-1}(s, s_2 - s)K_{ii}(s\;, t, \tau)]\,ds + \\
& + \int_t^{\min(T, t+\tau+h)} [K_i'(s, t)\,P_{i-1}(s)\,D_1(s)\,Q_{i-1}(s, t + \tau - s) + \\
& + 2\int_t^{s} ds_1\,K_i'(s_1, t)\,Q_{i-1}(s, s_1 - s)\,D_1(s)\,Q_{i-1}(s, t + \tau - s)]\,ds + \\
& + K_i'(T, t)\,HK_{ii}(T, \tau, t),
\end{aligned}
$$

where

$$
\begin{aligned}
K_{ii}(s, t, \tau) = & K_i(s, t + \tau + h)\,B(t + \tau \times h) - \\
& - \int_{-h}^{0} K_i(s, t + \tau - \varrho)\,D_1(t + \tau - \varrho)\,Q_{i-1}(t + \tau - \varrho), \varrho)\,d\varrho. \qquad (2.6)
\end{aligned}
$$

Finally $R_i(t, \tau, \varrho)$ can be determined by

$$
\begin{aligned}
R_i(t, \tau, \varrho) = & \int_t^T [K_{ii}'(s, t, \tau)\,(M(s) + P_{i-1}(s)\,D_1(s)\,P_{i-1}(s))\,K_{ii}'(s, t, \varrho) + \\
& + \int_{s-h}^{s} ds_1\,K_{ii}'(s, t, \tau)\,P_{i-1}(s)\,D_1(s)\,Q_{i-1}(s, s_1 - s)\,K_{ii}(s_1, t, \varrho) + \\
& + \int_{s-h}^{s}\int_{s-h}^{s} ds_1 ds_2\,K_{ii}'(s_1, t, \tau)\,Q_{i-1}'(s, s_1 - s)\,D_1(s)\,Q_{i-1}(s, s_2 - s)\,K_{ii}(s_2, t, \varrho)]\,ds + \\
& + \int_t^{\min[T, t+\tau+h]} Q_{i-1}'(s, t + \tau - s)\,D_1(s)\,P_{i-1}(s)\,K_{ii}(s, t, \varrho)\,ds + \qquad (2.7) \\
& + 2\int_t^{\min(T, t+\varrho+h)} ds \int_t^{s} ds_1\,K_{ii}'(s_1, t, \tau)\,Q_{i-1}'(s, s_1 - s)\,D_1(s)\,Q_{i-1}(s, t + \varrho - s) + \\
& + \int_t^{\min(T, t+\tau+h, t+\varrho+h)} Q_{i-1}'(s, t + \tau - s)\,D_1(s)\,Q_{i-1}(s, t + \varrho - s)\,ds + \\
& + K_{ii}'(T, t, \tau)\,HK_{ii}(T, t, \varrho).
\end{aligned}
$$

Note. In formulae (2.3)—(2.7) the solutions P_i, Q_i, R_i of the equations are given by partial derivatives with the boundary conditions (1.8). These fact is used throughout.

On the base of [4—6] the matrices given such a way satisfy the next inequality at arbitrary initial conditions $\varphi(t)$:

$$
\begin{aligned}
I_i &= \varphi'(0)\, P_i(t)\, \varphi(0) + \varphi'(0) \int_{-h}^{0} Q_i(t, \tau)\, \varphi(\tau)\, d\tau + \\
&+ \int_{-h}^{0} \varphi'(\tau)\, Q_i'(t, \tau)\, \varphi'(0)\, d\tau + \int_{-h}^{0} \int_{-h}^{0} \varphi'(\tau)\, R_i(t, \tau, \varrho)\, \varphi(\varrho)\, d\tau\, d\varrho \leq \\
&\leq I_{i-1} = \varphi'(0)\, P_{i-1}(t)\, \varphi(0) + \varphi'(0) \int_{-h}^{0} Q_{i-1}(t, \tau)\, \varphi(\tau)\, d\tau + \qquad (2.8) \\
&+ \int_{-h}^{0} \varphi'(\tau)\, Q_{i-1}'(t, \tau)\, \varphi'(0)\, d\tau + \int_{-h}^{0} \int_{-h}^{0} \varphi'(\tau)\, R_{i-1}(t, \tau, \varrho)\, \varphi(\varrho)\, d\tau d\varrho, \\
&0 \leq t \leq T; \quad i = 1,2, \ldots
\end{aligned}
$$

At the same time, in case of $t = 0$, quantity I_i is just the value of the functional (1.3) at the control $u(t)$ determined by (2.4).

In what follows, $P_0(t) = Q_0(t, \tau) = R_0(t, \tau, \varrho) = 0$ will be supposed. Then the matrix $K_1(s, t)$ is the fundamental solution of the system (1.1) if $u(t) \equiv 0$ can easily be found in direct analytical form (e.g. with "stepwise" integration [8]). The relation $P_i(t) \geq 0$, $0 \leq t \leq T$, $i = 1,2 \ldots$ follows from the determination of $P_i(t)$. This fact and (2.8) imply:

$$\| P_i(t) \| \leq \| P_1(t) \|, \qquad i = 1,2, \ldots, \tag{2.9}$$

where $P_1(t)$ is (by the determination of P_0, Q_0 and (2.3)):

$$P_1(t) = \int_{t}^{T} K_1'(s, t)\, M(s)\, K_1(s, t)\, ds + K_1'(T, t)\, H K_1(T, t), \tag{2.10}$$

i.e. $\| P_1(t) \|$ and all the $\| P_i(t) \|$ succeeded to estimate with the parameters of the initial system (1.1) and the qualitative functional (1.3)

The estimation of the norm $\| R_i(t, \tau, \varrho) \|$ follows. First of all, in accordance with (2.8), for arbitrary $0 \leq t \leq T$, $-h \leq \tau \leq 0$:

$$0 \leq R_k^{ii}(t, \tau, \tau) \leq R_1^{ii}(t, \tau, \tau), \quad k = 1,2, \ldots \tag{2.11}$$

Moreover, the following relation holds for any $k \geq 1$ and for arbitrary pairs (τ, ϱ), (i, j), any numbers q_1, q_2 (by the aid of the determinations of (2.2) and

the matrix $R_i(t, \tau, \varrho)$):

$$0 \leq q_1^2 R_k^{ii}(t, \tau, \tau) + R_k^{jj}(t, \varrho, \varrho)\, q_2^2 + (R_k^{ij}(t, \tau, \varrho)) + \\ + R_k^{ji}(t, \varrho, \tau)\, q_1 q_2;\ R_k^{ij}(t, \tau, \varrho) = R_k^{ji}(t, \varrho, \tau). \tag{2.12}$$

Hence for arbitrary, but fixed values t, τ, ϱ, i, j:

$$r^{11} = R_k^{ii}(t, \tau, \tau), \quad r^{22} = R_k^{jj}(t, \varrho, \varrho),$$
$$r^{12} = r^{21} = R_k^{ij}(t, \tau, \varrho)$$

a matrix r is defined by these entries, which is positive semidefinite. Hence

$$R_k^{ij}(t, \varrho, \tau) \mid \leq \max\, [R_k^{ii}(t, \tau, \tau), R_k^{jj}(t, \varrho, \varrho)].$$

Takint (2.11) into account

$$\| R_k(t, \tau, \varrho) \|_2 \leq \| R_1(t, \tau, \varrho) \|_2, \tag{2.13}$$

where, by (2.7)

$$R_1(t, \tau, \varrho) = \int_t^T K'_{11}(s, t, \tau)\, M(s)\, K'_{11}(s, t, \varrho)\, ds + \\ + K'_{11}(T, t, \tau)\, H K_{11}(T, t, \varrho). \tag{2.14}$$

At last function $Q_k(t, \tau)$ is examined. Analogously with (2.12) one can easily find, using (2.8), that values i, j, k, τ and q_1, q :

$$0 \leq q_1^2\, P_k^{ii}(t) + 2 q_1 q_2\, Q_k^{ij}(t, \tau) + R_k^{jj}(t, \tau).$$

Thus the matrix r_1 with the following entries is positive semidefinite

$$r_1^{11} = P_k^{ii},\ r_1^{22} = R_k^{jj}(t, \tau, \tau),\ r_1^{12} = r_1^{21} = Q_k^{ij}(t, \tau).$$

Hence

$$| Q_k^{ij}(t, \tau) | \leq \max\, [P_k^{ii}(t), R_k^{jj}(t, \tau, \tau)].$$

This relation and (2.9), (2.13) imply for any k:

$$\| Q_k(t, \tau) \|_1 \leq \max\, [\| P_1(t) \|,\ \| R_1(t, \tau, \tau) \|_2]. \tag{2.15}$$

2) The functions $\Delta_k^1(t), \Delta_k^2(t), \Delta_k^3(t, \tau, \varrho)$ are defined as follows:

$$\Delta_{k+1}^1(t) = -P_{k+1}(t) + P_k(t),\ \Delta_{k+1}^2(t, \tau) = -Q_{k+1}(t, \tau) + \\ + Q_k(t, \tau),\ \Delta_{k+1}^3(t, \tau, \varrho) = R_k(t, \tau, \varrho) - R_{k+1}(t, \tau, \varrho).$$

We shall prove that functions Δ_{k+1}^1 satisfy the same equations as the quantities P_{k+1} did at $M(t) \equiv 0$ with the homogeneous boundary conditions (1.8). The

homogeneousness of the boundary conditions is evident. The difference between the equations concerning P_k and P_{k+1} is:

$$\begin{aligned} \dot{\Delta}^1_{k+1}(t) + A'(t)\,\Delta^1_{k+1}(t) + \Delta^1_{k+1}(t)\,A(t) + \Delta^2_{k+1}(t,0) + \Delta^2_{k+1}(t,0) = \\ = \beta(t) = P_k(t)\,D_1(t)\,P_{k-1}(t) + P_{k-1}(t)\,D_1(t)\,P_k(t) - \\ - P_{k-1}(t)\,D_1(t)\,P_{k-1}(t) - P_{k+1}(t)\,D_1(t)\,P_k(t) - \\ - P_k(t)\,D_1(t)\,P_{k+1}(t) + P_k(t)\,D_1(t)\,P_k(t). \end{aligned} \tag{2.16}$$

The right hand side of Eq. (2.16) turns to be:

$$\begin{aligned} P_{k-1}(t)\,D_1(t)\,P_{k-1}(t) = \Delta^1_k(t)\,D_1(t)\,\Delta^1_k(t) + \\ + P_k(t)\,D_1(t)\,\Delta^1_k(t) + \Delta^1_k(t)\,D_1(t)\,P_k(t). \end{aligned} \tag{2.17}$$

Similarly:

$$P_k(t)\,D_1(t)\,P_{k-1}(t) - P_k(t)\,D_1(t)\,P_{k+1}(t) = P_k(t)\,(\Delta^1_{k+1}(t) + \Delta_k(t)).$$

Hence, applying also (2.17)

$$\beta(t) = P_k(t)\,D_1(t)\,\Delta^1_{k+1}(t) + \Delta^1_{k+1}(t)\,D_1(t)\,P_k(t) - \Delta^1_k(t)\,D_1(t)\,\Delta^1_k(t).$$

So the quantity $\Delta^1_{k+1}(t)$ satisfies the same equations as $P_{k+1}(t)$, in accordance with (2.16). $\Delta^2_{k+1}(t,\tau)$ and $\Delta^3_{k+1}(t,\tau,\varrho)$ satisfy the defining equations of $Q_{k+1}(t,\tau)$ and $R_{k+1}(t,\tau,\varrho)$ respectively as it can be proved analogously. The remarks at the solution of the system of Eqs. (2.3), (2.5), (2.7) of P_k, Q_k, R_k imply now that the integral representation of type (2.3), (2.5)—(2.7) still hold, if $M(t) \equiv 0$, $H = 0$ are substituted. E.g. (2.3) implies that:

$$\begin{aligned} \Delta^1_i(t) = \int_t^T [K_i^{-1}(s,t)\,\Delta^1_{i-1}(s)\,D_1(s)\,\Delta^1_{i-1}(s)\,K_i(s,t) + \\ + \int_{s-h}^{s} K_i'(s,t)\,\Delta^1_{i-1}(s)\,D_1(s)\,\Delta^2_{i-1}(s,\tau-s)\,K_i(\tau,t)\,d\tau + \\ + \int_{s-h}^{s}\int_{s-h}^{s} K_i'(\tau,t)\,\Delta^2_{i-1}(s,\tau-s)\,D_1(s)\,\Delta^2_{i-1}(s,\varrho-s)\,K_i(\varrho,t)\,d\tau d\varrho]\,ds. \end{aligned} \tag{2.18}$$

Δ^2_i and Δ^3_i can similarly be expressed by (2.5)—(2.7). If we use the same way to obtain the sequence Δ^2_j as we had for (2.13), (2.15), we can find:

$$\begin{aligned} \|\Delta^1_i(t)\| \leq \|P_1(t)\|, \quad \|\Delta^2_i(t,\tau)\|_1 \leq \max\,[\|P_1(t)\|, \|R_1(t,\tau,\tau)\|_2], \\ \|\Delta^3_i(t,\varrho,\tau)\|_2 \leq \|R_1(t,\tau,\varrho)\|_2, \qquad i = 1, 2, \ldots \end{aligned} \tag{2.19}$$

The functions $K_i(s,t)$ as the fundamental solution of the system of equations with uniformly bounded coefficients can easily be obtained in accordance

with (2.9), (2.13), (2.15), (2.3)—(2.6), (viz. [9], p. 153):

$$\| K_i(s, t) \| \leq \exp \int_t^s [\| B(\tau) \| + \| A(\tau) \| + \| N^{-1}(\tau) D'(\tau) P_1(\tau) \| + \\ + h \| Q_1(\tau, \varrho) \|_2] \, d\tau. \tag{2.20}$$

Let

$$\beta_1(t, s) = \max [\| K_i'(s, t) \| (\| D_1(s) \Delta_{i-1}^1(s) K_i(s, t) + \\ + \int_{s-h}^s D_1(s) \Delta_{i-1}^2(s, \tau - s) K_i(t, \tau) \, d\tau \|, \tag{2.21}$$

$$\| \int_{s-h}^s \int_{s-h}^s \| K_i'(\tau, t) \| D_1(s) \Delta_{i-1}^2(s, \varrho - s) K_i(\varrho, t) \, d\tau d\varrho \|],$$

$$i = 2, 3, \ldots$$

and substitute the value:

$$C_1 = \max_{t,s,i} \beta_1(t, s), \; 0 \leq t, s \leq T, \qquad i = 2,3, \ldots \tag{2.22}$$

Note, using Eqs. (2.19)—(2.22), that the constant C_1 can be expressed by the parameters of the system (1.1), (1.3).

Eqs. (2.18), (2.21), (2.22) imply

$$\| \Delta_i^1(t) \| \leq C_1 \int_t^T [\| \Delta_{i-1}^1(s) \| + \| \Delta_{i-1}^2(s, \tau) \|_1] \, ds. \tag{2.23}$$

Similarly using the expression of $\Delta_i^2(t, \tau)$ given by (2.5) one can verify the existence of the constant $C \geq 0$ such that:

$$\| \Delta_i^2(t, \tau) \|_1 \leq C_2 \int_t^T [\| \Delta_{i-1}^1(s) \| + \| \Delta_{i-1}^2(s, \tau) \|_2] \, ds. \tag{2.24}$$

The estimation of C_2 is similar to (2.22) and for the sake of briefness, will be omitted. We only remark that one can obtain it from the formula of $\Delta_i^2(t, \tau)$ exactly as (2.22) was calculated from (2.18). From (2.23), (2.24):

$$\| \Delta_i^1(t) \| + \| \Delta_i^2(t, \tau) \| \leq (C_1 + C_2) \int_t^T [\| \Delta_{i-1}^1(s) \| + \| \Delta_{i-1}^2(s, \tau) \|] \, ds.$$

It is approximately:

$$\| \Delta_{i-1}^1(t) \| + \| \Delta_{i+1}^2(t, \tau) \|_1 \leq \frac{(C_1 + C_2)^i C_3 (T - t)^i}{i!}, \tag{2.25}$$

$$i = 0,1, \ldots,$$

where (2.19) is taken into consideration for the constant C_3:

$$C_3 = \max [\max \| P_1(t) \|, \max_t \| R_1(t, \tau, \tau) \|_2], \quad 0 \leq t \leq T.$$

Furthermore, due to the pointwise convergence of $P_i(t)$ to quantity $P(t)$,

$$P(t) - P_i(t) = \sum_{j=1}^{\infty} (P_{j+1}(t) - P_j(t)) = -\sum_{j=1}^{\infty} \Delta_{j+1}^1(t).$$

Hence

$$\| P(t) - P_i(t) \| \leq \sum_{j=1}^{\infty} \frac{(C_1 + C_2)^j C_3 (T-t)^j}{j!}; \qquad i = 1, 2 \ldots \tag{2.26}$$

Similarly:

$$Q \| (t, \tau) - Q_i(t, \tau) \| \leq \sum_{j=1}^{\infty} \frac{(C_1 + C_2)^j C_3 (T-t)^j}{j!}, \quad i = 1, 2, \ldots \tag{2.27}$$

At last, Eq. $\Delta_i^3(t, \tau, \varrho)$ implies, if $i \geq 2$, that

$$\| \Delta_i^3(t, \tau, \varrho) \|_2 \leq C_4 \int_t^T [\| \Delta_{i-1}^1(s) \| + \| \Delta_{i-1}^2(s, \tau) \|] \, ds.$$

The constant C_4 is easily obtained here in terms of the parameters of (1.1) as it was the case for C_1. Similarly to (2.26), this result and (2.25) lead to

$$\| R(t, \tau, \varrho) - R_i(t, \tau, \varrho) \|_2 \leq \sum_{j=1}^{\infty} \frac{(C_1 + C_2)^{j-1} C_3 C_4 (T-t)^{j-1}}{(j-1)!}; \tag{2.28}$$

$$i = 1, 2, \ldots$$

(2.26)—(2.28) give the convergence-rate of the method of successive approximations at the tasks (1.1)—(1.3) of the optimal control.

Hence the following theorem is proved.

Theorem. Let the coefficients in (1.1), (1.3) and the initial function (1.2) be piecewise continuous. Now (2.26)—(2.28) determine the difference between the ith step of the successive approximation and between the values of the coefficients of the quality criteria with the constants $C_i \geq 0$ which are determined by the parameters of (1.1), (1.3).

The sums in (2.26)—(2.28) are the rests in the Taylor expansion of the function $C_3 \exp (C_1 + C_2)(T-t)$. Applying this remark and the formula for the rest in the Taylor expansion, according to the formula of Lagrange, we obtain

$$\sum_{j-1}^{\infty} \frac{(C_1 + C_2)^j C_3 (T-t)^j}{j!} \leq C_3 \frac{(C_1 + C_2)^i T^i}{i!} \exp (C_1 + C_2) T.$$

3. Some relations between the optimal values of the control and quality-criteria in problem 1, problem 2 will now be examined. The optimal control in the second problem is well known to be ([10]):

$$v_0(t) = -N^{-1}(t) \, D'(t) \, \bar{P}_1(t) \, y(t), \tag{3.1}$$

while the optimal value of the functional (1.5)

$$\varphi'(0)\,\bar{P}(0)\,\varphi(0),$$

where $\bar{P}(t)$ is the solution of the matrix-equation of Ricatti.

Let us evaluate at first:

$$\| u_0(t) - v_0(t) \| = \| N^{-1}(t)\, D'(t)\, (P(t)\,\varphi(0) + \int_{-h}^{0} Q(t,\tau)\,\varphi(\tau)\,d\tau - \bar{P}(t)\,\varphi(0) \|,$$

where $\varphi(\tau)$ is an arbitrary stepwise continuous function for $-h \leq \tau \leq 0$. This formula gives the difference between the optimal controls at the systems (1.1) and (1.6).

It is clear that

$$\begin{aligned} \| u_0(t) - v_0(t) \| \leq \| N^{-1}(t)\, D'(t) \int_{-h}^{0} Q(t,\tau)\,\varphi(\tau)\,d\tau \| + \\ + \| N^{-1}(t)\, D'(t)\, (P(t) - \bar{P}(t))\,\varphi(0) \| . \end{aligned} \tag{3.2}$$

Thus it is enough to determine the quantity $\| P(t) - \bar{P}(t) \|$, since the value of $\| Q(t,\tau) \|$ has been determined, and $\varphi(t)$ is supposed to be given.

Equations (1.7), (1.8) imply:

$$\begin{aligned} P(t) + (A'(t) + B'(t))\,P(t) + P(t)\,(A(t) + B(t)) + \\ + \left(\frac{\partial Q(t, -h\varepsilon)}{\partial \tau} + \frac{\partial Q'(t, -h\varepsilon)}{\partial \tau}\right) h = \\ = P'(t)\,D_1(t)\,P(t) - M(t), \qquad P(T) = H, \end{aligned} \tag{3.3}$$

where $\dfrac{\partial Q(t, -h\varepsilon)}{\partial \tau}$ is the derivative $\dfrac{\partial Q(t,\tau)}{\partial \tau}$ at the point $\tau = -h\varepsilon$, $0 \leq \varepsilon \leq 1$. Let us recollect [10] that the matrix $P(t)$ is the solution of the task (3.3) if $h = 0$.

The quantity $\left\| \dfrac{\partial Q(t,\tau)}{\partial \tau} \right\|_2$ is evaluated at first without writing formulae to majorize the functions by the parameters of the system since they were very difficult. Therefore we restrict our attention to the algorithms.

From (2.5) (i.e. from the pointwise convergence of P_i and Q_i to P and Q respectively and from the theorem of Lebesque referring to the limes of the integrands), we may conclude that the matrix $Q(t,\tau)$ satisfies the integral-identity (2.5), where all of the subscripts i, $i-1$ must be ignored. Derivate

the identity by τ. So using the inequalities (2.9), (2.15), (2.20), we get that at all the values $0 \leq h \leq h_0$:

$$\sup_t \left\| \frac{\partial Q(t,\tau)}{\partial \tau} \right\|_2 \leq C_1, \quad 0 \leq t \leq T; \tag{3.4}$$

and it is not difficult to evaluate the constant C_1 in terms of the parameters of the initial formulae (1.1), (1.3) by the presented method. Substitute: $\Delta t = P(t) - \bar{P}(t)$. Applying (2.10) and the analogous relation for $P(t)$:

$$\| \Delta(t) \| \leq C_2(t), \tag{3.5}$$

where $C_2(t)$ can be expressed by (2.10) just like the quantity $\| P(t) \|$ for which the following relations holds $(0 \leq t \leq T)$:

$$\| \bar{P}(t) \| \leq \| \int_t^T K'(s, t)\, M(s)\, K(s, t)\, ds + K'(T, t)\, HK(T, t) \|,$$

where $K(s, t)$ is the fundamental matrix of the system $\dot{x}(t) = A(t)\, x(t)$. The equations of $P(t)$ and (3.3) imply:

$$\begin{aligned} &\dot{\Delta}(t) + (A'(t) + B'(t))\, \Delta(t) + \Delta(t)\, (A(t) + B(t)) + \\ &+ h \left(\frac{\partial Q(t, -h\varepsilon)}{\partial \tau} + \frac{\partial Q'(t, -h\varepsilon)}{\partial \tau} \right) = \\ &= P(t)\, D_1(t)\, P(t) - \bar{P}(t)\, D_1(t)\, \bar{P}(t), \qquad \Delta(T) = 0. \end{aligned} \tag{3.6}$$

It may be easily checked that

$$P(t)\, D_1(t)\, P(t) = \Delta(t)\, D_1(t)\, \Delta(t) + P(t)\, D_1(t)\, \Delta(t) + \Delta(t)\, D_1(t)\, \bar{P}(t).$$

From this and (3.6) comes that

$$\Delta(t) = h \int_t^T Z'(t, s) \left[\frac{\partial Q(s, -h\varepsilon)}{\partial \tau} + \frac{\partial Q'(s, -h\varepsilon)}{\partial \tau} \right] Z(t, s)\, ds, \tag{3.7}$$

where

$$Z(t, s) = \exp \int_s^t \left[-A'(\tau) - B'(\tau) + \bar{P}(\tau)\, D_1(\tau) + \frac{1}{2} \Delta(\tau)\, D_1(\tau) \right] d\tau.$$

The latter equality and Eqs. (3.4)–(3.7) imply:

$$\sup_t \| \Delta(t) \| \leq hC_3, \qquad 0 \leq t \leq T,$$

where C_3 can be expressed by the parameters of the system (of (3.4), (3.5), (3.7)). So when the coefficients of (1.1), (1.3) satisfy the conditions of the theo-

rem and the elements of $B(t)$ have boundary-measured derivatives, inequality (3.8) holds.

Using the (3.2), (3.8) relations we can see that there exists such a constant C_3 that:

$$\| u_0(t) - v_0(t) \| \leq \bar{C}_3 h \sup_\tau \| \varphi(\tau) \|, \quad -h \leq \tau \leq 0.$$

With small modifications of the reasoning in proving (3.8) the next statements can be obtained:

1) If I_0 and I_1 are the optimal values of the functionals (1.3), (1.5) in problems 1 and 2, then for a certain value δ there is:

$$| I_0 - I_1 | \leq \delta h \, [\sup_\tau \| \varphi(\tau) \|]^2.$$

2) If $x_0(t)$ and $y_0(t)$ are optimal trajectories in the systems (1.1), (1.6) then:

$$\| x_0(t) - y_0(t) \| \leq C_7 h.$$

3) Let I_0 be the optimal value of the functional (1.3) in the 1st task and let $\bar{I}$ be the value of functional (1.3) in the case of the control

$$u(t) = -N^{-1}(t) D'(t) \bar{P}(t) x(t),$$

where $\bar{P}(t)$ is the solution of the task (3.3) at $h = 0$. Then:

$$| I_0 - I | = C_1 h \, [\sup_\tau \| \varphi(\tau) \|]^2.$$

References

1. *Салуквадзе М. Е.* К задаче синтеза оптимального регулятора в линейных системах с запаздыванием, подверженных постоянно действующим возмущениям. Автоматика и телемеханика, 1962, т. 23, № 2, 721—731.
2. *Репин Ю. М., Третьяков В. Е.* Решение задачи об аналитическом конструировании регулятора на моделирующих устройствах. Автоматика и телемеханика, 1963, т. 24, № 6, 738—743.
3. *Красовский Н. Н.* Iб аппроксимации одной задачи аналитического конструирования регуляторов в системе с запаздыванием. ПММ, 1964, т. 28, вып. 4, 319—326.
4. *Kushner, H. J., Barnea, D. I.*, On the control of a linear functional-differential equation with quadratic cost. Siam journ. of Control, 1970, vol. 8, No. 2, 156—174.
5. *Колмановский В. Б., Майзенберг Т. Л.* Синтез оптимальных процессов с последействием. Материалы первой конференции по оптимальному планированию и управлению народным хозяйством, Секция математических проблем, Изд-во ЦЭМИ АН СССР, 1971, 112—118.
6. *Колмановский В. Б., Майзенберг Т. Л.* Оптимальное управление стохастическими системами с последействием. Автоматика и телемеханика, 1973, № 1, 47—62.
7. *Красовский Н. Н.* Об аналитическом конструировании оптимального регулятора в системе с запаздываниями времени. ПММ, 1962, т. 26, вып. 1, 39—51.
8. *Эльсгольц Л. Э.* Введение в теорию дифференциальных уравнений с запаздывающим аргументом. М., «Наука», 1964.
9. *Красовский Н. Н.* Некоторые задачи теории устойчивости движения. М., Физматгиз, 1959.
10. *Летов А. М.* Аналитическое конструирование регулятора. Метод динамического программирования. Автоматика и телемеханика, 1961, т. 22, № 4, 387—398.

ANALYSIS OF MASS SERVICE SYSTEM WITH SWITCHING ON THRESHOLD

Yu. I. RIZHIKOV

(Leningrad)

An algorithm is proposed to calculate the stationary probabilities of the states of an $M/G/1$ type system. Processing in the system starts when $q > 1$ requests have accumulated and stops when the queue is completely exhausted. The generating function of the probabilities and the factorial moments of the queue length are produced. The Laplace transform of the waiting time distribution and the initial moments of this distribution are obtained and the selection method of the optimal q value is referred to.

In many mass service systems, especially in the case of small loads, the introduction of the "switching threshold" is feasible, i.e. processing in the system starts if the number of the requests is $q > 1$, and finishes when the queue is totally exhausted. In similar systems much more favourable conditions are assured for the processing organization and its personnel. On the other hand there is a certain degradation in the service of the customers at the same time, or more concretely the average number of requests in the system increases. In systems concerned with the replenishment of spares [1] the introduction of an activation threshold results in the necessity of increasing the reversible spares of aggregates. The quantitative evaluation of "threshold" systems is impossible without knowing the distribution of the number of requests in them. Below we propose an algorithm for the construction of such distributions for a single line system with the simplest request process, for any $B(\tau)$ function of the serving time distribution and an infinite queue.

The alogorithm presupposes a linear type Markov process.

The state of the system will be characterized by the number k of the requests in it ($k = 0,1, \ldots$) and for the activated processing by the time τ of serving the next request. The transitions can be described with the help of Fig. 1 (we have conditionally marked all characteristics, those belonging to the states without replenishment are marked with asterisk). Let λ denote the intensity of the arrival of requests and substitute $\mathfrak{B}(\tau) = 1 - B(\tau)$. Let us consider that $\mathfrak{B}(+0) = 1$ (this is equivalent with permitting the impossibility of instantaneous processing). Repeating the line of thought carried out in [1] for the above described processing organization (Fig. 1) we obtain the

following for stationary conditional densities:

$$\tilde{\varrho}_k(\tau) = \frac{\varrho(\tau)}{\mathcal{B}(\tau)}, \quad k = 1, 2, \ldots \tag{1}$$

For the probability of the states we obtain the following integro-differential system of equations:

$$\left.\begin{aligned} p_0^* &= \frac{1}{\lambda}\int_0^\infty \tilde{\varrho}_1(\tau)\, dB(\tau), \\ p_k^* &= \varrho_{k-1}^*, \quad k = 1, 2, \ldots, q^{-1}, \\ \frac{d\tilde{\varrho}_1(\tau)}{d\tau} &= -\lambda\tilde{\varrho}_1(\tau), \\ \frac{d\tilde{\varrho}_k(\tau)}{d\tau} &- -\lambda\tilde{\varrho}_k(\tau) + \lambda\tilde{\varrho}_{k-1}(\tau), \quad k = 2, 3, \ldots \end{aligned}\right\} \tag{2}$$

with the initial conditions:

$$\begin{aligned} \tilde{\varrho}_k(0) &= \int_0^\infty \tilde{\varrho}_{k+1}(\tau)\, dB(\tau), \quad k = 1, 2, \ldots; k \neq q \\ \tilde{\varrho}_q(0) &= \int_0^\infty \tilde{\varrho}_{q+1}(\tau)\, dB(\tau) + \lambda p_{q+1}^* . \end{aligned} \tag{3}$$

Solving the set of equations (2) beginning with the equation for $\tilde{\varrho}_1(\tau)$ and using the method of successive substitutions, we get:

$$\tilde{\varrho}_k(\tau) = e^{-\lambda t} \sum_{j=1}^{k} C_j \frac{(\lambda\tau)^{k-j}}{(k-\ )\,!} . \tag{4}$$

It can be seen that for any value of k: $\tilde{\varrho}_k(0) = C_k$. Let us substitute the following value:

$$A_j = \int_0^\infty \frac{(\lambda\tau)^j}{j\,!} e^{-\lambda\tau}\, dB(\tau), \; j = 0, 1, \ldots \tag{5}$$

this is just the probability of the arrival of exactly j new requests during the serving of a previous one. Thus:

$$\int_0^\infty \tilde{\varrho}_{k+1}(\tau)\, dB(\tau) = \int_0^\infty \sum_{j=1}^{k+1} \frac{(\lambda\tau)^{k+1-j}}{(k+1-j)!}\, dB(\tau) = \sum_{j=1}^{k+1} C_j A_{k+1-j},$$

and set (5) can be produced in the following form:

$$C_k = \sum_{j=1}^{k+1} C_j A_{k+1-j}, \; k = 1, 2, \ldots, q-1, q, q+1, q+2, \ldots,$$
$$C_q = \sum_{j=1}^{q+1} C_j A_{q+1-j} + \lambda p_0^*. \tag{6}$$

The stationary probabilities can be calculated on the basis of:

$$p_k = \int_0^\infty \varrho_k(\tau)\, d\tau = \int_0^\infty \tilde{\varrho}_k(\tau)\, \mathfrak{B}(\tau)\, d\tau.$$

It can be shown that:

$$\int_0^\infty \frac{(\lambda\tau)^k}{k!} e^{-\lambda\tau}\, \mathfrak{B}(\tau)\, d\tau = \frac{1}{\lambda}\left(1 - \sum_{z=0}^{k} A_z\right)$$

which, taking (4) into consideration, implies, that:

$$p_k = \frac{1}{\lambda} \sum_{j=1}^{k} C_j \left(1 - \sum_{z=0}^{k-j} A_z\right), \quad k = 1, 2, \ldots \tag{7}$$

A more simple representation of p_k is obtained. First of all:

$$\begin{aligned} p_k &= \frac{1}{\lambda}\left[\sum_{j=1}^{k-1} C_j \left(1 - \sum_{z=0}^{k-1-j} A_z - A_{k-j}\right) + C_k (1 - A_0)\right] = \\ &= \frac{1}{\lambda}\left[\sum_{j=1}^{k-1} C_j \left(1 - \sum_{z=0}^{k-1-j} A_z\right) - \sum_{j=1}^{k} C_j A_{k-j} + C_k\right] + \\ &= p_{k-1} + \frac{1}{\lambda}\left(C_k - \sum_{j=1}^{k} C_j A_{k-j}\right). \end{aligned} \tag{8}$$

According to (6) the sum in the brackets for all values of $k \neq q$ equals C_{k-1}. Accordingly:

$$p_k = p_{k-1} + \frac{1}{\lambda}(C_k - C_{k-1}), \; k = 2, 3, \ldots, k \neq q+1. \tag{9}$$

On the basis of (8) and (6) we get the following expression for C_q:

$$\begin{aligned} p_{q+1} &= p_q + \frac{1}{\lambda}\left(C_{q+1} - \sum_{j=1}^{q+1} C_j A_{q+1-j}\right) = p_q + \frac{1}{\lambda}(C_{q+1} - C_q + \lambda p_0^*) = \\ &= p_q + p_0^* + \frac{1}{\lambda}(C_{q+1} - C_q). \end{aligned} \tag{10}$$

Using direct substitution (setting $k = 1$ in (7)) we obtain the expression for p_1. Combining the results of (9) and (10) we get the following formula for the probabilities:

$$p_1 = \frac{1}{\lambda} C_1(1 - A_0);$$

$$p_k = p_{k-1} + \frac{1}{\lambda}(C_k - C_{k-1}), \quad k = 2, 3, \ldots; \quad k \neq q + 1, \tag{11}$$

$$p_{q+1} = p_q + p_0^* + \frac{1}{\lambda}(C_{q+1} - C_q).$$

In order to calculate the constants of (11) we add the equations in set (6)

$$\sum_{k=1}^{\infty} C_k = \lambda p_0^* + \sum_{k=1}^{\infty} \sum_{t=1}^{k+1} C_j A_{k+1-j} = \lambda p_0^* + \sum_{j=1}^{\infty} C_j \sum_{i=0}^{\infty} A_i - C_1 A_0 =$$

$$= \lambda p_0^* + \sum_{j=1}^{\infty} C_j - C_1 A_0$$

which leads to $C_1 = \dfrac{\lambda p_0^*}{A_0}$. The following constants can easily be obtained from a set of recursive formulae deducible from the same set (6)

$$C_1 = \frac{\lambda p_0^*}{A_0};$$

$$C_2 = \frac{1}{A_0}[C_1(1 - A_1)];$$

$$C_k = \left[C_{k-1}(1 - A_1) - \sum_{j=1}^{k-2} C_j A_{k-j}\right], \quad k = 3, 4, \ldots; \quad k \neq q + 1 \tag{12}$$

$$C_{q+1} = \frac{1}{A_0}\left[C_q(1 - A_1) - \lambda p_0^* - \sum_{j=1}^{q-1} C_j A_{q+1-j}\right] =$$

$$= \frac{1}{A_0}\left[C_q(1 - A_1) - \sum_{j=1}^{q-1} C_j A_{q+1-j}\right] - C_1.$$

To finish the method we have to show the way of calculating p_0^*. This probability can be obtained from the equation of equilibrium of the requests. On the average λT requests enter the system during time T and on the average the same number has to be served. The fraction of time of activating the server is: $1 - \sum_{k=0}^{q-1} p_k^* = 1 - q\,{}_0^*$. Considering the average service time to be τ, for the

equilibrium equation we get:

$$\lambda T = \frac{T(1 - qp_0^*)}{\tau}$$

which implies that:

$$p_0^* = \frac{1 - \lambda\tau}{q}.$$

The total probability of having k requests ($k = 1,2, \ldots q-1$) in the system can be determined by the expression:

$$P_k = p_0^* + p_k. \tag{14}$$

The method of calculating the constant coefficients $\{A_j\}$ can be found in [2], for example. To get the time descriptors of the functioning of the queueing system we must obtain the generating function of the stationary probabilities. Let us introduce the intensity of service:

$$r(\tau) = \frac{B'(\tau)}{\mathcal{B}(\tau)}.$$

Now the set of differential equations (2) describing the density of stationary probability can be transcribed to the form:

$$\begin{aligned} &\frac{d\varrho_1(\tau)}{d\tau} + [\lambda + r(\tau)]\,\varrho_1(\tau) = 0, \\ &\frac{d\varrho_k(\tau)}{d\tau} + [\lambda + r(\tau)]\,\varrho_k(\tau) = \lambda\varrho_{k-1}(\tau), \quad k = 2, 3, \ldots \end{aligned} \tag{15}$$

Let us multiply the equation for $\varrho_k(\tau)$ by u^k and sum for all k. We get

$$\sum_{k-1}^{\infty} u^k\,\varrho_k'(\tau) + [\lambda + r(\tau)] \sum_{k-1}^{\infty} u^k \varrho_k(\tau) = \lambda u \sum_{k-1}^{\infty} u^k \varrho_k(\tau). \tag{16}$$

However $\mathcal{G}(u, \tau) = \sum_{k=1}^{\infty} u^k \varrho_k(\tau)$ is, by definition the generating function. Hence Eq. (16) is equivalent with the differential equation:

$$\frac{\partial \mathcal{G}(u, \tau)}{\partial \tau} = [\lambda(u - 1) - r(\tau)]\,\mathcal{G}(u, \tau),$$

and therefore:

$$\mathcal{G}(u, \tau) = \mathcal{G}(u, 0) \exp\left[\lambda(u-1)\,\tau - \int_0^\tau r(\tau)\,d\tau\right].$$

As $\exp\left(-\int_0^\tau r(\tau)\,d\tau\right) = \exp[\ln \mathcal{B}(\tau)] = \mathcal{B}(\tau)$, the last equality leads to:

$$\mathcal{G}(u, \tau) = \mathcal{G}(u, 0)\,\mathcal{B}(\tau(\exp[\lambda\tau(u-1)]. \tag{17}$$

Let us determine the initial value $\mathcal{S}(u, 0)$. By the aid of (3) and the first equation of (2) we find:

$$\lambda p_0^* = \int_0^\infty \varrho_1(\tau)\, r(\tau)\, d\tau,$$

$$\varrho_k(0) = \int_0^\infty \varrho_{k+1}(\tau)\, r(\tau)\, d\tau, \quad k = 1,2, \ldots, \quad k \neq q,$$

$$\varrho_q(0) = \int_0^\infty \varrho_{q+1}(\tau)\, r(\tau)\, d\tau + \lambda p_{q-1}^* .$$

Applying the same transformation to this system as to (15) and considering the equations for p_0^* and p_{q-1}^* proved above we get:

$$u\mathcal{S}(u, 0) = \lambda u(u^q - 1)\, p_0^* + \int_0^\infty \mathcal{S}(u, \tau)\, r(\tau)\, d\tau.$$

Substituting the right-hand side of (17) into the integral, we find:

$$\int_0^\infty \mathcal{S}(u, \tau)\, r(\tau)\, d\tau = \mathcal{S}(u, 0)\, \beta(-\lambda u + \lambda),$$

where $\beta(s) = \int_0^\infty e^{-s\tau}\, dB(\tau)$ is the Laplace transform of the waiting time distribution $B(\tau)$. Thus $u\mathcal{S}(u, 0) = \lambda u(u^q - 1)\, p_0^* + \mathcal{S}(u, 0)\, \beta(\lambda - \lambda u)$, hence

$$\mathcal{S}(u, 0) = \frac{\lambda u(u^q - 1) p_0^*}{u - \beta(\lambda - \lambda u)}. \tag{18}$$

Using (17) and (18) the generating function of the stationary probabilities will be:

$$Q(u) = \sum_{k=0}^{q-1} p_k^* u^k + \int_0^\infty \mathcal{S}(u, \tau) d\tau = \frac{1 - \lambda\tau}{q} = \frac{u^q - 1}{\dfrac{u}{\beta(\lambda - \lambda u)} - 1}. \tag{19}$$

When $q = 1$ this formula coincides with the formula deduced in (3) for the usual single line system. With its help we can find the moments of the distribution of the number of requests. It is known that the kth factorial moment of the discrete distribution equals:

$$\nu_{[k]} = \frac{d^k}{du^k} Q(u)\big|_{u=1}, \tag{20}$$

and between the first two facforial and initial moments the following relations exists:

$$\nu_1 = \nu_{[1]},\ \nu_2 = \nu_{[2]} + \nu_{[1]}. \tag{21}$$

The direct differentiation of the right side of (19) and the ensuing substitution of $u = 1$ results in some difficulties and hardly avoidable uncertainties. Therefore let us expand (19) into a power series. We introduce the new variable $z = 1-u$. Thus, when $u \to 1$ $z \to 0$ (19) changes to the form:

$$Q(1-z) = p_0^* \frac{[(1-z)^q - 1]\,\beta(\lambda z)}{(1-z) - \beta(\lambda z)}.$$

Then we must expand the nominator and the denominator into a power series keeping only those members whose exponent in z is not higher than three (when looking for kth order moments, not higher than $k+1$) simplify the nominator and denominator by z, reverse the denominator and multiply the result with the numerator. Omitting these calculations, we present the final result:

$$Q(1-z) = p_0^* \frac{q}{\lambda b_1 - 1}\left\{-1 + \left[\frac{q-1}{2} + \lambda b_1 + \frac{\lambda^2 b_2}{2!\,(\lambda b_1 - 1)}\right] z + \right.$$
$$+ \left[-\frac{(q-1)(q-2)}{3!} - \frac{\lambda b_1}{2!}(q-1) - \frac{\lambda^2 b_2}{2!} + \frac{\lambda^3 b_3}{3!\,(\lambda b_1 - 1)} - \right.$$
$$\left.\left. - \frac{\lambda^4 b_2^2}{4(b_1-1)^2} + \frac{\lambda^2 b_2}{2!\,(\lambda b_1 - 1)}\left(\frac{q-1}{2!} + \lambda b_1\right)\right] z^2\right\} + o(z^2).$$

In this formula, and further on $\{b_k\}$ denotes the initial moments of the distribution of the service time $B(\tau)$. Considering (20) with the above executed exchange of variables, the expression for p_0^* and the equation $b_1 = \bar{\tau}$, we find the first two moments of the distribution of the number of requests in the system:

$$\nu_{[1]} = \frac{q-1}{2} + \lambda b_1 + \frac{\lambda^2 b_2}{2!\,(1 - \lambda b_1)},$$
$$\nu_{[2]} = (q-1)\left(\frac{q-2}{3} + \lambda b_1\right) + \frac{\lambda^2}{1 - \lambda b_1}\left[\frac{(q+1)b_2}{2} + \frac{(\lambda b_2)^2}{2(1-\lambda b_1)} + \frac{\lambda b_3}{3}\right]. \tag{22}$$

With these results we can find the distribution of waiting time before beginning the service of requests. If a request finds the system in state $E_k(\tau)$, it will be accepted only after having finished the service begun and after having completely served the $k-1$ requests which entered the system before the given request ($k = 1,2, \ldots$). But if the request found the system in the state E_k^*, i.e. when the service is switched off, then we have to wait for the arrival of

$(q-k^*-1)$ other requests, and the serve k^* $(k^* = 0,1, \ldots, q-1)$ requests. Thus for the Laplace–Stieltes transform of the stationary distribution of the waiting time we have:

$$\omega(s) = \int_0^\infty e^{-st} dW(t) =$$

$$= \sum_{k=0}^{q-1} p_k^* \left(\frac{\lambda}{\lambda+s}\right)^{q-k-1} \beta^k(s) + \int_0^\infty d\tau \sum_{k=1}^\infty \varrho_k(\tau)\, \beta^{k-1}(s) \int_0^\infty e^{-sv} \frac{B'(\tau+v)}{\mathfrak{B}(\tau)}\, dv.$$

Here $\dfrac{\lambda}{\lambda+s}$ is the Laplace transform of the interval distribution between the requests of the simplest process. The first sum in (23) is equal to:

$$\sum_{k=0}^{q-1} p_k^* \left(\frac{\lambda}{\lambda+s}\right)^{q-k+1} \beta^k(s) = p_0^* \left(\frac{\lambda}{\lambda+s}\right) \sum_{k=0}^{q-1} \left[\frac{(\lambda+s)\,\beta(s)}{\lambda}\right]^k =$$

$$= \frac{\beta^q(s) - \left(\dfrac{\lambda}{\lambda+s}\right)^q}{\beta(s) - \dfrac{\lambda}{\lambda+s}}\, p_0^*. \tag{24}$$

The second can be transcribed to the form $\dfrac{R[\beta(s)]}{\beta(s)}$ where:

$$R(u) = \int_0^\infty \frac{d\tau}{\mathfrak{B}(\tau)} \sum_{k=1}^\infty u^k\, \varrho_k(\tau) \int_0^\infty e^{-sv}\, B'(\tau+v)\, dv\,.$$

But $\sum_{k=1}^\infty u_k\, \varrho_k(\tau)$ is the generating function $\mathfrak{F}(u, \tau)$. Substituting the latter expression from (17), we find:

$$R(u) = \mathfrak{F}(u, 0) \int_0^\infty e^{-\lambda\tau(u-1)} \left[\int_0^\infty e^{-sv}\, B'(\tau+v)\, dv\right] d\tau =$$

$$= \mathfrak{F}(u, 0)\, \frac{\beta(s) - \beta(\lambda - \lambda u)}{\lambda - \lambda u - s}\,.$$

Substituting the above expression for $\mathfrak{F}(u, 0)$, we get:

$$R(u) = \frac{\lambda u(u^q - 1)\, p_0^*}{u - \beta(\lambda - \lambda u)} \cdot \frac{\beta(s) - \beta(\lambda - \lambda u)}{\lambda - \lambda u - s}\,.$$

Hence

$$\frac{R[\beta(s)]}{\beta(s)} = \frac{\lambda[\beta^q(s) - 1]}{\lambda - \lambda\beta(s) - s}\,.$$

Thus the Laplace transform of the distribution of the waiting time $W(t)$ will be:

$$\omega(s) = \frac{1-\lambda\tau}{q}\left\{\frac{\beta^q(s) - \left(\frac{\lambda}{\lambda+s}\right)^q}{\beta(s) - \frac{\lambda}{\lambda+s}} + \frac{\lambda[\beta^q(s)-1]}{\lambda - \lambda\beta(s) - s}\right\}. \tag{25}$$

At $q = 1$, (25) implies the well-known Polyatsek–Hintsin formula.

Let us find the moments of $W(t)$. On the basis of (23)

$$\omega(s) = 1 - \frac{w_1}{1!}s + \frac{w_2}{2!}s^2 - \frac{w_3}{3!}s^3 + \dots$$

where $\{w_k\}$ are essentially the initial moments of the distribution of $W(t)$. Thus w_k is the coefficient multiplied by $(-1)^k k!$ beside s^k in the power series of the right side of (25) according to the powers of s. Let us carry out this expansion, taking into account the results for w_1 and w_2. Substituting

$$\left.\begin{aligned}
A &= \frac{1}{\lambda} - b_1,\\
B &= \frac{b_2}{2!} + \frac{q-1}{2!}b_1 - \frac{q+1}{2!\,\lambda^2},\\
C &= \frac{(q+1)(q+2)}{3!\,\lambda^3} - \frac{b_3}{3!} - \frac{q-1}{2!}b_1 b_2 - \frac{(q-1)(q-2)}{3!}b_1^3.\\
D &= \frac{\frac{b_2}{2!} - \frac{1}{\lambda^2}}{\frac{1}{\lambda} - b_1},\\
E &= \frac{\frac{1}{\lambda^3} - \frac{b_3}{3!}}{\frac{1}{\lambda} - b_1}.
\end{aligned}\right\} \tag{26}$$

With these notations:

$$\frac{\beta(s) - \left(\frac{\lambda}{\lambda+s}\right)^q}{\beta(s) - \frac{\lambda}{\lambda+s}} = \frac{q^\lambda}{1-\lambda b_1}\cdot\frac{A + Bs + Cs^2 + \dots}{1 + Ds + Es^2 + \dots} = \tag{27}$$

$$= \frac{q^\lambda}{1-\lambda b_1}\{A + (B - AD)s + [C - BD + A(D^2 - E)]s^2 + \dots\}.$$

In an analogous way we produce the expansion:

$$\frac{\beta^q(s)-1}{1-\beta(s)-\frac{s}{\lambda}}=\frac{q^\lambda}{1-\lambda b_1}\{A_1-(B_1+A_1D_1)\,s+[C_1+B_1D_1+A_1(D_1^2+E)]\,s^2+\ldots\}, \tag{28}$$

where

$$\left.\begin{aligned}
A_1 &= b_1,\\
B_1 &= \frac{b_2}{2!}+\frac{q-1}{2!}\,b_1^2,\\
C_1 &= \frac{b_3}{3!}+\frac{q-1}{2!}\,b_1b_2+\frac{(q-1)(q-2)}{3!}\,b_1^3,\\
D_1 &= \frac{\frac{b_2}{2!}}{\frac{1}{\lambda}-b_1},\\
E_1 &= \frac{b_3}{3!\left(\frac{1}{\lambda}-b_1\right)}.
\end{aligned}\right\} \tag{29}$$

Substituting (27) and (28) into (25) we get sought expansion of $\omega(s)$ in terms of the powers of s, fromwhere we deduce the initial moments of the distribution of waiting time for the beginning of service, which we are looking for:

$$w_1=\lambda(B_1+A_1D_1-B+AD), \tag{30}$$

$$w_2=2\lambda[C-BD+A(D^2-E)+C_1+B_1D_1+A_1(D^2+E)]. \tag{31}$$

From formula (30), taking (26) and (29) into consideration, we can get:

$$w_1=\frac{q-1}{\lambda}+\frac{\lambda b_2}{2(1-\lambda b_1)},$$

which, in case of $q=1$, gives the known result for the traditional system $M/G/1$ ([3], p. 80).

Thus we obtained the major characteristics of the progress of requests in the system. On the basis of them we can solve the problem of the optimal value q.

Let, for example, the complex expression for the evaluation of the system have the form:

$$L(q)=g\cdot n(q)+h\cdot w_1(q),$$

where $n(q)$ is the number of activations of the system per unit time, g and h are the corresponding weight coefficients. It can be seen,

$$n(q) = \lambda p^*_{q-1} = \frac{\lambda}{q}(1 - \lambda b_1)$$

thus

$$L(q) = g\frac{\lambda}{q}(1 - \lambda b_1) + h\left[\frac{q-1}{\lambda} + \frac{\lambda b_2}{2(1 - \lambda b_1)}\right]. \tag{10}$$

Differentiating the result according to q, and equation the derivative with zero, we obtain the optimal activation threshold:

$$\tilde{q} = \lambda\sqrt{\frac{g}{h}(1 - \lambda b_1)}. \tag{32}$$

In practice between the integers $[\tilde{q}]$ and $[\tilde{q}] + 1$ we choose the one, which gives a lower value to the right-hand side of (32). It is understood that the choice of q can be based on other probability characteristics of the system.

In conclusion we once again underline the importance of the examined class of queueing systems. The "threshold" activation when $q > 1$ results in less frequent changes in the work phases of the system, increasing its reliability and decreasing the load of the operator in partially automated complexes, providing time for the latter to carry out other tasks.

Thus the application of the threshold systems of queueing theory (and their method of calculation developed in this paper) have to be considered very perspective.

References

1. *Рыжиков Ю. И.* Многономенклатурная задача о восстанавливаемом ЗИП. Известия АН СССР, «Техническая кибернетика», *6* (1969) 65—75.
2. *Рыжиков Ю. И.* Управление запасами. «Наука», М., 1969.
3. *Риордан Дж.* Вероятностные системы обслужисвания. «Связь», 1966.

NOTE TO CONTRIBUTORS

Two copies of the manuscripts (each duly completed by figures, tables and references) are to be sent either to

E. D. Teryaev coordinating editor

Department of Mechanics and Control Processes

Academy of Sciences of the USSR

Leninsky Prospect 14, Moscow V-71, USSR

or to

J. Kocsis coordinating editor

Department of Automation

Technical University

H–1111 Budapest XI., Goldman György tér 3, Hungary

The authors are requested to retain a third copy of the submitted typescript to be able to check the proofs against it

The papers, preferably in English or Russian, should be typed double-spaced on one side of good-quality paper, with wide margins (c. 4–5 cm) should carry the title of the contribution, the author(s)' name, and the name of the country. At the end of the typescript the name of that author who manages the proof-reading should also be given.

An abstract of about 50 to 100 words should head the paper.

The authors are encouraged to use the following headings: Introduction, (outlining the problem), Methods and results, Discussion, Conclusions, References. The entire material should not exceed 15 pages including tables and references. The proper location of the tables and figures must be indicated on the margins. Mathematical notations should follow up-to-date usage.

The summary – possibly in Russian if the paper is in English and *vice versa* – should contain a brief account of the proposition and indications of the formulas used and figures shown in the paper. The summary is not supposed to exceed 10–15 per cent of the paper.

The authors will be sent sheet-proofs which they are to return by next mail to the sender Regional Editorial Board.

Authors are entitled to 100 reprints free of charge. Rejected manuscripts will be returned to the authors.

К СВЕДЕНИЮ АВТОРОВ

Рукописи в двух экземплярах (каждый из которы должен содержать рисунки, таблицы и литературу направляются

Е. Д. Теряеву — Научный секретарь журнала

Отделение механики и процессов управления

Академия Наук СССР

Ленинский Проспект, 14, Москва В-71, СССР

или

Я. Кочишу — Научный секретарь журнала

Кафедра Автоматизации
Будапештского Технического
Университета, Будапешт XI, площадь
H–1111 Голдман Д., 3, Венгрия

Авторам рекомендуется оставлять у себя копию всех представленных ими материалов для справок при корректуре.

Статьи, желательно на русском или английском языках, отпечатанные на бумаге хорошего качества, с промежутком в два интервала и широкими (4–5 см) полями должны содержать наименование статьи, фамилию автора (авторов), название страны. В конце статьи необходимо также указать фамилию автора, ответственного за корректуру гранок.

Статье должна предшествовать аннотация объемом до 50–100 слов.

Авторы при написании статьи должны придерживаться следующей формы: введение (постановка задачи), основное содержание и результаты, обсуждение, выводы и литература. Объем статьи не должен превышать 15 печатных страниц, включая таблицы и ссылки. Последовательность таблиц и рисунков должна быть отмечена на полях. Математические обозначения рекомендуется давать в соответствии с современными требованиями и традициями.

К статье обязательно должно быть приложено резюме-реферат. Резюме — на русском языке, если статья написана на английском, и наоборот — должно содержать краткое изложение текста статьи со ссылками на необходимые формулы и графики, имеющиеся в основном тексте. Объем разюме не должен превышать 10–15% объема статьи.

Авторам высылаются гранки статьи, которые они должны незамедлительно возвратить в Региональную секцию Редколлегии журнала.

Авторам обеспечивается бесплатно 100 оттисков их статей. Рукописи непринятых статей возвращаются авторам.

CONTENTS ● СОДЕРЖАНИЕ

Сифоров В. И.: Развитие науки в СССР (к 250-летию Академии наук СССР) 3

Siforov V. I.: The development of science in the Soviet Union (The 250th anniversary of the foundation of the Academy of Sciences of the USSR) 7

Györfi, L.: On some estimation problems in pattern recognition (*Дерфи Л.:* Некоторые проблемы оценок в теории распознавания образов) 11

Nemetz, T.: Finite data reduction in the case of a long sequence of observations (*Немец Т.:* Коррекция конечного числа данных в случае продолжительной последовательности наблюдений) 19

Khazen, E. M.: Risk increase informational estimates under observed data reduction in sequential analysis (*Хазен Е. М.:* Информационные оценки для возрастания риска при редукции наблюдаемых данных в задачах последовательного анализа) 35

Belavkin, V. P.: Optimal linear randomized filtration of quantum boson signals (*Белавкин В. П.:* Оптимальная линейная рандомизированная фильтрация квантовых бозонных сигналов) 47

Колмановский В. Б.: Об аппроксимации линейных управляемых систем с последействием (*Kolmanovsky, V. B.:* On the approximation of linear control systems with time lag) 63

Tolmacheva, A. Yu.: Choice of base elements in structural synthesis algorithms (*Толмачева А. Ю.:* Выбор базовых элементов в алгоритмах структурного синтеза) 77

Рыжиков Ю. И.: Расчет системы массового обслуживания с порогом включения (*Rizhikov, Yu. I.:* Analysis of mass service system with switching on threshold) 85

6th Congress of the International Measurement Confederation 99

Book review 103

316.920

VOL. / TOM 3 * NUMBER / НОМЕР 2

ACADEMY OF SCIENCES OF THE USSR
HUNGARIAN ACADEMY OF SCIENCES

PROBLEMS OF CONTROL AND INFORMATION THEORY

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

АКАДЕМИЯ НАУК СССР
АКАДЕМИЯ НАУК ВЕНГРИИ

1974

AKADÉMIAI KIADÓ, BUDAPEST

PROBLEMS OF CONTROL AND INFORMATION THEORY

is an international quarterly sponsored jointly by the Presidium of the Academy of Sciences of the USSR and of the Hungarian Academy of Sciences. It offers publicity for original papers and short communications of the following topics:

— general theory of control systems and system theory

— theory of automata

— information theory

— operation research; theory of complex systems

— theory of economic control; system modelling

— theory and methods of adaptation, learning, identification and pattern recognition

— methods of information processing; application of digital computers in control and communication systems

— new physical principles in constructing technical devices for automation, control and information processing

The four issues published per year make up a volume of some 320 pp.

While this quarterly is mainly a publication forum of the research results achieved within the USSR and Hungary, also papers of international interest from other countries are welcome.

Distributor:

KULTURA

Hungarian Trading Co. for Books and Newspapers

H-1389 Budapest 62, P.O. Box 149, Hungary

The quarterly is published by

AKADÉMIAI KIADÓ

Publishing House of the Hungarian Academy of Sciences

H-1363 Budapest, P.O. Box 24, Hungary

Президиумами Академии Наук СССР и Академии Наук ВНР было принято решение о совместном издании журнала

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

Журнал создан для обеспечения быстрейшей публикации материалов о новейших результатах научных исследований, разработок и кратких сообщений о достижениях в следующих областях науки:

— общая теория процессов управления и теория систем
— теория автоматов
— теория информации
— исследование операций; теория сложных систем
— теория управления экономическими системами; модели систем
— теория и методы адаптации, обучения, идентификации и распознавания образов
— методы обработки информации; применение вычислительных машин в системах управления и передачи информации
— новые физические принципы создания технических средств автоматики, управления и передачи информации

Журнал издается четыре раза в год общим объёмом около 320 печатных страниц.

Журнал в первую очередь будет публиковать научные достижения обеих стран, а также статьи из других стран, представляющие международный интерес.

Распространитель:

KULTURA

Венгерское Общество по распространению книг и журналов

H-1389 Будапешт 62, п. о. 149, Венгрия

Издатель журнала:

AKADÉMIAI KIADÓ

Издательство Академии Наук Венгрии

H-1363 Будапешт, п. о. 24, Венгрия

PROBLEMS OF CONTROL AND INFORMATION THEORY

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

EDITORS

B. N. PETROV (Moscow)
F. CSÁKI (Budapest)

DEPUTY EDITORS

V. S. PUGACHEV (Moscow)
V. I. SIFOROV (Moscow)
S. CSIBI (Budapest)

COORDINATING EDITORS

S. V. EMELIANOV (Moscow)
L. KALMÁR (Szeged)

M. A. GAVRILOV (Moscow)
I. CSISZÁR (Budapest)

A. M. LETOV (Moscow)
A. PRÉKOPA (Budapest)

L. VARGA (Budapest)

E. D. TERYAEV (Moscow)
J. KOCSIS (Budapest)

РЕДАКТОРЫ ЖУРНАЛА

Б. Н. ПЕТРОВ (Москва)
Ф. ЧАКИ (Будапешт)

ЗАМЕСТИТЕЛИ РЕДАКТОРОВ

В. С. ПУГАЧЕВ (Москва)
В. И. СИФОРОВ (Москва)
Ш. ЧИБИ (Будапешт)

ЧЛЕНЫ РЕДАКЦИОННОЙ КОЛЛЕГИИ

С. В. ЕМЕЛЬЯНОВ (Москва)
Л. КАЛМАР (Сегед)

М. А. ГАВРИЛОВ (Москва)
И. ЧИСАР (Будапешт)

А. М. ЛЕТОВ (Москва)
А. ПРЕКОПА (Будапешт)

Л. ВАРГА (Будапешт)

Е. Д. ТЕРЯЕВ (Москва)
Я. КОЧИШ (Будапешт)

AKADÉMIAI KIADÓ

PUBLISHING HOUSE OF THE HUNGARIAN ACADEMY OF SCIENCES

BUDAPEST

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

Problems of Control and Information Theory, Vol. 3 (2), pp. 105–115 (1974)

ПРИМЕНЕНИЕ РЯДОВ ВОЛЬТЕРРА ДЛЯ ИССЛЕДОВАНИЯ НЕЛИНЕЙНЫХ СИСТЕМ

К. А. ПУПКОВ, В. И. КАПАЛИН

(Москва)

(Поступила в редакцию 30 апреля 1972 г.)

В статье рассматриваются вопросы применения рядов Вольтерра и теории многомерных интегральных преобразований в задачах анализа и синтеза нелинейных систем. Приводится алгоритм определения изображений ядер ряда Вольтерра по заданной структурной схеме. Рассматриваются методы определения выходного сигнала системы и решается задача нелинейной компенсации.

Введение

В работе рассматриваются методы анализа и синтеза нелинейных систем, основанные на применении рядов Вольтерра и теории многомерных интегральных преобразований. Эти методы позволяют получить аналитическое решение нелинейной задачи. Предлагается алгоритм определения изображений ядер ряда Вольтерра для систем, заданных структурной схемой. Приводятся методы определения реакции нелинейной системы на регулярные воздействия различных типов. В заключение рассматривается задача компенсации аналитических систем в предположении, что желаемая система является линейной. Для этого случая предлагается более простая методика определения ядер компенсатора по сравнению с общей методикой, разработанной в [1].

1. Ряды Вольтерра

Функционал $y(t) = F[x(\tau), \tau \leq t]$, описывающий связь между входным воздействием $x(t)$ на стационарную нелинейную систему и выходной реакцией системы и удовлетворяющий условию $F[0] \equiv 0$ для всех t, называется аналитическим, если его можно представить в виде ряда Вольтерра с ненулевым радиусом сходимости [1]

$$y(t) = F[x(\tau), \tau \leq t] = \sum_{n=1}^{\infty} \nu_n[x(\tau), \tau \leq t], \tag{1.1}$$

где $\nu_n[x(\tau), \tau \leq t]$, $n = 1, 2, \ldots$ — регулярный однородный функционал Вольтерра степени n

$$y_n(t) = \nu_n[x(\tau), \tau \leq t] = \int_{E_n} h_n(\tau_1, \ldots, \tau_n) \prod_{r=1}^{n} x(t - \tau_n)\, d\nu_\tau. \tag{1.2}$$

Здесь ядро $h_n(\tau_1, \ldots, \tau_n)$ обращается в нуль, если один из аргументов τ_j, $j = 1, 2, \ldots, n$ меньше нуля; E_n — область, определяемая так:

$$-\infty < \tau_j < \infty, \; j = 1, 2, \ldots, n; \; d\nu_\tau = d\tau_1 \ldots d\tau_n .$$

В работах [2], [3], [4], [5] были разработаны методы определения ядер ряда (1.1) для систем, описываемых нелинейными дифференциальными уравнениями. В данной работе рассматривается случай, когда нелинейная система, образованная соединением стационарных линейных инерционных звеньев истационарных безынерционных нелинейностей, задается структурной схемой. Для того, чтобы ряды Вольтерра можно было бы использовать при расчете такой системы, необходимо, чтобы все безынерционные нелинейности описывались либо полиномами, либо степенными рядами. Интеграл свертки, определяющий линейную систему, является регулярным однородным функционалом Вольтерра первой степени. Степенной ряд

$$y = \sum_{n=1}^{\infty} a_n x^n \tag{1.3}$$

тоже может рассматриваться как вырожденный вид ряда Вольтерра (1.1) с ядрами $h_n(\tau_1, \ldots, \tau_n) = a_n \prod_{r=1}^{n} \delta(\tau_r)$, $n = 1, 2, \ldots$, где $\delta(\tau_r)$ — δ-функция. Таким образом, задача определения ядер ряда Вольтерра для нелинейной системы, заданной структурной схемой, может быть сформулирована в общем виде так.

Дана система, состоящая из ряда подсистем, описания которых в виде рядов Вольтерра известны. Требуется определить ядра ряда Вольтерра, описывающего всю систему. Для нелинейных стационарных систем сформулированная задача проще всего решается с помощью многомерного преобразования Лапласа.

2. Многомерное преобразование Лапласа

Преобразованием Лапласа функции $f(t_1, \ldots, t_n)$ от n аргументов называется функция

$$F(s_1, \ldots, s_n) = \int_{E_n^+} \exp\left(-\sum_{i=1}^{n} s_i t_i\right) f(t_1, \ldots, t_n)\, d\nu_t , \tag{2.1}$$

где E_n^+ — область, определяемая так: $0 \leq t_i < \infty$, $i = 1, 2, \ldots, n$; s_i, $i = 1, 2, \ldots, n$ — комплексная переменная; $d\nu_t = dt_1 \ldots dt_n$. Если $f(t_1, \ldots, t_n) =$

$= 0$ для $t_i < 0$, $i = 1, 2, \ldots, n$, то обратное преобразование Лапласа существует и его можно получить по формуле

$$f(t_1, \ldots, t_n) = \left(\frac{1}{2\pi j}\right)^n \int\limits_{Br_n} F(s_1, \ldots, s_n) \exp\left(\sum_{i=1}^{n} s_i t_i\right) dv_s, \tag{2.2}$$

где Br_n — область, определяемая так: $c_i - j\infty < s_i < c_i + j\infty$, $i = 1, 2, \ldots, n$ и $dv_s = ds_1 \ldots ds_n$.

Соотношения между оригиналом $f(t_1, \ldots, t_n)$ и его изображением $F(s_1, \ldots, s_n)$ сокращенно будем записывать так:

$$f(t_1, \ldots, t_n) \Leftrightarrow F(s_1, \ldots, s_n).$$

Если $F_1(s_1, \ldots, s_n) \Leftrightarrow f_1(t_1, \ldots, t_n)$ и $F_2(s_1, \ldots, s_n) \Leftrightarrow f_2(t_1, \ldots, t_n)$, то справедлива формула многомерной свертки

$$F_1(s_1, \ldots, s_n)\, F_2(s_1, \ldots, s_n) \Leftrightarrow \int\limits_{E_n} f_1(\tau_1, \ldots, \tau_n)\, f_2(t_1 - \tau_1 \ldots t_n - \tau_n)\, dv_\tau .$$

Для того, чтобы формулу свертки можно было использовать в сочетании с рядами Вольтерра, сопоставим интегралу (1.2) выражение

$$y_n(t_1, \ldots, t_n) = \int\limits_{E_n} h_n(\tau_1, \ldots, \tau_n) \prod_{r=1}^{n} x(t_r - \tau_r)\, dv_\tau . \tag{2.3}$$

Применяя к выражению (2.3) формулу свертки, получим

$$Y_n(s_1, \ldots, s_n) = H_n(s_1, \ldots, s_n) \prod_{r=1}^{n} X(s_r) .$$

Чтобы определить отсюда $y_n(t)$, нужно взять обратное n-мерное преобразование, получить $y_n(t_1, \ldots, t_n)$ и положить $t_1 = \ldots = t_n = t$, поскольку $y_n(t_1, \ldots, t_n)\,|_{t_1 = \ldots = t_n = t} = y_n(t)$. Удобнее, однако, не определять n-мерный оригинал, а сразу перейти к одномерному изображению в комплексной области следующим образом. Положив в формуле (2.2) $t_i = t$, $i = 1, 2, \ldots, n$, получим

$$f(t_1, \ldots, t_n)\,\Big|_{t_1 = \ldots = t_n = t} = \left(\frac{1}{2\pi j}\right)^{n-1} \int\limits_{Br_{n-1}} \exp\left(t \sum_{i=1}^{n-2} s_i + ts\right) dv_s\, ds \cdot$$

$$\cdot \frac{1}{2\pi j} \int\limits_{c_{n-1} - j\infty}^{c_{n-1} + j\infty} F(s_1, \ldots, s_{n-1}, s - s_{n-1})\, ds_{n-1} . \tag{2.4}$$

Внутренний интеграл в формуле (2.4) берется с помощью теоремы о вычетах.

Повторяя этот процесс $n-1$ раз, переходим к $F(s)$ и по нему находим $f(t)$ посредством одномерного преобразования Лапласа. В дальнейшем оператор перехода к одной переменной в комплексной области будем обозначать знаком $\{\ \}^*$, т. е. $F(s) = \{F(s_1, \ldots, s_n\}^*$.

Рассмотрим некоторые следствия формулы (2.4).

а) Если $F(s_1, \ldots, s_n) = F(s_1 + \ldots + s_n)$, то для перехода к одной переменной следует положить $s = s_1 + \ldots + s_n$.

б) Если изображение имеет вид $F(s_1, \ldots, s_n) = F_1(s_1 + \ldots + s_n) F_2(s_1, \ldots, s_n)$, то применять формулу (2.4) следует только к $F_2(s_1, \ldots, s_n)$. Большинство n-мерных изображений, встречающихся на практике, имеет вид

$$F(s_1, \ldots, s_n) = P(s_1 + \ldots + s_n) \prod_{r=1}^{n} F(s_n).$$

Чтобы определить одномерный оригинал этого выражения можно поступить следующим образом. Найдем одномерные оригиналы от $F_r(s_r) - f_r(t)$, $r = 1, 2, \ldots, n$ и, перемножив их, получим

$$f_2(t) = \prod_{r=1}^{n} f_r(t).$$

Далее определим одномерное изображение от $f_2(t) - F_2(s)$. Умножим $F_2(s)$ на $P(s_1 + \ldots + s_n)$, в котором положим $s_1 + \ldots + s_n = s$. В результате получим требуемое изображение $F(s) = P(s) F_2(s)$.

в) Если $f(t) \Leftrightarrow \{F(s_1, \ldots, s_n)\}^*$, то

$$\{F(s_1 + a_1, \ldots, s_n + a_n)\}^* \Leftrightarrow f(t) \exp(-tA),$$

где $A = \sum_{i=1}^{n} a_i$.

г) Если $f(t) \Leftrightarrow \{F(s_1, \ldots s_n)\}^*$, то для любого $\tau > 0$

$$\left\{F(s_1, \ldots, s_n) \exp\left(-\tau \sum_{i=1}^{n} s_i\right)\right\}^* \Leftrightarrow f(t-\tau).$$

3. Определение изображений ядер ряда Вольтерра для систем, заданных структурной схемой

Используя теорию многомерного преобразования Лапласа, построим теперь методику определения изображений ряда Вольтерра для систем, заданных структурной схемой. В общем случае исследуемая система будет

содержать три типа соединения подсистем: параллельное, последовательное и обратное. Поэтому методика определения изображений ядер ряда Вольтерра основывается на формулах для этих трех типов соединений и является естественным развитием результатов работ [6] и [7].

Если две аналитические системы соединены параллельно, то, ввиду того, что многомерное преобразование Лапласа линейно, изображения ядер ряда одинакового порядка складываются.

Пусть теперь две аналитические системы соединены последовательно. Если первая из них — входная — имеет вид $z(t) = \sum_{j=1}^{\infty} \nu_i [x(\tau), \tau \leq t]$ с ядрами $h^j(\tau_1, \ldots, \tau_j)$, $j = 1, 2, \ldots$, а вторая — выходная — $y(t) = \sum_{i=1}^{\infty} \nu_i^*[z(\tau), \tau \leq t]$ с ядрами $k_i(\tau_1, \ldots, \tau_i)$, $i = 1, 2, \ldots$, то их последовательное соединение описывается рядом

$$y(t) = \sum_{i=1}^{\infty} \nu_i^* \left[\sum_{j=1}^{\infty} \nu_j [x(\tau)], \tau \leq t \right]. \tag{3.1}$$

Общий член этого ряда имеет вид

$$\int_{E_{N+2}} h_i(\tau_1, \ldots, \tau_i) k_{l_1}(\sigma - \tau_1, \ldots, \sigma_{l_1} - \tau_1) k_{l_2}(\sigma_{l_1+1} - \tau_2, \ldots, \sigma_{l_1+l_2} - \tau_2) \ldots$$

$$\ldots k_{l_p}(\sigma_{N-l_p+1} - \tau_i, \ldots, \sigma_N - \tau_i) \prod_{r=1}^{N} x(t - \sigma_r) \, d\nu_\sigma d\nu_\tau, \tag{3.2}$$

где $i = 1, 2, \ldots$; $p = 1, 2, \ldots$; $N = \sum_{i=1}^{p} l_i$.

Многомерное изображение от (3.2), как это можно показать, используя формулу свертки, имеет вид:

$$y(s_1, \ldots, s_n) = H_i \left(\sum_{j=1}^{l_1} s_j, \sum_{j=l_1+1}^{l_1+l_2} s_j, \ldots, \sum_{j=N-l_p+1}^{N} s_j \right) \cdot$$

$$\cdot K_{l_1}(s_1, \ldots, s_{l_1}) K_{l_2}(s_{l_1+1}, \ldots, s_{l_1+l_2}) \ldots K_{l_p}(s_{N-l_p+1}, \ldots, s_N) \prod_{j=1}^{N} X(s_j), \tag{3.3}$$

где $H_i(s_1, \ldots, s_i) \Leftrightarrow h_i(\tau_1, \ldots, \tau_i)$ и $K_{l_j}(s_1, \ldots, s_{l_j}) \Leftrightarrow k_{l_j}(\tau_1, \ldots, \tau_{i_j})$. Поясним применение формулы (3.3). Рассмотрим сначала в (3.1) слагаемое при $i = 1$. Положив в (3.2) и (3.3) $i = 1$, получим такие изображения ядер:

$$H_1(s_1) K_1(s_1), \; H_1(s_1 + s_2) k_2(s_1, s_2), \ldots .$$

Далее рассмотрим в (3.1) слагаемое при $i = 2$. Получим из (3.3)

$$H_2(s_1, s_2) K_1(s_1) K(s_2), \; H_2(s_1, s_2 + s_3) K_1(s_1) K(s_2, s_3), \ldots . \tag{3.4}$$

Группируя ядра одинакового порядка и обозначая изображения ядер последовательного соединения через $H_i^*(s_1, \ldots, s_i)$, получим

$$\left.\begin{aligned} &H^*(s_1) = H_1(s_1)\,K_1(s_1)\,; \\ &H_2^*(s_1, s_2) = H_1(s_1 + s_2)\,K_2(s_1, s_2) + H_2(s_1, s_2)\,K_1(s_1)\,K_1(s_2); \ldots \end{aligned}\right\}. \tag{3.5}$$

Две системы, соединенные последовательно, будут взаимно обратными, если ядра их последовательного соединения таковы:

$$H_1^*(s_1) = 1;\ H_i^*(s_1, \ldots s_i) = 0, \qquad i = 2, 3, \ldots .$$

Отсюда, изображения ядер обратной системы

$$\left.\begin{aligned} &K_1(s_1) = \frac{1}{H_1(s_1)}\,; \\ &K_2(s_1, s_2) = \frac{-H_2(s_1, s_2)}{H_1(s_1 + s_2)\,H_1(s_1)\,H_1(s_2)}\,; \ \ldots \end{aligned}\right\}. \tag{3.6}$$

Последовательно применяя полученные в этом разделе формулы, можно определить изображения ядер Вольтерра для аналитических систем, заданных структурной схемой любого вида.

4. Определение реакции нелинейной системы на регулярное входное воздействие

После определения сходимости ряда Вольтерра и изображений ядер можно перейти к определению выходного сигнала нелинейной системы. Пусть входное воздействие является регулярной (т. е. не случайной) функцией времени. Выходную реакцию нелинейной системы можно определить несколькими способами.

Если известно преобразование Лапласа от входного сигнала, то для определения выходного сигнала можно использовать формулу многомерной свертки, формулу (2.4) и следствия из нее, приведенные во втором разделе данной работы. Вычисления при этом сводятся к применению теоремы о вычетах. Для некоторых специальных видов многомерных изображений переход к одной переменной в комплексной области может быть осуществлен по таблицам, приведенным в [8].

Пусть теперь входное воздействие задано рядом Тейлора, построенным в окрестности точки t,

$$x(t - \tau) = \sum_{i=0}^{\infty} \frac{(-1)^i}{i!}\,\frac{d^i x(t)}{dt^i}\,\tau^i\,. \tag{4.1}$$

Установившийся сигнал в системе, используя разложение (4.1), можно записать так:

$$y(t) = \int\limits_{E_1} h_1(\tau)\, x(t-\tau)\, d\tau + \int\limits_{E_2} h_2(\tau_1, \tau_2)\, x(t-\tau_1)\, x(t-\tau_2)\, dv_\tau + \ldots =$$

$$= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \frac{d^i x(t)}{dt^i} \int\limits_{E_1^+} \tau_1^i\, h_1(\tau)\, d\tau + \tag{4.2}$$

$$+ \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^{i+j}}{i!\, j!} \frac{d^i x(t)}{dt^i} \frac{d^j x(t)}{dt^i} \int\limits_{E^+} \tau_1^i\, \tau_2^j\, h_2(\tau_1, \tau_2)\, dv_\tau + \ldots \; .$$

Величины

$$\mu_{ij\ldots l}^n = \int\limits_{E_n^+} \tau_1^i\, \tau_2^j \ldots \tau^l\, h_n(\tau_1, \ldots, \tau_n)\, dv_\tau \tag{4.3}$$

назовем моментами ядер системы. Моменты ядер системы легко можно вычислить, дифференцируя соотношение (2.1), т. е.

$$\frac{\partial^i \partial^j \ldots \partial^l \{H_n(s_1, \ldots, s_n)\}}{\partial s_1^i\, \partial s_2^j \ldots \partial s_n^l} =$$

$$= (-1)^{i+j+\ldots+l} \int\limits_{E_n^+} \tau_1^i\, \tau_2^j \ldots \tau_n^l\, h_n(\tau_1, \ldots, \tau_n) \exp\left(-\sum_{k=1}^{n} s_k\, \tau_k\right) dv_\tau . \tag{4.3}$$

Положим в (4.3) $s_k = 0$, $k = 1, 2, \ldots, n$. Тогда

$$\mu_{ij\ldots l}^n = (-1)^{i+j+\ldots+l} \left. \frac{\partial^i \partial^j \ldots \partial^l \{H_n(s_1, \ldots, s_n)\}}{\partial s_1^i\, \partial s_2^j \ldots \partial s_n^l} \right|_{s_1 = \ldots = s_n = 0} .$$

Пусть теперь входное воздействие задается своим рядом Фурье. Использование обычных рядов Фурье в сочетании с интегралами Вольтерра порядка выше первого неудобно, поскольку выходной сигнал непосредственно в виде ряда Фурье получить нельзя. Если же входной сигнал представлен рядом Фурье в комплексной форме, то удается получить алгебраические формулы, связывающие коэффициенты Фурье реакции системы с изображениями ядер и коэффициентами Фурье входного воздействия. Дадим эти формулы для воздействия частного вида

$$x(t) = c_{-1} \exp(-j\omega t) + c_1 \exp(+j\omega t) . \tag{4.4}$$

Подставим (4.4) в (1.1) и произведем необходимые выкладки. Если выходной сигнал искать в виде

$$y(t) = \sum_{-\infty}^{\infty} d_n \exp(j\omega_n t) , \tag{4.5}$$

то его коэффициенты в предположении, что все ядра ряда (1.1) симметричны, получаются такими

$$\left.\begin{aligned}
d_0(j\omega) &= 2c_1 c_{-1} H_2(j\omega, -j\omega) + \ldots \\
d_{-1}(j\omega) &= c_{-1} H_1(-j\omega) + 3c_{-1}^2 c_1 H_3(j\omega, -j\omega, -j\omega) + \ldots \\
d_1(j\omega) &= c_1 H_1(j\omega) + 3c_1^2 c_{-1} H_3(-j\omega, j\omega, j\omega) + \ldots \\
d_{-2}(j\omega) &= c_{-1}^2 H_2(-j\omega, -j\omega) + \ldots \\
d_2(j\omega) &= c_1^2 H_2(j\omega, j\omega) + \ldots \\
d_3(j\omega) &= c_{-1}^3 H_3(-j\omega, -j\omega, -j\omega) + \ldots \\
d_3(j\omega) &= c_1^3 H_3(j\omega, j\omega, j\omega) + \ldots
\end{aligned}\right\}. \qquad (4.6)$$

Здесь все коэффициенты $H_n(j\omega, \ldots, j\omega)$ получены из изображений многомерных ядер ряда Вольтерра по Фурье $H_n(j\omega_1, \ldots, j\omega_n)$ подстановкой $\omega_1 = \ldots = \omega_n = \omega$. Пусть ряд (1.1) нелинейной системы содержит лишь функционалы Вольтерра нечетных степеней. Тогда коэффициент $d_1(j\omega)$ в (4.6) получится в таком виде:

$$d_1(j\omega) = c_1 H_1(j\omega) + 3c_1^2 c_{-1} H_3(-j\omega, j\omega, j\omega) + \\ + 8c_1^3 c_{-1}^2 H_5(-j\omega, -j\omega, j\omega, j\omega, j\omega).$$

Разделив $d_1(j\omega)$ на c_1, получим эквивалентный частотный коэффициент $d_э(j\omega)$ нелинейной аналитической системы по первой гармонике

$$d_э(j\omega) = H_1(j\omega) + 3c_1 c_{-1} H_3(-j\omega, j\omega, j\omega) + \\ + 8(c_1 c_{-1})^2 H_5(-j\omega, -j\omega, j\omega, j\omega, j\omega). \qquad (4.7)$$

Из (4.7) следует, что в отличие от частотной характеристики линейной системы коэффициент $d_э(j\omega)$ зависит от амплитуды входного воздействия.

5. Компенсация нелинейных систем

Сформулируем задачу нелинейной компенсации. Пусть задана структурная схема синтезируемой системы, параметры неизменяемых элементов и желаемая передаточная функция, полученная каким-либо из известных методов синтеза линейных систем. Требуется определить ядра компенсатора таким образом, чтобы расстояние между желаемым $y_d(t)$ и действительным $y(t)$ выходами было бы малым в пространстве непрерывных функций $c(T)$. При решении указанной задачи предположим, что все линейные подсистемы устойчивые и минимально фазовые.

Предлагаемая методика компенсации основывается на алгоритме определения изображений ядер ряда Вольтерра, приведенном в третьем разделе и не зависит от того, является ли нелинейная система инерционным или безынерционным элементом. Поэтому ограничимся случаем, когда нелинейная система безынерционна. Теорию компенсации удобнее рассмотреть на конкретном примере системы с обратной связью, представленной на рис. 1, где

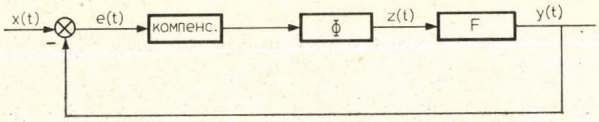


Рис. 1

нелинейный блок F описывается полиномом $y = \sum_{n-1}^{N} a_n z^n$, линейный блок Φ имеет передаточную функцию $H(s)$. Компенсатор будем строить в виде ряда Вольтерра. Обозначим изображения его ядер через $\{K_i(s_1, \ldots, s_i)\}$. Применяя методику определения изображения ядер для соединения компенсатора и линейной системы, можно записать

$$K_1(s)\,H(s);\;\; K_2(s_1, s_2)\,H(s_1 + s_2);\;\; K_3(s_1, s_2, s_3)\,H(s_1 + s_2 + s_3);\;\; \ldots . \qquad (5.1)$$

Определим изображения ядер всей прямой ветви. Они получаются такими

$$a_1 K_1(s)\,H(s);\;\; a_2 K_1(s_1)\,K_1(s_2)\,H(s_1)\,H(s_2) + a_1\,K_2(s_1, s_2)\,H(s_1 + s_2);\;\; \ldots \qquad (5.2)$$

Обозначим желаемую передаточную функцию прямой ветви через $H_d(s)$. Отсюда получим систему уравнений относительно изображений ядер компенсатора

$$\left.\begin{aligned} &a_1\,K_1(s)\,H(s) = H_d(s);\\ &a_2\,K_1(s_1)\,K_1(s_2)\,H(s_1)\,H(s_2) + a_1\,K_2(s_1, s_2)\,H(s_1 + s_2) = 0\,;\;\ldots \end{aligned}\right\} \qquad (5.3)$$

Решая систему (5.3), находим

$$\left.\begin{aligned} &K_1(s) = \frac{H_d(s)}{a_1\,H(s)}\,;\\ &K_2(s_1, s_2) = -\,\frac{a_2\,H_d(s_1)\,H_d(s_2)}{a_3\,H(s_1 + s_2)}\,;\;\ldots \end{aligned}\right\} \qquad (5.4)$$

Аналогично решается задача компенсации, когда компенсатор находится в цепи обратной связи. Вся методика компенсации остается без изменений и в том случае, если компенсируемый нелинейный элемент можно представить рядом Вольтерра. Вид всех формул в этом случае остается таким

же, только вместо коэффициентов a_i появятся изображения многомерных ядер $H_i(s_1, \ldots, s_i)$. Как следует из формул (5.4), компенсатор получается сепарабельным [1] и будет физически реализуем, если реализуемо $\frac{H_d(s)}{H(s)}$.

Предложенное решение задачи компенсации удобно тем, что оно дает изображения ядер компенсатора. Вопрос их реализации был рассмотрен в [9], где показано, что ядра не очень высокого порядка можно синтезировать сравнительно простым образом.

Выводы

Функциональные ряды Вольтерра представляют собой математический аппарат, пригодный для решения довольно широкого круга нелинейных задач. Предложенный в статье метод определения изображений ядер Вольтерра и методы определения выходного сигнала позволяют в известной мере упростить решение задачи анализа аналитических систем.

Метод синтеза, рассмотренный в работе, удобен для построения линейной системы из заданной нелинейной, хотя он может быть использован как и общая методика, построенная в [1] для синтеза оптимальных нелинейных систем. Однако задача нелинейной компенсации в последней, общей формулировке имеет небольшую практическую ценность по следующим причинам. Во-первых, полученные ядра компенсатора высших порядков могут оказаться (и, как правило, оказываются) физически нереализуемыми, даже если ядро первого порядка реализуемо [10]. Во-вторых, как было показано в [11], прямое определение изображений ядер желаемой системы возможно лишь для одного специального типа воздействий, на практике не встречающегося. При практическом применении методов анализа и синтеза нелинейных систем, рассмотренных в статье, необходимо контролировать сходимость полученных рядов Вольтерра. Для этого можно использовать метод, предложенный в [1].

Литература

1. *Ван-Трис Г.* Синтез оптимальных нелинейных систем управления. «Мир», 1964.
2. *Van Trees, H.*, Functional techniques for the analysis of nonlinear behaviour of phase locked loops. Proc. IEEE **52** (1964) 8.
3. *Flake, R.*, Volterra series representation of nonlinear systems. IEEE Trans. on Appl. and Ind. **82** (1963).
4. *Flake, R.*, Volterra series representation of time-varying nonlinear systems. Proc. of Second IFAC Congress, London, 1963.
5. *Ku, Y., Wolf, A.*, Volterra Wiener functionals for the analysis of nonlinear systems. J. Franklin Inst. (Jan. 1966).
6. *Barrett, J.*, The use of functionals in the analysis of nonlinear physical systems. J. of Electronics and Control **15** (1963).

Kielkewicz, M., Cascade connection of nonlinear systems. Int. Journ. of Control **11** (1970) 6.
8. *Lubbock, J., Bansal, V.*, Multidimensional Laplace transforms for solution of nonlinear equations. Proc. IEE **116** (1969) 12.
9. *Schetzen, M.*, Synthesis of a class of nonlinear systems. Int. Journ. of Electronics (1966) **5.**
10. *Киселев О. Н.* Условия физической реализуемости при статистической оптимизации нелинейных систем. Автоматика и телемеханика (1967) **9.**
11. *Katzenelson, J., Gould, L.*, A spectrum factorization method of calculation of nonlinear filters of Volterra type. Inf. and Control **8** (1965) 3.

The use of Volterra series in nonlinear system designing

K. A. PUPKOV, V. I. KAPALIN

(Moscow)

This paper considers the methods of Volterra series and multidimensional Laplace transforms application for nonlinear system analysis and synthesis. We give a brief summary of the paper in the introduction.

In the first part of the paper a general problem of defining Volterra kernels on the given block diagram of the time-invariant analytical system is to be formulated. This problem is solved by using multidimensional Laplace transforms while some formulas of this theory are to be given in the second part of the paper. In this section of the paper we give general formulas of direct (2.1), inverse (2.2) transforms, convolution formula, formula of association of variables (2.4) and some application results of the last formula for special types of transforms.

The third part of the paper concerns the method of defining Laplace transforms of Volterra kernels for nonlinear systems by using their block diagrams. Three types of connection are supposed to be in this system: parallel, cascade, inverse. That is why, the method which is based on the results of [6], [7] includes three formulas for these types of connection.

The fourth part of the paper considers the problems of defining the response of nonlinear analytical systems to regular inputs. For this purpose, we use multidimensional Laplace transform theory, introduce the Volterra kernel moments (4.3), and apply Fourier transform and equivalent frequency coefficient.

The last part of the paper discusses the problem of nonlinear compensation by using the method for defining Volterra kernels which was presented in the third part of the given paper. We suppose the desired system to be linear and time-invariant. The realizability of compensator is also discussed.

As a result the authors conclude that the methods suggested in the paper allow to find an analytical solution of nonlinear problems, if Volterra series constructed for nonlinear systems is converged.

Problems of Control and Information Theory, Vol. 3 (2), pp. 117—125 (1974)

РАЗБИЕНИЕ ЦИКЛИЧЕСКИХ КОДОВ НА РАВНОВЕСНЫЕ КЛАССЫ

С. Ш. ОГАНЕСЯН, В. И. ТАИРЯН, В. Г. ЯГДЖЯН

(Ереван)

(Поступила в редакцию 25 мая 1972 г.)

Выводятся формулы для выражений циклических и p-ичных представителей произвольных циклических кодов над конечными полями $GF(q)$.

Исследование корректирующих возможностей циклических кодов непосредственно связано с анализом распределения весов кодовых векторов, т. е. с нахождением весовых спектров кодов. Рациональность определения весовых спектров заключается в вычислении весов не всех кодовых векторов, а некоторого возможно меньшего их подмножества [1—5]. Это позволяет существенно сократить время вычисления весовых спектров на ЭВМ.

До настоящего времени делались попытки нахождения циклических представителей для произвольных кодов над полями Галуа $GF(q)$, где $(n, q) = 1$ (q — степень простого числа p) [1], [2], [5], но, к сожалению, не было найдено[1] общего аналитического выражения для таких представителей [5].

В данной работе приводятся формулы для выражений не только циклических представителей произвольных кодов, но и для, так называемых, p-ичных представителей,[2] которые объединяют циклические представители по одинаковым весам посредством операции возведения в степень p.

Известно, что циклический код можно рассматривать как множество полиномов из кольца полиномов R_n^q по модулю $X^n - 1$ над полем коэффициентов $GF(q)$ ($q = p^\mu$, μ — натуральное число), которые делятся на фиксированный полином $g(x)$. Такое множество полиномов называется идеалом, а $g(x)$ — порождающим полиномом этого идеала.

Пусть $\mathcal{U}$ — произвольный идеал над $GF(q)$ и полином $a(x) = \sum_{i=0}^{n-1} a_i x^i$ принадлежит идеалу $\mathcal{U}$. Число S ненулевых коэффициентов $a(x)$ называется весом полинома $a(x)$, т. е.

$$S = \sum_{i=0}^{n-1} |a|, \text{ где } |a_i| = \begin{cases} 0, & \text{если } a_i = 0 \\ 1, & \text{если } a_i \neq 0. \end{cases}$$

[1] Для случая $q = 2$ задача решена в [3].

[2] Для случая $q = 2$ при некоторых ограничениях квадратичные представители найдены в [4].

Ясно, что веса полиномов $a(x)$ и $x^{\mu} a(x) \in \mathcal{U}$ одинаковы. Следовательно, можно разбить все полиномы идеала $\mathcal{U}$ на классы – циклические классы равновесных полиномов.

Каждому циклическому классу поставим в соответствие число b, называемое циклическим периодом или, короче, периодом класса, равное количеству полиномов в классе и число S, называемое весом класса, равное весу полиномов этого класса. Тогда каждому полиному идеала $\mathcal{U}$ соответственно будет сопоставлено число b — период циклического класса, к которому этот полином принадлежит. Каждый циклический класс может быть представлен циклическим представителем, а весь идеал — множеством циклических представителей.

Пусть $\{a_1(x)\} = \{a_1(x), xa_1(x), \ldots, x^{b_1-1} a_1(x)\}$ и $\{a_2(x)\} = \{a_2(x), xa_2(x), \ldots$ $\ldots, x^{b_2-1} a_2(x)\}$ представляют собой соответственно некоторые циклические классы произвольных идеалов $\mathcal{U}_1$ и $\mathcal{U}_2$, не содержащих общих элементов, т. е. $\mathcal{U}_1 \cap \mathcal{U}_2 = 0$. При этом циклические представители $a_1(x)$ и $a_2(x)$ классов $\{a_2(x)\}$ и $\{a_2(x)\}$ имеют периоды b_1 и b_2 соответственно. Множество всех элементов вида $x^i a_1(x) + x^j a_2(x)$, где $i = \overline{0, b_1 - 1}$, $j = \overline{0, b_2 - 1}$, есть прямая сумма циклических классов $\{a_1(x)\}$ и $\{a_2(x)\}$ идеалов $\mathcal{U}_1$ и $\mathcal{U}_2$ соответственно.

Теорема 1. Прямая сумма циклических классов $\{ a_1(x) \}$ и $\{ a_2(x) \}$ распадается на (b_1, b_2) циклических классов периода $T = [b_1, b_2]$. При этом элементы

$$x^d a_1(x) + a_2(x), \quad d = \overline{0, (b_1, b_2) - 1} \tag{1}$$

находятся в различных циклических классах и могут быть выбраны в качестве их циклических представителей.

Доказательство. Определим период T элементов вида (1), т. е. наименьшее натуральное число, при котором справедливо сравнение

$$x^T(x^d a_2(x) + a_2(x) \equiv x^d a_1(x) + a_2(x) \bmod X^n - 1 . \tag{2}$$

По условию $\mathcal{U}_1 \cap \mathcal{U}_2 = 0$, поэтому из (2) имеем

$$x^{d+T} a_1(x) \equiv x^d a_1(x) \bmod X^n - 1, \tag{3}$$

$$x^T a_2(x) \equiv a_2(x) \bmod X^n - 1 . \tag{4}$$

Периоды $a_1(x)$ и $a_2(x)$ соответственно равны b_1 и b_2, поэтому из сравнений (3) и (4) получим

$$d + T \; d \bmod b_1, T \equiv 0 \bmod b_2 .$$

Следовательно, $T \equiv 0 \bmod [b_1, b_2]$, значит $T = [b_1, b_2]$. Дальнейшее доказательство теоремы сводится к доказательству двух положений.

1. Все элементы вида (1) находятся в различных циклических классах т. е. из выполнения сравнения

$$x^{\mu}(x^d a_1(x) + a_2(x)) \equiv x^{d'} a_1(x) + a_2(x) \bmod X^n - 1, \tag{5}$$

где $d, d' < (b_1, b_2)$, следует $d = d'$.

2. Циклические классы с представителями вида (1) содержат все элементы прямой суммы циклических классов $\{a_1(x)\}$ и $\{a_2(x)\}$.

1. Из сравнения (5), учитывая условие $\mathcal{U}_1 \cap \mathcal{U}_2 = 0$, получим:

$$x^{d+\mu} a_1(x) \equiv x^{d'} a_1(x), \; x^{\mu} a_2(x) \equiv a_2(x) \bmod X^n - 1.$$

Так как периоды $a_1(x)$ и $a_2(x)$ соответственно равны b_1 и b_2, то $d + \mu \equiv \equiv d' \bmod b_1$, $\mu \equiv 0 \bmod b_2$, т. е. $d \equiv d' \bmod (b_1, b_2)$, но $d, d' < (b_1, b_2)$, поэтому $d = d'$.

2. Количество элементов прямой суммы циклических классов $\{a_1(x)\}$ и $\{a_2(x)\}$ равно $N_1 = b_1 \cdot b_2$. Количество элементов циклических классов с представителями вида (1) равно $N_2 = (b_1, b_2)\,[b_1, b_2] = b_1 \cdot b_2$, т. е. $N_1 = N_2$, что и требовалось доказать.

Пусть $A_1 = \{\{a_1^{(1)}(x)\}, \ldots, \{a_i^{(1)}(x)\}, \ldots, \{a_f^{(1)}(x)\}\}$ и $A_2 = \{\{a_1^{(2)}(x)\}, \ldots \ldots, \{a_j^{(2)}(x)\}, \ldots, \{a_l^{(2)}(x)\}\}$ — множества всех циклических классов произвольных идеалов $\mathcal{U}_1$ и $\mathcal{U}_2$, которые не содержат общих элементов ($\mathcal{U}_1 \cap \mathcal{U}_2 = 0$), а циклические представители $a_i^{(1)}(x)$ и $a_j^{(2)}(x)$ имеют периоды $b_i^{(i)}$ и $b_j^{(2)}$ соответственно. Очевидно, что прямые суммы произвольных различных пар циклических классов идеалов $\mathcal{U}_1$ и $\mathcal{U}_2$ не содержат общих элементов ($\mathcal{U}_1 \cap \mathcal{U}_2 = 0$). Исходя из этого, на основании теоремы 1 определяются циклические представители циклических классов прямой суммы идеалов $\mathcal{U}_1$ и $\mathcal{U}_2$, т. е. идеала $\mathcal{U} = \mathcal{U}_1 \otimes \mathcal{U}_2$ по следующей теореме.

Теорема 2. Циклические представители циклических классов идеала $\mathcal{U}$ имеют вид

$$x^{d_{ij}} a_i^{(2)} + a_j^{(2)}, \quad d_{ij} = \overline{0, (b_i^{(1)}, b_j^{(2)}) - 1}, \quad i = \overline{1, f}, \quad j = \overline{1, l}.$$

Пусть A_i представляют собой множества[3] всех циклических классов минимальных идеалов $\mathcal{U}_i$, $i = \overline{1, t}$. Тогда, опираясь на теорему 2, методом полной математической индукции определяются циклические представители циклических классов прямой суммы идеалов $\mathcal{U}_i$, $i = 1, t$, идеала $\mathcal{U} = \mathcal{U}_1 \oplus \oplus \mathcal{U}_2 \oplus \ldots \oplus \mathcal{U}_i \oplus \ldots \oplus \mathcal{U}_i$ по теореме.

[3] Метод определения множеств A_i приведён в работе [1].

Теорема 3. Циклические представители циклических классов идеала $\mathfrak{U}$ имеют вид:

$$\sum_{i=1}^{t-1} x^{\sum_{j=i}^{t-1} d_j} a_i(x) + a_t(x), \tag{1}$$

где $a_i(x)$ — представитель класса $\{a_i(x)\} \in A_i$; d_j выбираются при фиксированном $a_i(x)$, $i = \overline{1, t}$ из множества $D_j = \{0, 1, \ldots, ([b_1, b_2, \ldots, b_j], b_{j+1}) - 1\}$. При этом период T циклического представителя равен $[b_1, b_2, \ldots, b_t]$.

Пусть $A = \{\{a_1(x)\}, \ldots, \{a_i(x)\}, \ldots, \{a_j(x)\}, \ldots, \{a_t(x)\}\}$ есть множество всех циклических классов произвольного идеала $\mathfrak{U}$. Отметим правомерность следующей операции над классами: $\{a_i(x)\}\{a_j(x)\} = \{a_i(x)\, a_j(x)\}$. Поэтому $\{a_i(x)\}^p = \{a_i^p(x)\}$, и так как характеристика поля коэффициентов полиномов идеала $\mathfrak{U}$ равна p, то вес циклического класса $\{a_i^p(x)\}$ равен весу класса $\{a_i(x)\}$.

Разобьем множество A посредством возведения в степень p на новые классы, которые назовём p-ичными классами или классами II порядка (собственно циклические классы суть классы I порядка). Обозначим множество классов II порядка идеала $\mathfrak{U}$ через R, т.е. $R = \{\{r_1(x)\}, \ldots, \{r_i(x)\}, \ldots, \{r_e(x)\}\}$, где $\{r_j(x)\} = \{\{a_j(x)\}, \{a_j(x)\}^p, \ldots, \{a_j(x)\}^{p^{c_i}}\}$, c_i — количество классов I порядка в i-ом классе II порядка, называемое его p-периодом.

Теорема 4. Циклические периоды классов $\{a(x)\}$ и $\{a(x)\}^p$ идеала $\mathfrak{U}$ одинаковы.

Доказательство. Утверждение теоремы очевидно для минимального идеала, учитывая [1]. Так как произвольный идеал $\mathfrak{U}$ представляет собой прямую сумму некоторого числа минимальных идеалов, т. е. $\mathfrak{U} = \mathfrak{U}_1 \oplus \mathfrak{U}_2 \oplus \ldots \oplus \mathfrak{U}_i \oplus \ldots \oplus \mathfrak{U}_t$, $t \geq 1$, то любой элемент $a(x)$ идеала $\mathfrak{U}$ представляется в виде $a(x) = \sum_{i=1}^{t} a_i(x)$, где $a_i(x) \in \mathfrak{U}_i$. При этом период T элемента $a(x)$ равен $[b_1, b_2, \ldots, b_t]$ (теорема 3). Период T_1 циклического класса $\{a(x)\}^p$ определится из сравнения:

$$a_t^{T_1}(a_1(x) + a_2(x) + \ldots + a_t(x))^p \equiv (a_1(x) + a_2(x) + \ldots + a_t(x) \bmod X^n - 1. \tag{6}$$

Учитывая, что $\mathfrak{U}_i \cap \mathfrak{U}_j = 0$ по всем $i \neq j$ и справедливость теоремы для любого минимального идеала, из сравнения (6) имеем:

$$T_1 \equiv 0 \bmod b_i, \qquad i = \overline{1, t},$$

т. е. $T_1 = [b_1, b_2, \ldots b_t]$ и $T_1 = T$, что и требовалось доказать.

Любой полином $a(x)$ идеала $\mathcal{U}$ характеризуется весом S, числом b, равным периоду класса I порядка, содержащего $a(x)$, числом c, равным p-периоду класса II порядка, в который входит класс I порядка, содержащий полином $a(x)$. Каждый класс II порядка может быть представлен p-ичным представителем и весь идеал — множеством p-ичных представителей.

Пусть $R_1 = \{\{r_1(x)\}, \{r_1(x)\}^p, \ldots, \{r_1(x)\}^{p^{c_1-1}}\}$ и $R_2 = \{\{r_2(x)\}, \{r_2(x)\}^p, \ldots, \{r_2(x)\}^{p^{c_2-1}}\}$ — соответственно некоторые классы II порядка произвольных идеалов $\mathcal{U}_1$ и $\mathcal{U}_2$, не содержащих общих элементов, т. е. $\mathcal{U}_1 \cap \mathcal{U}_2 = 0$. При этом p-ичные представители $r_1(x)$ и $r_2(x)$ классов R_1 и R_2 имеют периоды b_1 и b_2 соответственно и p-периоды c_1 и c_2.

Прямая сумма классов II порядка R_1 и R_2 есть совокупность прямых сумм классов I порядка $\{r_1(x)\}^{p^i}$ и $\{r_2(x)\}^{p^j}$ по всем $i = \overline{0, c_1 - 1}$ и $j = \overline{0, c_2 - 1}$, т. е.

$$R_1 \oplus R_2 = \bigcup_{i,j} \left(\{r_1(x)\}^{p^i} \oplus \{r_2(x)\}^{p^j}\right).$$

Разобьем множество $D = \{0, 1, \ldots, (b_1, b_2) - 1\}$ на p-циклы посредством подстановки $\omega \to p\omega \bmod t = (b_1, b_2)$, где $\omega \in D$. Количество p-циклов обозначим через K; количество чисел в i-ом p-цикле, называемое p-периодом — через t_i, а наименьший представитель p-цикла — через ω_i. Обозначим $F_i = x^{\omega_i}$.

Теорема 5. Прямая сумма классов II порядка R_1 R_2 идеалов $\mathcal{U}_1$ и $\mathcal{U}_2$ распадается на $\sum_{i=1}^{k} (c_1, t_i)([c_1, t_i], c_2)$ классов II порядка p-периода $T_p = [c_2, t_i, c_2]$, $i = \overline{1, k}$. Элементы

$$r_1^{p^{\varphi+\gamma}}(x) \cdot F_i^{p^\gamma} + r_2(x), \tag{7}$$

где $\varphi = \overline{0, (c_1, t_i) - 1}$, $\gamma = \overline{0, ([c_1, t_i], c_2) - 1}$, $i = \overline{1, k}$ находятся в различных классах II порядка и могут быть выбраны в качестве p-ичных представителей.

Доказательство. Определим p-период T_p классов II порядка с представителями вида (7), т. е. наименьшее натуральное число, при котором справедливо сравнение:

$$x^\nu \left(r_1^{p^{\gamma+\varphi}}(x) F_i^{p^\gamma} + r_2(x)\right)^{p^{T_p}} \equiv r_1^{p^{\gamma+\varphi}}(x) F_i^{p^\gamma} + r_2(x) \bmod X^n - 1. \tag{8}$$

Так как $F_i = x^{\omega_i}$ и $\mathcal{U}_1 \cap \mathcal{U}_2 = 0$, то из (8) получим:

$$r_1^{p^{\gamma+\varphi+T_p}}(x) \cdot x^{\omega_i p^{\gamma+T_p}} + \nu \equiv r_1^{p^{\varphi+\gamma}}(x) x^{\omega_i p^\gamma}, \quad r_2^{p^{T_p}}(x) x^\nu \equiv r_2(x) \bmod X^n - 1. \tag{9}$$

p-периоды $r_1(x)$, $r_2(x)$, ω_i и периоды $r_1(x)$, $r_2(x)$ соответственно равны c_1, c_2, t_i и b_1, b_2, поэтому из сравнений (9) имеем: $T_p \equiv 0 \bmod c_1$, $T_p \equiv 0 \bmod c_2$, $T_p \equiv 0 \bmod t_i$, т. е. $T_p = [c_1, t_i, c_2]$.

Дальнейшее доказательство теоремы сводится к доказательству двух положений.

1. Все элементы вида (7) находятся в различных классах II порядка, т. е. из выполнения сравнения:

$$x^{\nu}\left(r_1^{p^{\varphi+\gamma}}(x)\,F_i^{p^{\gamma}}+r_2(x)\right)^{p^{\mu}}\equiv r_1^{p^{\varphi'+\gamma'}}(x)\,F_j^{p^{\gamma'}}+r_2(x)\ \bmod\ X^n-1\,, \qquad (10)$$

где ν, μ — натуральные числа; $i, j<k$; $\varphi<(c_1, t_i)$; $\gamma<([c_1, t_i], c_2)$; $\varphi'<$ $<(c_1, t_i)$; $\gamma'<([c_1, t_j], c_2)$, следует $i=j$, $\varphi=\varphi'$, $\gamma=\gamma'$.

2. Классы II порядка с представителями вида (7) содержат все элементы прямой суммы классов II порядка R_1 и R_2.

1. Учитывая, что $F_i=x^{\omega_i}$, $F_j=x^{\omega_j}$ и $\mathfrak{U}_1\cap\mathfrak{U}_2=0$, из сравнения (10) получим:

$$r_1^{p^{\varphi+\gamma+\mu}}(x)\,x^{\omega_i p^{\varphi+\mu}+\nu}\equiv r_1^{p^{\varphi'+\gamma'}}(x)\,x^{\omega_j p^{\gamma'}}\ \bmod\ X^n-1\,, \qquad (11)$$

$$r_2^{p^{\mu}}(x)\,x^{\nu}\equiv r_2(x)\ \bmod\ X^n-1\,. \qquad (12)$$

Здесь p-периоды и периоды $r_1(x), r_2(x)$ соответственно равны c_1, c_2 и b_1, b_2, поэтому из сравнений (11) и (12) имеем:

$$\varphi+\gamma+\mu\equiv\varphi'+\gamma'\bmod c_1, \qquad (13)$$

$$\mu\equiv 0\bmod c_2, \qquad (14)$$

$$\omega_i\,p^{\gamma+\mu}+\nu\equiv\omega_j\,p^{\gamma'}\bmod b_1, \qquad (15)$$

$$\nu\equiv 0\bmod b_2. \qquad (16)$$

Из сравнений (15) и (16) следует:

$$\omega_i\,p^{\gamma+\mu}\equiv\omega_j\,p^{\gamma'}\bmod t=(b_1, b_2), \qquad (17)$$

т. е. ω_i и ω_j находятся в одном p-цикле по модулю t, и так как они суть наименьшие представители p-цикла, то $i=j$. Тогда, исходя из того, что p-период ω_i равен t_i, из (17) получим:

$$\gamma+\mu\equiv\gamma'\bmod t_i. \qquad (18)$$

Из сравнений (13) и (18) следует: $\varphi\equiv\varphi'\bmod(c_1, t_i)$, т. е. $\varphi=\varphi'$. Тогда из сравнений (13), (18) и (14) имеем:

$$\gamma\equiv\gamma'\bmod([c_1, t_i], c_2),\quad \text{значит}\quad \gamma=\gamma'.$$

2. Количество элементов прямой суммы классов II порядка R_1 и R_2 равно $N_3=c_1\cdot b_1\cdot c_2\cdot b_2$. Количество же элементов в классах II порядка с представителями вида (7) при фиксированном F_i равно $(c_1, t_i)([c_1, t_i], c_2)$

$[c_1, t_i, c_2]\,[b_1, b_2] = c_1 \cdot c_2 \cdot [b_1, b_2] \cdot t_i$. Их общее количество равно $N_4 = \sum_{i=1}^{k} c_1 \cdot c_2 \cdot [b_1, b_2]\, t_i = c_1 \cdot c_2 \cdot [b_1 \cdot b_2] \cdot \sum_{i=1}^{k} t_i = c_1 \cdot c_2 \cdot [b_1, b_2]\,(b_1, b_2) = c_1 \cdot c_2 \cdot b_1 \cdot b_2$, т. е. $N_3 = N_4$, что и требовалось доказать.

Пусть $R^{(1)} = \{\{r_2^{(1)}(x)\}, \ldots, \{r_i^{(1)}(x)\}, \ldots, \{r_d^{(1)}(x)\}\}$ и $R^{(2)} = \{\{r_1^{(2)}(x)\}, \ldots, \{r_j^{(2)}(x)\}, \ldots, \{r_n^{(2)}(x)\}\}$ суть множества всех классов II порядка произвольных идеалов $\mathcal{U}_1$ и $\mathcal{U}_2$, которые не содержат общих элементов ($\mathcal{U}_1 \cap \mathcal{U}_2 = 0$) и p-ичные представители $r_i^{(i)}(x)$, $r_j^{(2)}(x)$ имеют периоды $b_i^{(1)}, b_j^{(2)}$ и p-периоды $c_1^{(1)}, c_j^{(2)}$ соответственно.

Очевидно, что прямые суммы произвольных различных классов II порядка идеалов $\mathcal{U}_1$ и $\mathcal{U}_2$ не содержат общих элементов ($\mathcal{U}_2 \cap \mathcal{U}_2 = 0$).

Для каждой пары классов II порядка $\{r_i^{(1)}(x)\}$ и $\{r_j^{(2)}(x)\}$ образуем множество $D_{ij} = \{0, 1, \ldots, (b_i^{(1)}, b_i^{(2)}) - 1\}$. Обозначим количество чисел в D_{ij} через t_{ij} и разобьём каждое множество на непересекающиеся $(p, t_{ij}) = 1$ циклы посредством подстановки $\omega \to p\omega \bmod t_{ij}$. Множество p-ичных представителей этих циклов обозначим через W_{ij}, а множество $\{x^{\omega_{ij}}\}$, где $\omega_{ij} \in W_{ij}$ — через $R = \{r_{11}(x), \ldots, r_{ij}(x), \ldots, r_{dh}(x)\}$, $i = \overline{1, d}$, $j = \overline{1, h}$. Количество чисел в ij-ом цикле обозначим через c_{ij}. Тогда на основании теоремы 5 p-ичные представители классов II порядка прямой суммы $\mathcal{U} = \mathcal{U}_1 \oplus \mathcal{U}_2$ идеалов $\mathcal{U}_1$ и $\mathcal{U}_2$ определяются по теореме.

Теорема 6. p-ичные представители классов II порядка идеала $\mathcal{U}$ имеют вид:

$$\left(r_i^{(2)}(x)\right)^{p^{\varphi+\gamma}} \cdot r_{ij}^{p^{\gamma}} + r_j^{(2)}(x),$$

где $\varphi = \overline{0, (c_i^{(1)}, c_{ij}) - 1}$, $\gamma = \overline{0, ([c_i^{(1)}, c_{ij}], c_j^{(2)}) - 1}$, $i = \overline{1, d}$, $j = \overline{1, h}$.

Согласно [1], множество циклических представителей классов I порядка минимального идеала $\mathcal{U}_i$ есть $\{0, \alpha_i^0(x)\,\varepsilon_i(x), \alpha_i^1(x)\,\varepsilon_i(x), \ldots, \alpha_i^{m_i-1}(x)\,\varepsilon_i(x)\}$, поэтому подстановка $\omega \to p\omega \bmod m_i$, где $(p, m_i) = 1$, $\omega \in \{0, 1, \ldots, m_i - 1\}$ разбивает множество классов I порядка на классы II порядка.

Выберем в качестве p-ичных представителей классов II порядка циклические представители классов I порядка в наименьшей степени для каждого класса II порядка соответственно.

Пусть R_{2i-1}, $i = \overline{1, t}$ суть множества p-ичных представителей классов II порядка минимальных идеалов $\mathcal{U}_i$. Образуем множество всевозможных, различных наборов вида $(r_1(x), \ldots, r_{2i-1}(x), \ldots, r_{2i-1}(x))$, где $r_{2i-1}(x) \in R_{2i-1}$ $i = \overline{1, t}$. Для каждого такого набора построим семейство множеств $\{D_2, \ldots, D_{2i}, \ldots, D_{2t-2}\}$, где $D_{2i} = \overline{0, ([b_1, \ldots, b_i], b_{i+1}) - 1}$, $i = \overline{1, t}$. Количество чисел в D_{2i} обозначим через t_i и разобьём каждое множество чисел D_{2i} на непересекающиеся p-циклы посредством подстановки $\omega \to p\omega \bmod t_i ((t_i, p) = 1)$.

Множество p-ичных представителей этих циклов обозначим через W_{2i} множество $\{r_{2i}(x) = x^{\omega_{2i}}\}$, где $\omega_{2i} \in W$ — через R_{2i}. Каждое $r_{2i}(x)$, называемое также p-ичным представителем, характеризуется числом C_{2i}, равным количеству чисел в p-цикле с p-ичным представителем ω_{2i}. Множество всевозможных различных наборов вида $(r_2(x), \ldots, r_{2i}(x), \ldots, r_{2t-2}(x))$ обозначим через M.

В указанных обозначениях, методом полной математической индукции, на основании теоремы 6, определяются p-ичные представители классов II порядка прямой суммы минимальных идеалов $\mathcal{U}_i$, $i = \overline{1, t}$, идеала $\mathcal{U} = \mathcal{U}_1 \oplus \mathcal{U}_2 \oplus \ldots \oplus \mathcal{U}_i \oplus \ldots \oplus \mathcal{U}_t$ по теореме.

Теорема 7. p-ичные представители классов II порядка идеала $\mathcal{U}$ можно задать выражением вида

$$\sum_{i=1}^{t-1} \left(r_{2i-1}^{\sum_{j=2i-1}^{2t-2} \varphi_j}(x) \prod_{v=i}^{t-1} r_{2i}^{\sum_{j=2v}^{2t-2} \varphi_j}(x) \right) + r_{2t-1}(x), \tag{II}$$

где $r_1(x), \ldots, r_{2t-1}(x)$ — элементы набора $\omega \in W$, $r_2(x), \ldots, r_{2i-2}(x)$ — элементы набора, принадлежащего множеству M при фиксированном ω, а $\varphi_1, \ldots, \varphi_i, \ldots, \varphi_{2t-2}$ выбираются из множества чисел $\{0, 1, \ldots, ([c_1, \ldots, c_i], c_{i+1}) - 1\}$, $i = \overline{1, 2t-2}$ при фиксированных r_i, $i = \overline{1, 2t-1}$. p-период T_p p-ичных представителей равен $[c_1, c_2, \ldots, c_{2t-2}]$.

Литература

1. *Goethals, J. M.*, Analysis of weight distribution in binary cyclic codes. IEEE Trans. **IT-12** (1966) **3**, 401—402.
2. *Мак-Вильямс Дж.* Структура и свойства бинарных циклических алфавитов. Кибернетический сборник (1967) **4**, 7—42.
3. *Оганесян С. Ш., Ягджян В. Г.* Нахождение циклических представителей в бинарных циклических алфавитах. Труды ВЦ АН Арм. ССР и ЕГУ (1969) **6**, 35—38.
4. *Оганесян С. Ш., Ягджян В. Г.* Объединение циклических представителей по одинаковым весам. Труды ВЦ АН Арм. ССР и ЕГУ (1969) **6**, 39—48.
5. *Tavares, S. E., Allard, P. E., Shiva, S. G. S.*, On the decomposition of cyclic codes into cyclic classes. Information and Control, New York (1971) **18** *4*, 342—354.

Decomposition of cyclic codes into equal-weight classes

S. SH. OGANESYAN, V. I. TAIRYAN, V. G. YAGDZYAN

(Erevan)

Investigation of correcting possibilities of the cyclic codes is connected immediately with the analysis of weight distribution of code vector that is with computation of the code weight spectrum.

In the determination of weight spectrum it is reasonable to compute the weights not of all the code vectors, but of a certain their subset being as small as possible [1—5]. It allows to reduce considerably the time of spectrum calculation on electronic computers.

Up to now attempts have been made to find cyclic representatives for arbitrary codes over $GF(q)$, where $(n, q) = 1$ (q is the power of prime number) [1], [2], [5], but unfortunately the common analytical expression for such representatives have not been found[1] [5]. In this paper formulas are given to obtain not only the cyclic representatives,[2] connecting the cyclic representatives of the same weight by means of operation of raising to the power of "p".

Let A_i[3] be the set of all cyclic representatives of minimal ideals $\mathfrak{U}_i$, $i = \overline{I, t}$. Then cyclic representatives of the direct sum of ideals $\mathfrak{U}_i$, $i = \overline{I, t}$, of the ideal $\mathfrak{U} = \mathfrak{U}_1 \oplus \mathfrak{U}_2 \oplus \oplus \ldots \oplus \mathfrak{U}_i \oplus \ldots \oplus \mathfrak{U}_i$ are determined with the theorem:

Theorem 3. Cyclic representatives of ideals $\mathfrak{U}$ have a form (I).

Theorem 7. "p" representatives of ideals $\mathfrak{U}$ have a form (II).

С. Ш. Оганесян
Вычислительный центр АН Арм. ССР
СССР, Ереван 44, ул. Паруира Севака, 1

[1] For $q = 2$ the problem is solved in [3].
[2] For $q = 2$ with some limitations quadratic representatives are found in [4].
[3] Method of determination of sets A_i is given in [1].

Problems of Control and Information Theory, Vol. 3 (2), pp. 127—136 (1974)

ЛИНЕЙНОЕ КОДИРОВАНИЕ ПРИ РЕТРАНСЛЯЦИИ СИГНАЛА

Н. К. МИЛЕНИН

(Рязань)

(Поступила в редакцию 23 марта 1971 г.)

Рассматривается задача синтеза оптимальных линейных кодирующих и декодирующих фильтров в линии связи с многократной ретрансляцией сигнала. Получены формулы для оптимальных передаточных функций оконечных и промежуточных звеньев при условии, что средние мощности передатчиков одинаковы, а энергетические спектры шумов на всех участках линии связи совпадают.

1. Введение

Для передачи информации на большие расстояния применяется многократная или однократная ретрансляция сигнала, например в радиорелейных и кабельных линиях связи, а также в системах связи через искусственные спутники Земли. Упрощенная блок-схема линии связи с многократной ретрансляцией сигнала показана на рисунке, где введены следующие обозначения: $S(\omega)$ — энергетический спектр полезного сигнала, $K_1(j\omega)$ и $W(j\omega)$ — передаточные функции оконечных кодирующего и декодирующего фильтров соответственно, $K_2(j\omega)$, $K_3(j\omega)$, . . ., $K_n(j\omega)$ — передаточные функции промежуточных кодирующих фильтров (ретрансляторов), $N_i(\omega)$ — энергетический спектр аддитивного флуктуационного шума, воздействующего на i-том участке линии связи, $i = 1, 2, 3, \ldots, n$, где n — общее число участков. Причем предполагается, что ослабление полезного сигнала на каждом i-том участке линии связи уже скомпенсировано в последующем ретрансляторе, а энергетический спектр шума $N_i(\omega)$ также соответствующим образом пересчитан ко входу этого ретранслятора.

Цель данной работы заключается в отыскании оптимальных передаточных функций кодирующих $[K_i(j\omega)_0$, $i = 1, 2, \ldots, n$, и декодирующего $W_0(j\omega)$ фильтров без учета условия их физической осуществимости. При этом считается, что полезный сигнал и шумы являются независимыми стационарными случайными процессами с равными нулю математическими ожиданиями, а все звенья линии связи линейны.

В качестве критерия оптимальности используется предложенный в работе [1] частотно-взвешенный среднеквадратичный критерий с неодинаковыми весовыми функциями $A^2(\omega)$ и $M^2(\omega)$ для динамических и случайных

ошибок, т. е.

$$\sigma^2 = \sigma_g^2 + \sigma_ш^2 = \frac{1}{2\pi} \int\limits_{\omega_н}^{\omega_в} S_0(\omega)\, A^2(\omega)\, d\omega + \frac{1}{2\pi} \int\limits_{\omega_н}^{\omega_в} N_0(\omega)\, M^2(\omega)\, d\omega = \min, \quad (11)$$

где $\sigma_д^2$ и $\sigma_ш^2$, $S_0(\omega)$ и $N_0(\omega)$ — взвешенные дисперсии и энергетические спектры динамических и случайных ошибок соответственно, σ^2 — дисперсия результирующей взвешенной ошибки, $f_в = \omega_в/2\pi$ и $f_н = \omega_н/2\pi$ — верхняя и нижняя границы полосы пропускания линии связи, которые находятся из условия: $A(\omega) = 0$ при $\omega \geq \omega_в$) и $\omega \leq \omega_н$. Причем считается, что динамические ошибки возникают за счет отличия выходного полезного сигнала от входного, а случайные ошибки определяются результирующим уровнем шумов на выходе линии связи.

2. Нахождение оптимальной передаточной функции декодирующего фильтра

Используя принцип суперпозиции и рисунок, нетрудно найти энергетические спектры динамических и случайных ошибок,

$$S_0(\omega) = S(\omega) \left| W(j\omega) \prod_{i=1}^{n} K_i(j\omega) - 1 \right|^2, \quad (2.1)$$

$$N_0(\omega) = |W(j\omega)|^2 \sum_{k=1}^{n} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2, \quad (2.2)$$

где при $k = n$ $\prod\limits_{i=k+1}^{n} |K_i(j\omega)|^2 = 1$.

Подставляя выражения (2.1) и (2.2) в (1.1), получаем

$$\sigma^2 = \frac{1}{2\pi} \int\limits_{\omega_н}^{\omega_в} \left[S(\omega)\, A^2(\omega) \left| W(j\omega) \prod_{i=1}^{n} K_i(j\omega) - 1 \right|^2 + \right.$$

$$\left. + M^2(\omega) |W(j\omega)|^2 \sum_{k=1}^{n} N_k(j\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2 \right] d\omega. \quad (2.3)$$

Подберем передаточную функцию декодирующего фильтра так, чтобы при фиксированных значениях передаточных функций всех кодирующих фильтров $K_i(j\omega)$, $i = 1, 2, \ldots, n$ величина дисперсии σ^2 (2.3) суммарной взвешенной ошибки была минимальна. Для этой цели фазовую характеристику декодирующего фильтра ($\arg W(j\omega)$) нужно выбрать такой, чтобы передаточная функция всей линии связи была действительным числом, т. е.

$$W(j\omega) \prod_{i=1}^{n} K_i(j\omega) = \left| W(j\omega) \prod_{i=1}^{n} K_i(j\omega) \right|. \quad (2.4)$$

Оптимальное значение передаточной функции декодирующего фильтра $W_0(j\omega)$ можно найти, если составить уравнение Эйлера для подынтегральной функции $R[W(j\omega)]$ в выражении (2.3)

$$\partial R\,[W(j\omega)]/\partial W(j\omega) = 0 \tag{2.5}$$

и решить это уравнение при выполнении условия (2.4).

Из выражений (2.3)—(2.5) получаем

$$W_0(j\omega) = \frac{S(\omega) \prod\limits_{i=1}^{n} K_i^*(j\omega)}{S(\omega) \prod\limits_{i=1}^{n} |K_i(j\omega)|^2 + L^2(\omega) \sum\limits_{k=1}^{n} N_k(\omega) \prod\limits_{i=k+1}^{n} |K_i(j\omega)|^2}, \tag{26}$$

где $K_i^*(j\omega)$ — функции, сопряженные передаточным функциям $K_i(j\omega)$, а

$$L(\omega) = \frac{M(\omega)}{A(\omega)}.$$

Подставляя соотношение (2.6) в (2.3), находим, что минимальная величина дисперсии σ_0^2 суммарной взвешенной ошибки при фиксированных значениях передаточных функций всех кодирующих фильтров $K_i(j\omega)$ равна:

$$\sigma_0^2 = \frac{1}{2\pi} \int\limits_{\omega_{\text{н}}}^{\omega_{\text{в}}} F[|K_i(j\omega)|]\,d\omega =$$

$$= \frac{1}{2\pi} \int\limits_{\omega_{\text{н}}}^{\omega_{\text{в}}} \frac{S(\omega)\,M^2(\omega) \sum\limits_{k=1}^{n} N_k(\omega) \prod\limits_{i=k+1}^{n} |K_i(j\omega)|^2}{S(\omega) \prod\limits_{i=1}^{n} |K_i(j\omega)|^2 + L^2(\omega) \sum\limits_{k=1}^{n} N_k(\omega) \prod\limits_{i=k+1}^{n} |K_i(j\omega)|^2}\,d\omega. \tag{2.7}$$

3. Синтез оптимальных передаточных функций кодирующих фильтров

Будем считать, что средняя мощность Q_s любого передатчика (оконечного или промежуточного) с номером s ограничена, т. е.

$$Q_s = \frac{1}{2\pi} \int\limits_{\omega_{\text{н}}}^{\omega_{\text{в}}} \Phi_s[|K_i(j\omega)|]\,d\omega =$$

$$= \frac{1}{2\pi} \int\limits_{\omega_{\text{н}}}^{\omega_{\text{в}}} \left[S(\omega) \prod\limits_{i=1}^{s} |K_i(j\omega)|^2 + \sum\limits_{k=1}^{s-1} N_k(\omega) \prod\limits_{i=k+1}^{s} |K_i(j\omega)|^2 \right] d\omega, \tag{3.1}$$

где $s = 1, 2, \ldots, n$. Тогда оптимальные значения модулей передаточных

функций кодирующих фильтров $|K_r(j\omega)|_0$, $r = 1, 2, \ldots, n$, обеспечивающих при ограниченных средних мощностях передатчиков Q_s (3.1) минимум дисперсии взвешенной ошибки σ_0^2 (2.6), можно найти, если решить изопериметрическую задачу вариационного исчисления [2]. Для этой цели нужно составить следующую систему из n уравнений Эйлера

$$\frac{\partial}{\partial |K_r(j\omega)|}\left\{F[|K_i(j\omega|] + \sum_{s=1}^{n} \lambda_s \Phi_s [|K_i(j\omega)|]\right\} = 0 \Bigg| \qquad (3.2)$$

$$r = 1, 2, \ldots, n,$$

где $F[|K_i(j\omega)|]$ и $\Phi_s[|K_i(j\omega)|]$ – подынтегральные функции в формулах (2.7) и (3.1) соответственно, а λ_s — некоторые постоянные множители, которые находятся с помощью подстановки оптимальных значений модулей передаточных функций кодирующих фильтров $|K_i(j\omega)|_0$, $i = 1, 2, \ldots, n$, в начальные условия (3.1). Причем фазовые характеристики кодирующих фильтров ($\arg [K_i(j\omega)]_0$, $i = 1, 2, \ldots, n$) можно выбрать произвольными, так как выполнение условия (2.4) обеспечивается оптимальным выбором фазовой характеристики декодирующего фильтра, т. е. $\arg W_0(j\omega)$.

После дифференцирования из соотношений (2.7, 3.1, 3.2) получаем следующую систему из n алгебраических уравнений для определения оптимальных значений модулей передаточных функций кодирующих фильтров:

$$\frac{S^2(\omega) M^2(\omega) \prod_{i=1}^{n} |K_i(j\omega)|^2 \sum_{k=r}^{n} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2}{\left[S(\omega) \prod_{i=1}^{n} |K_i(j\omega)|^2 + L^2(\omega) \sum_{k=1}^{n} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2\right]^2} =$$

$$= \sum_{S=r}^{n} \lambda_s \left[S(\omega) \prod_{i=1}^{S} |K_i(j\omega)|^2 + \sum_{k=1}^{r-1} N_k(\omega) \prod_{i=k+1}^{S} |K_i(j\omega)|^2\right] \Bigg| \qquad (3.3)$$

$$r = 1, 2, \ldots, n,$$

где при $K = n \prod_{i=k+1}^{n} |K_i(j\omega)|^2 = 1$, а $\sum_{k=1}^{r-1} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2 = 0$, если $r = 1$.

В процессе решения системы уравнений (3.3) необходимо учитывать очевидное условие $|K_i(j\omega)|_0 \geq 0$ для всех значений $i = 1, 2, \ldots, n$. В частности, при $n = 1$, т. е. в линии связи без ретрансляции сигнала, из соотношения (3.3) находим, что

$$|K_1(j\omega)|_0^2 = \begin{cases} \sqrt{\dfrac{N_1(\omega) M^2(\omega)}{\lambda_1 S(\omega)}} - \dfrac{L^2(\omega) N_1(\omega)}{S(\omega)} & \text{при} \quad h_1(\omega) \leq 1, \\ 0 & \text{при} \quad h_1(\omega) > 1, \end{cases} \qquad (3.9)$$

где

$$h_1(\omega) = \frac{\lambda_1 L^4(\omega) N_1(\omega)}{S(\omega) M^2(\omega)}.$$

Подставив выражение (3.4) в формулы (2.7) и (3.1), находим наименьшее значение дисперсии $\sigma^2_{\min}$ суммарной взвешенной ошибки при оптимальных передаточных функциях декодирующего $W_0(j\omega)$ (2.6) и оконечного кодирующего $|K_1(j\omega)|_0$ (3.4) фильтров, а также постоянную λ_1 в линии связи без ретрансляции сигнала (при $n = 1$)

$$\sigma^2_{\text{мин}} = \frac{1}{2\pi} \int\limits_{\Delta\omega_1} \sqrt{\lambda_1 S(\omega) N_1(\omega) M^2(\omega)}\, d\omega + \frac{1}{2\pi} \int\limits_{\Delta\omega_2} S(\omega) A^2(\omega)\, d\omega, \qquad (3.5)$$

$$\sqrt{\lambda_1} = \frac{\int\limits_{\Delta\omega_1} \sqrt{N_1(\omega) S(\omega) M^2(\omega)}}{\int\limits_{\Delta\omega_2} L^2(\omega) N_1(\omega)\, d\omega + \int\limits_{\omega_{\text{н}}}^{\omega_{\text{в}}} S(\omega)\, d\omega}, \qquad (3.6)$$

где $\Delta\omega_1$ и $\Delta\omega_2$ — области частот, расположенные в полосе пропускания линии связи $\Delta\omega = \omega_{\text{в}} - \omega_{\text{н}} = \Delta\omega_1 + \Delta\omega_2$, каждая из которых определяется неравенством $h_1(\omega) \leq 1$ и $h_1(\omega) > 1$ соответственно.

Соотношения (3.4, 3.5, 3.6) были получены ранее: в наиболее общем виде — в работе [1], при $L(\omega) = M(\omega) = 1$ — в работах [3, 4], а при $L(\omega) = 0$ — в работе [5]. Они будут использованы далее для сравнения со случаем многократной ретрансляции сигнала.

К сожалению, в общем случае (при $n \geq 2$) разрешить в радикалах систему уравнений (3.3) не удается. Здесь для её решения можно использовать один из приближенных методов [6].

4. Один частный случай

С целью упрощения системы уравнений (3.3) потребуем, чтобы подынтегральные выражения в соотношении (3.1) для средних мощностей передатчиков при любых значениях $s = 1, 2, \ldots, n$ совпадали между собой, т. е. будем считать, что

$$\Phi_s[|K_i(j\omega)|] = \Phi_1[|K_i(j\omega)|] = S(\omega)\,|K_1(j\omega)|^2, \quad s = 2, 3, \ldots, n. \qquad (4.1)$$

Очевидно, что при выполнении условия (15) средние мощности оконечного и всех промежуточных передатчиков должны быть равными друг другу, т. е.

$$Q_s = Q_1 = \frac{1}{2\pi} \int\limits_{\omega_{\text{н}}}^{\omega_{\text{в}}} S(\omega)\,|K_1(j\omega)|^2\, d\omega, \qquad s = 2, 3, \ldots, n. \qquad (4.2)$$

С другой стороны, из соотношения (3.1) следует, что

$$\Phi_s[|K_i(j\omega)|] = S(\omega)\prod_{i=1}^{s}|K_i(j\omega)|^2 + \sum_{k=1}^{s-1} N_k(\omega)\prod_{i=k+1}^{s}|K_i(j\omega)|^2 =$$

$$= |K_s(j\omega)|^2\left[S(\omega)\prod_{i=1}^{s-1}|K_i(j\omega)|^2 + \sum_{k=1}^{s-2} N_k(\omega)\prod_{i=k+1}^{s-1}|K_i(j\omega)|^2 + N_{s-1}(\omega)\right] =$$

$$= |K_s(j\omega)|^2\left[\Phi_{s-1}[|K_i(j\omega)|] + N_{s-1}(\omega)\right]. \tag{4.3}$$

Подставляя условие (4.1) в формулу (4.3), находим связь между модулями передаточных функций оконечного и промежуточного (с номером $s = i$) кодирующих фильтров

$$|K_s(j\omega)|^2 = \frac{S(\omega)\,|K_1(j\omega)|^2}{S(\omega)\,|K_1(j\omega)|^2 + N_{s-1}(\omega)}, \qquad s = 2, 3, \ldots, n. \tag{4.4}$$

Для простоты в дальнейшем будем считать, что энергетические спектры шумов на всех участках линии связи совпадают между собой, т. е.

$$N_s(\omega) = N_i(\omega) = N(\omega), \qquad s = i = 1, 2, \ldots, n. \tag{4.5}$$

В этом случае модули передаточных функций (4.4) всех ретрансляторов становятся одинаковыми

$$D(\omega) = |K_i(j\omega)|^2 = \frac{S(\omega)\,|K_1(j\omega)|^2}{S(\omega)\,|K_1(j\omega)|^2 + N(\omega)}, \qquad i = 2, 3, \ldots, n. \tag{4.6}$$

Кроме того, нетрудно показать, что условия (4.1, 4.2, 4.5) приводят к равенству друг другу всех постоянных λ_s, т. е.

$$\lambda_s = \lambda, \qquad s = 1, 2, \ldots, n. \tag{4.7}$$

Подставляя соотношения (4.6) и (4.7) в (3.3) и учитывая при этом, что

$$\sum_{k=1}^{n} N_k(\omega)\prod_{i=k+1}^{n}|K_i(j\omega)|^2 = N(\omega)\sum_{k=1}^{n}\prod_{i=k+1}^{n} D(\omega) =$$

$$= N(\omega)\sum_{k=1}^{n} D^{n-k}(\omega) = N(\omega)\,\frac{D^n(\omega) - 1}{D(\omega) - 1}$$

как сумма n членов геометрической прогрессии, получаем следующее уравнение для определения модуля оптимальной передаточной функции $|K_1(j\omega)|_0$ оконечного кодирующего фильтра

$$\frac{N(\omega)\,M^2(\omega)}{\lambda S(\omega)}\,|K_1(j\omega)|^{2(n-1)} =$$

$$= \left[|K_1(j\omega)|^2 + \frac{N(\omega)}{S(\omega)}\right]^{n+1}\left\{L^2(\omega) + [1 - L^2(\omega)]\,\frac{|K_1(j\omega)|^{2n}}{\left[|K_1(j\omega)|^2 + \dfrac{N(\omega)}{S(\omega)}\right]^n}\right\}^2 \tag{4.8}$$

Уравнение (4.8) разрешимо в радикалах при $n=1$, 2, и 3. При других значениях n для его решения можно воспользоваться одним из приближенных методов [6]. Причем это уравнение существенно упрощается, если выбран частотно-взвешенный критерий $(L(\omega)=1)$ или критерий максимума отношения сигнал/шум $(L(\omega)=0)$ [1]. При этом оно оказывается разрешимым в радикалах, когда $n=1$, 2, 3, 5 и 7.

Задаваясь конкретными значениями n и $L(\omega)$, из уравнения (4.8) получаем:

1. Если $L(\omega)=0$ и $n=2$, то

$$|K_1(j\omega)|_0^2=\sqrt{\frac{4N(\omega)\,M^2(\omega)}{3\,\lambda\,S(\omega)}}\left[\begin{array}{ll}\cos\dfrac{\varphi_1(\omega)}{3} & \text{при}\quad h_2(\omega)\leq 1,\\[2ex] \operatorname{ch}\dfrac{\varphi_2(\omega)}{3} & \text{при}\quad h_2(\omega)>1\end{array}\right], \tag{4.9}$$

где $\cos\varphi_1(\omega)=\operatorname{ch}\varphi_2(\omega)=h_2(\omega)$; $h_2(\omega)=\sqrt{\dfrac{27\,\lambda\,N(\omega)}{4\,S(\omega)\,M^2(\omega)}}$.

2. При условии, что $L(\omega)=0$ и $n=3$,

$$|K_1(j\omega)|_0^2=\sqrt{\frac{N(\omega)\,M^2(\omega)}{\lambda\,S(\omega)}}\left[\frac{1}{2}+\sqrt{\frac{1}{4}+\sqrt{\frac{\lambda\,N(\omega)}{S(\omega)\,M^2(\omega)}}}\right]. \tag{4.10}$$

3. Когда $L(\omega)=1$ и $n=3$, то

$$|K_1(j\omega)|_0^2=\begin{cases}\left(\sqrt{\dfrac{N(\omega)\,M^2(\omega)}{\lambda\,S(\omega)}}-\dfrac{N(\omega)}{S(\omega)}\right)\left[1+\dfrac{\sqrt{\dfrac{1}{4}-B(\omega)}-\dfrac{1}{2}}{1-B(\omega)}\right] & \text{при } B(\omega)<\dfrac{1}{4},\\[2ex] 0 \quad \text{при} \quad B(\omega)>\dfrac{1}{4},\end{cases} \tag{4.11}$$

где

$$B(\omega)=\sqrt{\frac{\lambda\,N(\omega)}{S(\omega)\,M^2(\omega)}}.$$

Подобные (но более громоздкие) соотношения для $|K_1(j\omega)|_0^2$ можно получить и для других значений n и $L(\omega)$.

Первый сомножитель в формулах (4.9, 4.10, 4.11) с точностью до постоянного множителя совпадает с квадратом модуля оптимальной передаточной функции оконечного кодирующего фильтра в линии связи без ретрансляции сигнала ($n=1$), а второй сомножитель (в квадратных скобках) характеризует поправку, учитывающую влияние ретрансляции сигнала. При-

чем, из соотношения (4.8) следует, что в общем случае этой поправкой можно пренебречь только при условии, что $S(\omega)M^2(\omega) \gg n^2\lambda N(\omega)$, т. е. при очень малых шумах.

После определения из уравнения (4.8) оптимального значения модуля передаточной функции $|K_1(j\omega)|_0$ оконечного кодирующего фильтра нетрудно

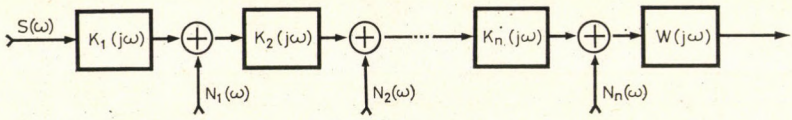


Рис. 1

найти квадраты модулей оптимальных передаточных функций промежуточных кодирующих фильтров (ретрансляторов) $D_0(\omega)$ и декодирующего фильтра $|W_0(j\omega)|^2$, а также вычислить наименьшую дисперсию $\sigma^2_{\min}$ суммарной взвешенной ошибки при оптимальных передаточных функциях всех кодирующих и декодирующих фильтров. Используя для этой цели соотношения (2.6), (2.7), (4.6) и (4.8), получаем

$$D_0(\omega) = |K_i(j\omega)|_0^2 = \frac{S(\omega)\,|K_1(j\omega)|_0^2}{S(\omega)\,|K_1(j\omega)|_0^2 + N(\omega)}, \qquad i = 2, 3, \ldots, n,$$

$$|W_0(j\omega)|^2 = \frac{\lambda S(\omega)\,|K_1(j\omega)|_0^2}{N(\omega)\,M^2(\omega)},$$

$$\delta^2_{\text{мин}} = \frac{1}{2\pi}\int\limits_{\omega_{\text{н}}}^{\omega_{\text{в}}} \frac{S(\omega)\,M^2(\omega)\,[1 - D_0^n(\omega)]}{L^2(\omega) + [1 - L^2(\omega)]\,D_0^n(\omega)}\,d\omega\,.$$

Постоянная λ находится с помощью подстановки конкретного значения функции $|K_1(j\omega)|_2^0$ в начальное условие (4.2).

5. Заключение

В работе рассмотрена задача построения оптимальных линейных кодирующих и декодирующих фильтров в линии связи с многократной ретрансляцией сигнала. В качестве критерия оптимальности применен наиболее общий частотно-взвешенный среднеквадратичный критерий с неодинаковыми весовыми функциями для динамических и случайных ошибок [1], частным случаем которого являются широко известные критерии минимума среднего квадрата ошибки [3, 4] и максимума отношения сигнал/шум [5].

Найдены общие выражения (2.6) и (2.7) для оптимальной передаточной функции оконечного декодирующего фильтра и минимальной дисперсии взвешенной ошибки при фиксированных передаточных функциях всех кодирующих фильтров. Эти выражения дают возможность решить задачу оптимальной фильтрации сигнала на оконечной приемной станции.

Сформулирована вариационная задача для отыскания оптимальных передаточных функций оконечного и промежуточных кодирующих фильтров при ограниченных средних мощностях передатчиков. Составлена система алгебраических уравнений (3.3), позволяющая вычислить эти передаточные функции.

Подробно исследован важный для практики случай, когда средние мощности всех передатчиков одинаковы, а энергетические спектры на всех участках линии связи также совпадают. При этом найдены достаточно простые соотношения для оптимальных передаточных функций декодирующего и кодирующих фильтров. Причем, частотные характеристики ретрансляторов должны быть равномерными только при очень малых шумах, когда выполняется условие $S(\omega)M^2(\omega) \gg \lambda n^2 N(\omega)$. В противном случае форма оптимальных частотных характеристик ретрансляторов существенно отличается от равномерной в рабочей полосе частот. Последнее обстоятельство необходимо учитывать при проектировании оптимальных линий связи с многократной ретрансляцией сигнала.

Литература

1. *Миленин Н. К.* Линейное кодирование в системе передачи информации с дополнительными источниками шумов. II Международный симпозиум по теории информации, г. Цахкадзор, Арм. ССР. Тезисы докладов, Москва—Ереван, 1971, стр. 175—178.
2. *Эльсгольц Л. Э.* Дифференциальные уравнения и вариационное исчисление. «Наука», М., 1969.
3. *Costas, J. P.*, Coding with Linear Systems. Proc. IRE **40** (1952) *9*, 1101—1103.
4. *Цыбаков Б. С.* Линейное кодирование сообщений. Радиотехника и электроника (1962) **7** *1*, 25—38.
5. *Штейн В. М.* О расчете предыскажающих и корректирующих устройств. Радиотехника **11** (1956) 2, 60—63.
6. *Мишина А. П., Проскуряков И. В.* Высшая алгебра. «Физматгиз», М., 1962.

Linear coding in the retransmission of the signal

N. K. MILENIN

(Ryazan')

The synthesis of optimum coding and decoding linear filters in the communication line with multiple signal retransmission is carried out. The criterion of the minimum of the weighing dispersion error is offered to use. That error is equal to the sum of the dynamic and random errors, which have been weighed of different laws. The optimum transfer functions of the input, output and intermediate links are found in that case, when mean powers of the transmitters are equal and the power density noise spectrum on all sections of communication line is coincided.

Н. К. Миленин
Радиотехнический институт
СССР Рязань 24, ул. Гагарина, 59/1

Problems of Control and Information Theory, Vol. 3 (2), pp. 137—145 (1974)

GLOBAL CONTROLLABILITY OF PERTURBED QUASI-LINEAR SYSTEMS

A. G. KARTSATOS

(Tampa)

(Received December 6, 1973)

Conditions are given on the matrices $A(t, x, u)$, $B(t, x, u)$ and the vector $(Q(t, x, u)$, ensuring the controllability of the system

$$(*) \qquad x' = A(t, x, u)x + B(t, x, u)\, u + Q(t, x, u).$$

The control is continuous and can be chosen with initial value equal to zero. Recent results of Davison and Kunze, Lukes are extended.

1. Introduction

Consider the differential control system

$$(*) \qquad x' = A(t, x, u)x + B(t, x, u)u + Q(t, x, u)$$

where A, B are matrices of size $n \times n$, $n \times r$ respectively, and Q is an n-vector.

The aim of this paper is to provide conditions on A, B, Q such that the system (*) is *controllable* in the following sense:

"*for every* $t_0, t_1 \in R_1 = (-\infty, \infty)$, $x_0, x_1 \in R^n$ *there exists a continuous control* $u : [t_0, t_1] \to R^r$ *with* $u(t_0) = 0$, *which produces a response* $x(t)$ *satisfying the boundary conditions* $x(t_0) = x_0$, $x(t_1) = x_1$."

The method employed here is based on the fact that if the *linear* systems

$$(*)_{fg} \qquad x' = A(t, f(t), g(t))\, x + B(t, f(t), g(t))u + Q(t, f(t), g(t))$$

(where (f, g) lies in a suitable subset of a Banach space) are controllable, then it is possible to show the controllability of (*) by use of a fixed point theorem. This idea was exploited by Davison and Kunze [4], who considered the case $Q \equiv 0$ and $A \equiv A(t, x), B \equiv B(t, x)$. However, their method cannot be applied in the case (*), since the integral equation, to which the problem is reduced in [4], would now contain $u(t)$ besides the solution $x(t)$.

On the other hand, Lukes [6] has shown that in the case $A =$ const., $B =$ const., the problem of the controllability of (*) can be reduced to the problem of solving a system of two integral equations in $u(t)$ and $x(t)$. He actually

proved a stronger result w.r.t. u. Here we use the method developed by Lukes in [6] in order to extend to the case (*) the main results in [4], [6].

For other related results the reader is referred to Dauer [1—3]. The author wishes to thank Prof. Lukes for some helpful suggestions.

Instead of the points t_0, t_1 in the definition of controllability, we will be considering the points $0, T$ respectively, and this does not affect the generality. We also put $x(T) = x_T$.

2. Preliminaries

In what follows in this section, $I = [0, T]$, where T is a fixed positive constant, $R = (-\infty, \infty)$, and $R_+ = [0, \infty)$. For a vector $x \in R^k$, $\|x\|_k$ denotes the Euclidean norm and for an $m \times k$ matrix $A = [A_{ij}]$ we put

$$\| A \| = \sup_{\|x\|=1} \| Ax \| .$$

A^* will denote the transpose of A. By C_T^j we denote the space of all R^j-valued continuous functions f on I with norm

$$\| f \|_j = \sup_{t \in I} \| f(t) \| .$$

Now assume that the $n \times n$ matrix A, the $n \times r$ matrix B, and the vector Q in (*) are defined and continuous on $R \times R^n \times R^n$, and let $(f, g) \in C_T^n \times C_T^r$.

Then by $X(t; f, g)$ we denote the fundamental matrix of the linear system

$$x' = A(t, f(t), g(t))\, x, \quad t \in I, \tag{2.1}$$

which satisfies $X(0; f, g) = E$ (the $n \times n$ identity matrix). Now we are ready for the following lemma:

Lemma 2.1. Let

$$U(t; f, g) = \int_{T-t}^{T} X(T; f, g)\, X^{-1}(s; f, g),\, B(s, f(s), g(s))\, ds \tag{2.2}$$

for every $t \in I$, and let

$$S_T(f, g) = \int_0^T U(s; f, g)\, U^*(s; f, g)\, ds . \tag{2.3}$$

Assume that $S_T(f, g)$ has an inverse $S_T^{-1}(f, g)$, and let

$$V_T(t, f, g) = \int_{T-t}^{T} U^*(s; f, g)\, ds, \tag{2.4}$$

$$W_T(t, f, g) = \int_0^t X(t; f, g)\, X^{-1}(s; f, g)\, B(s, f(s), g(s))\, V_T(s; f, g)\, ds. \tag{2.5}$$

Then any solution to the system

$$x(t) = X(t; x, u)x_0 + W_T(t; x, u)\, y(x, u) +$$

$$+ \int_0^t X(t; x, u)\, X^{-1}(s; x, u)\, Q(s, x(s), u(s))\, ds, \tag{2.6}$$

$$u(t) = V_T(t; x, u)\, y(x, u), \tag{2.7}$$

where

$$y(x, u) = S_T^{-1}(x, u)\, [x_T - X(T; x, u)\, x_0 -$$

$$- S_T^{-1}(x, u) \int_0^T X(T; x, u)\, X^{-1}(s; x, u)\, Q(s, x(s), u(s))\, ds, \tag{2.8}$$

is a solution to the boundary value problem

$$(*) \qquad x' = A(t, x, u)\, x + B(t, x, u)\, u + Q(t, x, u)$$

$$x(0) = x_0,\ x(T) = x_T,\ u(0) = 0. \tag{2.9}$$

Proof. We first show that $W_T(T; f, g) = S_T(f, g)$.

In fact,

$$W_T(T; f, g) = \int_0^T X(T; f, g)\, X^{-1}(s; f, g)\, B(s, f(s), g(s))\, V_T(s; f, g)\, ds =$$

$$= \int_0^T X(T; f, g)\, X^{-1}(s; f, g)\, B(s, f(s), g(s)) \int_{T-s}^T U^*(u; f, g)\, du\, ds =$$

$$= \int_0^T \int_{T-\tau}^T X(T; f, g)\, X^{-1}(u; f, g)\, B(u, f(u), g(u))\, du\, U^*(\tau; f, g)\, d\tau =$$

$$= \int_0^T U(\tau; f, g)\, U^*(\tau; f, g)\, d\tau = S_T(f, g). \tag{2.10}$$

Using the above equality we easily obtain $x(0) = x_0$, $x(T) = x_T$, $u(0) = 0$. Now we have

$$\int_0^t X(t; x, u)\, X^{-1}(s; x, u)\, B(s, x(s), u(s))\, u(s)\, ds = \tag{2.11}$$

$$= \int_0^t X(t; x, u)\, X^{-1}(s; x, u)\, B(s, x(s), u(s))\, V_T(s; x, u)\, y(x, u)\, ds = W_T(t; x, u)\, y.$$

Thus, from (2.6) we obtain

$$x(t) = X(t\,;x,u)\,x_0 + \tag{2.12}$$

$$+\int_0^t X(t\,;x,u)\,X^{-1}(s\,;x,u)\,[B(s,x(s),u(s))\,u(s) + Q(s,x(s),u(s))]\,ds,$$

which is obviously the general formula of the solutions of (*).

3. Main result

The main result of this paper is contained in the following theorem. In view of the last remark in the introduction, the variable t will be considered only in the interval $R_+ = [0, \infty)$.

Theorem 3.1. For the system (*) *assume the following:*

(i) *A, B are of size $n\times n$ and $n\times r$ resp., and are defined and continuous on $R_+\times R^n\times R^r$ and such that there exist two continuous functions $p, q : R_+ \to R_+$ such that $\| A(t, v_1\, v_2) \| \leq p(t)$, $\| B(t, v_1, v_2) \| \leq q(t)$ for every $(t, v_1, v_2) \in R_+$;*

(ii) *Q is an n-vector defined and continuous on $R_+\times R^n\times R^r$ and such that for every $T \in R_+$,*

$$\liminf_{m\to\infty} \int_0^T \sup_{(v_1,v_2)\in\Sigma_m} \| Q(s, v_1, v_2) \|\, ds = 0, \tag{3.1}$$

where $\Sigma_m = \{(v_1, v_2) \in R^n\times R^r\,;\ \| v_1 \|_n + \| v_2 \|_r \leq m\}$[*)].

(iii) Det $[S_T(f,g)] \geq M_T > 0$ *for all* $(f,g) \in C_T^n\times C_T^r$ *and all* $T \in R_+$, *where M_T depends only on T.*

Then the system (*) *is controllable.*

Proof. Fix $T > 0$, and let $B = C_T^n\times C_T^r$ have the norm

$$\| (f,g) \|_B = \| f \|_n + \| g \|_r\,. \tag{3.2}$$

Then B is Banach space. Now consider the operators $T_1 : B \to C_T^n$, $T_2 : B \to C_T^r$ given by

$$[K_1(f,g)]\,(t) = X(t\,;f,g)\,x_0 + W_T(t\,;f,g)\,y(f,g) +$$

$$+\int_0^t X(t\,;f,g)\,X^{-1}(s\,;f,g)\,Q(s,f(s),g(s))\,ds \tag{3.3}$$

$$[K_2(f,g)]\,(t) = V_T(t;f,g)\,y(f,g)\,. \tag{3.4}$$

Then the problem of controllability of (*) reduces to finding a fixed point of the operator $K : B \to B$, given by

$$K(f, g) = (T_1(f, g),\ T_2(f, g))\,;\ (f, g) \in B. \tag{3.5}$$

In order to achieve this (by Tychonov's theorem), we show first that there is a closed ball $D \subset B$ containing the origin and such that $KD \subseteq D$.

In fact assume that this is not the case, and let

$$D_m = \{(f, g) \in B;\ \|(f, g)\|_B \leq m\}, \quad m = 1, 2, \ldots .$$

Then for each $m = 1, 2, \ldots$ there exists at least one $(f_m, g_m) \in D_m$ such that $\| K(f_m, g_m) \| > m$. Now let $0 \leq s \leq t \leq T$, and consider the matrix $X(t; f, g)\, X^{-1}(s; f, g)$ for some $(f, g) \in B$. This matrix is the unique solution to the matrix system

$$X' = A(t, f(t), g(t))\, X, \quad X(s) = E. \tag{3.6}$$

Thus, fixing s, integrating (3.6) from s to $t \geq s$, and applying Gronwall's inequality, we get

$$\| X(t; f, g)\, X^{-1}(s; f, g) \| \leq \exp\Big\{ \int_0^T p(\tau)\, d\tau \Big\} = \lambda_T > 0. \tag{3.7}$$

Now, by (iii), $S_T^{-1}(f, g)$ exists for every $(f, g) \in B$ and $\| S_T^{-1}(f, g) \| \leq N_T$ (a constant depending on T) (cf. Davison and Kunze [4], p. 492).

Combining this fact, (3.7) and (i), we obtain from (3.3) and (3.4):

$$\| K_1(f, g) \|_n \leq \alpha_T + \beta_T \int_0^T \| Q(s, f(s), g(s)) \|\, ds, \tag{3.8}$$

$$\| K_2(f, g) \|_r \leq \gamma_T + \delta_T \int_0^T \| Q(s, f(s), g(s)) \|\, ds, \tag{3.9}$$

where α_T, β_T, λ_T, δ_T are nonnegative constants independent of (f, g). Letting $\alpha_T + \gamma_T = \varepsilon_T$, and $\beta_T + \delta_T = \zeta_T$, we obtain

$$\begin{aligned} \| K(f, g) \|_B &= \| K_1(f, g) \|_n + \| K_2(f, g) \|_r \\ &\leq \varepsilon_T + \zeta_T \int_0^T \| Q(s, f(s), g(s)) \|\, ds, \end{aligned} \tag{3.10}$$

and, in particular,

$$m < \| K(f_m, g_m) \|_B \leq \varepsilon_T + \zeta_T \int_0^T \| Q(s, f_m(s), g_m(s))\, ds \| . \tag{3.11}$$

Dividing by m we obtain

$$1 < \frac{\varepsilon_T}{m} + \frac{\zeta_T}{m} \int_0^T \sup_{(v_1, v_2) \in \Sigma_m} \| Q(s, v_1, v_2) \| \, ds \,, \tag{3.12}$$

which implies a contradiction in view of (3.1). Thus, there exists an m, say m_0, such that $TD_{m_0} \subseteq D_{m_0}$.

We put $D_{m_0} = D$. It is easy to see now that the set TD is relatively compact. ($\overline{TD}$ is compact, where the closure is taken w.r.t. the norm is B.) In fact, there are constants η_T, θ_T such that, for each $(f, g) \in D$, $t, t' \in I$,

$$\| [K_1(f, g)] (t) - [K_1(f, g) (t')] \| \leq \eta_T \mid t - t' \mid , \tag{3.13}$$

$$\| [K_2(f, g) (t) - [K_1(f, g) (t')] \| \leq \theta_T \mid t - t' \mid . \tag{3.14}$$

Thus, the sets $K_1 D$, $K_2 D$ are relatively compact in C_T^n, C_T^r respectively, and $KD = (K_1 D, K_2 D)$. To show the continuity of K on D, let $(f_i, g_i) \in D$, $(f, g) \in D$ with

$$\lim_{i \to \infty} \| (f_i, g_i) - (f, g) \|_B = 0.$$

This implies that

$$\lim_{i \to \infty} \| f_i - f \|_n = 0, \ \lim_{i \to \infty} \| g_i - g \|_r = 0. \tag{3.15}$$

Since $K(f_i, g_i) \in KD$, which is a relatively compact set, there exists a subsequence $\{(f_{i_k}, g_{i_k})\}$, $k = 1, 2, \ldots$, of the sequence $\{(f_i, g_i)\}$ such that

$$\lim_{k \to \infty} \| K(f_{i_k}, g_{i_k}) - (f_0, g_0) \|_B = 0, \tag{3.16}$$

where $(f_0, g_0) \in D$. Now we show that $K(f, g) = (f_0, g_0)$. We first notice that

$$X'(t \,; f_{i_k}, g_{i_k}) = A(t, f_{i_k}(t), g_{i_k}(t)) \, X(t \,; f_{i_k}, g_{i_k}) \tag{3.17}$$

$$X'(t \,; f, g) \quad = A(t, f(t), g(t)) \, X(t \,; f, g). \tag{3.18}$$

Subtracting (3.18) from (3.17), integrating from 0 to $t \geq 0$, and then taking norms and applying Gronwall's inequality, we obtain

$$\| X(t \,; f_{i_k}, g_{i_k}) - X(t \,; f, g) \| \leq$$

$$\leq \lambda_T^2 \int_0^T \| A(s, f_{i_k}(s), g_{i_k}(s)) - A(s, f(s), g(s)) \| \, ds. \tag{3.19}$$

Since the integrand in (3.19) converges to zero uniformly on I,

$$\lim_{t\to\infty} \| X(t\,; f_{i_k}, g_{i_k}) - X(t; f, g) \| = 0 \tag{3.20}$$

uniformly on I. To show that (3.20) is also true if X is replaced by X^{-1}, let $X_k \equiv X(t\,; f_{i_k}, g_{i_k})$, $X(t) \equiv X(t\,; f, g)$.

Then it is easy to see that

$$[X_k^{-1}(t)]' = - X_k^{-1}(t)\, A(t, f(t), g(t))\, X_k^{-1}(t); \;\; t \in I. \tag{3.21}$$

Following the above steps with $X_k^{-1}(t)$ instead of $X_k(t)$, we obtain

$$\lim_{t\to\infty} \| X_k^{-1}(t) - X^{-1}(t) \| = 0, \text{ uniformly on } I. \tag{3.22}$$

Consequently,

$$\| X(t, f_{i_k}, g_{i_k})\, X^{-1}(s, f_{i_k}, g_{i_k}) - X(t\,; f, g)\, X^{-1}(s\,; f, g) \| \tag{3.23}$$

converges to zero uniformly on the set $\{(t, s) \in I^2; t \geq s\}$. This immediately implies that

$$\lim_{k\to\infty} S_T(f_{i_k}, g_{i_k}) = S_T(f, g). \tag{3.24}$$

We show that

$$\lim_{k\to\infty} S_T^{-1}(f_{i_k}, g_{i_k}) = S_T^{-1}(f, g). \tag{3.25}$$

In fact, let $S_k = S_T(f_{i_k}, g_{i_k})$, $S = S_T(f, g)$. Then we have

$$\| S_k^{-1} - S^{-1} \| = \| S^{-1}(S - S_k)\, S_k^{-1} \| \leq \| S^{-1} \| \, \| S_k^{-1} \| \, \| S - S_k \|, \tag{3.26}$$

from which it follows that $\| S^{-1} - S_k^{-1} \| \to 0$ as $k \to \infty$, because $\| S^{-1} \| \leq \leq N_T$, $\| S_k^{-1} \| \leq N_T$. The rest of the proof of the fact $K(f, g) = (f_0, g_0)$ follows now easily from (3.23), (3.24), (3.26) and Eqs (3.3) and (3.4). Since we could have started with any subsequence of $\{K(f_i, g_i)\}$, instead of $\{K(f_i, g_i)\}$ itself, we have actually shown that every subsequence of $\{K(f_i, g_i)\}$ contains a subsequence converging to $K(f, g)$. It follows that

$$\lim_{i\to\infty} \| K(f_i, g_i) - K(f, g) \| = 0, \tag{3.27}$$

which proves the continuity of K on D. By Tychonov's theorem, there is $(x, u) \in D$ such that $(x, u) = K(x, u)$. Since $T > 0$ is arbitrary, the proof is complete.

Remark 3.1. It should be noted that Lukes [6] considered the case in which $Q(t, v_1, v_2)$ is bounded on $[0, \infty) \times R^n \times R^r$, or Lipschitzian with respect to v_1 and v_2. It is obvious that Th. 3.1 allows perturbations $Q(t, v_1, v_2)$ satisfying

$$\| Q(t, v_1, v_2) \| \leq \lambda_1(t) \| v_1 \|_n^{\sigma} + \lambda_2(t) \| v_2 \|_r^{\sigma_2}, \tag{3.28}$$

where $0 < \sigma_1, \sigma_2 < 1$, and λ_1, λ_2 are nonnegative and continuous on $[0, \infty)$.*

Now consider the differential control system

$$x' = A_0(t, x, u) + B_0(t, x, u), \tag{3.29}$$

where A_0, B_0 are n-vectors defined and continuous on $[0, \infty) \times R^n \times R^r$. Assume further that A_0 is continuously differentiable with respect to x, and B_0 is continuously differentiable with respect to u. Then there exist two matrices $A(t, x, u)$, $B(t, x, u)$ of sizes $n \times n$, $n \times r$ respectively, defined and continuous on $[0, \infty) \times R^n \times R^r$ and such that

$$A_0(t, x, u) = A(t, x, u)x + A_0(t, 0, u), \tag{3.30}$$

$$B_0(t, x, u) = B(t, x, u)u + B_0(t, x, 0).^{**} \tag{3.31}$$

Now we have the following

Corollary 3.1. Assume that the vectors $A_0(t, x, u)$, $B_0(t, x, u)$ in (3.29) *ary defined and continuous on $[0, \infty) \times R^n \times R^r$ and, moreover, that the vector A_0 is continuously differentiable with respect to x, while the vector B_0 is continuousle differentable with respect to u. Then if the matrices A, B in* (3.30), (3.31) *and the vector $Q(t, x, u) \equiv A_0(t, 0, u) + B(t, x, 0)$ satisfy the hypotheses of Th.* 3.1, *the system* (3.29) *is controllable.*

It is clear that Examples 1, 2 in Davison and Kunze [4] can be modified so as to satisfy the hypotheses of Th. 3.1.

* However, for this reason the hypothesis (ii) in Th. 3.1 must be modified in a simpler way: the equation

$$\lim_{m \infty \to} \inf \frac{1}{m} \int_0^T \sup_{(v_1, v_2) \in \Sigma_m} \| Q(s, v_1, v_2) \| \, ds = 0$$

must be held with $0 < \alpha < 1 - \max(\sigma_1 \sigma_2)$. (Remark of the Editorial Board.)

** Cf. Lakshmikantham and Leela ([5,] p. 73).

References

1. *Dauer, J. D.*, Perturbations of linear control systems, SIAM J. Control **9** (1971), 393—400.
2. *Dauer, J. D.*, A note on bounded perturbations of controllable systems, J. Math. Anal. Appl. **42** (1973), 221—225.
3. *Dauer, J. D.*, On controllability of systems of the form $x_3 = A(t)x + g(t, u)$, J. Optimization Th. Appl. **11** (1973), 132—138.
4. *Davison, E. J., Kunze, E. G.*, Some sufficient conditions for the global and local controllability of nonlinear time-varying systems, SIAM J. Control **8** (1970), 489—497.
5. *Lakshmikantham, V., Leela, S.*, Differential and integral inequalities. Theory and applications, Vol. 1, Academic Press, New York 1969.
6. *Lukes, D. L.*, Global controllability of nonlinear systems. **10** (1972), 112—126.

Управляемость в целом квазилинейных систем в условиях возмущений

А. Г. КАРТСАТОС

(Тампа)

Получены условия для матриц $A(t, x, u)$, $B(t, x, u)$ и вектора $Q(t, x, u)$, обеспечивающие управляемость системы

$$(*) \qquad x' = A(t, x, u)\, x + B(t, x, u)\, u + Q(t, x, u).$$

Управление предполагается непрерывным и обладающим нулевым начальным значением. Найденные условия являются обобщением известных результатов Дэвизона—Кунце, Лукеса.

A. G. Kartsatos
Department of Mathematics
University of South Florida
Tampa, Florida 33620, USA

Problems of Control and Information Theory, Vol. 3 (2), pp. 147–155 (1974)

ON THE α-DIVERGENCE RATE FOR MARKOV-DEPENDENT HYPOTHESES*

T. NEMETZ

(Budapest)

(Received November 13, 1973)

In testing statistical hypotheses Rényi's α-divergence plays a very important role. In this paper we give an asymptotic expression for the α-divergence of two probability measures associated with Markov chains and discuss its connection with the error probability.

1. Introduction

The discussion in this paper is motivated mainly by the following simple vs. simple hypothesis testing problem, where the Bayesian approach is accepted.

Let ϑ be a parameter-random variable with the a priori probability distribution $w_0 = \text{Prob}\ (\vartheta = \vartheta_0) = 1 - \text{Prob}\ (\vartheta = \vartheta_1) = 1 - w_1$, where $w_0 \cdot w_1 > 0$.

Let ξ be a random variable defined on a certain probability space and taking its values on the sample space $\{X, \mathcal{X}\}$. Denote the probability measure generated by ξ on $\{X, \mathcal{X}\}$ under the hypotheses $\mathcal{H}_0 := \{\vartheta = \vartheta_0\}$ and $\mathcal{H}_1 := \{\vartheta = \vartheta_1\}$ by P and Q respectively, and let the corresponding density func) tions with respect to a given dominating σ-finite measure μ be $p(x)$ and $q(x$- respectively.

Let us consider an arbitrary decision $d(\xi)$ on the basis of the observation ξ. The error probability of the decision rule $d(.)$ is defined as the probability of wrong decision:

$$e(d) = w_0 P(d(\xi) = \vartheta_1) + w_1 Q(d(\xi) = \vartheta_0).$$

The minimal error prob ability

$$e(w_0, P, Q) = \inf_{(d)} e(d)$$

will be referred simply as error probability. It is well known (see e.g. Rényi [9]) that the minimal error probability may always be achieved.

* An earlier version of this paper was presented at the Sixth Prague Conference on Information Theory, Prague, September, 1971.

It is also known that the error probability may be bounded both from below and above by the Hellinger integral of order α. The Hellinger integral of order α of two probability measures P_1 and P_2 is defined as follows:

$$H_\alpha(P_1, P_2) = \int_X \left[\frac{dP_1}{d\mu}\right]^\alpha \left[\frac{dP_2}{d\mu}\right]^{1-\alpha} d\mu\,, \quad 0 < \alpha < 1\,. \tag{1.1}$$

There is a monotone one-to-one correspondence between $H_\alpha(P_1, P_2)$ and the divergence of order α defined by A. Rényi as $\frac{1}{\alpha - 1} \log H_\alpha(P_1, P_2)$ (see Csiszár [1]). Therefore we will refer to $H_\alpha(.\,,.)$ itself as α-divergence hoping that this will not cause misunderstanding.

The following known inequalities make use of α-divergences:

$$e(\Delta) = e(w_0, P, Q) \leq \frac{1}{2} \inf_{0<\alpha<1} H_\alpha(P, Q) \tag{1.2}$$

(see e.g. Hellman and Raviv [3]).

$$H_{1/2}(P, Q) \leq \sqrt{\frac{e(\Delta)}{w_0 \cdot w_1}} \tag{1.3}$$

(see Nemetz [6]).

$$\frac{1}{2} H_\alpha(P, Q) \leq [e(\Delta)]^{\min(\alpha, 1-\alpha)} [1 - e(\Delta)]^{\max(\alpha, 1-\alpha)} \tag{1.4}$$

(see Vajda [10]).

Noting that $H_\alpha(.\,,.)$ is multiplicative on independent experiments, it follows immediately that if ξ is a sequence of i.i.d. random variables under both hypotheses, then the error probability tends to zero exponentially, and we have an upper estimation for the error exponent which is the same as the Chernoff-bound. J. F. Korsh [5] has obtained that for ergodic Markov-experiments the error probability goes to zero exponentially, too. The Markov-case was also dealt with by L. H. Koopmans [4] who, using deep techniques, has succeded in determining the error exponent for ergodic Markov-processes with continuous state-space under certain additional restrictions.

The inequalities above and the mentioned investigations show that the calculation of the α-divergence rate is of importance. In this paper we will show a simple way of doing this for Markov chains with finite state space, while we need not assume even the ergodicity of the chains.

We should like just mention that an investigation of the amount of information in an experiment also offers a motivation for this paper. Namely, there are estimations for the amount of the missing information (also called equivocation) similar to (1.2)—(1.4), see e.g. Vajda [11], Nemetz [7].

2. On the exponential convergence of the α-divergence

Let $\xi_0, \xi_1, \ldots$ be a Markov chain with finite state space, under both possible hypotheses. Without any loss of generality we can assume that ξ_i's can take on the values $0, 1, \ldots, r-1$ only. Let the initial probabilities be

$$\mathbf{P}_0 = \{p_0, p_1, \ldots, p_{r-1}\} \text{ and } \mathbf{Q}_0 = \{q_0, q_1, \ldots, q_{r-1}\} \tag{2.1}$$

under the hypothesis $\mathcal{H}_0 : \{\vartheta = \vartheta_0\}$ and $\mathcal{H}_1 : \{\vartheta = \vartheta_1\}$, respectively. Similarly let the matrix of transition probabilities be

$$\mathbf{P} = \begin{vmatrix} p_{0,0} & \cdots & p_{0,r-1} \\ \cdot & \cdots & \cdot \\ p_{r-1,0} & \cdots & p_{r-1,r-1} \end{vmatrix} \text{ and } \mathbf{Q} = \begin{vmatrix} q_{0,0} & \cdots & q_{0,r-1} \\ \cdot & \cdots & \cdot \\ q_{r-1,0} & \cdots & q_{r-1,r-1} \end{vmatrix}. \tag{2.2}$$

Our derivation is based on an enumeration of the possible values of the vector $\xi^n = \{\xi_0, \xi_1, \ldots, \xi_n\}$. We shall say that the vector $\varepsilon^n = \varepsilon^n(k) = \{\varepsilon_0, \varepsilon_1, \ldots, \varepsilon_n\}$ is the kth value of ξ_n, iff

$$k = \varepsilon_0 + \varepsilon_1 r + \ldots + \varepsilon_n r^n$$

where $0 \leq \varepsilon_i \leq r-1$, and $k = 0, 1, \ldots, r^{n+1} - 1$. It can be easily checked that

$$\varepsilon^n(k) = \{\varepsilon^{n-1}(k - ir^n), i\} \text{ if } i \cdot r^n \leq k < (i+1) r^n, \; i = 0, 1, \ldots, r-1.$$

Having this, the following equality is obvious:

$$P\{\xi^{n+1} = (\varepsilon^n(k), j)\} = P\{\xi^n = \varepsilon(k)\} \cdot p_{ij} \quad \text{if}$$
$$ir^n \leq k < (i+1) r^n, \quad i = 0, 1, \ldots, r-1.$$

Let us denote the probability that the sequence ξ^n takes on its kth value, by $P_{n,k}$ and $Q_{n,k}$, under the hypothesis $\mathcal{H}_0$ and $\mathcal{H}_1$ respectively, i.e.

$$P_{n,k} = P\{\xi^n = \varepsilon^n(k)\} \text{ and } Q_{n,k} = Q\{\xi^n = \varepsilon^n(k)\}.$$

Then we have:

$$\left.\begin{aligned} P_{n+1,k} &= P_{n,k-j\cdot r^{n+1}} \cdot p_{i,j} \\ Q_{n+1,k} &= Q_{n,k-j\cdot r^{n+1}} \cdot q_{i,j} \end{aligned}\right\} \text{ if } \begin{aligned} jr^{n+1} + ir^n \leq k < jr^{n+1} + (i+1) r^n, \\ i, j = 0, 1, \ldots, r-1. \end{aligned}$$

For the α-divergence of the probability distributions P_n and Q_n this yields the following:

$$H_{\alpha,n+1} \stackrel{\text{def}}{=} H_\alpha(P_{n+1}, Q_{n+1}) = \sum_{k=0}^{r^{n+2}-1} P^\alpha_{n+1,k}\, Q^{1-\alpha}_{n+1,k} =$$

$$= \sum_{j=0}^{r-1} \sum_{i=0}^{r-1} p^\alpha_{ij}\, q^{1-\alpha}_{ij} \sum_{k=jr^{n+1}+ir^n}^{j\cdot r^{n+1}+(i+1)r^n-1} P^\alpha_{n,k-jr^{n+1}} \cdot Q^{1-\alpha}_{n,k-jr^{n+1}}\,.$$

By introducing the notations

$$X^{(\alpha)}_{n,i} = \sum_{k=i\cdot r^n}^{(i+1)r^n-1} P^\alpha_{n,k} \cdot Q^{1-\alpha}_{n,k}$$

we get

$$H_{\alpha,n+1} = \sum_{i=0}^{r-1} X^{(\alpha)}_{n,i} \cdot \sum_{j=0}^{r-1} p^\alpha_{i,j}\, q^{1-\alpha}_{i,j} \stackrel{\text{def}}{=} \sum_{i=0}^{r-1} X^{(\alpha)}_{n,i} \sum A^{(\alpha)}_{j,i}\,.$$

In addition we have

$$X^{(\alpha)}_{n+1,i} = \sum_{j=0}^{r-1} X^{(\alpha)}_{n,j} A^{(\alpha)}_{j,i}$$

or, in matrix form:

$$\mathbf{X}^{(\alpha)}_{n+1} = \mathbf{X}^{(\alpha)}_n \cdot \mathbf{A}_\alpha$$

where

$$\mathbf{A}_\alpha = |A^{(\alpha)}_{i,j}| = \begin{vmatrix} p^\alpha_{0,0} \cdot q^{1-\alpha}_{0,0} & , \ldots , & p^\alpha_{0,r-1} \cdot q^{1-\alpha}_{0,r-1} \\ \cdot & , \ldots , & \cdot \\ p^\alpha_{r-1,0} \cdot q^{1-\alpha}_{r-1,0}, & \ldots , & p^\alpha_{r-1,r-1} \cdot q^{1-\alpha}_{r-1,r-1} \end{vmatrix}. \tag{2.3}$$

This recursion results in the following:

$$\mathbf{X}^{(\alpha)}_n = \mathbf{X}^{(\alpha)}_0 \cdot \mathbf{A}^n_\alpha\,.$$

Taking into account that

$$H_{\alpha,n} = \sum_{i=0}^{r-1} X^{(\alpha)}_{n,i}$$

we can restrict our investigation to the examination of the eigenvalues of the matrix $\mathbf{A}_\alpha$. In order to do this let us express the nth power of an arbitrary $r \times r$ matrix $\mathbf{M}$ in terms of its distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_m$:

$$\mathbf{M}^n = \sum_{k=1}^{m} \sum_{i=0}^{r_k-1} \frac{n!}{(n-i)!}\, \lambda_k^{n-i}\, \mathbf{Z}_{ki}\,,$$

where r_k denotes the multiplicity of the root λ_k of the minimal polynomial of $\mathbf{M}$ (and thus $r_k > 0$), and the matrices $\mathbf{Z}_{k,i}$ do not depend on n. In addition

the constituent matrices $\mathbf{Z}_{k,i}$ are linearly independent. (See [2], Chapter V, § 3.) Since all the elements of $\mathbf{A}_\alpha$ are non-negative (and therefore so are those of $\mathbf{A}_\alpha^n$) we may consider matrices $\mathbf{M}$ with non-negative elements only: It is well known, that such a matrix always has a non-negative eigenvalue ϱ such that the moduli of all the eigenvalues do not exceed ϱ. In addition all the eigenvalues with moduli ϱ are of the form

$$\lambda_k = \varrho \cdot e^{i \frac{2\pi s_k}{h_k}}.$$

(See [2], Chapter XIII, §§ 2—3.)

Let us order the eigenvalues of $\mathbf{M}$

$$\lambda_1 = \varrho, \lambda_2, \ldots, \lambda_q, \lambda_{q+1}, \ldots \lambda_m$$

such that $|\lambda_k| = \varrho$ if $1 \leq k \leq q$, and

$$|\lambda_k| < \varrho \quad \text{if} \quad q < k.$$

Let h be the least common multiple of h_k's, $k=1, 2, \ldots, q$ and $t + 1 = \max_{1 \leq k \leq q} r_k$. (We need not go into the details whether λ_k's are simple roots of the minimal polynomial or not.) We divide the summation in the fundamental formula (2.4) into four parts

$$\mathbf{M}^n = \sum_{\substack{k=1 \\ r_k=t+1}}^{q} \varrho^{n-t} \cdot \frac{n!}{(n-t)!} e^{i \frac{2\pi(n-t)\cdot s_k}{h_k}} \mathbf{Z}_{k,t} + \sum_{\substack{k=1 \\ r_k=t+1}}^{q} \sum_{l=1}^{t-1} \frac{n!}{(n-l)!} \lambda_k^{n-l} \mathbf{Z}_{k,l} +$$

$$+ \sum_{\substack{k=1 \\ r_k \leq t}}^{q} \sum_{l=1}^{r_k-1} \frac{n!}{(n-l)!} \lambda_k^{n-l} \mathbf{Z}_{k,l} + \sum_{k=q+1}^{m} \sum_{l=1}^{r_k+1} \frac{n!}{(n-l)} \lambda_k^{n-l} \mathbf{Z}_{k,l}. \tag{2.5}$$

Considering the subsequence $n = t + hj + g$; $j = 1, 2, \ldots,$ $0 \leq g < h$ fix, we observe that the numbers

$$\mu_k^g = e^{i \frac{2\pi(n-t)s_k}{h_k}}$$

do not depend on t. So in this case the first term of $\mathbf{M}^{hj+g+t}$ has the following form:

$$\mathbf{M}^{hj+g+t} = \varrho^{hj+g} \frac{n!}{(n-t)!} \cdot \sum_{\substack{k=1 \\ r_k=t+1}}^{q} \mu_k^g \mathbf{Z}_{k,l} + \ldots$$

where at least one of the coefficients μ_k differs from 0. Since the constituent matrices are linearly independent it follows from here immediately that the matrix

$$\mathbf{Z}_g = \sum_{(k)} \mu_k^g \mathbf{Z}_{k,t}$$

is not identically zero. Taking into account that $\mathbf{M}^n$ is a non-negative matrix a routine comparison of the magnitude of the additional elements in (2.5) shows that all elements of $\mathbf{Z}_g$ must be non-negative real numbers. This means that at least one of the scalar products of the columns (and those of the rows) of $\mathbf{Z}_g$ and an arbitrary positive vector $(x_1, x_2, \ldots, x_r)$ must be positive. This means that the numbers

$$\mu^{(g)} = [x_1, x_2, \ldots x_r] \left[\mathbf{Z}_g\right] \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix}, \qquad g = 0, 1, \ldots, h-1$$

are positive provided $x_k > 0$, $k = 1, 2, \ldots, r$. It is obvious from here that there exist numbers $0 < K_1 < K_2$ such that

$$K_1 \varrho^n \frac{n!}{(n-t)!} \leq [x_1, x_2, \ldots, x_r] \left[\mathbf{M}^n\right] \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix} \leq K_2 \varrho^n \frac{n!}{(n-t)!}$$

if n is large enough. And since $\sqrt[n]{K \frac{n!}{(n-t)!}} \to 1$ for every fixed $K > 0$ and t, we can deduce the following

Theorem. If the initial probabilities under both hypotheses form positive vectors (i.e. $p_i > 0$, $q_i > 0$, $i = 0, 1, \ldots, r-1$), then the α-divergence of the probability measures P_n and Q_n associated with the Markov chains given by (2.1) and (2.2) satisfies

$$\lim_{n \to +\infty} \sqrt[n]{H_\alpha(P_n, Q_n)} = \lambda_\alpha \tag{2.6}$$

where λ_α is the maximal eigenvalue of the matrix given in (2.3).

Remark. We note that there is no assumption made on the transition probabilities, and the assumption made for the initial probabilities is also very weak: any of them may be arbitrary small, provided $p_0 + p_1 + \ldots + p_{r-1} = 1$ and $q_0 + q_1 + \ldots + q_{r-1} = 1$. Of course, on the basis of our derivation one can formulate another version of this theorem, in which the positivity of the initial probability vectors may be dropped and can be restricted to a certain set of states depending on the transition probabilities. The reason, why our assumption cannot be dropped without introducing another one is, that

although $H_\alpha(P_n, Q_n)$ is a continuous function of the initial probabilities for every fixed n, the α-entropy rate $\lim_{(n)} \sqrt[n]{H_\alpha(P_n, Q_n)}$ may not be continuous at certain points, where $p_i = 0$ or $q_i = 0$. This is the case, for example, if

$$\mathbf{P} = \begin{bmatrix} 1/3 & 2/3 & 0 \\ 2/3 & 1/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \qquad \mathbf{Q} = \begin{bmatrix} 2/3 & 1/3 & 0 \\ 1/3 & 2/3 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

In this example it holds for the α-entropy rate:

$$\lim_{n\to\infty} \sqrt[n]{H_\alpha(P_n, Q_n)} = \begin{cases} 1, & \text{if } p_2 \cdot q_2 > 0 \\ 1/3\,(2^\alpha + 2^{1-\alpha}) & \text{if } p_2 \cdot q_2 = 0\,. \end{cases} \tag{2.7}$$

We mention also that the α-entropy rate, α given, is a continuous function of the initial probabilities at every point in the open set, defined by $p_i \cdot q_i > 0$, $i = 0, 1, \ldots, r-1$.

On account of (1.2) a simple consequence of our theorem is the following

Corollary. For a Markovian experiment with arbitrary initial probabilities and transition probabilities (2.2) the inequality

$$e(\Delta) \leq \frac{1}{2} \cdot \lambda_0^n$$

holds with

$$\lambda_0 = \inf_{0<\alpha<1} \lambda_\alpha .$$

3. Remarks

3.1. Let

$$H_\alpha(k) = \sum_{i=0}^{r-1} p_{k,i}^\alpha \cdot q_{k,i}^{1-\alpha} .$$

Because of a well-known theorem of Frobenius we have

$$\min_{(k)} H_\alpha(k) \leq \lambda_\alpha \leq \max_{(k)} H_\alpha(k) \tag{3.1}$$

and thus

$$0 \leq \lambda_0 \leq 1.$$

3.2. If $\lambda_0 = 1$, then, in general the probability of error does not go to zero owing to (1.9). This is conceivable because we did not assume ergodicity. This means that if the error probability tends to zero in the case of a Markovian experiment, then the convergence is exponential.

3.3. Recently A. Perez [8] obtained a lower and an upper error bound for Markovian experiments. He proved that there exist constants C_1 and C_2 such that

$$C_1 \lambda_1^n \leq e(\Delta) \leq C_2 \lambda_n^2$$

holds. Here

$$\lambda_2 = \max_{(k)} \inf_{(\alpha)} H_\alpha(k)$$

and

$$\lambda_1 = \inf_{0<\alpha<1} \sum_{k=0}^{r-1} H_\alpha(k) \cdot \beta_k ,$$

where $\{\beta_k\}$ is the stationary distribution. The inequality (3.1) tells us that

$$\lambda_0 \leq \lambda_2$$

always holds. Considering the matrices of transition probabilities

$$\mathbf{P} = \begin{vmatrix} 1/2 & 1/4 & 1/4 \\ 1/4 & 1/2 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{vmatrix}, \qquad \mathbf{Q} = \begin{vmatrix} 1/4 & 1/2 & 1/4 \\ 1/2 & 1/4 & 1/4 \\ 1/4 & 1/4 & 1/2 \end{vmatrix}$$

we have an example where

$$\lambda_1 < \lambda_0 < \lambda_2.$$

References

1. *Csiszár, I.*, Information-type measures of divergence of probability distributions and indirect observations, Studia Sci. Math. Hung. **2** (1967), 299–318.
2. *Gantmacher, F. R.*, The theory of matrices, vols I, II. Chelsea, New York, 1959.
3. *Hellman, M. E., Raviv, J.*, Probability of error, equivocation, and the Chernoff-bound, IEEE **IT-16,** *4* (1970), 368–372.
4. *Koopmans, L. H.*, Asymptotic rate of discrimination for Markov processes. Ann. Math. Stat. **31** (1960), 982–994.
5. *Korsh, J. P.*, An extension of Rényi's exponential bound for probability of error. Information and Control, to be published.
6. *Nemetz, T.*, Information theory and testing hypotheses. In: Proc. Coll. on Inf. Theory, Debrecen, 1967. Bolyai Society, Budapest, 1969, vol II, 283–294.
7. *Nemetz, T.*, Információelmélet és hipotézisvizsgálat (Information theory and hypothesis testing). Matematikai Lapok **20** (1969), 241–275. (In Hungarian.)
8. *Perez, A.*, Lecture, presented at the 2nd International Conference on Information Theory, Tsahkadsor, 1971.
9. *Rényi, A.*, Statistics and information theory. Studia Sci. Math. Hung. **2** (1967), 249–256.
10. *Vajda, I.*, Limit theorems for total variation of Cartesian Product Measures. Studia Sci. Math. Hung. **6** (1971), 317–333.
11. *Vajda, I.*, On the convergence of information contained in a sequence of observations. In: Proc. Coll. on Inf. Theory, Debrecen, 1967. Bolyai Society, Budapest, 1969, vol. II, 489–502.

Об асимптотическом исследовании α-расходимости мер, соответствующих двум однородным марковским процессам

Т. НЕМЕЦ

(Будапешт)

При использовании подхода Бейеса для оценки простых гипотез важно знать верхнее и нижнее значения вероятности ошибки. Для получения соответствующих оценок с помощью (1.1) используются интегралы Геллингера степени α. Из уравнений (1.2)—(1.4) непосредственно вытекает экспоненциальная сходимость вероятности ошибки при последовательности независимых наблюдений. Аналогичным свойством обладают егродические марковские процессы [4]. С помощью интегралов Геллингера получены асимптотические выражения для α-расходимости мер, соответствующих двум однородным конечным марковским процессам. Доказательство полученных выражений основано на специальной нумерации пространства состояния марковских цепей. При этом предполагается, что каждое начальное состояние имеет ненулевую вероятность по этим обеим гипотезам. Последовательность интегралов Геллингера стремится к нулю с показателем экспоненты $\lambda_0 < 1$, где $\lambda_0 = \inf_{0<\alpha<1} \lambda_\alpha$ является наибольшим собственным значением матрицы (2.3).

T. Nemetz
Mathematical Research Institute
Hungarian Academy of Sciences
1053 Budapest, V. Reáltanoda u. 13—15
Hungary

Problems of Control and Information Theory, Vol. 3 (2), pp. 157—166 (1974)

A FINITE PRIMAL METHOD FOR THE ASSIGNMENT PROBLEM

É. KOMÁROMI

(Budapest)

(Received November 13, 1973)

Two basic methods exist for solving the assignment problem together with (and by means of) its dual: the Hungarian method [2] which is of dual type and the simplex method [1] which is of primal type. From the point of view of computational efficiency, the difference between the two methods is not significant. However, the simplex algorithm has a disadvantage: its finiteness is not proved, even practical sample of its possible infiniteness has not yet been found. Variations of the simplex method ensuring finiteness have been developed but mostly at the expense of the computational efficiency. The present paper gives a finite primal method solving the problem which works with bases of special structure and also involves some simplifications in computation comparing to the simplex method.

1. The problem

The problem is to find $X = \{x_{ij}\}$, set of numbers which minimizes the

$$C(X) = \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij}\, x_{ij} \tag{1}$$

objective function, subject to

$$\sum_{j=1}^{n} x_{ij} = 1 \qquad i = 1, 2, \ldots, n;$$

$$\sum_{i=1}^{n} x_{ij} = 1 \qquad j = 1, 2, \ldots, n; \tag{2}$$

$$x_{ij} = 1 \text{ or } 0.$$

Here we assume that c_{ij} is integer for $i, j = 1, \ldots, n$.

The solution $X = \{x_{ij}\}$ satisfying the conditions (2) is called feasible solution while that minimizing the objective function (1), too, is called optimal solution.

For the assignment problem as a linear programming problem of special structure, the *optimality criterium* is as follows: A solution $X = \{x_{ij}\}$ is optimal if and only if such u_i, v_j potentials can be given to it ($i, j = 1, \ldots, n$) by which

the conditions

$$v_j - u_i \leqq c_{ij} \tag{3}$$

$$v_j - u_i = c_{ij} \qquad \text{if} \quad x_{ij} = 1$$

are satisfied.

2. Definitions

Let us assume that we have a feasible solution and for the sake of simplicity (and also without loss of generality) we suppose that it is of the following form:

$$X = \left\{ \begin{matrix} x_{ij} = 1 & \text{if} & i = j \\ & & \\ x_{ij} = 0 & \text{if} & i \neq j \end{matrix} \quad i, j = 1, 2, \ldots, n \right\}.$$

Now we introduce the following terms: *Chain* is such an ordered set of pairs (i, j) of indices

$$\mu_{i_1 j_s} = \{(i_1, j_1), (i_2, j_1), (i_2, j_2), \ldots, (i_s, j_{s-1}), (i_s, j_s)\}$$

for which

$$\begin{array}{lll} i_k \leqq n, & j_k \leqq n, & k = 1, \ldots, s; \\ i_k = j_k, & \text{i.g. } x_{i_k j_k} = 1, & k = 1, \ldots, s; \\ i_k \neq i_r, & r = 1, \ldots, k-1, & k = 1, \ldots, s. \end{array}$$

Loop is a chain $\mu_{i_1 j_s}$ completed with (i_1, j_s).

Feasible basis B_n is called here a set of pairs (i, j) of indices $i \in \{1, 2, \ldots, n\}$, $j \in \{1, 2, \ldots, n\}$ which

(a) contains all (i, i), that is: $(i, j) \in B_n$ if $x_{ij} = 1$;
(b) contains for every index $1 < i \leqq n$ exactly one (i, j) for which $x_{ij} = 0$ (i.g. $j \neq i$);
(c) does not contain any loop.

Let us note that a feasible basis B_n is a tree with common roote $(1, 1)$; this means that

(d) for each index $1 \leqq j \leqq n$ B_n contains one (and by definition only one) chain

$$\overline{\mu}_{1j} = \{(1, 1) = (i_1, j_1), (i_2, j_1), \ldots, (i_{r+1}, j_r), \ldots, (i_s, j_{s-1}), (i_s, j_s)\}$$

where $j_s = j$;

(e) fixing the value u_1, we can determine potentials in a unique way by means of the above chains $\overline{\mu}_{1j}$ as follows:

$$v_{j_k} = u_{i_k} + c_{i_k j_k} \qquad (i_k j_k) \in \mu_{1 j_k}$$

$$u_{i_{k+1}} = v_{j_k} - c_{i_{k+1} j_k} \qquad (i_{k+1}, j_k) \in \mu_{1 j_{k+1}}$$

(f) for these potentials it also holds via the chain $\overline{\mu}_{1j}$ that

$$v_j = u_{i_r} + c_{i_r j_r} - c_{i_{r+1} j_r} + \ldots + c_{i_s j_s} \quad \text{for} \quad 1 \leq r \leq s, j_s = j.$$

A feasible basis is called *optimal basis* if the potentials determined to it satisfy conditions (3).

3. Example 1

Let us consider the problem with the following matrix $C = \{c_{ij}\}$ and with the following feasible basis:

	v_j	1	0	-2	0	1
	j	1	2	3	4	5
u_i	i					
0	1	$\boxed{1}$	7	3	4	2
-1	2	2	$\boxed{1}$	1	1	2
-3	3	4	5	$\boxed{1}$	5	1
-1	4	3	2	1	$\boxed{1}$	2
0	5	1	2	3	2	$\boxed{1}$

Here the elements of B_5 are denoted by a square where $x_{ij} = 1$ and by a circle where $x_{ij} = 0$. Fixing $u_1 = 0$, the potentials for B_5 are calculated via the chains:

$$\{(1, 1), (2, 1), (2, 2)\}$$

$$\{(1, 1), (3, 1), (3, 3)\}$$

$$\{(1, 1), (5, 1), (5, 5), (4, 5), (4, 4)\}.$$

Obviously the solution $x_{ii} = 1$; $x_{ij} = 0$ if $i \neq j$; $i = 1, \ldots, 5, j = 1, \ldots, 5$; is optimal while the above basis B_5 is not optimal. However, an optimal basis can be also constructed, for instance that consisting of the chain

$$\{(1, 1), (5, 1), (5, 5), (3, 5), (3, 3), (4, 3), (2, 4), (4, 2), (2, 2)\}$$

for which the potentials: $u_1 = 0$, $v_j = 1$; $i, j = 1, \ldots, 5$.

4. The algorithm

The main idea of our method is that we try to build up an optimal basis and thereby to prove that the solution is optimal — while if we did not succeed, we can determine a new feasible solution for which the value of the objective function decreases. While constructing B_m we assume that we have constructed an optimal basis B_{m-1} for the problem which includes the first $m-1$ rows and columns of our original problem.

Step 0: Initially, set $m = 2$ and create

$$B_1 = \{(1, 1)\},\ u_1 = 0,\ v_1 = c_{11}.$$

Here the basis B_1 satisfies all the conditions required.

Basing on B_{m-1}, B_m is built up as follows:

Step 1 (linking the mth row and column): Let be

$$u_m = \max_{j<m} (v_j - c_{mj}) = v_{j_0} - c_{mj_0}$$

$$v_m = u_m + c_{mm}$$

$$B_m = B_{m-1} \cup \{(m, m), (m, j_0)\}.$$

Introduce the set $V_m^0 = \{m\}$ and also give $k = 1$.

Step 2 (checking the basis and the solution): Compute

$$\varepsilon^k = \max_{i,j\leq m} (v_j - u_i - c_{ij}) = v_{j^k} - u_{i^k} - c_{i^k j^k}.$$

If $\varepsilon^k = 0$ then B_m is optimal, the procedure of determining B_m is completed. If $m = n$, the procedure is finished. If $m < n$, increase the value m by 1 and continue at Step 1. If $\varepsilon^k > 0$, two cases are possible:

— $(i^k, i^k) \notin \overline{\mu}_{1m}$; in this case perform Step 3.

— $(i^k, i^k) \in \overline{\mu}_{1m}$; in this case perform Step 4.

(As a consequence of Lemmas 1 and 2 below, we here exploit the fact that $(i^k, i^k) \in \overline{\mu}_{1m}$ if and only if $(i^k, i^k) \in \overline{\mu}_{1j^k}$.)

Step 3 (change of basis): By denoting $J(i^k)$ all the subscripts j for which $(i^k, i^k) \in \overline{\mu}_{1j^k}$ (including $j^k = i^k$ itself) let

$$\left.\begin{aligned} u_j &= u_j + \varepsilon^k \\ v_j &= v_j + \varepsilon^k \end{aligned}\right. \quad \text{for all} \quad j \in J(i^k).$$

Replace the sole element (i^k, j) of B_m, for which $x_{i^k j} = 0$, by (i^k, j^k), that is, let

$$B_m = B_m \setminus \{(i^k, j)\} \cup \{(i^k, j^k)\}$$

as well as let $V_m^k = V_m^{k-1} \cup \{J(i^k)\}$.

Increase the value k by 1 and repeat Step 2.

Step 4 (loop searching and improving the solution): Consider the unique chain

$$\mu_{i^k j^k} = \{(i^k, i^k) = (i_1, j_1), (i_2, j_1), \ldots, (m, m), \ldots, (i_s, j_{s-1}), (i_s, j_s) = j^k, j^k)\}$$

where

$$(i_r, j_r) \in B_m \qquad \text{for } r = 1, \ldots, s;$$

$$(i_{r+1}, j_r) \in B_m \qquad \text{for } r = 1, \ldots, s-1.$$

Modify the solution as follows. Let

$$x_{i_r j_r} = 0 \qquad \text{for } r = 1, \ldots, s;$$

$$x_{i_{r+1} j_r} = 1 \qquad \text{for } r = 1, \ldots, s-1;$$

$$x_{i^k j^k} = 1$$

x_{ij} unchanged otherwise.

Rearrange the subscripts according to the new solution to satisfy the assumption made for X. Repeat the procedure, starting with Step 0.

Let us note that in all stages of the procedure, B_m is a feasible solution.

5. Verification

Lemma 1. In the course of constructing B_m, the case $\varepsilon^k > 0$ can occur only for such a subscript j^k for which $j^k \in V_m^{k-1}$.

Proof. Assume that, contrary to the statement, $j^k \notin V_m^{k-1}$. This means that the value of v_{j^k} is unchanged compared with its value defined to the basis B_{m-1}. But then

$$\varepsilon_i = v_{j^k} - u_i - c_{ij^k}$$

(a) is non-positive in case of such $i < m$ for which the value of u_i is unchanged compared with its value defined to B_{m-1};

(b) is negative in case of such $i < m$ for which the value of u_i has changed, that is, has increased by a positive number compared with its value defined to B_{m-1};

(c) is non-positive if $i = m$ and the value of u_m is that one defined in Step 1;

(d) is negative, if $i = m$ and the value of u_m has changed, that is, has increased by a positive number compared with its initial value.

This contradicts the assumption that

$$\varepsilon^k = \max_{i \leq m} \varepsilon_i > 0.$$

Lemma 2. In the course of constructing B_m, $\varepsilon^k > 0$ can occur for such a subscript i^k only for which $i^k \notin V_m^{k-1} \setminus \{m\}$.

Proof. Assume that, contrary to the statement, $i^r \notin V_m^{r-1} \setminus \{m\}$ for $r = 1, 2, \ldots, k-1$ but $i^k \in V_m^{k-1} \setminus \{m\}$. We can make the following statements:

(a)
$$\varepsilon^1 \geq \varepsilon^2 \geq \ldots \geq \varepsilon^{k-1}.$$

This follows from the way of defining ε^r and from the assumption made for B_{m-1}.

(b) i^k could not become the element of $V_m^{k-1} \setminus \{m\}$ earlier than j^k. (Here the statement involves that $j^k = m$ is impossible.) Otherwise, supposing that i^k and j^k became elements of $V_m^{k-1} \setminus \{m\}$ at the pth and rth iterations ($p < r$) respectively, the values of their potentials defined to B_{m-1}

$$\hat{u}_{i^k} = u_{i^k} - \varepsilon^p$$

$$\hat{v}_{j^k} = v_{j^k} - \varepsilon^r.$$

Then

$$v'_{j^k} - \hat{u}_{i^k} = v_{j^k} - u_{i^k} - \varepsilon^r + \varepsilon^p > c_{i^k j^k}.$$

It contradicts the assumptions made for B_{m-1}.

(c) nor could i^k become an element of $V_m^{k-1} \setminus \{m\}$ later than j^k became element of V_m^{k-1}. Otherwise, supposing that i^k became element of $V_m^{k-1} \setminus \{m\}$ and j^k became element of V_m^{k-1} at the rth and pth iterations ($p < r$) respectively, and denoting the potentials defined to B_{m-1} by $\hat{u}_i$, $\hat{v}_j$:

$$\varepsilon^r = v_{j^r} - u_{i^r} - c_{i^r j^r} \geq v_{j^k} - \hat{u}_{i^k} - c_{i^k j^k}$$

$$\varepsilon^k = v_{j^k} - u_{i^k} - c_{i^k j^k} = v_{j^k} - \hat{u}_{i^k} - c_{i^k j^k} \leq 0$$

in contradiction with the assumption $\varepsilon^k > 0$.

The statement of Lemma 2 follows therefore from Lemma 1 and from the above statements.

Lemma 3. If in the course of constructing B_m $\varepsilon^k > 0$ and $(i^k, i^k) \in \overline{\mu}_{1m}$, then X is not an optimal solution.

Proof. Assume that $(i^k, i^k) \in \overline{\mu}_{1m}$. Lemmas 1 and 2 imply that $(m, m) \in \overline{\mu}_{1j^k}$. From the fact that B_m satisfies condition (c) it also holds: $(i^k, i^k) \in \overline{\mu}_{ij^k}$. Let us consider the chain

$$\mu_{i^k j^k} = \{(i^k, i^k) = (i_1, j_1), (i_2, j_1), \ldots, (m, m), \ldots, (i_s, j_{s-1}), (i_s, j_s) = (j^k, j^k)\}$$

which is a part of $\overline{\mu}_{1j^k}$ and for which

$$v_{j^k} = u_{i^k} + c_{i_1 j_1} - c_{i_2 j_1} + \ldots + c_{mm} - \ldots + c_{i_s j_s}.$$

If we modify the solution X as we did at Step 4:

$$\begin{aligned} \overline{x}_{i_r j_r} &= 0 & & r = 1, \ldots, s, \\ \overline{x}_{i_{r+1} j_r} &= 1 & & r = 1, \ldots, s-1, \\ \overline{x}_{i^k j^k} &= 1 & & \\ \overline{x}_{ij} &= x_{ij} & & \text{otherwise} \end{aligned}$$

then the difference between the objective function values:

$$\begin{aligned} \delta &= \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} x_{ij} - \sum_{i=1}^{n} \sum_{j=1}^{n} c_{ij} \overline{x}_{ij} = \\ &= \sum_{(i,j) \in \mu_{i^k j^k}} c_{ij}(x_{ij} - \overline{x}_{ij}) + c_{i^k j^k}(x_{i^k j^k} - \overline{x}_{i^k j^k}) = \\ &= c_{i_1 j_1} - c_{i_2 j_1} + \ldots + c_{mm} - \ldots + c_{i_s j_s} - c_{i^k j^k} = \\ &= v_{j^k} - u_{i^k} - c_{i^k j^k} = \varepsilon^k > 0. \end{aligned}$$

Consequence 1. We can improve the computation in Step 2 by replacing the definition of ε^k by the following one:

$$\varepsilon^k = \max_{i \notin V_m^{k-1} \setminus \{m\};\, j \in V_m^{k-1}} (v_j - u_i - c_{ij}) = v_{j^k} - u_{i^k} - c_{i^k j^k}.$$

Consequence 2. Constructing the basis B_m, having had already B_{m-1}, requires at most $m - 2$ iterations.

Consequence 3. All corrections of the solution result in the decrease of the objective function and each correction is preceded by at most

$$\sum_{m=3}^{n} m - 2 = \binom{n-1}{2}$$

empty changes of basis (which do not modify the solution).

Due to the special structure of the basis, both an empty change of basis and an improvement of the solution, requires a number of operations in the order of n.

Here we remark that the idea of the method can be easily applied for the transportation problem, too. In that case, however, an improvement of the solution requires a number of operation in the order of n^2.

6. Example 2

Let us apply the algorithm to solve the problem we dealt with in Example 1. Its matrix, the basis which is denoted by squares and circles, respectively, and the potentials as they will be calculated in the course of the procedure are tabled below:

u_i \ v_j			1	1	1		
			1	0	-2	-1	1
		j / i	1	2	3	4	5
	0	1	$\boxed{1}$	7	3	4	2
0	-1	2	$\cancel{2}$	$\boxed{1}$	1	1	2
0	-3	3	$\cancel{4}$	5	$\boxed{1}$	5	1
0	-2	4	3	$\cancel{2}$	1	$\boxed{1}$	2
	0	5	1	2	3	2	$\boxed{1}$

The initial solution: $x_{ii} = 1$; $x_{ij} = 0$ if $i \neq j$.

Step 0: Set $m = 1$

$$B_1 = \{(1, 1)\};\ u_1 = 0;\ v_1 = u_1 + c_{11} = 1.$$

Step 1: $u_2 = v_1 - c_{21} = -1$; $v_2 = u_2 + c_{22} = 0$;

$$B_2 = \{(1, 1)\} \cup \{(2, 1), (2, 2)\}$$

$$V_2^0 = \{2\}; \qquad \text{set } k = 1.$$

Step 2: $\varepsilon^1 = \max\limits_{i \leq 2} (v_2 - u_i - c_{i2}) = 0$; set $m = 3$.

Step 1: $u_3 = \max\limits_{j<3} (v_j - c_{3j}) = v_1 - c_{31} = -3$; $v_3 = -2$;

$$B_3 = \{(1, 1), (2, 1), (2, 2)\} \cup \{(3, 1), (3,3)\}$$

$$V_3^0 = \{3\}; \text{ set } k = 1.$$

Step 2: $\varepsilon^1 = \max\limits_{j \in V_3^0;\, i \leq 3} (v_j - u_i - c_{ij}) = 0$; set $m = 4$.

Step 1: $u_4 = \max\limits_{j<4} (v_j - c_{4j}) = v_2 - c_{42} = -2;\ v_4 = -1;$

$B_4 = \{(1, 1), (2, 1), (2, 2), (3, 1), (3, 3)\} \cup \}(4, 2), (4, 4)\}$

$V_4^0 = \{4\};$ set $k = 1.$

Step 2: $\varepsilon^1 = \max\limits_{j \in V\ ;\ i \leq} (v_j - u_i - c_{ij}) = 0;$ set $m = 5.$

Step 1: $u_5 = \max\limits_{j<5} (v_j - c_{5j}) = v_1 - c_{51} = 0;$ $v_5 = 1;$

$B_5 = \{(1, 1), (2, 1), (2, 2), (3, 1), (3, 3), (4, 2), (4, 4)\} \cup \{(5, 1), (5, 5)$

$V_5^0 = \{5\};$ set $k = 1.$

Step 2: $\varepsilon^1 = \max\limits_{j \in V_5^0;\ i \leq 5} (v_j - u_i - c_{ij}) = v_5 - u_3 - c_{35} = 3$

$(3,3) \notin \overline{\mu}_{15} = \{(1, 1), (5, 1), (5, 5)\}.$

Step 3: $u_3 = -3 + 3 = 0;\ v_3 = -2 + 3 = 1;$

$B_5 = \{(1, 1), (2, 1), (2, 2), (3, 5), (3, 3), (4, 2), (4, 4), (5, 1), (5, 5)\}$

$V_5^1 = \{5, 3\};$ set $k = 2.$

Step 2: $\varepsilon^2 = \max\limits_{j \in V_5^1;\ i \in \{1, 2, 4, 5\}} (v_j - u_i - c_{ij}) = v_3 - u_4 - c_{43} = 2;$

$(4, 4) \notin \overline{\mu}_{15}.$

Step 3: $u_4 = -2 + 2 = 0;\quad v_4 = -1 + 2 = 1;$

$B_5 = \{(1, 1), (2, 1), (2, 2), (3, 5), (3, 3), (4, 3), (4, 4), (5, 1), (5, 5)\}$

$V_5^2 = \{5, 3, 4\};$ set $k = 3.$

Step 2: $\varepsilon^3 = \max\limits_{j \in V^2;\ i \in \{1, 2, 5\}} (v_j - u_i - c_{ij}) = v_4 - u_2 - c_{24} = 1$

$(2, 2) \notin \overline{\mu}_{15}.$

Step 3: $u_2 = -1 + 1 = 0;\quad v_2 = 0 + 1 = 1;$

$B_5 = \{(1, 1), (2, 4), (2, 2), (3, 5), (3, 3), (4, 3), (4, 4), (5, 1), (5, 5)\}$

$V_5^3 = \{5, 3, 4, 2\};$ set $k = 4.$

Step 2: $\varepsilon^4 = \max\limits_{j \in V_5^3;\ i \in \{1, 5\}} (v_j - u_i - c_{ij}) = 0.$

The procedure is finished, B_5 is an optimal basis.

References

1. *Dantzig, G. B.*, Linear Programming and Extensions, Princeton University Press, 1963.
2. *Kuhn, H. W.*, The Hungarian Method for the Assignment Problem, Naval Research Logistics Quarterly, 1955.

Конечный примальный метод для решения транспортной задачи

Е. КОМАРОМИ

(Будапешт)

Предложен новый метод, в котором применяются базисы специальной структуры для нахождения оптимального базиса. Под допустимым базисом B_n, принадлежащего к допустимому решению $X = \{x_{ij}\}$, подразумевается такое множество пар индексов (i, j) $(i \in \{1, 2, \ldots, n\}, j \in \{1, 2, \ldots, n\})$, для которых действительны следующие условия:

а) $(i, j) \in B_n$, если $x_{ij} = 1$,

б) базис B_n содержит именно одну такую пару индексов (i, j) для каждого индекса $1 < i \leq n$, при которой $x_{ij} = 0$,

в) в базисе нет петли.

Заметим, что определенный таким образом базис имеет еще и следующие свойства:

г) базис B_n содержит именно одну цепь типа

$$\overline{\mu}_{1j} = \{(1, j_1), (i_2, j_1), (i_2 j_2), \ldots, (i_s, j_{s-1}), (i_s, j_s) = (i_s, j)\}$$

для всех индексов $1 \leq j \leq n$,

д) если зафиксировать значение u_1, то значения потенциалов u_i, v_j $(i, j = 1, \ldots, n)$ могут быть определены однозначно по цепям $\overline{\mu}_{1j}$, следующим образом:

$$v_{j_k} = u_{i_k} + c_{i_k j_k}; \qquad k = 1, 2, \ldots, s; \qquad i_1 = 1;$$

$$u_{i_{k+1}} = v_{j_k} - c_{i_{k+1} j_k}; \qquad k = 1, 2, \ldots, s - 1.$$

Базис является оптимальным, если выполняются условия $v_j - u_i \leq c_{ij}$; $v_j - u_i = c_{ij}$ если $x_{ij} = 1$.

При описании предложенного алгоритма (без нарушения общности) предполагается, что начальное решение задачи имеет следующий вид:

$$X = \begin{cases} x_{ij} = 1 & \text{если } i = j \\ x_{ij} = 0 & \text{если } i \neq j \end{cases} \qquad i, j = 1, 2, \ldots, n.$$

Основная идея алгоритма заключается в том, что базис B_m $(m \leq n)$ составляется на основе применения оптимального базиса B_{m-1} такой частной задачи, которая содержит первые $(m - 1)$ строк и столбцов исходной задачи.

É. Komáromi
Cumputer and Automation Institute
Hungarian Academy of Sciences
1502 Budapest, XI., Kende u. 13–17.
Hungary

Problems of Control and Information Theory, Vol. 3 (2), pp. 167–170 (1974)

ON THE ENTROPY OF THE SUM OF IID RANDOM VARIABLES

J. BECK

(Budapest)

(Received November 13, 1973)

If $(\xi_n)_{n=1}^{\infty}$ is any sequence of real valued random variables such that the range of each ξ_n is the same finite set X, then it can be easily shown that $H(\xi_1 + \xi_2 + \ldots + \xi_n) < c \log n$, where c is independent of n.

If the range X of the ξ_n's is infinite, no similar bound is true, in general, except for the trivial one

$$H(\xi_1 + \xi_2 + \ldots + \xi_n) \leqq H(\xi_1, \xi_2, \ldots, \xi_n). \tag{1}$$

A. Rényi has asked, what can be asserted on the asymptotic behaviour of $H(\xi_1 + \xi_2 + \ldots + \xi_n)$ if the ξ_i's are IID (independent and identically distributed) with $H(\xi_1) < +\infty$ and X is countable.

In this note we prove the following

Proposition. For any sequence of positive numbers $\varepsilon(n) \to 0$ there exists a sequence of IID random variables with $H(\xi_1) < +\infty$ and

$$\overline{\lim_n} \frac{1}{\varepsilon(n)n} H(\xi_1 + \xi_2 + \ldots + \xi_n) > 0\,.$$

Proof. Without any loss of generality we may assume that $\varepsilon(n) \geqq \frac{1}{n}$. Let $\{R_n\}_{n=1}^{\infty}$ be a sequence of natural numbers, for which $R_n \geqq 2$ and $\varepsilon(R_n) \leqq \frac{1}{n^2}$ hold, i.e.

$$\frac{1}{R_n} \leqq \varepsilon(R_n) \leqq \frac{1}{n^2}. \tag{2}$$

Since

$$0 < \sum_{n=1}^{\infty} \varepsilon(R_n) \leqq \sum_{n=1}^{\infty} \frac{1}{n^2} < +\infty$$

for a proper choice of c

$$h_n = \frac{c\,\varepsilon(R_n)}{R_n \log R_n} \tag{3}$$

is a probability distribution, i.e. $\sum_{n=1}^{\infty} h_n = 1$. Let $\mathcal{S} = \{p_1, p_2, \ldots\}$ be the distribution defined by

$$p_m = \frac{h_n}{R_n^{2R_n}}, \text{ if } R_{n-1}^* = \sum_{i=1}^{n-1} R_i^{2Ri} < m \leq \sum_{i=1}^{n} R_i^{2R_i} = R_n^*. \tag{4}$$

Then, using (2), (3) and (4) we obtain

$$H(\mathcal{S}) = -\sum_{m=1}^{\infty} p_m \log p_m = -\sum_{n=1}^{\infty} \left(\sum_{R_{n-1}^* < m \leq R^*} p_m \log p_m \right) =$$

$$= -\sum_{n=1}^{\infty} h_n \log \frac{h_n}{R_n^{2R_n}} = \sum_{n=1}^{\infty} \frac{c\,\varepsilon(R_n)}{R_n \log R_n} \left(\log \frac{R_n \log R_n}{c\,\varepsilon(R_n)} + \log R_n^{2R_n} \right) =$$

$$= \sum_{n=1}^{\infty} \frac{c\,\varepsilon(R_n)}{R_n \log R_n} \left(O(\log R_n) + 2\,R_n \log R_n \right) < +\infty .$$

Let the distribution of the ξ_i's as defined by $P\{\xi_i = a_m\} = p_m$ where $\{a_m\}_{m=1}^{\infty}$ are linearly independent over the rational numbers; e.g., one can choose $a_m = \log q_m$ where q_m is the mth prime number, because then $\sum_{i=1}^{n} c_i a_i = 0$ with integer c_i's and $\sum_{i=1}^{n} c_i^2 \neq 0$ would imply $\prod_{i=1}^{n} q_i^{c_i} = 1$ contradicting to the unique factorization theorem. If $\xi_k = a_{i_K}$ $K = 1, 2, \ldots, n$ we have

$$\zeta_n = \sum_{K=1}^{n} \xi_K = \sum_{K=1}^{n} a_{i_K} \overset{\text{def}}{=} b .$$

Observe that $\sum_{K=1}^{n} a_{i_K} = \sum_{K=1}^{n} a_{j_K} = b$ holds if and only if there exists a permutation π of the first n natural numbers for which

$$a_{i_{\pi(K)}} = a_{j_K} \qquad K = 1, 2, \ldots, n .$$

This is a trivial corollary of the linear independence of the a_i's. Thus the number of representations of any b in the form $\sum_{K=1}^{n} a_{i_K}$ is not greater than $n!$, hence

$$P\{\zeta_n = b\} \leq n! \prod_{K=1}^{n} P\{\xi_k = a_{i_K}\} .$$

We apply the previous arguments choosing $n = R_n$. Let B_n be the set of those numbers b which can be represented as $\sum_{K=1}^{K_n} a_{i_K} = b$ where at least one

subscript is satisfies $R^*_{n-1} < i_K \leq R^*_n$. Then

$$P\{\zeta_{R_n} = b\} \leq R_n! \prod_{K: R^*_{n-1} < i_K \leq R^*} P\{\xi_K = a_{i_K}\} \leq R_n! \frac{h_n}{R_n^{2R_n}} \leq \frac{1}{R_n^{R_n}}. \tag{5}$$

$$(b \in B_n).$$

Further,

$$P\{\zeta_{R_n} \in B_n\} = 1 - P\{\zeta_{R_n} \notin B_n\} = 1 - \Big(1 - \sum_{R^*_{n-1} < m \leq R^*_n} p_m\Big)^{R_n} = 1 - (1 - h_n)^{R_n}.$$

We want to estimate this quantity. For $X > 1$ we have $\left(1 - \frac{1}{x}\right)^x < e^{-1} <$ $< \left(1 - \frac{1}{x}\right)^{x-1}$ thus $e^{-1} = \left(1 - \frac{1}{x}\right)^{x-\vartheta_x}$ where $0 < \vartheta_x < 1$. With $X = \frac{R_n \log R_n}{c\varepsilon(R_n)}$ and $\overline{X} = \frac{c\varepsilon(R_n)}{R_n \log R_n}$ we obtain (see (3))

$$(1 - h_n)^{R_n} = \left(1 - \frac{1}{x}\right)^{X\overline{X}} = e^{-\frac{x}{x-\vartheta_x}\overline{X}},$$

Since $\lim_{n\to\infty} \frac{x}{x - \vartheta_x} = 1$ and $\lim_{n\to\infty} \overline{x} = 0$ we hawe $c^* \overline{x} < c^* \frac{x}{x - \vartheta_x} \overline{x} < 1 -$ $- e^{-\frac{x}{x-\vartheta_x}\overline{x}}$, where $c^* > 0$ if $n \geq n_0$ yielding $c^* \frac{c\varepsilon(R_n)}{\log R_n} < 1 - (1 - h_n)^{R_n}$, i.e.

$$c_1 \frac{\varepsilon(R_n)}{\log R_n} < P\{\zeta_{R_n} \in B_n\} \qquad \text{if} \quad n \geq n_0, \tag{6}$$

where $c_1 = c^* c > 0$. (5) and (6) give rise to

$$H(\xi_1 + \xi_2 + \ldots + \xi_{R_n}) = H(\zeta_{R_n}) \geq$$

$$\geq - \sum_{b \in B_n} P\{\zeta_{R_n} = b\} \log P\{\zeta_{R_n} = b\} \geq P\{\zeta_{R_n} \in B_n\} R_n \log R_n \geq$$

$$\geq c_1 \frac{\varepsilon(R_n)}{\log R_n} R_n \log R_n = c_1 \varepsilon(R_n) R_n \qquad \text{if } n \geq n_0,$$

completing the proof.

Remark. It remains open whether $\overline{\lim_n} \frac{1}{n} H(\xi_1 + \xi_2 + \ldots + \xi_n) > 0$ can be achieved, too.

Об энтропии одинаково распределенных независимых случайных величин

Й. БЕК

(Будапешт)

Доказывается, что для произвольной прогрессии положительных чисел $\varepsilon(n) \to 0$ существует последовательность дискретных, независимых случайных величин $(\xi_n)_{n=1}^{\infty}$, имеющих одинаковые распределения, при этом выполняются условия $H(\xi_1) < +\infty$ и

$$\overline{\lim_n} \frac{H(\xi_1 + \xi_2 + \ldots + \xi_n)}{\varepsilon(n) \cdot n} > 0.$$

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

Printed in Hungary
AKADÉMIAI NYOMDA, BUDAPEST

PROBLEMS OF CONTROL AND INFORMATION THEORY VOL. 3, NO. 2
ENGLISH TRANSLATION OF THE COMMUNICATIONS IN RUSSIAN

APPLICATION OF VOLTERRA SERIES FOR THE INVESTIGATION OF NONLINEAR SYSTEMS

K. A. PUPKOV, V. I. KAPALIN

(Moscow)

In the paper we examine the questions of the application of Volterra series and the theory of multivariable integral transformations in the tasks of the analysis and synthesis of nonlinear systems. The algorithm which determines the transforms of the kernels of the Volterra series is introduced for the case of a given structural scheme. The methods for the determination of the output signal of the system are examined and a nonlinear compensation task is solved.

Introduction

In the present paper the methods of analysis and synthesis of nonlinear systems based on the theory of multi-variable integral transformations and on the application of Volterra series are examined. With these methods the analytic solution of the nonlinear problem can be found. An algorithm is proposed which defines the transforms of the Volterra series kernels for systems with a given structure. The methods for determining the response of the nonlinear system to different regular signals are shown. Finally, a compensation problem of analytic systems is examined, supposing that the required system is linear. To determine the kernels of the compensator in this case, we propose a simpler method than that elaborated in [1].

1. Volterra series

The function $y(t) = F[x(\tau), \tau \leq t]$, describing the connection between the input signal $x(t)$ exerted on the stationary nonlinear system and the output response of the system, which satisfies the condition $F[0] \equiv 0$ for all values of t, is called analytic if it can be produced as a Volterra series with a convergence cycle of nonzero radius ([13]):

$$y(t) = F[x(t), \tau \leq t] = \sum_{n=1}^{\infty} v_n[x(\tau), \tau \leq t], \tag{1.1}$$

where $v_n[x(\tau), \tau \leq t]$, $n = 1, 2 \ldots$ is a regular, homogeneous, nth order Volterra function

$$y(t) = v_n[x(\tau), \tau \leq t] = \int_{E_n} h_n(\tau_1, \ldots, \tau_n) \prod_{r=1}^{n} x(t - \tau_r) dv_r.$$

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

Here the kernel $h_n(\tau_1, \ldots, \tau_n)$ becomes zero when one of the arguments τ_j, $j = 1, 2, \ldots, n$ is smaller than zero; the range E_n can be determined in the following way:

$$-\infty < \tau_j; < \infty, \; j = 1, 2, \ldots, n; \; dv_\tau = d\tau_1 \ldots d\tau_n.$$

In works [2]—[5] the determination method for the kernels of series (1.1) has been elaborated for systems described by nonlinear differential equations. In the present work the case is examined in which the nonlinear system resulting from the combination of stationary, linear inert members with stationary, inertialess nonlinearities is given by its structural scheme. In order to be able to use the Volterra series for the calculation of such a system it is necessary that all inertialess nonlinearities could be described by either polynomials or power series. The convolution integral describing the linear system is a regular, homogeneous, first order Volterra series. The power series

$$y = \sum_{n=1}^{\infty} a_n x^n \tag{1.3}$$

can also be examined as a degenerated (1.1) type Volterra series with kernels $h_n(\tau_1, \ldots, \tau_n) = a_n \prod_{r=1}^{n} \delta(\tau_r)$, $n = 1,2$, where $\delta(\tau_r)$ is a δ function. Thus the task of determining the Volterra series kernels for a nonlinear system with a given structure can be formulated in the general case as follows.

A system is given which consists of a sequence of subsystems whose Volterra series description is known. The kernels of the Volterra series describing the entire system have to be determined. The formulated task can be solved for stationary nonlinear systems most simply by a multi-variable Laplace transformation.

2. Multi-variable Laplace transformation

The following function is called the Laplace transform of the function $f(t_1, \ldots, t_n)$ depending on n arguments in total:

$$F(s_1, \ldots, s_n) = \int_{E_n^+} \exp\left(-\sum_{i=1}^{n} s_i t_i\right) f(t_1, \ldots, t_n) dv_t, \tag{2.1}$$

where the range E_n^+ can be determined in the following way: $0 \leq t_i \leq \infty$, $i = 1, 2, \ldots, n$; s_i $i = 1, 2 \ldots n$ is a complex variable; $dv_t = dt_1 \ldots dt_n$.

If $f(t_1, \ldots, t_n) = 0$ is valid for $t_i < 0$ $(i = 1, 2, \ldots, n)$, then the inverse Laplace transform exists, and can be obtained from the following formula:

$$f(t_1, \ldots, t_n) = \left(\frac{1}{2\pi j}\right)^n \int\limits_{Br_n} F(s_1, \ldots, s_n) \tag{2.2}$$

where the range Br_n, can be determined in the following way: $C_i - j\infty < s_i < C_i + j\infty$, $i = 1, 2, \ldots n$ and $dv_s = ds_1 \ldots ds_n$. The connection between the original function $f(t_1, \ldots, t_n)$ and the transformed function $F(s_1, \ldots, s_n)$ is written shortly in the form:

$$f(t_1, \ldots, t_n) \Leftrightarrow F(s_1, \ldots, s_n).$$

If $F_1(s_1, \ldots, s_n) \Leftrightarrow f_1(t_1, \ldots, t_n)$ and $F_2(s_1, \ldots, s_n) \Leftrightarrow f_2(t_1, \ldots t_n)$ then the formula of multi-variable convolution is valid:

$$F_1(s_1, \ldots, s_n)\, F_2(s_1, \ldots, s_n) \Leftrightarrow \int\limits_{E_n} f_1(\tau_1, \ldots, \tau_n) f_2\,(t_1 - \tau_1, \ldots, t_n - \tau_1)\, dv_\tau .$$

In order to use the convolution formula together with Volterra series, let us compare the integral (1.2) with the following expression:

$$y_n(t_1, \ldots, t_n) = \int\limits_{E_n} h_n(\tau_1, \ldots, \tau_n) \prod_{r=1}^{n} x(t_r - \tau_r)\, dv_\tau. \tag{2.3}$$

Applying the convolution formula to expression (2.3)

$$y_n(s_1, \ldots, s_n) = H_n(s_1, \ldots, s_n) \prod_{r=1}^{n} X(s_r).$$

In order to obtain the quantity $y_n(t)$ from here, we have to form the n-dimensional inverse transform, have to get the value $y_n(t_1, \ldots, t_n)$ and substitute $t_1 = \ldots = t_n = t$, as $y_n(t_1, \ldots, t_n)\big|_{t_1 = \ldots = t_n = t} = y_n(t)$.

However it is more convenient not to find the original n-dimensional function but to change to the one dimensional transform in the complex space in the following way. Let us substitute the value $t_i = t$, $i = 1, 2, \ldots, n$ in the expression (2.2), thus

$$f(t_1, \ldots, t_n)\big|_{t = \ldots = t_n = t} = \tag{2.4}$$

$$= \left(\frac{1}{2\pi j}\right)^{n-1} \int\limits_{Br_{n-1}} \exp\left(t \sum_{i=1}^{n-2} s_i + ts\right) dv_s\, ds \frac{1}{2\pi j} \int\limits_{C_{n-1} - j\infty}^{C_{n-1} + j\infty} \left(F(s_1, \ldots, s_{n-1}, s - s_{n-1})\right) ds_{n-1}.$$

In formula (2.4) the inner integral can be obtained by the theorem for substraction.

Repeating the process $n - 1$ times we change to $F(s)$ and find the $f(t)$ for this by one dimensional Laplace transformation. Further on we shall denote the change to one dimension in the complex space by $\{\quad\}^*$, that is $F(s) = \{F(s_1, \ldots, s_n)\}^*$.

Let us examine a few consequences of formula (2.4).

a) If $F(s_1, \ldots, s_n) = F(s_1 + \ldots + s_n)$, then to change to one variable the value $s = s_1 + \ldots + s_n$ has to be substituted.

b) If the transform has the form: $F(s_1, \ldots, s_n) = F_1(s_1 + \ldots + s_n)F_2(s_1, \ldots, s_n)$ then we apply the formula (2.4) only to $F_2(s_1, \ldots, s_n)$.

The majority of n-dimensional transforms occurring in practice has the form:

$$F(s_1, \ldots, s_n) = P(s_1 + \ldots + s_n) \prod_{r=1}^{n} F(s_n) .$$

In order to determine the one-variable original function of this expression we have to act in the following way. Let us find the one-variable original functions belonging to the functions $F_r(s_r) - f$, $r = 1, 2, \ldots$ and let us multiply them together:

$$f_2(t) = \prod_{r=1}^{n} f_r(t) .$$

Then we determine the one-variable transform $F_2(s)$ from the function $f_2(t)$. Let us multiply the functions $F_2(s)$ with $P(s_1 + \ldots + s_n)$ by substituting $s_1 + \ldots + s_n = s$ in the latter. As a result we get the required transform $F(s) = P(s)F_2(s)$.

c) If $f(t) \Leftrightarrow \{F(s_1, \ldots, s_n)\}^*$, then

$$\{F(s_1 + a_1, \ldots, s_n + a_n\}^* \Leftrightarrow f(t) \exp(-tA),$$

where $A = \sum_{i=1}^{n} a_i$.

d) If $f(t) \Leftrightarrow \{F(s_1, \ldots s_n)\}^*$, then for any value $\tau > 0$

$$\{F(s_1, \ldots, s_n \exp(-\tau \sum_{i=1}^{n} s_i)\}^* \Leftrightarrow f(t - \tau) .$$

3. Determination of the transforms of Volterra series kernels for systems given by their structural scheme

Using the theory of multi-variable Laplace transformation let us now construct the method for determining the transforms of the Volterra series for systems given by their structural scheme. In the general case the system

examined will contain three types of connections in the subsystems: parallel, serial and feedback. Therefore the method of determination of the transforms of Volterra series kernels is based on the formulae for these three connection types and can be considered as natural development of the results in works [6] and [7].

When two analytic systems are connected in parallel, then, considering that the multi-variable Laplace transformation is linear, the transforms of series kernels with equal order numbers are added. Let us now connect the two analytic systems in series. If the first of them, the input has the form $z(t) = \sum_{i=1}^{\infty} v_j[x(\tau), \tau \leq t]$ with kernels $h_j(\tau_1, \ldots, \tau_j)$, $j = 1, 2$, the second, the output has the form $y(t) = \sum_{i=1}^{\infty} v_i^*[z(\tau), \tau \leq t]$ with kernels $K_i(\tau_1, \ldots \tau_i)$, $i = 1, 2 \ldots$, then their series connection is described by the series:

$$y(t) = \sum_{i=1}^{\infty} v_1^* \Big[\sum_{j=1}^{\infty} v_j [x(\tau)], \tau \leq t \Big] . \tag{3.1}$$

The general member of this series is:

$$\int_E h_i(\tau_1, \ldots, \tau_i)\, K_{l_1(\sigma-\tau_1, \ldots, \sigma_{l_1}-\tau_1)}\, K_{l_2(\sigma_{l_1+1}-\tau_2, \ldots, \sigma_{l_1+l_2}-\tau_2)} \cdots$$

$$\cdots K_{l_p(\sigma_N+l_{p+1}-\tau_i, \ldots, \sigma_{N_m}-\tau_i)} \prod_{r=1}^{N} x(t-\sigma_r)\, dv_r , \tag{3.2}$$

where $i = 1, 2 \ldots$; $p = 1, 2 \ldots$; $N = \sum_{i=1}^{p} l_i$. It can be shown by using the convolution formula that the multi-variable transform of (3.2) is:

$$y(s_1, \ldots, s_s) = H_i \Big(\sum_{j=1}^{l} s_j, \sum_{j=l_1+1}^{l_1+l_2} s_j, \ldots, \sum_{j=N-l_p}^{N} s_j \Big) .$$

$$K_{l_1(s_i, \ldots, s_{l_1})}\, K_{l_2(s_{l_1}+1, \ldots, s_{l_1+l_2})} \cdots K_{l_p(s_{N-l_p+1}, \ldots, s_N)} \prod_{j=1}^{N} X(s_j) , \tag{3.3}$$

where $H_i(s_1, \ldots, s_i) \Leftrightarrow h_i(\tau_1, \ldots, \tau_i)$ and $K_{l_p(s_1, \ldots, s_{l_j})} \Leftrightarrow K_{l_j(\tau_1, \ldots, \tau_{l_j})}$. Let us explain the use of formula (3.3). Let us examine first the component of (3.1) for the case $i = 1$. Let us substitute the value $i = 1$ in (3.2) and (3.3) so that we get the transforms of the kernels:

$$H_1(s_1)\, K_1(s_1),\ H_1(s_1 + s_2)\, K_2(s_1, s_2) , \ldots$$

Let us also examine the components in (3.1) for the case $i = 2$. From (3.3):

$$H_2(s_1, s_2)\, K_1(s_1)\, K(s_2),\ H_2(s_1, s_2 + s_3)\, K_1(s_1)\, K(s_2, s_3) . \tag{3.4}$$

If we group the kernels of equal order and denote the transforms of the serially connected kernels by $H_i^*(s_1, \ldots, s_i)$ we obtain the following:

$$K(s_1) = H_1(s_1)\, K_1(s_1);$$

$$H_2^*(s_1, s_2) = H_1(s_1 + s_2)\, K_2(s_1, s_2) + H_2(s_1, s_2)\, K_1(s_1)\, K_1(s_2)\,; \ldots \tag{3.5}$$

Two serially connected systems are considered to be mutually invertible, if their serially connected kernels:

$$H_1^*(s_1) = 1;\; H_i^*(s_1), \ldots, s_i) \equiv 0,\; i = 2, 3, \ldots$$

From here the transform of the kernels of the inverse system:

$$K_1(s_1) = \frac{1}{H_1(s_1)}\,;$$

$$K_2(s_1, s_2) = \frac{H_2(s_1, s_2)}{H_1(s_1 + s_2)\, H_1(s_1)\, H_1(s_2)}\,; \ldots \tag{3.6}$$

Applying one after another the formulae obtained in this section, the transforms of the Volterra kernels can be determined for arbitrary type analytic systems given by their structural scheme.

4. Determination of the response of a nonlinear system to a regular input signal

After the determination of the convergence of the Volterra series and the convergence of the kernels transforms we can start with the determination of the output signal of the nonlinear system. Let the input signal be a regular (i.e. not random) time function. The response of the nonlinear system can be found in several ways.

If the Laplace transform of the input signal is known we can use for the determination of the output signal the formula for multi-variable convolution, formula (2.4) and its consequences described in the second part of this paper. At the same time the calculations can be led back to the application of the substraction theorem. For some special kinds of multi-variable transforms the change to one variable in the complex space can be realized on the basis of the tables shown in [8].

Let the input effect be given in the form of a Taylor series, constructed in the vicinity of point t:

$$x(t - \tau) = \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \frac{d^i x(t)}{dt^i} \tau^i\,. \tag{4.1}$$

The steady state signal of the system, using the series (4.1), can be written in the form:

$$y(t) = \int_{E_1} h_1(\tau)\, x(t-\tau)\, d\tau + \int_{E_2} h_2(\tau_1, \tau_2)\, x(t-\tau_1)\, x(t-\tau_2)\, dv_\tau + \ldots =$$

$$= \sum_{i=0}^{\infty} \frac{(-1)^i}{i!} \frac{d^i x(t)}{dt^i} \int_{E_1^+} \tau^i\, h_1(\tau)\, d\tau +$$

$$+ \sum_{i=0}^{\infty} \sum_{j=0}^{\infty} \frac{(-1)^{i+j}}{i!\, j!} \frac{d^i x(t)}{dt^i} \frac{d^j x(t)}{dt^j} \int_{E_2^+} \tau_1^i\, \tau_2^j\, h_2(\tau_1, \tau_2)\, dv_\tau + \ldots .$$

The quantities

$$\mu_{ij\ldots l}^n = \int_{E_n^+} \tau_1^i\, \tau_2^j \ldots \tau^l\, h_n(\tau_1, \ldots, \tau_n)\, dv_\tau$$

are called the moments of the kernels of the system. The moments of the system kernels can be calculated easily if we differentiate the connection (2.1), that is

$$\frac{\partial^i \partial^j \ldots \partial^l \{H_n(s_1, \ldots, s_n\}}{\partial s_1^i\, \partial s_2^j \ldots \partial s_n^l} = (-1)^{i+j+\ldots+l} \int_{E_n^+} \tau_1^i\, \tau_2^j \ldots \tau_n^l\, h_n(\tau_1, \ldots, \tau_n) \cdot$$

$$\cdot \exp\left(-\sum_{k=1}^{n} s_k \tau_k\right) dv_\tau . \tag{4.3}$$

Let us substitute the values $s_k = 0$; $k = 1, 2, \ldots, n$ in (4.3), then

$$\mu_{ij\ldots l}^n = (-1)^{i+j+\ldots+l} \left. \frac{\partial^i \partial^j \ldots \partial^l \{H_n(s_1, \ldots, s_n)\}}{\partial s_1^i\, \partial s_2^j \ldots \partial s_n^l} \right|_{s_1 = \ldots s_n = 0} .$$

Let the input effect now be given by its Fourier series. It is inconvenient to apply the usual Fourier series together with Volterra integrals if the order of the latter is higher than one, as the output signal cannot directly be obtained in the form of a Fourier series. However if the input signal is given in the form of complex Fourier series, then algebraic formulae can be obtained which connect the Fourier coefficients of the response of the system with the transforms of the kernels and the Fourier coefficients of the input effect. These formulae are given for the case of a special effect:

$$x(t) = C_1 \exp(-j\omega t) + C_1 \exp(+j\omega t) . \tag{4.4}$$

Let us substitute (4.4) into (1.1) and carry out the necessary factoring. If the output signal is sought in the form:

$$y(t) = \sum_{-\infty}^{\infty} d_n \exp(j\omega_n t), \tag{4.5}$$

then, supposing that every kernel of series (1.1) is symmetric, the coefficients are obtained in the following form:

$$\left.\begin{aligned}
d_0(j\omega) &= 2C_1C_{-1}H_2(j\omega, -j\omega) + \dots \\
d_{-1}(j\omega) &= C_{-1}H_1(-j\omega) + 3C_{-1}^2 C_1 H_3(j\omega, -j\omega, -j\omega) + \dots \\
d_1(j\omega) &= C_1H_1(j\omega) + 3C_1^2C_{-1}H_3(-j\omega, j\omega, j\omega) + \dots \\
d_{-2}(j\omega) &= C_1^2H_2(-j\omega, -j\omega) + \dots \\
d_{-3}(j\omega) &= C_{-1}^3H_3(-j\omega, -j\omega, -j\omega) + \dots \\
d_3(j\omega) &= C_1H_3(j\omega, j\omega, j\omega) + \dots
\end{aligned}\right\} \tag{4.6}$$

Here every coefficient $H_n(j\omega, \dots, j\omega)$ is obtained from the transforms of the Fourier type Volterra series multi-variable kernels $H_n(j\omega_1, \dots, j\omega_n)$ by substituting $\omega_1 = \dots = \omega_n = \omega$. Let all the Volterra functions have an odd exponent in series (1.1) of the nonlinear system. Then the coefficient $d_1(j\omega)$ in (4.6) takes the following form:

$$\begin{aligned}
d_1(j\omega) = C_1H_1(j\omega) &+ 3C_1^2C_{-1}H_3(-j\omega, j\omega, j\omega) + \\
&+ 8C_1^3C_{-1}^2H_5(-j\omega, -j\omega, j\omega, j\omega, j\omega).
\end{aligned}$$

Let us divide the expression for $d_1(j\omega)$ by C_1, thus we obtain the equivalent frequency coefficient $d_e(j\omega)$ of the nonlinear system for the first harmonic

$$\begin{aligned}
d_e(j\omega) = H_1(j\omega) &+ 3C_1C_{-1}H_3(-j\omega, j\omega, j\omega) + \\
&+ 8(C_1C_{-1})^2H_5(-j\omega, -j\omega, j\omega, j\omega, j\omega, j\omega).
\end{aligned} \tag{4.7}$$

It follows from (4.7) that, contrary to the frequency characteristic of the linear system, the coefficient $d_e(j\omega)$ depends on the amplitude of the input effect.

5. Compensation of nonlinear systems

Let us formulate the task of nonlinear compensation. Let the structural scheme of the system forming the subject of synthesis as well as the parameters of the non-changing elements be given.

Let the parameters of the required linear system also be given. The kernels of the compensator have to be determined so that the distance between

the requored $y_d(t)$ and the actual output $y(t)$ be small in the space of the continuous functions $C(T)$. During the solution of the given task we suppose that all linear subsystems are stable and have minimal phase.

We shall consider that the required transfer function is known, obtained by one of the synthesis methods of the theory of linear systems. The proposed compensation method is based on the determination algorithm of the transforms of Volterra series kernels derived in the third section and does not depend on the linear system having inertia or not. Therefore we limit ourselves to the case when the nonlinear system is not inert. It is more convenient to examine the theory of compensation on Fig. 1 illustrating a concrete example of a feedback system, where the nonlinear block F is described by the polynomial $y = \sum_{n=1}^{N} a_n z^n$ and the linear block Φ has a transfer function $H(s)$. The compensator is constructed as a Volterra series. Let us denote the transforms of its kernel by $\{K_i(s_1, \ldots, s_i)\}$. Applying the method of determination of the kernel transforms for the connection between the compensator and the linear system we can write:

$$K_1(s)\,H(s);\;\; K_2(s_1, s_2)\,H(s_1 + s_2)\,;\;\; K_3(s_1, s_2, s')\,H(s_1 + s_2 + s');\ldots \tag{5.1}$$

Let us determine the transforms of the kernels of the entire forward branch. These are:

$$\begin{aligned} &a_1 K_1(s)\,H(s) = H_d(s)\,; \\ &a_2 K_1(s_1)\,K_1(s_2)\,H(s_1)\,H(s_2) + a_1 K_2(s_1, s_2)\,H(s_1 + s_2) = 0;\ldots \end{aligned} \tag{5.2}$$

Let us denote the required transfer function of the forward branch by $H_d(s)$. From here we get the following set of equations for the transforms of the compensator kernels:

$$\begin{aligned} &a_1 K_1(s)\,H(s) = H_d(s); \\ &a_2 K_1(s_1)\,K_1(s_2)\,H(s_1)\,H(s_2) + a_1 K_2(s_1, s_2)\,H(s_1 + s_2) = 0;\ldots \end{aligned} \tag{5.3}$$

Solving the set (5.3):

$$K_1(s) = \frac{H_d(s)}{a_1 H(s)}\,;$$

$$K_2(s_1, s_2) = -\frac{a_2 H_d(s_1)\,H_d(s_2)}{a^3 H(s_1 + s_2)}\,;\ldots \tag{5.4}$$

The task of compensation can be solved in an analogous way when the compensator is in the feedback branch. The entire method of compensation in-

variably remains valid even in the case when the nonlinear element to be compensated can be generated by a Volterra series. In this case the form of every formula remains the same, but instead of the coefficients a_i the transforms $H_i(s_1, \ldots, s_i)$ of the multi-variable kernels appear. As it follows from the nature of formula (5.4), every kernel of the compensator will be physically realizable if the first order kernel is realizable.

The proposed solution of the compensation task is convenient as it produces the transforms of the compensator kernels. The question of realization was investigated in [9] where it is shown that the kernels of not too high an order can be synthesized in a relatively simple way.

General conclusions

The functional Volterra series represent a mathematical apparatus capable of solving a fairly wide range of nonlinear problems. The method of determination of the transforms of Volterra series kernels and of the output signal, which we have proposed in the paper, make it possible to reduce significantly the problem of analysis of analytic systems.

The method of synthesis examined in this paper is convenient for the construction of linear systems from given nonlinear systems, even though it can be used as a general method constructed in [1] for the synthesis of optimal nonlinear systems. However, the task of nonlinear compensation in this latter formulation does not have a great practical value because of the following reasons. First of all the higher order compensator kernels produced can be (and usually they are) physically non-realizable, even though the first order kernel is realizable [10]. Second, as is shown in [11], the direct determination of the transforms of the kernels of the required system can be carried out only for a special effect not occurring in practice. During the practical application of the analysis and synthesis methods of nonlinear systems examined in the paper the convergence of the Volterra series obtained has to be checked. The method proposed in [1] can be applied to this effect.

References

1. *Ван-Трис, Г.*, Синтез оптимальных нелинейных систем управления. «Мир», 1964.
2. *Van Trees, H.*, Functional techniques for the analysis of nonlinear behaviour of phase locked loops. Proc. IEEE, **52** (1964) 8.
3. *Flake, R.*, Volterra series representation of nonlinear systems. IEEE Trans. on Appl. and Ind., **82** (1963).
4. *Flake, R.*, Volterra series representation of time-varying nonlinear systems. Proc. of Second IFAC Congress, London, 1963.
5. *Ku, Y., Wolf, A.*, Volterra Wiener functionals for the analysis of nonlinear systems. J. Franklin Inst., Jan. 1966.

6. *Barrett, J.*, The use of functionals in the analysis of nonlinear physical systems. Jour. of Electronics and Control, **15** (1963).
7. *Kielkewicz, M.*, Cascade connection of nonlinear systems. Int. Journ. of Control **11** (1970) 6.
8. *Lubbock, J.*, *Bansal, V.*, Multidimensional Laplace transforms for solution of nonlinear equations. Proc. IEE, **116** (1969) 12.
9. *Schetzen, M.*, Synthesis of a class of nonlinear systems. Int. Journ. of Electronics, **5** (1966).
10. *Киселева, О. Н.*, Условия физической реализуемости при статистической оптимизации нелинейных систем. Автоматика и телемеханика, **9** (1967).
11. *Katzenelson, J.*, *Gould, L.*, A spectrum factorization method of calculation of nonlinear filters of Volterra type. Inf. and Control, **8** (1965) 3.

DECOMPOSITION OF CYCLIC CODES INTO EQUAL-WEIGHT CLASSES

S. SH. OGANESYAN, V. I. TAIRYAN, V. G. YAGDZYAN

(Erevan)

Formulae are derived for cyclic codes over finite $GF(q)$ fields and for the p-tuple representatives of arbitrary cyclic codes.

The search into the correction capabilities of cyclic codes is directly connected with the analysis of the weight distribution of the code vectors, i.e. with finding the weight spectrum of the codes. The reasonableness of the determination of the weight spectrum consists in that the weights are not calculated for every code vector, but for a certain, possibly smaller subset [1—5]. Thereby the time required for the computation of the weight spectrum by the computer is significantly reduced.

The experiments to find the cyclic representatives had been carried out until now for arbitrary codes interpreted over $GF(q)$ Galois fields, where $(n, q = 1)$ (q is the degree of the primitive number p) [1, 2, 5] but unfortunately no general analytic expression was found for such representatives.[1]

In the present paper we derive expressions not only for the cyclic representatives of arbitrary codes, but also for the so called p-tuple representatives[2] which connect the cyclic codes on the basis of equal weight, with the help of the operation of raising to the power p.

It is known that the cyclic code can be considered to be the set of polynomial formed from the ring of R_n^q polynomials according to module $x^n - 1$ over the $GF(q)$ field of the coefficients ($q = p^\mu$, μ is a natural number) which can be divided to fixed $q(x)$ polynomials. Such a polynomial set is called an ideal, and $g(x)$ is the generator polynomial of this ideal.

Let u be an arbitrary ideal over $GF(q)$ and let the polynomial $a(x) = \sum_{i=0}^{n=1} a_i x_i$ belong to the ideal u. The number S of the nonzero coefficients of $a(x)$ is called the weight of the polynomial $a(x)$, i.e.

$$S = \sum_{i=0}^{n-1} |a_i| \quad \text{where} \quad |a_i| = \begin{cases} 0 & \text{if} \quad a_i = 0, \\ 1 & \text{if} \quad a_i \neq 0. \end{cases}$$

[1] For the case $q = 2$ the problem had been solved in [3].
[2] For $q = 2$ with some limitations quadratic representatives are found in [4].

It is clear that the weights of the polynomials $a(x)$ and $x^{\mu}a(x) \in u$ are the same. As a consequence all the polynomials of the ideal u can be divided into classes, the cyclic classes of equal weight polynomials.

Let us assign a suitable number b to every cyclic class which is called the cyclic period of the class, or shortly the period, which corresponds to the number of polynomials in the class. Let us further assign a number S to every cyclic class, called the weight of the class which corresponds to the weight of the polynomials of this class. Then, corresponding to every polynomial of the ideal u such a number b, the period of the cyclic class can be assigned to it, to which this polynomial belongs. Every cyclic class can be generated by its cyclic representative, and the entire ideal by the set of the cyclic representatives.

Let the expressions $\{a_1(x)\} = \{a_1(x),\ xa_1(x) \ldots x^{b_1-1}a_1(x)\}$ and $\{a_2(x)\} = \{a_2(x), xa_2(x) \ldots x^{b_2-1}a_2(x)\}$ represent certain cyclic classes of the arbitrary ideals u_1 and u_2 respectively, which do not contain common elements ($u_1 \cap u_2 = 0$). The classes $\{a_1(x)\}$ and $\{a_2(x)\}$, corresponding to the cyclic representatives $a_1(x)$ and $a_2(x)$ have periods b_1 and b_2. The set of all elements, in the form $x^i(a_1(x) + x^j a_2(x)$, where $i = \overline{0,\ b_1 - 1}$, $j = \overline{0, b_2 - 1}$ is nothing else but the direct sum of the respective cyclic classes $\{a_1(x)\}$ and $\{a_2(x)\}$ of the ideals u_1 and u_2.

Theorem 1. The direct sum of the cyclic classes $\{a_1(x)\}$ and $\{a_2(x)\}$ can be decomposed to $(b_1,\ b_2)$ cyclic classes with period $T = [b_1, b_2]$. Meanwhile the elements

$$x^d a_1(x) + a_2(x),\ d = \overline{0, (b_1 b_2) - 1} \tag{1}$$

can be found in different classes and they can be selected as cyclic representatives.

Proof. Let us determine the period T of elements of type (1), i.e. the smallest natural number for which the following comparison is valid:

$$x^T(x^d a_1(x) + a_2(x)] \equiv x^d a_1(x) + a_2(x) \bmod X^n - 1. \tag{2}$$

According to the condition $u_1 \cap u_2 = 0$ it follows from (2) that:

$$x^{d+T} a_1(x) \equiv x^d a_2(x) \mod X^n - 1, \tag{3}$$

$$x^T a_2(x) \equiv a_2(x) \mod X^n - 1. \tag{4}$$

The periods of $a_1(x)$ and $a_2(x)$ are b_1 and b_2 respectively, therefore from (3) and (4)

$$d + T \equiv d \bmod b_1 \qquad T \equiv 0 \bmod b_2.$$

As a consequence $T \equiv 0 \bmod [b_1, b_2]$, therefore $T = [b_1, b_2]$. The further proof of the theorem leads to the proof of two statements.

1. All type (1) elements can be found in different classes, i.e. from the fulfilment of the following comparison

$$x^{\mu}(x^d a_1(x) + a_2(x)] \equiv x^{d'} a_1(x) + a_2(x) \bmod X^n - 1. \tag{5}$$

where $d, d' < (b_1, b_2)$ it follows that $d = d'$.

2. The cyclic classes characterized by (1) type representatives contain all elements of the direct sum of the cyclic classes $\{a_1(x)\}$ and $\{a_2(x)\}$.

1. From comparison (5), considering the condition $u_1 \cap u_2 = 0$:

$$x^{d+\mu} a_1(x) \equiv x^{d'} a_1(x), \; x^{\mu} a_2(x) \equiv a_2(x) \qquad \bmod X^n - 1.$$

As the periods of $a_1(x)$ and $a_2(x)$ are b_1 and b_2 respectively, therefore $d + \mu = d' \bmod b_1$, $u \equiv 0 \bmod b_2$, that is $d = d' \bmod (b_1, b_2)$, but $d, d' < (b_1, b_2)$ therefore $d = d'$.

2. The number of elements of the direct sum of cyclic classes $\{a_1(x)\}$ and $\{a_2(x)\}$ is $N_1 = b_1 \cdot b_2$. The number of elements of cyclic classes characterized by (1) type representatives is $N_2 = (b_1, b_2)[b_1, b_2] = b_1 \cdot b_2$, that is $N_1 = N_2$, and this is what had to be proved.

Let $A_1 = \{\{a_1^{(1)}(x)\}, \ldots, \{a_i^{(1)}(x)\}, \ldots, \{a_f^{(1)}(x)\}\}$ and $A_2 = \{\{a_1^{(2)}(x)\}, \ldots, \{a_j^{(2)}(x)\}, \ldots, \{a_l^{(2)}(x)\}\}$ be the set of all cyclic classes of the arbitrary ideals u_1 and u_2 which do not contain common elements ($u_1 \cap u_2 = 0$), the cyclic representatives $a_i^{(1)}(x)$ and $a_j^{(2)}(x)$ have periods $b_i^{(1)}$ and $b_j^{(2)}$ respectively It is trivial that the direct sums of different arbitrary pairs of cyclic classes of ideals u_1 and u_2 do not contain common elements ($u_1 \cap u_2 = 0$). On this basis from Theorem 1 the cyclic representatives of the cyclic classes of theses direct sum of the ideals u_1 and u_2 can be found that is of the ideal $u = u_1 \oplus u_2$, on the basis of the next theorem.

Theorem 2. The cyclic representatives of the cyclic classes of ideal u have the following form:

$$x^{d_{ij}} a_i^{(1)} + a_j^{(2)}, \quad d_{ij} = \overline{0, (b_i^{(1)}, b_j^{(2)}) - 1},$$
$$i = \overline{1, f}, \qquad j = \overline{1, l}.$$

Let A_i be the set of all cyclic classes of minimal ideals u_i, $i = \overline{1, t}$.[3] Thus, relying on Theorem 2, with the method of mathematical induction and on the basis of the next theorem we can determine the cyclic representatives of the

[3] The method of determination of sets A_i can be found in [1].

cyclic classes of the direct sum of the ideals u_i, $i = \overline{1, t}$, of the ideal $u = u_1 \oplus \oplus u_2 \oplus \ldots \oplus u_i \oplus \ldots \oplus u_t$.

Theorem 3. The cyclic representatives of cyclic classes of the ideal u have the form:

$$\sum_{i=1}^{t-1} x^{\sum_{j=i}^{t-1} d_j} a_i(x) + a_t(x), \tag{I}$$

where $a_i(x)$ is the representative of class $\{a_i(x)\} \in A_i$, the value d_j is chosen at fixed $a_i(x)$, $i = \overline{1, t}$ from the set $D_j = \{0, 1 \ldots ([b_1, b_2 \ldots b_j], b_{j+1}) - 1\}$. The period T of the cyclic representative is $[b_1, b_2, \ldots b_t]$.

Let $A = \{\{a_1(x)\}, \ldots, \{a_i(x)\}, \ldots, \{a_j(x)\} \ldots \{a_t(x)\}\}$ be the set of all cyclic classes of the arbitrary ideal u. Let us note the validity of the following operation for the following classes: $\{a_i(x)\}\{a_j(x)\} = \{a_i(x)\, a_j(x)\}$. Therefore $\{a_i(x)\}^p = \{a_i^p(x)\}$, and as the characteristic curve of the coefficient field of the polynomials of ideal u is of order p, therefore the weight of the cyclic class $\{a_i^p(x)\}$ equals the weight of class $\{a_i(x)\}$.

Let us decompose the set A to new classes with the help of raising to the power p, which we shall call p-tuple classes, or second order classes (practically the cyclic classes are the first order classes). Let us denote the set of second order classes belonging to ideal u by R, i.e. $R = \{\{r_1(x)\}, \ldots \ldots, \{r_i(x)\}, \ldots, \{r_e(x)\}\}$, where $\{r_i(x)\} = \{\{a_j(x)\}, \{a_j(x)\}^p, \ldots, \{a_j(x)\}^{p^{c_i}}\}$, c_i is the number of first order classes in the second order ith class, this number is called p-period.

Theorem 4. The cyclic periods of the classes $\{a(x)\}$ and $\{a(x)\}^p$ of ideal u are equal.

Proof. The proof of the theorem is trivial for the minimal ideal, considering [1]. As the arbitrary ideal u is the direct sum of a certain number of minimal ideals, i.e. $u = u_1 \oplus u_2 \oplus \ldots \oplus u_i \oplus \ldots u_t$, $t \leq 1$, therefore an arbitrary element $a(x)$ of the ideal u can be presented in the form $a(x) = \sum_{i=1}^{t} a_i(x)$ where $a_i(x) \in u_i$. At the same time the period T of element $a(x)$ has the value $[b_1, b_2 \ldots b_t]$ (Theorem 3). The period T_1 of the cyclic class $\{a(x)\}^p$ is determined from the comparison:

$$x^{T_1}(a_1(x) + a_2(x) + \ldots + a_t(x))^p \equiv$$
$$\equiv (a_1(x) + a_2(x) + \ldots + a_t(x))^p \qquad \mathrm{mod}\ X^n - 1. \tag{6}$$

Considering that $u_i \cap u_j = 0$ for all $i \neq j$ and the truth of the theorem for any minimal ideal, from (6):

$$T_1 \equiv 0 \qquad \bmod b_i, \qquad i = \overline{1, t},$$

that is $T_1 = [b_1, b_2, \ldots, b_t]$ and $T_1 = T$, and this is what had to be proved. Any polynomial $a(x)$ of ideal u can be described by the weight S, the number b, which equals the period of the first order class containing $a(x)$, and further by the number c, which equals the p-period of the second order class to which the first order class belongs, which in turn contains the polynom $a(x)$. Every second order class can be generated by a p-tuple representative, and the entire ideal by the set of p-tuple representatives.

Let $R_1 = \{\{r_1(x)\}, \{r_1(x)\}^p, \ldots \{r_1(x)\}^{p^{c_1-1}}\}$ and $R_2 = \{\{r_2(x)\}, \{r_2(x)\}^p, \ldots \{r_2(x)\}^{p^{c_2-1}}\}$ be the second order classes of certain arbitrary ideals u_1 and u_2 respectively, which do not contain common elements, that is $u_1 \cap u_2 = 0$. At the same time the p-tuple representatives $r_1(x)$ and $r_2(x)$ of classes R_1 and R_2 have periods b_1 and b_2 respectively, and p-periods of size c_1 and c_2.

The direct sum of the second order classes R_1 and R_2 is the set of the direct sums of the first order classes $\{r_1(x)\}^{p^i}$ and $\{r_2(x)\}^{p^j}$ for all values $i = \overline{0, c_1 - 1}$ and $j = \overline{0, c_2 - 1}$ that is

$$R_1 \oplus R_2 = \bigcup_{i,j} (\{r_1(x)\}^{p^i} \oplus \{r_2(x)\}^{p^j}).$$

Let us decompose the set $\mathfrak{D} = \{0, 1 \ldots (b_1, b_2) - 1\}$ to p-cycles by substituting the value $\omega \to p\omega \bmod t = (b_1, b_2)$, where $\omega \in \mathfrak{D}$. The number of p-cycles is denoted by K; the number of numbers in the ith p-cycle, which is called p-period, by t_i; and the smallest representative of the p-cycle by ω_i. Let us denote:

$$F^i = x^{\omega_i}.$$

Theorem 5. The direct sum of the second order classes R_1 and R_2 of ideals u_1 and u_2 decomposes to second order $\sum_{i=1}^{k}$ class with p-period $T_p = [c_1, t_i, c_2]$, $i = \overline{1.K}$. The elements

$$r_1^{p^{\varphi+\gamma}}(x) \cdot F_i^{p^\gamma} + r_2(x) \tag{7}$$

where $\varphi = \overline{0, (c_1, t_i) - 1}$, $\gamma = \overline{0, ([c_1, t_i], c_2) - 1}$, $i = \overline{1, K}$ can be found in different second order classes and can be selected as p-tuple representatives.

Proof. Let us determine the p-period T_p of the second order classes which can be represented by (7) type representatives, that is the smallest

natural number for which the following comparison holds:

$$x^{\nu}\left(r_1^{p^{\gamma+\varphi}}(x)\, F_i^{p^{\gamma}} + r_2(x)\right)^{p^{T_p}} \equiv r_1^{p^{\gamma+\varphi}}(x)\, F_i^{p^{\gamma}} + r_2(x) \bmod X^n - 1\,. \tag{8}$$

As $F_i = x^{\omega_i}$ and $u_1 \cap u_2 = 0$, therefore from (8) we obtain the following:

$$r_1^{p^{\varphi+\gamma+T_p}}(x)\cdot(x)^{\omega_i p^{\gamma+T_p}+\nu} \equiv r_1^{p^{\varphi+\gamma}}(x)\cdot(x)^{\omega_i p^{\gamma}},\ r_2^{p^{T_p}}(x)\, x^{\nu} \equiv r_2(x)\,. \tag{9}$$

The p-periods of $r_1(x)$, $r_2(x)$, ω_i and the periods of $r_1(x)$, $r_2(x)$ have the values c_1, c_2, t_i and b_1, b_2 respectively, therefore from comparison (9):

$T_p \equiv 0 \mod c_1$, $T_p \equiv 0 \mod c_2$, $T_p \equiv 0 \mod t$, that is $T_p = [c_1, t_i, c_2]$.

The further proof of the theorem leads to the proof of two statements.

1. All (7) type elements can be found in different second order classes, that is can be found as a result of the fulfilment of the following comparison:

$$x^{\nu}\left(r_1^{p^{\varphi+\gamma}}(x)\, F_i^{p^{\gamma}} + r_2(x)\right)^{p^{\mu}} \equiv r_1^{p^{\varphi'+\gamma'}}(x)\, F_j^{p^{\gamma'}} + r_2(x) \bmod X^n - 1\,, \tag{10}$$

where γ, μ are natural numbers; $i, j < K$;

$$\gamma < (c_1, t_i);\ \gamma < ([c_1, t_i], c_2);\ \varphi' < (c_2, t_j);\ \gamma' < ([c_1, t_j], c_2)$$

it follows that $i = j$, $\varphi = \varphi'$, $\gamma = \gamma'$.

2. The second order classes which can be characterized by (7) type representatives contain all the elements of the direct sum of second order classes R_1 and R_2.

1. Considering that $F_i = x^{\omega_i}$, $F_j = x^{\omega_j}$ and $u_1 \cap u_2 = 0$, it follows from comparison (10):

$$r_1^{p^{\varphi+\varphi+\mu}}(x)^{\omega_i p^{\gamma+\mu}} \equiv r_1^{p^{\varphi'+\gamma'}}(x)\, x^{\omega_j p^{\gamma'}} \bmod X^n - 1\,; \tag{11}$$

$$r_2^{p^{\mu}}(x)\, x^{\nu} \equiv r_2(a) \bmod X^n - 1\,. \tag{12}$$

Here the value of the p-periods and of the $r_1(x)$, $r_2(x)$ periods is c_1, c_2 and b_1, b_2 respectively, therefore from comparisons (11) and (12):

$$\varphi + \gamma + \mu \equiv \varphi' + \gamma' \mod c_1, \tag{13}$$

$$\mu \equiv 0 \mod c_2, \tag{14}$$

$$\omega_i p^{\gamma+\mu} \equiv \omega_j p^{\gamma'} \bmod b_1\,, \tag{15}$$

$$\nu \equiv 0 \mod b_2. \tag{16}$$

It follows from (15) and (16) that:

$$\omega_i p^{\gamma+\mu} \equiv \omega_j p^{\gamma'} \bmod t = (b_1, b_2)\,, \tag{17}$$

that is ω_i and ω_j can be found in one p-cycle in module t, and as they are essentially the smallest representatives of the p-cycle, therefore $i = j$. Thus, considering that the p-period of ω_i is t_i, from (17):

$$\gamma + \mu \equiv \gamma' \mod t_i. \tag{18}$$

From (13) and (18) it follows that $\varphi \equiv \varphi' \mod (c_1, t_i)$, that is $\varphi = \varphi'$. Then from (13), (18) and (14):

$$\gamma \equiv \gamma' \mod ([c_1, t_i], c_2),$$

therefore $\gamma = \gamma'$.

2. The number of elements in the direct sum of the second order classes R_1 and R_2 is $N_3 = c_1 \cdot b_1 \cdot c_2 \cdot b_2$. However in second order classes characterized by (7) type representatives the number of elements for fixed F_i is

$$(c_1, t_i)([c_1, t_i], c_2)\,[c_1, t_i, c_2]\,[b_1, b_2] = c_1 \cdot c_2[b_1, b_2]t_i.$$

The total number of these is $N_4 = \sum_{i=1}^{k} c_1 \cdot c_2[b_1, b_2]t_i = c_1\, c_2[b_1\, b_2] \sum_{i=1}^{k} t_i = = c_1\, c_2[b_1, b_2](b_1, b_2) = c_1 \cdot c_2 \cdot b_1 \cdot b_2$, that is $N_3 - N_4$, and this is what had to be proved.

Let $R^{(1)} = \{\{r_1^{(1)}(x)\}, \ldots, \{r_1^{(1)}(x)\}, \ldots, \{r_d^{(1)}(x)\}\}$ and $R^{(2)} = \{\{r_1^{(2)}(x)\}, \ldots \ldots, \{r_j^{(2)}(x)\}, \ldots, \{r_n^{(2)}(x)\}\}$ be the set of all second order classes of the arbitrary ideals u_1 and u_2 not containing common elements $(u_1 \cap u_2 = 0)$ and the p-tuple representatives $r_i^{(1)}(x)$, $r_j^{(2)}(x)$ have periods $b_i^{(1)}$, $b_j^{(2)}$, and p-periods $c_i^{(1)}$, $c_j^{(2)}$ respectively.

It is trivial that the direct sums of the arbitrary, different second order classes of ideals u_1 and u_2 do not contain common elements $(u_1 \cap u_2) = 0$. For every pair of the second order classes $\{r_1^{(1)}(x)\}$ and $\{r_j^{(2)}(x)\}$ let us form the set $\mathfrak{D}_{ij} = \{0, 1, \ldots (b_i^{(1)}, b_j^{(2)} - 1\}$. Let us denote the number of numbers in $\mathfrak{D}_{ij}$ by t_{ij} and decompose every set to non-cutting cycles $[(p, t_{ij}) = 1]$ by substituting $\omega \to p\omega \mod t_{ij}$. Let us denote the set of p-tuple representatives of these cycles by $\mathfrak{W}_{ij}$, and the set $\{x^{\omega_{ij}}\}$, where $\omega_{ij} \in \mathfrak{W}_{ij}$, by $R = = \{r_{11}(x), \ldots r_{ij}(x), \ldots, r_{dh}(x)\}$, $i = \overline{1, d}$, $j = \overline{1, h}$. The number of numbers in the ijth cycle is denoted by c_{ij} Then according to Theorem 5 the p-tuple representatives of the second order classes of the direct sum $u = u_1 \oplus u_2$ of ideals u_1 and u_2 is determined by the theorem.

Theorem 6. The p-tuple representatives of the second order classes of ideal u have the following form:

$$\left(r_i^{(1)}(x)\right)^{p^{\varphi+\gamma}} \cdot r_{ij}^{p^{\gamma}} + r_j^{(2)}(x)\,,$$

where $\overline{\varphi = 0, (c_i^{(1)}, c_{ij}) - 1}$, $\gamma = \overline{0, ([c_1^{(1)}, c_{ij}], c_j^{(2)}) - 1}$, $i = \overline{1, d}$, $j = \overline{1, h}$.

On the basis of [1] the set of cyclic representatives of the first order classes of minimal ideal u_i is $\{0, \alpha_i^0 \varepsilon_i(x), \alpha_i^{(1)}(x)\, \varepsilon_i(x), \ldots \alpha_i^{m_i-1}(x)\, \varepsilon_i(x)\}$ therefore the substitution $\omega \to p\omega \mod m_i$, where $(p, m_i) = 1$, $\omega \in \{0, 1 \ldots m_i - 1\}$ decomposes the set of first order classes to second order classes.

Let us choose the cyclic representatives of lowest degree first order classes as the p-representatives of second order classes, suitably for every second order class.

Let R_{2i-1}, $i = \overline{1,t}$ be the sets of p-tuple representatives of second order classes of minimal ideals u_i. Let us form $(r_1(x), \ldots, r_{2i-1}(x), \ldots, r_{2t-1}(x))$ type sets where $r_{2i-1}(x) \in R_{2i-1}$, $i = \overline{1, t}$, are generated by all possible choices.

For every such version let us compose the family of sets $\{\mathfrak{D}_2, \ldots \mathfrak{D}_{2i}, \ldots \ldots, \mathfrak{D}_{2+2}\}$, where $\mathfrak{D}_{2i} = \overline{0, ([b_1, \ldots b_i]\, b_{i+1}) - 1}$, $i = \overline{1, t}$. Let us denote the number of numbers in $\mathfrak{D}_{2i}$ by t_i, and decompose every set of numbers $\mathfrak{D}_{2i}$ to non-cutting p-cycles by substituting $\omega \to p\omega \mod t_i((t_i, p) = 1)$. The set of p-tuple representatives of these cycles is denoted by $\mathfrak{W}_{2i}$, the set $\{r_{2i}(x) = = x^{\omega_{2i}}\}$, where $\omega_{2i} \in \mathfrak{W}$ by R_{2i}. Every $r_{2i}(x)$, which is also called p-tuple representative, can be characterized by the number c_{2i}, which agrees with the number of numbers in p-cycle having p-tuple representative ω_{2i}. The set of all possible variants of the type $(r_2(x), \ldots, r_{2i}(x), \ldots r_{2t-2}(x))$ is denoted by M.

With the above notations, using the method of mathematical induction we determine the p-tuple representatives of the second order class of the direct sum of ideal $u = u_1 \oplus u_2 \oplus \ldots \oplus u_i \oplus \ldots \oplus u_t$ of the smallest minimal ideals u_i, $i = \overline{1, t}$, on the basis of the theorem.

Theorem 7. The p-tuple representatives of the second order classes of ideal u can be given by the expression:

$$\sum_{i=1}^{t-1} r_{2i-1}^{\sum_{j=2i-1}^{2t-2} \varphi_j}(x) \prod_{\nu=i}^{t-1} r_{2i}^{\sum_{j=2\nu}^{2t-2} \varphi_j}(x) + r_{2t-1}(x)\,, \tag{II}$$

where $r_1(x), \ldots, r_{2t-1}(x)$ are the elements of variant $\omega \in \mathfrak{W}$, $r_2(x), \ldots r_{2t-2}(x)$ are the elements of the variant which belongs to set M for fixed ω, the values $\varphi_1, \ldots, \varphi_i, \ldots, \varphi_{2t-2}$ have to be chosen from the set of numbers $\{0, 1, \ldots, ([c_1, \ldots c_i]c_{i+1}) - 1\}$, $i = \overline{1, 2t-2}$ at fixed values of r_i. The T_p-period of the p-tuple representatives is: $[c_1, c_2, \ldots c_{2t-2}]$.

References

1. *Goethals, J. M.*, Analysis of weight distribution in binary cyclic codes. IEEE Trans. **IT-12** (1966) *3*, 401–402.
2. *Мак-Вильямс Дж.* Структура и свойства бинарных циклических алфавитов. Кибернетический сборник (1967) *4*, 7—42.
3. *Оганесян С. Ш., Ягджян В. Г.* Нахождение циклических представителей в бинарных циклических алфавитах. Труды ВЦ АН Арм. СССР и ЕГУ (1969) *6*, 35–38.
4. *Оганесян, С. Ш., Ягджян, В. Г.* Объединение циклических представителей по одинаковым весам. Труды ВЦ АН Арм. ССР и ЕГУ (1969) *6*, 39–48.
5. *Tavares, S. E., Allard, P. E., Shiva, S. G. S.*, On the decomposition of cyclic codes into cyclic classes. Information and Control, New York **18** (1971) *4*, 342–354.

LINEAR CODING IN THE RETRANSMISSION OF THE SIGNAL

N. K. MILENIN

(Ryazan)

The task of synthesis of optimal linear encoding and decoding filters on channels provided with multiple retransmission of the signal is examined. Formulae have been produced for the optimal transfer functions of the final and intermediate members with the condition that the average power of the transmitters is equal and the energy spectra of the noise coincide on all sections of the channel.

1. Introduction

At the transmission of information to great distances the multiple or single retransmission of the signal is applied, for example in radio relay and cable connections, furthermore in connection systems established through the artificial satellites of the Earth. The figure shows the simplified block diagram of a connection using multiple retransmission of the signal. The following notations are introduced: $S(\omega)$ is the energy spectrum of the useful signal, $K_1(j\omega)$ and $W(j\omega)$ are respectively the transfer functions of the encoding and decoding filters at the end, $K_2(j\omega)$, $K_3(j\omega)$, . . . , $K_n(j\omega)$ are the transfer functions of the intermediate encoding filters (retranslators), $N_i(\omega)$ is the energy spectrum of additive, fluctuating noise exerting its effect on the ith section of the connection, $i = 1, 2, 3, \ldots, n$, where n is the total number of sections. At the same time we suppose that the attenuation of the useful signal has already been compensated for on each section i of the connection by the following retransmitter, and that the energy spectrum $N_i(\omega)$ of the noise has been suitably recalculated for the input of this retransmitter.

The aim of the present work is to find the optimal transfer functions of the coding filter $[K_i(j\omega)]_0$, $i = 1, 2 \ldots n$ and the decoding filter $W_0(j\omega)$, without considering their physical realisability. At the same time we suppose that the useful signal and the noises are independent random processes with a zero expected value and that every section of the connection is linear.

As an optimisation criterion we apply the frequency weighted quadratic mean value criterion with unequal weighting functions $A^2(\omega)$ and $M^2(\omega)$ for the dynamic and random errors, i.e

$$\sigma^2 = \sigma_d^2 = \sigma_e^2 = \frac{1}{2\pi} \int_{\omega_a}^{\omega_f} S_0(\omega)\, A^2(\omega)\, d\omega + \frac{1}{2\pi} \int_{\omega_a}^{\omega_f} N_0(\omega)\, M^2(\omega)\, d\omega = \min\,, \quad (1.1)$$

where σ_d^2 and σ_e^2, $S_0(\omega)$ and $N_0(\omega)$ are the weighted deviations and energy

spectra of the dynamic and random errors respectively, σ^2 is the deviation of the resulting weighted error, $f_f = \omega_f/2\pi$ and $f_a = \omega_a/2\pi$ are the upper and lower bounds of the transmission band of the connection obtained from the following condition: $A(\omega) = 0$ if $\omega < \omega_f$ and $\omega < \omega_a$. At the same time we suppose that the dynamic error results at the expense of the deviation of the useful output signal from the input, and that random errors are determined by the resultant level of noises appearing at the output.

2. Optimal transfer function of the decoding filter

Applying the principle of superposition and the figure it is not difficult to find the energy spectra of the dynamic and random errors:

$$S_0(\omega) = S(\omega) \left| W(j\omega) \prod_{i=1}^{n} K_i(j\omega) - 1 \right|^2, \tag{2.1}$$

$$N_0(\omega) = |W(j\omega)|^2 \sum_{k=1}^{n} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2, \tag{2.2}$$

where for the case $k = n$ $\prod_{i=k+1}^{n} |K_i(j\omega)|^2 = 1$. Let us substitute the expressions (2.1) and (2.2) into (1.1):

$$\sigma^2 = \frac{1}{2\pi} \int_{\omega_a}^{\omega_f} \left[S(\omega) A^2(\omega) \left| W(j\omega) \prod_{i=1}^{n} K_i(j\omega) - 1 \right|^2 + \right.$$

$$\left. + M^2(\omega) |W(j\omega)|^2 \sum_{k=1}^{n} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2 \right] d\omega . \tag{2.3}$$

Let us choose the transfer function of the decoding filter so that at fixed values of the transfer functions of all the encoding filters $K_i(j\omega)$; $i = 1, 2 \ldots n$ the deviation σ^2 (2.3) of the added weighted error be minimal. To achieve this aim the phase characteristic arg $W(j\omega)$ of the decoding filter has to be chosen thus that the transfer function of the entire connection be a real number, i.e.

$$W(j\omega) \prod_{i=1}^{n} K_i(j\omega) = \left| W(j\omega) \prod_{i=1}^{n} K_i(j\omega) \right| . \tag{2.4}$$

The optimal value of the transfer function $W_0(j\omega)$ of the decoding filter can be found by writing the Euler equation for the function $R[W(j\omega)]$ in the integral in expression (2.3):

$$\partial R[W(j\omega)]/\partial W(j\omega) = 0 \tag{2.5}$$

and solve this equation so that the condition (2.4) should be met.

From expressions (2.3)–(2.5):

$$W_0(j\omega) = \frac{S(\omega) \prod_{i=1}^{n} K_i^*(j\omega)}{S(\omega) \prod_{i=1}^{n} |K_i(j\omega)|^2 + L^2(\omega) \sum_{k=1}^{n} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2}, \tag{2.6}$$

where the functions $K_i^*(j\omega)$ are the conjugates of functions $K_i(j\omega)$ and

$$L(\omega) = \frac{M(\omega)}{A(\omega)}.$$

By substituting (2.6) into (2.3) we obtain for the minimal value of the deviation of the added weighted error σ_0^2, keeping the value of all the transfer functions $K_i(j\omega)$ of the encoding filters fixed:

$$\sigma_0^2 = \frac{1}{2\pi} \int_{\omega_a}^{\omega_f} F[|K_i(j\omega)|]\, d\omega =$$

$$= \frac{1}{2\pi} \int_{\omega_a}^{\omega_f} \frac{S(\omega) M^2(\omega) \sum_{k=1}^{n} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2}{S(\omega) \prod_{i=1}^{n} |K_i(j\omega)|^2 + L^2(\omega) \sum_{k=1}^{n} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2}. \tag{2.7}$$

3. Synthesis of the optimal transfer function of coding filters

We shall consider the mean power Q_s of the arbitrary transmitter with order number S (intermediate or end) to be limited, i.e.

$$Q_s = \frac{1}{2\pi} \int_{\omega_a}^{\omega_f} \Phi_s[|K_i(j\omega)|]\, d\omega =$$

$$= \frac{1}{2\pi} \int_{\omega_a}^{\omega_f} \left[S(\omega) \prod_{i=1}^{s} |K_i(j\omega)|^2 + \sum_{k=1}^{s-1} N_k(\omega) \prod_{i=k+1}^{s} |K_i(j\omega)|^2 \right] d\omega, \tag{3.1}$$

where $s = 1, 2, \ldots, n$. Thus the optimal value of the modules $|K_r(j\omega)|_0$, $r = 1, 2, \ldots, n$ of the transfer function of the coding filters, obtained at the limited mean power of the transmitters Q_s according to (3.1), the minimum of the deviation of the weighted error σ_0^2 according to (2.6) can be found by

solving the isoperimetric problem of variation calculus [2]. Therefore the following set of n Euler equations has to be established:

$$\frac{\partial}{\partial |K_r(j\omega)|} \left\{ F[|K_i(j\omega)|] + \sum_{s=1}^{n} \lambda_s \Phi_s [|K_i(j\omega)|] \right\} = 0 \Bigg|_{r=1,2,\ldots n},$$

where $F[|K_i(j\omega)|]$ and $\Phi_s[|K_i(j\omega)|]$ are the functions in the integral in (2.7) and (3.1), λ_s are certain constant multipliers obtained by substituting the optimal absolute value of the transfer functions $|K_i(j\omega)_0|$ of the coding filters into the initial conditions (3.1) ($i = 1, 2, \ldots, n$). At the same time the phase characteristics arg $[K_i(j\omega)]$; $i = 1, 2, \ldots n$ of the coding filters can be chosen arbitrarily, as the fulfilment of condition (2.4) is assured by the optimal choice of the phase characteristic of the decoding filter, i.e. of arg $W_e(j\omega)$.

From equations (2.7), (3.1), (3.2) after differentiation the following set consisting of n algebraic equations is obtained for the determination of the optimal absolute value of the transfer function of the coding filters:

$$\frac{S^2(\omega)\, M^2(\omega) \prod_{i=1}^{n} |K_i(j\omega)|^2 \sum_{k=r}^{n} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2}{\left[S(\omega) \prod_{i=1}^{n} |K_i(j\omega)|^2 + L^2(\omega) \sum_{k=1}^{n} N_k(\omega) \prod_{i=k+1}^{n} |K_i(j\omega)|^2\right]^2} = \tag{3.3}$$

$$= \sum_{s=r}^{n} \lambda_s \left[S(\omega) \prod_{i=1}^{s} |K_i(j\omega)|^2 + \sum_{k=1}^{r-1} N_k(\omega) \prod_{i=k+1}^{s} |K_i(j\omega)|^2 \right] \Bigg|_{r=1,2,\ldots,n},$$

where, for $k = n$, $\prod_{i=k+1}^{n} |K_i(j\omega)|^2 = 1$, $\sum_{k=1}^{r-1} N_k(\omega) \prod_{i=k+1}^{n} |K_1(j\omega)|^2 = 0$, for $r = 1$.

During the solution of the equation set (3.3) the trivial condition $|K_i(j\omega)|_0 \geq 0$ for all $i = 1, 2, \ldots, n$ has to be considered. Concretely for $n = 1$, that is on a connection operating without the retransmission of the signal, it follows from (3.3) that

$$|K_1(j\omega)|_0^2 = \begin{cases} \sqrt{\dfrac{N_1(\omega)\, M^2(\omega)}{\lambda_1 S(\omega)}} - \dfrac{L_2(\omega)\, N_1(\omega)}{S(\omega)} & \text{for} \quad h_1(\omega) \leq 1, \\ 0 & \text{for} \quad h_1(\omega) \leq 1, \end{cases} \tag{3.4}$$

where

$$h_1(\omega) = \frac{\lambda_1 L^4(\omega)\, N_1(\omega)}{S(\omega)\, M^2(s)}.$$

Let us substitute expression (3.4) into formulae (2.7) and (3.1), then we find the minimum of the deviation $\sigma^2_{\min}$ of the added weighted error for the optimal transfer functions of the decoding filter $W_0(j\omega)$ (2.6) and of the

final coding filter $|K_1(j\omega)|_0$ (3.4), as well as the constant λ_1 for the connection operating without the retransmission of the signal (for $n = 1$):

$$\delta^2_{\min} = \frac{1}{2\pi} \int\limits_{\Delta\omega_1} \sqrt{\lambda_1 S(\omega)\, N^1(\omega)\, M^2(\omega)}\, d\omega + \frac{1}{2\pi} \int\limits_{\Delta\omega_2} S(\omega)\, A^2(\omega)\, d\omega, \tag{3.5}$$

$$\sqrt{\lambda_1} = \frac{\int\limits_{\Delta\omega_1} \sqrt{N_1(\omega)\, S(\omega)\, M^2(\omega)}\, d\omega}{\int\limits_{\Delta\omega_2} L^2(\omega)\, N_1(\omega)\, d\omega + \int\limits_{\omega_a}^{\omega_f} S(\omega)\, d\omega}, \tag{3.6}$$

where $\Delta\omega_1$ and $\Delta\omega_2$ are frequency ranges located in the transmission band $\Delta\omega = \omega_f - \omega_a = \Delta\omega_1 + \Delta\omega_2$ of the connection, suitably determined by the inequality $h_1(\omega) \leq 1$.

Connections (3.4), (3.5) and (3.6) had been obtained earlier: in the most general form in [1]; for $L(\omega) = M(\omega) = 1$ in [3, 4]; for $L(\omega) = 0$ in [5]. These will be used further on for comparison with the case of multiple retransmission of the signal.

Unfortunately in the general case ($n \geq 2$) the system of equations (3.3) cannot be radically solved. Here some approximations can be applied [6].

4. A special case

In order to simplify the set of equations (3.3) we shall require that in expression (3.1) the values of the mean power of the transmitters under the integral be equal for all values $s = 1, 2, \ldots, n$, that is we shall consider that

$$\Phi_s[|K_i(j\omega)|] = \Phi_1[|K_i(j\omega)|] = S(\omega)\, |K_1(j\omega)|^2 \quad s = 2, 3, \ldots, n. \tag{4.1}$$

It can be seen that for the case of fulfilment of condition (5.5) the mean power of the final and all intermediate transmitters have to be equal with each other, that is

$$Q_s = Q_1 = \frac{1}{2\pi} \int\limits_{\omega_a}^{\omega_f} S(\omega)\, |K_1(j\omega)|^2\, d\omega, \quad s = 2, 3, \ldots, n. \tag{4.2}$$

On the other hand it follows from (3.1) that

$$\Phi_s\,[|K_i(j\omega)|] = S(\omega) \prod_{i=1}^{s} |K_i(j\omega)|^2 + \sum_{k=1}^{s-1} N_k(\omega) \prod_{i=k+1}^{s} |K_i(j\omega)|^2 =$$

$$= |K_s(j\omega)|^2\, [S(\omega) \prod_{i=1}^{s-1} |K_i(j\omega)|^2 + \sum_{k=1}^{s-2} N_k(\omega) \prod_{i=k+1}^{s-1} |K_i(j\omega)|^2 +$$

$$+ N_{s-1}(\omega)] = |K_s(j\omega)|^2\, [\Phi_{s-1}[|K_i(j\omega)|] + N_{s-1}(\omega)]. \tag{4.3}$$

Let us substitute condition (4.1) into formula (4.3), thus we obtain the connection between the absolute value of the transfer functions of the final and intermediate filters (number $s = i$):

$$|K_s(j\omega)|^2 = \frac{S(\omega)\,|K_1(j\omega)|^2}{S(\omega)\,|K_1(j\omega)|^2 + N_{s-1}(\omega)}, \quad s = 2, 3, \ldots, n. \tag{4.4}$$

For the sake of simplicity from here on we shall consider that on all sections of the connection the energy spectra of the noises coincide, that is:

$$N_s(\omega) = N_i(\omega) = N(\omega), \qquad s = i = 1, 2, \ldots, n. \tag{4.5}$$

In this case the absolute values of the transfer functions of all the transmitters are equal:

$$D(\omega) = |K_i(j\omega)|^2 = \frac{S(\omega)\,|K_1(j\omega)|^2}{S(\omega)\,|K_1(j\omega)|^2 + N(\omega)}, \quad i = 2, 3, \ldots, n. \tag{4.6}$$

Besides it is not difficult to show that conditions (4.1), (4.2) and (4.5) lead to the equality of all constants λ_s, that is

$$\lambda_s = \lambda; \qquad s = 1, 2, \ldots, n. \tag{4.7}$$

By substituting connections (4.6) and (4.7) into (3.3) and considering

$$\sum_{k=1}^{n} N_k(\omega)\,|K_i(j\omega)|^2 = N(\omega) \sum_{k=1}^{n} \prod_{i=k+1}^{n} D(\omega) = N(\omega) \sum_{k=1}^{n} D^{n-k}(\omega) =$$

$$= N(\omega)\,\frac{D^n(\omega) - 1}{D(\omega) - 1}$$

as the sum of n elements of the geometric series, we obtain the following equation for the determination of the absolute value $|K_1(j\omega)|_0$ of the optimal transfer function of the final coding filter:

$$\frac{N(\omega)\,M^2(\omega)}{\lambda S(\omega)}\,|K_1(j\omega)|^{2(n-1)} = \left[|K_1(j\omega)|^2 + \frac{N(\omega)}{S(\omega)}\right]^{n+1} \times$$

$$\times \left\{L^2(\omega) + [1 - L^2(\omega)]\,\frac{|K_1(j\omega)|^{2n}}{\left[K_1(j\omega)|^2 + \dfrac{N(a)}{S(\omega)}\right]^n}\right\}^2. \tag{4.8}$$

Equation (4.8) can be solved in close form for the values $n = 1$, 2 and 3. For other values of n some approximation method can be applied to obtain a solution [6]. However, the equation becomes much more simple if we choose the frequency weighted criterion ($L(\omega) = 1$) or the criterion of maximum

signal to noise ratio ($L(\omega) = 0$) [1]. The equation can be solved in closed form for $n = 1, 2, 3, 5$ and 7.

By assigning concrete values to n and $L(\omega)$, the following are obtained from equation (4.8):

1. For $L(\omega) = 0$ and $n = 2$

$$|K_1(j\omega)|_0^2 = \sqrt{\frac{4N(\omega)\,M^2(\omega)}{3\lambda S(\omega)}} \begin{bmatrix} \cos\dfrac{\varphi_1(\omega)}{3}, & \text{if } h_2(\omega) \leq 1 \\ \operatorname{ch}\dfrac{\varphi_2(\omega)}{3}, & \text{if } h_2(\omega) > 1 \end{bmatrix},$$

where $\cos\varphi_1(\omega) = \operatorname{ch}\varphi_2(\omega) = h_2(\omega)$; $h_2(\omega) = \sqrt{\dfrac{27\lambda N(\omega)}{4S(\omega)\,M^2(\omega)}}$.

2. Supposing that $L(\omega) = 0$ and $n = 3$

$$|K_1(j\omega)|_0^2 = \sqrt{\frac{N(\omega)\,M^2(\omega)}{\lambda S(\omega)}} \left[\frac{1}{2} + \sqrt{\frac{1}{4}\sqrt{\frac{\lambda N(\omega)}{S(\omega)\,M^2(\omega)}}}\right]. \tag{4.10}$$

3. For $L(\omega) = 1$ and $n = 3$

$$|K_1(j\omega)|_0^2 = \begin{cases} \left(\sqrt{\dfrac{N(\omega)\,M^2(\omega)}{\lambda S(\omega)}} - \dfrac{N(\omega)}{S(\omega)}\right)\left[1 + \dfrac{\sqrt{\dfrac{1}{4} - B(\omega)} - \dfrac{1}{2}}{1 - B(\omega)}\right], & \text{if } B(\omega) \leq \dfrac{1}{4} \\ 0, \quad \text{if } B(\omega) > \dfrac{1}{4} \end{cases} \tag{4.11}$$

where $B(\omega) = \sqrt{\dfrac{\lambda N(\omega)}{S(\omega)M^2(\omega)}}$.

Similar (though more cumbersome) expressions are obtained for $|K_1(j\omega)|_0^2$ for other values of n and $L(\omega)$.

In formulae (4.9), (4.10) and (4.11) the first member of the product with the precision of the constant multiplier corresponds to the square of the absolute value of the optimal transfer function of the final coding filter on the connection operating without the retransmission of the signal ($n = 1$), the second product (in square brackets) characterises the correction which takes the retransmission of the signal into account. However it follows from (4.8) that in the general case this correction can be neglected only on the condition $S(\omega)\,M^2(\omega) \gg n^2\lambda N(\omega)$, that is for the case of very small noises only.

From equation (4.8) for the optimal value of $|K_1(j\omega)|_0$, the absolute value of the transfer function of the final coding filter, it is not difficult to

find the square of the absolute value $D_0(\omega)$ and $|W_0(j\omega)|^2$ of the optimal transfer functions of the intermediate coding filters (retransmitters) and of the final decoding filter, and to calculate the smallest deviation $\sigma^2_{\min}$ of the weighted resultant error for the optimal transfer function of all coding and decoding filters. Applying for this purpose equations (2.6), (2.7), (4.6) and(4.8):

$$D_0(\omega) = |K_i(j\omega)|_0^2 = \frac{S(\omega)\,|K_1(j\omega)|_0^2}{S(\omega)\,|K_1(j\omega)|_0^2 + N(\omega)}\,, \qquad i = 2, 3, \ldots, n,$$

$$|W_0(j\omega)|^2 = \frac{\lambda S(\omega)\,|K_1(j\omega)|_0^2}{N(\omega)\,M^2(\omega)}\,,$$

$$\sigma^2_{\min} = \frac{1}{2\pi}\int\limits_{\omega_a}^{\omega_f} \frac{S(\omega)\,M^2(\omega)\,[1 - D_0^n(\omega)]}{L^2(\omega) + [1 - L^2(\omega)]\,D_0^n(\omega)}\,d\omega\,.$$

The constant λ can be found by substituting the concrete value of the function $|K_1(j\omega)|_0^2$ into the initial condition (4.2).

5. Conclusion

In the present paper we have examined the task of optimal linear construction of the coding and decoding filters on connections operating with multiple retransmission of the signal. As optimisation criterion the most general frequency weighted mean square criterion was applied with unequal weighting functions for the dynamic and random errors [1], the special cases of which are the widely used criterion of minimal mean square error [3, 4] and the criterion of maximal signal to noise ratio [5].

The general expressions (2.6) and (2.7) were obtained for the optimal transfer function of the final decoding filters and for the minimal deviation of the weighted error in the case when the transfer functions of all coding filters are fixed. With the help of these expressions the problem of optimal filtration of the signal can be solved at the final receiving station.

A variation task was formulated to find the optimal transfer function of the intermediate and final coding filters for the case of limited mean power of the transmitter. The system of algebraic equations (3.3) was constructed with which these transfer functions can be calculated.

The practically important case when the mean power of all the transmitters is equal and the energy spectra coincide on all sections of the connection was examined in detail. For this case relatively simple expressions were found for the optimal transfer functions of the coding and decoding

filters. At the same time the frequency curves of the retransmitters have to be uniform for very small noises only when the condition $S(\omega)\,M^2(\omega) \gg \lambda n^2 N(\omega)$ is satisfied. Otherwise the shape of the optimal frequency curve of the retransmitters is significantly far from being uniform in the range of the working frequencies. This latter circumstance has to be taken into consideration when designing optimal connections operating with the multiple retransmission of the signal.

References

1. *Миленин, Н. К.*, Линейное кодирование в системе передачи информации с дополни, тельными источниками шумов. II Международный симпозиум по теории информации- г. Цахкадзор, Арм. ССР. Тезисы докладов, Москва—Ереван, 1971, стр. 175—178.
2. *Эльсгольц Л. Э.* Дифференциальные уравнения и вариационное исчисление. «Наука», М., 1969.
3. *Costas, J. P.*, Coding with Linear Systems. Proc. IRE, 1952, **40**, *9*, 1101—1103.
4. *Цыбаков Б. С.* Линейное кодирование сообщений. Радиотехника и электроника, **7** (1962) *1*, 25—38.
5. *Штейн, В. М.* О расчете предыскажающих и корректирующих устройств. Радиотехника, **11** (1956) *2*, 60—63.
6. *Мишина А. П.* и *Проскуряков И. В.* Высшая алгебра. «Физматгиз», М., 1962.

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

NOTE TO CONTRIBUTORS

Two copies of the manuscripts (each duly completed by figures, tables and references) are to be sent either to

E. D. Teryaev coordinating editor

Department of Mechanics and Control Processes

Academy of Sciences of the USSR

Leninsky Prospect 14, Moscow V-71, USSR

or to

J. Kocsis coordinating editor

Department of Automation

Technical University

H–1111 Budapest XI., Goldman György tér 3, Hungary

The authors are requested to retain a third copy of the submitted typescript to be able to check the proofs against it.

The papers, preferably in English or Russian, should be typed double-spaced on one side of good-quality paper, with wide margins (c. 4–5 cm) should carry the title of the contribution, the author(s)' name, and the name of the country. At the end of the typescript the name of that author who manages the proof-reading should also be given.

An abstract of about 50 to 100 words should head the paper.

The authors are encouraged to use the following headings: Introduction, (outlining the problem), Methods and results, Discussion, Conclusions, References. The entire material should not exceed 15 pages including tables and references. The proper location of the tables and figures must be indicated on the margins. Mathematical notations should follow up-to-date usage.

The summary – possibly in Russian if the paper is in English and *vice versa* – should contain a brief account of the proposition and indications of the formulas used and figures shown in the paper. The summary is not supposed to exceed 10–15 per cent of the paper.

The authors will be sent sheet-proofs which they are to return by next mail to the sender Regional Editorial Board.

Authors are entitled to 100 reprints free of charge. Rejected manuscripts will be returned to the authors.

К СВЕДЕНИЮ АВТОРОВ

Рукописи в двух экземплярах (каждый из которы должен содержать рисунки, таблицы и литературу направляются

Е. Д. Теряеву — Научный секретарь журнала

Отделение механики и процессов управления

Академия Наук СССР

Ленинский Проспект, 14, Москва В-71, СССР

или

Я. Кочишу — Научный секретарь журнала

Кафедра Автоматизации
Будапештского Технического
Университета
Н–1111 Будапешт XI, площадь Голдман Д., 3,
Венгрия

Авторам рекомендуется оставлять у себя копию всех представленных ими материалов для справок при корректуре.

Статьи, желательно на русском или английском языках, отпечатанные на бумаге хорошего качества, с промежутком в два интервала и широкими (4–5 см) полями должны содержать наименование статьи, фамилию автора (авторов), название страны. В конце статьи необходимо также указать фамилию автора, ответственного за корректуру гранок.

Статье должна предшествовать аннотация объемом до 50–100 слов.

Авторы при написании статьи должны придерживаться следующей формы: введение (постановка задачи), основное содержание и результаты, обсуждение, выводы и литература. Объем статьи не должен превышать 15 печатных страниц, включая таблицы и ссылки. Последовательность таблиц и рисунков должна быть отмечена на полях. Математические обозначения рекомендуется давать в соответствии с современными требованиями и традициями.

К статье обязательно должно быть приложено резюме-реферат. Резюме — на русском языке, если статья написана на английском, и наоборот — должно содержать краткое изложение текста статьи со ссылками на необходимые формулы и графики, имеющиеся в основном тексте. Объем разюме не должен превышать 10–15% объема статьи.

Авторам высылаются гранки статьи, которые они должны незамедлительно возвратить в Региональную секцию Редколлегии журнала.

Авторам обеспечивается бесплатно 100 оттисков их статей. Рукописи непринятых статей возвращаются авторам.

CONTENTS • СОДЕРЖАНИЕ

Пупков К. А.—Капалин В. И.: Применение рядов Вольтерра для исследования нелинейных систем (*Pupkov, K. A.—Kapalin, V. I.:* The use of Volterra series in nonlinear system designing) 105

Оганесян С. Ш.—Таирян В. И.—Ягджян В. Г.: Разбиение циклических кодов на равновесные классы (*Oganesyan, S. Sh.—Tairyan, V. I.—Yagdzyan, V. G.:* Decomposition of cyclic codes into equal-weight classes) 117

Миленин Н. К.: Линейное кодирование при ретрансляции сигнала (*Milenin, N. K.:* Linear coding in the retransmission of the signal) 127

Kartsatos, A. G.: Global controllability of perturbed quasi-linear systems (*Картсатос А. Г.:* Управляемость в целом квазилинейных систем в условиях возмущений) 137

Nemetz, T.: On the α-divergence rate for Markov-dependent hypotheses (*Немец Т.:* Об асимптотическом исследовании α-расходимости мер, соответствующих двум однородным марковским процессам) 147

Komáromi, É.: A finite primal method for the assignment problem) (*Комароми Е.:* Конечный примальный метод для решения транспортной задачи) 157

Beck, J.: On the entropy of the sum of IID random variables (*Бек Й.:* Об энтропии одинаково распределенных независимых случайных величин) 167

316.920

VOL. / ТОМ 3 * NUMBER / НОМЕР 3

ACADEMY OF SCIENCES OF THE USSR
HUNGARIAN ACADEMY OF SCIENCES

PROBLEMS OF CONTROL AND INFORMATION THEORY

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

АКАДЕМИЯ НАУК СССР
АКАДЕМИЯ НАУК ВЕНГРИИ

1974

AKADÉMIAI KIADÓ, BUDAPEST

PROBLEMS OF CONTROL AND INFORMATION THEORY

is an international quarterly sponsored jointly by the Presidium of the Academy of Sciences of the USSR and of the Hungarian Academy of Sciences. It offers publicity for original papers and short communications of the following topics:

— general theory of control systems and system theory

— theory of automata

— information theory

— operation research; theory of complex systems

— theory of economic control; system modelling

— theory and methods of adaptation, learning, identification and pattern recognition

— methods of information processing; application of digital computers in control and communication systems

— new physical principles in constructing technical devices for automation, control and information processing

The four issues published per year make up a volume of some 320 pp.

While this quarterly is mainly a publication forum of the research results achieved within the USSR and Hungary, also papers of international interest from other countries are welcome.

The quarterly is published by
AKADÉMIAI KIADÓ
Publishing House of the Hungarian Academy of Sciences
H—1363 Budapest P.O. B. 24, Hungary

Президиумами Академии Наук СССР и Академии Наук ВНР было принято решение о совместном издании журнала

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

Журнал создан для обеспечения быстрейшей публикации материалов о новейших результатах научных исследований, разработок и кратких сообщений о достижениях в следующих областях науки:

— общая теория процессов управления и теория систем
— теория автоматов
— теория информации
— исследование операций; теория сложных систем
— теория управления экономическими системами; модели систем
— теория и методы адаптации, обучения, идентификации и распознавания образов
— методы обработки информации; применение вычислительных машин в системах управления и передачи информации
— новые физические принципы создания технических средств автоматики, управления и передачи информации

Журнал издается четыре раза в год общим объёмом около 320 печатных страниц.

Журнал в первую очередь будет публиковать научные достижения обеих стран, а также статьи из других стран, представляющие международный интерес.

Издатель журнала:
AKADÉMIAI KIADÓ
Издательство Академии Наук Венгрии
H—1363 Будапешт п. о. 24, Венгрия

Distributor for:

Albania, Bulgaria, China, Cuba, Czechoslovakia, German Democratic Republic, Korean People's Republic, Mongolia, Poland, Roumania, USSR, People's Republic of Viet-Nam, Yugoslavia

KULTURA
Hungarian Trading Co. for Books and Newspapers
H—1389 Budapest 62, P.O.B. 149, Hungary

Distributor for all other countries:

PLENUM PUBLISHING CO. LTD
4a Lower John Street
London. W1R 3PD
ENGLAND

Annual subscription price: US $ 35,– plus postage of US $1,80

PROBLEMS OF CONTROL AND INFORMATION THEORY

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

EDITORS

B. N. PETROV (Moscow)
F. CSÁKI (Budapest)

DEPUTY EDITORS

V. S. PUGACHEV (Moscow)
V. I. SIFOROV (Moscow)
S. CSIBI (Budapest)

COORDINATING EDITORS

S. V. EMELIANOV (Moscow)
L. KALMÁR (Szeged)

M. A. GAVRILOV (Moscow)
I. CSISZÁR (Budapest)

A. M. LETOV (Moscow)
A. PRÉKOPA (Budapest)

L. VARGA (Budapest)

E. D. TERYAEV (Moscow)
J. KOCSIS (Budapest)

РЕДАКТОРЫ ЖУРНАЛА

Б. Н. ПЕТРОВ (Москва)
Ф. ЧАКИ (Будапешт)

ЗАМЕСТИТЕЛИ РЕДАКТОРОВ

В. С. ПУГАЧЕВ (Москва)
В. И. СИФОРОВ (Москва)
Ш. ЧИБИ (Будапешт)

ЧЛЕНЫ РЕДАКЦИОННОЙ КОЛЛЕГИИ

С. В. ЕМЕЛЬЯНОВ (Москва)
Л. КАЛМАР (Сегед)

М. А. ГАВРИЛОВ (Москва)
И. ЧИСАР (Будапешт)

А. М. ЛЕТОВ (Москва)
А. ПРЕКОПА (Будапешт)

Л. ВАРГА (Будапешт)

Е. Д. ТЕРЯЕВ (Москва)
Я. КОЧИШ (Будапешт)

AKADÉMIAI KIADÓ

PUBLISHING HOUSE OF THE HUNGARIAN ACADEMY OF SCIENCES

BUDAPEST

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

Problems of Control and Information Theory, Vol. 3 (3), pp. 171—182 (1974)

GAME-THEORETIC CONTROL AND PROBLEMS OF STABILITY

N. N. KRASOVSKII

(Sverdlovsk)

(Received February 5, 1974)

The paper is devoted to the discussion of control problems for an ordinary dynamical system under conditions of conflict or uncertainty. A stable game-theoretic scheme of control is constructed with a theoretical foundation being described. The scheme which includes an imitating model plant — the so-called leader — is based on a mathematical formalization of the problems under consideration as introduced in papers [1, 2]. However in the sequel the "control with a leader" of papers [2, 3] arrives in the form of a refined and stabilized procedure. Thus the scheme proposed in this paper for the construction of the control actions unites the methods of the theory of stability of motion with those of the theory of differential games.

1. A considerable number of problems of control under conditions of conflict or uncertainty for systems described by ordinary differential equations may be reduced to the following mathematical model of the approach-evasion game. Assume that the motions of the dynamical system under consideration are determined by the vector differential equation

$$\dot{x} = f(t, x, u, v) , \tag{1.1}$$

where x is the n-dimensional phase vector of the system, u and v are the r-dimensional and n-dimensional vectors of the control actions, governed by the first and second *players* respectively and subjected to the constraints

$$u \in P , \quad v \in Q . \tag{1.2}$$

Here P and Q are closed and bounded sets in the corresponding vector spaces $\{u\}$ and $\{v\}$. (In all the coming formulas involving vectors the latter are considered to be vector columns if no special reservation is made.)

The conditions of the problem include the description of closed sets M and N in the space $\{t, x\}$, the first of which is the *target* for the first player and the second one constitutes his *living space*. The problem for the first player is to specify a *strategy* which ensures that motion $x[t]$ encounters set M under the condition that it remains within set N until the moment τ of encounter, no matter what are the actions of the second player.

When speaking of the given conflict control problem it is necessary to underline that the notions of the first and second player are a reflection of

actually existing counteracting sides which both govern the system and which tend to directly opposite aims. On the other hand if the given problem is that of a single-sided single-directed control (e.g. by means of force u), complicated however by an uncertain disturbance (e.g. by the force v) then the role of the opposing player is addressed to a circumstantial action which reflects a blindfold formation of the disturbance.

In this article we shall restrict ourselves to the discussion of the problem for the first player which forms the control u on base of the information on the state space variables for system (1.1) only and which does not utilize any stochastic procedures of control. Then according to articles [1, 4] his behaviour in the *formalized* game will be determined by pure strategies $U \div u(t, x)$ each of which is identified with the function $u(t, x) \in P$. Then the *motions* $x[t, t_*, x_*, U]$ $(t_* \leq t \leq \vartheta^*)$, derived by the strategy $U \div u(t, x)$ from position $\{t_*, x_*\}$ are defined formally as the limits for the uniformly converging sequences of *Euler broken lines* $x_\Delta[t, t_*, x_*^{(k)}, U, v_*^{(k)}[\,\cdot\,]]$ $(k = 1, 2, \ldots)$ each of which is the solution of the differential equation

$$\dot{x}_\Delta = f(t, x_\Delta, u(\tau_i, x_\Delta[\tau_i]), v[t]) \qquad (1.3)$$
$$(\tau_i \leq t \leq \tau_{i+1}) .$$

Here the *division* $\{\tau_i\}$ $(i = 0, 1, 2, \ldots;\ \tau_0 = t_*)$ is a certain unboundedly increasing sequence, $v[t]$ $(t_* \leq t \leq \vartheta_*)$ is a measurable function of t with range Q. In each of the converging sequences $x_\Delta^{(k)}[t] = x_\Delta[t, t_*, x_*^{(k)}, U, v^{(k)}[\,\cdot\,]]$ $(k = 1, 2, \ldots)$ which determine the limiting motions the realizations $v^{(k)}[t]$ may vary in general with the variation of k. The division $\{\tau_i^{(k)}\}$ must satisfy the condition

$$\lim_{k \to \infty} \delta^{(k)} = \limsup_i (\tau_{i+1}^{(k)} - \tau_i^{(k)}) = 0. \qquad (1.4)$$

In order to estimate the abilities of the second player in the game under consideration it is possible to make use of the notion of *counterstrategy* $V^{(u)} \div \div v \subset t, x, u)$ [1, 4]. We shall identify the latter with the function $v \subset t, x, u) \in Q$ which satisfies the condition that with $u = u[t]$ being a measurable function on $t_* \leq t \leq \vartheta^*$ with range P the function $v[t] = v(t^*, x^*, u[t])$ comes to be measurable whatever are the fixed values t^* and x^*. Here *the motions* $x[t, t_*, x_*, V^{(u)}]$ $(t_* \leq t \leq \vartheta^*)$ derived by the counterstrategy $V^{(u)} \div v(t, x, u)$ from position $\{t_*, x_*\}$ are formally defined as the limits of converging sequences of *Euler broken lines* $x_\Delta[t, t_*, x_*^{(k)}, V^{(u)}, u^{(k)}[\cdot]]$ $(k = 1, 2, \ldots)$ each of which is the solution of the differential equation

$$\dot{x}_\Delta = f(t, x_\Delta, u[t], v(\tau_i x_\Delta[\tau_i], u[t])) \qquad (1.5)$$
$$(\tau_i \leq t < \tau_{i+1}) .$$

In the abstract framework of the *ideal* motions $x[t, t_*, x_*, U]$ and $x[t, t_*, x_*, V^{(u)}$ the following main theorem *on the alternative* is true (see e.g. the versions of the theorem in [1, 4]).

Whatever is the initial position $\{t_0, x_0\}$ and the moment of time $\vartheta > t_*$ there either exists a strategy U^0 which guarantees the encounter

$$\{\tau, x[\tau]\} \in M, \ \{t, x[t]\} \in N, \tag{1.6}$$
$$(t_0 \leq t \leq \tau \leq \vartheta)$$

with set M within N when $\tau \leq \vartheta$, for any motion $x[t] = x[t, t_0, x_0, U^0]$ or there exist certain neighbourhoods $\mathfrak{I}(M)$ and $\mathfrak{H}(N)$ of sets M and N and a counterstrategy $V^{(u)0}$ which excludes the encounter

$$\{\tau, x[\tau]\} \in \mathfrak{I}(M), \ \{t, x[t]\} \in \mathfrak{H}(N) \tag{1.7}$$
$$(t_0 \leq t \leq \tau \leq \vartheta)$$

with $\mathfrak{I}(M)$ within $\mathfrak{H}(N)$ when $\tau \leq \vartheta$ for any motion $x[t, t_0, x_0, V^{(u)0}]$.

In compliance with the abstract character of the ideal motions $x[t, t_*, x_*, U]$ and $x[t, t_*, x_*; V^{(u)}]$ the given theorem arrives as a strict assertion which characterizes the abstract mathematical model of the differential game. The aim of the present paper is to transform the given result and its abstract foundation into the terms of a practically realizable and stable control procedure.

2. The base of the control procedure described in the sequel is the introduction of a precision model into the closed loop of the control circuit of the first player. The model is to imitate the actual system (1.1) and is realizable by means of a computer technique. Within the model which is characterized by its own state space vector w and which includes the control actions u^* and v^* a certain game in between u^* and v^* does hold. Thus we arrive at two games held simultaneously (as on "two boards") — one in the actual system where the first player plays with its antagonist, the second player, and another in the model plant where the first player plays with itself. Under the first condition of the alternative the control procedure proposed here admits the specification of the controls $\{u, u^*, v^*\}$ governed by the first player so as to achieve two aims. The first is to transfer the state space vector $w[t]$ of the model plant to the state desired so that

$$\{\tau, w[\tau]\} \in M, \quad \{t, w[t]\} \in N, \quad (t_0 \leq t \leq \tau \leq \vartheta). \tag{2.1}$$

At the same time the second aim is to force the actual motion $x[t]$ and the motion $w[t]$ of the model plant to track each other. Hence the actual

motion $x[t]$ is held within a minor neighbourhood of N until the moment τ and at moment τ it is transferred to a sufficiently small neighbourhood of the target set M. The practical aim of the control process is thus achieved. It is noteworthy to say however that a similar procedure of control with a leader $w[t]$ as described in [2, 3] has a certain drawback. This is due to the appearance of a multiplier of type $\exp \lambda(t - t_0)$ in the estimate for the deviation $x[t] - w[t]$.

If the transitional control process holds for a sufficiently large period of time the given estimate may happen to be inadequate. Hence we shall introduce a refined form of control with a leader defined with the application of methods of the theory of stabilization of motions [5] (see Supplement IV).

We shall confine ourselves to the case when the equations of motion (1.1) may be presented in the form

$$\dot{x} = A(t)x + f(t, u, v) + \mu g(t, x, u, v) , \tag{2.2}$$

where $A(t)$ is a continuous matrix function, f and g are continuous and the function $g(t, x, u, v)$ is uniformly bounded for all t, x, u, v; μ is a small parameter.

For the *model* plant we shall select the system characterized by an equation

$$\dot{w} = A(t)w + f^*(t, u^*, v^*) , \tag{2.3}$$

which approximates equation (2.2). Here f^* is also continuous and the control vectors u^* and v^* are restricted by the constraints

$$u^* \in P^*, \quad v^* \in Q^*. \tag{2.4}$$

The sets P^* and Q^* are closed and bounded in the vector spaces $\{u^*\}$ and $\{v^*\}$. For the sake of determination we assume the matrices $A(t)$ in (2.2) and (2.3) to be the same, although in general this assumption is not inevitable.

We finally assume that the information on the current values of the state space vector $x[t]$ which arrives at the controller of the first player results in an inaccurate value $x^*[t]$ so that only the estimate

$$||x[t] - x^*[t]|| \leq \xi \tag{2.5}$$

is known exactly. (The symbol $||x||$ stands for the Euclidean norm of x.) The information on the current values of $w[t]$ is assumed to be accurate.

3. The main assumptions for the validity of the procedure proposed consist in the following.

Condition A. This is said to be fulfilled if it is possible to select a continuous matrix function $B(t) = \{b_{ij}(t)\}$ $(i = 1, \ldots, n;\ j = 1, \ldots, k)$ such that for $t \geq t_0$ the following two conditions are fulfilled.

($\mathrm{I_A}$) The system described by the differential equation

$$\dot{y} = A(t)\,y + B(t)\,h \tag{3.1}$$

is stabilizable [5], i.e. it is possible to select a vector function

$$h = H(t)\,y \tag{3.2}$$

which after being substituted in (3.1) yields an asymptotically stable system

$$\dot{y} = A(t)\,y + B(t)H(t)\,y\ ; \tag{3.3}$$

($\mathrm{II_A}$) Whatever is the n-dimensional vector s of unit norm the following inequality is true

$$\min_{u^* \in P}\ \max_{v^* \in Q}\ s'f^*(t, u^*, v^*) - \min_{u \in P}\ \max_{v \in Q}\ s'f(t, u, v) \geq \alpha \max_{||h||=1} s'B(t)\,h, \tag{3.4}$$

where $\alpha > 0$ is constant. Here and everywhere below the prime denotes transposition. The essence of these conditions is in the point that the resources of control v^* in comparison with those of u^* in the model plant are greater in some sense connected with conditions of stabilizability than the resources of v when compared with those of u in the actual system.

Example. A simple illustration for the fulfilment of condition A is given by a system formed of two material points $m^{(1)}$ and $m^{(2)}$ whose motions are in the vicinity of each other and both lie in the neighbourhood of an equilibrated circular orbit of radius r_0, surrounding a fixed center of attraction 0 with mass M. Points $m^{(1)}$ and $m^{(2)}$ are controlled by forces $F^{(1)}$ and $F^{(2)}$ respectively each of which lies within the plane orthogonal to the radius vector $\bar{r}_*^{(i)}$ of the corresponding point $m^{(i)}$ the radius being traced from point 0. The forces $F^{(i)}$ are restricted by the inequalities $||F^{(i)}|| \leq \varkappa^{(i)}$. The first player which governs the force $F^{(1)}$ may appoint the direction of to accurate $F^{(1)}$ an uncontrollable clearance the latter being described by angle Ψ, subjected to the constraint $|\psi| \leq \psi_0 < \pi/2$.

Assume $r^{(i)}$ is the value of the projection of radius vector $\bar{r}_*^{(i)}$ for the ith point on plane Π of the basic orbit, $\varphi^{(i)}$ is the polar angle of this projection, $\eta^{(i)}$ is the coordinate of the ith point along the axis orthogonal to Π; r and φ are the polar coordinates for a fictitious point m, moving along a circular orbit. Denote $x_1 = r^{(1)} - r$, $x_2 = r_0(\varphi^{(1)} - \varphi)$, $x_3 = \eta^{(1)}$; $x_7 = r^{(2)} - r$, $x_8 = r_0(\varphi^{(2)} - \varphi)$, $x_9 = \eta^{(2)}$; $F_1^{(1)} = u_1 \cos v_3 - u_2 \sin v_3$, $F_2^{(1)} = u_1 \sin v_3 + u_2 \cos v_3$; $F_1^{(2)} = v_1$, $F_2^{(2)} = v_2$.

Then a standard reduction of equation (2.2) carried out in an appropriate scale leads to the system of equations

$$\begin{aligned}
\dot{x}_1 &= \dot{x}_4,\ \dot{x}_2 = x_5,\ \dot{x}_3 = x_6\,,\\
\dot{x}_4 &= \alpha x_1 + \beta x_5 + g_4\,,\\
\dot{x}_5 &= \gamma x_4 + u_1 \cos v_3 - u_2 \sin v_3 + g_5\,,\\
\dot{x}_6 &= \delta x_3 + u_1 \sin v_3 + u_2 \cos v_3 + g_6\,,\\
\dot{x}_7 &= x_{10},\ \dot{x}_8 = x_{11},\ \dot{x}_9 = x_{12}\,,\\
\dot{x}_{10} &= \alpha x_7 + \beta x_{11} + g_{10}\,,\\
\dot{x}_{11} &= \gamma x_{10} + v_1 + g_{11}\,,\\
\dot{x}_{12} &= \delta x_9 + v_2 + g_{12}\,.
\end{aligned} \tag{3.5}$$

where α, β, γ, δ are constants, the symbol g_i denotes members of higher order than the x'. The constraints on u and v obtain the form

$$[u_1^2 + u_2^2]^{\frac{1}{2}} \leq \varkappa^{(1)},\ [v_1^2 + v_2^2]^{\frac{1}{2}} \leq x^{(2)},\ |v_3| \leq \psi_0\,. \tag{3.6}$$

For the model plant satisfying Condition A we may select the system which imitates the motions of $m^{(1)}$ and $m^{(2)}$ within the linear approximation. We then arrive at the following equations of motion for the model

$$\begin{aligned}
\dot{w}_1 &= w_4,\ \dot{w}_2 = w_5,\ \dot{w}_3 = w_6\,,\\
\dot{w}_4 &= \alpha w_1 + \beta w_5,\ \dot{w}_5 = \gamma w_4 + u_1^*,\ \dot{w}_6 = \delta w_3 + u_2^*\\
\dot{w}_7 &= w_{10},\ \dot{w}_8 = w_{11},\ \dot{w}_9 = w_{12}\\
\dot{w}_{10} &= \alpha w_7 + \beta w_{11},\ \dot{w}_{11} = \gamma w_{10} + v_1^*,\ \dot{w}_{12} = \delta w_9 + v_2^*
\end{aligned} \tag{3.7}$$

with constraints

$$\begin{aligned}
&[u_1^{*2} + u_2^{*2}]^{\frac{1}{2}} \leq \varkappa^{(1)*},\\
&[v_1^{*2} + v_2^{*2}]^{\frac{1}{2}} \leq \varkappa^{(2)*}.
\end{aligned} \tag{3.8}$$

Let us select

$$\varkappa^{(1)*} < \varkappa^{(1)} \cos \psi_0,\ \varkappa^{(2)*} > \varkappa^{(2)}. \tag{3.9}$$

Then on the basis of well-known stabilizability conditions [5] it is possible to verify that Condition A is satisfied by matrix $\beta = \{b_{5,1} = b_{6,2} = b_{11,3} = b_{12,4} = 1$, $b_{ij} = 0$ for other values of $i, j\}$ with h selected as a four-dimensional vector $h = \{h_1, h_2, h_3, h_4\}$.

Surely the given model is efficient only if the motions of $m^{(1)}$ and $m^{(2)}$ are close enough to that of the fictitious point m.

4. Let us describe the construction of controls $\{u, u^*, v^*\}$ of the first player. We shall assume that these are formed with the aid of a digital computing device. Hence we shall construct the control procedure as a scheme discrete in time and determined by a certain division $\{\tau_i\}$ $(i = 0, 1, 2, \ldots;$ $\tau_0 = t_0)$ of the semiaxis $t_0 \leq t < \infty$.

The control actions u and v^* are determined from the condition of stable tracking of motions $x[t]$ and $w[t]$ by each other. This is carried out as follows. Assume that control (3.2) stabilizes the system (3.1) in the sense induced by Condition A. Then due to a well-known theorem of Lyapunov (see e.g. [5]) we find that for an arbitrary positive definite quadratic form

$$\xi(t, y) = \sum_{i,j=1}^{n} \xi_{ij}(t)\, y_i\, y_j, \tag{4.1}$$

there exists a positive definite form

$$\lambda(t, y) = \sum_{i,j=1}^{n} \lambda_{ij}(t)\, y_i\, y_j \tag{4.2}$$

whose total derivative $(d\lambda/dt)$ (3.3) along the solutions of systems (3.3), satisfies the equality

$$\left(\frac{d\lambda}{dt}\right)_{(3.3)} = -\xi(t, y). \tag{4.3}$$

Suppose the pair of functions $\xi(t, y)$ amd $\lambda(t, y)$ is fixed. Also suppose that at moment $t = \tau_i$ the actual system and the model produce the values $x^*[\tau_i]$ and $w[\tau_i]$. Denote

$$s[\tau_i] = \left[\frac{\partial\lambda}{\partial y}\right]_{(t=\tau_i,\, y=x_*[\tau_i]-w[\tau_i])} \tag{4.4}$$

where symbol $[\partial\lambda/\partial y]$ stands for the gradient vector $\{\partial\lambda/\partial y_i\}$ $(i = 1, \ldots, n)$. Now select the vector $u[\tau_i]$ and the vector-function $v^*[\tau_i, u^*]$ from the conditions

$$\max_{v \in Q} s'[\tau_i] f(\tau_i, u[\tau_i], v) = \min_{u \in P} \max_{\vartheta \in Q} s'[\tau_i] f(\tau_i, u, v). \tag{4.5}$$

$$s'[\tau_i] f^*(\tau_i, u^*, v^*[\tau_i, u^*]) = \max_{v^* \in Q^*} s'[\tau_i] f(\tau_i, u^*, v^*). \tag{4.6}$$

Assume that in the semiinterval $\tau_i \leq t < \tau_{i+1}$ the system (2.2) and the model plant (2.3) are influenced by the controls $u = u[\tau_i]$, $v^* = v^*[\tau_i, u^*]$ and also by a certain measurable control $v = v[t] \in Q$ and a certain piecewise constant control $u^* = u^*[t, \tau_i] \in P$. Then after making a calculation of the total derivative $(d\lambda/dt)_{(2.2)(2.3)}$ of function $\lambda[t] = \lambda \in t,\ x[t] - w[t])$ along the solutions of the so obtained equations of motion (2.2) and (2.3) and taking

into account the estimate (2.5), the condition (3.4), the equality (4.3) and the conditions (4.5) and (4.6) a number of standard calculations related to the theory of stability will bring us to the following estimate

$$\left(\frac{d\lambda}{dt}\right) \leq -\xi(t, x[t] - w[t]) + \varphi(\mu, \xi, \delta). \tag{4.7}$$

Here $\lim \varphi = 0$ with $\{\mu, \xi, \delta\} \to 0$, and $\delta = \sup_i (\tau_{i+1} - \tau_i)$. Under sufficiently small values of μ, ξ and δ the given estimate does ensure that motions $x[t]$ and $w[t]$ will lie close enough to each other in time t.

We will now describe the construction of the controls $u^*[t, \tau_i]$ $(\tau_i \leq \leq t < \tau_{i+1})$ in the model plant. Suppose that the initial position $\{t_0, x_0^*\}$ available for the first player is such that after selecting $\{t_0, w_0\} = \{t_0, x_0^*\}$ it is possible for the player to solve within the model plant (2.3) the problem of encounter with M within set N until the moment $\vartheta > t_0$. Then following paper [1] it is possible to assert that in the space $\{t, x\}$ there exists a *stable bridge* W which includes position $\{t_0, w_0\}$, which lies entirely in set N and whose end lies at moment ϑ in the set M. Let us assume that the bridge W is fixed. Denote with $L(t)$ the section acquired with the aid of a hyperplane $t = \text{const}$ for a set L in space $\{t, x\}$. Take $w[\tau_i] \in W(\tau_i)$. Then the stability property for the set W denotes according to paper [1] that with function $v^*[\tau_i, u^*]$ being given, it is possible for $\gamma > 0$ arbitrary and chosen in advance to select a piecewise constant function $u^* = u[t, \tau_i]$ $(\tau_i \leq t < \tau)$ such that the following property does hold: for any motion $w[t]$ of system (2.3) starting from position $\{\tau_i, w[\tau_i]\}$ and being directed by controls $u^* = u^*[t, \tau_i]$, $v^* = v^*[\tau_i, u^*[t, \tau_i]]$ it is either that

$$\varrho[w[t], W(t)] < \gamma, \quad \tau_i \leq t < \tau_{i+1}$$

or

$$\varrho[w[t], W(t)] < \gamma, \quad \tau \leq t \leq \tau \tag{4.8}$$

$$\varrho[w[\tau], M(\tau)] < \gamma \text{ with } \tau \in [\tau_i, \tau_{i+1}], \tag{4.9}$$

where symbol $\varrho[w, L(t)]$ stands for the distance from point w to set $L(t)$. Hence taking $\gamma > 0$ to be given in advance we shall select the function $u^* = u^*[t, \tau_i]$ $(\tau_i \leq t < \tau_{i+1})$ from the conditions given above.

We have thus described the choice of controls $u = u[\tau_i]$, $v^* = v^*[\tau_i, u^*]$ and $u^* = u^*[t, \tau_i]$ with $\tau_i \leq t < \tau_{i+1}$ .It now remains to describe the way the motions $x[t]$ (2.2) and $w[t]$ (2.3) are realized in time. According to the conditions of the problem the motion $x[t]$ of the initial system starts from position $\{t_0, x_0\}$ and is generated in time under the actions of the control $u = u[\tau_i]$ $(\tau_i \leq t < \tau_{i+1}, i = 0, 1, 2, \ldots)$ for the first player and of a certain measurable

realization $v = v[t] \in Q$ $(t_0 \leq t)$ due to the second one. Here it is understood that the latter generates $v[t]$ in time t under any desirable control law based on any type of information conceived. The motion $w[t]$ of the model starts at the initial position $\{t_0, w_0\} = \{t_0, x_0^*\}$ and is formed with respect to equation (2.3) under the actions of two controls governed by the first player. These are $v^* = v^*[\tau_i, u^*]$ and $u^* = u^*[t, \tau_i]$ $(\tau_i \leq t < \tau_{i+1}, i = 0, 1, 2, \ldots)$. Here also at moments of time $t = \tau_i$ $(i = 1, 2, \ldots)$ the motion $w[t]$ is subjected to finite discontinuities caused by the choice for the value $w[\tau_i]$ of the point of $W(\tau_i)$ nearest to $w[\tau_i - 0]$. In view of condition (4.8) the value $w[\tau_i] - w[\tau_i - 0]$ does not exceed the value γ. Finally it is noteworthy to declare that the control process comes to an end at moment $\tau \leq \vartheta$ when condition (4.9) is fulfilled at first. The existence of such a moment $\tau \leq \vartheta$ is ensured by the stability property of the bridge W.

Finally having the previous suggestions in view and making use of the estimate (4.7) and of a standard consideration in the stability theory we arrive at the following main assertion.

Suppose that Condition A is fulfilled. If for the initial position $\{t_0, w_0\} = \{t_0, x_0^*\}$ of model (2.3) the problem of encounter of $\{t, w[t]\}$ with set M within set N until the moment ϑ is solvable for the first player, then for any assigned $\varepsilon > 0$ it is possible to select sufficiently small values of $\mu > 0$, $\zeta > 0$, $\delta > 0$ and further on the value $\gamma > 0$, such that with conditions (2.5) being fulfilled and $\tau_{i+1} - \tau_i = \delta$ the procedure described ensures for all $t \geqq t_0$ the inequality $\| x[t] - w[t] \| < \varepsilon/2$ and the construction of motion ensures that there exists a moment of time $t = \tau \leq \vartheta$ when the inequality $\varrho(w[\tau], M(\tau)) < \varepsilon/2$, $\varrho(w[t], N(t)) < \varepsilon/2$, $t_0 \leq t \leq \tau$ will be attained and this will denote that the game is practically solved for the first player.

Hence the fulfilment of the given conditions leads under the described procedure to the appearance of a regime of sliding for the position $\{t, w[t]$ along the bridge W till the encounter of $\{t, x\}$ with M at $\tau \leq \vartheta$. A stable procedure which holds the actual motion $x[t]$ and the model motion $w[t]$ in the vicinity of each other is realized throughout the whole process.

It necessary to note that the estimate for the deviation of $x[t]$ from $w[t]$ is not essentially connected with the duration of the control process and the given method is thus workable enough for time intervals $t_0 \leq t \leq \vartheta$ of arbitrary length including the idealized problems where the control process is expanded throughout the entire semiinterval $t_0 \leq t < \infty$ (if through the entire interval the members μg in equation (2.2) are sufficiently small). Problems of this kind may obtained by eliminating for the first player the requirement that $\{\tau, x(\tau)\}$ is to arrive at M and by leaving for the player the only requirement that $\{t, x[t]\}$ is to remain within N (now however for all

values of $t \in [t_0, \infty)$). In the latter case we obtain for the first player a version of the problem of evasion within the infinite interval of time.

5. Let us continue the discussion of the example of § 3. Suppose that the first player is to solve the problem of evasion for point $m^{(1)}$ from $m^{(2)}$ throughout a time interval as large as possible. We are assuming that in the selected model (3.7), (3.8) the resources of the control for the second player $\varkappa^{(2)*}$ are greater than those of the first player $\varkappa^{(1)*}$ so that $\varkappa^{(2)*} - \varkappa^{(1)*} = \varkappa > 0$.

Suppose the conditions of the problems are such that the first player intends to evade point $m^{(1)}$ from an intersection with a given ν — neighbourhood of point $m^{(2)}$ throughout a time interval as large as possible ($\nu \geq 0$). According to the suggestions of § 4 for ensuring the desired evasion condition $(x_1 - x_7)^2 + (x_2 - x_8)^2 + (x_3 - x_9)^2 \geqq \nu$ in the actuating system (3.5), (3.6) within a certain interval of time it is necessary to ensure an evasion condition for the linear model (3.7), (3.8), i.e. the inequality $(w_1 - w_7)^2 + (w_2 - w_8)^2 + (w_3 - w_9)^2 - (\nu + \varepsilon)^2$.

The latter condition is to hold within the same time interval as the former. According to the results of the pursuit-evasion problem theory for linear *plants of similar type* [6] the least duration ever possible for the process of evasion is given by the value $T^0 = \vartheta^0 - t_0$ determined by solving an auxiliary time-optimal control problem $T^0 = \vartheta^0 - t_0 = \min_h (\vartheta - t_0)$ with boundary conditions

$$\begin{aligned} &z_j(t_0) = w_j(t_0) - w_{j+6}(t_0) = z_{j0} \\ &z_j(\vartheta) = w_i(\vartheta) - w_{j+6}(\vartheta) \qquad (j = 1, 2, \ldots, 6) \\ &z_1^2(\vartheta) + z_2^2(\vartheta) + z_3^2(\vartheta) \leq (\nu + \varepsilon)^2 \end{aligned} \tag{5.1}$$

for the system described by the equations

$$\begin{aligned} &\dot{z}_1 = z_4, \dot{z}_2 = z_5, \dot{z}_3 = z_6 \\ &\dot{z}_4 = \alpha z_1 + \beta z_5, \dot{z}_5 = \gamma z_4 + \sigma_1 \\ &\dot{z}_6 = \delta z_3 + \sigma_2 \end{aligned} \tag{5.2}$$

with constraints

$$\sigma_1^2 + \sigma_2^2 \leq \varkappa^2 . \tag{5.3}$$

Therefore suppose that $T^0 = \vartheta^0 - t_0$ is the solution of this problem. Then for the model system it is possible to solve the problem of encounter of motion $w[t]$ with set $M = [t = \vartheta^0]$ within set $N = [(w_1 - w_7)^2 + (w_2 - w_8)^2 + (w_3 - w_9)^2 \geq (\nu + \varepsilon)^2,\ t_0 \leq t \leq \vartheta^0]$. Thus, the problem of our interest is transformed to the latter problem for the first player. For keeping the motion $w[t]$ within the bridge W (which is here formed of position $\{t, w\}$ given by the

relation $w_j = w_{j+6} = z_j(t_1, t_0, z_0)$ $(j = 1, \ldots, 6)$ where $z(t, t_0, z_0)$ are the motions of (5.2), (5.3) suffices to select the control $u^* = u^*[t, \tau_i]$ so that conditions $\| v^*[\tau_i] - u^*[t, \tau_i] \| \leq \varkappa$ would be fulfilled. In other words it is sufficient that $u^*[t, \tau_i] = \varkappa^{(1)*}$, $v^*[\tau_i]/\varkappa^{(2)*}$. (One is to bear in mind that in the given particular case the value $v^*[\tau_i, u^*]$ (4.6) is actually independent of u^*. See e.g. an analogous case in [2, 3].) Further on, the stabilizing vector valued function $h = H(t)y = Hy$ (see Condition (1_A) and the quadratic forms (4.1) and (4.2) $\xi(t, y) = \xi(y)$ and $\lambda(t, y) = \lambda(y)$ are determined via well-known methods of theory of stability and stabilization of motions [5]. With form $\lambda(y) = \sum_{i,j=1}^{12} \lambda_{ij}\, y_i\, y_j$ being determined the control actions v^* and u are being found due to (4.5) and (4.6) from the following relations

$$u_1 = -\frac{\varkappa^{(1)}\partial\lambda/\partial y_5}{\left[\left(\frac{\partial\lambda}{\partial y_5}\right)^2 + \left(\frac{\partial\lambda}{\partial y_6}\right)^2\right]^{1/2}}, \qquad u_2 = -\frac{\varkappa^{(1)}\partial\lambda/\partial y_6}{\left[\left(\frac{\partial\lambda}{\partial y_5}\right)^2 + \left(\frac{\partial\lambda}{\partial y_6}\right)^2\right]^{1/2}}$$

$$v_1^* = \frac{\varkappa^{(2)*}\partial\lambda/\partial y_{11}}{\left[\left(\frac{\partial\lambda}{\partial y_{11}}\right)^2 + \left(\frac{\partial\lambda}{\partial y_{12}}\right)^2\right]^{1/2}}, \qquad v_2^* = \frac{\varkappa^{(2)*}\partial\lambda/\partial y_{12}}{\left[\left(\frac{\partial\lambda}{\partial y_{11}}\right)^2 + \left(\frac{\partial\lambda}{\partial y_{12}}\right)^2\right]^{1/2}} \tag{5.4}$$

$$(y = x[\tau_i] - w[\tau_i]).$$

On the contrary assume that the first player intends to ensure an encounter of point $m^{(1)}$ with a given $(\nu + \varepsilon)$-neighbourhood of point m^2. Suppose we have $\varkappa^{(1)*} - \varkappa^{(2)*} = \varkappa > 0$. Then it is known again [6] that the desired approach resulting in condition $(w_1 - w_7)^2 + (w_2 - w_8)^2 + (w_3 - w_9)^2 \leq \nu^2$ for the model is realizable only until moment ϑ^0 where ϑ^0 is again the solution of the open loop time-optimal control problem for systems (5.2), (5.3) with boundary condition (5.1) being substituted for $z_1^2(\vartheta) + z_2^2(\vartheta) + z_3^2(\vartheta) \leq \nu^2$. Having the solution ϑ° obtained we may solve for the model plant the first player's problem of encounter of motion $w[t]$ with set $M = [(w_1 - w_7)^2 + (w_2 - w_8)^2 + (w_3 - w_9)^2 \leq \nu^2,\ t = \vartheta^0$.

The role of bridge W is now addressed to the surface $W = [w_j - w_{j+6} = z_j(t),\ j = 1, 2, \ldots, 6,\ t_0 \leq t \leq \vartheta^0]$, where $z^0(t)$ is the optimal motion for the open-loop problem of above. The control actions u and v^* are again determined by equalities (5.4) and the control $u^* = u^*[t, \tau_i]$ is determined from the condition of conserving position $\{t, w[t]\}$ within the given bridge W, i.e. by the equalities

$$u_1^* = \dot{z}_5^0(t) - \gamma z_4^0(t) + v_1^*$$

$$u_2^* = \dot{z}_6^0(t) - \delta z_3^0(t) + v_2^*. \tag{5.5}$$

References

1. *Krasovskii, N. N.*, Differential game of approach-evasion I. Izv. Acad. Sci. USSR. Technical Cybernetics (1973) **2**.
2. *Krasovskii, N. N.*, Differential game of approach-evasion II. Izv. Acad. Sci. USSR. Technical Cybernetics (1973) **3**.
3. *Krasovskii, N. N.*, *Subbotin, A. I.*, Approximation in a differential game. Prikl. Math. i Mech. **37** (1973) 2.
4. *Krasovskii, N. N.*, *Subbotin, A. I.*, *Ushakov, V. N.*, A minimax-type differential game. Dokl. Acad. Sci. USSR **206** (1972) **2**.
5. *Malkin, J. G.*, Theory of stability of motion. "Nauka", M., 1966.
6. *Krasovskii, N. N.*, Game problems on the encounter of motions. "Nauka", M., 1970.

Игровое управление и задачи устойчивости

Н.Н. КРАСОВСКИЙ

(Свердловск)

Рассматриваются задачи управления в конфликтных ситуациях или в условиях неопределенности для динамических систем, описываемых обыкновенными дифференциальными уравнениями. Конструируется и теоретически обосновывается устойчивая игровая схема формирования управлений, которые обеспечивают должные значения функции цели.

В предлагаемом решении, именуемом «управлением с поводырем», к исходной системе присоединяется имитирующая модельная система. В результате формируется обобщенная дифференциальная игра, рассматриваемая в рамках формализации [1, 2], в ходе которой конструируются как искомая стратегия управления для реальной системы, так и управления для модели. Указанное построение осуществляется таким образом, чтобы привести движение модели в желаемое состояние и в то же время заставить реальное движение и движение модели взаимно отслеживать друг друга, с тем, чтобы движения модели и реального объекта были близки. Показано, что при определенных условиях указанные управления могут быть построены таким образом, чтобы осуществлялось отфильтровывание воздействий, вызываемых неизвестными возмущениями или противодействующими неопределенными факторами.

В результате указанное «управление с поводырем» приводит к улучшенной и устойчивой процедуре управления. Таким образом, в предлагаемой схеме соединяются методы теории устойчивости движения и теории дифференциальных игр.

Н. Н. Красовский
Институт математики и механики УНЦ АН СССР
СССР, Свердловск К—66,
ул. С. Ковалевской, д. 16.

Problems of Control and Information Theory, Vol. 3 (3), pp. 183—196 (1974)

ON COMPLEXITY OF SWITCHING CIRCUITS

T. G. TARJÁN

(Budapest)

(Received December 6, 1973)

1. Introduction

Let $\mathcal{G}$ be the set of circuits, an element G of which realizes the truth function $f_G \in \mathcal{F} = \{0, 1\}^n \to \{0, 1\}$ and consists of $e(G)$ components. Let us call $e(G)$ the complexity of G. Let the complexity of f in $\mathcal{G}$ be

$$\lambda_{\mathcal{G}}(f) = \min_{\substack{G \in \mathcal{G} \\ f_G = f}} e(G) \tag{1}$$

and let $\lambda_{\mathcal{G}}(n) = \max_{f \in \mathcal{F}} \lambda_{\mathcal{G}}(f)$ be the Shannon function, that is, $\lambda_{\mathcal{G}}(n)$ is defined as the least number such that any truth function of n variables can be realized with not more than $\lambda_{\mathcal{G}}(n)$ components.*

Thus any truth function of n variables can be realized with $\lambda_{\mathcal{G}}(n)$ components and at least one of them with no less.

Shannon has introduced in [1] the set of two-terminal switching circuits $\mathcal{G}_s$ and the set of series-parallel two-terminal circuits $\mathcal{G}_{\pi}$, nevertheless, the conventions of this paper agree with Lupanov's paper [4].

Shannon has proved in [3] for $\mathcal{G}_s$ that for arbitrary $\varepsilon > 0$, $n > n(\varepsilon)$

$$(1 - \varepsilon)\frac{2^n}{n} < \lambda_s(n) < (1 + \varepsilon)\frac{2^{n+1}}{n}$$

and for sufficiently large n all truth functions of n variables excepting a fraction δ require at least $(1 - \varepsilon)\dfrac{2^n}{n}$ components where ε and δ are arbitrary positive numbers, that is, nearly all the truth functions' complexity are near to the maximum. Lupanov has shown in [4] that also $\lambda_s(n) \sim \dfrac{2^n}{n}$ is true. But there is not known $h \in \mathcal{F}$ truth function series for which $\lambda_s(h) \gtrsim n^2$ would be true.**

* The complexity introduced by Shannon in [3] is slightly different from (1): It is $\dfrac{\lambda_{\mathcal{G}}(f)}{\lambda_{\mathcal{G}}(n)}$ for $f \in \mathcal{F}$.

** $a_n \gtrsim b_n$ if $\overline{\lim\limits_{n \to \infty}} \dfrac{b_n}{a_n} \leq 1$ holds.

Shannon has proved in [1] for $\mathcal{G}_{\mathfrak{J}\pi}$ that

$$\lambda_\pi(n) \leq 3 \cdot 2^{n-1} - 2 \tag{2}$$

and Riordan and Shannon have shown in [2] that

$$\lambda_\pi(n) \gtrsim \frac{2^n}{\log_2 n}$$

and for sufficiently large n all truth functions of n variables excepting a fraction δ require at least $(1 - \varepsilon)\dfrac{2^n}{\log_2 n}$ components where ε and δ are arbitrary positive numbers, that is, nearly all the truth functions' complexity are near to the maximum. Lupanov has shown in [4] that also $\lambda_\pi(n) \sim \dfrac{2^n}{\log_2 n}$ holds. But there is not known $h \in \mathfrak{F}$ truth function series for which $\lambda_\pi(h) \gtrsim n^2$ would be true. In these classes of circuits we do not know anything about the structure of the extremal functions. Therefore I tried to find subclasses of $\mathcal{G}_{\mathfrak{J}\pi}$ for which the determination of the extremal functions is possible. Such results can help us to conjecture something on the extremal functions of larger classes. The first class $\mathcal{G}_{\mathfrak{J}0}$ is the set of disjunctive normal forms,* $\mathcal{G}_{\mathfrak{J}1}$ and $\mathcal{G}_{\mathfrak{J}2}$ are rather natural extensions of $\mathcal{G}_{\mathfrak{J}0}$, for which

$$\mathcal{G}_{\mathfrak{J}0} \subset \mathcal{G}_{\mathfrak{J}1} \subset \mathcal{G}_{\mathfrak{J}2} \subset \mathcal{G}_{\mathfrak{J}\pi} \subset \mathcal{G}_{\mathfrak{J}s} \quad \text{is true.}$$

The most complex truth function series in $\mathcal{G}_{\mathfrak{J}k}$ $k = 0, 1, 2$ is

$$h(x_1, \ldots, x_n) = \begin{cases} 0 & \text{if } \sum_{j=1}^{n} x_j \text{ is even} \\ 1 & \text{otherwise} \end{cases} \tag{3}$$

which occurs in [1, p. 718]. The idea of Shannon** is working for the proof lo the upper bounds of $\lambda_k(n)$ $k = 0, 1, 2$ but it is not constructive and by the he p of my proof the unicity of (3) follows easily. Thus, the estimation (2) is the best possible for the class $\mathcal{G}_{\mathfrak{J}2}$; leaving $\mathcal{G}_{\mathfrak{J}2}$ it probably breaks down. I shall show that — in a certain sense — $\mathcal{G}_{\mathfrak{J}2}$ is the largest class for which h is the most complex truth function.

* The terms in connection with this are found for example in [5].

** In Shannon's paper [1] the statement is formulated for $\mathcal{G}_{\mathfrak{J}\pi}$. It is not true, as Riordan has pointed out later, but his proof works for $\mathcal{G}_{\mathfrak{J}0}$, $\mathcal{G}_{\mathfrak{J}1}$, $\mathcal{G}_{\mathfrak{J}2}$.

2. On complexity of Boolean formulas

Let $\mathfrak{F}$ denote the set of n variables truth functions where $\mathfrak{F} \in \{0, 1\}^n \to \to \{0, 1\}$ and let us interpret the next three operations among the elements of $\mathfrak{F}$ in the following way:

negation: $\bar{f} = 1 - f$
disjunction: $f \vee g = f + g - f \cdot g$
conjunction: $f \wedge g = f \cdot g$ where f and g are arbitrary truth functions.

It is easy to show that $\mathfrak{F}$ forms a Boolean algebra by the operations of negation, disjunction and conjunction. A formula is defined as sequence formed from the elementary (truth) functions $x_1, \bar{x}_1, \ldots, x_n, \bar{x}_n$ $(X = \{x_1, \bar{x}_1, \ldots, x_n, \bar{x}_n\})$ by the help of disjunctions and conjunctions where $x_j(\delta_1, \ldots, \delta_n) = \delta_j$, $1 \leq j \leq n$. Let us define the set of elementary conjunctions $\mathfrak{E}$ for which

(i) $X \subset \mathfrak{E}$
(ii) $\mathfrak{E} \wedge \mathfrak{E} \subset \mathfrak{E}$

hold* and the set of disjunctive normal forms $\mathfrak{D}$ for which

(i) $\mathfrak{E} \subset \mathfrak{D}$
(ii) $\mathfrak{D} \vee \mathfrak{D} \subset \mathfrak{D}$ are true.*

Let us denote by $L(\mathfrak{A})$ the length of the formula $\mathfrak{A}$ where $L(\mathfrak{A})$ is the number of elementary functions in $\mathfrak{A}$ (e.g. $L(x_1 \wedge (x_1 \vee \bar{x}_2)) = 3$).

Problem. Let us determine the truth function $h^{\mathfrak{D}} \in \mathfrak{F}$ with the property that its shortest disjunctive normal form is the longest possible, that is, look for $h^{\mathfrak{D}} \in \mathfrak{F}$ for which

$$\lambda_{\mathfrak{D}}(h^{\mathfrak{D}}) = \lambda_{\mathfrak{D}}(n)$$

holds where

$$\lambda_{\mathfrak{D}}(f) = \min_{\substack{\mathfrak{A}=f \\ \mathfrak{A} \in \mathfrak{D}}} L(\mathfrak{A}),\ f \in \mathfrak{F}$$

and

$$\lambda_{\mathfrak{D}}(n) = \max_{f \in \mathfrak{F}} \lambda_{\mathfrak{D}}(f).$$

Theorem 1. If

$$h(\delta_1, \ldots, \delta_n) = \begin{cases} 1 \text{ if } \sum_{j=1}^{n} \delta_j \text{ is even} \\ 0 \text{ otherwise} \end{cases} \tag{4}$$

* $\varepsilon \wedge \varepsilon = \{\mathfrak{A}_1 \wedge \mathfrak{A}_2 : \mathfrak{A}_1, \mathfrak{A}_2 \in \varepsilon\}$
$\mathfrak{D} \vee \mathfrak{D} = \{\mathfrak{A}_1 \vee \mathfrak{A}_2 : \mathfrak{A}_1, \mathfrak{A}_2 \in \mathfrak{D}\}$

then

$$\lambda_{\mathfrak{D}}(h) = \lambda_{\mathfrak{D}}(\bar{h}) = \lambda_{\mathfrak{D}}(n) = n \cdot 2^{n-1}$$

and

$$\lambda_{\mathfrak{D}}(f) = n \cdot 2^{n-1}, \; f \in \mathfrak{F} \text{ implies } f = h \text{ or } \bar{h}.$$

Proof

Upper bound: Let $f \in \mathfrak{F}$ be an arbitrary truth function and let us consider the following representation of f which is known as full disjunctive normal form of f

$$\mathfrak{A} = \bigvee_{\substack{\delta_1,\ldots,\delta_n \in \{0,1\}^n \\ f(\delta_1\ldots,\delta_n)=1}} x_1^{\delta_1} \wedge \ldots \wedge x_n^{\delta_n} \text{ where } x^\delta = \begin{cases} x & \text{if } \delta = 1 \\ \bar{x} & \text{otherwise}. \end{cases} \tag{5}$$

If for $\delta_1, \delta_2, \ldots, \delta_n \in \{0, 1\}^n$ there exist both $x_1^{\delta_1} \wedge x_2^{\delta_2} \wedge \ldots \wedge \bar{x}_n$ and $x_1^{\delta_1} \wedge x_2^{\delta_2} \wedge \ldots \wedge x_n$ in $\mathfrak{A}$ then let us delete them from $\mathfrak{A}$ and write the term $x_1^{\delta_1} \wedge x_2^{\delta_2} \wedge \ldots \wedge x_{n-1}^{\delta_{n-1}}$ instead of them. Let $\mathfrak{A}'$ denote the new disjunctive normal form which is also trivially a representation of f, thus for arbitrary $f \in \mathfrak{F}$, $\lambda_{\mathfrak{D}}(f) \leq L(\mathfrak{A}') \leq n \cdot 2^{n-1}$ are true.

Lower bound: Let us consider an $\mathfrak{A} \in \mathfrak{D}$ disjunctive normal form which represents h, that is, $\mathfrak{A} = h$ holds. Let $\Delta = \delta_1, \delta_2, \ldots, \delta_n \in \{0, 1\}^n$ for which $\sum_{j=1}^{n} \delta_j$ is even, then $\mathfrak{A}$ has a term $\mathfrak{A}_\Delta$ which is equal to 1 on $\delta_1, \delta_2, \ldots, \delta_n$. $x_j^{1-\delta_j}$, $1 \leq j \leq n$ does not occur in $\mathfrak{A}_\Delta$ because $\delta_j^{(1-\delta_j)} = 0$, but $x_j^{\delta_j}$ $1 \leq j \leq n$ must occur in $\mathfrak{A}_\Delta$ otherwise $\mathfrak{A}_\Delta$ would be equal to 1 also on $\delta_1, \ldots, 1 - \delta_j, \ldots, \delta_n$ but $1 - 2\delta_j + \sum_{k=1}^{n} \delta_k$ is odd which would be a contradiction.

Thus $L(\mathfrak{A}_\Delta) \geq n$, on the other hand, if $\Delta, \Delta' \in \{0, 1\}^n$, $\Delta \neq \Delta'$ and $h(\Delta) = h(\Delta') = 1$ hold then $\mathfrak{A}_\Delta \neq \mathfrak{A}_{\Delta'}$ is true. Consequently $L(\mathfrak{A}) \geq n \cdot 2^{n-1}$ and $\mathfrak{A}$ was arbitrary, therefore also $\lambda_{\mathfrak{D}}(h) \geq n \cdot 2^{n-1}$ holds.

Accordingly for arbitrary $f \in \mathfrak{F}$

$$\lambda_{\mathfrak{D}}(f) \leq n \cdot 2^{n-1} \leq \lambda_{\mathfrak{D}}(h) \leq \max_{f \in \mathfrak{F}} \lambda_{\mathfrak{D}}(f)$$

are true then

$$\max_{f \in \mathfrak{F}} \lambda_{\mathfrak{D}}(f) \leq n \cdot 2^{n-1} \leq \lambda_{\mathfrak{D}}(h) \leq \max_{f \in \mathfrak{F}} \lambda_{\mathfrak{D}}(f)$$

holds which implies that $\lambda_{\mathfrak{D}}(h) = \lambda_{\mathfrak{D}}(n) = n \cdot 2^{n-1}$. The proof is the same for $\bar{h}$.

Unicity. At the beginning of our proof $f \in \mathfrak{F}$ was an arbitrary truth function for which we have constructed $\mathfrak{A}'$ and have obtained that $\lambda_{\mathfrak{D}}(f) \leq$ $\leq L(\mathfrak{A}') \leq n \cdot 2^{n-1}$ are true.

If we now suppose that $\lambda_{\mathfrak{D}}(f) = n \cdot 2^{n-1}$ is true we get the next parities: $\lambda_{\mathfrak{D}}(f) = L(\mathfrak{A}') = n \cdot 2^{n-1}$. Thus by the construction of $\mathfrak{A}'$ we know that

$$\mathfrak{A}' = \bigvee_{\delta_1, \ldots, \delta_{n-1} \in \{0,1\}^{n-1}} x_1^{\delta_1} \wedge x_2^{\delta_2} \wedge \ldots x_n^{\lambda} \text{ - shaped}$$

where $\lambda = \lambda(\delta_1, \ldots, \delta_{n-1})$. Now let us consider the chromatic graph C in which the vertices are the terms of $\mathfrak{A}$, their colour is red if $\lambda + \sum_{j=1}^{n-1} \delta_j$ is even, and blue otherwise. Two vertices $x_1^{\delta_1} \wedge x_2^{\delta_2} \wedge \ldots \wedge x_n^{\lambda}$ and $x_1^{\varepsilon_1} \wedge x_2^{\varepsilon_2} \wedge \ldots \wedge x_n^{\lambda}$ are adjacent if $\sum_{j=1}^{n-1} |\delta_j - \varepsilon_j| = 1$.

It is obvious that C is connected, therefore, all the vertices of C are either red or blue which implies the last statement of our theorem. Let us suppose that v_1 is a red and v_2 is a blue vertex of C. Since C is connected, there is a path which joines v_1 and v_2. At this path there are vertices of C, v_3 and v_4, which are adjacent, and of different colours. So $\mathfrak{A}'$ is not the shortest, because instead of the terms v_3 and v_4 we can write only one term and this is a contradiction with the parity $\lambda_{\mathfrak{D}}(f) = \lambda(\mathfrak{A}')$. Q.e.d.

3. On complexity of two-terminal graphs

$G = \{V, E, \varrho, v^+, v^-\}$ is called a two-terminal graph of n variables where V is the set of vertices, E is the set of edges, $\varrho \in E \to X$ if $X = \{x_1, \bar{x}_1, \ldots, x_n, \bar{x}_n\}$ is the set of variables, $v^+ \in V$ is the positive terminal and $v^- \in V$ is the negative terminal. E.g.

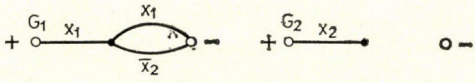


Fig. 1

These graphs are called switching circuits by Shannon in [1]. Let us denote by $\mathcal{G}_s$ the set of two-terminal graphs of n variables. Let G_1 and G_2 be elements of $\mathcal{G}_s$. We say that G_1 is a subgraph of G_2, $G_1 \subset G_2$, if $V_1 \subset V_2$, $E_1 \subset E_2$, $\varrho_1 = \varrho_2 \wedge E_1$ hold where ϱ_1 is the restriction of ϱ_2 on E_1. If $v_1^+ = v_2^+$ and $v_1^- = v_2^-$ are also true we say that G_1 is less then G_2, that is, $G_1 < G_2$.

$P \in \mathcal{G}_s$ is a path, if the vertices of P have a sequence $v^+, v_1, \ldots, v_m, v^-$ such that exactly the successive vertices of G are adjacent. This sequence is uniquely determined and determines a natural oder among the edges of P, namely let e_1 be incident with v_j, v_{j+1} and e_2 with v_k, v_{k+1} then $e_1 < e_2$ if $j < k$ where $v_0 = v^+$, $v_{m+1} = v^-$.

Let $\mathcal{S}$ denote the set of two-terminal graphs of n variables which are paths.

Let $\mathcal{F}$ denote the set of truth functions, $\mathcal{F} = \{0, 1\}^n \to \{0, 1\}$. In case of $G \in \mathcal{G}_s$ let $f_G \in \mathcal{F}$ be 1 on $\delta_1, \delta_2, \ldots, \delta_n \in \{0, 1\}^n$ if there exists a path $P \in \mathcal{S}$ which is less then G, $P < G$, and for which all the edges have the value 1 in case of the replacement $x_j = \delta_j$, $\bar{x}_j = 1 - \delta_j$, $1 \leq i \leq n$ or if $v^+ = v^-$, and 0 otherweise, that is,

$$f_G(\delta_1, \ldots, \delta_n) = \begin{cases} 1 \text{ if in case of the replacement } x_j = \delta_j, x_j = 1 - \delta_j, \\ \quad 1 \leq j \leq n \text{ there exists a } P \in \mathcal{S},\ P < G \text{ for which} \\ \quad \varrho_p(e) = 1 \text{ holds if } e \in E_p \text{ or } v^+ = v^- \\ 0 \text{ otherwise} \end{cases}$$

where $G \in \mathcal{G}_s$ and $\delta_1, \ldots, \delta_n \in \{0, 1\}^n$. Let $f_G \in \mathcal{F}$ be called the truth function realized by G. Let $e(G)$ denote the number of edges of $G \in \mathcal{G}_s$ namely $e(G) = |E|$. Let $\mathcal{G}_j$ be an arbitrary subset of $\mathcal{G}_s$ and $f \in \mathcal{F}$ an arbitrary truth function, then let us denote by $\lambda_j(f)$ the complexity of f in $\mathcal{G}_j$ where

$$\lambda_j(f) = \min \{e(G) : f_G = f, G \in \mathcal{G}_j\}.$$

Then let $\lambda_j(n)$ be the complexity of the most complex truth function which is called the Shannon function of $\mathcal{G}_j$, that is $\lambda_j(n) = \max_{f \in \mathcal{F}} \lambda_j(f)$. Let $G_1 \vee G_2$ be the following two-terminal graph G if $G_1, G_2 \in \mathcal{G}_s$: V is the union of V_1, V_2 in such a way that in V, $v^+ = v_1^+ = v_2^+$, $v^- = v_1^- = v_2^-$ are true and all the other elements of $V_j, j = 1, 2$ are elements of V but in V they differ from all the elements of V_{3-j}. E is the union of E_1, E_2 in such a way that all the elements of $E_j, j = 1, 2$ are elements of E but in E they differ from all the elements of $E_{3-j} \cdot \varrho(e)$ is equal to $\varrho_1(e)$ if e was an element of E_1, otherwise is equal to $\varrho_2(e)$. E.g. for G_1, G_2 being on Fig. 1 $G_1 \vee G_2$ is the following:

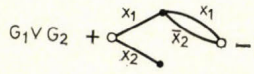


Fig. 2

Let $G_1 \wedge G_2$ be the following two-terminal graph G if $G_1, G_2 \in \mathcal{G}_s$: V is the union of V_1, V_2 in such a way that $v_1^- = v_2^+$ in V and the other elements of V_j, $j = 1, 2$ are elements of V but in V they differ from all the elements of V_{3-j}. Let $v^+ = v_1^+$ and $v^- = v_2^-$ be true. E is the union of E_1, E_2 such that all the elements of E_j, $j = 1, 2$, are elements of E but in E they differ from all the ele-

ments of $E_{3-j} \cdot \varrho(e)$ is equal to $\varrho_1(e)$ if e was an element of otherwise it is equal to $\varrho_2(e)$. E.g. for G_1, G_2 being on Fig. 1 $G_1 \wedge G_2$ is the following:

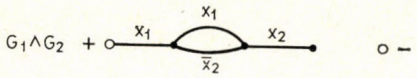


Fig. 3

Let $\mathcal{G}_0 \subset \mathcal{G}_s$ such that

(i) $\mathcal{S} \subset \mathcal{G}_0$

(ii) $\mathcal{G}_0 \vee \mathcal{G}_0 \subset \mathcal{G}_0$ hold.*

Consequently $G \in \mathcal{G}_0$ has arbitrary "disjoint" paths between the terminals. E.g.

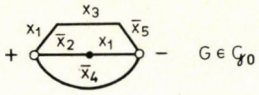


Fig. 4

Let us denote $\mathcal{G}_0$ also by $\mathcal{G}_{0/2}$. It is obvious that $\mathcal{G}_0$ corresponds to $\mathfrak{D}$, the set of disjunctive normal forms in § 2.

Let $\mathcal{G}_{1/2} \subset \mathcal{G}_s$ such that

(i) $\mathcal{S} \subset \mathcal{G}_{1/2}$

(ii) $\mathcal{G}_{1/2} \vee \mathcal{G}_{1/2} \subset \mathcal{G}_{1/2}$

(iii) $\mathcal{S} \wedge \mathcal{G}_0 \subset \mathcal{G}_{1/2}$ hold.**

Let $\mathcal{G}_1 \subset \mathcal{G}_s$ such that:

(i) $\mathcal{S} \subset \mathcal{G}_1$

(ii) $\mathcal{G}_1 \vee \mathcal{G}_1 \subset \mathcal{G}_1$

(iii) $\mathcal{S} \wedge \mathcal{G}_0, \mathcal{G}_0 \wedge \mathcal{S} \subset \mathcal{G}_1$ hold.

Consequently $G \in \mathcal{G}_{1/2}$ is a "special tree" between the terminals. E.g.

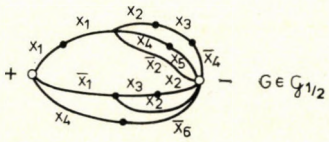


Fig. 5

Let $\mathcal{G}_{2/2} \subset \mathcal{G}_s$ such that

(i) $\mathcal{S} \subset \mathcal{G}_{2/2}$

(ii) $\mathcal{G}_{2/2} \vee \mathcal{G}_{2/2} \subset \mathcal{G}_{2/2}$

(iii) $\mathcal{S} \wedge \mathcal{G}_{2/2} \subset \mathcal{G}_{2/2}$ hold.**

Let $\mathcal{G}_2 \subset \mathcal{G}_s$ such that

(i) $\mathcal{S} \subset \mathcal{G}_2$

(ii) $\mathcal{G}_2 \vee \mathcal{G}_2 \subset \mathcal{G}_2$

(iii) $\mathcal{G}_2 \wedge \mathcal{S}, \mathcal{S} \wedge \mathcal{G}_2 \subset \mathcal{G}_2$ hold.

Consequently $G \in \mathcal{G}_{2/2}$ is an "arbitrary tree" between the terminals.

* $\mathcal{G} \vee \mathcal{G} = \{G_1 \vee G_2 : G_1, G_2 \in \mathcal{G}\}$ if $\mathcal{G} \subset \mathcal{G}_s$

** $\mathcal{G} \wedge \mathcal{G}' = \{G_1 \wedge G_2 : G_1 \in \mathcal{G}, G_2 \in \mathcal{G}'\}$ if $\mathcal{G}, \mathcal{G}' \subset \mathcal{G}_s$.

Let $\mathcal{G}_{\mathcal{J}\pi} \subset \mathcal{G}_{\mathcal{J}s}$ be the set of two-terminal series-parallel graphs such that

(i) $\mathcal{S} \subset \mathcal{G}_{\mathcal{J}\pi}$
(ii) $\mathcal{G}_{\mathcal{J}\pi} \vee \mathcal{G}_{\mathcal{J}\pi} \subset \mathcal{G}_{\mathcal{J}\pi}$
(iii) $\mathcal{G}_{\mathcal{J}\pi} \wedge \mathcal{G}_{\mathcal{J}\pi} \subset \mathcal{G}_{\mathcal{J}\pi}$.

E.g.

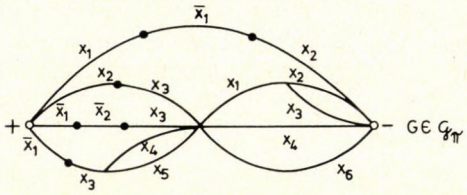


Fig. 6

Theorem 2. If

$$h(\delta_1, \ldots, \delta_n) = \begin{cases} 1 \text{ if } \sum_{j=1}^{n} \delta_j \text{ is even} \\ 0 \text{ otherwise} \end{cases}$$

then

$$\lambda_0(h) = \lambda_0(\bar{h}) = \lambda_0(n) = n \cdot 2^{n-1}$$
$$\lambda_1(h) = \lambda_1(\bar{h}) = \lambda_1(n) = \alpha(n)2^{n-1}, \quad \text{where} \quad \alpha(n) = \min_{1 \le k \le n-2} k \cdot 2^{k+1-n} + (n-k).$$
$$\lambda_2(h) = \lambda_2(\bar{h}) = \lambda_2(n) = 3 \cdot 2^{n-1} - 2$$

and $\lambda_k(f) = \lambda_k(n)\, f \in \mathcal{F}$ implies $f = h$ or $\bar{h}$, $k = 0, 1, 2$.

Lemma 1. In case of $G_1, G_2 \in \mathcal{G}_{\mathcal{J}s}$

$$e(G_1 \vee G_2) = e(G_1 \vee G_1),\ f_{G_1 \vee G_2} = f_{G_2 \vee G_1}$$
$$e(G_1 \wedge G_2) = e(G_2 \wedge G_1),\ f_{G_1 \wedge G_2} = f_{G_2 \wedge G_1} \text{ hold.}$$

Proof: trivial.

Lemma 2. Let $G_k, G'_k \in \mathcal{G}_{\mathcal{J}s}$, $k = 1, 2$, such that $e(G_k) = e(G'_k)$, $f_{G_k} = f_{G_k}$ are true, then

$$e(G_1 \vee G_2) = e(G'_1 \vee G'_2),\ f_{G_1 \vee G_2} = f_{G'_1 \vee G'_2}$$
$$e(G_1 \wedge G_2) = e(G'_1 \wedge G'_2),\ f_{G_1 \wedge G_2} = f_{G'_1 \wedge G'_2} \text{ hold.}$$

Proof: trivial.

Lemma 3. For arbitrary $G \in \mathcal{G}_{\mathcal{J}j}$ there exists a $G' \in \mathcal{G}_{\mathcal{J}j/2}$ for which $f_G = f_{G'}$, $e(G) = e(G')$ hold ($j = 1, 2$).

Proof:

(i) if $G \in \mathcal{S}$ holds then $G \in \mathcal{G}_{j/2}$ is true so $G' = G$ will be right.

(ii) if $G_k \in \mathcal{G}_j$ ($k = 1, 2$; $j = 1$ or 2) and there exists $G'_k \in \mathcal{G}_{j/2}$ for which $e(G_k) = e(G'_k)$ and $f_{G_k} = f_{G'_k}$ are true, then by Lemma 2 also $e(G_1 \vee G_2) = e(G'_1 \vee G'_2)$, $f_{G_1 \vee G_2} = f_{G'_1 \vee G'_2}$ hold and by definition $G'_1 \vee G'_2 \in \mathcal{G}_{j/2}$ is true. Thus, for $G_1 \vee G_2$, the $G'_1 \vee G'_2$ will be right.

(iii) if $G_1 \in \mathcal{S}$, $G_2 \in \mathcal{G}_0$ when $j = 1$; $G_2 \in \mathcal{G}_2$ when $j = 2$ and there exists a $G'_2 \in \mathcal{G}_{j/2}$, $j = 1$, or 2 for which $e(G_2) = e(G'_2)$ and $f_{G_2} = f_{G'}$ are true then

(iii.1) by Lemma 2 $e(G_1 \wedge G_2) = e(G_1 \wedge G'_2)$, $f_{G_1 \wedge G_2} = f_{G_1 \wedge G'_2}$ hold and by definition $G_1 \wedge G'_2 \in \mathcal{G}_{j/2}$ so for $G_1 \wedge G_2$ the $G_1 \wedge G'_2$ will be right.

(iii.2) by Lemmata 2 and 1 $e(G_2 \wedge G_1) = e(G'_2 \wedge G_1) = e(G_1 \wedge G'_2)$, $f_{G_2 \vee G_1} = f_{G'_2 \wedge G_1} = f_{G_2 \wedge G'_1}$ hold and by definition $G_1 \wedge G'_1 \in \mathcal{G}_{j/2}$ so for $G_2 \wedge G_1$ the $G_2 \wedge G'_1$ will be right. Q.e.d.

Lemma 4. Let $G = \{V, E, \varrho, v^+, v^-\} \in \mathcal{G}_{2/2}$. If $v \in V \setminus \{v^-\}$ hold then there is at most one $P_v = \{V_v, E_v, \varrho_v, v_v^+, v_v^-\}$ for which $P_v \subset G$, $P_v \in \mathcal{S}$, $v_v^+ = v^+$, $v_v^- = v$ are true.

Proof:

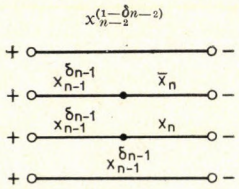
(i) if $G \in \mathcal{S}$ holds then the statement is trivial.

(ii) if $G_1, G_2 \in \mathcal{G}_{2/2}$ are like that the statement is true, v is a vertex of $G = G_1 \vee G_2$ and $P_v \subset G$ then by the definition of $G_1 \vee G_2$, $V_v \subset V_j$, $V_v \cap V_{3-j} = (v^+)$ $j = 1$ or 2 and for G_j there is at most one P_v.

(iii) if $G_2 \in \mathcal{G}_{2/2}$ is like that the statement is true, $G_1 \in \mathcal{S}$, v is a vertex of $G = G_1 \wedge G_2$ and $P_v \subset G$, then
— in case if v was a vertex of G_1, the statement is trivial.
— in case if v was a vertex of G_2, by the definition of $G_1 \wedge G_2$ $P_v = P_v^{(1)} \wedge P_v^{(2)}$ holds where $P_v^{(1)} = G_1$; $P_v^{(2)} \in \mathcal{S}$, $P_v^{(2)} \subset G_2$ and for G_2 there is at most one $P_v^{(2)}$. Q.e.d.

Proof of Theorem 2

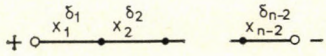
Upper bounds: Let $f \in \mathcal{F}$ be an arbitrary truth function and let us consider the following realizations $G_{\Delta(0)}$ of f based on the sequence $1, 2, \ldots, n$. For all $\Delta(n-1) = \delta_1, \ldots, \delta_{n-1} \in \{0, 1\}^{n-1}$ let $G_{\Delta(n-1)}$ be the following graph:



if $f(\delta_1, \ldots, \delta_{n-1}, 0) = 0$ and $f(\delta_1, \ldots, \delta_{n-1}, 1) = 0$
if $f(\delta_1, \ldots, \delta_{n-1}, 0) = 1$ and $f(\delta_1, \ldots, \delta_{n-1}, 1) = 0$
if $f(\delta_1, \ldots, \delta_{n-1}, 0) = 0$ and $f(\delta_1, \ldots, \delta_{n-1}, 1) = 1$
if $f(\delta_1, \ldots, \delta_{n-1}, 0) = 1$ and $f(\delta_1, \ldots, \delta_{n-1}, 1) = 1$

$$\text{where } x^\delta = \begin{cases} \bar{x} \text{ if } \delta = 0 \\ x \text{ otherwise.} \end{cases}$$

(0.2) For all $\Delta(n-1) = \delta_1, \ldots, \delta_{n-1} \in \{0, 1\}^{n-1}$ let $P_{\Delta(n-1)}$ be the following graph:



$$\text{where } x^\delta = \begin{cases} \bar{x} \text{ if } \delta = 0 \\ x \text{ otherwise.} \end{cases}$$

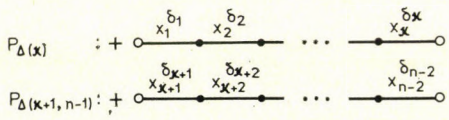
Then let $G_{\Delta(0)}$ let be the next one:

$$G_{\Delta(0)} = \bigvee_{\Delta(n-1) \in \{0,1\}^{n-1}} (P_{\Delta(n-1)} \wedge G_{\Delta(n-1)}).$$

It is obvious that $G_{\Delta(0)} \in \mathcal{G}_0$, $f_{G\Delta(0)} = f$ and

$$\lambda_0(f) \leq e(G_{\Delta(0)}) \leq n \cdot 2^{n-1} \text{ are true.}$$

(1.2) For all $\Delta(k) = \delta_1, \ldots, \delta_k \in \{0, 1\}^n$ and $\Delta(k+1, n-1) = \delta_{k+1}, \ldots, \delta_{n-1} \in \{0, 1\}^{n-1-k}$ let us consider the following graphs where $\min\limits_{1<k<n-2} [k \cdot 2^{k+1-n} + (n-k)] = \alpha(n)\, \varkappa\, 2^{\varkappa-1-n} + n - \varkappa$



Then let $G_{\Delta(0)}$ be the next graph:

$$G_{\Delta(0)} = \bigvee_{\Delta(\varkappa) \in \{0,1\}^\varkappa} \{P_{\Delta(x)} \wedge [\bigvee_{\Delta(\varkappa+1, n-1) \in \{0,1\}^{n-1-\varkappa}} (P_{\Delta(\varkappa+1, n-1)} \wedge G_{\Delta(\varkappa)\Delta(\varkappa+1, n-1)})]\}.$$

It is obvious that $G_{\Delta(0)} \in \mathcal{G}_{1/2}$, $f_{G\Delta(0)} = f$ and $\lambda_{1/2}(f) \leq e(G_{\Delta(0)}) \leq \alpha(n) \cdot 2^{n-1}$.

(2.2) Let us suppose we have already defined $G_{\Delta(i+1)}$ for every $\Delta(i+1) \in \{0,1\}^{i+1}$.

For $\Delta(i) = \delta_1, \ldots, \delta_i \in \{0, 1\}^i$, $i \geq 0$

let $G_{\Delta(i)} = P_{\Delta(i)} \wedge (G_{\Delta(i),0} \vee G_{\Delta(i),1})$ be where $P_{\Delta(i)}$: $\begin{cases} + \circ \overset{x_i^{\delta_i}}{\text{———}} \circ - & \text{if } i \geq 1 \\ + \circ - & \text{if } i = 0. \end{cases}$

It is obvious that

$$G_{\Delta(0)} \in \mathcal{G}_{\mathfrak{J}2/2}, f_{G_{\Delta(0)}} = f \qquad \text{and}$$
$$\lambda_{2/2}(f) \leq e\,(G_{\Delta(0)}) \leq 3 \cdot 2^{n-1} - 2 \quad \text{are true.}$$

Lower bounds. Let us consider a $G = \{V, E, \varrho, v^+, v^-\} \in \mathcal{G}_{\mathfrak{J}k/2}$, $(k = 0, 1, 2)$ which realizes h, that is $f_G = h$ is true and let $\Delta = \delta_1, \delta_2, \ldots, \delta_n \in \{0, 1\}^n$ arbitrary for which $\sum_{j=1}^{n} \delta_j$ is even. Then there exists a path $P_\Delta = \{V_\Delta, E_\Delta, \varrho_\Delta, v^+, v^-\} \in \mathfrak{S}$, $P_\Delta < G$ in which all the edges have the value 1 in case of the replacement $x_j = \delta_j$, $\bar{x}_j = 1 - \delta_j$, $1 \leq j \leq n$. $x_i^{1-\delta_i}$ $(1 \leq i \leq n)$ does not occur in P_Δ because $\delta_j^{1-\delta_j} = 0$, but $x_j^{\delta_j}$ must occur in P_Δ otherwise $h(\delta_1, \ldots, 1 - \delta_j, \ldots, \delta_n) = 1$ would be true but $1 - 2\delta_j + \sum_{k=1}^{n} \delta_k$ is odd which would be a contradiction. Thus $e(P_\Delta) \geq n$, on the other hand if $\Delta, \Delta' \in \{0, 1\}^n$, $\Delta \neq \Delta'$ and $h(\Delta) = h(\Delta') = 1$ then $P_\Delta \neq P_{\Delta'}$. Let be

$$G' = \{ \bigcup_{\substack{h(\Delta)=1 \\ \Delta \in \{0,\}^n}} V_\Delta, \bigcup_{\substack{h(\Delta)=1 \\ \Delta \in \{0,1\}^n}} F_\Delta, \varrho \wedge (\bigcup_{\substack{h(\Delta)=1 \\ \Delta \in \{0,1\}^n}} E_\Delta), v^+, v^- \}$$

for which $e(G') \leq e(G)$ is trivially true.

In case of $e \in E_\Delta$, $h(\Delta) = 1$, $\Delta \in \{0, 1\}^n$ let $\mu(e)$ denote the number of P_Δ's which contain e, that is $\mu(e) = |\{\Delta : e \in E_\Delta, f(\Delta) = 1, \Delta \in \{0, 1\}^n\}|$ it is obvious that

$$e(G') = |\bigcup_{\substack{h(\Delta)=1 \\ \Delta \in \{0,1\}^n}} E_\Delta| = \sum_{\substack{h(\Delta)=1 \\ \Delta \in \{0,1\}^n}} \sum_{e \in E\Delta} \frac{1}{\mu(e)}.$$

In the same case let $\nu_\Delta(e)$ denote the number of different variables of X which belong to the edges of P_Δ and are less than e by the natural order, namely

$$\nu_\Delta(e) = |\{e' : e' \leq e, \nexists e'' \leq e' \text{ for which } \varrho_\Delta(e') = \varrho_\Delta(e'')\}|.$$

By Lemma 4 and by $G' < G$, $\nu_\Delta(e)$ is independent from Δ that is $\nu_\Delta(e) = \nu(e)$.

(0.2) Let us suppose that $G \in \mathcal{G}_{\mathfrak{J}0/2}$ was true then since $P_\Delta \neq P_{\Delta'}$ holds if $\Delta \neq \Delta'$ and $\mu(e) = 1$ when $e \in E_\Delta$ holds. Thus $e(G) \geq e(G') \geq n \cdot 2^{n-1}$ and so $\lambda_0(n) \geq n \cdot 2^{n-1}$.

(1.2) Let us suppose that $G \in \mathcal{G}_{\mathfrak{J}1/2}$ was true then $\mu(e)$ has 2 values at most. There are $e', e'' \in E_\Delta$ for which $\mu(e') \neq \mu(e'')$ and $k = \nu(e') = \nu(e'') - 1$ then $\mu(e) \leq \begin{cases} 2^{n-k-1} & \text{if } e \leq e' \\ 1 & \text{otherwise} \end{cases}$ because $P_\Delta \neq P_{\Delta'}$ if $\Delta \neq \Delta'$. Since

also $k \leq n-2$ holds, thus $\sum_{e \in E_\triangle} \frac{1}{\mu(e)} \geq k \cdot 2^{n-1-k} + (n-k) \geq \alpha(n)$ and so $e(G) \geq e(G') \geq 2^{n-1}\alpha(n)$ which implies that

$$\lambda_{1/2}(h) \geq 2^{n-1}\alpha(n)$$

is true. By Lemma 3

$$\lambda_1(h) \geq 2^{n-1}\alpha(n)$$

holds too.

(2.2) Let us suppose that $G \in \mathcal{G}_{j2/2}$ was true then, since $P_\Delta \neq P_{\Delta'}$ holds if $\Delta \neq \Delta'$,

$$\mu(e) \leq \begin{cases} 2^{n-1-\nu(e)} & \text{if } \nu(e) \leq n-1 \\ 1 & \text{otherwise} \end{cases}$$

is true and for every j, $1 \leq j \leq n$ there exists $e \in E_\Delta$ for which $\nu(e) = j$. Thus

$$e(G) \geq e(G') = \sum_{\substack{h(\Delta)=1 \\ \Delta \in \{0,1\}^n}} \sum_{e \in E_\Delta} \frac{1}{\mu(e)} \geq \sum_{\substack{h(\Delta)=1 \\ \Delta \in \{0,1\}^n}} \left(1 + \sum_{i=1}^{n-1} 2^{i-1-n}\right) =$$

$$= 2^{n-1} + \sum_{i=1}^{n=1} 2^i = 3 \cdot 2^{n-1} - 2$$

and so $\lambda_{2/2}(h) \geq 3 \cdot 2^{n-1} - 2$ holds. By Lemma 3 $\lambda_2(h) \geq 3 \cdot 2^{n-1} - 2$ is also true. The theorem is true for $\bar{h} = 1 - h$ in the same way.

Unicity. At the beginning of our proof $f \in \mathfrak{F}$ was an arbitrary truth function of n variables for which we have constructed $G_{\Delta(0)} \in \mathcal{G}_{jk/2}$, $k = 0, 1, 2$ and have obtained that $\lambda_{k/2}(f) \leq e(G_{\Delta(0)}) \leq \lambda_{k/2}(n)$ are true. If we now suppose that $\lambda_{k/2}(f) = \lambda_{k/2}(n)$ we get the next parities:

$$\lambda_{k/2}(f) = e(G_{\Delta(0)}) = \lambda_{k/2}(n).$$

Thus by the construction of $G_{\Delta(0)}$ we know that all paths of $G_{\Delta(0)}$ have exactly the series of variables $x_1^{\delta_1}, x_2^{\delta_2}, \ldots, x_{n-1}^{\delta_{n-1}}, x_n^{\lambda}$ where $\delta_1, \ldots, \delta_{n-1} \in \{0, 1\}^{n-1}$ and $\lambda = \lambda(\delta_1, \ldots, \delta_{n-1}) \in \{0, 1\}^{n-1} \to \{0, 1\}$ is a fixed function. Now let us consider the chromatic graph C in which the vertices are the series of $\delta_1, \ldots, \delta_{n-1}$, λ defined above. Their colour is red if $\lambda + \sum_{j=1}^{n-1} \delta_j$ is even and blue otherwise. Two vertices $\delta_1, \ldots, \delta_{n-1}, \lambda$; $\varepsilon_1, \ldots, \varepsilon_{n-1}, \lambda$ are adjacent if $\sum_{j=1}^{n-1} |\delta_j - \varepsilon_j| = 1$.

It is obvious that C is connected. Therefore, all the vertices of C is either red or blue because let us suppose that v_1 would be red and v_2 a blue vertex

of C. Since C is connected, there is a path which joines v_1 and v_2 and so there are the vertices $v_3 = \delta_1, \ldots, \delta_{n-1}, \lambda$ and $v_4 = \varepsilon_1, \ldots, \varepsilon_{n-1}, \lambda$ which are adjacent, and of different colours.

Thus $\sum_{j=1}^{n-1} | \delta_j - \varepsilon_j | = 1$ and so there exists exactly one i $(1 \leq i \leq n-1)$ for which $\delta_i \neq \varepsilon_i$ and $\lambda(\delta_0, \ldots, \delta_{n-1}) = \lambda(\varepsilon_1, \ldots, \varepsilon_{n-1})$ because v_3 and v_4 are of different colours. If we construct $G_{\Delta(0)}$ on the basis of the sequence $1, \ldots, i-1, n, i+1, \ldots, n-1, i$ and we denote this by $G'_{\Delta(0)}$ then $\lambda_{k/2}(f) \leq \leq e(G'_{\Delta(0)}) < \lambda_{k/2}(n)$ and this is a contradiction. Q.e.d.

Remark. Let $\mathcal{G}_{0/0} \subset \mathcal{G}_s$ be the following set of graphs

(i) $\mathcal{G}_0 \wedge \mathcal{G}_0 \subset \mathcal{G}_{0/0}$.

We will present a $G \in \mathcal{G}_{0/0}$ for which $f_G = h$ and $e(G) = \left[\frac{h}{2}\right] 2^{\left[\frac{n}{2}\right]} + \left\{\frac{n}{2}\right\} \cdot 2^{\left\{\frac{n}{2}\right\}}$ hold.

Let $k = [\frac{n}{2}]$ and for all $\Delta_1 = \delta_1, \ldots, \delta_k \in \{0, 1\}^k$ let be

$$P_{\Delta_1} : + \circ \overset{x_1^{\delta_1}}{\text{———}} \cdot \overset{x_2^{\delta_2}}{\text{———}} \cdot \text{———} \cdots \text{———} \cdot \overset{x_k^{\delta_k}}{\text{———}} \circ -$$

and for all $\Delta_2 = \delta_{k+1}, \ldots, \delta_n \in \{0, 1\}^{n-k}$ let be

$$P_{\Delta_2} : + \circ \overset{x_{k+1}^{\delta_{k+1}}}{\text{———}} \cdot \overset{x_{k+2}^{\delta_{k+2}}}{\text{———}} \cdot \text{———} \cdots \text{———} \cdot \overset{x_k^{\delta_n}}{\text{———}} \circ - .$$

Let G be the following graph:

$$G = [(\bigvee_{\substack{\Delta_1 \in \{0,1\}^k \\ h(\Delta_1)=1}} P_{\Delta_1}) \vee (\bigvee_{\substack{\Delta_2 \in \{0,1\}^{n-k} \\ h(\Delta_2)=0}} P_{\Delta 2})] \wedge [(\bigvee_{\substack{\Delta_1 \in \{0,1\}^k \\ h(\Delta_1)=1}} P_{\Delta_1}) \vee (\bigvee_{\substack{\Delta_2 \in \{0,1\}^{n-k} \\ h(\Delta_2)=1}} P_{\Delta_2})] .$$

It is obvious that $G \in \mathcal{G}_{0/0}$, $f_G = h$ and $e(G) \leq \left[\frac{n}{2}\right] \cdot 2^{\left[\frac{n}{2}\right]} + \left\{\frac{n}{2}\right\} \cdot 2^{\left\{\frac{n}{2}\right\}}$.

*

I am greatly indebted to G. O. H. Katona for some helpful comments.

References

1. *Shannon, C. E.*, A symbolic analysis of relay and switching circuits, Trans. AIEE, **57** (1938) 713—722.
2. *Riordan, J., Shannon, C. E.*, The number of two-terminal series-parallel networks, J. Math. and Phys. **21**, *2* (1942) 83—93.
3. *Shannon, C. E.*, The synthesis of two-terminal switching circuits, Bell Syst. Techn. J., **28**, *1* (1949) 59—98.
4. *Лупанов О. Б.*, О синтезе некоторых классов управляющих систем, Сб. проблемы кибернетики, вып. 10, Физматгиз, М., 1963, 64—94.
5. *Ádám, A.*, Truth functions and the problem of their realization by two-terminal graphs. Akadémiai Kiadó, Budapest.

О сложности релейно-контактных схем

Т. Г. ТАРЯН

(Будапешт)

Шенноном были введены понятия двухполюсной релейно-контактной схемы и последовательно-параллельной схемы ([1], [3]). Обозначим класс этих множеств символами $\mathcal{G}_s$ и $\mathcal{G}_\pi$; связь между ними характеризуется соотношением $\mathcal{G}_\pi \subset \mathcal{G}_s$. Введем обозначения для количества контактов схемы $G \in \mathcal{G}_s : e(G)$; причем для некоторой реализации $G : f_G \in \{0,1\}^n \to \{0, 1\}$ (n — некоторое натуральное число).

Пусть будет любой класс множества $\mathcal{G} \subset \mathcal{G}_s$. Тогда под термином сложности функции правдоподобия $f \in \{0, 1\}^n \to \{0, 1\}$ подразумевается такое минимальное число $\lambda_{\mathcal{G}}(f)$, для которого можно найти схему $G \in \mathcal{G}$, при которой соотношения $f_G = f$ и $e(G) = \lambda_{\mathcal{G}}(f)$ выполняются. Далее пусть число $\lambda_{\mathcal{G}}(n)$ представляет собой такое минимальное число, при котором функция правдоподобия $f \in \{0,1\}^n \to \{0, 1\}$ реализуется с помощью схемы, содержащей по крайней мере $\lambda_{\mathcal{G}}(n)$ контактов.

Известно ([1], [2], [4]), что $\lambda_{\mathcal{G}_\pi}(n) \leq 3 \cdot 2^{n-1} - 2$ для любого числа n: $\lambda_{\mathcal{G}_\pi}(n) \sim \sim \dfrac{2^n}{\log_2 n}$ и при числах $\forall\, \varepsilon, \delta > 0$ можно найти такое значение, при котором будет выполнено условие $\lambda_{\mathcal{G}_\pi}(f) \geq (1 - \varepsilon) \dfrac{2^n}{\log_2 n}$ за исключением по крайней мере единственного значения δ-функции правдоподобия $f \in \{0,1\}^n \to \{0,1\}$. В то же время ничего неизвестно об экстремальной функции $h_n \in \{0,1\}^n \to \{0,1\}$, $\lambda_{\mathcal{G}_\pi}(n) = \lambda_{\mathcal{G}_\pi}(n)$ в данном классе схем. Поэтому была сделана попытка найти такие подклассы класса $\mathcal{G}_\pi$, для которых экстремальная функция задается. Такими подклассами являются следующие:

$\mathcal{G}_0$: между двумя терминалами возможен любой «дизъюнктивный путь», для которого $\lambda_{G_0}(n) = n \cdot 2^{n-1}$;

$\mathcal{G}_{2/2}$: между двумя терминалами возможно любое «дизъюнктивное дерево», для которого $\lambda_{G_{2/2}}(n) = 3 \cdot 2^{n-1} - 2$.

В обоих случаях экстремальная функция имеет следующий вид:

$$h_n(x_1, \ldots, x_n) = \begin{cases} 1 \text{ если } \sum\limits_{i=1}^{n} x_i \text{ четно} \\ 0 \text{ если } \sum\limits_{i=1}^{n} x_i \text{ нечетно.} \end{cases}$$

Показано, что в определенном смысле класс $\mathcal{G}_{2l2}$ является самым общим классом, для которого h_n экстремальна.

T. G. Tarján
Institute of Economics
Hungarian Academy of Sciences
1051 Budapest, V., Münnich F. u. 7
Hungary

Problems of Control and Information Theory, Vol. 3 (3), pp. 197–207 (1974)

ОПРЕДЕЛЕНИЕ ФУНКЦИИ РАСПРЕДЕЛЕНИЯ ВХОДЯЩЕГО ПОТОКА В ОДНОЛИНЕЙНОЙ СИСТЕМЕ МАССОВОГО ОБСЛУЖИВАНИЯ

А. А. ВОРОНОВ, В. В. СКАЛЕЦКИЙ, Ю. В. ЧИСТЯКОВ

(Владивосток)

(Поступила в редакцию 28 ноября 1972 г.)

Рассматривается задача определения функции распределения входящего потока в однолинейной системе массового обслуживания при известных функциях распределения выходящего потока и обслуживания. Получены аналитические выражения в виде интегральных соотношений, решение которых сводится к известным численным методам массового обслуживания.

Многие экономические задачи, такие как управление запасами, совершенствование структуры управления и документооборота предприятия и т. п. могут быть интерпретированы как однолинейная система массового обслуживания. Среди них имеются такие, для решения которых необходимо знать вероятностную структуру входящего потока информации. Однако не всегда эти характеристики могут быть определены экспериментальным путем. Часто из практических соображений бывает удобнее определить характеристики выходящего потока и обслуживания, а уже через них косвенным путем найти закон, описывающий входящий поток. Для случая простейшей системы эта задача решена в работе [1]. Здесь решается аналогичная проблема для случая произвольных законов выходящего потока и обслуживания.

Постановка задачи. Рассматривается однолинейная система массового обслуживания, в которой известными являются функции распределения $B(t)$ — времени обслуживания и $C(t)$ — промежутка времени между двумя соседними моментами окончания обслуживания. Ставится задача отыскания функции распределения $A(t)$ — времени между моментами поступления вызовов в систему, т. е. требуется по выходящему потоку и обслуживанию определить входящий поток.

Решение поставленной задачи. Пусть $a(s)$, $c(s)$, $b(s)$ — преобразования Лапласа–Стилтьеса функций $A(t)$, $C(t)$, $B(t)$. Тогда при достаточно общих предположениях верно следующее равенство:

$$a(-j\omega) = \left\{ \frac{s_0 F_2^+(0)}{j\omega - s_0} + F_2^-(j\omega) \right\} e^{-F_1^-(j\omega)}, \tag{1}$$

где

$$F_1^-(j\omega) = -\frac{1}{2\pi j}\int\limits_{-\infty}^{\infty} \frac{\ln[\alpha(jx)(jx - s_0)/c]}{x - \omega}\,dx - \frac{1}{2}\ln\left[\frac{\alpha(j\omega)(j\omega - s_0)}{c}\right]; \quad (2)$$

$$F_2^{\pm}(j\omega) = \frac{1}{2\pi j}\int\limits_{-\infty}^{\infty}\left\{\frac{\exp\{F_1^-(jx)\}}{b(jx)} - \frac{jx}{B'(0)} - k\right\}\frac{dx}{x - \omega} \mp \quad (3)$$

$$\mp \frac{1}{2}\left\{\frac{\exp\{-F_1^-(j\omega)\}}{b(j\omega)} \pm \frac{j\omega}{B'(0)} \pm K\right\}.$$

Обозначения и ограничения, налагаемые на исходные функции, формулируются ниже в процессе доказательства основных соотношений.

Доказательство основных соотношений. Для эффективного решения рассматриваемого здесь класса задач будем использовать широко известные частотные методы [2]. С этой целью решение поставленной задачи выполним в следующей последовательности. Сначала получим преобразования Лапласа–Стилтьеса от искомых функций распределения по контуру, наиболее близкому к мнимой оси (совпадающему с мнимой осью за исключением конечного числа окрестностей «нежелательных» точек). Затем в полученных выражениях выполним предельный переход к контуру, совпадающему с мнимой осью. И, наконец, перейдем к значению найденных выражений на мнимой оси.

Получение основного уравнения. Как показано в [2], функции $a(s)$, $b(s)$, $c(s)$ и $w(s)$, где $w(s)$ — преобразование Лапласа–Стилтьеса функции распределения времени ожидания $\neq(t)$, удовлетворяют во всех точках, где соответствующие функции определены, равенству

$$a(-s) = \frac{c(-s) - b(-s)}{b(-s)\,b(s)\,w(s)} + \frac{1}{b(s)}. \quad (4)$$

Уравнение (4) будем называть основным. Фактически имеется одно уравнение с двумя неизвестными $w(s)$ и $a(s)$. Для решения уравнения (4) воспользуемся аналитическими свойствами входящих в него функций, используя для этого известное решение краевой задачи Римана [3] или Гильберта–Привалова [4].

Выбор контура. Желательно в качестве контура интегрирования, используемого в краевой задаче Гильберта–Привалова, выбрать мнимую ось. Однако на мнимой оси, в частности, при $\sigma \to \infty$ функция $\frac{1}{b(\sigma)}$ не удовлетворяет условию Гельдера.

Так как точка $\sigma = \infty$ является «нежелательной», исключим ее из рассмотрения, соединив концы отрезка $[-jR, jR]$ дугой радиуса R. Такую дугу можно провести в левой или правой полуплоскости, обозначая ее в этих случаях через C_R^- или C_R^+, соответственно. Полученные при этом контуры будем обозначать соответственно через $\overline{C}_R^-$ и $\overline{C}_R^+$.

Рассмотрим возможность использования обоих контуров для решения поставленной задачи.

Использование контура $\overline{C}_R^+$. Преобразуем выражение (4) к виду

$$f^-(s) = \alpha(s) f^+(s) + \beta(s), \tag{5}$$

где

$$\alpha(s) = \frac{c(-s) - b(-s)}{sb(-s)}; \quad \beta(s) = \frac{1}{b(s)}; \tag{6}$$

$$f^+(s) = \frac{s}{b(s)\, w(s)}; \quad f^-(s) = a(-s).$$

Рассмотрим следующую краевую задачу: найти функцию $f^+(s)$, аналитическую вне контура $\overline{C}_R^+$ и функцию $f^+(s)$, аналитическую внутри контура $\overline{C}_R^+$, удовлетворяющие на $\overline{C}_R^+$ граничному условию (5). Заметим, что для решения рассматриваемой здесь задачи достаточно найти $f^-(s)$ при $R \to \infty$.

Предполагая, что: 1) существует R_0 такое, что для всех $R \geq R_0$ $|a(-s)\, b(s)| < 1$ на $\overline{C}_R^+$; 2) $w(s) \neq 0$ в правой полуплоскости; 3) $b(s) \neq 0$ на $\overline{C}_R^+$, докажем, что сформулированная краевая задача есть задача Гильберта–Привалова, и для описания ее решения, следовательно, можно воспользоваться известными формулами [3]. Действительно, из доказанных в [2] соотношений следует, что

$$\alpha(s) = \frac{b(s)\, a(-s) - 1}{s}\, w(s); \tag{7}$$

$w(s) \neq 0$ по нашему предположению, $b(s)\ a(-s) - 1 = 0$ только при $s = 0$, т. к. при остальных s

$$|b(s)\, a(-s)| < 1,$$

$$\left.\frac{b(s)\, a(-s) - 1}{s}\right|_{s=0} = \int_0^\infty t\, dA(t) - \int_0^\infty t\, dB(t) > 0, \tag{8}$$

откуда $\alpha(s) \neq 0$ нигде на $\overline{C}_R^+$. Согласно доказанному в [2], функция $(b(s)\, a(-s) - 1)/s$ удовлетворяет условию Гельдера с показателем $\mu = 1$ на $\overline{C}_R^+$, откуда $\alpha(s)$ удовлетворяет этому условию. $\beta(s) = \dfrac{1}{b(s)}$ удовлетворяет условию Гельдера в силу сделанных относительно $b(s)$ предположений.

Для того, чтобы применить формулы решения задачи Гильберта–Привалова, осталось выяснить, каков индекс функции $\alpha(s)$ относительно контура $\overline{C}_R^+$.

Заметим сначала, что изменение аргумента $w(s)$ при обходе $\overline{C}_R^+$ равно нулю, так как $w(s)$ не имеет нулей ни внутри, ни на $\overline{C}_R^+$, и, следовательно, достаточно выяснить изменение аргумента функции $k(s) = \dfrac{b(s)\,a(-s)-1}{s}$ при обходе нашего контура.

При близких к нулю ω

$$\frac{b(j\omega)\,a(-j\omega)-1}{j\omega} = C_1 + C_2\,j\omega + 0(\omega)\,; \qquad \text{где}$$

$$C_1 = a - b\,;$$

$$C_2 = \frac{1}{2}\,[d_A + d_B + (a-b)^2] > 0\,;$$

d_A и d_B — дисперсии случайных величин с функциями распределения $A(t)$ и $B(t)$ соответственно, a и $в$ — соответствующие их математические ожидания. Отсюда

$$\lim_{\omega\to+0} \arg k(j\omega) = 0; \quad \lim_{\omega\to-0} \arg k(j\omega) = 2\pi\,.$$

Далее, при отличных от нуля s, лежащих на контуре $\overline{C}_R^+$, $\arg k(s) \neq 0$, так как при $s \neq 0$

$$\arg k(s) = \arg\,[b(s)\,a(-s) - 1] - \arg S\,;$$

$[b(s)a(-s)-1]$ лежит строго в левой полуплоскости, и s лежит в правой полуплоскости. Таким образом, при обходе $\overline{C}_R^+$ вектор $k(s)$ делает ровно один оборот в направлении часовой стрелки, на контуре $\overline{C}_R^+$ функция $\arg k(s)$ — непрерывна и не пересекает луч $[0, \infty]$. Следовательно, индекс $k(s)$ равен -1.

Выявленные свойства функций $\alpha(s)$ и $\beta(s)$ позволяют использоватъ решение задачи Гильберта–Привалова с двумя неопределенными константами, используя которое и получим

$$a(-s) = \left\{A_0 + \frac{A_1}{z-s_0} + F_2^-(s)\right\} \exp\{-F_1^-(s)\}\,, \quad s_0 \text{ внутри } S_R^+\,; \tag{9}$$

$$f^+(s) = \{A_0(s-s_0) + A_1 + (s-s_0)\,F_2^+(s)\} \exp\{-F_1^+(s)\}\,; \tag{10}$$

$$F_1(s) = \frac{1}{2\pi j} \int\limits_{\overline{c}^+} \frac{\ln\,[\alpha(\zeta)\,(\zeta-s_0)]\,d\zeta}{\zeta - s_0}\,; \quad F_2(s) = -\frac{1}{2\pi j} \int\limits_{\overline{c}^+} \frac{\exp\{F_1^-(\zeta)\}\,d\zeta}{b(\zeta)\,(\zeta-s)}\,.$$

MAGYAR TUDOMANYOS AKADEMIA KONYVTARA

Использование контура $\overline{C}_R^-$. Преобразуем выражение (5) к виду

$$\varphi^+(s) = \alpha(s)\,\varphi^-(s) + \beta(s)\,, \tag{11}$$

где

$$\varphi^+(s) = a(-s), \quad \varphi^-(s) = \frac{s}{b(s)\,w(s)}\,.$$

Рассмотрим следующую краевую задачу: найти функцию $\varphi^-(s)$, аналитическую вне контура $\overline{C}_R^-$, удовлетворяющую на $\overline{C}_R^-$ граничному условию (11). Напомним, что $\alpha(s) = k(s)\,w(s)$.

Индекс $k(s)$ равен нулю. Действительно, как отмечалось ранее, $k(0) = C_1 > 0$, $\arg k(0) = 0$. Покажем, что ни при каком отличном от нуля s, принадлежащем $\overline{C}_R^-$, $\arg k(s) \neq \pi$. При $s \neq 0$

$$\arg k(s) = \arg\,[b(s)\,a(-s) - 1] - \arg s\,;$$

$b(s)\,a(-s) - 1$ лежит строго в левой полуплоскости, s также лежит в левой полуплоскости, и модуль разности их аргументов строго меньше π. Следовательно, при обходе контура $\overline{C}_R^-$ вектор $k(s)$ не может сделать ни одного оборота, т. е. индекс $k(s)$ равен нулю. Учитывая это обстоятельство, хотелось бы поставленную нами задачу свести к краевой задаче Римана именно для этого контура (если бы это удалось, то можно было бы получить решение с одной неопределенной константы).[1]

Для того, чтобы воспользоваться решением задачи Римана, необходимо, чтобы функция $\varphi^+(s)$ была аналитична внутри контура $\overline{C}_R^-$, а $\varphi^-(s)$ — вне его. Однако $\varphi^-(s)$ аналитична только в конечной части правой полуплоскости, а в бесконечности имеет полюс второго порядка. Можно попытаться избавиться от особенности в бесконечности, разделив обе части равенства (11) на s^2, тогда $\varphi_1^+(s) = \dfrac{\varphi^+(s)}{s^2}$ и $\varphi_1^-(s) = \dfrac{\varphi^-(s)}{s^2}$ будут аналитичны в соответствующих областях, но $\beta_1(s) = \beta(s)/s^2$ не удовлетворяет условию Гельдера на $\overline{C}_R^-$ (при $s = 0$).

Таким образом, воспользоваться контуром $\overline{C}_R^-$ не удается.

Сведение к интегрированию по мнимой оси. Непосредственное вычисление интегралов (10) в большинстве случаев связано со значительными трудностями, поэтому желательно иметь приближенные численные методы для их нахождения.

[1] В работе [2] имеет место неточность. Контур $\overline{C}_R$ следует рассматривать именно лежащим в левой, а не в правой полуплоскости, в связи с чем выражение (4.17) принимает вид:

$$w(s) = k^{1/2}(0)\,\exp\left\{\frac{s}{2\pi j}\int\limits_{\overline{C}_R^-}\frac{\ln k(\xi)\,d\xi}{\xi(\xi - s)}\right\}.$$

Рассмотрим сначала интеграл:

$$\int\limits_{\overline{C}_R^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s}, \qquad s \quad \text{вне} \quad \overline{C}_R^+.$$

Для любого s из левой полуплоскости все особенности подинтегральной функции при некотором R_0 находятся внутри контура $\overline{C}_{R_0}^+$, поэтому для любых $R_1, R_2 \geq R_0$

$$\int\limits_{\overline{C}_{R_1}^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s} = \int\limits_{\overline{C}_{R_2}^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s},$$

и, следовательно, существует предел

$$\lim_{R\to\infty} \int\limits_{\overline{C}^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s} = \int\limits_{\overline{C}_{R_0}^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s}.$$

Заметим далее, что существует конечный предел $\lim\limits_{|\zeta|\to\infty} \ln[(\zeta - s_0)\,\alpha(\zeta)]$. Действительно,

$$(\zeta - s_0)\,\alpha(\zeta) = \frac{c(-\zeta) - b(-\zeta)}{b(-\zeta)} \cdot \frac{(\zeta - s_0)}{\zeta},$$

$$\lim_{|\zeta|\to\infty} (\zeta - s_0)\,\alpha(\zeta) = \lim_{|\zeta|\to\infty} \frac{c(-\zeta) - b(-\zeta)}{b(-\zeta)} = \frac{C'(0) - B'(0)}{B'(0)} = C,$$

$$\lim_{|\zeta|\to\infty} \ln[(\zeta - s_0)\,\alpha(\zeta)] = \ln C.$$

Далее

$$\int\limits_{\overline{C}_R^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s} = \int\limits_{\overline{C}^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - \ln C}{\zeta - s}\,d\zeta,$$

в силу того, что s лежит вне контура $\overline{C}_R^+$, и

$$\int\limits_{\overline{C}_R^+} \frac{d\zeta}{\zeta - s} = 0.$$

Итак, доказано следующее равенство:

$$\lim_{R\to\infty} \int\limits_{\overline{C}^+} \frac{\ln[(\zeta - r_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s} = \lim_{R\to\infty} \int\limits_{\overline{C}_R^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - \ln C}{\zeta - s}\,d\zeta.$$

Но $$\int\limits_{\overline{C}_R^+} \frac{\ln[(\zeta - s_0)\alpha(\zeta)] - \ln C}{\zeta - s}\, d\zeta = \int\limits_{\overline{C}_R^+} \frac{\ln[(\zeta - s_0)\alpha(\zeta)] - \ln C}{\zeta - s_0}\, d\zeta +$$

$$+ \int\limits_{jR}^{-jR} \frac{\ln[(\zeta - s_0)\alpha(\zeta)] - \ln C}{\zeta - s}\, d\zeta\,;$$

$$\lim_{R\to\infty} \int\limits_{C_R^+} \frac{\ln[(\zeta - s_0)\alpha(\zeta) - \ln C}{\zeta - s}\, d\zeta = 0\,,$$

т. к. на $\overline{C}_R^+$ при фиксированном s, $\operatorname{Re} s < 0$, причем

$$\left| \int\limits_{C_R^+} \frac{\ln[(\zeta - s_0)\alpha(\zeta)] - \ln C}{\zeta - s}\, d\zeta \right| =$$

$$= \left| \int\limits_{-\pi/2}^{\pi/2} \frac{\{\ln[(Re^{j\varphi} - s_0)\alpha(Re^{j\varphi})] - \ln C\}\, je^{j\varphi}}{e^{j\varphi} - s/R}\, d\varphi \right| \leq$$

$$\leq 2\pi \max_{C_R^+} |\ln[(Re^{j\varphi} - s_0)\alpha(Re^{j\varphi})] - \ln C|$$

для $R \geq 2\,|s|$.

Но $|\ln[(Re^{j\varphi} - s_0)\alpha(Re^{j\varphi})] - \ln C| \to 0$

при $R \to \infty$, что и требовалось показать. Таким образом, окончательно получаем, что для любого $R_1 \geq R_0$

$$\int\limits_{\overline{C}_{R_1}^+} \frac{\ln[(\zeta - s_0)\alpha(\zeta)]}{\zeta - s}\, d\zeta = \lim_{R\to\infty} \int\limits_{\overline{C}^+} \frac{\ln[(\zeta - s_0)\alpha(\zeta)]}{\zeta - s}\, d\zeta = \tag{12}$$

$$= \int\limits_{-j\infty}^{j\infty} \frac{\ln[(\zeta - s_0)\alpha(\zeta)] - \ln C}{\zeta - s}\, d\zeta, \quad \text{для} \quad Re s < 0.$$

Далее при $s = j\omega$ по формуле Сохоцкого получаем

$$F_1^-(j\omega) = -\frac{1}{2\pi j} \int\limits_{-j\infty}^{j\infty} \frac{\ln[(\zeta - s_0)\alpha(\zeta)] - \ln C}{\zeta - j\omega}\, d\zeta - \frac{1}{2} \ln \frac{(j\omega - s_0)\alpha(j\omega)}{C}\,. \tag{13}$$

Итак, интегрирование по $\overline{C}_R^+$ сведено к интегрированию по мнимой оси, для которого разработаны приближенные численные методы [2]. Аналогичным

образом поступим со вторым интегралом. Для того, чтобы свести его вычисления к интегрированию по оси, докажем существование конечного предела

$$\lim_{|s|\to\infty} s\left\{\frac{\exp\{+F_1^-(s)\}}{sb(s)}-\frac{1}{B'(0)}\right\}=K. \tag{14}$$

Действительно,

$$F_1(s)=\frac{1}{2\pi j}\int\limits_{\overline{C}^+}\frac{\ln[(\zeta-s_0)\alpha(\zeta)]}{\zeta-s}\,d\zeta.$$

Но, согласно ранее доказанному, для любого $R>R_0$

$$\int\limits_{\overline{C}_R^+}\frac{\ln[(\zeta-s_0)\alpha(\zeta)]}{\zeta-s}\,d\zeta=\int\limits_{\overline{C}_{R_0}^+}\frac{\ln[(\zeta-s_0)\alpha(\zeta)]}{\zeta-s}\,d\zeta.$$

Последний интеграл оценивается следующим образом:

$$\left|\frac{1}{2\pi j}\int\limits_{\overline{C}_{R_0}^+}\frac{\ln[(\zeta-s_0)\alpha(\zeta)]}{\zeta-s}\,d\zeta\right|\le\frac{|\overline{C}_{R_0}^+|}{2\pi}\max|\ln[(\zeta-s_0)\alpha(\zeta)]|\frac{2}{|s|}=\frac{C_0}{|s|} \tag{α}$$

при достаточно больших $|s|$ (точнее, при $|s|\ge 2R_0$). Иначе говоря, $|F_1(s)|\le\frac{C_0}{|s|}$ для любых $s:|s|\ge 2R_0$. Отсюда

$$\exp\{F_1^-(s)\}=1+F_1^-(s)+0\left(\frac{1}{|s|}\right).$$

Теперь, переходя к пределу, найдем

$$\lim_{s\to\infty} s\left\{\frac{\exp\{F_1^-(s)\}}{sb(s)}-\frac{1}{B'(0)}\right\}=\frac{\Phi B'(0)-B''(0)}{[B'(0)]^2}.$$

В результате чего (10) можно считать доказанным. Теперь перепишем выражение для $F_2(s)$ так:

$$F_2^-(s)=-\frac{1}{2\pi j}\int\limits_{\overline{C}_R^+}\left\{\frac{\exp\{F_1^-(\zeta)\}}{b(\zeta)}-\frac{\zeta}{B'(0)}-k\right\}\frac{d\zeta}{\zeta-s}+$$

$$-\frac{1}{2\pi j}\int\limits_{C_R^+}\left\{\frac{\zeta}{B'(0)}+k\right\}\frac{d\zeta}{\zeta-s}.$$

Теперь докажем, что существует конечный предел $\lim\limits_{|s|\to\infty} s\,F_1^-(s)$.

Действительно,

$$sF_1^-(s)=\frac{1}{2\pi j}\int\limits_{\overline{C}_{R_0}^+}\frac{s}{\zeta-s}\ln[(\zeta-s_0)\alpha(\zeta)]\,d\zeta.$$

Отсюда, переходя к пределу под знаком интеграла, получаем

$$\lim_{|s|\to\infty} sF_1^-(s) = -\frac{1}{2\pi j}\int\limits_{\overline{C}^+} \ln\left[(\zeta - s_0)\,\alpha(\zeta)\right] d\zeta = \Phi .$$

Далее докажем, что если почти всюду существуют ограниченные три первые производные $B(t)$, то

$$sb(s) = B'(0) + \frac{B''(0)}{s} + 0\left(\frac{1}{|s|}\right).$$

По определению имеем:

$$b(s) = \int\limits_0^\infty e^{-st}\, dB(t) = \int\limits_0^\infty e^{-st} B'(t)\, dt ,$$

откуда

$$sb(s) = B'(0) + \frac{1}{s} B''(0) + \frac{1}{s}\int\limits_0^\infty e^{-st} B'''(t)\, dt ,$$

Значит, если $|B'''(t)| \leq M$, то

$$\left| sb(s) - B'(0) - \frac{B''(0)}{s} \right| \leq \frac{M}{|s|^2},$$

что и требовалось доказать.

Перейдем к доказательству существования предела (14). Имеем

$$s\left\{\frac{\exp\{F_1^-(s)\}}{sb(s)} - \frac{1}{B'(0)}\right\} = \frac{sF_1^-(s)\,B'(0) - B''(0) + s_0\left(\frac{1}{|s|}\right)}{sb(s)\,B'(0)} .$$

Но $\int\limits_{\overline{C}_k^+} \left\{\frac{\zeta}{B'(0)} + k\right\} \frac{d\zeta}{\zeta - s} = 0$, так как подинтегральная функция не имеет особенностей внутри контура интегрирования, и, следовательно,

$$F_2^-(s) = -\frac{1}{2\pi j}\int\limits_{\overline{C}_k^+} \left\{\frac{\exp\{F_1^-(\zeta)\}}{b(\zeta)} - \frac{\zeta}{B'(0)} - k\right\} \frac{d\zeta}{\zeta - s} .$$

Аналогично предыдущему случаю можно показать, что

$$\int\limits_{\overline{C}_k^+} \left[\frac{\exp\{F_1^-(\zeta)\}}{b(\zeta)} - \frac{\zeta}{B'(0)} - k\right] \frac{d\zeta}{\zeta - s} \underset{R\to\infty}{\to} 0$$

$$F_2^-(s) = +\frac{1}{2\pi j}\int\limits_{-j\infty}^{j\infty}\left[\frac{\exp\{F_1^-(\zeta)\}}{b(\zeta)} - \frac{\zeta}{B'(0)} - k\right]\frac{d\zeta}{\zeta - s}, \tag{15}$$

где s лежит в левой полуплоскости.

Для того, чтобы вычислить $F_2(s)$ при s, лежащих на мнимой оси ($s = j\omega$), воспользуемся формулами Сохоцкого [3]. Согласно этим формулам

$$F_2^-(j\omega) = \frac{1}{2\pi j}\int\limits_{-\infty}^{\infty}\left[\frac{\exp\{F_1^-(jx)\}}{b(jx)} - \frac{jx}{B'(0)} - k\right]\frac{dx}{x-\omega} + \\ + \frac{1}{2}\left\{\frac{\exp\{F_1^-(j\omega)\}}{b(j\omega)} - \frac{j\omega}{B'(0)} - k\right\}. \tag{16}$$

Вычисление констант

Так как $a(\infty) = 0$, то (9) примет вид

$$[A_0 + F_2^-(\infty)]\, e^{-F_1^-(\infty)} = 0.$$

Но, согласно (α) $F_1^-(\infty) = 0$, аналогично устанавливаем, что $F_2^-(\infty) = 0$ и, следовательно, $A_0 = 0$.

С другой стороны

$f^+(0) = 0$ и (9) принимает вид

$$\{A_1 - s_0 F_2^+(0)\}\, e^{-F_1^+(0)} = 0.$$

В силу того, что $F_1^+(0)$ — конечная величина, то

$$A_1 = s_0 F_2^+(0).$$

В результате получаем выражение

$$a(-s) = \left\{\frac{s_0 F_2^+(0)}{s - s_0} + F_2^-(s)\right\} e^{-F_1^-(s)}.$$

Откуда, учитывая выражения (9) и (10), получаем основные соотношения (1)—(3).

Пример. Пусть $b(s) = \dfrac{\mu}{\mu + s}$; $c(s) = \dfrac{\lambda}{\lambda + s}$. Тогда $\alpha(s) = \dfrac{\mu - \lambda}{\mu(\lambda - s)}$.

Положим $s_0 = \lambda$, тогда

$$a(-s) = \left\{A_0 + \frac{A_1}{s - \lambda} + F_2^-(s)\right\} \exp\left\{-\frac{1}{2\pi j}\int\limits_{\overline{C}_R^+}\frac{\ln\dfrac{\lambda - \mu}{\mu}\, d\zeta}{\zeta - s}\right\};$$

для s, лежащих вне контура, $F_1^-(s) = 0$ по формуле Коши,

$$F_2^-(s) = \frac{1}{2\pi j} \int\limits_{\overline{C}_R^+} \frac{e^{+F_1^-(\zeta)}(\mu + \zeta)}{\mu(\zeta - s)} d\zeta = \frac{1}{2\pi j} \int\limits_{\overline{C}_R^+} \frac{\mu + \zeta}{\mu(\zeta - s)} d\zeta = 0$$

по той же причине, и $a(-s) = A_0 + \dfrac{A_1}{s - \lambda}$ для всех $s : Res < 0$. Но в силу того, что $\lim\limits_{s \to \infty} a(-s) = 0$, имеем $A_3 = 0$. Наконец, т. к. $a(0) = 1$, $\dfrac{A_1}{-\lambda} = 1$, $A_1 = -\lambda$. Значит $a(-s) = \dfrac{\lambda}{\lambda - s}$, $a(s) = \dfrac{\lambda}{\lambda + s}$ для $s : Res > 0$, что совпадает с результатами, приведенными в [1].

Литература

1. *Clarke, A. B.*, Maximum Likelihood Estimates in a Simple Queue. Ann. Math. Statist. (1957) **28**.
2. *Воронов А. А., Чистяков Ю. В.* Приближенные методы описания работы однолинейной системы массового обслуживания. Problems of Control and Information Theory. Akadémiai Kiadó, Budapest, **1** (1972) 1.
3. *Гахов К. Д.* Краевые задачи. Физматгиз, М., 1958.
4. *Лаврентьев М. А., Шабат Б. В.* Методы теории функции комплексного переменного «Наука», М., 1965.

Determination of the input distribution in a one-channel queue system

A. A. VORONOV, V. V. SKALETSKII, Yu. V. CHISTYAKOV

(Vladivostok)

This paper deals with single-server queueing system in which the distribution functions of $B(t)$, of the service time, and of $C(t)$, of the time interval between two neighbouring service ends, are known. The problem is to find the distribution function of $A(t)$, of the time interval between the calls entering the system.

The problem is to be solved in the following way. In work [2], the relationship (6) is given, through which we obtain the Laplace—Stieltjes transformations from the distribution functions sought (9)—(10) by integrating in the contour closest to the imaginary axis (coinciding with the imaginary axis excluding the finite number of the neighbourhoods of "undesirable" points). For this purpose, the problem is reduced to the Gilbert—Privalov boundary-value problem [4] with the contour $\overline{C}_R^+$. Then in the expressions obtained, first a limit transition to the contour coinciding with the imaginary axis is achieved (12)—(15) and, finally, with the aid of Sokhotsky's formulae [3], the transition to the value on the imaginary axis (13), (16).

Furthermore, an example is given where in $a(s)$ is calculated by means of formulae (12) and (15).

А. А. Воронов, В. В. Скалецкий, Ю. В. Чистяков
Институт автоматики и процессов
управления ДВНЦ АН СССР
СССР, Владивосток, ул. Суханова, 5а

Problems of Control and Information Theory, Vol. 3 (3), pp. 209—226 (1974)

TRANSFORMATION METHODS FROM PHASE-VARIABLE FORM TO CANONICAL FORMS

F. CSÁKI

(Budapest)

(Received March 15, 1974)

1. Introduction

In our age, with the broad and widening use of state-variable formulations a great deal of interest has appeared for similarity transformations. One of the basic problems is perhaps the transformation from phase-variable form to canonical forms with explicit eigenvalues. Tou [1, 2] has shown first that, when starting with the phase-variable form, transformation by means of the Vandermonde matrix results in the desired diagonal form in the system matrix for the case of distinct eigenvalues. Moreover, Tou has also given formulas for determining the inverse of the Vandermonde matrix. Many other papers were dealing with the same problem [3—7] giving some improvements to the method. For the cases involving repeated eigenvalues, the modified (or confluent) Vandermonde matrix transforms the system matrix to a Jordan canonical form. Now, the question arises how to determine the inverse of the confluent Vandermonde matrix.

Many suggestions were made for the computation of the confluent inverse Vandermonde matrix [8—16] the latter being a so-called modal matrix for the companion matrix, that is, for the system matrix in phase-variable form. Although inversion can be accomplished by usual methods, a direct evaluation of the inverse matrices in terms of system parameters is desirable. At the same time an insight into the composition of matrix elements would also be gained. In some cases one or an other method may also have considerable practical advantage.

The purpose of this paper is to present a comparison between the proposed methods. At the same time we will endeavour to show the connection links between the various methods.

In the second part of the paper we will demonstrate that inversion of the Vandermonde matrix is only an intermediary step and further transformation have to be taken into account in order to obtain the final results. An illustrative example is also given.

2. Problem formulation

Let us start from the well-known phase-variable form

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ \cdot & \cdot & \cdot & & \cdot & \cdot \\ 0 & 0 & 0 & \dots & 0 & 1 \\ -a_0 & -a_1 & -a_2 & \dots & -a_{n-2} & -a_{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \cdot \\ \cdot \\ \cdot \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix} u. \quad (1)$$

$$y = [\ b_0,\ b_1,\ b_2, \dots,\ b_{n-2}\ \ b_{n-1}\ ]\ \mathbf{x}\,.$$

The corresponding signal-flow graph is depicted in Fig. 1. Equation (1) put in short-hand notation reads

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}_0 \mathbf{x} + \mathbf{b}_0 u \\ y &= \mathbf{c}_0^{\mathrm{T}} \mathbf{x}\,. \end{aligned} \quad (2)$$

We completely omitted the term $d\,u$ in Eqs (1) and (2) for the sake of simplicity. It should be mentioned that the latter term means a parallel branch from node u to node v in Fig. 1 not influencing otherwise the main part of the signal-flow graphs or the corresponding equations.

It is to be emphasized that many other phase-variable forms exist. The form given in Eq. (1) will be called the *principle variant*, whereas any other *particular variant* may be obtained by an appropriate transformation $\mathbf{x} = \mathbf{K}\mathbf{x}'$ or by renumbering the phase variables.

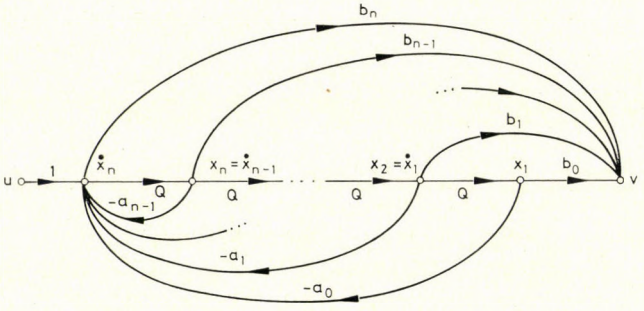


Fig. 1. State diagram of Eq. (1). Q means here an integrator

By introducing an appropriate linear transformation

$$\mathbf{x} = \mathbf{L}\mathbf{z}, \quad \mathbf{z} = \mathbf{L}^{-1}\mathbf{x} \tag{3}$$

from Eq. (2) we may obtain the Jordan canonical form (see also Fig. 2):

$$\begin{aligned} \dot{\mathbf{z}} &= \mathbf{J}\mathbf{z} + \mathbf{b}u \\ y &= \mathbf{c}^{\mathrm{T}}\mathbf{z} \,. \end{aligned} \tag{4}$$

Here, the Jordan matrix is given in pseudodiagonal form: as a hypermatrix

$$\mathbf{J} = \mathbf{L}^{-1}\mathbf{A}_0\mathbf{L} = \mathbf{diag}\,[\mathbf{J}_1, \mathbf{J}_2, \ldots, \mathbf{J}_m] \tag{5}$$

where $k_1 + k_2 + \ldots + k_m = n$. Furthermore,

$$\mathbf{b} = \mathbf{L}^{-1}\mathbf{b}_0, \quad \mathbf{c}^{\mathrm{T}} = \mathbf{c}_0^{\mathrm{T}}\mathbf{L} \,. \tag{6}$$

In Eq. (5) the $k_i \times k_i$ matrix

$$\mathbf{J}_i = \begin{bmatrix} \lambda_i & 1 & & & & \\ & \lambda_i & 1 & & & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & \lambda_i & 1 \\ & & & & & \lambda_i \end{bmatrix} \quad (i = 1, 2, \ldots m) \tag{7}$$

is a so-called Jordan block containing the eigenvalue λ_i with multiplicity k_i.

If some of the eigenvalues is distinct ($k_i = 1, \exists\, i$) then the corresponding Jordan block reduces to a scalar quantity which is not else than the eigenvalue itself, $\mathbf{J}_i = \lambda_i, \exists\, i$.

For short, transformation (3) with result (5) will be called *modal transformation*, and $\mathbf{L}$ is then called a *modal matrix* $\mathbf{M}$, $\mathbf{L} = \mathbf{M}$.

Attention is focused on the problem how to obtain $\mathbf{L}$ and furthermore $\mathbf{L}^{-1}$ in order to get Eq. (5).

As it is well known the original Vandermonde matrix is a modal matrix yielding a diagonal form, furthermore an appropriate modal transformation matrix for satisfying Eq. (5) is the confluent Vandermonde matrix:

$$\mathbf{V} = [\mathbf{V}_1, \mathbf{V}_2, \ldots, \mathbf{V}_m] \tag{8}$$

where

$$\mathbf{V}_i = \begin{bmatrix} 1 & 0 & 0 & \dots \\ \lambda_i & 1 & 0 & \dots \\ \lambda_i^2 & 2\lambda_i & 1 & \\ \vdots & \vdots & \vdots & \\ \lambda_i^{n-1} & (n-1)\,\lambda_i^{n-2} & \frac{1}{2!}(n-1)(n-2)\,\lambda^{n-3} & \dots \end{bmatrix} \widehat{=}$$

$$\widehat{=} \begin{bmatrix} \mathbf{v}_i, & \mathbf{v}_1', & \frac{1}{2!}\mathbf{v}_i'', & \dots \end{bmatrix} \tag{9}$$

is an $n \times k_i$ Vandermonde block with corresponding $n \times 1$ vectors $\mathbf{v}_i$ and its derivatives.

Relatively few is spoken about $\mathbf{b}$, that is, how to determine $\mathbf{L}$ and $\mathbf{L}^{-1}$ in order that when satisfying Eq. (5) at the same time also $\mathbf{b}$ should get a certain special form. The latter problem was treated in [8] in some extent. It can be shown that $\mathbf{b}$ may get the very simple form:

$$\mathbf{b} = \begin{bmatrix} \mathbf{j}_1 \\ \mathbf{j}_2 \\ \vdots \\ \mathbf{j}_{m-1} \\ \mathbf{j}_m \end{bmatrix} = \mathbf{j}\,; \tag{10}$$

with

$$\mathbf{j}_i = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \qquad (i = 1, 2, \dots, m) \tag{11}$$

where $\mathbf{j}_i$ is a $k_i \times 1$ matrix.

For Eqs (10), (11), (5) and (7) we obtain the *principal variant* of the Jordan form. Any other case will be called a *particular case.*

It should, however, be emphasized that Vandermonde matrices $\mathbf{V}$, or other modal matrices $\mathbf{M}$, do not fulfil automatically neither Eq. (10) nor Eq. (11):

$$\mathbf{V}^{-1}\mathbf{b}_0 \neq \mathbf{b}, \quad \mathbf{M}^{-1}\mathbf{b}_0 \neq \mathbf{b}\,,$$

For this purpose special auxiliary matrices $\mathbf{T}$ must be introduced such that in general

$$\mathbf{L} = \mathbf{MT}, \quad \mathbf{L}^{-1} = \mathbf{T}^{-1}\mathbf{M}^{-1} \tag{12a}$$

or specifically

$$\mathbf{L} = \mathbf{VT}, \quad \mathbf{L}^{-1} = \mathbf{T}^{-1}\mathbf{V}^{-1} \tag{12b}$$

with Vandermonde matrix $\mathbf{V}$ as given in Eq. (8). In such cases

$$\mathbf{T}^{-1}\mathbf{M}^{-1}\mathbf{b}_0 = \mathbf{b} \quad \text{or} \quad \mathbf{T}^{-1}\mathbf{V}^{-1}\mathbf{b}_0 = \mathbf{b} \tag{13}$$

where $\mathbf{b}$ is given in Eqs (10) and (11).

After this introductory remarks we will concentrate our attention to the determination of the Jordan canonical form and especially to the computation of $\mathbf{V}^{-1}$.

3. Inversion of confluent Vandermonde matrices

In the last few years many successful attempts have been made concerning the inversion of the so-called Vandermonde matrix [11—16]. Some of the methods [11, 13] are based on Hermite polynomials aiming at the determination of the rows of the inverse matrix. (The problem is skilfully solved in [11] by recursive formulas.) Other methods (e.g. [9]) are based on Laplace transforms endeavouring the determination of the columns of the inverse matrix.

In this paper we will show that the various methods, appearing at first sight quite different, have the same explanation. The general formulation given in this paper contains all the results obtained earlier (e.g. in [9, 11, 13]).

As concerns the Vandermonde matrix, we start from the hyper-matrix form as given in Eqs (8) and (9).

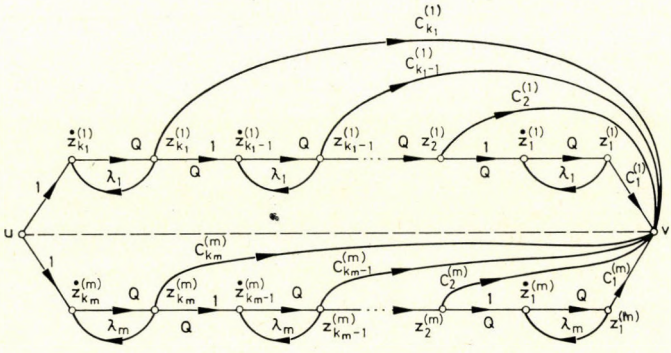


Fig. 2. State diagram of Eqs (4) and (10). Q means here an integrator

The inverse of the Vandermonde matrix is also required in a hypermatrix form

$$\mathbf{W} = \mathbf{V}^{-1} = \begin{bmatrix} \mathbf{W}_1 \\ \mathbf{W}_2 \\ \vdots \\ \mathbf{W}_i \\ \vdots \\ \mathbf{W}_m \end{bmatrix} \tag{14}$$

where

$$\mathbf{W}_i = \begin{bmatrix} \mathbf{w}_1^{T(i)} \\ \mathbf{w}_2^{T(i)} \\ \vdots \\ \mathbf{w}_{k_i}^{T(i)} \end{bmatrix} \tag{15}$$

and $\mathbf{w}_j^T$ $(j = 1, 2, \ldots, k_i)$ are the $1 \times n$ row vectors of the block $\mathbf{W}_i$.

The characteristic equation belonging to Eq. (1) or Eq. (2) can be expressed by $|s\mathbf{I} - \mathbf{A}_0| = 0$ as

$$D(s) \mathrel{\widehat{=}} s^n + a_{n-1} s^{n-1} + \ldots a_1 s + a_0 \tag{16}$$

or

$$D(s) \mathrel{\widehat{=}} \prod_{i=1}^{m} (s - \lambda_i)^{k_i} \tag{17}$$

where λ_i, $i = 1, 2, \ldots, m$ are the eigenvalues with multiplicity k_i.

Let us now define the so-called truncated polynomials by

$$N_l(s) = s^{n-l} + a_{n-1} s^{n-l-1} + \ldots + a_{l+1} s + a_l \tag{18}$$
$$(l = 1, 2, \ldots, n) .$$

By partial fraction expansion of $N_l(s)/D(s)$ it can be shown (see the Appendix) that the element in the jth row and lth column of the ith block, $\mathbf{W}_i$, is

$$w_{jl}^{(i)} = \frac{1}{(k_i - j)!} \frac{d^{k_i - j}}{ds^{k_i - j}} \frac{N_l(s)}{D_i(s)} \bigg|_{s=\lambda_i} \tag{19}$$

where

$$D_i(s) \mathrel{\widehat{=}} \frac{D(s)}{(s - \lambda_i)^{k_i}} . \tag{20}$$

This is essentially the method proposed in [9] but obtained here by a somewhat modified procedure of thinking. Let us define the inverse of the expression in Eq. (20)

$$\Delta_i(s) = \frac{1}{D_i(s)}. \tag{21}$$

Applying the Leibnitz rule of derivation the element of the inverse matrix can also be expressed as

$$w_{jl}^{(i)} = \sum_{k=1}^{k_i-j} \left[\frac{1}{(k_i - j - k)!} N_l^{(k_i-j-k)}(s) \right] \left[\frac{1}{k!} \Delta_i^{(k)}(s) \right] \Bigg|_{s=\lambda_i}. \tag{22}$$

Let us define furthermore

$$d_k^{(i)} \triangleq \frac{1}{(k-1)!} D_i^{(k-1)}(s) \Bigg|_{s=\lambda_i} \tag{23}$$

and

$$\delta_k^{(i)} \triangleq \frac{1}{(k-1)!} \Delta_i^{(k-1)}(s) \Bigg|_{s=\lambda_i}. \tag{24}$$

By the way, the determination of $d_k^{(i)}$ being easier, the following recurrent relationship can also be applied

$$\delta_j^{(i)} = -\frac{1}{d_1^{(i)}} \sum_{p=1}^{j-1} d_{j+1-p}^{(i)} \delta_p^{(i)} \tag{25}$$

$$(j = 2, 3, \ldots, k_i)$$

and

$$\delta_1^{(i)} \triangleq \frac{1}{d_1^{(i)}}. \tag{26}$$

Based on the expression of $w_{jl}^{(i)}$ in Eq. (22), the inverse Vandermonde matrix can be expressed in the form

$$\mathbf{W} = \mathbf{D}^{-1} \mathbf{Q} \tag{27}$$

where

$$\mathbf{Q} = \begin{bmatrix} \mathbf{Q}_1 \\ \mathbf{Q}_2 \\ \vdots \\ \mathbf{Q}_i \\ \vdots \\ \mathbf{Q}_m \end{bmatrix} \tag{28}$$

and

$$\mathbf{Q}_i = \begin{bmatrix} N_1(\lambda_i), & \dots & N_2(\lambda_i) & & N_n(\lambda_i) \\ N_1'(\lambda_i), & \dots & N_2'(\lambda_i) & & N_n'(\lambda_i) \\ \vdots & \dots & \vdots & & \vdots \\ \frac{1}{(k_i-1)!} N_1^{(k_i-1)}(\lambda_i), & & \frac{1}{(k_i-1)!} N_2^{(k_i-1)}(\lambda_i), & \dots, & \frac{1}{(k_i-1)!} N_n^{(k_i-1)}(\lambda_i) \end{bmatrix}. \tag{29}$$

Furthermore, $\mathbf{D}^{-1}$ is a pseudodiagonal matrix

$$\mathbf{D}^{-1} = \operatorname{diag}[\mathbf{D}_1^{-1}, \mathbf{D}_2^{-1}, \dots, \mathbf{D}_i^{-1}, \dots, \mathbf{D}_m^{-1}] \tag{30}$$

where $\mathbf{D}_i^{-1}$ $(i = 1, 2, \dots, m)$ are special upper left-hand triangular matrices with equal elements in the subsidiary diagonals

$$\mathbf{D}_i^{-1} = \begin{bmatrix} \delta_{k_i}^{(i)} & \dots & \delta_2^{(i)} & \delta_1^{(i)} \\ \vdots & \ddots & \ddots & \\ \delta_2^{(i)} & \delta_1^{(i)} & & \\ \delta_1^{(i)} & & & \end{bmatrix}. \tag{31}$$

By introducing

$$q_{n-l}^{(p)}(\lambda_i) = \frac{1}{p!} N_l^{(p)}, \quad (p = 0, 1, 2, \dots, k_i - 1) \tag{32}$$

we may obtain the result exactly in the same form as given in [11].

Finally, we remark that the elements of the inverse Vandermonde matrix can also be obtained from the coefficients of special Hermite interpolation polynomials

$$P_{j(i)}(s) = \sum_{l=1}^{n} w_{jl}^{(i)} s^{l-1}. \tag{33}$$

We will call these polynomials Hermite—Kronecker polynomials because of the following characteristics:

$$\frac{1}{q!} P_{j(i)}^{(q)}(\lambda_k) = \delta_{ik}\, \delta_{j,q+1} \tag{34}$$

$$(i, k = 1, 2, \dots, m), \ (j = 1, 2, \dots, k_i), \ (q = 0, 1 \dots, k_i - 1)$$

where δ_{ik} is the Kronecker symbol ($\delta_{ik} = 0$ for $i \neq k$, $\delta_{ik} = 1$ for $i = k$) and superscript q means derivation. These properties of the Hermite—Kronecker polynomials are based on $\mathbf{WV} = \mathbf{V}^{-1}\mathbf{V} = \mathbf{I}$, or

$$\mathbf{w}_j^{T(i)} \frac{1}{(h-1)!} \mathbf{v}_k^{(h-1)} = \delta_{ik}\,\delta_{jh} \tag{35}$$

$$(i = 1, 2, \ldots, m),\ (h, j, k = 1, 2, \ldots, k_i)\,.$$

Assuming,

$$\frac{P_{j(i)}(s)}{D(s)} = \sum_{i'=1}^{m} \sum_{j'=1}^{k_i} \frac{A_{j,j'}^{(i,i')}}{(s-\lambda_{i'})^{k_{i'}-j'+1}} \tag{36}$$

or

$$P_{j(i)}(s) = \sum_{i'=1}^{m} \sum_{j'=1}^{k_i} A_{j,j'}^{(i,i')}(s-\lambda_{i'})^{j'-1} D_{i'}(s) \tag{37}$$

the coefficients can be obtained by Eq. (34).

It can easily be shown that the coefficients other than zero can be expressed as

$$A_{1,k}^{(i,i)} = A_{2,k+1}^{(i,i)} = \ldots = A_{j,k+j-1}^{(i,i)} = \delta_k^{(i)} = \frac{1}{(k-1)!} \Delta_i^{(k-1)}(\lambda_i) \tag{38}$$

$$(j, k = 1, 2, \ldots, k_i)\,.$$

Expanding Eq. (37) according to the powers of s, Eq. (33) and the elements of the inverse Vandermonde matrix are obtained. This is essentially the same method as given in [13].

As a final conclusion, Eq. (22) can be considered the most general expression for determining the elements of the inverse Vandermonde matrix. The formula in Eq. (22) contains all the previously published results obtained by quite different methods.

For distinct eigenvalues Eq. (22) reduces to

$$w_{1l}^{(i)} = N_l(\lambda_i)\,\Delta_i(\lambda_i) = \frac{N_l(\lambda_i)}{D_i(\lambda_i)} \tag{39}$$

where for distinct eigenvalues,

$$D_i(\lambda_i) = \frac{1}{\Delta_i(\lambda_i)} = \prod_{\substack{j=1 \\ j\neq i}}^{n} (s_i - \lambda_j)\,.$$

4. Determination of final transformation matrices

Let us return to Eqs (12) and (13). As was shown we have to introduce an auxiliary matrix $\mathbf{T}$ and its inverse such that Eqs (12b) and (13) hold with $\mathbf{b} = \mathbf{j}$ as given in Eqs (10) and (11).

Furthermore,

$$\mathbf{T}^{-1}\,\mathbf{J}\mathbf{T} = \mathbf{J} \tag{40}$$

or

$$\mathbf{J}\mathbf{T} = \mathbf{T}\mathbf{J}, \quad \mathbf{T}^{-1}\,\mathbf{J} = \mathbf{J}\mathbf{T}^{-1} \tag{41}$$

must also hold. According to Eq. (40) we will call auxiliary matrix $\mathbf{T}$ also as *commutativity matrix.*

As the Jordan matrix is given in pseudodiagonal form so are also $\mathbf{T}$ and $\mathbf{T}^{-1}$

$$\begin{aligned} \mathbf{T} &= \mathbf{diag}\,[\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_m] \\ \mathbf{T}^{-1} &= \mathbf{diag}\,[\mathbf{T}_1^{-1}, \mathbf{T}_2^{-1}, \ldots, \mathbf{T}_m^{-1}] \end{aligned} \tag{42}$$

with

$$\mathbf{T}_i^{-1}\,\mathbf{J}_i\,\mathbf{T}_i = \mathbf{J}_i \tag{43}$$

and

$$\mathbf{J}_i\,\mathbf{T}_i = \mathbf{T}_i\,\mathbf{J}_i, \quad \mathbf{T}_i^{-1}\,\mathbf{J}_i = \mathbf{J}_i\,\mathbf{T}_i^{-1}. \tag{44}$$

Evaluating the latter relationships in Eq. (43), for the elements $t_{kl}^{(i)}$ of $\mathbf{T}_i$ and elements $\tau_{kl}^{(i)}$ of $\mathbf{T}^{-1}$ it becomes obvious that both $\mathbf{T}_i$ and $\mathbf{T}_i^{-1}$ are upper right-hand triangular matrices, with equal elements in the main diagonals

$$\mathbf{T}_i = \begin{bmatrix} t_1^{(i)} & t_2^{(i)} & \cdots & t_{k_i}^{(i)} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \vdots \\ & & t_1^{(i)} & t_2^{(i)} \\ & & & t_1^{(i)} \end{bmatrix}, \tag{45}$$

$$\mathbf{T}_i^{-1} = \begin{bmatrix} \tau_1^{(i)} & \tau_2^{(i)} & \cdots & \tau_{k_i}^{(i)} \\ & \ddots & \ddots & \vdots \\ & & \ddots & \vdots \\ & & \tau_1^{(i)} & \tau_2^{(i)} \\ & & & \tau_1^{(i)} \end{bmatrix}. \tag{46}$$

After some computations, based on Eq. (13), it can be shown that

$$\begin{aligned} t_1^{(i)} &= w_{k_i,n}^{(i)} \\ t_2^{(i)} &= w_{k_i-1,n}^{(i)} \\ &\vdots \\ t_{k_i-1}^{(i)} &= w_{2n}^{(i)} \\ t_{k_i}^{(i)} &= w_{1n}^{(i)} . \end{aligned}$$

The elements of the inverse matrix can be obtained by recurrent relationships

$$\tau_1^{(i)} = \frac{1}{t_1^{(i)}}$$

$$\tau_j^{(i)} = -\frac{1}{t_1^{(i)}} \sum_{l=1}^{j-1} t_{l+1}^{(i)} \tau_{j-l}^{(i)} \qquad (j = 2, 3, \ldots, k_i) \tag{48}$$

or

$$\tau_1^{(i)} = \frac{1}{w_{k_i,n}^{(i)}}$$

$$\tau_j^{(i)} = -\frac{1}{w_{k_i,n}^{(i)}} \sum_{k=1}^{j-1} w_{k_i-h,n}^{(i)} \tau_{j-h}^{(i)} \quad (j = 2, 3, \ldots k_i) . \tag{49}$$

Holding Eqs (45) and (46), from Eq. (13) the final transformation matrix, $\mathbf{L}$ and its inverse, $\mathbf{L}^{-1}$ can be obtained. Then, Eq. (6) yields the input and output matrices, respectively.

Starting from the characteristic equation (16) and definitions (20) and (21), after some computation it can be shown that

$$\tau_k^{(i)} = \frac{1}{(k-1)!} D_i^{(k-1)}(\lambda_i) \tag{50}$$

and

$$t_k^{(i)} = \frac{1}{(k-1)!} \Delta_i^{(k-1)}(\lambda_i) \tag{51}$$

where superscripts $(k - 1)$ mean derivations.

Comparing the latter results with Eqs (23) and (24), the following results come out

$$\begin{aligned} \tau_k^{(i)} &= d_k^{(i)} \\ t_k^{(i)} &= \delta_k^{(i)} . \end{aligned} \tag{52}$$

*

Because of these properties

$$\mathbf{T}_i^{-1}\mathbf{D}_i^{-1} = \begin{bmatrix} \tau_1^{(i)} & \tau_2^{(i)} & \cdots & \tau_{k_i}^{(i)} \\ \ddots & \ddots & & \vdots \\ & \ddots & \ddots & \vdots \\ & & \tau_1^{(i)} & \tau_2^{(i)} \\ & & & \tau_1^{(i)} \end{bmatrix} \begin{bmatrix} \delta_{k_i}^{(i)} & \cdots & \delta_2^{(i)} & \delta_1^{(i)} \\ \vdots & \ddots & \ddots & \\ \vdots & \ddots & & \\ \delta_2^{(i)} & \delta_1^{(i)} & & \\ \delta_1^{(i)} & & & \end{bmatrix} = \tag{53}$$

$$= \begin{bmatrix} & & & & 1 \\ & & & \cdot & \\ & & \cdot & & \\ & \cdot & & & \\ 1 & & & & \\ 1 & & & & \end{bmatrix} \widehat{=} \mathbf{E}_i$$

the latter being a row changing matrix and

$$\mathbf{T}^{-1}\mathbf{D}^{-1} = \mathbf{diag}\,[\mathbf{E}_1, \mathbf{E}_2, \ldots, \mathbf{E}_m]. \tag{54}$$

The inverse Vandermonde matrix is expressed by Eq. (27) and so

$$\mathbf{L}^{-1} = \mathbf{T}^{-1}\mathbf{W} = \mathbf{T}^{-1}\mathbf{D}^{-1}\mathbf{Q} = \begin{bmatrix} \mathbf{N}_1 \\ \mathbf{N}_2 \\ \vdots \\ \mathbf{N}_m \end{bmatrix} \tag{55}$$

where

$$\mathbf{N}_i = \begin{bmatrix} \frac{1}{(k_i-1)!} N_1^{(k_i-1)}(\lambda_i), & \frac{1}{(k_i-1)!} N_2^{(k_i-1)}(\lambda_i), \ldots, & \frac{1}{(k_i-1)!} N_n^{(k_i-1)}(\lambda_i) \\ \vdots & \vdots & \vdots \\ N_1'(\lambda_i), & N_2'(\lambda_i), \quad \ldots, & N_n'(\lambda_i) \\ N_1(\lambda_i), & N_2(\lambda_i), \quad \ldots, & N_n(\lambda_i) \end{bmatrix}. \tag{56}$$

On the other hand by Eqs (12b) and (42)

$$\mathbf{L} = \mathbf{VT} = \mathbf{V}\,\mathbf{diag}\,[\mathbf{T}_1, \mathbf{T}_2, \ldots, \mathbf{T}_m]. \tag{57}$$

Although **L** in Eq. (57) is more complicated than the original Vandermonde matrix **V**, its inverse, as given in Eqs (55) and (56), is much easier obtainable.

5. An illustrative example

For the case

$$D(s) = (s+1)^3(s+2)^2(s+3)$$

the original confluent Vandermonde matrix is given as

$$\mathbf{V} = \begin{bmatrix} 1 & 0 & 0 & 1 & 0 & 1 \\ -1 & 1 & 0 & -2 & 1 & -3 \\ 1 & -2 & 1 & 4 & -4 & 9 \\ -1 & 3 & -3 & -8 & 12 & -27 \\ 1 & -4 & 6 & 16 & -32 & 81 \\ -1 & 5 & -10 & -32 & 80 & -243 \end{bmatrix}$$

and its inverse is (see [11])

$$\mathbf{W} = \begin{bmatrix} 16.5 & 58.0 & 83.125 & 56.375 & 17.875 & 2.125 \\ -9.0 & -36.0 & -52.25 & -34.75 & -10.75 & -1.25 \\ 6.0 & 20.0 & 25.5 & 15.5 & 4.5 & 0.5 \\ -15.0 & -56.0 & -80.0 & -54.0 & -17.0 & -2.0 \\ -6.0 & -23.0 & -34.0 & -24.0 & -8.0 & -1.0 \\ -0.5 & -2.0 & -3.125 & -2.375 & -0.875 & -0.125 \end{bmatrix}.$$

By Eq. (51) we have

$$t_1^{(1)} = 0.5; \quad t_2^{(1)} = -1.25; \; t_3^{(1)} = 2.125$$
$$t_1^{(2)} = -1.0; \; t_2^{(2)} = -2.0$$
$$t_1^{(3)} = -0.125$$

and by Eq. (57) we obtain

$$\mathbf{L} = \begin{bmatrix} 0.5 & -1.25 & 2.125 & -1.0 & -2.0 & -0.125 \\ -0.5 & 1.75 & -3.375 & 2.0 & 3.0 & 0.375 \\ 0.5 & -2.25 & 5.125 & -4.0 & -4.0 & -1.125 \\ -0.5 & 2.75 & -7.375 & 8.0 & 4.0 & 3.375 \\ 0.5 & -3.25 & 10.125 & -16.0 & 0.0 & -10.125 \\ -0.5 & 3.75 & -13.375 & 32.0 & -16.0 & 30.375 \end{bmatrix}.$$

On the other hand, Eqs (55) and (56) yield

$$\mathbf{L}^{-1} = \begin{bmatrix} 12 & 16 & 7 & 1 & 0 & 0 \\ 12 & 28 & 23 & 8 & 1 & 0 \\ 12 & 40 & 51 & 31 & 9 & 1 \\ 3 & 10 & 12 & 6 & 1 & 0 \\ 6 & 23 & 34 & 24 & 8 & 1 \\ 4 & 16 & 25 & 19 & 7 & 1 \end{bmatrix}.$$

In the first case we have $\mathbf{b} = [2.125, -1.25, 0.5, -2.0, -1.0, -0.125]^T$ whereas in the second case $\mathbf{b} = [0, 0, 1, 0, 1, 1]^T$. The Jordan matrices are the same in both cases.

6. Final conclusions

The purpose of this paper was to compare the various methods performing the transformation from phase-variable form to the Jordan canonical form, in the case of multiple eigenvalues, or to diagonal form, in the case of distinct eigenvalues. In such cases the confluent or original Vandermonde matrix may play an important role: that of a modal matrix. The determination of the inverse Vandermonde matrix is not a trivial one. Previously three various methods were elaborated to obtain it.

In this paper a general formula is given in Eq. (19) or in detailed form in Eg. (22) which contains all the previous results.

It is also shown that the Vandermonde matrix itself is not the best transformation matrix and introduction of auxiliary commutativity matrices becomes necessary.

The final transformation matrix is given in Eq. (57) whereas its inverse is shown in Eqs (55) and (56).

A worked example illustrates the theoretical considerations.

Appendix 1

Let us start from the phase-variable form given as

$$\dot{\mathbf{x}}(t) = \mathbf{A}_0\mathbf{x}(t) \,. \tag{A-1}$$

Taking the Laplace transform

$$s\mathbf{X}(s) = \mathbf{A}_0\,\mathbf{X}(s) + \mathbf{x}(0) \tag{A-2}$$

Applying the modal transformation

$$\begin{gathered} \mathbf{x}(t) = \mathbf{V}\,\mathbf{z}(t), \quad \mathbf{z}(t) = \mathbf{V}^{-1}\mathbf{x}(t) = \mathbf{W}\,\mathbf{x}(t) \\ \mathbf{X}(s) = \mathbf{V}\,\mathbf{Z}(s), \; \mathbf{Z}(s) = \mathbf{V}^{-1}\mathbf{X}(s) = \mathbf{W}\,\mathbf{X}(s) \end{gathered} \tag{A-3}$$

we may write

$$s\mathbf{Z}(s) = \mathbf{J}\,\mathbf{Z}(s) + \mathbf{W}\,\mathbf{x}(0) \tag{A-4}$$

where $\mathbf{J} = \mathbf{V}^{-1}\mathbf{A}_0\mathbf{V} = \mathbf{W}\,\mathbf{A}_0\mathbf{V}$. The above matrix equation can be decomposed for each Jordan block yielding

$$s\mathbf{Z}^{(i)}(s) = \mathbf{J}_i\,\mathbf{Z}^{(i)}(s) + \mathbf{W}_i\,\mathbf{x}(0) \,. \tag{A-5}$$

Written in detailed form reads

$$\begin{aligned} sZ_1^{(i)}(s) &= \lambda_i Z_1^{(i)}(s) + Z_2^{(i)}(s) + \mathbf{w}_1^{T(i)}\,\mathbf{x}(0) \\ sZ_2^{(i)}(s) &= \lambda_i Z_2^{(i)}(s) + Z_3^{(i)}(s) + \mathbf{w}_2^{T(i)}\,\mathbf{x}(0) \\ &\vdots \\ sZ_{k_i}^{(i)}(s) &= \lambda_i Z_{k_i}^{(i)}(s) + \mathbf{w}_{k_i}^{T(i)}\,\mathbf{x}(0)\,. \end{aligned} \tag{A-6}$$

Successive substitution yields

$$Z_1^{(i)}(s) = \sum_{j=1}^{k_i} \frac{\mathbf{w}_j^{T(i)}\,\mathbf{x}(0)}{(s-\lambda_i)^j}\,. \tag{A-7}$$

Now, taking the structure of the Vandermonde matrix into consideration, the modal transformation yields

$$X_1(s) = \sum_{i=1}^{m} Z_1^{(i)}(s) \tag{A-8}$$

or

$$X_1(s) = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \frac{\mathbf{w}_j^{T(i)}\,\mathbf{x}(0)}{(s-\lambda_i)^j} \tag{A-9}$$

which can also be expressed as

$$X_1(s) = \sum_{i-1}^{m} \sum_{j-1}^{k_i} \frac{\sum_{l=1}^{n} w_{jl}^{(i)}\,x_l(0)}{(s-\lambda_i)^j}\,. \tag{A-10}$$

On the other hand, from the Laplace transform of the phase-variable form, by successive substitutions we can obtain

$$X_1(s) = \frac{\sum_{l=1}^{n} N_l(s)\,x_l(0)}{D(s)}\,. \tag{A-11}$$

Setting $x_k(0) = 0$ for $k \neq l$, and $x_k(0) = 1$ for $k = l$ as well as comparing the two expressions of $X_1(s)$ we obtain

$$\frac{N_l(s)}{D(s)} = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \frac{w_{jl}^{(i)}}{(s-\lambda_i)^j}\,. \tag{A-12}$$

Thus the elements of the inverse Vandermonde matrix can be calculated by direct conventional methods based on partial fraction expansion techniques, as shown in Eq. (19).

Finally, we do mention that inverse Laplace transform yields

$$x_1(t) = \sum_{i=1}^{m} \sum_{j=1}^{k_i} \sum_{l=1}^{n} w_{jl}^{(i)} \frac{t^{j-l}}{(j-1)!} e^{\lambda_i t} x_l(0) . \tag{A-13}$$

It is readily seen that from a physical point of view $w_{jl}^{(i)}$ is the coefficient of the ith modal transient process subjected to the initial condition $x_l(0)$ and belonging to the jth multiplicity.

Appendix 2

By the way the elements of matrix $\mathbf{Q}_i$ in Eq. (29) can obtained according to Horner's scheme by

$$\begin{aligned} N_n(\lambda_i) &= 1 \\ N_l(\lambda_i) &= N_{l+1}(\lambda_i)\lambda_i + a_l \\ (l &= n-1, \ldots, 2, 1) \end{aligned} \tag{A-14}$$

and, furthermore,

$$\begin{aligned} N_p^{(p)}(\lambda_i) &= 1 \\ N_l^{(p)}(\lambda_i) &= N_{l+1}^{(p)}(\lambda_i)\,\lambda_i + N_{l+1}^{(p-1)}(\lambda_i) \\ (p = 1, 2, \ldots, k_i - 1) \quad &(l = n-2, \ldots, p) . \end{aligned} \tag{A-15}$$

References

1. *Tou, J. T.*, Determination of the inverse Vandermonde matrix. IEEE Trans. on Automatic Control, **AC—9** (1964, July) (Correspondence.)
2. *Tou, J. T.*, Modern Control Theory. McGraw-Hill, 1964.
3. *Brulé, J. D.*, A note on the Vandermonde determinant. IEE Trans. on Automatic Control, **AC—9** (1964, July) 314—315. (Correspondence.)
4. *Wertz, H. J.*, On the numerical inversion of the recurrent problem: the Vandermonde matrix. IEEE Trans. on Automatic Control. **AC—10** (1965, Oct.) 492. (Correspondence.)
5. *Wu, S. H.*, On the inverse of Vandermonde matrix. IEEE Trans. on Automatic Control, **AC—11** (1966, Oct.) 769. (Correspondence.)
6. *Reis, G. C.*, A matrix formulation for the inverse Vandermonde matrix. IEEE Trans. on Automatic Control, **AC—12** (1967, Dec.) 793. (Correspondence.)
7. *Syslo, M.*, The inverse for the real similarity transform of a companion matrix. IEEE Trans. on Automatic Control, **AC—16** (1971, Oct.) 491—492.
8. *Csáki, F.*, Some remarks on modal transformations. Periodica Polytechnica, Electrical Engineering, **16** (1972) 333—346.
9. *Beck, R., Lance, G. M.*, Direct and inverse transformations between phase variable and canonical forms. Joint Automatic Control Conference of the American Automatic Council. Stanford, 1972, pp. 133—137.
10. *Beck, R.*, Analysis and pole-zero synthesis using state variable techniques. Ph. D. Dissertation, The University of Iowa. May, 1972.

11. *Kaufman, I.*, The inversion of the Vandermonde matrix and the transformation to the Jordan canonical form. IEEE Trans. on Automatic Control, **AC—14** (1969, Dec.) 774—777. (Correspondence.)
12. *Kan, P. E.*, An inversion procedure of the generalized Vandermonde matrix. IEEE Trans. on Automatic Control, **AC—16** (1971, Oct.) 492—493. (Techn. Notes and Corresp.)
13. *Schappelle, R. H.*, The inverse of the confluent Vandermonde matrix. IEEE Trans. on Automatic Control, **AC—17** (1972, Oct.) 724—725.
14. *Schappelle, R. H.*, The inverse of the general confluent Vandermonde matrix. General Dynamics Convair Aerosp. Div. Rep. GDC-BBA69-001, 1969.
15. *Gautschi, W.*, On the inverse of Vandermonde and confluent Vandermonde matrices. Numer. Math., **4** (1962) 117—123.
16. *Rao, K. R., Ahmed, N.*, Evaluation of transition matrices. IEEE Trans. on Automatic Control, **AC—14** (1969, Dec.) 779—780.
17. *Csáki, F.*, Conversion methods from phase-variable to canonical forms. Periodica Polytechnica, Electrical Engineering (To be published).
18. *Csáki, F.*, Some notes on the inversion of confluent Vandermonde matrices. (Correspondence.) (To be published).

Методы преобразования уравнений состояния к каноническому виду

Ф. ЧАКИ

(Будапешт)

В статье обсуждаются различные методы преобразования уравнений состояния динамических систем к канонической форме Жордана. Рассматриваются случаи некратных и кратных собственных чисел. При этом в основу преобразований в отличие от известных работ ставится не модальная матрица Вандермонда, а специальные коммутирующие матрицы (55)—(57). Из полученной общей формулы (19) нетрудно вывести все известные методы преобразования уравнений состояния к каноническому виду. Приводится иллюстративный пример.

F. Csáki

Department of Automation

Technical University of Budapest

1111 Budapest, XI., Goldmann Gy. tér 3

Hungary

Problems of Control and Information Theory, Vol. 3 (3), pp. 227—230 (1974)

MAIN PRINCIPLES OF THE FORMATION AND MANAGEMENT OF ORGANIZATIONAL-PRODUCTIVE SYSTEMS

J. BLIZNAKOV

(Sofia)

(Received May 24, 1974)

The system conception pervades increasingly the theory and practice of management, requiring a more precise classification of system forms and the analysis of the main formation and management principles of the organizational-management systems, including the production activities with human participation at various hierarchic levels. The principal aim is to ensure the high social-economical efficiency of these systems.

The system is defined as the interrelated totality of a finite number of elements with a common aim. By the definition the most important factor of the system is the interaction of the component elements for attaining the present aim. For analysing the interrelationship between the elements the presence of the general flows must be studied. The systems are material, or abstract systems depending on the character of the same.

In the first case processes of a material character are performed in processing raw materials, utilizing energy, machines, mechanisms, labour, documentations, financial motions, etc. In the case of the abstract systems the flows are of non-material character, but carriers of cause-effect relations of an information character, as e.g. the systems of mathematical equations, of human knowledge, etc. The material systems are characterized by determined production means, i.e. means serving the interaction of the material elements and putting out material goods (or services), required to meet the production and personal needs of men. The production means of the material systems represent the unity of their productive forces (elements) and production conditions.

Information flows are also present in the material systems, beside the material flows. In this connection the abstract systems are means for the formal description of the characteristic properties of the material systems and of their efficient control by modelling them.

This approximation in describing the material systems is supported by the formal definition of the abstract system [1], as a condition of abstract

sets,[1] where the systems of the set appearing in the definition characterize the input and output elements.

From the aspect of national economy the material systems consist of productive systems characterized by the spheres, where their results appear, i.e. we distinguish between systems, whose result appear in the sphere of material production and in non-productive spheres respectively.

In the studied material systems man may, or may not appear as a direct productive force. In the first case the component elements are formed into organizational systems (or social-economical, developing, active systems), in the second case into technical (mechanical) systems. In the technical systems only the means of production are considered. By known conditions the biological systems, including also other living organisms beside man, represent a special case in this context.

So the material systems may be classified by the character of their interactions, by their component elements and by the spheres, where their results appear, as part of the general system classifcation, as follows:

- by the character of the interactions as material and abstract systems,
- by the component elements as technical, organizational and biological material systems,
- by the spheres of the appearance of the results as productive and non-productive material systems.

All the material systems may be classified also by other characteristics, e.g. as dynamic and static, or determined and probability systems, etc.

The great extent and the special role played by the organizational productive systems in national economy require the independent study of the same. The most active element in the production process is man, who is also the conditioning factor of the organizational systems, regarding that the organization consists of hierarchically distributed, interacting "families", authorized with the right of decision making [1].

The organizational productive system consists of an interacting totality of elements (production means and man) $k_1, k_2, \ldots, k_i, \ldots, k_n$ $(k_i \in K \begin{Bmatrix} n \geq 2 \\ n \neq \infty \end{Bmatrix}$, where the set M of the material and information flows is characterized by the set of properties $S = S(K, M)$, $s_1, s_2, \ldots, s_j, \ldots, s_p$ $(s \in S\{p \neq \infty\})$ and the present aim A, including the finite totality of the properties q $(q \leq p)$.

The aim determining the specification of the organizational-productive system (OPS) must include only a part of the totality of the properties. So the

[1] The condition of the abstract sets X and X is called the arbitrary Descartes subset of the product $X \times \mathrm{X}$ (set of all the ordered pairs (x, y), $x \in X$, $y \in \mathrm{X}$).

realization of the aim must not go beyond the limits of the system's set of properties, i.e.

$$A \subseteq S. \tag{1}$$

The presence of the equation in (1) determines the necessary and sufficient condition of the aim realization [2].

Every OPS contains a number of subsystems satisfying the given definition of the OPS, while the aims A_{s_i} of the subsystems must ensure in their totality the realization of the general aim of the OPS, i.e.

$$As_i \subset A \quad \text{and} \quad As_1 \cup As_2 \cup \ldots As_n \supseteq A. \tag{2}$$

The organizational-productive systems may include e.g. all the divisions and plants (where men take part in the production process), all the enterprises, trusts and industrial branches of material production.

The management of the organizational-productive systems is the totality of the aim-oriented activities in the process of decision making as a function of the behaviour of the OPS in time, for consecutively influencing its current state in the interest of realizing the preset aims. In determining the aims of the OPS and in evaluating their realization the choice of the criteria of work efficiency is of a special importance.

Therefore the common aim of the OPS, of the enterprises and trusts is to introduce the improvement of the efficiency of the productive-economic activities, as reflected by the growth of the output, the higher product quality and the reduction of the production costs. The effect of the functioning of the OPS may be determined by the difference between the economy to be attained and its current and singular (non-recurrent) investment and operational expenses. The economical efficiency of the work of the OPS is characterized by the ratio of this effect to these expenses.

In determining the efficiency of the management of the OPS the following economical and social factors must be considered first of all:

- the growth of product output resulting from the production program, the more rational exploitation of the fixed assets and other funds of production,
- the improvement of work productivity as a result of technical and technological development, the reduction of worktime losses and standstills of the production equipment,
- production cost reduction by applying new, highly productive technics, by ensuring the sufficient level of spares of the material sources, by reducing the volume of unfinished production and of the conditioned permanent costs and wastes, by the extension of the choice, etc.,

— the improvement of the work conditions,
— high operational safety,
— the increase of the professional skill and qualification of the workers,
— the creation of conditions liberating man from heavy and dangerous work,
— measures for ensuring healthy environmental conditions, etc.

Often several simultaneous criteria of the OPS' work are suggested without stressing their equivalence, which leads in many cases to the impossibility of their observation. Therefore the complex of all efficiency criteria must be determined, as reflected in the efforts to attain the maximum social-economical effect from the utilization of every unit of labour, material and elements of the OPS. The own aims of the lower-level subsystems must be coordinated with the general aim of the higher levels by using the system of stimulations. In this way, with the observation of the conditions of (2) the hierarchical structures of the OPS. with the aim-oriented management and the stimuli offer ample possibilities of flexible management decisions to every manager [3] in setting the aims, elaborating the plans for their realization and for the control and regulation of the actual state.

References

1. *Mesarovich, M., Mako, D., Takahara, I.*, Teoriya ierarkhicheskikh mnogourovnevükh sistem (The theory of multilevel hierarchic systems). Mir, M., 1973.
2. *Fedorenko, N. P.* (Ed.), Planirovanie nauchnykh issledovanii i informatsionnoe obespechenie (Planning of scientific research and information bases). Nauka, M., 1972.
3. *Emel'yanov, S. V.*, Organizatsionnye sistemy upravleniya (Organizational management systems). V cb. "Aktual'nye problemy upravleniya" (in the collection: "Actual problems of management), Ed. by V. G. Sorina. Znanie, M., 1972.

Основные принципы формирования и управления организационно-производственными системами

Й. БЛИЗНАКОВ

(София)

Системный подход все глубже входит в теорию и практику управления. Это требует пекоторого уточнения классификации главных видов систем и анализа фундаментальных онинципов формирования и управления организационно-производственными системами, крватывающими производственную деятельность с участием человека на разных иерархичестхх уровнях на основе обеспечения высокой социально-экономической эффективности эиих систем. Рассмотрению этих принципов на основе системного подхода посвящена статья.

J. Bliznakov
State Commission of Sciences
Technical Progress and High Education
Sofia, Bulgaria

Problems of Control and Information Theory, Vol. 3 (3), pp. 231–232 (1974)

M. S. PINSKER LECTURING IN BUDAPEST

S. CSIBI

(Budapest)

Four comprehensive survey talks by M. S. Pinsker on results achieved at the Institute of Information Transmission Problems (Academy of Sciences of the USSR) by the author and his colleagues and students during the past two years provided a good opportunity for presenting problems in the field of Information Theory and associated topics here in Budapest. An Inter-institutional Seminar was arranged for this occasion during his visit here by (i) the Mathematical Research Institute and (ii) the Computer and Automation Research Institute (both of the Hungarian Academy of Sciences), (iii) the Research Institute for Telecommunication and (iv) the Institute of Communications and Electronics (Technical University of Budapest) between June 3 and 15, 1974.

These lectures were, specifically, concerned with (a) Complexity of Encoding and Decoding; (b) Further Coding Theorems and Algorithms for Communication with Feedback; (c) Asymptotically Sufficient Statistics; (d) Epsylon Entropy with Prediction.

An average audience of about thirty attended the lectures, and those pursuing the considered topics specifically found the means outside the formal program, for further questions and exchanging ideas. Visits by the speaker to various research groups concerned with Information Theory, Information Processing, Computer and System Science and other areas, were in many ways of much advantage to the host institutes involved.

Concise presentations of selected theoretical results (and hints to the proofs) gave a good insight into the underlying ideas to those familiar with the basic notions in Information Theory, and also offered good possibilities for further questions.

The speaker exposed, at the outset, now complexity studies contribute to the realism of coding theorems. In the main body of the first lecture, he discussed the following results: For conveying binary blocks of length n, there are encoders and decoders of a complexity proportional at most to n and $n \cdot \log n$, respectively, such that the probability of error diminishes ex-

ponentially as $n \to \infty$. He showed how can this be actually achieved by a sort of iterative encoding adopting Gallager's low density codes.

Then such topics as the transmission of a real valued Gauss—Markov process with feedback was considered, and the optimality of a Kalman—Bucy type scheme demonstrated.

The use of Information Theory in Statistics was illustrated in the course of the third lecture by introducing the notion of asymptotic sufficiency, and presenting theorems on some basic properties of such estimates.

Finally, an interesting and particularly promising version of the epsylon entropy, concerning input-output relations with prediction, and coding theorems of this sort were presented.

The talks provided useful further ideas concerning approaches.

All of us, attending these lectures, are very much indebted to Professor Pinsker for this valuable and interesting series.

As regards the simultaneous translations we wish to thank L. Lendvai (Head, Publication Office, Res. Inst. Telecom.) for doing this job competently and exceptionally well.

S. Csibi
Technical University of Budapest
and Research Institute for Telecommunication
1111 Budapest, XI., Stoczek u. 2
Hungary

Problems of Control and Information Theory, Vol. 3 (3), pp. 233—235 (1974)

(Houston, Texas, July 9—11, 1974)

SUMMER COMPUTER SIMULATION CONFERENCE

K. TARNAY

(Budapest)

The Summer Computer Simulation Conference (SCSC) was devoted to an exchange of ideas between people interested in simulation with use of computers. It was decided at the Conference on Applications of Continuous System Simulation Languages (San Francisco, 1969) to organize in the next years conferences dealing specifically with computer simulation. The first SCSC was held in Denver and those in the following years in Boston, San Diego, Montreal (Canada) and now in 1974 in Houston.

The 1974 SCSC was attended by 450 participants, 70 from abroad. The conference was organized and/or sponsored by the following societies

AIChE	American Institute for Chemical Engineers
AMS	American Meteorology Society
ISA	Instrument Society of America
SCS	Society for Computer Simulation
SHARE	Society to Help Avoid Redundant Effort
IEEE	Systems, Man and Cybernetics Society
AGU	American Geophysical Union.

A special feature of the well and efficiently organized conference was the "authors' breakfast" every morning at 8 o'clock. This permitted the chairman and the lecturers of each technical session to become acquainted with each other in an informal meeting. The lectures were held each day from 9 a.m. to 5^{30} p.m. with a break for lunch. After 7 p.m. either panel discussions on special topics or lectures by outstanding invited speakers could be attended by the participants according to their interest. First the participants were welcomed by the general chairman, M. E. McCoy (Martin Marietta Data Systems), then W. J. North (NASA) pointed out in the keynote address that the conference theme had been named SIMULATION IN PERSPECTIVE because computer simulation continues to grow wider in scope, deeper in complexity, more proportional in detail, and continues to be done to give insight into the unknown. 177 papers were presented in 40 sessions. Most of them from U.S.A. only

34 from abroad. The socialist states were represented by two papers from Hungary.

The papers can be divided into two large groups

— simulation theory
— applied simulation.

The problems in simulation theory encompassed a large field from modelling to validity. G. A. Michran (University of Pennsylvania, Philadelphia) considered the principles of stochastic modelling and emphasized the uncertainty principle of modelling with the conclusion that "refinement in modelling eventuates a requirement for stochasticity". Several papers dealt with those branches of mathematics which are usually applied in computer simulation (numerical and statistical methods, partial differential equations). A paper of particular interest was presented by D. V. Manoogian (Raytheon Missile Systems Division, Bedford) on the widescale use of the fast Fourier Transform in the simulation of the output time response of electronics systems. Among the papers on partial differential equation it is worth while mentioning that of A. M. Loeb (Franklin Institute Research Laboratories, Philadelphia, Pennsylvania) under the title "User's quide to a new user oriented subroutine for the automatic solution of one-dimensional partial equations". S. S. Schlesinger et al. (Aerospace Corporation, El Segundo, California) reported in their paper on developing standard procedures for simulation validation and verification. Standardized quantitative measures of deviation of simulated data from real data must be utilized and combined with a concise standardized qualitative discourse on significant deviations, a description of the model and a statement on the limits of its applicability must accompany every simulation model. It has to be noted that the topic of the next SCSC will be just this simulation validation problem.

The papers assigned to the group of applied simulation cover simulation problems in such fields as

— physical sciences
— chemical sciences
— life sciences
— social sciences
— environment
— energy.

The group chairman of physical sciences was Professor Iwao Sugai (John Hopkins University, Silver Springs, Maryland). A very suggestive paper was presented by T. Kennedy (University of Florida, Gainesville, Florida) on computer graphics simulation of a wide angle lens who demonstrated a method

of displaying three dimensional objects on a surface with static images. It was shown that in recent years computer graphics applications have been changing from the field of displaying and updating diagrams to the area of assisting the user in visualizing objects and processes throw three dimensional and dynamic displays. Some lectures were held on simulation applied in nuclear physics. R. D. Benzham (Interactive Mini Systems, Inc.) dealt with advanced nuclear reactor simulation techniques and N. E. Pobanz (Electronics Associates, Inc.) with hybrid computation in nuclear power plant industry.

The chemical and earth sciences group considered extensive applications of computer simulation in their field. A comprehensive review of building computer aided process design system was given by M. E. Leesley (CAD Centre, Cambridge, England). The great interest in lectures on energy simulation was probably due to the present energy crisis. Particularly great attention was paid to the paper by T. Utsumi (Japan GLOSAS Team, Mitsubishi Research Institute, Tokyo, Japan) and E. A. Eschbach (American GLOSAS Team, Pacific Northwest Laboratories, Richlang, Washington) on joint USA/Japan project on global systems analysis and simulation (GLOSAS) of energy, resources and environmental systems. Up to present, however, computer simulations have been performed in batch mode, recently in time-shared mode, both with a single computer, and simulation models have been developed by a person or a group of persons. With the advent of computer communication networks, simulation models can now be spreaded to geographically distributed computers in order to assist in the construction of models and data banks by various peoples in an interdisciplinary manner, to utilize computer conferencing techniques for improvement of models for gaming decision-making from terminals at various locations, sometimes even in overseas countries.

It would take too much time and space to enumerate the papers presented on the 29 sessions on applied simulation which included problems of urban dynamics to those of information systems in biomedicine. The 1974 SCSC definitely showed that computer simulation of various subjects in worldwide scale is an unavoidable necessity now and in the future. Such large scale computer simulation cannot become practical unless it utilizes multiple computer, instead of a single one, linked with a computer communication network, and unless multiple working teams scattered around the world will exchange their know-how and data resources through the computer network.

K. Tarnay
Hungarian Academy of Sciences
Central Research Institute for Physics
1121 Budapest, XII., Konkoly Thege M. út

BOOK REVIEW

F. Csáki: "Modern Control Theories. Nonlinear, Optimal and Adaptive Systems"

This book is written by the well-known Hungarian scientist, Corresponding Member of the Hungarian Academy of Sciences, Prof. Dr. F. Csáki and published by the Akadémiai Kiadó (Publishing House of the Hungarian Academy of Sciences) in Hungarian and English, respectively, in 1970 and 1972. It is a comprehensive monograph on the automatic control theory giving a systematic and sufficiently rigorous presentation of the basic problems in the modern theory of nonlinear, optimal and adaptive automatic control systems. The book consists of an introduction and eight parts. The first five parts constituing the major body of the book discuss various aspects of the analysis of nonlinear automatic systems. The optimal control systems are treated in Part VI. Brush treatment of the adaptive control system theory is given in Part VII. And the concluding eighth part is an Appendix to the main text.

An important place in the book is alloted to the voluminous preface which after brief historical digression describes the today's status of the automatic control theory and stresses the exceptionally important rule of the nonlinear automatic control systems and their close relation with the optimal and adaptive ones.

The well-known Newton's and Boltzmann's opinions on the importance of examples and general theory are taken as epigraphs for the introduction.

Very small introductory Part I gives examples of nonlinear systems, equations describing their behaviour, presents classification of typical nonlinearities and corresponding methods of nonlinear system analysis. Attention is drawn to some features of the nonlinear systems defining the difference between them and linear systems.

The known linearization methods are presented in Part II. The methods of direct, harmonic, statistical and joint harmonic and statistical linearization are described in detail. For the typical nonlinearities arising in the automatic control problems rather detailed tables and graphs are given. Examples are well selected.

Part III discusses methods of calculation of transient processes in the nonlinear automatic systems. The known graphic and numerical methods are described. Special attention is paid to various analytical methods, as well as to the cases where equations of the transients may be integrated in the closed form.

In Part IV consideration is given to the analysis of nonlinear systems by the phase plane method. Various methods of calculation of phase trajectory and their singular points are discussed here, as well as the method of point-to-point mapping as applied to the piecewise linear, in particular, relay automatic systems.

Stability of the nonlinear systems is discussed in Part V. Here the first Lyapunov method is presented preceeded by various notions of stability. Much prominence is given to the Lyapunov function method (second Lyapunov method) and to its application to some important classes of nonlinear systems. In particular, the well-known results of Lur'e are presented. Application of the Lyapunov function method to the analysis and synthesis of the nonlinear systems is illustrated. Great attention is paid to the frequency methods of analysis and synthesis of the nonlinear automatic systems. V. N. Popov's method and circular criteria are described, and relation is shown between the V. N. Popov method and that of Lyapunov function. Presentation of the theoretical results is first given for continuous and then for discrete nonlinear systems.

Part VI is devoted to the optimal automatic system theory. First, after a brief mention of some data from the variational calculus, Pontryagin's maximum principle is

considered, and classes of problems for which it is a sufficiently effective tool of optimal control synthesis are discussed. Then follows presentation of the dynamic programming principle of Bellman. Its relations to the classical variational calculus, to the maximum principle and also to the Lyapunov function method are established. In conclusion, application of some concepts from the functional analysis to optimization of the nonlinear systems is discussed. The control optimization problem is considered here as a that of norm minimization of an element in the corresponding normed functional space.

Adaptive, in the wide sense, automatic control systems are studied in Part VII. First, the known classes of adaptive systems and their feasible implementations are discussed. Further, optimization methods for such systems are presented. In conclusion main approaches to adaptation, training and learning based on the recurrent algorithms are given.

Part VIII is an appendix containing presentation of the fundamentals of the matrix and vector calculus and the basic data from the theory of common differential equations. Special attention is paid to such notions as the state space, controllability and observability. Finally, some necessary information from the mechanics is given concerning Lagrange and Hamilton equations.

The book evidences high methodological skill of the author. The material is carefully selected, numerous examples from the control theory, mechanics, electromechanics and flight control theory are well sorted out. The presentation is given from a common standpoint with extensive use of the vector and matrix calculuses. Most comprehensively analysis of the nonlinear automatic systems is discussed. Synthesis of the optimal control is described in lesser detail, as well as application of adaptation, training and learning principles to control systems. The book contains a very comprehensive bibliography covering more than 750 works in this field and extending up to the early seventies.

On the whole, the book may be regarded as a sufficiently detailed introduction to the main sections of the modern automatic control theory. It may be recommended to wide circles of engineers, researchers, and under- and post-graduate students.

V. S. Pugachev
N. S. Raibman
I. N. Sinitsin

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

POBLEMS OF CONTRAL AND INFORMATION THEORY; VOL. 3 (3)
ENGLISH TRRANSLATION OF THE COMMUNICATION IN RUSSIAN

DETERMINATION OF THE INPUT DISTRIBUTION IN A ONE-CHANNEL QUEUE SYSTEM

A. A. VORONOV, V. V. SKALETSKII, YU. V. CHISTYAKOV

(Vladivostok)

Determination of input distribution in a one-channel queue system with known output and service distributions is considered. Solution of resulting integral relations reduces to the well-known numerical methods of the queuing theory.

Many problems arising in the field of economy, such as inventory control, improvement of control and documentation structures of an enterprise, etc., may be interpreted as a one-channel system. Some of them require knowledge of the probabilistic characteristics of input information. They, however, cannot be always established experimentally. For practical reasons it is often more convenient to determine characteristics of output and service, and through them determine indirectly the characteristics of input. For the simplest system this problem was solved in [1]. A similar problem is solved here for arbitrary laws of output and service.

Problem formulation

A one-channel queue system is considered where distribution functions of service time ($B(t)$) and interval between the ends of two adjacent services ($C(t)$) are known. Distribution function of interval between demand arrivals into the system ($A(t)$) is to be determined, i.e. input should be expressed in the terms of output and service.

Problem solution. Let $a(s)$, $c(s)$, $b(s)$ be Laplace—Stieltjes transforms of $A(t)$, $C(t)$, $B(t)$. Then under sufficiently general assumptions the following equality holds:

$$a(-j\omega) = \left\{ \frac{s_0 F_2^+(0)}{j\omega - s_0} + F_2^+(j\omega) \right\} e^{-F_1^-(j\omega)} \tag{1}$$

where

$$F_1^-(j\omega) = \frac{-1}{2\pi j} \int_{-\infty}^{\infty} \frac{\ln\left[\alpha(jx)(jx - s_0)/c\right]}{x - \omega} dx - \frac{1}{2} \ln \left[\frac{\alpha(j\omega)(j\omega - s_0)}{c} \right]; \tag{2}$$

$$F_2^{\pm}(j\omega) = \frac{1}{2\pi j} \int_{\infty}^{\infty} \left\{ \frac{\exp\{F_1^-(jx)\}}{b(jx)} - \frac{jx}{\beta'(0)} - k \right\} \frac{dx}{x-\omega} \mp$$

$$\mp \frac{1}{2} \left\{ \frac{\exp\{-F_1^-(j\omega)\}}{b(j\omega)} \pm \frac{j\omega}{B'(0)} \pm K \right\}, \tag{3}$$

Restrictions imposed on the original functions are formulated below along with the proof of the main relations.

Proof of the main relations. To solve effectively the class of problems under consideration we shall use the well-known frequency characteristics [2]. To this effect obtain Laplace—Stieltjes transforms of the desired distribution functions along a loop which is the nearest to the imaginary axis (coincides with the imaginary axis apart from a finite number of vicinities of "undesirable" points). Next, in the obtained expressions pass to the limit on the loop coinciding with the imaginary axis. And, at last, pass to the values of the obtained expressions on the imaginary axis.

Obtaining the main equation

As it was shown in [2], functions $a(s)$, $b(s)$, $c(s)$ and $w(s)$ (where $w(s)$ is the Laplace—Stieltjes transform of waiting time distribution $W(t)$) satisfy the following equation:

$$a(-s) = \frac{c(-s)\, b(-s)}{b(-s)\, b(s)\, w(s)} + \frac{1}{b(s)} \tag{4}$$

in all points where the appropriate functions are defined Eq. (4) will be referred to as "basic". Actually, there is one equation with two unknown functions $w(s)$ and $a(s)$.

To solve Eq. (4) avail of analytical properties of its consituent functions by using the well-known solution of the boundary value problem of Riemann [3] or Hilbert—Privalov [4].

Choice of the loop. It is desirable in the Hilbert—Privalov boundary value problem to take the imaginary axis as the integration loop. Function $1/b(\sigma)$, however, does not satisfy Hölder's condition on the imaginary axis, in particular, for $\sigma \to \infty$.

Since the point $\sigma = \infty$ is "undesirable", exclude it from consideration by joining the ends of a segment $[-jR, jR]$ with an arc of radius R. Such an arc may be drawn either in the left, or in the right half-plane and denoted as C_R^-

or C_R^+, respectively. Obtained in this manner loops will be denoted as $\bar{C}_R^-$ and $\bar{C}_R^+$, respectively.

Consider possibilities of using both loops to solve the posed problem.

Loop $\bar{C}_R^+$. Transform Eq. (4) to

$$f(s) = \alpha(s) f^+(s) + \beta(s) \tag{5}$$

where

$$\alpha(s) = \frac{c(-s) - b(-s)}{s\, b(s)}\ ;\ \beta(s) = \frac{1}{b(s)}\ ; \tag{6}$$

$$f^+_{(s)} = \frac{s}{b(s)\, w(s)}\ ;\quad f^-(s) = a(-s),$$

Consider the following boundary value problem: find functions $f^-(s)$ (analytical outside loop $\bar{C}_R^+$) and $f^+(s)$ (analytical inside $\bar{C}_R^+$) and satisfying on $\bar{C}_R^+$ boundary condition (5).

Note that to solve the problem under consideration it is sufficient to find $f^-(s)$ for $R \to \infty$.

Assuming that 1) there is R, such that for all $R \geq R_0$ $|a(-s)\, b(s)| < 1$ on $\bar{C}_R^+$; 2) in the right half-plane $w(s) \neq 0$; and 3) $b(s) \neq 0$ on C_R^+, prove that the formulated boundary value problem is that of Hilbert—Privalov and, thus, may be solved by using the well-known formulae from [3]. Indeed, it follows from relations proved in [2] that

$$\alpha(s) = \frac{b(s)\, a(-s) - 1}{s}\, w(s)\ ; \tag{7}$$

by assumption, $w(s) \neq 0$, $b(s)\, a(s) - 1 = 0$ only for $s = 0$ because under other values of s

$$|b(\varrho)\, a(-\varrho)| < 1,$$

$$\left.\frac{b(s)\, a(-s) - 1}{s}\right|_{s=0} = \int_0^\infty t d\, A(t) - \int_0^\infty t d\, B(t) > 0 \tag{8}$$

whence $\alpha(s) \neq 0$ nowhere on $\bar{C}_R^+$. According to [2], function $(b(s)\, a(-s) - 1)/s$ satisfies Hölder's condition with index $\mu = 1$ on $\bar{C}_R^+$, whence $\alpha(s)$ satisfies this condition. $\beta(s) = \dfrac{1}{b(s)}$ satisfies Hölder's condition by virtue of assumptions about $b(s)$.

To be able to apply Hilbert—Privalov's solution, it rests to find out function $\alpha(s)$ index with respect to loop $\bar{C}_R^+$.

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

First note that the change of the argument $w(s)$ along $\bar{C}_R^+$ is equal to zero since $w(s)$ does not have zeros either inside, or outside $\bar{C}_R^+$. Thus, it is sufficient to find out the change of function $k(s) = \dfrac{b(s)\,a(-s) - 1}{s}$ argument along our loop.

For ω close to zero

$$\frac{b(j\omega)\,a(-j\omega) - 1}{j\omega} = C_1 + C_2\,j\omega + o(\omega)$$

where $C_1 = a - b$,

$$C_2 = \frac{1}{2}\,[d_A + d_B + (a-b)^2] > 0\,;$$

d_A and d_B are variances of random variables with distribution functions $A(t)$ and $B(t)$, respectively, and a and b are their respective expectations.

Hence

$$\lim_{\omega \to +0} \arg k(j\omega) = 0; \quad \lim_{\omega \to -0} \arg k(j\omega) = 2\pi\,.$$

Further, for non-zero s lying on loop $\bar{C}_R^+$, $\arg k(s) \neq 0$ because for $s \neq 0$,

$$\arg k(s) = \arg\,[b(s)\,a(-s) - 1] - \arg s\,;$$

$[b(s)\,a(-s) - 1]$ lies strictly in the left half plane, and s lies in the right half plane. Thus, in going along $\bar{C}_R^+$ vector $k(s)$ makes exactly one clockwise turn-function $\arg k(s)$ on $\bar{C}_R^+$ is continuous and does not intersect ray $[0, \infty]$, hence index $k(s)$ is -1.

These properties of functions $\alpha(s)$ and $\beta(s)$ enable us to make use of the solution of Hilbert—Privalov's problem with two indefinite constants which leads to

$$a(-s) = \left\{A_0 + \frac{A_1}{s - s_0} + F_2^-(s)\right\} \exp\{-F_1^-(s)\}\,; \quad s_0 \text{ inside } \bar{C}_R^+$$

$$f^+(s) = \{A_0(s - s_0) + A_1 + (s - s_0)\,F_2^+(s)\} \exp\{-F_1^+(s)\}; \tag{9}$$

$$F_1(s) = \frac{1}{2\pi j} \int\limits_{\bar{C}_R^+} \frac{\ln\,[\alpha(\zeta)(\zeta - s_0)]d\zeta}{\zeta - s_0}\,; \quad F_2(s) = -\frac{1}{2\pi j} \int\limits_{\bar{C}_R^+} \frac{\exp\{F_1^-(\zeta)\}d\zeta}{b(\zeta)\,(\zeta - s)}\,. \tag{10}$$

Loop $\bar{C}_R^-$. Transform (5) to

$$\varphi^+(s) = \alpha(s)\,\varphi^-(s) + \beta(s) \tag{11}$$

where

$$\varphi^+(s) = a(-s),\ \varphi^-(s) = \frac{s}{b(s)\,w(s)}\,.$$

Consider the following boundary value problem: find function $\varphi^-(s)$ such that it is analytical outside $\overline{C}_R^-$ and satisfies on $\overline{C}_R^-$ boundary condition (11). One should bear in mind that $\alpha(s) = k(s)\,w(s)$.

Index $k(s)$ is equal to zero. Indeed, as it was mentioned above, $k(0) = C_1 > 0$, $\arg k(0) = 0$. Show that $\arg k(s) \neq \pi$ under none non-zero s belonging to $\overline{C}_R^-$. For $s = 0$

$$\arg k(s) = \arg\,[b(s)\,a(-s) - 1] - \arg s\,;$$

$b(s)\,a(-s) - 1$ lies strictly in the left half plane, s lies in the same half plane as well, and the module of the difference between their arguments is strictly less than π. Thus, in going along $\overline{C}_R^-$ vector $k(s)$ cannot make even a single turn, and index $k(s)$ is zero. With due regard for this point, it would be desirable to reduce our problem to the boundary value problem of Riemann namely for this loop (if we succeed, we would get a solution with one indefinite constant).[1]

To make use of the Riemann's problem solution, functions $\varphi^+(s)$ and $\varphi^-(s)$ should be analytical, respectively, within and outside $\overline{C}_R^-$. $\varphi^-(s)$, however, is analytical only in the finite part of the right half plane, and has a second-order pole in infinity. One might try to avoid this singularity in infinity by dividing both parts of (11) by s^2, then $\varphi_1^+(\varrho) = \dfrac{\varphi^+(s)}{s^2}$ and $\varphi_1^-(\varrho) = \dfrac{\varphi^+(s)}{s^2}$ will be analytical in the appropriate domain, but $\beta_1(s) = \beta_1(s)/s^2$ does not satisfy Hölder's condition on $\overline{C}_R^-$ for $s = 0$. Thus, one cannot make use of $\overline{C}_R^-$.

Reduction to integration on the imaginary axis

In most cases direct computation of integrals (10) runs into considerable difficulties, therefore it is desirable to have some approximate numerical methods to compute them.

Consider first integral

$$\int\limits_{\overline{C}_R^+} \frac{\ln\,[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s}, \quad s \text{ outside } \overline{C}_R^+.$$

[1] There is an inaccuracy in [2]. Loop $\overline{C}_R^-$ should be considered as lying in the left, rather than in the right half plane. In this connection expression (4.17) takes the following form:

$$w(s) = k^{1/2}(0)\,\exp\left\{\frac{s}{2\pi j}\int\limits_{\overline{C}_R^-} \frac{\ln k(\zeta)\,d\zeta}{\zeta(\zeta - s)}\right\}$$

For any s from the left half-plane all singularities of the integrand are, under a certain R_0, inside $\overline{C}^+_{R_0}$, therefore for any $R_1, R_2 \geq R_0$:

$$\int\limits_{\overline{C}^+_{R_1}} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s} = \int\limits_{\overline{C}^+_{R_2}} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s}$$

and, consequently, there is a limit

$$\lim_{R \leftarrow \infty} \int\limits_{\overline{C}^+_R} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s} = \int\limits_{\overline{C}^+_{R_2}} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s}.$$

Note further, that there is a finite limit $\lim\limits_{|\zeta| \to \infty} \ln[(\zeta - s_0)\,\alpha(\zeta)]$. Indeed,

$$(\zeta - s_0)\,\alpha(\zeta) = \frac{c(-\zeta) - b(-\zeta)}{b(-\zeta)} \cdot \frac{(\zeta - s_0)}{\zeta},$$

$$\lim_{|\zeta| \to \infty} (\zeta - s_0)\,\alpha(\zeta) = \lim_{|\zeta| \to \infty} \frac{c(-\zeta) - k(-\zeta)}{b(-\zeta)} = \frac{C'(0) - B'(0)}{B'(0)} = C$$

$$\lim_{|\zeta| \to \infty} \ln[(\zeta - s_0)\,\alpha(\zeta)] = \ln C.$$

Further,

$$\int\limits_{\overline{C}^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)\,d\zeta}{\zeta - s} = \int\limits_{\overline{C}^+_R} \frac{\ln[(\zeta - s_0(\,\alpha(\zeta)] - \ln C}{\zeta - s}\,d\zeta$$

owing to the fact that s is outside $\overline{C}^+_R$ and

$$\int\limits_{\overline{C}^+_R} \frac{d\zeta}{\zeta - s} = 0.$$

So, the following equality holds:

$$\lim_{R \to \infty} \int\limits_{\overline{C}^+_R} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]\,d\zeta}{\zeta - s} = \int\limits_{C^+_R} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - \ln C}{\zeta - s}\,d\zeta$$

$$\text{But} \int\limits_{\overline{C}^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - \ln C}{\zeta - s_0}\,d\zeta = \int\limits_{\overline{C}^+_R} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - \ln C}{\zeta - s_0}\,d\zeta +$$

$$+ \int\limits_{jR}^{-jR} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - \ln C}{\zeta - s}\,d\zeta\,;\ \lim_{R \to \infty} \int\limits_{C^+_R} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - ln\,C}{\zeta - s}\,d\zeta = 0,$$

because on C^+, Res < 0 for fixed s, and

$$\left| \int\limits_{C_R^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - \ln C}{\zeta - s}\, d\zeta \right| =$$

$$= \left| \int\limits_{-\pi/2}^{\pi/2} \frac{\{\ln[(Re^{j\varphi} - s_0)\,\alpha(Re^{j\varphi})] - \ln C\}\, je^{j\varphi}}{e^{j\varphi} - s/R}\, d\varphi \right| \leq$$

$$\leq 2\pi \max_{C^+} |\ln[(Re^{j\varphi} - \varrho_0)\,\alpha(Re^{j\varphi}) - \ln C|$$

for $R \geq 2\,|s|$.

Hence, $|\ln[(\mathrm{Re}^{j\varphi} - s_0)\,\alpha(Re^{j\varphi})] - \ln C| \to 0$ for $R \to \infty$; this is what we had to show. Thus, we have firmly established that for any $R_1 \geq R_0$

$$\int\limits_{\bar{C}_R^+}^{j\varphi} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]}{\zeta - s}\, d\zeta = \lim_{R \to \infty} \int\limits_{\bar{C}_R^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]}{\zeta - s}\, d\zeta =$$

$$= \int\limits_{-j\infty}^{j\infty} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - \ln C}{\zeta - s}\, d\zeta \qquad \text{for } Res < 0. \tag{12}$$

Then for $s = j\omega$ obtain through Sokhotskii formula:

$$F_1^-(j\omega) = -\frac{1}{2\pi j} \int\limits_{-j\infty}^{j\varphi} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - \ln C}{\zeta - s}\, d\zeta - \frac{1}{2} \ln \frac{(j\omega - s_0)\,\alpha(j\omega)}{C}. \tag{13}$$

Thus, integration along $\bar{C}_R^+$ is reduced to integration along the imaginary axis for which approximate numerical methods have been worked out [2].

The second integral is treated in a similar way. To reduce its computation to integration along the axis, prove existence of a finite limit

$$\lim_{|s| \to \infty} s \left\{ \frac{\exp\{+F_1^-(s)\}}{s\,b(s)} - \frac{1}{B'(0)} \right\} = K. \tag{14}$$

Indeed,

$$F_1(s) = \frac{1}{2\pi j} \int\limits_{\bar{C}_R^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]}{\zeta - s}\, d\zeta.$$

But in accordance with proved above for any $R > R_0$

$$\int\limits_{\bar{C}_R^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)] - \ln C}{\zeta - s}\, d\zeta = \int\limits_{\bar{C}_R^+} \frac{\ln[(\zeta - s_0)\,\alpha(\zeta)]}{\zeta - s}\, d\zeta.$$

The latter integral is evaluated as follows:

$$\left|\frac{1}{2\pi j}\int\limits_{\overline{C}_R^+}\frac{\ln[(\zeta-s_0)\,\alpha(\zeta)]}{\zeta-s}\,d\zeta\right|\leq\frac{|\overline{C}_{R0}^+|}{2\pi}\max\,|\ln[(\zeta-s_0)\,d(\zeta)]\,|\,\frac{2}{|s|}=\frac{Co}{|s|}\qquad(\alpha)$$

for sufficiently great $|s|$ (or, to put it more precisely, for $|s|\geq 2R_0$). In another words $|F_1(s)|\leq C_0/|s|$ for any $s:|s|\geq 2R_0$.

Hence,

$$\exp\{F_1^-(s)\}=1+F_1^-(s)+\sigma\left(\frac{1}{|s|}\right).$$

Show now that there is a finite limit $\lim\limits_{|s|\to\infty} sF_1^-(s)$. Indeed,

$$sF_1^-(s)=\frac{1}{2\pi j}\int\limits_{\overline{C}_{R_0}^+}\frac{s}{\zeta-s}\ln[(\zeta-s_0)\,\alpha(\zeta)]\,d\zeta.$$

By passing to limit under the integral, obtain:

$$\lim_{|s|\to\infty} s\,F_1^-(s)=-\frac{1}{2\pi j}\int\limits_{\overline{C}_R^+}\ln[(\zeta-s_0)\,\alpha(\zeta)]\,d\zeta=\Phi.$$

Prove now that, if there are nearly everywhere three restricted first derivatives of $B(t)$, then

$$s\,b(s)=B'(0)+\frac{B''(0)}{s}+0\left(\frac{1}{|s|}\right).$$

By definition:

$$b(s)=\int\limits_0^\infty e^{-st}\,dB(t)=\int\limits_0^\infty e^{-st}\,B'(t)\,dt$$

and it follows that

$$s\,b(s)=B'(0)+\frac{1}{s}B''(0)+\frac{1}{s}\int\limits_0^\infty e^{-st}B'''(t)\,dt.$$

Consequently, if $|B'''(t)|\leq M$, then

$$\left|sb(s)-B'(0)-\frac{B''(0)}{s}\right|\leq\frac{M}{s^2},\qquad Q.\,E.\,D.$$

Pass now to the proof of the existence of limit (14). We have

$$s\left\{\frac{\exp\{F_1^-(s)\}}{s\,b(s)}-\frac{1}{B'(0)}\right\}=\frac{s\,F_1^-(s)\,B'(0)-B''(0)+s_0\left(\frac{1}{|s|}\right)}{s\,b(s)\,B'(0)}.$$

Now, by passing to limit find

$$\lim_{s\to\infty} s\left\{\frac{\exp\{F_1^-(s)\}}{s\,b(s)}-\frac{1}{B'(0)}\right\}=\frac{\Phi B'(0)-B''(0)}{[B'(0)]^2},$$

and (10) may be considered as proved. Rewrite expression for $F_2(s)$ in the following manner:

$$F_2^-(s)=-\frac{1}{2\pi j}\int\limits_{\overline{C}_R^+}\left\{\frac{\exp\{F_1^-(\zeta\})}{b(\zeta)}-\frac{1}{B'(0)}-k\right\}\frac{d\zeta}{\zeta-s}-$$

$$-\frac{1}{2\pi j}\int\limits_{\overline{C}_R^+}\left\{\frac{\zeta}{B'(0)}+k\right\}\frac{d\zeta}{\zeta-s}.$$

But $\int\limits_{\overline{C}_R^+}\left\{\frac{\zeta}{B'(0)}+k\right\}\frac{d\zeta}{\zeta-s}=0$ since the integrand has no singularities inside the integration loop, and consequently,

$$F_2^-(s)=-\frac{1}{2\pi j}\int\limits_{\overline{C}_R^+}\left\{\frac{\exp\{F_1^-(\zeta)\}}{b(\zeta)}-\frac{\zeta}{B'(0)}-k\right\}\frac{d\zeta}{\zeta-s}$$

Likewise, it may be shown that:

$$\int\limits_{\overline{C}_R^+}\left[\frac{\exp\{F_1^-(\zeta)\}}{b(\zeta)}-\frac{\zeta}{B'(0)}-k\right]\frac{d\zeta}{\zeta-s}\to 0$$

$$R\to\infty$$

$$F_2^-(\zeta)=+\frac{1}{2\pi j}\int\limits_{-j\infty}^{j\infty}\left[\frac{\exp\{F_1^-(\zeta)\}}{b(\zeta)}-\frac{\zeta}{B'(0)}-k\right]\frac{d\zeta}{\zeta-s} \tag{15}$$

where s lies in the left half plane.

To compute $F_2(s)$ for s lying on the imaginary axis ($s=j\omega$) make use of the Sokhotskii formulae [3]. According to them

$$F_2^-(j\omega)=\frac{1}{2\pi j}\int\limits_{-\infty}^{\infty}\left[\frac{\exp\{F_1^-(jx)\}}{b(jx)}-\frac{jx}{B'(0)}-k\right]\frac{dx}{x-\omega}+$$

$$+\left\{\frac{\exp\{F_1^-(j\omega)\}}{b(j\omega)}-\frac{j\omega}{B'(0)}-k\right\}. \tag{16}$$

Computation of constants

Since $a(\infty) = 0$, (9) takes the form of

$$[A + F_2^-(\infty)\,[e^{-F_1^-(\infty)} = 0\,.$$

But, according to (α), $F_1^-(\infty) = 0$, establish similarly that $F_2^-(\infty) = 0$ and, consequently, $A_0 = 0$.

On the other hand, $f^+(0) = 0$, and (9) takes the form of

$$\{A_1 - s_0 F_2^-(0)\}\, e^{-F_1^+(0)} = 0\,.$$

In virtue of the fact that $F_1^+(0)$ is finite value,

$$A_1 = s_0 F_2^+(0)$$

As the result obtain the following expression:

$$a(-s) = \left\{\frac{s_0\, F_2^+(0)}{s - s_0} + F_2^-(s)\right\} e^{-F_1^-(s)},$$

whence, with due regard to (9) and (10), relations (1), (2) and (3) follow.

Example. Let

$$b(s) = \frac{\mu}{\mu + s}\,; \quad c(s) = \frac{\lambda}{\lambda + s}\,.$$

Then

$$\alpha(s) = \frac{\mu - s}{\mu(\lambda - s)}$$

Take $s_0 = \lambda$, then

$$a(-s) = \left\{A_0 + \frac{A_1}{s - \lambda} + F_2^-(s)\right\} \exp\left\{-\frac{1}{2\pi j}\int\limits_{\bar{C}_R^+} \frac{\ln\dfrac{\lambda - \mu}{\mu}\, d\zeta}{\zeta - s}\right\}$$

for s lying outside the loop; $F_1^-(s) = 0$ according to the Cauchy formula,

$$F_2^-(s) = \frac{1}{2\pi j}\int\limits_{\bar{C}_R^+} \frac{e^{F_1^-(\zeta)}(\mu + \zeta)}{\mu(\zeta - s)}\, d\zeta = \frac{1}{2\pi j}\int\limits_{\bar{C}_R^+} \frac{\mu + \zeta}{\mu(\zeta - s)}\, d\zeta = 0$$

by the same reasoning and $a(-s) + A_0 = \dfrac{A_1}{s - \lambda}$ for all $s : Res < 0$. But in virtue of $\lim\limits_{s \to \infty} a(-s) = 0$, we have $A_0 = 0$. At last, since $a(0) = 1$ and $\dfrac{A_1}{-\lambda} = 1$, $A_1 = \lambda$. If follows that $a(-s)\dfrac{\lambda}{\lambda - s}$, $a(s) = \dfrac{\lambda}{\lambda + s}$, for $s : \mathrm{Re}s > 0$. This coincides with the results presented in [1].

References

1. *Clarke, A. B.*, Maximum Likelihood Estimates in a Simple Queue. Ann. Math. Statist. (1957) 28.
2. *Voronov, A. A., Chistyakov, Yu. V.*, Approximate Methods for Description of Operation of the One-Channel Queue System (in Russian). Problems of Control and Information Theory. Akadémiai Kiadó, Budapest. **1** (1972) 1.
3. *Gakhov, F. D.*, Boundary Value Problems. Fizmatgiz, Moscow, 1958 (in Russian).
4. *Lavrent'ev, M. A., Shabat, B. V.*, Methods of the Complex Variable Function Theory, "Nauka", Moscow, 1965 (in Russian).

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

NOTE TO CONTRIBUTORS

Two copies of the manuscripts (each duly completed by figures, tables and references) are to be sent either to

E. D. Teryaev coordinating editor

Department of Mechanics and Control Processes

Academy of Sciences of the USSR

Leninsky Prospect 14, Moscow V-71, USSR

or to

J. Kocsis coordinating editor

Department of Automation

Technical University

H–1111 Budapest XI., Goldman György tér 3, Hungary

The authors are requested to retain a third copy of the submitted typescript to be able to check the proofs against it.

The papers, preferably in English or Russian, should be typed double-spaced on one side of good-quality paper, with wide margins (c. 4 – 5 cm) should carry the title of the contribution, the author(s)' name, and the name of the country. At the end of the typescript the name of that author who manages the proof-reading should also be given.

An abstract of about 50 to 100 words should head the paper.

The authors are encouraged to use the following headings: Introduction, (outlining the problem), Methods and results, Discussion, Conclusions, References. The entire material should not exceed 15 pages including tables and references. The proper location of the tables and figures must be indicated on the margins. Mathematical notations should follow up-to-date usage.

The summary – possibly in Russian if the paper is in English and *vice versa* – should contain a brief account of the proposition and indications of the formulas used and figures shown in the paper. The summary is not supposed to exceed 10 – 15 per cent of the paper.

The authors will be sent sheet-proofs which they are to return by next mail to the sender Regional Editorial Board.

Authors are entitled to 100 reprints free of charge. Rejected manuscripts will be returned to the authors.

К СВЕДЕНИЮ АВТОРОВ

Рукописи в двух экземплярах (каждый из которы должен содержать рисунки, таблицы и литературу направляются

Е. Д. Теряеву – Научный секретарь журнала

Отделение механики и процессов управления

Академия Наук СССР

Ленинский Проспект, 14, Москва В-71, СССР

или

Я. Кочишу – Научный секретарь журнала

Кафедра Автоматизации
Будапештского Технического
Университета
H–1111 Будапешт XI, площадь Голдман Д., 3,
Венгрия

Авторам рекомендуется оставлять у себя копию всех представленных ими материалов для справок при корректуре.

Статьи, желательно на русском или английском языках, отпечатанные на бумаге хорошего качества, с промежутком в два интервала и широкими (4 – 5 см) полями должны содержать наименование статьи, фамилию автора (авторов), название страны. В конце статьи необходимо также указать фамилию автора, ответственного за корректуру гранок.

Статье должна предшествовать аннотация объемом до 50 – 100 слов.

Авторы при написании статьи должны придерживаться следующей формы: введение (постановка задачи), основное содержание и результаты, обсуждение, выводы и литература. Объем статьи не должен превышать 15 печатных страниц, включая таблицы и ссылки. Последовательность таблиц и рисунков должна быть отмечена на полях. Математические обозначения рекомендуется давать в соответствии с современными требованиями и традициями.

К статье обязательно должно быть приложено резюме-реферат. Резюме — на русском языке, если статья написана на английском, и наоборот — должно содержать краткое изложение текста статьи со ссылками на необходимые формулы и графики, имеющиеся в основном тексте. Объем разюме не должен превышать 10 – 15% объема статьи.

Авторам высылаются гранки статьи, которые они должны незамедлительно возвратить в Региональную секцию Редколлегии журнала.

Авторам обеспечивается бесплатно 100 оттисков их статей. Рукописи непринятых статей возвращаются авторам.

CONTENTS • СОДЕРЖАНИЕ

Krasovskii, N. N.: Game-theoretic control and problems of stability (*Красовский Н. Н.* Игровое управление и задачи устойчивости) 171

Tarján, T. G.: On complexity of switching circuits (*Тарян Т. Г.:* О сложности релейно-контактных схем) 183

Воронов А. А.—Скалецкий В. В.—Чистяков Ю. В. Определение функции распределения входящего потока в однолинейной системе массового обслуживания (*Voronov, A. A.—Skaletskii, V. V.—Chistyakov, Yu. V.:* Determination of the input distribution in a one-channel queue system) 197

Csáki, F.: Transformation methods from phase-variable form to canonical forms (*Чаки Ф.* Методы преобразования уравнений состояния к каноническому виду) 209

Bliznakov, J.: Main principles of the formation and management of organizational-productive systems (*Близнаков Й.* Основные принципы формирования и управления организационно-производственными системами) 227

Csibi, S.: M. S. Pinsker lecturing in Budapest 231

Tarnai, K.: Summer Computer Simulation Conference 233

Book Review 237

316.920

VOL. / TOM 3 * NUMBER / НОМЕР 4

ACADEMY OF SCIENCES OF THE USSR
HUNGARIAN ACADEMY OF SCIENCES

PROBLEMS OF CONTROL AND INFORMATION THEORY

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

АКАДЕМИЯ НАУК СССР
АКАДЕМИЯ НАУК ВЕНГРИИ

1974

AKADÉMIAI KIADÓ, BUDAPEST

2

PROBLEMS OF CONTROL AND INFORMATION THEORY

is an international quarterly sponsored jointly by the Presidium of the Academy of Sciences of the USSR and of the Hungarian Academy of Sciences. It offers publicity for original papers and short communications of the following topics:

— general theory of control systems and system theory

— theory of automata

— information theory

— operation research; theory of complex systems

— theory of economic control; system modelling

— theory and methods of adaptation, learning, identification and pattern recognition

— methods of information processing; application of digital computers in control and communication systems

— new physical principles in constructing technical devices for automation, control and information processing

The four issues published per year make up a volume of some 320 pp.

While this quarterly is mainly a publication forum of the research results achieved within the USSR and Hungary, also papers of international interest from other countries are welcome.

Distributor:

KULTURA

Hungarian Trading Co. for Books and Newspapers

H-1389 Budapest 62, P.O. Box 149, Hungary

The quarterly is published by

AKADÉMIAI KIADÓ

Publishing House of the Hungarian Academy of Sciences

H-1363 Budapest, P.O. Box 24, Hungary

Президиумами Академии Наук СССР и Академии Наук ВНР было принято решение о совместном издании журнала

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

Журнал создан для обеспечения быстрейшей публикации материалов о новейших результатах научных исследований, разработок и кратких сообщений о достижениях в следующих областях науки:

— общая теория процессов управления и теория систем
— теория автоматов
— теория информации
— исследование операций; теория сложных систем
— теория управления экономическими системами; модели систем
— теория и методы адаптации, обучения, идентификации и распознавания образов
— методы обработки информации; применение вычислительных машин в системах управления и передачи информации
— новые физические принципы создания технических средств автоматики, управления и передачи информации

Журнал издается четыре раза в год общим объёмом около 320 печатных страниц.

Журнал в первую очередь будет публиковать научные достижения обеих стран, а также статьи из других стран, представляющие международный интерес.

Распространитель:

KULTURA

Венгерское Общество по распространению книг и журналов

H-1389 Будапешт 62, п. о. 149, Венгрия

Издатель журнала:

AKADÉMIAI KIADÓ

Издательство Академии Наук Венгрии

H-1363 Будапешт, п. о. 24, Венгрия

PROBLEMS OF CONTROL AND INFORMATION THEORY

ПРОБЛЕМЫ УПРАВЛЕНИЯ И ТЕОРИИ ИНФОРМАЦИИ

EDITORS

B. N. PETROV (Moscow)
F. CSÁKI (Budapest)

DEPUTY EDITORS

V. S. PUGACHEV (Moscow)
V. I. SIFOROV (Moscow)
S. CSIBI (Budapest)

COORDINATING EDITORS

S. V. EMELIANOV (Moscow)
L. KALMÁR (Szeged)

M. A. GAVRILOV (Moscow)
I. CSISZÁR (Budapest)

A. M. LETOV (Moscow)
A. PRÉKOPA (Budapest)

L. VARGA (Budapest)

E. D. TERYAEV (Moscow)
J. KOCSIS (Budapest)

РЕДАКТОРЫ ЖУРНАЛА

Б. Н. ПЕТРОВ (Москва)
Ф. ЧАКИ (Будапешт)

ЗАМЕСТИТЕЛИ РЕДАКТОРОВ

В. С. ПУГАЧЕВ (Москва)
В. И. СИФОРОВ (Москва)
Ш. ЧИБИ (Будапешт)

ЧЛЕНЫ РЕДАКЦИОННОЙ КОЛЛЕГИИ

С. В. ЕМЕЛЬЯНОВ (Москва)
Л. КАЛМАР (Сегед)

М. А. ГАВРИЛОВ (Москва)
И. ЧИСАР (Будапешт)

А. М. ЛЕТОВ (Москва)
А. ПРЕКОПА (Будапешт)

Л. ВАРГА (Будапешт)

Е. Д. ТЕРЯЕВ (Москва)
Я. КОЧИШ (Будапешт)

AKADÉMIAI KIADÓ

PUBLISHING HOUSE OF THE HUNGARIAN ACADEMY OF SCIENCES

BUDAPEST

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

ALEKSANDR MIKHAILOVICH LETOV

On September 30, 1974 after a long and heavy illness departed from this life Professor Aleksandr Mikhailovich Letov, the distinguished Soviet scientist, D. Sc. (phys. math.), Corresponding Member of the Academy of Sciences of the USSR, Vice-director of the International Institute for Applied Systems Analysis, Head of Laboratory at the Institute of Control Science (automation and telemechanics), Member of the Editorial Board of an international quaterly "Problems of Control and Information Theory" and deputy editor-in-chief of "Automation and Remote Control", journal of the Academy of Sciences of the USSR.

A. M. Letov was born on November 24, 1911. He was graduated in mechanics and mathematics at the Moscow State University (MGU) in 1937. His results in mathematical theory of control processes are of great value for the theory of automatic control and design of control systems. He obtained fundamental results in the theory of stability of nonlinear systems which is the core of theory of analytical regulator design. A. M. Letov worked out new methods of analysis of control processes in large-scale systems and methods of system analysis.

A. M. Letov was an outstanding organizer as a great scientist. He was one of the initiators and one of the presidents of the International Federation of Automatic Control (IFAC). International collaboration between scientists of

different countries owes much to his activity in establishment and development of international scientific contacts.

The inexhaustible energy of A. M. Letov, his constant creative activity and love of humanity made him a person meeting whom was joy and pleasure. These qualities won him general love, authority and respect of the Soviet and foreign scientists.

Science has lost one of its best men, an extraordinary personality, whose life to his last days was devoted to serve humanity.

АЛЕКСАНДР МИХАЙЛОВИЧ ЛЕТОВ

30 сентября 1974 года после продолжительной тяжелой болезни скончался Александр Михайлович Летов, член-корреспондент Академии наук СССР, доктор физико-математических наук, профессор, Вице-директор Международного института прикладного системного анализа, заведующий лабораторией Института проблем управления (автоматики и телемеханики), член редколлегии журнала «Проблемы управления и теории информации», заместитель главного редактора журнала Академии наук СССР «Автоматика и телемеханика».

Выдающийся советский ученый А. М. Летов родился 24 ноября 1911 г. В 1937 г. окончил механико-математический факультет Московского государственного университета. Его труды в области математической теории процессов управления имеют первостепенное научное значение для теории автоматического управления и практики проектирования систем управления. Он внес значительный вклад в теорию устойчивости нелинейных систем, заложил основы теории аналитического конструирования регуляторов. В последние годы А. М. Летов разрабатывал новые методы исследования процессов управления большими системами и методы системного анализа.

Наряду с научное работой А. М. Летов вел большую научно-общественную и научно-организационную работу. Он был одним из основателей Международной федерации по автоматическому управлению (ИФАК), одним из ее президентов. Его деятельность в области установления и укрепления международных научных связей во многом способствовала развитию международного сотрудничества ученых разных стран.

Неистощимая энергия А. М. Летова, его жизнерадостность, постоянная творческая активность и человеколюбие делали его обаятельным человеком, общение с которым всегда доставляло радость и наслаждение. Этими своими качествами он заслужил всеобщую любовь, авторитет и уважение со стороны как советских, так и иностранных ученых.

В лице А. М. Летова наука потеряла одного из своих лучших представителей, а общество — замечательного человека, до последнего момента отдававшего свои силы служению человечеству.

Problems of Control and Information Theory, Vol. 3 (4), pp. 241—246 (1974)

УНИВЕРСАЛЬНЫЕ АЛГЕБРАИЧЕСКИЕ КОДЫ ДЛЯ СИММЕТРИЧНЫХ КАНАЛОВ

В. Д. ГОППА

(Владивосток)

(Поступила в редакцию 3 апреля 1973 г.)

Вводится понятие метрики Фитингофа и обобщается граница Гилберта на случай этой метрики. Доказывается существование универсальных (L, g)-кодов для симметричных каналов без памяти.

1. Пусть E^n — векторное пространство размерности n над $GF(q)$, $q = p^l$, p — простое число. Линейный код — подпространство E^n. Будем рассматривать специальные каналы, для которых эффективны алгебраические коды (каналы с независимым аддитивным шумом). Любой такой канал представляет собой совокупность двух источников — входного источника и источника шума, каждый из которых независимо друг от друга порождает векторы $x, y \in E^n$ в соответствии с некоторыми распределениями $p(x)$ (входа) и $p(y)$ (шума). Эти векторы складываются покоординатно, образуя вектор $z = x + y$ на выходе. Переходная матрица для таких каналов обладает специфической симметричной структурой, так что для краткости будем называть такие каналы *симметричными*, добавляя слово «без памяти» в том случае, когда распределение шума полиномиально.

2. Пусть S — некоторое семейство дискретных каналов без памяти, $I_s(x, y)$ — взаимная информация канала $s \in S$. Величина

$$\underline{C} = \sup_{p \in P} \inf_{s \in S} I_s(x, y), \tag{1}$$

где sup берется по всем распределениям на входе, называется *пропускной способностью класса каналов* S. Известно, что существует код, единый для всего класса S, со скоростью передачи R, как угодно близкой к $\underline{C}$ и вероятностью ошибки, как угодно близкой к нулю [1]. Кроме того, при $R > \underline{C}$ такого кода не существует.

В общем случае

$$\underline{C} \leq \bar{C} = \inf_{s \in S} \sup_{p \in P} I_s(x, y), \tag{2}$$

однако, в частном случае, когда S представляет собой множество всех симметричных каналов без памяти, пропускная способность которых не меньше

некоторой величины C_0, имеет место равенство

$$\underline{C} = \overline{C} = C_o \,. \tag{3}$$

Таким образом, существует универсальный оптимальный код для симметричных каналов без памяти. В настоящей заметке показано, что при поиске такого кода можно ограничиться сравнительно узким подклассом линейных кодов—множеством неприводимых (L, g)-кодов. Кроме того, вводится понятие метрики Фитингофа и обобщается граница Гилберта на случай этой метрики.

3. Пусть в векторе $x \in E^n$ буква $a_i \in E$ встречается m_i раз $(i = 1, 2, \ldots, q)$. Величина $\Phi(x) = H(m_1/n, \ldots, m_q/n)$ называется *весом Фитингофа* этого вектора. Если $n = r\nu$, то можно разбить вектор на r частей, отождествить каждый блок из ν элементов поля E с элементом нового алфавита и трактовать слово x длины n как слово z длины r в новом алфавите. Величину $\Phi(z)$ будем называть *ν-весом Фитингофа вектора* x и обозначать $\Phi_\nu(x)$. Ясно, что $\Phi(x) = \Phi_1(x)$.

Теорема 1. Вес Фитингофа обладает следующими свойствами:

1) *Для любого стационарного распределения на* E^n

$$M\,[\Phi_\nu(x)] \leq H(E^\nu), \tag{4}$$

где M обозначает математическое ожидание.

2) *Для любого эргодического распределения на* E^∞ *величина* $\Phi_\nu(x)/\nu$ *сходится почти всюду к* $H(E)$, *если* ν *выбрано так, что* $n \to \infty \Rightarrow \nu \to \infty$.

3) *Пусть* E_0 *обозначает подпространство векторов* $x \in E^n$, *для которых* $\Phi_\nu(x) = 0$. $\Phi_\nu(x)$ *является нормой на факторпространстве* E^n/E_0 *и определяет, таким образом, метрику* $\Phi_\nu(x - y)$ *на* E^n *(метрика Фитингофа).*

4) *Если* $V_\nu(w)$ — *шар радиуса* w *в метрике Фитингофа, т. е.* $V_\nu(w) = \{x : \Phi_\nu(x) \leq w\}$ *и* ν *выбрано так, что* $\log_q n - \log_q \log_q n - \nu \to \infty$ *при* $n \to \infty$, *то*

$$\log \mid V_\nu(w) \mid = n\big(w/\nu + 0(1)\big). \tag{5}$$

Доказательство. Свойство 1) следует из выпуклости энтропии. Свойство 2) тотчас же следует из определений. Свойство 3) менее тривиально и не было известно ранее. Отметим, прежде всего, что вектор из E_0 при $\nu = 1$ имеет вид $(a, a, \ldots, a)$, так что E_0 — действительно подпространство, причем ν-мерное. Далее, пусть $x, y \in E^n$, $x = (x_1, \ldots, x_n)$, $y = (y_1, \ldots, y_n)$. Введем вектор $x \otimes y$, координатами которого являются пары (x_i, y_i):

$$x \otimes y = \big((x_1\, y_1), \ldots, (x_n\, y_n)\big)\,. \tag{6}$$

Пусть $\{m_i\}$, $\{n_j\}$, $i, j = 1, \ldots, q$ — композиции векторов x и y соответственно, $\{m_{ij}\}$ — композиция вектора $x \otimes y$ (в алфавите $(E \times E)$. Ясно, что $\sum_i m_i = \sum n_j = \sum_{i,j} m_{ij} = n$,

$$\sum_i m_{ij} = n_j, \quad \sum_j m_{ij} = m_i .$$

Так как для любого распределения на $Y \times X$

$$H(X \otimes Y) \leq H(X) + H(Y), \tag{7}$$

то $\Phi(x \otimes y) \leq \Phi(x) + \Phi(y)$.

Далее, $\Phi(x + y) \leq \Phi(x \otimes y)$, так как композиция вектора $x + y$ получается из композиции $x \otimes y$ в результате отождествления некоторых пар $(x_i y_i)$ (например, если $x_1 + y_1 = x_2 + y_2$, $x_1 \neq x_2$, то в $x \otimes y$ координаты $(x_1 y_1)$ и $(x_2 y_2)$ будут считаться различными, а в $x + y$ — одинаковыми).

Таким образом, справедливо неравенство треугольника:

$$\Phi_\nu(x + y) \leq \Phi_\nu(x) + \Phi_\nu(y). \tag{8}$$

Остается заметить, что $\Phi_\nu(x)$ является функцией классов смежности по E_0:

$$x \equiv y \pmod{E_0} \Rightarrow \Phi_\nu(x) = \Phi_\nu(x). \tag{9}$$

Свойство 4) использовалось Фитингофом [2] в несколько другой формулировке. Оно является следствием разложения Стирлинга полиномиального слагаемого $n!/m_1! \ldots m_q!$ и приведено, например, в книге Фано [3] (для $\nu = 1$):

$$\log | V_1(w) | = n(w + 0(1)), \; 0(1) = (q \log n) \; n. \tag{10}$$

Чтобы перейти к $| V_\nu(w) |$, нужно подставить в это выражение q^ν вместо q и $r = n/\nu$ вместо n.

4. Обычно в теории кодирования стараются построить коды с большим весом Хэмминга. Такая постановка оправдана для двоичного симметричного канала, но уже в недвоичном случае, не говоря уже о каналах с памятью, метрика Хэмминга становится неудовлетворительной. Коды же с большим весом Фитингофа пригодны для более широкого класса каналов.

Теорема 2 (аналог границы Гилберта). *Существует неприводимый* (L, g)-*код длины* n *с минимальным весом Фитингофа, равным* w *и скоростью передачи*

$$k/n > 1 - (\log_q 2) w/\nu - \varepsilon \tag{11}$$

для любого $\varepsilon > 0$ *и всех достаточно больших* n. Доказательство аналогично доказательству последней теоремы из статьи [4]. Заметим, что в случае ДСК

можно выбрать $\nu = 1$, тогда приведенное неравенство совпадает с границей Гилберта, поскольку в этом случае $w = H(d/n)$, где d — минимальный вес Хэмминга кода.

5. В работе [5] было введено *декодирование по весу Фитингофа*, при котором в каждом смежном классе линейного кода выбирается вектор с наименьшим весом Фитингофа.

Теорема 3. Пусть $S(C_0)$ обозначает класс всех симметричных каналов без памяти, пропускная способность которых не меньше C_0. Существует последовательность неприводимых (L, g)-кодов со скоростью передачи $R = (\log q)(k/n)$, такая, что вероятность ошибки декодирования по весу Фитингофа для этих кодов стремится к 0 с ростом n равномерно по всем каналам из $S(C_0)$ при любом $C_0 > R$.

Доказательство. Методом случайного кодирования автором [5] была получена следующая оценка для вероятности ошибки, осредненной по всему классу неприводимых (L, g)-кодов при использовании декодирования по весу Фитингофа:

$$\overline{p}_e \leq 2^{-n\left(\frac{C-R}{2}\right)} + P\left[\,|\Phi(x) - H(p)| > \frac{C-R}{2}\right]. \tag{12}$$

Здесь $C = \log q - H(p)$, $H(p)$ — энтропия шума, R — скорость передачи кода. Скорость сходимости $m_i/n \to p_i$ легко оценивается с помощью перекоса распределений [3]:

$$P[\,|\,m_1/n - p_1\,| \leq \varepsilon, \ldots, |\,m_{q-1}/n - p_{q-1}\,| \leq \varepsilon] \geq 1 - 2^q \cdot 2^{-nE(\overline{p}, p)}$$

где
$$\overline{p}_i = p_i + \varepsilon,\ i = 1, 2, \ldots, q-1,\ \varepsilon < 1/q,$$

$$E(\overline{p}, p) = H(p) + \varepsilon H'(p) - H(\overline{p}) = H(p_1, \ldots, p_{q-1}) +$$
$$+ \varepsilon \sum_i \frac{\partial H(p_1, \ldots, p_{q-1})}{\partial p_i} - H(p_1 + \varepsilon_1 \ldots, p_{q-1} + \varepsilon).$$

Но если $|\,m_i/n - p_i\,| \leq \varepsilon < 1/2$, $i = 1, \ldots, q-1$, то $|\,H(m_1/n, \ldots, m_{q-1}/n) - H(p_1, \ldots, p_{q-1})\,| \leq -q\,\varepsilon \log \varepsilon \leq q\sqrt{\varepsilon}$. ([1]).

Поэтому

$$P\left[\,|\Phi(x) - H(p)| > \frac{C-R}{2}\right] \leq 2^q \cdot 2^{-nE(\overline{p}, p)}, \tag{13}$$

где

$$\overline{p}_i = p_i + \left(\frac{C-R}{2q}\right)^2.$$

Таким образом, получена оценка снизу для экспоненты вероятности ошибки

$$-\lim_{n\to\infty}\frac{1}{n}\log\overline{p}_e \geq E(\overline{p}, p)\,, \tag{14}$$

где функция надежности $E(\overline{p}, p)$ вычисляется при

$$\varepsilon = \overline{p}_i - p_i = \left(\frac{C-R}{2q}\right)^2 \quad i = 1, \ldots, q-1\,. \tag{15}$$

Функция $(E(\overline{p}, p)$ всегда больше нуля при $C - R > 0$, поэтому для любого $C_0 > R$

$$\min_{S(C_0)} E(\overline{p}, p) > 0$$

(здесь минимум берется по всем распределениям из $S(C_0)$). Таким образом, существует код, удовлетворяющий условиям теоремы.

6. Теория оптимального кодирования для симметричных каналов имеет специфический характер. Для того чтобы построить для такого канала код, удовлетворяющий основной теореме Шеннона, не нужно знать распределения вероятностей шума, достаточно знать одно число—пропускную способность канала.

Появившееся при доказательстве теоремы 1 произведение двух слов $x \otimes y$ позволяет ввести невероятностный аналог «взаимной информации» и обобщить полученные результаты на произвольные каналы без памяти. Этой теме будет посвящена следующая статья.

Литература

1. *Blackwell, D., Breiman, L., Thomasian, A. J.*, The Capacity of a class of channels. Annals of Math. Stat. **30** (1959) **4**, 1229—1241.
2. *Фитингоф Б. М.* Оптимальное кодирование при неизвестной и меняющейся статистике сообщений. Проблемы передачи информации **2** (1966), 2, 3—11.
3. *Фано Р.* Передача информации. «Мир», М., 1965.
4. *Гоппа В. Д.* Рациональное представление кодов и (L, g)-коды. Проблемы передачи информации **7** (1971), 3, 41—49.
5. *Гоппа В. Д.* Универсальное декодирование для симметричных каналов. Проблемы передачи информации (1975) *1*.

Universal algebraic codes for symmetric channels

V.D. GOPPA

(Vladivostok)

Let x be a sequence of n letters and m_i be the number of letters $a_i \in X$ in this sequence ($i = 1, 2, \ldots, N = |X|$). Fitingoff weight of x is a real number $\Phi(x) = H(m_1/n, \ldots, m_N/n)$.

Let X be a finite field $GF(q)$ and V^n be a vector space of n-tuples over X. Consider the set V_0 of all vectors $x \in V^n$ with $\Phi(x) = 0$. V_0 is a subspace of V^n and $\Phi(x)$ is a function of cosets (9).

Theorem 1. $\Phi(x)$ is a norm on quotent space V^n/V_0 and defines a metric (Fitingoff distance).

Theorem 2 (analogue of Gilbert bound). It is possible to construct a linear code of length n and minimum Fitingoff distance w with k information symbols satisfying for any $\varepsilon > 0$ and $n \to \infty$ inequality (11).

Define now Fitingoff weight decoding for a linear code choosing each coset leader to have minimum Fitingoff weight in its coset.

Theorem 3. Let $S(C_0)$ be a class of all symmetric channels without memory with capacity not less than C_0. There exists an irreducible (L, g)-code of the rate R with a vanishingly small probability of error of Fitingoff weight decoding simultaneously for all channels from $S(C_0)$ for any $C_0 > R$.

В. Д. Гоппа
Институт автоматики и процессов управления
с вычислительным центром ДВНЦ АН СССР
СССР, Владивосток, ул. Суханова, 5а

Problems of Control and Information Theory, Vol. 3 (4), pp. 247–260 (1974)

ADAPTIVE DUAL EXTREMUM CONTROL BY FINITE ORDER VOLTERRA MODEL

R. HABER, L. KEVICZKY

(Budapest)

(Received June 15, 1974)

A nonlinear dynamic model, the so-called finite order Volterra model is considered, which may be separated into a linear dynamic term and a nonlinear one with memory. This model is linear in parameters of finite number and has, nevertheless, infinite memory. This approximation represents more exactly a general nonlinear dynamic model described by the Volterra series, than the Hammerstein model, having already been investigated in details in several earlier papers [6], [7], [10]. On the basis of this new model an adaptive extremum control algorithm is formulated according to the dual control theory. The steps of the on-line identification and control are computed by stochastic approximation algorithms optimized in quadratic sense. The noises, which disturb the plant are reduced to the output and are assumed to be a linear autoregressive stochastic process. (Here the equation error is also an autoregressive one.) The elaborated algorithms are supported by simulation results.

1. Introduction

In certain industrial control problems the task is not to control a constant or time-varying reference value as it is usual, but to ensure the optimum value of the controlled variable. Today when the aim is to exploit the industrial equipments as much as perfectly, the extremum controls have come into the limelight. Such a task occurs e.g. at the maximization of the crossflow-rate of a grinding mill (since if the grist is zero then the milling product is also zero, or if the input quantity is large enough then the mill goes to saturation, i.e. the product decreases).

As the industrial processes are generally complicated — or at least unknown — the optimum seeking is made by fitting a model of known structure in the environment of a given point of the input/output surface and steps are taken according to the estimated parameters of this approximating model. Thus the optimization of the process can be separated to two connected parts; to the recognition or identification and to the optimization of a given cost function. The above control strategy based on the adaptive model is called dual control after Feldbaum [1].

In spite of the fact, that the industrial processes are described by dynamic equations, mostly static optimization methods have been used which are based upon only the static characteristic of the process [5], [11]. It is obvious

that taking the dynamic feature of the plant into account, the control is quicker. The transient properties of the process characterizing the dynamics may be described in different ways. Tsypkin constructed the system by cascade connection of a linear dynamic and a nonlinear static element, where the linear term was assumed to be known [2]. Kocsis proposed an iterative scheme for the identification of the parameters of the dynamic and static parts [3]. In both references the dynamical part was assumed to be described by its weighting function series. Keviczky and Haber applied the generalized Hammerstein model having finite number of parameters and, nevertheless, an infinite memory for the control task [10]. As this model is linear in parameters, all parameters could be estimated simultaneously. Further improvement of the approximating modell will be given in this paper.

2. The finite order Voltera model

A nonlinear static function may be approximated by a Taylor series and a linear dynamic system by an impulse response series. Generally the so-called Volterra integral operator is used to describe nonlinear, dynamic, continuous systems. Since the steady state characteristics of the single input plant are preferably approximated by a quadratic parabola in the vicinity of the working point, and the discrete form has advantages for data processing by digital computers, further on we deal with the Volterra series of quadratic order kernels:

$$y(t+d) = r_0 + \sum_{i=0}^{\infty} w_i\, u(t-i) + \sum_{i=0}^{\infty}\sum_{j=0}^{\infty} w_{ij}\, u(t-i)\, u(t-j)\,. \tag{1}$$

Here $u(t)$ and $y(t)$ are the input and output signals, respectively in the tth moment, assuming the sampling time $T = 1$, $(d-1)$ is the dead time, and r_0, w_i, w_{ij} are the kernels. Let us introduce the following symbols:

$$B_1(z^{-1}) = b_{10} + b_{11}z^{-1} + \ldots + b_{1m}z^{-m} \tag{2}$$

and

$$\mathbf{B}_2 = \begin{bmatrix} b_{200} & b_{201} & \ldots & b_{20m} \\ 0 & b_{211} & \ldots & b_{21m} \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \ldots & b_{2mm} \end{bmatrix} \tag{3}$$

where z^{-1} is the backward shift operator and

$$\mathbf{f}^T(u(t)) = [u(t), \ldots, u(t-m)]\,. \tag{4}$$

Rewriting Eq. (1), we get:

$$y(t+d) = r_0 + B_1(z^{-1})\,u(t) + \mathbf{f}^T(u(t))\,\mathbf{B}_2\,\mathbf{f}(u(t)). \tag{5}$$

This is the so-called finite order Volterra weighting function model and has only finite memory. Extending (5) with the older values of the output we get the finite order Volterra model [4]:

$$y(t+d) = -\widetilde{A}(z^{-1})\,y(t+d) + r_0^* + B_1(z^{-1})\,u(t) + \mathbf{f}^T(u(t))\,\mathbf{B}_2\mathbf{f}(u(t)) \tag{6}$$

where

$$A(z^{-1}) = 1 + \widetilde{A}(z^{-1}) = 1 + a_1 z^{-1} + \ldots + a_n z^{-n}. \tag{7}$$

This model is shown in Fig. 1, where the nonlinear memory can be well seen. Assuming that $\mathbf{B}_2$ is a diagonal matrix, we get the well-known generalized Hammerstein model. Contrary to the Hammerstein model the finite order Volterra model considers the crossproducts between the input signal terms of different time moments as well, therefore it gives a better approximation for the general Volterra series. Its finite order feature ($m < \infty$) causes that only those crossproduct terms occur, which belong to time points of interval less or equal to m. Increasing m or n the model gives better and better approach of a general nonlinear dynamic plant. Further extensions and details of the model can be found in [4].

The form (6) is linear in the parameters, i.e. the output can be given as a scalar product of the parameter vector $\mathbf{p}$ and the situation vector $\mathbf{g}$:

$$y(t+d) = \mathbf{g}^T(t)\,\mathbf{p} \tag{8}$$

where

$$\mathbf{p} = [r_0^*, b_{10}, b_{11}, \ldots, b_{1m}, b_{200}, b_{201}, \ldots, b_{2mm}, a_1, \ldots, a_n]^T \tag{9}$$

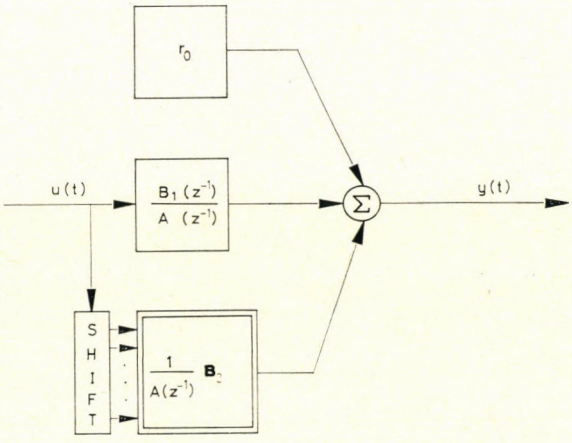


Fig. 1

and

$$\mathbf{g}^T(t) = [1, u(t), u(t-1), \ldots, u(t-m), u^2(t), u(t)\,u(t-1), \ldots$$

$$\ldots, u^2(t-m), -y(t+d-1), \ldots, -y(t+d-n)]. \tag{10}$$

Here

$$r_0^* = A(1)\,r_0\,. \tag{11}$$

3. Adaptive extremum control

The goal of the control may be the extremization of a state variable or the output of the plant. We disregard the restrictions [3]. Assume that our task is to extremize the output of the plant on the basis of the approximating model (6):

$$y(t+d) = \frac{r_0^*}{A(1)} + \frac{B_1(z^{-1})}{A(z^{-1})}\,u(t) + \frac{1}{A(z^{-1})}\,\mathbf{f}^T\big(u(t)\big)\,\mathbf{B}_2\,\mathbf{f}\big(u(t)\big) \tag{12}$$

and

$$E\{y(t)\} \Rightarrow \underset{u(t)}{\mathrm{EXTR}}. \tag{13}$$

It seems to have to be controlled for $\dfrac{dy(t+d)}{du(t)} = 0$, but this one-stage strategy would extremize only an inner state variable. Namely by getting the control steady, the surface to be extremized is a static one:

$$Y = \frac{r_0^*}{A(1)} + \frac{B_1(1)}{A(1)}\,U + \frac{B_2(1)}{A(1)}\,U^2 = c_0 + c_1\,U + c_2\,U^2 \tag{14}$$

where

$$B_2(1) = \sum_{i=0}^{m}\sum_{j=0}^{m} b_{2ij} \tag{15}$$

and $U = u(\infty)$, $Y = y(\infty)$.

Now the task is to control for $\dfrac{dY}{dU} = 0$. Multiplying (14) by $A(1)$, we have to ensure $\dfrac{dY\,A(1)}{dU} = A(1)\,\dfrac{dY}{dU} = 0$, which gives the same final result. In this suboptimal case the computation algorithm is a bit easier but practically there is no difference.

Assume, that the input $(u(t))$ and output signals $(y(t))$ of the plant are measured in the tth moment, on the basis of which a new estimation is made

for the parameters $(\mathbf{p}(t-1) \Rightarrow \mathbf{p}(t))$, the new control signal is calculated and immediately put on the plant $(u(t) \Rightarrow u(t+1))$. (See Fig. 2.)

The control problem can be solved by a stochastic approximation algorithm [2]:

$$u(t+1) = u(t) - r_c(t) \frac{dY(\hat{\mathbf{p}}(t), u(t))}{dU} \tag{16}$$

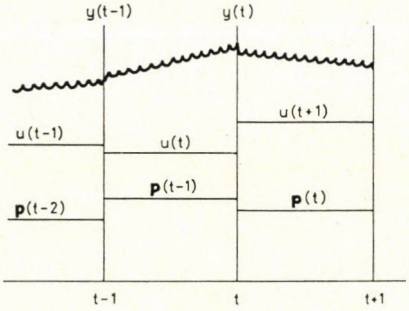


Fig. 2

where the necessary terms may be computed from the estimated $(\hat{v})$ value of (14). Thus

$$\frac{dY(\hat{\mathbf{p}}(t), u(t))}{dU} = \hat{c}_1(t) + 2\hat{c}_2(t)\, u(t) \,. \tag{17}$$

Optimizing (16) in quadratic sense the deterministic convergence coefficient is:

$$r_{cd}(t) = \frac{1}{2\hat{c}_2(t)} \,. \tag{18}$$

Assuming that the measurements and the plant are disturbed by noise we use the stochastic convergence coefficient:

$$r_{cs}(t) = \sum_{k=1}^{t} \frac{1}{\hat{c}\,(k)} = \frac{r_{cs}(t-1)}{1 + 2\hat{c}_2(t)\, r_{cs}(t-1)} \,. \tag{19}$$

The deterministic control means that the local extremum is sought only in possession of the actual information:

$$u(t+1) = -\frac{\hat{c}_1(t)}{2\hat{c}_2(t)} \tag{20}$$

which is the extremum point of (14).

4. On-line identification

In the tth moment the best estimate of the parameters can be calculated — having no a priori information — by the least squares (LS) method. This minimizes the expected value of the performance index:

$$E\{Q(t)\} = E\{Q(u(t-d), y(t), \hat{\mathbf{p}}(t-l))\} =$$
$$= E\left\{\frac{1}{2}(y(t) - \hat{y}(t))^2\right\} \Rightarrow \underset{\hat{\mathbf{p}}(t)}{\mathrm{MIN}}. \tag{21}$$

Applying again the stochastic approximation:

$$\hat{\mathbf{p}}(t) = \hat{\mathbf{p}}(t-1) - \mathbf{R}_{id}(t)\frac{dQ(t)}{d\hat{\mathbf{p}}(t-1)}. \tag{22}$$

Here

$$\frac{dQ(t)}{d\hat{\mathbf{p}}(t-1)} = -(y(t) - \hat{y}(t))\,\mathbf{g}(t-d) \tag{23}$$

as $\mathbf{g}(t-d)$ includes terms $y(t-j)$, $j \geq 1$ and $u(t-i)$, $i \geq d$, which do not depend on $\hat{\mathbf{p}}(t-1)$ because the simulation of continuous dynamic systems by discrete difference equation is made generally by $d \geq 1$. Otherwise a complementary product — including the so-called sensitivity value — may be taken into account on the right side of Eq. (23) [10], [11].

If in Eq. (22) $\mathbf{R}_{id}$ is optimized in quadratic sense then the stochastic approximation gives the same result as the generally known recursive LS method. In this case the optimum weighting matrix can be computed recursively:

$$\mathbf{R}_{id}(t) = \mathbf{R}_{id}(t-1) - \frac{\mathbf{R}_{id}(t-1)\,\mathbf{g}(t-d)\,\mathbf{g}^T(t-d)\,\mathbf{R}_{id}(t-1)}{1 + \mathbf{g}^T(t-d)\,\mathbf{R}_{id}(t-1)\,\mathbf{g}(t-d)} \tag{24}$$

where $\mathbf{R}_{id}(0)$ can be chosen in different ways e.g. as a diagonal matrix of a high (10^3—10^5) constant value or as the result of an initial least squares estimation.

5. Filtering the output correlated noise

In industrial environment the plant and the measurements may be disturbed by noises. Reducing them to the output we get:

$$y(t+d) = \mathbf{g}^T(t)\,\mathbf{p} + \lambda C(z^{-1})\,e(t+d) \tag{25}$$

regarding only a linear noise model. In the case when the noise model, reduced to the output, is an autoregressive one ($C(z^{-1}) = 1$, white noise equation error) the LS method may be applied to ensure asymptotically unbiased parameter estimates. Considering, that there is no proper on-line algorithm for the maxi-

mum likelihood estimation, we use the Hastings—James recursive generalized LS (GLS) method for the identification of the finite-order Volterra model [8].

The moving average process $(C(z^{-1})e(t))$ is supposed to be approximated by an autoregressive filter $(1/H(z^{-1})\,e(t))$, and increasing the degree of the filter a satifactory estimate can be obtained.

$$H(z^{-1}) \approx \frac{1}{C(z^{-1})}\cdot \tag{26}$$

where

$$H(z^{-1}) = 1 + h_1 z^{-1} + \ldots + h_s z^{-s}. \tag{27}$$

Let us suppose that an initial LS estimate has been performed. Let us compute the equation error

$$\hat{\varepsilon}(t) = y(t) - \mathbf{g}^T(t-d)\,\hat{\mathbf{p}}(t)\,. \tag{28}$$

This will be approximated according to (25) and (26):

$$\hat{\varepsilon}(t) = \lambda\,\hat{C}(z^{-1})\,e(t) = \lambda\,\frac{1}{\hat{H}(z^{-1})}\,e(t)\,. \tag{29}$$

Introducing the vectors

$$\mathbf{h} = [h_1, h_2, \ldots, h_s]^T \tag{30}$$

and

$$\mathbf{f}^T(t) = [-\hat{\varepsilon}(t-1),\ -\hat{\varepsilon}(t-2), \ldots, -\hat{\varepsilon}(t-s)]. \tag{31}$$

The

$$\lambda e(t) = \hat{\varepsilon}(t) - \mathbf{f}^T(t)\,\mathbf{h}(t-1) \tag{32}$$

is the source of the white noise. Therefore the cost function, to be minimized is

$$E\{Q_f(t)\} = E\{Q_f(\hat{\varepsilon}(t), \hat{\mathbf{h}}(t-1))\} =$$

$$= E\left\{\frac{1}{2}\left(\hat{\varepsilon}(t) - \mathbf{f}^T(t)\,\hat{\mathbf{h}}(t-1)\right)^2\right\} \Rightarrow \underset{\hat{\mathbf{h}}(t)}{\mathrm{MIN}}\,. \tag{33}$$

Applying again the stochastic approximation:

$$\hat{\mathbf{h}}(t) = \hat{\mathbf{h}}(t-1) - \mathbf{R}_f(t)\,\frac{dQ_f(t)}{d\hat{\mathbf{h}}(t-1)} \tag{34}$$

where

$$\frac{dQ_f(t)}{d\hat{\mathbf{h}}(t-1)} = -\,[\hat{\varepsilon}(t)\,\mathbf{f}^T(t) - \hat{\mathbf{h}}(t-1)]\,\mathbf{f}(t) \tag{35}$$

and

$$\mathbf{R}_f(t) = \mathbf{R}_f(t-1) - \frac{\mathbf{R}_f(t-1)\,\mathbf{f}(t)\,\mathbf{f}^T(t)\,\mathbf{R}_f(t)}{1 + \mathbf{f}^T(t)\,\mathbf{R}_f(t-1)\,\mathbf{f}(t)} \tag{36}$$

if the algorithm is optimized in quadratic sense.

After computing the filter parameters let us multiply (25) by $\hat{H}(z^{-1})$ and introduce the following notations for the filtered values:

$$\begin{aligned} y^F(t) &= \hat{H}(z^{-1})\, y(t) \\ u^F(t) &= \hat{H}(z^{-1})\, u(t) \\ u^{2F}(t-i, t-j) &= \hat{H}(z^{-1})\, u(t-i)\, u(t-j) \\ \mathbf{1}^F(t) &= \hat{H}(z^{-1})\mathbf{1} = H(1). \end{aligned} \tag{37}$$

The filtered system equation is

$$y^F(t+d) = \mathbf{g}^{FT}(t)\, \hat{\mathbf{p}}(t) + \lambda C(z^{-1})\, \hat{H}(z^{-1})\, e(t+d) \tag{38}$$

where

$$\begin{aligned} \mathbf{g}^{FT}(t) = [&\hat{H}(1), u^F(t), u^F(t-1), \ldots, u^F(t-m), u^{2F}(t,t), \\ &u^{2F}(t, t-1), \ldots, u^{2F}(t-m, t-m), -y^F(t+d-1), \ldots \\ &\ldots, -y^F(t+d-n)]. \end{aligned} \tag{39}$$

If the filtering is efficient then $C(z^{-1})\, \hat{H}(z^{-1})\, e(t)$ is less correlated then $C(z^{-1})\, e(t)$, so the estimation of the parameters of Eq. (38) gives better result. The algorithm is the same as the above on-line LS estimation, only the measured values have to be replaced by their filtered ones. Similarly to (22), (23) and (24) the procedure is the following:

$$\hat{\mathbf{p}}(t) = \hat{\mathbf{p}}(t-1)\, \mathbf{R}^F_{id}(t)\, \frac{dQ^F(t)}{d\hat{\mathbf{p}}(t-1)} \tag{40}$$

where

$$\frac{dQ^F(t)}{d\hat{\mathbf{p}}(t-1)} = -[y^F(t) - \mathbf{g}^F(t-d)\, \hat{\mathbf{p}}(t)]\, \mathbf{g}^F(t-d) \tag{41}$$

and

$$\mathbf{R}^F_{id}(t) = \mathbf{R}^F_{id}(t-1) - \frac{\mathbf{R}^F_{id}(t-1)\, \mathbf{g}^F(t-d)\, \mathbf{g}^{FT}(t-d)\, \mathbf{R}^F_{id}(t-1)}{1 + \mathbf{g}^{FT}(t-d)\, \mathbf{R}^F_{id}(t-1)\, \mathbf{g}^F(t-d)}. \tag{42}$$

On the basis of the obtained parameter vectors new estimation can be made for the filter parameters and so on. It is important that $\hat{\varepsilon}(t)$ has to be computed from the original measured values (28).

Before presentig some simulation results, some remarks have to be made. The first is that the GLS method theoretically does not ensures the true estimate of the parameters by all means [9], but practically it is a good and efficient method and suitable for on-line computation [8]. The goodness of the method for nonlinear plants had also been supported by several simulation results [6], [7]. Besides the separated control and on-line identification, the

algorithm proved to be excellent at the general dual control, i.e. at the simultaneous identification and control, when each identification step is followed by a control one. In spite of the fact that the input values depend on the noises, the estimation gives an unbiased result because $u(t)$ depends only on $e(t-1)$, and — assuming white noise equation error and LS method — their cross-products occur only in form $u(t-i)\,e(t)$, $i \geq 0$, but $e(t)$ and $e(t-j)$, $j \neq 0$ are not correlated. This fact can be seen easily by substituting $\mathbf{p}(t)$ by the real $\mathbf{p}$ and

$$y(t) = \mathbf{g}^T(t-d)\,\mathbf{p} + \lambda e(t) \tag{43}$$

in Eqs (22) and (23):

$$\mathbf{p} = \mathbf{p} + \mathbf{R}_{id}(t)\,\lambda e(t)\,\mathbf{g}(t-d) \tag{44}$$

and considering the probability limes of the bias:

$$\mathrm{p}\lim_{N\to\infty} \{\mathbf{R}_{id}(t)\,\lambda\, e(t)\,\mathbf{g}(t-d)\} = \mathbf{0}\,. \tag{45}$$

In filtered case the same is true for $u^F(t-i)e(t)$. Of course, while the equation error in the filtered system equation is not sufficiently uncorrelated, the parameter estimation and hereby the control, is biased. The rightness of the presented algorithm is supported by several simulation results.

6. Simulation results

The programs of every algorithm mentioned in this paper have been made. We present only two from the several experimental tests. First let us consider the second order Volterra system equation (6) with parameters

$$r_0 = 2$$

$$B_1(z^{-1}) = 1 + 0.5\,z^{-1}$$

$$B_2 = \begin{bmatrix} 1 & 0.2 \\ 0 & 0.5 \end{bmatrix}$$

$$A(z^{-1}) = 1 - 1.5\,z^{-1} + 0.7\,z^{-2}$$

$$d = 1\,.$$

The results of the on-line LS identification of the noiseless system are presented in Fig. 3 with $u(0, 1)$ and starting value $\mathbf{R}_{id}(0) = 10^4$. The steady state characteristics is

$$Y = 2 + 7.5\,U + 8.5\,U^2$$

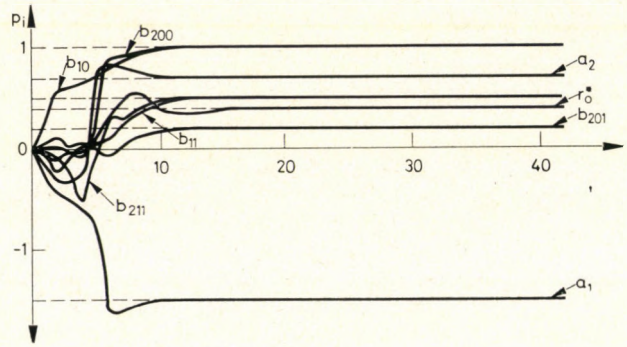


Fig. 3

which has a minimum at $U^0 = -0.4412$ and here the value of the function is $Y^0 = 0.3456$. The simulation results of the deterministic dual control of the same system are shown in Fig. 4, where the model output $(\hat{y}(t) = \mathbf{g}^T(t-d)\,\hat{\mathbf{p}}(t-1))$ and the control input $(u(t))$ are presented.

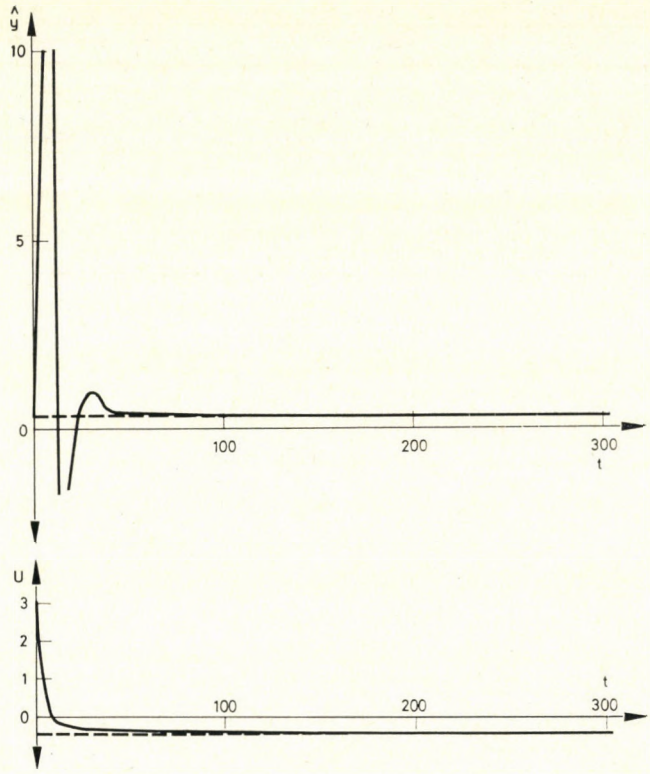


Fig. 4

As a second example let us consider the first order Volterra system — which is equal to the Hammerstein one — with parameters

$$r_0 = 2$$

$$B_1(z^{-1}) = 0.4$$

$$B_2(z^{-1}) = 0.2$$

$$A(z^{-1}) = 1 - 0.8\, z^{-1}$$

$$d = 1.$$

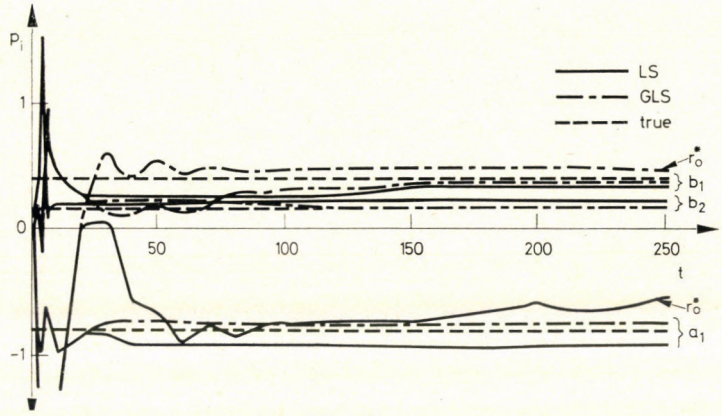


Fig. 5

The plant was disturbed by correlated output noise of very high value

$$\lambda \frac{1}{A(z^{-1})\, C(z^{-1})}\, e(t) = 0.6 \frac{1}{A(z^{-1})\,(1 - 0.8z^{-1})}\, e(t)$$

where $e(t) = e(0, 1)$. The results of the on-line GLS identification are presented in Fig. 5 with $u(0, 2)$ and starting value $\mathbf{R}_{id}^{F}(0) = 10^5$. The steady state characteristics is

$$Y = 2 + 2U + U^2$$

which has a minimum at $U^0 = -1$ and here the value of the function is $Y^0 = 1$. The simulation results of the stochastic dual control — including the filtering of the noise — of the same system are shown in Fig. 6, where the model output $(\hat{y}(t))$ and the control input $(u(t))$ are presented.

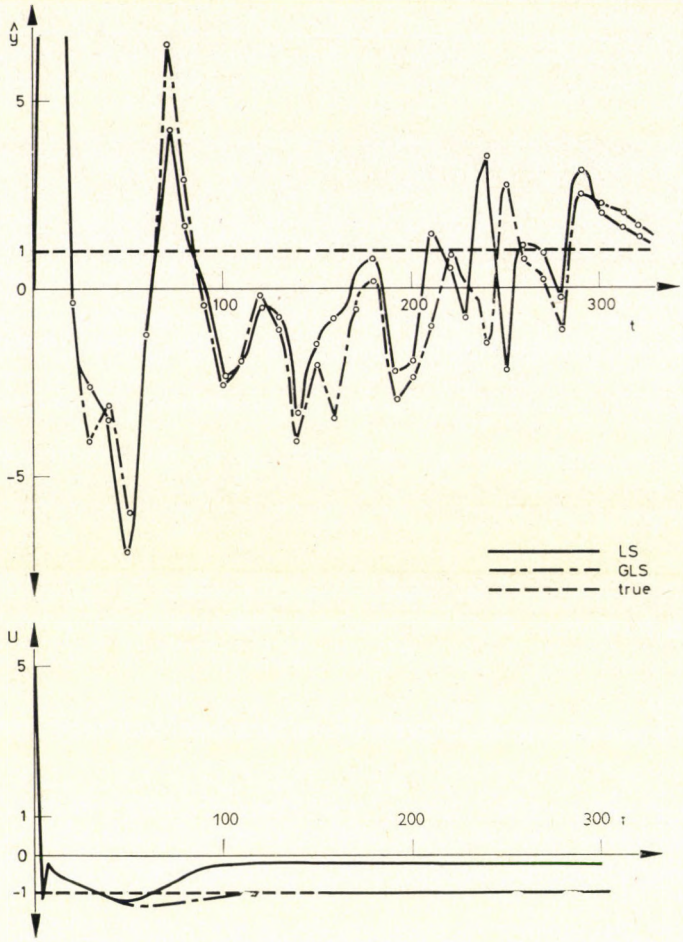


Fig. 6

7. Conclusions

In the present work an adaptive dual extremum control formed on the basis of a discrete time finite order Volterra model is considered. This model is a further generalization of the Hammerstein one and gives better approximation for a general nonlinear dynamic plant described by its Volterra series. The most general case of the dual control, the simultaneous identification and control are investigated. The on-line algorithms of the adaptive control and identification — including the filtering of the noises — are given and optimized in quadratic sense.

From several simulation results we have drawn the conclusion that the finite order Volterra model gives a good descripton for the general nonlinear dynamic plant, though the task of the researchers remains to find better and better models — both linear and nonlinear in parameters — for the control of industrial plants.

8. Acknowledgement

The authors are grateful to J. Peter, graduate of Technical University of Budapest for helping in simulation.

References

1. *Feldbaum, A. A.*, Theory of Dual Control I—II—III—IV. Avtomatika i Telemekhanika, **21** (1960) 9, 11; **22** (1961) 1, 2.
2. *Tsypkin, Ya. Z.*, Adaptation and Learning in Automatic Control Systems. Nauka, Moscow, 1968.
3. *Kocsis, J.*, A Possible Use of Adaptive Programming. Problems of Control and Information Theory, **1** (1972) 153—162.
4. *Haber, R., Keviczky, L.*, Nonlinear Structures for System Identification. Periodica Polytechnica, Electrical Engineering, 1974 (In Press)
5. *Keviczky, L.*, On Some Questions of Adaptive Process Control. Measurements and Automation (in Hungary), **XIX**, *11*, 417—422.
6. *Bányász, Cs., Haber, R., Keviczky, L.*, Some Estimation Methods for Nonlinear Discrete Time Identification. 3rd IFAC Symposium, Hague/Delft, 1973.
7. *Bányász, Cs., Keviczky, L., Haber, R.*, Identification of Discrete Dynamic Systems with Separable Nonlinearity. 3rd All Union Conference on Statistical Methods in Control Processes, Vilnius, 1973.
8. *Hastings-James, R., Sage, M. W.*, Recursive Generalized Least Squares Procedure for On-line Identification of Process Parameters. Proceedings of the IEE, **116** (1969) 12, 2057—2062.
9. *Söderström, T.*, On the Convergence Properties of the Generalized Least Squares Identification Method, Report 7228, Division of Automatic Control, Lund Institute of Technology, 1972.
10. *Keviczky, L., Haber, R.*, Adaptive Dual Control by Hammerstein Model. IFAC Symposium on Stochastic Control, Budapest, 1974.
11. *Haber, R., Keviczky, L.*, On Simultaneous Identification and Optimization of a Static Response Surface, IFAC Symposium on Optimization Methods, Varna, 1974.

Адаптивное дуальное экстремальное управление при помощи конечной модели Вольтерра

Р. ХАБЕР, Л. КЕВИЦКИ

(Будапешт)

Рассматривается нелинейная динамическая модель, так называемая модель Вольтерра конечного порядка. Данная модель может быть разделена на линейную, относительно ограниченного количества параметров, динамическую часть и нелинейную часть с памятью бесконечной емкости.

При описании обобщенной нелинейной динамической модели, построенной с применением рядов Вольтерра, обеспечивается более точное приближение, чем при использовании модели Гаммерштайна, детально рассмотренной в некоторых предыдущих ра-

ботах [6, 7, 10]. На основе применения предлагаемой новой модели разработан алгоритм адаптивного экстремального управления в соответствии с принципами дуального управления. Расчёты отдельных этапов процесса он-лайн идентификации и управления проводятся при помощи применения алгоритмов стохастической аппроксимации, являющихся оптимальными в квадратичном смысле. Предполагается, что помехи, воздействующие на объект управления и приведенные к выходу объекта, представляют собой линейные авторегрессивные случайные процессы и погрешности в уравнениях системы являются также авторегрессивными.

Приводятся результаты моделирования разработанных алгоритмов.

R. Haber and
L. Keviczky
Department of Automation
Technical University of Budapest
1111 Budapest, XI., Goldmann Gy. tér 3.
Hungary

Problems of Control and Information Theory, Vol. 3 (4), pp. 261—276 (1974)

ВЫЧИСЛЕНИЕ ПЕРЕХОДНЫХ ПРОЦЕССОВ В КОМБИНАЦИОННЫХ КОНЕЧНЫХ АВТОМАТАХ

В. И. ЛЕВИН

(Рига)

(Поступила в редакцию 18 октября 1972 г.)

Предложена методика исследования переходных процессов в комбинационных конечных автоматах без памяти. Она основана на переносе задержек элементов на входы схемы, моделировании получившейся двухступенчатой схемы и последующем получении переходных процессов в многовходовых конъюнкторах и дизъюнкторе.

1. Введение

В теории конечных автоматов принято считать, что изменение состояния автомата осуществляется мгновенно [1]. В то же время на практике переключение технических устройств происходит в течение конечного времени и образует некоторый переходный процесс. Поэтому изучение переходных процессов в конечных автоматах представляет большой практический интерес. С другой стороны, такое изучение интересно и для самой теории автоматов, поскольку оно приводит к совершенно новым представлениям и задачам.

Впервые переходные процессы в автоматах стали рассматриваться американскими учеными [2—4]. При этом объектом изучения были не сами процессы, а некоторые их нежелательные последствия (состязания), определяющие надежность автомата. И лишь сравнительно недавно стали изучаться собственно переходные процессы [5]. При этом обнаружились трудности, связанные с отсутствием специализированного математического аппарата. В данной работе предлагается один из возможных путей преодоления этих трудностей.

2. Постановка задачи

Ограничимся рассмотрением конечных автоматов без памяти, заданных в виде структурных схем без обратных связей. Логические элементы схем (по предпо ложению) работают в двузначном алфавите, реализуя некоторую булеву функцию от входных сигналов и задерживая выходной сигнал на определенное время. При этом как число входов элемента, так и его функция могут быть произвольными. В рассматриваемых схемах есть и специальный эле-

мент задержки, функция которого — сдвиг во времени момента изменения сигнала ($0 \to 1$ или $1 \to 0$) на постоянную величину T без изменения формы сигнала.

На каждый вход схемы поступает процесс, складывающийся из постоянных начального и конечного значений (1 или 0) и последовательности мгновенных изменений сигнала ($1 \to 0$ и $0 \to 1$) между этими значениями. Эти воздействия вызывают на выходах всех элементов схемы процессы изменения сигнала; их будем называть переходными процессами в элементах. Совокупность переходных процессов в элементах схемы автомата назовем переходным процессом в автомате.

Математической моделью реального логического элемента считаем последовательное соединение идеального логического элемента (реализующего некоторую булеву функцию от мгновенных значений входных сигналов) с элементом задержки. Отсюда ясно, что переходные процессы в элементах схемы имеют тот же вид, что и воздействия на входах схемы. В качестве переходного процесса в произвольном логическом элементе будем указывать процесс на выходе его идеальной части, поскольку последующая паразитная задержка не искажает формы процесса, а лишь сдвигает его по времени.[1]

Задача состоит в том, чтобы научиться выражать в аналитической форме переходные процессы в логических элементах и схемах. При этом в случае схемы нас будет интересовать, как правило, только выходной процесс.

Некоторые обозначения.

1 — постоянный сигнал 1,	$1'_\alpha$	— $1'$ в момент α,	
0 — постоянный сигнал 0,	$0'_\beta$	— $0'$ в момент β,	
$1'$ — изменение сигнала $0 \to 1$,	$1(\alpha, \beta)$	— импульс	$1'_\alpha\ 0'_\beta$,
$0'$ — изменение сигнала $1 \to 0$,	$0(\alpha, \beta)$	— пауза	$0'_\alpha\ 1'_\beta$.

Процессы на входах схемы и в её элементах будем записывать в виде последовательности скачков (изменений сигнала) с указанием момента скачка, либо в виде последовательности импульсов и пауз. В последнем случае для экономии опускаются начальное и оконечное постоянные значения сигнала, а моменты промежуточных изменений сигнала указываются только один раз. Например, $y(t) = 1(\alpha, \beta)\, 0\,(-, \sigma)$ означает процесс, в котором до момента α (начало процесса) сигнал имеет постоянное значение 0, в интервале (α, β)

[1] Принятая модель реального логического элемента не учитывает его фильтрующих свойств. Возможность такого учета с помощью используемого в данной работе математического аппарата выяснена в [8].

равен 1, в интервале (β, σ) снова равен 0 и после момента σ (окончание процесса) принимает постоянное значение 1.

В качестве математического аппарата используем бесконечнозначную логику [6—7]. Возможность этого была впервые выяснена в работе [8]. Необходимые нам операции бесконечнозначной логики определяются следующим образом

$$\alpha_1 \vee \alpha_2 = \max(\alpha_1, \alpha_2), \qquad \alpha_1, \alpha_2 \in A, \tag{2.1}$$

$$\alpha_1 \alpha_2 \quad = \min(\alpha_1, \alpha_2), \qquad \alpha_1, \alpha_2 \in A, \tag{2.2}$$

где A — произвольное бесконечное множество, для элементов которого определены отношения «больше», «меньше», «равно». Операции (2.1) и (2.2) называются дизъюнкцией и конъюнкцией и обобщают соответствующие операции в двузначной логике, где $A = \{0,1\}$.

Основные законы бесконечнозначной логики совпадают с соответствующими в двузначной логике:

$$\alpha \vee \alpha = \alpha, \qquad \alpha\alpha = \alpha; \tag{2.3}$$

$$(\alpha \vee \beta) \vee \sigma = \alpha \vee (\beta \vee \sigma), \qquad (\alpha\beta)\sigma = \alpha(\beta\sigma); \tag{2.4}$$

$$\alpha \vee \beta = \beta \vee \alpha \quad , \qquad \alpha\beta = \beta\alpha; \tag{2.5}$$

$$\alpha(\beta \vee \sigma) = \alpha\beta \vee \alpha\sigma, \qquad \alpha \vee \beta\sigma = (\alpha \vee \beta)(\alpha \vee \sigma); \tag{2.6}$$

$$\overline{\alpha \vee \beta} = \bar{\alpha}\bar{\beta} \quad , \qquad \overline{\alpha\beta} = \bar{\alpha} \vee \bar{\beta}; \tag{2.7}$$

$$\alpha \vee \alpha\beta = \alpha \quad , \qquad \alpha(\alpha \vee \beta) = \alpha. \tag{2.8}$$

Для нас важны также следующие законы:

$$\alpha\beta < \alpha, \beta, \qquad \alpha \vee \beta \geq \alpha, \beta; \tag{2.9}$$

$$\alpha_1 \ldots \alpha_{i-1} \alpha_i \alpha_{1+1} \ldots \alpha_n = \alpha_1 \ldots \alpha_{i-1} \alpha_{i+1} \ldots \alpha_n \text{ если } \alpha_i \geq \text{какого-то } \alpha_k \tag{2.10}$$

$$\alpha_1 \vee \ldots \vee \alpha_{i-1} \alpha_i \vee \alpha_{i+1} \vee \ldots \vee \alpha_n = \alpha_1 \vee \ldots \vee \alpha_{i-1} \vee \alpha_{i+1} \vee \ldots \vee \alpha_n,$$
$$\text{если } \alpha_i \leq \text{какого-то } \alpha_k. \tag{2.11}$$

3. Переходные процессы в элементах

Переходные процессы в одновходовых элементах (повторитель, инвертор и элемент задержки) выражаются через соответствующие входные воздействия непосредственно. Поэтому перейдем к многовходовым элементам.

Прежде всего, заметим, что задержка τ сигнала на выходе элемента эквивалентна задержкам на такую же величину сигналов на всех входах. Поэтому для нахождения переходного процесса в реальном элементе достаточно вычислить переходный процесс, возникающий в соответствующем

идеальном элементе при входных воздействиях, полученных из исходных воздействий сдвигом по времени на τ.

Идеальный логический элемент всегда может быть представлен в виде эквивалентной логической схемы, содержащей лишь инверторы и двухвходовые дизъюнкторы и конъюнкторы. Справедливость этого утверждения в случае, когда рассматриваемый элемент — многовходовой дизъюнктор или конъюнктор, вытекает из возможности группировки членов в дизъюнкции и конъюнкции:

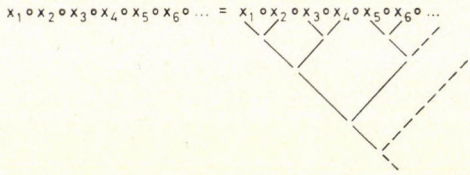


Рис. 1

Здесь $\circ$ — символ $\vee$ или $\wedge$. То же самое утверждение в общем случае (когда рассматриваемый многовходовой элемент реализует произвольную булеву функцию) следует из представимости произвольной булевой функции в дизъюнктивной нормальной форме [1] и формулы (3.1). Таким образом, отыскание переходного процесса в произвольном многовходовом логическом элементе сводится к нахождению переходного процесса в эквивалентной логической схеме из инверторов и двухвходовых дизъюнкторов и конъюнкторов. Последнее можно выполнить методом подстановок, т. е. последовательно подставляя найденные процессы на выходах одних элементов в качестве входных воздействий на последующие элементы схемы. При этом необходимо считаться с эффектом дробления, заключающимся в том, что число скачков сигнала на выходе элемента может быть больше аналогичных чисел на его входах. Иными словами, при расчете схемы большой длины приходится находить реакции элементов на входные воздействия с большим числом скачков, даже если числа скачков во входных воздействиях схемы малы.

Проблема отыскания реакции двухвходового дизъюнктора или конъюнктора на входные воздействия с произвольным конечным числом скачков была решена в работе [8], где для этой цели был предложен метод декомпозиции, сущность которого состоит в следующем. Один из двух входных процессов элемента $x(t)$ или $y(t)$ (например, $x(t)$) делится точкой скачка на два последовательных процесса $x_1(t)$ и $x_2(t)$. Находятся реакции $f_1(t)$ и $f_2(t)$ элемента на входные воздействия $\{x_1(t), y(t)\}$ и $\{x_2(t), y(t)\}$ (частичные реакции). Если $f_1(t)$ и $f_2(t)$ не пересекаются во времени (т. е. $f_1(t)$ оканчивается не позднее, чем $f_2(t)$ начинается), то реакция $f(t)$ элемента на первоначальное входное воз-

действие $\{x(t), y(t)\}$ определится, очевидно, как последовательность $f_1(t)$ и $f_2(t)$. Можно показать [9], что указанное непересечение имеет место, если процесс $x_1(t)$ оканчивается: а) в случае конъюнктора — точкой вида $0'_\beta$, б) в случае дизъюнктора — точкой вида $1'_\beta$. Метод декомпозиции позволяет заменить отыскание переходного процесса в элементе с заданными входными воздействиями отысканием двух подпроцессов, отвечающих более простым воздействиям. Последовательное применение этого метода сводит задачу к расчету реакции элемента на простейшие входные воздействия, что выполнимо непосредственно в терминах бесконечнозначной логики.

Пример 1. Найдем переходный процесс $f(t) = 0'_\alpha\, 1'_\beta \wedge 1'_\gamma\, 0'_\delta$, образующийся на выходе конъюнктора при подаче на его входы процессов $x(t) = 0'_\alpha\, 1'_\beta$ и $y(t) = 1'_\gamma\, 0'_\delta$.

Разделим процесс $x(t)$ на процессы $x_1(t) = 0'_\alpha$ и $x_2(t) = 1'_\beta$. Тогда частичные реакции конъюнктора определяются как

$$f_1(t) = 0'_\alpha \wedge 1'_\gamma\, 0'_\delta\,, \qquad f_2(t) = 1'_\beta \wedge 1'_\gamma\, 0'_\delta\,,$$

Найдем реакцию $f_1(t)$. Метод декомпозиции здесь уже не применим, так как единственная имеющаяся «потенциальная» точка деления $(1'_\gamma)$ не имеет требуемого вида $0'_\beta$. Однако, ввиду малого числа скачков во входных воздействиях, $f_1(t)$ можно вычислить непосредственно. Действительно, рассматривая три возможных положения скачка $0'_\alpha$ относительно импульса $1'_\gamma\ 0'_\delta$, можно записать

$$f_1(t) = \begin{cases} 1(\gamma, \alpha), & \gamma \leq \alpha \leq \delta\,, \\ 1(\gamma, \delta), & \gamma < \delta \leq \alpha\,, \\ 0 = 1\,(\alpha, \alpha)\,, & \alpha \leq \gamma < \delta\,. \end{cases}$$

Введя единую запись в терминах бесконечнозначной логики, получаем $f_1(t) = 1(\alpha\gamma, \alpha\delta)$. Аналогично находим $f_2(t) = 1(\beta \vee \gamma,\ \beta \vee \delta)$. Поскольку $\alpha\delta \leq \alpha$, а $\beta \vee \gamma \geq \beta > \alpha$, то процессы $f_1(t)$ и $f_2(t)$ не пересекаются во времени, так что искомый переходный процесс $f(t)$ определится как последовательность процессов $f_1(t)$ и $f_2(t)$. Учитывая, что между непересекающимися импульсами $f_1(t)$ и $f_2(t)$ должна быть пауза, находим

$$f(t) \equiv 0'_\alpha\, 1'_\beta \wedge 1'_\gamma\, 0'_\delta = 1\,(\alpha\gamma, \alpha\delta)\, 0(—, \beta \vee \gamma)\, 1\,(—, \beta \vee \delta)\,.$$

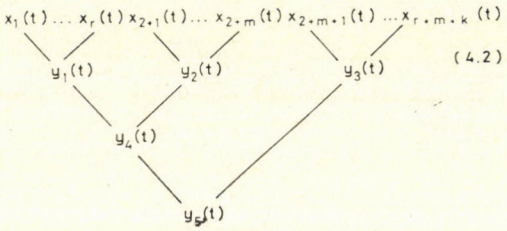
В заключение раздела отметим, что хотя сведение многовходового логического элемента к эквивалентной схеме из одно- и двухвходовых элементов всегда позволяет найти его переходный процесс, такое сведение не всегда выгодно с точки зрения объема вычислений. Часто оказывается целесообразнее сводить многовходовой элемент к эквивалентной схеме из многовходовых же (но с меньшим числом входов) элементов (разделы 4,6).

4. Переходные процессы в многовходовых конъюнкторах, дизъюнкторах и элементах И—НЕ, ИЛИ—НЕ

Рассмотрим n-входовой конъюнктор, воздействие на входах которого содержат не более одного скачка каждый и заданы в виде

$$\begin{aligned}
&x_1(t) &&= 1'_{\alpha_1}, \ldots, x_r(t) &&= 1'_{\alpha_r} && \text{(1-ая группа)},\\
&x_{r+1}(t) &&= 0'_{\beta_1}, \ldots, x_{r+m}(t) &&= 0'_{\beta_m} && \text{(2-ая группа)},\\
&x_{r+m+1}(t) &&= 0\ , \ldots, x_{r+m+k}(t) &&= 0 && \text{(3-ая группа)},\\
&x_{r+m+k+1}(t) &&= 1\ , \ldots, x_{r+m+k+p=n}(t) &&= 1 && \text{(4-ая группа)},
\end{aligned} \tag{4.1}$$

Поскольку $1 \wedge x = x$, входы с воздействиями 4-й группы можно исключить из дальнейшего рассмотрения. Конъюнктор с оставшимися $r + m + k$-входами представим следующей эквивалентной трехступенчатой схемой из двух двухвходовых и трех многовходовых конъюнкторов



(4.2)

Рис. 2

Процессы на выходах элементов первой ступени выражаются очевидными формулами

$$y_1(t) = 1'_{\alpha_1} \wedge \ldots \wedge 1'_{\alpha_r} = 1'_{\alpha_1 \vee \ldots \vee \alpha_r}, \tag{4.3}$$

$$y_2(t) = 0'_{\beta_1} \wedge \ldots \wedge 0'_{\beta_m} = 0'_{\beta_1 \vee \ldots \vee \beta_m}, \tag{4.4}$$

$$y_3(t) = \overbrace{0 \wedge \ldots \wedge 0}^{k} = \begin{cases} 0, & k \geq 1, \\ 1, & k = 0. \end{cases} \tag{4.5}$$

Процесс на выходе второй ступени $y_4(t) = y_1(t) \wedge y_2(t)$ после подстановки выражений (4.3), (4.4) и использования формулы (32) из [8] будет иметь следующее представление:

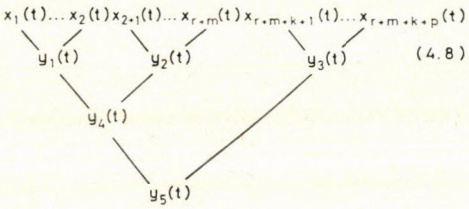
$$y_4(t) = 1'_{\alpha_1 \vee \ldots \vee \alpha_r} \wedge 0'_{\beta_1 \ldots \beta_m} = \begin{cases} 1(\alpha_1 \vee \ldots \vee \alpha_r, \alpha_1 \vee \ldots \vee \alpha_r \vee \beta_1 \ldots \beta_m), & m \geq 1, \quad r \geq 1, \\ 1'_{\alpha_1 \vee \ldots \vee \alpha_r}, & r \geq 1, \quad m = 0, \\ 0'_{\beta_1 \ldots \beta_m}, & r = 0, \quad m \geq 1. \end{cases} \tag{4.6}$$

Процесс на выходе третьей ступени $y_5(t) = y_3(t) \wedge y_4(t)$ и представляет собой переходный процесс $z(t)$ в рассматриваемом n-входовом конъюнкторе. Используя (4.5), (4.6), находим

$$Z(t) = 1'_{\alpha_1} \wedge \ldots \wedge 1'_{\alpha_r} \wedge 0'_{\beta_1} \wedge \ldots \wedge 0'_{\beta_m} \wedge \overbrace{0 \wedge \ldots \wedge 0}^{k} \wedge \overbrace{1 \wedge \ldots \wedge 1}^{p=n-r-m-k} =$$

$$= \begin{cases} 1(\alpha_1 \vee \ldots \vee \alpha_r, \alpha_1 \vee \ldots \vee \alpha_r \vee \beta_1 \ldots \beta_m), & k = 0, \ r \geq 1, \ m \geq 1, \\ 1'_{\alpha_1 \vee \ldots \vee \alpha_r}, & k = 0, \ r \geq 1. \ m \geq 0, \\ 0'_{\beta_1 \ldots \beta_m}, & k = 0, \ r = 0, \ m \geq 1, \\ 0, & k \geq 1. \end{cases} \tag{4.7}$$

Теперь рассмотрим n-входовой дизъюнктор с входными воздействиями (4.1). Поскольку $0 \vee x = x$, входы с воздействиями 3-й группы можно исключить из рассмотрения. Эквивалентная схема дизъюнктора с оставшимися $n-k$-входами имеет три ступени и содержит два двухвходовых и три многовходовых дизъюнктора



(4.8)

Рис. 3

Очевидны процессы на выходах первой ступени:

$$y_1(t) = 1'_{\alpha_1} \vee \ldots \vee 1'_{\alpha_r} = 1'_{\alpha_1 \ldots \alpha_r}, \tag{4.9}$$

$$y_3(t) = \overbrace{1 \vee \ldots \vee 1}^{p} = \begin{cases} 1, & p \geq 1, \\ 0, & p = 0. \end{cases} \tag{4.10}$$

$$y_2(t) = 0'_{\beta_1} \vee \ldots \vee 0'_{\beta_n} = 0'_{\beta_1 \vee \ldots \vee \beta_m}. \tag{4.11}$$

Процесс на входе второй ступени $y_4(t) = y_1(t) \vee y_2(t)$ с использованием выражений (4.9), (4.10) и формулы (32) из [8] получается в виде

$$y_4(t) = 1'_{\alpha_1 \ldots \alpha_r} \vee 0'_{\beta_1 \vee \ldots \vee \beta_m} = \begin{cases} 0(\beta_1 \vee \ldots \vee \beta_m, \beta_1 \vee \ldots \vee \beta_m \vee \alpha_1 \ldots \alpha_r), & r \geq 1, \ m \geq 1, \\ 1'_{\alpha_1 \ldots \alpha_r}, & r \geq 1, \ m = 0, \\ 0'_{\beta_1 \vee \ldots \vee \beta_m}, & r = 0, \ m \geq 1, \end{cases} \tag{4.12}$$

Процесс на выходе третьей ступени $y_5(t) = y_3(t) \vee y_4(t)$ является переходным процессом $z(t)$ в нашем n-входовом дизъюнкторе. С помощью (4.11) и (4.12) находим

$$Z(t) = 1_{\alpha_1} \vee \ldots \vee 1'_{\alpha_r} \vee 0'_{\beta_1} \vee \ldots \vee 0'_{\beta_m} \vee \overbrace{0 \vee \ldots \vee 0}^{k} \vee \overbrace{1 \vee \ldots \vee 1}^{p=n-r-m-k} =$$

$$= \begin{cases} 0(\beta_1 \vee \ldots \vee \beta_m, \beta_1 \vee \ldots \vee \beta_m \vee \alpha_1 \ldots \alpha_r), & p = 0,\ r \geq 1,\ m \geq 1, \\ 1'_{\alpha_1 \ldots \alpha_r} & , \quad p = 0,\ r \geq 1,\ m = 0, \\ 0'_{\beta_1 \vee \ldots \vee \beta_m} & , \quad p = 0,\ r = 0,\ m \geq 1, \\ 1 & , \quad p \geq 1 \end{cases} \tag{4.13}$$

Легко понять, что переходный процесс в n-входовом элементе И—НЕ при входных воздействиях (4.1) получается инвертированием (т. е. заменами $1 \rightleftarrows 0$, $1'_\alpha \rightleftarrows 0'_\alpha$) соответствующего этим воздействиям процесса (4.7) в n-входовом элементе И. Аналогично переходный процесс в n-входовом элементе ИЛИ—НЕ при входных воздействиях (4.1) получается инвертированием процесса (4.13) от воздействий (4.1) в n-входовом элементе ИЛИ.

5. Определение переходного процесса в логической схеме

Для нахождения переходного процесса на выходе произвольной логической схемы (входные воздействия которой содержат не более одного скачка каждый) предлагается следующий алгоритм.

1-й шаг. Заменяем задержку на выходе каждого элемента такими же по величине задержками на всех его входах [5]. Эти замены производятся при последовательном движении от выхода схемы к её входам с суммированием последовательно соединенных задержек. В результате, если в исходной схеме не было разветвлений, все задержки оказываются перенесенными на входы схемы. Пусть имеется разветвление канала i с сигналом $y_i(t)$ на m каналов, содержащих задержки $\tau_1, \ldots, \tau_m$. Прямое перенесение этих задержек в канал i возможно лишь в случае $\tau_1 = \tau_2 = \ldots = \tau_m$. Поэтому в общем случае подсхему, вырабатывающую $y_i(t)$, надо заменить m её копиями и перенести на выход первой задержку τ_1, на выход второй — задержку τ_2 и т. д.

В результате выполнения 1-го шага из исходной схемы получается эквивалентная ей, содержащая идеальную логическую схему с задержками только на входах. Если в исходной схеме не было разветвлений, то полученная идеальная схема совпадает с исходной, в которой исключены задержки. Если же имелись разветвления, то идеальная схема отличается от исходной повторением некоторых подсхем, входы которых являются входами всей исходной схемы. При этом задержки, полученные на соответственных вхо-

дах повторяющихся подсхем, неодинаковы. Таким образом, в эквивалентной схеме каждому входу исходной схемы с сигналом $x_i(t)$ в общем случае соответствует несколько входов, сигналы в которых получены из $x_i(t)$ различными временными сдвигами.

Преобразование процесса задержкой очевидно. Поэтому после выполнения первого шага задача сводится к нахождению переходного процесса в идеальной логической схеме, реализующей некоторую булеву функцию $f(x_1, \ldots, x_n)$.

2-ой шаг. Представляем функцию $f(x_1, \ldots, x_n)$ в дизъюнктивной нормальной форме. Соответственно этому заменяем идеальную логическую схему, полученную на 1-ом шаге, эквивалентной трехступенчатой схемой (ступени: 1) инверторов, 2) конъюнкторов, 3) дизъюнкторов). Поскольку преобразование процесса инвертором очевидно, можно считать схему состоящей из ступени конъюнкторов и заключительного дизъюнктора.

3-й шаг. Поскольку ни задержки, ни инверторы не меняют числа скачков преобразуемого процесса, то воздействия на входах ступени конъюнкторов содержат не более одного скачка каждый. Поэтому процессы на выходах конъюнкторов вычисляем по формуле (4.7). Процесс на выходе заключительного дизъюнктора (являющийся искомым переходным процессом) вычисляем по формуле (4.13), если выходы конъюнкторов (т. е. входы дизъюнктора) несут процессы не более чем с одним скачком каждый. В общем случае эти процессы могут содержать до двух скачков каждый [формула (4.7)]. Поэтому в общем случае искомый переходный процесс может находиться универсальным методом разложения многовходового дизъюнктора на двухвходовые (раздел 3). Однако более эффективным в этом случае оказывается специальный метод итерации (раздел 6).

6. Итерационный метод вычисления переходного процесса в многовходовом дизъюнкторе

Рассмотрим идеальный дизъюнктор, входные воздействия которого содержат не свыше двух скачков каждый, точнее, имеют один из видов: $1'_\alpha$, $0'_\alpha$, $1'_\alpha\, 0'_\beta$ (еще одно теоретически возможное воздействие $0'_\alpha\, 1'_\beta$ в силу формулы (4.7) никогда не появляется на входах заключительного дизъюнктора двухступенчатой схемы). Число входов s дизъюнктора считаем переменным. Покажем, каким образом, зная переходный процесс в s-входовом дизъюнкторе, найти соответствующий процесс в $(s+1)$-входовом дизъюнкторе.

Пусть имеется некоторый процесс

$$x(t) = a(\alpha_1, \alpha_2)\, a(-, \alpha_\beta) \ldots \tilde{a}(-, \alpha_m)\,, \tag{6.1}$$

где $\tilde{a}$ означает a или $\bar{a}$. Сдвинутый влево до β процесс $x(t)$ определяется следующим образом через конъюнкции в бесконечнозначной логике

$$x(t\beta) = a(\alpha_1 \beta, \alpha_2 \beta)\, \bar{a}(-, \alpha_3 \beta) \ldots \tilde{a}(-, \alpha_m \beta)\,, \tag{6.2}$$

а сдвинутый вправо до β процесс $x(t)$ — через дизъюнкции в этой же логике

$$x(t \vee \beta) = a(\alpha_1 \vee \beta, \alpha_2 \vee \beta)\, \bar{a}(-, \alpha_3 \vee \beta) \ldots \tilde{a}(-, \alpha_m \vee \beta)\,. \tag{6.3}$$

Достроенный справа до γ процесс $x(t)$ определяется таким образом

$$x_\gamma(t) = \begin{cases} x(t) & , \quad \tilde{a} = 0, \\ x(t)\, 0(\alpha_m, \gamma), & \tilde{a} = 1, \quad \gamma > \alpha_m\,. \end{cases} \tag{6.4}$$

Достроенный слева до γ процесс $x(t)$ определяется как

$$_\gamma x(t) = \begin{cases} x(t) & , \quad a = 0, \\ 0(\gamma, \alpha_1)\, x(t), & a = 1, \quad \gamma < \alpha_1\,. \end{cases} \tag{6.5}$$

Если над процессом выполняются обе операции (сдвиг и достройка), то первой условимся выполнять достройку, а второй — сдвиг.

Предложение 1. Пусть $\alpha^{(s)}(t)$ — переходный процесс в s-входовом дизъюнкторе с произвольными входными воздействиями, заданный в форме (6.1). Добавим к дизъюнктору $(s+1)$-й вход с воздействием $1'_\alpha$. Тогда переходный процесс $x^{(s+1)}(t)$ в получившемся $(s+1)$-входовом дизъюнкторе определяется формулой

$$x^{(s+1)}(t) = x_\gamma^{(s)}(t\alpha)\,, \tag{6.6}$$

в которой γ — произвольная постоянная, удовлетворяющая условию $\gamma \geq \alpha$.

Доказательство. Представим $(s+1)$-входовой дизъюнктор двухступенчатой схемой (в 1-й ступени — s-входовой, во 2-ой — 2-входовой дизъюнкторы). Тогда можно записать

$$x^{(s+1)}(t) = x^{(s)}(t) \vee 1'_\alpha = \begin{cases} x^{(s)}(t)\,, & t \leq \alpha\,, \\ 1 \quad , & t > \alpha\,. \end{cases} \tag{6.7}$$

Рассмотрим теперь три случая:

1) $\alpha < \alpha_1$. При этом $\alpha_1 \alpha = \alpha_2 \alpha = \ldots = \alpha_m \alpha = \alpha$, так что все скачки процесса $x^{(s)}(t\alpha)$ совместятся в один скачок в точке α. Отсюда

$$x^{(s)}(t\alpha) = \begin{cases} \bar{a} = x^{(s)}(t)\,, & t \leq \alpha, \\ \tilde{a} \quad , & t > \alpha. \end{cases} \tag{6.8}$$

Если $a = 0$, то с помощью (6.8) находим

$$x_\gamma^{(s)}(t\alpha) = x_{(s)}(t\alpha) = \begin{cases} x^{(s)}(t), & t \leq \alpha, \\ 1 \quad , & t > \alpha. \end{cases} \tag{6.9}$$

Сравнив (6.7) и (6.9), убеждаемся в справедливости (6.6).

Если же $\tilde{a} = 1$, то

$$x_\gamma^{(s)}(t\alpha) = x^{(s)}(t\alpha)\, 0(\alpha_m \alpha, \gamma\alpha). \tag{6.10}$$

По условию, $\alpha < \alpha_m$, $\alpha \leq \gamma$, так что $\alpha_m \alpha = \gamma\alpha = \alpha$, пауза 0 (.) вырождается в скачок $1'_\alpha$, и мы получаем из (6.10), используя (6.8),

$$x_\gamma^{(s)}(t\alpha) = \begin{cases} x^{(s)}(t\alpha), & t \leq \alpha \\ 1 \quad , & t > \alpha \end{cases} = \begin{cases} x^{(s)}(t), & t \leq \alpha, \\ 1 \quad , & t > \alpha. \end{cases} \tag{6.11}$$

Сравнив (6.7) и (6.11), снова убеждаемся в справедливости (6.6).

2) $\alpha > \alpha_m$. При этом $x^{(s)}(t\alpha) = x^{(s)}(t)$. Если $\tilde{a} = 0$, то, с одной стороны,

$$x_\gamma^{(s)}(t\alpha) = x^{(s)}(t\alpha) = x^{(s)}(t), \tag{6.12}$$

с другой стороны, из (6.7) с учетом того, что при $\tilde{a} = 0$, $\alpha > \alpha_m$ и $t > \alpha$, справедливо $x^{(s)}(t) = 1$, получаем

$$x^{(s+1)}(t) = x^{(s)}(t). \tag{6.13}$$

Теперь сравнение (6.12) с (6.13) дает требуемую формулу (6.6).

Если $\tilde{a} = 1$, то справедливо (6.10). По условию, $\alpha_m < \alpha \leq \gamma$. Следовательно, $\alpha_m \alpha = \alpha_m$, $\gamma\alpha = \alpha$, и $x^{(s)}(t\alpha) = x^{(s)}(t)$, так что (6.10) переходит в

$$x_\gamma^{(s)}(t\alpha) = x^{(s)}(t)\, 0(\alpha_m, \alpha) = \begin{cases} x^{(s)}(t), & t \leq \alpha, \\ 1 \quad , & t > \alpha. \end{cases} \tag{6.14}$$

Сравнив (6.7) и (6.14), снова получим требуемое.

3) $\alpha_i \leq \alpha \leq \alpha_{i+1}$. Запишем

$$x^{(s)}(t\alpha) = a(\alpha_1 \alpha, \alpha_2 \alpha)\, \bar{a}(-, \alpha_3 \alpha) \ldots a(-, \alpha_m \alpha).$$

В рассматриваемом случае

$$\alpha_1 \alpha = \alpha_1, \ldots, \alpha_i \alpha = \alpha_i; \quad \alpha_{i+1} \alpha = \ldots = \alpha_m \alpha = \alpha. \tag{6.15}$$

Следовательно,

$$x^{(s)}(t\alpha) = \begin{cases} x^{(s)}(t), & t \leq \alpha, \\ - \quad , & t > \alpha. \end{cases} \tag{6.16}$$

Теперь, если $\tilde{a} = 0$, то

$$x_\gamma^{(s)}(t\alpha) = x^{(s)}(t\alpha) = \begin{cases} x^{(s)}(t), & t \leq \alpha, \\ 1 \quad , & t > \alpha. \end{cases} \tag{6.17}$$

Сравнив (6.7) с (6.17), получим (6.6).

Если $\tilde{a} = 1$, то справедливо (6.10). В силу того, что $\gamma \geq \alpha$ и (6.15), имеем $\alpha_m \alpha = \gamma\alpha = \alpha$, и пауза 0 (.) в (6.10) вырождается в скачок $1'_\alpha$. Учитывая еще (6.16), находим

$$x_\gamma^{(s)}(t\alpha) = \begin{cases} x^{(s)}(t), & t \leq \alpha, \\ 1 \quad , & t > \alpha. \end{cases} \tag{6.18}$$

Сравнение (6.7) с (6.18) снова дает (6.6).

Предложение 2. Пусть $x^{(s)}(t)$ — переходный процесс в s-входовом дизъюнкторе с произвольными входными воздействиями, заданный в форме (6.1). Добавим к дизъюнктору $(s+1)$-й вход с воздействием $0'_\alpha$. Тогда переходный процесс в получившемся $(s+1)$-входовом дизъюнкторе определяется формулой

$$x^{(s+1)}(t) = {}_\gamma x^{(s)}(t \vee \alpha), \tag{6.19}$$

в которой γ — произвольная постоянная, удовлетворяющая условию $\gamma \leq \alpha$.

Доказательство этого предложения следует плану доказательства предложения 1. Как и выше, представляем $(s+1)$-входовой дизъюнктор двухступенчатой схемой и получаем

$$x^{(s+1)}(t) = x^{(s)}(t) \vee 0'_\alpha = \begin{cases} 1 \quad , & t \leq \alpha, \\ x^{(s)}(t), & t > \alpha. \end{cases} \tag{6.20}$$

После этого рассматриваем те же три случая:

1) $\alpha < \alpha_1$. При этом также $\alpha < \alpha_2, \ldots, \alpha < \alpha_m$. Отсюда

$$x^{(s)}(t \vee \alpha) = x^{(s)}(t). \tag{6.21}$$

Теперь, если $a = 0$, то

$${}_\gamma x^{(s)}(t \vee \alpha) = x^{(s)}(t \vee \alpha) = x^{(s)}(t). \tag{6.22}$$

Но из (6.20) следует (с учетом того, что при $a = 0$, $\alpha < \alpha_1$ и $t < \alpha$ справедливо $x^{(s)}(t) = 1$):

$$x^{(s+1)}(t) = x^{(s)}(t). \tag{6.23}$$

Сравнив (6.22) с (6.23), получаем требуемую формулу (6.19).

Если же $a = 1$, то

$$_{\gamma}x^{(s)}(t \vee \alpha) = 0(\gamma \vee \alpha, \alpha_1 \vee \alpha)\, x^{(s)}(t \vee \alpha). \tag{6.24}$$

По условию, $\alpha_1 > \alpha \geq \gamma$, поэтому $\gamma \vee \alpha = \alpha$, $\alpha_1 \vee \alpha = \alpha_1$, пауза 0 (.) превращается в 0 (α, α_1), и получаем из (6.24) с помощью (6.21)

$$_{\gamma}x^{(s)}(t \vee \alpha) = \begin{cases} 1 & , \quad t \leq \alpha, \\ x^{(s)}(t \vee \alpha), & t > \alpha, \end{cases} = \begin{cases} 1 & , \quad t \leq \alpha, \\ x^{(s)}(t), & t > \alpha. \end{cases} \tag{6.25}$$

Сравнение (6.20) с (6.25) дает (6.19).

2) $\alpha > \alpha_m$. В этом случае $\alpha > \alpha_1, \ldots, \alpha > \alpha_m$, так что $\alpha_1 \vee \alpha = \ldots = \alpha_m \vee \alpha = \alpha$, и потому все скачки процесса $x^{(s)}(t \vee \alpha)$ совместятся в один скачок в точке α. Отсюда

$$x^{(s)}(t \vee \alpha) = \begin{cases} \bar{a} & , \quad t \leq \alpha, \\ \tilde{a} = x^{(s)}(t), & t > \alpha. \end{cases} \tag{6.26}$$

Если $a = 0$, то с помощью (6.26) находим

$$_{\gamma}x^{(s)}(t \vee \alpha) = x^{(s)}(t \vee \alpha) = \begin{cases} 1 & , \quad t \leq \alpha, \\ x^{(s)}(t), & t > \alpha. \end{cases} \tag{6.27}$$

Сравнение (6.20) с (6.27) снова дает (6.19).

Если $a = 1$, то справедливо (6.24) (по условию $\alpha > \alpha_1$, $\alpha \geq \gamma$). Поэтому $\gamma \vee \alpha = \alpha_1 \vee \alpha = \alpha$, пауза 0 (.) в (6.24) вырождается в скачок $0'_{\alpha}$ и, используя (6.26), находим

$$_{\gamma}x^{(s)}(t \vee \alpha) = \begin{cases} 1 & , \quad t \leq \alpha, \\ x^{(s)}(t \vee \alpha), & t > \alpha, \end{cases} = \begin{cases} 1 & , \quad t \leq \alpha, \\ x^{(s)}(t), & t > \alpha. \end{cases} \tag{6.28}$$

Сравнив (6.20) с (6.28), опять убеждаемся в справедливости (6.19).

3) $\alpha_i \leq \alpha \leq \alpha_{i+1}$. Запишем

$$x^{(s)}(t \vee \alpha) = a(\alpha_1 \vee \alpha, \alpha_2 \vee \alpha)\, \bar{a}(—, \alpha_3 \vee \alpha) \ldots \tilde{a}(—, \alpha_m \vee \alpha).$$

В данном случае

$$\alpha_1 \vee \alpha = \ldots = \alpha_i \vee \alpha = \alpha; \alpha_{i+1} \vee \alpha = \alpha_{i+1}, \ldots, \alpha_m \vee \alpha = \alpha_m. \tag{6.29}$$

Следовательно,

$$x^{(s)}(t \vee \alpha) = \begin{cases} \bar{a} & , \quad t \leq \alpha, \\ x^{(s)}(t). & t > \alpha. \end{cases} \tag{6.30}$$

Теперь, если $a = 0$, то

$$_{\gamma}x^{(s)}(t \vee \alpha) = x^{(s)}(t \vee \alpha) = \begin{cases} 1 \quad , & t \leq \alpha , \\ x^{(s)}(t) , & t > \alpha . \end{cases} \tag{6.31}$$

Сравнивая (6.20) с (6.31), получаем (6.19).

Если $a = 1$, то имеет место (6.24). Поскольку по условию $\alpha_1 \leq \alpha$ $\gamma \leq \alpha$, то $\gamma \vee \alpha = \alpha_1 \vee \alpha = \alpha$, пауза 0 (.) в (6.24) вырождается в скачок $0'_\alpha$ и, используя (6.30), получаем

$$_{\gamma}x^{(s)}(t \vee \alpha) = \begin{cases} 1 \quad , & t \leq \alpha , \\ x^{(s)}(t \vee \alpha) , & t > \alpha . \end{cases} = \begin{cases} 1 \quad , & t \leq \alpha , \\ x^{(s)}(t) , & t > \alpha . \end{cases} \tag{6.32}$$

Сравнение (6.20) с (6.32) снова дает (6.19).

Предложение 3. Пусть $x^{(s)}(t)$ — переходный процесс в s-входовом дизъюнкторе с произвольными входными воздействиями. Добавим к дизъюнктору $(s + 1)$-й вход с воздействием $1'_\alpha 0'_\beta$. Тогда переходный процесс $x^{(s+1)}(t)$ в получившемся $(s + 1)$-входовом дизъюнкторе определяется в виде последовательности двух подпроцессов:

$$x^{(s+1)}(t) = x^{(s)}_{\gamma_1}(t\alpha)\, _{\gamma_2}x^{(s)}(t \vee \beta) , \tag{6.33}$$

где γ_1, γ_2 — произвольные постоянные, удовлетворяющие условиям

$$\alpha \leq \gamma_1 , \qquad \gamma_2 \leq \beta . \tag{6.34}$$

Доказательство. Представив $(s + 1)$-входовой дизъюнктор двухступенчатой схемой, можно записать

$$x^{(s+1)}(t) = x^{(s)}(t) \vee 1'_\alpha 0'_\beta . \tag{6.35}$$

Применим метод декомпозиции (раздел 3). Разделим процесс $1'_\alpha\, 0'_\beta$ на процессы $1'_\alpha$ и $0'_\beta$. Тогда частичные реакции $(s + 1)$ -входового дизъюнктора находятся при помощи предложения 1:

$$x_1^{(s+1)}(t) = x^{(s)}(t) \vee 1'_\alpha = x^{(s)}_{\gamma_1}(t\alpha) \qquad \gamma_1 \geq \alpha \tag{6.36}$$

и предложения 2

$$x_2^{(s+1)}(t) = x^{(s)}(t) \vee 0'_\beta = {}_{\gamma_2}x^{(s)}(t \vee \beta) , \qquad \gamma_2 \geq \beta . \tag{6.37}$$

Из (6.36) видно, что $x_1^{(s+1)}$ оканчивается не позже момента α, а из (6.37) — что $x_2^{(s+1)}(t)$ начинается не раньше момента β. Но $\alpha < \beta$ (иначе заданное воздействие $1'_\alpha 0'_\beta$ не существует). Поэтому первая частичная реакция оканчивается до начала второй. Отсюда переходный процесс в $(s + 1)$-входовом дизъюнкторе можно представить в виде последовательности его частичных реакций

$$x^{(s+1)}(t) = x_1^{(s+1)}(t)\, x_2^{(s+1)}(t) . \tag{6.38}$$

Подставив сюда выражения (6.36), (6.37), получим (6.33).

7. Заключение

Трудоемкость изложенной методики определяется, в первую очередь, числом входных наборов, обращающих реализуемую схемой функцию в единицу, и в гораздо меньшей степени — длиной схемы. Методика позволяет анализировать вручную схемы с числом входов до 5—7 (при практически неограниченной длине схемы) и может быть легко машинизирована. Она также полезна при качественном изучении переходных процессов в логических схемах (сбои, риски, состязания и т. д.).

Литература

1. *Глушков В. М.* Синтез цифровых автоматов. Физматгиз, М., 1962.
2. *Huffman, D. A.* The design and of hazard-free switching networks. J. Assoc. Comput. Mashinery **4** (1957) 1.
3. *Mecluskey, E.*, Transients in combinatorial logic circuits, from redundancy techniques for computing system. Spartan Book Co, 1962.
4. *Eichelberger, E. B.* Hazard detection in combinational and sequential switching circuits. IBHI. Res. and Development. **9** (1965) 2.
5. *Рогинский В. Н.* Переходные процессы в релейных устройствах (дискретных автоматах). Сб. «Сети передачи информации и их автоматизация». «Наука», М., 1965.
6. *Mc. Naughton, R.* A theorem about infinite valued sentential logic. J. Symb. Logic **16** (1951) 1. (Русский перевод: *Мак-Нотон Р.* Теорема о бесконечнозначной логике высказываний. Кибернетич. сб., вып. 3, ИЛ, М., 1961).
7. *Гинзбург С. А.*, Математическая непрерывная логика и изображение функций. «Энергия», М., 1968.
8. *Левин В. И.* Бесконечнозначная логика и переходные процессы в конечных автоматах. Автоматика и вычислительная техника (1972) 6.
9. *Левин В. И.* Декомпозиция переходных процессов в конечных автоматах. Автоматика и вычислительная техника (1973) 3.
10. *Anger, S. H.*, Hazards and Delays in Asynchronous Sequential Switching Circuits, IRE Trans. on Circuits Theory **6** (1959) 1.
11. *Моисил ГР. К.* Алгебраическая теория дискретных автоматических устройств. ИЛ, М., 1963.

Calculation of the transient processes in the combinational finite automata

V. I. LEVIN

(Riga)

Finite automata without memory are examined, which are given by the logical circuits without feedback, in which the synchronization of the elements' operation is absent. Transient processes are being studied, which appear as a result of switching the signals at the inputs of the circuit and which are determined by the said switching taking place not at the same time and by delays of the signals, when they pass through the elements. It is supposed that all the signals are rectangular. The logical element is simulated as serial coupling of the ideal (without delays) logical element with pure delay. The following methods for the analysis of the transient processes are proposed: 1. transference of the delays of all the elements to the inputs of the circuit; 2. representation of the resulting ideal circuit as the equivalent two-stage circuit (the stage of the many input AND plus final many input OR); 3. successive analysis of the processes at the

outputs of the first and the second stages. The general formulas are derived, which express the process taking place at the output of the element AND with the arbitrary number of inputs and as a result of the arbitrary switching of signals at these inputs. For the final OR of the two-stage realization the formulas are received, which permit to find the output process of the circuit with many inputs with the help of iteration. The proposed methods are based on the apparatus of the infinite-valued logic.

В. И. Левин
Институт электроники и вычислительной техники АН Латв. ССР
СССР, Рига 61,
ул. Академияс, 14

Problems of Control and Information Theory, Vol. 3 (4), pp. 277–284 (1974)

HOMOMORPHISMS OF STOCHASTIC AND NONDETERMINATE AUTOMATA

B. KWIATEK, R. ZIEBA

(Krakow)

(Received April 20, 1974)

The purpose of the present paper is to introduce stochastic and nondeterminate automata[1] and also their homomorphisms and their quasi-strong and strong homomorphisms.

The notions of stochastic and nondeterminate automata as well as their homomorphisms and their strong homomorphisms of [2] are particular cases of respective notions of this paper. Furthermore, properties of quasi-strong homomorphisms with their proofs are similar to those of strong homomorphisms of [1].

1. Stochastic automata and their homomorphisms

Definition 1. Let X, Y, Z be arbitrary nonempty sets and B a nonempty subset of $Z\times X$. For a fixed $[z, x]\in B$ we denote by $H[z, x]$ the function having discrete measures over $X\times Z$ as its values [2]. If $[z, x]\notin B$ the function H is not defined. The quadruple $\langle X, Y, Z, H\rangle$ is called a stochastic automaton or an S-automaton shortly.

For simplicity of notations we will write $H[z, x](y, z')$ instead of $H[z, x](\{[y, z']\})$.

Let $S^i = \langle X^i, Y^i, Z^i, H^i\rangle$ for $i = 1, 2, 3$ be S-automata, denoted in the sequel by S^i, and let ζ_i, η_i, ω_i for $i = 1, 2$ be arbitrary functions that map resp. X^i, Y^i, Z^i into X^{i+1}, Y^{i+1}, Z^{i+1}. If we are to consider a homomorphism of two S-automata we will write $[\zeta, \eta, \omega]$ instead of $[\zeta_1, \eta_1, \omega_1]$.

Definition 2. The triple $[\zeta, \eta, \omega]$ is said to be a homomorphism of S^1 into S^2, and we write $[\zeta, \eta, \omega]$: $S^1 \xrightarrow{h} S^2$, if $(H^1\,[z, x]$ defined $\Longrightarrow$ $\Longrightarrow \bigwedge_{[z,x]\in Z^1\times X^1} \bigwedge_{[\bar{y},\bar{z}]\in\eta(Y^1)\times\omega(Z^1)} \{H^2[\omega(z), \zeta(x)]$ defined and

$$H^2[\omega(z), \zeta(x)](\bar{y}, \bar{z}) = \sum_{\eta(y)=\bar{y},\ \omega(z')=\bar{z}} H^1[z,x]\,(y,z')\}).$$

Definition 3. The triple $[\zeta, \eta, \omega]$ is said to be a quasi-strong homomorphism of S^1 into S^2, and we write $[\zeta, \eta, \omega]$: $S^1 \xrightarrow{qsh} S^2$, if the following two conditions are fulfilled:

[1] Throughout this paper by an automaton we mean an arbitrary automaton which is not necessarily complete [1], i.e., an automaton for which we do not assume that function H (resp. h) is defined everywhere.

a) $[\zeta, \eta, \omega]\colon S^1 \xrightarrow{h} S^2$,

b) $\bigwedge_{[z',x'] \in \omega(Z^1)\times\zeta(X^1)} \bigvee_{[z,x] \in Z^1\times X^1} \{z' = \omega(z), x' = \zeta(x), (H^2[z', x'] \text{ defined} \Longrightarrow$
$\Longrightarrow H^1[z, x] \text{ defined})\}$.

Definition 4. The triple $[\zeta, \eta, \omega]$ is said to be a strong homomorphism of S^1 into S^2, and we write $[\zeta, \eta, \omega]\colon S^1 \xrightarrow{sh} S^2$, if the following two conditions are fulfilled:

a) $[\zeta, \eta, \omega]\colon S^1 \xrightarrow{qsh} S^2$,

b) $\bigwedge_{[z,x] \in Z^1\times X^1} \bigwedge_{[y,z'] \in Y^1\times Z^1} \{H^1[z, x] \text{ defined} \Longrightarrow (H^2[\omega(z), \zeta(x)]\,(\eta(y), \omega(z')) =$
$= H^1[z, x]\,(y, z'))\}$.

In the sequel we will need the following

Theorem 1. For each homomorphism $[\zeta, \eta, \omega]$ of S^1 into S^2 and whenever $H^1[z, x]$ is defined we have the following equality:

$$H^2[\omega(z), \zeta(x)]\,(\eta(Y^1)\times\omega(Z^1)) = 1\,.$$

Proof

$$H^2[\omega(z), \zeta(x)]\,(\eta(Y^1)\times\omega(Z^1)) = \sum_{[\bar{y},\bar{z}] \in \eta(Y^1)\times\omega(Z^1)} H^2[\omega)z),\zeta(x)]\,(\bar{y}, \bar{z}) =$$

$$= \sum_{[\bar{y},\bar{z}] \in \eta(Y^1)\times\omega(Z^1)} \sum_{\eta(y)=\bar{y},\ \omega(z')=\bar{z}} H^1[z, x]\,(y, z') = \sum_{[y,z'] \in Y^1\times Z^1} H^1[z, x]\,(y, z') = 1\,.$$

Corollary 1. If $[\zeta, \eta, \omega]$ is a homomorphism of S^1 into S^2 and whenever $H^1[z, x]$ is defined we have

$$H^2[\omega(z), \zeta(x)]\,(Y^2\times Z^2\setminus\eta(Y^1)\times\omega(Z^1)) = 0\,,$$

i.e. for the automaton S^2 the probability of the fact that either we get $y \notin \eta(Y^1)$ as an output or S^2 goes to a state $z^* \notin \omega(Z^1)$ is equal to zero.

Condition b) in Definition 4 may be replaced by an equivalent condition. Namely, we have the following

Theorem 2. For each strong homomorphism $[\zeta, \eta, \omega]$ of S^1 into S^2 and for all $z, z_1, z_2 \in Z^1$, $x \in X^1$, $y_1, y_2 \in Y^1$ such that card $(\{y_1, y_2\}\times\{z_1, z_2\}) \geq$ ≥ 2, $\omega(z_1) = \omega(z_2)$, $\eta(y_1) = \eta(y_2)$ and whenever $H^1[z, x]$ is defined:

The necessary and sufficient condition for the equality

$$H^2[\omega(z), \zeta(x)]\,(\eta(y_1), \omega(z_1)) = H^1[z, x](y_1, z_1)$$

is that

$$H^1[z, x]\,(\{y_1, y_2\}\times\{z_1, z_2\}) = 0\,.$$

Proof of the necessity. Suppose the contrary, i.e., that there exist $y_1, y_2 \in Y^1$, $z, z_1, z_2 \in Z^1$, $x \in X^1$ such that $y_1 \neq y_2$ and $H^1[z, x](y_1, z_1) > 0$. (If $y_1 = y_2$ and $z_1 \neq z_2$ the proof runs similarly.) For $i = 1, 2$ we have

$$H^1[z, x](y_2, z_i) = H^2[\omega(z), \zeta(x)](\eta(y_2), \omega(z_i)) = H^1[z, x](y_1, z_1) > 0$$

hence

$$1 = \sum_{[y,z'] \in Y^1 \times Z^1} H^1[z, y](y, z') = \sum_{[y,z'] \in Y^1 \times Z^1} H^2[\omega(z), \xi(x)](\eta(y), \omega(z')) >$$
$$> \sum_{[\bar{y},\bar{z}] \in \eta(Y^1) \times \omega(Z^1)} H^2[\omega(z), \zeta(x)](\bar{y}, \bar{z}) = 1 .$$

Proof of the sufficiency. We choose any pair $[\bar{y}, \bar{z}] \in \eta(Y^1) \times \omega(Z^1)$. Two cases are possible: either

1° $\{[y, z'] \in Y^1 \times Z^1 : \eta(y) = \bar{y}$ and $\omega(z') = \bar{z}\}$ is a one-element set or

2° $\{[y, z'] \in Y^1 \times Z^1 : \eta(y) = \bar{y}$ and $\omega(z') = \bar{z}\}$ is not a one-element set.

In the first case

$$H^2[\omega(z), \zeta(x)](\bar{y}, \bar{z}) = \sum_{\eta(y)=\bar{y}, \omega(z')=\bar{z}} H^1[z, x](y, z') = H^1[z, x](y, z')$$

and in the second one

$$H^2[\omega(z), \zeta(x)](\bar{y}, \bar{z}) = \sum_{\eta(y)=\bar{y}, \omega(z')=\bar{z}} H^1[z, x](y, z') = 0 ,$$

thus

$$H^2[\omega(z), \zeta(x)](\bar{y}, \bar{z}) = H^1[z, x](y, z') .$$

Now we will prove that the composition of two homomorphisms of S-automata is also a homomorphism.

Theorem 3. If $[\zeta_i, \eta_i, \omega_i]$ is a homomorphism of S^i into S^{i+1}, $i = 1, 2$, then $[\zeta_2\zeta_1, \eta_2\eta_1, \omega_2\omega_1]$ is a homomorphism of S^1 into S^3.

Proof. For each pair $(y'', z'') \in \eta_2\eta_1(Y^1) \times \omega_2\omega_1(Z^1)$ and whenever the function $H^1[z, x]$ is defined we have in view of Theorem 1 and Definition 2:

$$H^3[\omega_2\omega_1(z), \zeta_2\zeta_1(x)](y'', z'') = \sum_{\eta_2(y')=y'', \omega_2(z')=z''} H^2[\omega_1(z)\, \zeta_1(x)](y', z') =$$
$$= \sum_{\eta_2(y')=y'', \omega_2(z')=z''} \sum_{\eta_1(y)=y', \omega_1(z_1)=z'} H^1[z, x](y, z_1) = \sum_{\eta_2\eta_1(y)=y'', \omega_2\omega_1(z_1)=z''} H^1[z, x](y, z_1)$$

which completes the proof.

This is no longer true for quasi-strong homomorphisms. The following example illustrates this statement.

Example 1. Figure 1 represents the graphs of three S-automata.

$[I_x, I_y, \omega_1]$, where I_x and I_y are the identity maps, $\omega_1(a_i) = b_i$, $i = 1, 2, 3$, is a quasi-strong homomorphism of S^1 into S^2, and $[I_x, I_y, \omega_2]$ is a quasi-strong homomorphism of S^2 into S^3 whenever I_x and I_y are the identity maps, $\omega_2(b_i) = c_i$, $i = 1, 2, 3$, $\omega_2(b_4) = c_2$.

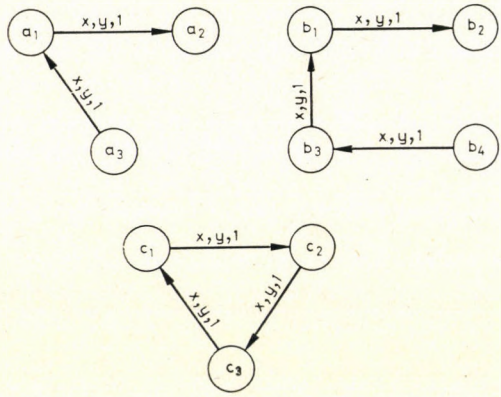


Fig. 1

The triple $[I_x, I_y, \omega_2\omega_1]$ is not a quasi-strong homomorphism of S^1 into S^3, because the pair $[c_2, x]$ for which H^3 is defined is the image of the unique pair $[a_2, x]$ for which H^1 is not defined.

We have, however, the following

Theorem 4. If $[\zeta_i, \eta_i, \omega_i]$ is a quasi-strong homomorphism of S^i into S^{i+1} for $i = 1, 2$, and either ζ_1 and ω_1 are surjective or ζ_2 and ω_2 are injective, then $[\zeta_2\zeta_1, \eta_2\eta_1, \omega_2\omega_1]$ is a quasi-strong homomorphism of S^1 into S^3.

Proof in the case where ζ_1 and ω_1 are surjective. Condition a) of Definition 3 is fulfilled.

For any pair $[z_3, x_3] \in \omega_2\omega_1(Z^1) \times \zeta_2\zeta_1(X^1)$ there is a pair $[z_2, x_2]$ such that $[z_2, x_2] \in Z^2 \times X^2$, $z_3 = \omega_2(z_2)$, $x_3 = \zeta_2(x_2)$ and whenever $H^3[z_3, x_3]$ is defined $H^2[z_2, x_2]$ is also defined. For the pair $[z_2, x_2]$ there is such a pair $[z_1, x_1] \in \in Z^1 \times X^1$ for which: $z_2 = \omega_1(z_1)$, $x_2 = \zeta_1(x_1)$ and $H^1[z_1, x_1]$ is defined, thus Condition b) of Definition 3 is fulfilled.

If ζ_2 and ω_2 are injective then Condition a) of Definition 3 is fulfilled. To prove Condition b) of Definition 3 we choose any pair $[z_3, x_3] \in \omega_2\omega_1(Z^1) \times \times \zeta_2\zeta_1(X^1)$ for which $H^3[z_3, x_3]$ is defined. There is exactly one pair

$$[z_2, x_2] \in \omega_1(Z^1) \times \zeta_1(X^1) \tag{1}$$

such that we have:

$$\omega_2(z_2) = z_3,\ \zeta_2(x_2) = x_3 \text{ and } H^2[z_2, x_2] \text{ is defined.}$$

By (1) and since $[\zeta_1, \eta_1, \omega_1]$ is the quasi-strong homomorphism, we get the existence of a pair $[z_1, x_1] \in Z^1 \times X^1$ such that we have:

$$\omega_1(z_1) = z_2,\ \zeta_1(x_1) = x_2 \text{ and } H^1[z_1, x_1] \text{ is defined.}$$

Example 1 shows also that the composition of two strong homomorphisms need not be a strong homomorphism of S-automata. We give now a sufficient condition for this composition to be a strong homomorphism.

Theorem 5. If $[\zeta_i, \eta_i, \omega_i]$ is a strong homomorphism of S^i into S^{i+1} for $i = 1, 2$, and either ζ_1 and ω_1 are surjective or ζ_2 and ω_2 are injective, then $[\zeta_2\zeta_1, \eta_2\eta_1, \omega_2\omega_1]$ is a strong homomorphism.

Proof. The hypotheses of Theorem 4 are fulfilled, therefore the composition $[\zeta_2\zeta_1, \eta_2\eta_1, \omega_2\omega_1]$ is a quasi-strong homomorphism. Since for all $z, z' \in Z^1$, $x \in X^1$, $y \in Y^1$ for which $H^1[z, x]$ is defined we have:

$$H^1[z, x](y, z') = H^2[\omega_1(z), \zeta_1(x)](\eta_1(y), \omega_1(z')) =$$
$$= H^3[\omega_2\,\omega_1(z), \zeta_2\,\zeta_1(x)]\,(\eta_2\,\eta_1(y), \omega_2\,\omega_1(z'))\,,$$

so the proof is completed.

2. Nondeterminate automata and their homomorphisms

Definition 5. Let X, Y, Z be arbitrary nonempty sets and B a nonempty subset of $Z \times X$. We denote by h the function defined on B whose values lie in a set $2^{Y \times Z} \setminus \{\varnothing\}$. The quadruple $\langle X, Y, Z, h\rangle$ is called a nondeterminate automaton or an N-automaton shortly.

Each S-automaton generates a certain N-automaton if we accept the following

Definition 6. For any automaton $S = \langle X, Y, Z, H\rangle$

$$N(H) \stackrel{\text{def}}{=} \begin{cases} \{[y, z'] : H[z, x]\,(y, z') > 0\} & \text{if } H[z, x] \text{ defined,} \\ \text{not defined} & \text{otherwise.} \end{cases}$$

This definition and Definition 5 imply the following

Corollary 2. $\langle X, Y, Z, N(H)\rangle$ is an N-automaton. The automaton $N(S) = \langle X, Y, Z, N(H)\rangle$ is called an N-automaton generated by the automaton $S = \langle X, Y, Z, H\rangle$.

Let $N^i = \langle X^i, Y^i, Z^i, h^i \rangle$ for $i = 1, 2, 3$ be N-automata, denoted in the sequel by N^i, and let $\zeta_i, \eta_i, \omega_i$ for $i = 1, 2$ be arbitrary functions that map resp. X^i, Y^i, Z^i into $X^{i+1}, Y^{i+1}, Z^{i+1}$. If we are to consider a homomorphism of two N-automata we will write $[\zeta, \eta, \omega]$ instead of $[\zeta_1, \eta_1, \omega_1]$.

Definition 7. The triple $[\zeta, \eta, \omega]$ is said to be a homomorphism of N^1 into N^2, and we write $[\zeta, \eta, \omega] : N^1 \xrightarrow{h} N^2$, if

$$\bigwedge_{[z,x] \in Z^1 \times X^1} \bigwedge_{[\bar{y},\bar{z}] \in \eta(Y^1) \times \omega(Z^1)} (h^1[z, x] \text{ defined} \Longrightarrow \{h^2[\omega(z), \zeta(x)] \text{ defined and}$$

$$h^2[\omega(z), \zeta(x)] = \{[y', z''] \in Y^2 \times Z^2 : \bigvee_{y \in Y^1} \bigvee_{z^* \in Z^1} ([y, z^*] \in h^1[z, x] \text{ and}$$

$$\eta(y) = y' \text{ and } \omega(z^*) = z'')\}\}).$$

Definition 8. The triple $[\zeta, \eta, \omega]$ is said to be a quasi-strong homomorphism of N^1 into N^2, and we write $[\zeta^4, \omega] : N^1 \xrightarrow{qsh} N^2$, if the following two conditions are fulfilled:

a) $[\zeta, \eta, \omega] : N^1 \xrightarrow{h} N^2$,

b) $$\bigwedge_{[z',x'] \in \omega(Z^1) \times \zeta(X^1)} \bigvee_{[z,x] \in Z^1 \times X^1} (z' = \omega(z) \text{ and } x' = \zeta(x)$$

and $[h^2[z', x'] \text{ defined} \Longrightarrow h^1[z, x] \text{defined}])$.

Definition 9. The triple $[\zeta, \eta, \omega]$ is said to be a strong homomorphism of N^1 into N^2, and we write $[\xi, \eta, \omega] : N^1 \xrightarrow{sh} N^2$, if the following two conditions are fulfilled:

a) $[\zeta, \eta, \omega] : N^1 \xrightarrow{qsh} N^2$,

b) $$\bigwedge_{[z,x] \in Z^1 \times X^1} \bigvee_{[y,z'] \in Y^1 \times Z^1} (h^1[z, x] \text{ defined} \Rightarrow$$

$$\Rightarrow \{[y, z'] \in h^1[z, x] \Leftrightarrow [\eta(y), \omega(z')] \in h^2[\omega(z), \zeta(x)]\}).$$

It follows directly from Definition 7

Theorem 6. If $[\zeta_i, \eta_i, \omega_i]$ is a homomorphism of N^i into N^{i+1} for $i = 1, 2$, then $[\zeta_2\zeta_1, \eta_2\eta_1, \omega_2\omega_1]$ is a homomorphism of N^1 into N^3.

This is no longer true for quasi-strong homomorphisms as shows the following example.

Example 2. Let $\zeta_i, \eta_i, \omega_i$ be the functions of Example 1, and let $N(S^i)$ be N-automata generated by automata S^i of that example; $[\zeta_i, \eta_i, \omega_i]$ is a quasi-strong homomorphism (a strong homomorphism) of $N(S^i)$ into $N(S^{i+1})$

for $i = 1, 2$, and the composition $[\zeta_2\zeta_1, \eta_2\eta_1, \omega_2\omega_1]$ is not a quasi-strong homomorphism (a strong homomorphism) of $N(S^1)$ into $N(S^3)$.

We have however, the following

Theorem 7. If $[\zeta_i, \eta_i, \omega_i]$ is a quasi-strong homomorphism of N^i into N^{i+1} for $i = 1, 2$, and either ζ_1 and ω_1 are surjective or ζ_2 and ω_2 are injective, then $[\zeta_2\zeta_1, \eta_2\eta_1, \omega_2\omega_1]$ is a quasi-strong homomorphism of N^1 into N^3.

Proof in the case where ζ_1 and ω_1 are surjective. On account of Theorem 6 we infer that the Condition a) of Definition 8 is fulfilled. To prove Condition b) of Definition 8 we choose any pair $[z_3, x_3] \in \omega_2\omega_1(Z^1) \times \zeta_2\zeta_1(X^1)$ for which h^3 is defined. For this pair there is a pair $[z_2, x_2] \in Z^2 \times X^2$ such that we have:

$$z_3 = \omega_2(z_2),\ x_3 = \zeta_2(x_2) \text{ and } h^2[z_2, x_2] \text{ is defined.}$$

For the pair $[z_2, x_2]$ there is a pair $[z_1, x_1] \in Z^1 \times X^1$ for which:

$$z_2 = \omega_1(z_1),\ x_2 = \zeta_1(x_1) \text{ and } h^1[z_1, x_1] \text{ is defined.}$$

Consequently $[\zeta_2\zeta_1, \eta_2\eta_1, \omega_2\omega_1]$ is a quasi-strong homomorphism.

The proof in the other case runs similarly.

For strong homomorphisms we have the following theorem that results directly from Definition 9 and Theorem 7.

Theorem 8. If $[\zeta_i, \eta_i, \omega_i]$ is a strong homomorphism of N^i into N^{i+1} for $i = 1, 2$, and either ζ_1 and ω_1 are surjective or ζ_2 and ω_2 are injective, then $[\zeta_2\zeta_1, \eta_2\eta_1, \omega_2\omega_1]$ is a strong homomorphism of N^1 into N^3.

Given any homomorphism (quasi-strong homomorphism, strong homomorphism) of $S^1 = \langle X^1, Y^1, Z^1, H^1\rangle$, into $S^2 = \langle X^2, Y^2, Z^2, H^2\rangle$. It is easy to show that $[\zeta, \eta, \omega]$ is a homomorphism (quasi-strong homomorphism, strong homomorphism) of $N(S^1)$ into $N(S^2)$.

References

1. *Gancarzewicz, J.*, Remarks on algebraic properties of sequential machines. Zesztyty Naukowe UJ, Prace Matematyczne **15** (1971) 47—49.
2. *Starke, P. H.*, Abstrakte Automaten. VEB Deutscher Verlag der Wissenschaften, Berlin, 1969.

Гомоморфизм стохастических и недетерминированных автоматов

Б. Квиатек, Р. Зиеба

(Краков)

В статье обобщаеся понятие сильной гомоморфии, введенное Ганцарзевицем в случае не вполне определенных стохастических и недетерминированных автоматов. Доказано, что гомоморфизм таких автоматов при опеределенных условиях образует полугруппу.

B. Kwiatek, R. Zieba
Wyzsza Szkoła Ekonomiczna
Krakow
ul. Rakowicka 27/310
Poland

BOOK REVIEW

EYKHOFF P.: System Identification. Parameter and State Estimation. John Wiley and Sons, London, New York, Sydney, Toronto, 1974

In the last years system parameter estimation and system state estimation became more and more important from theoretical as well as practical point of view. This development has not arisen only from the need for better system performance but the possibilities were also increased by the application of modern computing facilities.

This book aims at giving a unified, concise and systematic treatise of the main problems in parameter and state estimation.

The first two chapters serve as a foundation of the most important approches to the solution of estimation problems.

The part called Prolegomena summarizes the basic concepts about signals, dynamic systems, dynamic processes and models. Some topics of probability theory statics and optimization theory are also mentioned.

The part entitled Estimation of Parameters contains six chapters which are devoted to explicit (open-loop) methods as well as implicit methods, that is, model adjustment techniques (closed-loop methods), with sampled signals and/or continuous signals. Here linear and nonlinear problems are also treated. The parameter estimation is partly based on least-squares and Markov estimates, partly on maximum likelihood estimates and Bayes' estimation.

The next part is called Simultaneous Estimation of Parameters and States. The first chapter in this part reviews the state estimation problem starting from the Wiener filtering problem and turning later to the Kalman–Bucy filters. The second chapter is devoted to the combined parameter and state estimation techniques.

The last part of the book, Applications, gives a survey of estimation cases from various fields: physical, chemical processes; nuclear reactors; power plants and distribution systems; telecommunication systems; air and space vehicles, biological objects; socioeconomic systems, learning and pattern recognition systems are mentioned and discussed in some detail.

The Appendix contains e.g. a list of symbols, a summary of notions, summary of matrix calculus. Author's and Subject Index completes the book, consisting of 555 pages.

This book is a very good summary of the state of art and the state of science in the field of identification, parameter and state estimation. Further studies may be facilitated by references and additional literature being at the end of each chapter.

Cs. F.

J. CHINAL: Design Methods for Digital Systems. (Translated from French, Springer-Verlag, Berlin Heidelberg New York, Heidelberg 1973, 506 pp)

This book constitutes an introduction to the theory of binary switching networks (binary logic circuits) such as are encountered in industrial automatic systems, in communications networks and, more particularly, in digital computers.

In general, the book covers the exposition of the algebraic techniques necessary for the analysis and synthesis of logic networks. No previous knowledge of this field of activity is necessary: any technician or engineer possessing an elementary knowledge of mathematics and electronics can undertake its reading. Mathematicians will discover,

apart from recent algebraic theories, a large number of applications, mostly borrowed from the field of computers, and thus an exposition of the practical motivations of these theories, which constitute an ever more important aspect of applied mathematics. A large number of recent methods taken from technical literature are also discussed. An effort has been made to give some aspects to the presentation and to the proofs of the corresponding theorems and at the same time to integrate the results of diverse origins and of unequal theoretical levels into a single work. On the whole, no systematic effort has been made to give a strict mathematical rigour to the proofs. In every cases, however, the reader will find a reasonable justification of the results stated.

The first chapter is devoted to some reminders of the concepts of sets, of relations, functions and algebraic structure which prove to be useful in the following chapters.

In the second chapter the numeration systems, mainly the binary and related numeration systems, are described. A detailed discussion is given to the latter and it is shown how it can be manipulated. The chapter reviews the binary addition, substraction, multiplication and division. This is followed by the concepts of systems with bases 2^p ($p < 1$) and the conversion methods of different numeration systems are also given.

Chapter 3 treats the codes used in swithing systems, the relations between different codes, the choice of codes and their adaptation to digital operations.

Before dealing with the Boolean algebra in chapters 4 and 5 diverse algebrae have been introdced: the algebra of contacts (chapter 4), the algebra of classes and the algebra of logic (chapter 5). It is shown that there are striking resemblances between these three types of algebra. And this approach is a great help for the reader studying the Boolean algebra.

Chapter 6 describes the Boolean algebra and it is shown in a more precise manner how the three algebrae can be considered as particular cases. It makes systematically clear the differences between ordinary algebra and the Boolean algebra. A detailed discussion is given to the basic Boolean elements: set, operations, axioms and fundamental relations. This chapter includes the principle of duality and the Boolean variables in general cases (for n variables) too.

The next (seventh) chapter is devoted to the Boolean functions. The definitions of the individual functions are mostly illustrated by means of their truth tables. Other means of defining a Boolean function are introduced as the characteristic vector and the decimal representation. Beginning with two-variable functions the chapter goes into further details of n-variable functions and some distinguished functions are also described as the Pierce, Sheffer and threshold functions. In the final part of this chapter the determination of the canonical forms and some operations performed with them are also given.

Chapter 8 deals with the geometric representaton of Boolean functions and shows the method of simplification of algebraic expressions by the Karnaugh diagram method.

Chapter 9 gives a short survey of the physical representations of logic circuits (DCTL, DTL) and a number of examples (binary adder, substracter, parity checker etc.) illustrate the methods introduced in the preceeding chapters.

Chapter 10 treats first the simplification criterion introducing the cost function then goes into details regarding the method which can be used to simplify the networks. The algebraic method, the Karnaugh diagram method, the Quine—McCluskey method and the functional decomposition method are introduced. The latter two methods, because of their practical importance, are discussed in more detail. After their axiomatic discussion practical examples make them easy to understand and use.

In chapter 11 the sequential networks are introduced with the help of a number of examples while in chapter 12 a general model is described characterising the binary sequential networks (included synchronous and asynchronous networks). New representations are also introduced: the transition table, the graph of the sequential network and the sequence diagram.

Chapter 13 gives a short survey of regular expressions and regular events. The definitions and operations discussed here give the base to the usage of the method discussed in chapter 15.

Chapter 14 is devoted to the simplification of sequential networks and minimization of transition tables. The problem is posed in similar terms for the sequential networks as for the combinational ones. All the methods connected with the minimization (completely specified table, incompletely specified table, compatible pairs etc.) are illustrated by examples.

MTA Könyvtára
Periodikus 1076/19 75 sz.

Chapter 15 gives the synthesis of synchronous sequential networks. The direct method and the state diagram method are described. Examples make the theoretical considerations understandable. It covers the determination of the state diagram, the binary coding of the states and the determination of the control functions for the flip-flops and the output functions. Algorithms for the synthesis of a Moore and a Mealy machine are also given.

Chapter 16 is a short summary of the counters constituting a special type of sequen tial networks. Binary, decimal, up-down counters are described by examples.

This fundamental textbook may be helpful for anyone who is interested in design- and analysis of digital systems.

T. Gál

MAGYAR
TUDOMÁNYOS AKADÉMIA
KÖNYVTÁRA

Printed in Hungary
AKADÉMIAI NYOMDA, BUDAPEST

PROBLEMS OF CONTROL AND INFORMATION THEORY VOL. 3, No. 4
ENGLISH TRANSLATION OF THE COMMUNICATIONS IN RUSSIAN

UNIVERSAL ALGEBRAIC CODES FOR SYMMETRIC CHANNELS

V. D. GOPPA

(Vladivostok)

We introduce the idea of metrics according to Fitinghof and generalize the Hilbert boundary for this metrics. We proof the existence of universal (L, g) codes for symmetric, memoryless channels.

1. Let E^n be an n-dimensional vectorspace in $GF(q)$; $q = p^l$; p is a prime number. E^n is a linear code subspace. We examine such special channels for which algebraic codes are efficient (channels with independent additive noise). Every such channel represents the combination of two sources, the input and the noise, both of which independently produce vectors $x, y \in E^n$ according to distribution $p(x)$ (input), and distribution $p(y)$ (noise). We add these vectors by coordinates, obtaining at the output the vector $z = x + y$. For such channels the transfer matrix has a specific symmetric structure, thus for the sake of being short we call these channels symmetric, adding the expression "memory-less" when the noise distribution is polinomial.

2. Let S be a certain family of mamory-less discrete channels, $I_n(x, y)$ is the mutual information of channel $s \in S$.

$$C = \sup_{p \in P} \inf_{s \in S} I_s(x, y) \tag{1}$$

quantity, where the value of *sup* is taken at the input for the entire distribution, is called the transfer capacity of the channel class S. It is known that there exists a code, which is uniform for the whole class S, with a transfer rate R, which is arbitrarily close to C, and its error probability is arbitrarily close to zero [1]. For $R > C$ such a code does not exist.

Generally

$$C \leq \bar{C} = \inf_{s \in S} \sup_{p \in P} I_s(x, y) \tag{2}$$

but in the special case when S represents the set of all memory-less symmetric channels whose transfer capacity is not smaller than a certain quantity C_0, the following equation holds:

$$C = \bar{C} = C_0 \,. \tag{3}$$

Thus there exists a universal optimal code for memoryless symmetric channels. In the present paper we show that upon searching for such a code we can limit ourselves to the set of non-reducible (L, g) codes, a relatively narrow subclass of linear codes. Besides we introduce the idea of Fitinghof's metrics and generalize the Hilbert boundary for this metrics.

3. Let us suppose that in the vector $x_i \in n$ the letter $a \in E$ occurs m_i times $(i = 1, 2, \ldots, q)$. We call the quantity $\Phi(x) = H(m_1/n, \ldots . m_q/n)$ the Fitinghof weight of this vector. If $n = rv$, then we can divide the vector into r parts, identify every element v in the field E with an element of the new alphabet and consider the word x with length;n as word r with length z in the new alphabet. The quantity $\Phi(z)$ is called the Fitinhof v-weight of the vector x and is designated by $\Phi_v(x)$. It is clear that $\Phi(x) = \Phi_1(x)$.

Theorem 1. The Fitinghof weight;has the folowing characteristics:

1) In any E^n for stationary distribution

$$M\,[\Phi_v(x)] \leq H(E^v), \tag{4}$$

where M designates the expected value.

2) In any E^∞ for ergodic distribution the quantity $\Phi_v(x)/v$ almost everywhere converges to $H(E)$, if the value of v had beenc hosen thus that $n \to \infty \Rightarrow \Rightarrow v \to \infty$.

3) Let us denote by E_0 the subspace of those vectors $x \in E_v$ for which $\Phi_v(x) = 0$. $\Phi_v(x)$ forms a norm in the factor space E^n/E_0 and thus determines over E^n the metrics $\Phi_v(x - y)$ (Fitinghof metrics).

4) If $V_v(w)$ represents a sphere of radius w in the Fitinghof metrics, i. e. $V_v = \{x\colon \Phi_v(x)\ \ w\}$ and the value v had been chosen thus that $\log_q n - \log_q$ for $n \to \infty$, then

$$\log |V_v(w)| = n\big(w/v + 0(1)\big)\,. \tag{5}$$

Proof. The 1) characteristic follows from the convex nature of entropy. The 2) characteristic directly follows from the definitions. The 3) characteristic is less trivial and was not known earlier. We first of all remark that the vector derived from E_0 in the case $v = 1$ has the form: $(a, a, \ldots, a)$ thus E_0 is really a subspace, but at the same time v-dimensional. Let $x, y \in E^n$, $x = = (x_1, \ldots, x_n)$, $x = (y_1, \ldots, y_n)$. Let us introduce the vector $x \otimes y$, whose coordinates are the (x_i, y_i) pairs:

$$x \otimes y = \big((x_1\, y_1), \ldots, (x_n\, y_n)\big). \tag{6}$$

Let $\{m_i\}$, $\{n_j\}$; $i, j = 1, \ldots, q$ the compositions of the x and y vectors, $\{m_{ij}\}$ is the composition of vector $x \otimes y$ (in the $(E \times E)$ alphabet). It is evident that

$$\sum_i m_i = \sum_j m_j = \sum_{i,j} m_{ij} = n\,,$$

$$\sum_i m_{ij} = n_j\,,\ \sum_j m_{ij} = m_i\,.$$

As for any distribution in $X \times Y$

$$H(X \otimes Y) \otimes H(X) + H(Y)\,, \tag{7}$$

therefore

$$\Phi(x \otimes y) = \Phi(x) + \Phi(y)\,.$$

Further on $\Phi(x + y) \leq \Phi(x \otimes y)$ as the composition of vector $x + y$ is obtained from the $x \otimes y$ composition as a result of identifying certain (x_i, y_i) pairs (for example if $x_1 + y_1 = x_2 + y_2$, $x_1 \neq x_2$ then the coordinates (x_1, y_1) and (x_2, y_2) shall be considered different in $x \otimes y$ and identical in $x + y$).

Thus the triangle ingequality is valid:

$$\Phi_\nu(x + y) \leq \Phi_\nu(x) + \Phi_\nu(y)\,. \tag{8}$$

We have to remark that $\Phi_\nu(x)$ is a function of the boundary classes of E_0 :

$$x \equiv y \mod E_0 \Rightarrow \Phi_\nu(x) = \Phi_\nu(y)\,. \tag{9}$$

The 4) characteristic had been applied by Fitinghof [2] in a slightly different form. This is the result of the Stirling decomposition of the polinomial product $n!/m_1! \ldots mq!$ and was described for example in Fano's book [3]/for the case $\nu = 1$):

$$\log |V_1(w)| = n\big(w + 0(1)\big), \quad 0(1) = q \log n)/n\,. \tag{10}$$

In order to change to $|V_\nu(\omega)|$ we have to substitute q^ν for q and $r = n/\nu$ for n in this expression.

4. In coding theory usually codes of a great Hammingweight are built. This is true for the binary symmetric channel, but for the non-binary case, not to mention channels having memory, the Hamming metrics becomes insufficient. However codes having great Fitinghof weight are suitable for a wider class of channels.

Theorem 2. (the analogue of the Hilbert boundary). There exists a non-reducible (l, g) code having a minimal w value of Fitinghof weight, length n, whose transfer rate is

$$k/n > 1 - (\log_q 2)\ w/v - \varepsilon \tag{11}$$

for an arbitrary value $\varepsilon > 0$ and for all values of n which are sufficiently great.

MAGYAR TUDOMÁNYOS AKADÉMIA KÖNYVTÁRA

The proof is similar to the proof of the last theorem of article [4]. We remark that for the case of discrete symmetric code we can chose $\nu = 1$, then the derived inequality corresponds with the Hilbert boundary, as in this case $w = H(d/n)$, where d is the minimal weight of the Hamming code.

5. In [5] decoding by the Fitinghof weight had been introduced, where in any boundary class of the linear code the vector can be chosen with the smallest Fitinhof weight.

Theorem 3. Let $S(C_0)$ denote the class of all memoryless channels whose transfer capacity is not smaller than C_0. There exists a series of non-reducible (L, g) codes with a transmission rate;of $R = (\log q)\ (k/n)$ such that the error probability of decoding by the Fitinghof weight for any $C_0 > R$ from $S(C_0)$ uniformly approches zero upon increasing n for these codes.

Proof. We obtained by the random coding method proposed by the author [5] the following estimate for the error which was averaged for the entire class of non-reducible (L, g) codes upon the application of decoding by the Fitinghof weight:

$$\overline{p}_e \leq 2^{-n\left(\frac{C-R}{2}\right)} + p\,[\,|\Phi(x) - H(p)| > \frac{C-R}{2}\,. \tag{12}$$

Here $C = \log q - H(p)$, $H(p)$ is the entropy of noise, R is the transmission rate of the code. We can easily evaluate the rate of convergence $m_i/n_i \to p$ with the help of the cutset; of the distributions [3]:

$$P\,[\,|m_1/n - p_1| \leq \varepsilon, \ldots, |m_{q-1}/n - p_{q-1}| \leq \varepsilon] \geq 1 - 2^q \cdot 2^{-nE(\overline{p}, p)},$$

where

$$\overline{p}_i = p_i + \varepsilon,\ \ i = 1, 2, \ldots, q-1,\ \ \varepsilon < 1/q,$$

$$E(\overline{p}, p) = H(p) + \varepsilon H'(p) - H(\overline{p}) = H(p_1, \ldots, p_{q-2}) +$$

$$+ \varepsilon \sum_i \partial H \frac{(p_1, \ldots, p_{p-1})}{\partial p_i} - H(p_1 + \varepsilon, \ldots, p_{q-1} + \varepsilon)\,.$$

However if $|\,m_i/n - p_i\,| \leq \varepsilon < 1/2,\ i = 1, \ldots, q-1$, then

$$|H\,(m_1/n, \ldots, m_{q-1}/n) - H(p_1, \ldots, p_{q-1})| \leq q\ \varepsilon \log \varepsilon) \leq q\ \sqrt{\varepsilon}.\ ([1]).$$

Therefore

$$P\,[\,|\Phi(x) - H(p)| > \frac{c-R}{2} \leq 2^q \cdot 2^{-nE(p, p)}, \tag{13}$$

where

$$\overline{p}_i = p_i + \left(\frac{c-R}{2q}\right)^2,$$

thus we obtained a lower estimate for the exponent of error probability:

$$-\lim_{n\to\infty} \frac{1}{n} \log \overline{p}_e \geq E(\overline{p}, p)\,, \tag{14}$$

where the $E(\overline{p}, p)$ reliability function can be calculated

$$\varepsilon = \overline{p}_i - p_i = \left(\frac{c-R}{2q}\right)^2, \quad i = 1, \ldots, q-1. \tag{15}$$

The $E(\overline{p}, p)$ function is always greater than zero in the case $C - R > 0$, therefore for any $C_0 > R$:

$$\min_{S(C_0)} E(\overline{p}, p) > 0$$

(here the minimum is calculated for all the distributions from $S(C_0)$). Thus there exists a code which satisfies the conditions of the theorem.

6. Optimal coding theory for symmetric channels clearly expressed is not probabilistic by nature. More precisely, to build a code for such a channel which will satisfy the main theorem of Shannon, the probability distributions of noise need not be known, it capacity of the channel.

The product of the two $x \otimes y$ words which appeared at the proof of the first theorem makes it possible to introduce the non-probabilistic analogue of "mutual information" and to generalize the results obtained for arbitrary memory-less channels. We devote the next paper to this topic.

References

1. *Blackwell, D., Breiman, L., Thomasian, A. J.*, The capacity of a class of channels. Annals of Math. Stat., **30** (1959) **4**, 1229–1241.

2. *Фитингоф Б. М.* Оптимальное кодирование при неизвестной и меняющейся статистике сообщений. Проблемы передачи информации (1966) **2,** 3–11.

3. *Фано Р.* Передача информации. «Мир», М., 1965.

4. *Гоппа В. Д.* Рациональное представление кодов и (L, g)-коды. Проблемы передачи информации **7,** (1971) 41–49.

5. *Гоппа В. Д.* Универсальное декодирование для симметричных каналов. Проблемы передачи информации (1975) *1*.

COMPUTATION OF TRANSIENTS IN COMBINATIONAL FINITE AUTOMATA

V. I. LEVIN

(Riga)

A technique is suggested for studying transients in combinational memoriless finite automata by putting element delays at the circuit input and modelling the resulting circuit by a two-stage circuit with subsequent determination of transients in multi-input conjunctors and a disjunctor.

1. Introduction

Although the finite automata theory assumes that the automaton state changes instantly [1], in physical devices states change during a finite time interval and form a transient. That is why study of transients in finite automata is of great practical importance. On the other hand, this study is promising for the automata theory itself because it leads to absolutely new concepts and problems.

American scientists were the first to consider transients in automata [2—4]. In so doing they were studying not the processes but some of their undesirable consequences (hazards) deteriorating automaton reliability. And only quite recently transients proper were considered [5], but their study was handicapped by the lack of a special purpose mathematical apparatus. The present paper suggests one of possible ways to overcome these difficulties.

2. Problem formulation

Confine ourselves to memoriless finite automata given as structural diagrams without feed-backs. By assumption, the logic elements in circuits work with a two-valued alphabet and implement a Boolean function of inputs with certain delay of the output signal. Number of inputs and function implemented by the element may be arbitrary. The circuits under consideration contain also a special delay element shifting in time the moment of signal change ($0 \to 1$ or $1 \to 0$) by a constant value τ without distorting the signal form.

At each of its inputs the circuit receives a process consisting of constant initial and final values (1 or 0) and a sequence of instantaneous changes of the signal ($1 \to 0$ and $0 \to 1$) in between them. This leads to signal change processes at outputs of all circuit elements that will be referred to as transient processes (on transients). Their totality will be referred to as transients in the automaton.

A serial connection of the ideal logic element (implementing a Boolean function of momentary values of its inputs) with a delay is regarded as a mathematical model of physical logic element. Thus, it stands to reason that transients in circuit elements are of the same kind that actions at circuit inputs. Process at the output of the ideal part of an arbitrary logic element is regarded as its transient because the following parasite delay does not distort the process but only shifts it in time.[1]

The problem is to find an analytical representation of transients in logic elements and circuits. In the case of circuit we shall be interested, as a rule, only in the output process.

Some notations:

1 — constant signal 1,	$1'_\alpha \to 1'$ at moment α,
0 — constant signal 0,	$0'_\beta \to 0'$ at moment β,
$1'$ — signal change $0 \to 1$,	$1(\alpha, \beta)$ — pulse $1'_\alpha 0'_\beta$,
$0'$ — signal change $1 \to 0$.	$0(\alpha, \beta)$ — pause $0'_\alpha 1'_\beta$.

Processes at the inputs of a circuit and its elements will be represented as a sequence of signal changes with indication of the moments of the changes, or as a sequence of pulses and pauses. In the latter case, the initial and finite constant signal values are omitted, and the moments of intermediate signal changes are indicated, for the sake of brevity only once. For example, $y(t) = 1(\alpha, \beta)0(-, \sigma)$ means that a process where before moment α (beginning of the process) the signal had constant value 0, is 1 during interval (α, β), again 0 during (β, σ) and after moment σ (process completion) takes constant value 1.

As a mathematical apparatus the infinite valued logic is used [6, 7]. The possibility to do so was first established in [8]. The necessary operation of the infinite valued logic are defined as follows:

$$\alpha_1 \vee \alpha_2 = \max(\alpha_1, \alpha_2), \quad \alpha_1, \alpha_2 \in A \tag{2.1}$$

$$\alpha_1 \alpha_2 = \min(\alpha_1, \alpha_2), \quad \alpha_1, \alpha_2 \in A, \tag{2.2}$$

where A is an arbitrary infinite set for whose elements relations "greater", "less", "equal to" are defined. Operations (2.1) and (2.2) are called, respectively, disjunction and conjunction and generalize the appropriate two-valued logic operations where $A = \{0, 1\}$.

[1] The accepted model of the physical logic element does not account for its filtering properties. The possibility to take them into account by means of the mathematical apparatus used in the present paper was elucidated in [8].

The fundamental laws of the infinite valued logic coincide with those of the two-valued logic:

$$\alpha \vee \alpha = \alpha, \qquad \alpha\alpha = \alpha; \tag{2.3}$$

$$(\alpha \vee \beta) \vee \sigma = \alpha \vee (\beta \vee \delta), \qquad (\alpha\beta)\sigma = \alpha(\beta\sigma); \tag{2.4}$$

$$\alpha \vee \beta = \beta \vee \alpha \quad , \qquad \alpha\beta = \beta\alpha; \tag{2.5}$$

$$\alpha(\beta \vee \sigma) = \alpha\beta \vee \alpha\sigma, \qquad \alpha \vee \beta\sigma = (\alpha \vee \beta)(\alpha \vee \sigma); \tag{2.6}$$

$$\bar{\alpha} \vee \bar{\beta} = \overline{\alpha\beta} \quad , \qquad \overline{\alpha\beta} = \bar{\alpha} \vee \bar{\beta} \tag{2.7}$$

$$\alpha \vee \alpha\beta = \alpha \quad , \qquad \alpha(\alpha \vee \beta) = \alpha. \tag{2.8}$$

The following laws are also of importance to us:

$$\alpha\beta \leq \alpha, \beta, \qquad \alpha \vee \beta \geq \alpha, \beta \tag{2 9}$$

$$\alpha_1 \ldots \alpha_{i-1}\alpha_i\alpha_{i+1} \ldots \alpha_n = \alpha_1 \ldots \alpha_{i-1}\alpha_{i+1} \ldots \alpha_m \text{ if } \alpha_i \leqq \text{some } \alpha_k. \tag{2.10}$$

$$\alpha_1 \vee \ldots \vee \alpha_{i-1} \vee \alpha_i \vee \alpha_{i+1} \vee \ldots \vee \alpha_n = \alpha_1 \vee \ldots \vee \alpha_{i-1} \vee \alpha_{i+1} \vee \ldots \vee \alpha_n$$
$$\text{if } \alpha_i \leq \text{some } \alpha_k. \tag{2.11}$$

3. Transients in elements

Since transients in one-input elements (repeater, inverter and delay) are expressed directly through the appropriate inputs, go immediately to the multi-input elements.

First note that signal delay τ at element output is equivalent to the same delay of all inputs. Therefore, to find the transient in a physical element, it is sufficient to compute the transient in the corresponding ideal element for inputs obtained from the initial ones by shifting them in time by τ.

The ideal element may be always represented as an equivalent logic circuit consisting of inverters and two-input disjunctors and conjunctors only. This statement holds for the case where the element under consideration is a multi-input disjunctor or conjunctor because the terms may by grouped in sums and products: (Sec Fig. 1)

Here O stands for $\vee$ or $\wedge$. In the general case where the multi-input element under consideration implements and arbitrary Boolean function, the same follows from the representability of an arbitrary Boolean function in the normal disjunctive form [1] and from (3.1). Thus, determination of the transient in an arbitrary multi-input logic element boils down to determination of the transient in an equivalent logic circuit consisting of inverters and two-input disjunctors and conjunctors. This may be done by substituting processes de-

termined at outputs of some elements as inputs of subsequent ones. In so doing one should take into account that the number of signal changes at element output may be greater than that at its input. In other words, when dealing with long circuits one has to find out responses of elements to inputs with many changes even if the number of changes in input actions of the circuit is small.

Response of two-input disjunctor/conjunctor to inputs with an arbitrary finite number of changes was found in Ref. [8] by means of a decomposition method. It is essentially as follows: One of two element inputs $x(t)$ or $y(t)$ (e.g. $x(t)$) is divided by a point of change into two cascaded processes $x_1(t)$ and $x_2(t)$. Responses $f_1(t)$ and $f_2(t)$ to inputs $\{x_1(t), y(t)\}$ and $\{x_2(t), y(t)\}$ (partial responses) are determined. If $f_1(t)$ and $f_2(t)$ do not overlap in time (i.e $f_1(t)$ does not end later than $f_2(t)$ begins), element response $f(t)$ to the initial input $\{x(t), y(t)\}$ may be represented as a sequence of $f_1(t)$ and $f_2(t)$. It may be shown [9] that overlapping does not take place if $x_1(t)$ ends a) in the case of conjunctor, by a point of $0'_\beta$ kind, and b) in the case of disjunctor, by a point of $1'_\beta$ kind. The decomposition method enables one to use determination of two sub-transients responding to simpler actions instead of determination of a transient in an element with given inputs. Systematic applicaton of the method reduces the problem to finding element response to the simplest inputs, this being feasible directly in the terms of the infinite valued logic.

Example 1. Find transient at conjunctor output $f(t) = 0'_\alpha 1_\beta \wedge 1'_\gamma 0'_\delta$ under inputs $x(t) = 0'_\alpha 1'_\beta$ and $y(t) = 1'_\gamma 0'_\delta$.

Divide $x(t)$ into processes $x_1(t) = 0'_\alpha$ and $x_2(t) = 1'_\beta$. Then conjunctor's partial responses are defined as $f_1(t) = 0'_\alpha \wedge 1'_\gamma 0'_\delta$ and $f_2(t) = 1'_\beta \wedge 1'_\gamma 0'_\delta$. Find response $f_1(t)$. Decomposition is not applicable in this case because the only "potential" point of division $(1'_\gamma)$ is not of the required kind $0'_\beta$. However, $f_1(t)$ may be computed directly owing to the small number of changes in inputs. Indeed, by considering the three possible positions of change $0'_\alpha$ with respect to pulse $1'_\gamma 0'_\delta$, one may write the following:

$$f_2(t) = \begin{cases} 1(\gamma, \alpha), & \gamma \leq \alpha \leq \delta\,, \\ 1(\gamma, \delta), & \gamma < \delta \leq \alpha\,, \\ 0 = 1(\alpha, \alpha), & \alpha \leq \gamma < \delta\,. \end{cases}$$

By introducing a single writing in the terms of the infinite valued logic, obtain $f_1(t) = 1(\alpha\gamma, \alpha\delta)$. Similarly find $f_2(t) = 1(\beta \vee \gamma, \beta \vee \delta)$. Since $\alpha\delta \leq \alpha$ and $\beta \vee \gamma \geq \beta > \alpha$, processes $f_1(t)$ and $f_2(t)$ do not overlap in time, and the sought for transient $f(t)$ will be defined as a sequence of transients $f_1(t)$ and

$f_2(t)$. Taking into account that there should be a pause between two non-overlapping pulses $f_1(t)$ and $f_2(t)$, find that

$$f(t) = 0'_\alpha 1'_\beta \wedge 1'_\gamma 0'_\sigma = 1(\alpha\gamma, \alpha\delta)\, 0\,(-, \beta \vee \gamma)\, 1\,(-, \beta \vee \delta)\,.$$

To conclude this section note that although reduction of a multi-input logic element to an equivalent circuit of one- and two-input elements always enables one to find its transient, this reduction is not always resonable in the terms of amount of computation. It is more reasonable often to reduce a multi-input element to an equivalent circuit consisting of multi-input elements (with less number of inputs) (see Sections 4 and 6).

4. Transients in multi-input conjunctors, disjunctors, NANDs and NORs

Consider an n-input conjunctor whose inputs have each at most one change and are given in the following form:

$$\begin{aligned}
&x_1(t) &&= 1'_{\alpha 1}, \ldots, x_r(t) &&= 1'_{\alpha r} && \text{(1st group)} \\
&x_{r+1}(t) &&= 0'_{\beta 1}, \ldots, x_{r+m}(t) &&= 0_{\beta m} && \text{(2nd group)} \\
&x_{r+m+1}(t) &&= 0, \ldots, x_{r+m+k}(t) &&= 0 && \text{(3rd group)} \\
&x_{r+m+k+1}(t) &&= 1, \ldots, x_{r+m+k+p=n}(t) &&= 1 && \text{(4th group)}
\end{aligned} \tag{4.1}$$

Since $1 \wedge x = x$, inputs with actions of the 4th group may be excluded from consideration. Conjunctor with remaining $r + m + k$ inputs will be represented by an equivalent three-stage circuit of two two-input and three multi-input conjunctor: (Sec Fig. 2)

Outputs of the first stage elements are expressed by evident formulae:

$$y_1(l) = 1'_{\alpha_1} \wedge \ldots \wedge 1'_{\alpha r} = 1'_{\alpha_1 \vee \ldots \vee \alpha_r} \tag{4.3}$$

$$y_2(t) = 0'_{\beta_1} \wedge \ldots \wedge 0'_{\beta m} = 0'_{\beta_1 \ldots \beta m} \tag{4.4}$$

$$y_3(t) = \overbrace{0 \wedge \ldots \wedge 0}^{k} = \begin{cases} 0, & k \geq 1\,, \\ 1, & k = 0\,. \end{cases} \tag{4.5}$$

The output process of the second stage $y_4(t) = y_1(t) \wedge y_2(t)$ after substitution of (4.3) and (4.4) and using (32) of Ref. [8] will have the following representation

$$y_4(t) = 1'_{\alpha_1 \vee \ldots \vee \alpha_r} \wedge 0'_{\beta_1 \ldots \beta_m} = \begin{cases} 1(\alpha_1 \vee \ldots \vee \alpha_r, \alpha_1 \vee \ldots \vee \alpha_r \vee \beta_1 \ldots \beta_m), & m \geq 1,\ r \geq 1, \\ 1'_{\alpha_1 \vee \ldots \vee \alpha_r}\,, & \tau \geq 1, m = 0, \\ 0'_{\beta_1 \ldots \beta_m}\,, & \tau = 0, m \geq 1. \end{cases} \tag{4.6}$$

The process at the output of the third stage $y_5(t) = y_3(t) \wedge y_4(t)$ is the transient $z(t)$ in the n-input conjunctor under consideration. By using (4.5) and (4.6) find:

$$Z(t) = 1'_{\alpha_1} \wedge \ldots \wedge 1'_{\alpha_r} \wedge 0'_{\beta_1} \wedge \ldots \wedge 0'_{\beta_m} \wedge \overbrace{0 \wedge \ldots \wedge 0}^{k} \wedge \overbrace{1 \wedge \ldots \wedge 1}^{p=n-r-m-k} =$$

$$= \begin{cases} 1(\alpha_1 \vee \ldots \vee \alpha_r, \alpha_1 \vee \ldots \vee \alpha_r \vee \beta_1 \ldots \beta_m), & k = 0, \ r \geq 1, \ m \geq 1, \\ 1'_{\alpha_1 \vee \ldots \vee \alpha_r} \quad , & k = 0, \ r \geq 1, \ m = 1, \\ 0'_{\beta_1 \ldots \beta_m} \quad , & k = 0, \ r = 0, \ m \geq 1, \\ 0 \quad , & k \geq 1. \end{cases} \tag{4.7}$$

Consider now an n-input disjunctor with inputs (4.1). Since $0 \vee x = x$ inputs with actions of the third group may be ruled out of consideration. The equivalent disjunctor circuit with remaining $n - k$ inputs has three stages and contains two two-input and three multi-input disjunctors
(Sec Fig. 3)

Processes at the outputs of the first stage are evident:

$$y_1(t) = 1'_{\alpha_1} \vee \ldots \vee 1'_{\alpha_\tau} = 1'_{\alpha_1 \ldots \alpha_\tau}, \tag{4.9}$$

$$y_3(t) = \overbrace{1 \vee \ldots \vee 1}^{k} = \begin{cases} 1, & p \geq 1, \\ 0, & p = 0, \end{cases} \tag{4.10}$$

$$y_2(t) = 0'_{\beta_1} \vee \ldots \vee 0'_{\beta_m} = 0'_{\beta_1 \vee \ldots \vee \beta_m}. \tag{4.11}$$

The process at the input of the second stage $y_4(t) = y_1(t) \vee y_2(t)$ after application of (4.9) and (4.10), and (32) of [8] takes the following form

$$y_4(t) = 1'_{\alpha_1 \ldots \alpha_r} \vee 0_{\beta_1 \vee \ldots \vee \beta_m} = \begin{cases} 0(\beta_1 \vee \ldots \vee \beta_m, \beta_1 \vee \ldots \vee \beta_m \vee \alpha^2 \ldots \alpha_\tau), & r \geq 1, \ m \geq 1, \\ 1'_{\alpha_1 \ldots \alpha_r} \quad , & r \geq 1, \ m = 0 \\ 0'_{\beta_1 \vee \ldots \vee \beta_m}, & r = 0, \ m \geq 1. \end{cases} \tag{4.12}$$

The output process of the third stage $y_5(t) = y_3(t) \vee y_4(t)$ is the transient $z(t)$ of our n-input disjunctor. By means of (4.11) and (4.12) find that

$$Z(t) = 1'_{\alpha_1} \vee \ldots \vee 1'_{\alpha_r} \vee 0'_{\beta_1} \vee \ldots 0'_{\beta_m} \vee \overbrace{0 \vee \ldots \vee 0}^{k} \vee \overbrace{1 \vee \ldots \vee 1}^{p=n-r-m-k} =$$

$$= \begin{cases} 0(\beta_1 \vee \ldots \vee \beta_m, \beta_1 \vee \ldots \vee \beta_m \vee \alpha_1 \ldots \alpha_r), & p = 0, \ r \geq 1, \ m \geq 1, \\ 1'_{\alpha_1 \ldots \alpha_r} \quad , & p = 0, \ r \geq 1, \ m = 0, \\ 0'_{\beta_1 \vee \ldots \vee \beta_m} \quad , & p = 0, \ r = 0, \ m \leq 1, \\ 1 \quad , & p \geq 1. \end{cases} \tag{4.13}$$

It is easy to see that the transient in n-input NAND under actions (4.1) is obtained through inversion of the corresponding process (4.7) in n-input AND (i.e. by interchanges $1 \rightleftarrows 0$, $1'_\alpha \rightleftarrows 0'_\alpha$). Similarly, transient in n-input NOR under input actions (4.1) is obtained by inversion of (4.13) under (4.1) in n-input OR.

5. Determination of the transient in a logic circuit

The following algorithm is suggested for finding the transient at the output of an arbitrary logic circuit whose input actions have each one change at most.

Step 1. The delay at the output of each element is omitted and delays of the same value are inserted at all inputs of the element [5]. This process begins at the circuit output and all serially connected delays are summed as one moves towards circuit inputs. As the result, if the initial circuit did not include branches, all delays appear at the circuit inputs. Let channel i carrying signal $y_i(t)$ is branched into m channels with respective delays $\tau_1, \ldots, \tau_m$. All delays may by transferred directly into channel i only if $\tau_1 = \tau_2 = \ldots = \tau_m$. Therefore, in the general case, the sub-circuit generating $y_i(t)$ should be replaced by m copies, and delay τ_1 should be put at the output of the first copy, delay τ_2 put at the output of the second copy, etc.

Thus, after Step 1 the initial circuit is transformed into an equivalent one consisting of the ideal logic circuit with delay at inputs only. If the initial circuit was branchless, the resulting ideal circuit coincides with the initial one with the exception of delays. If there are branches, the ideal circuit differs from the initial one in that some sub-circuits whose inputs are inputs of the initial circuit as a whole, are repeated. In this case delays at the corresponding inputs of repeated circuits are not equal. Thus, in the general case, to each input of the initial circuit with signal $x_i(t)$ several inputs correspond whose signals are obtained from $x_i(t)$ by means of various time shifts.

It is evident, that delay transforms the process. Therefore, after Step 1 the problem reduces to finding the transient in the ideal logic circuit implementing a Boolean function $f(x_1, \ldots, x_n)$.

Step 2. Function $f(x_1, \ldots, x_n)$ is represented in the normal disjunctive form. The ideal circuit obtained in Step 1 is replaced, correspondingly, by a three-stage circuit (stages or: 1. inverters, 2. conjunctors, 3. disjunctors). Since process transformation in the inverter is evident, the circuit may be considered as sonsisting of the conjunctor stage and the output disjunctor.

Step 3. Since neither delays, nor inverters influence the number of changes in the process being transformed, actions at the inputs of the conjunctor stage have each one change at most. Therefore processes at conjunctor outputs are computed through (4.7). The disjunctor output process which is the desired transient is computed by (4.13) if conjunctor outputs carry processes with one change at most. In the general case, these processes may have up to two changes each (formula (4.7)). Therefore, in the general case, the sought for transient may be found by the universal method of multi-input disjunctor decomposition into two-input ones (see Section 3). However, a special iteration procedure (see Section 6) has proved to be more effective.

6. Interation procedure for computing multi-input disjunctor transient

Consider the ideal disjunctor whose input actions have at most two changes each, or, to put in more precisely, they have one of the following kinds: $1'_\alpha$, $0'_\alpha$, $1'_\alpha 0'_\beta$ (another theoretically possible by virtue of (4.7) action 01M never appears at the inputs of the final disjunctor of the two-stage circuit). The number of disjunctor inputs ϱ is regarded as variable. Show how one can find the transient in an $(s+1)$-input disjunctor if that in the s-input disjunctor is known.

Let there be a process

$$x(t) = a(\alpha_1, \alpha_2)\, \bar{a}(-, \alpha_3) \ldots \tilde{a}(-, \alpha_m)\,, \tag{6.1}$$

where $\tilde{a}$ stands for a and $\bar{a}$. In the infinite valued logic process $x(t)$ left-shifted by β is defined through conjunctions as follows:

$$x(t\beta) = a(\alpha_1\beta, \alpha_2\beta)\, \bar{a}\,(-, \alpha_3\beta) \ldots \tilde{a}(-, \alpha_m\beta) \tag{6.2}$$

and process $x(t)$ right-shifted by β is defined through disjunctions:

$$x(t \vee \beta) = a(\alpha_1 \vee \beta, \alpha_2 \vee \beta)\, \bar{a}\,(-, \alpha_3 \vee \beta) \ldots \tilde{a}(-, \alpha_m \vee \beta)\,. \tag{6.3}$$

Process $x(t)$ completed up to γ from the right is defined as follows:

$$x_\gamma(t) = \begin{cases} x(t) & \tilde{a} = 0 \\ x(t)\, 0(\alpha_m, \gamma) & \tilde{a} = 1\,, \quad \gamma > \alpha_m\,, \end{cases} \tag{6.4}$$

Process $x(t)$ completed up to γ from the left is defined as follows:

$${}_\gamma x(t) = \begin{cases} x(t) & , \quad a = 0 \\ 0(\gamma, \alpha_1)\, x(\gamma), & a = 1\,, \quad \gamma < \alpha_1\,. \end{cases} \tag{6.5}$$

If both operations are performed over a process (shift and completion), agree to perform first completion and second shift.

Statement 1. Let $x^{(s)}(t)$ be a transient in s-input disjunctor with arbitrary input actions given in the form of (6.1). Add to the disjunctor $(s+1)$-st input with action $1'_\alpha$. Then transient $x^{(s+1)}(t)$ in the resulting $(s+1)$-input disjunctor is defined by the following formula:

$$x^{(s+1)}(t) = x^{(s)}_\gamma(t\alpha) \tag{6.6}$$

where γ is an arbitrary constant satisfying condition $\gamma \geq \alpha$.

Proof. Represent $(s+1)$-input disjunctor by a two-stage circuit having in stage 1 an s-input, and in stage 2 a 2-input disjunctors. One may write then:

$$x^{(s+1)}(t) = x^{(s)}(t) \vee 1'_\alpha = \begin{cases} x^{(s)}(t), & t \leq \alpha \\ 1 \quad , & t < \alpha. \end{cases} \tag{6.7}$$

Consider now three cases:

1) $\alpha < \alpha_1$. In this case $\alpha_1\alpha = \alpha_2\alpha = \ldots = \alpha_m\alpha = \alpha$ and all changes in procoess $x^{(s)}(t\alpha)$ will superpose in one change in point α. Hence,

$$x^{(s)}_\gamma(t\alpha) = \begin{cases} \bar{a} = x^{(s)}(t), & t \leq \alpha \\ \tilde{a} \quad , & t > \alpha. \end{cases} \tag{6.8}$$

If $a = 0$, finc with (6.8)

$$x^{(s)}_\gamma(t\alpha) - x^{(s)}(t\alpha) = \begin{cases} x^{(s)}(t), & t \leq \alpha \\ 1 \quad , & t > \alpha. \end{cases} \tag{6.9}$$

Having compared (6.7) and (6.9) see that (6.6) holds. If $\tilde{a} = 1$ then

$$x^{(s)}_\varphi(t\alpha) = x^{(s)}(t\alpha)\, 0(\alpha_m\alpha, \gamma\alpha) \tag{6.10}$$

By condition, $\alpha < \alpha_m$, $\alpha \leq \gamma$ therefore $\alpha_m\alpha = \gamma\alpha = \alpha$ and pause $0(.)$ degenerates into change $1'_\alpha$ and we obtain from (6.10) by using (6.8)

$$x^{(s)}_\gamma(t\alpha) = \begin{cases} x^{(s)}(t\alpha), & t \leq \alpha \\ 1 \quad , & t > \alpha \end{cases} = \begin{cases} x^{(s)}(t), & t \leq \alpha, \\ 1 \quad , & t > \alpha. \end{cases} \tag{6.11}$$

By comparing (6.7) and (6.11) once more see that (6.6) holds.

2) $\alpha > \alpha_m$. In this case $x^{(s)}(t\alpha) = x^{(s)}(t)$. If $\tilde{a} = 0$, then, on the one hand,

$$x^{(s)}_\gamma(t\alpha) = x^{(s)}(t\alpha) = x^{(s)}(t) \tag{6.12}$$

and, on the other hand, with due regard to the fact that under $\tilde{a} = 0$, $\alpha > \alpha_m$ and $t > \alpha$, $x^{(s)}(t) = 1$ holds, obtain

$$x^{(s+1)}(t) = x^{(s)}(t) \tag{6.13}$$

Now comparison of (6.12) results in the sought for formula (6.6).

If $\tilde{a} = 1$, (6.10) holds. By condition, $\alpha_m < \alpha \leq \gamma$. Consequently, $\alpha_m \alpha = \alpha_m$, $\gamma\alpha = \alpha$ and $x^{(s)}(t\alpha) = x^{(s)}(t)$, therefore (6.10) turns into

$$x_\gamma^{(s)}(t\alpha) = x^{(s)}(t)\, 0(\alpha_m, \alpha) = \begin{cases} x^{(s)}(t)\,, & t \leq \alpha\,, \\ 1 \quad , & t > \alpha\,. \end{cases} \tag{6.14}$$

Comparing (6.7) with (6.14) obtain the required.

3) $\alpha_i \leq \alpha \leq \alpha_{i+1}$. Write

$$x^{(s)}(t\alpha) = a(\alpha_1\, \alpha, \alpha_2\, \alpha)\, \bar{a}(-, \alpha_3\, \alpha) \ldots \tilde{a}(-, \alpha_m\, \alpha)\,.$$

In the case under consideration

$$\alpha_1\, \alpha = \alpha_1, \ldots \alpha_1\, \alpha = \alpha_i; \alpha_{i+1}\, \alpha = \ldots = \alpha_m\, \alpha = \alpha\,. \tag{6.15}$$

Hence,

$$x^{(s)}(t\alpha) = \begin{cases} x^{(s)}(t)\,, & t \leq \alpha\,, \\ \bar{\tilde{a}} \quad , & t > \alpha\,. \end{cases} \tag{6.16}$$

Now, if $\tilde{a} = 0$, then

$$x_\gamma^{(s)}(t\alpha) = x^{(s)}(t\alpha) = \begin{cases} x^{(s)}(t)\,, & t \leq \alpha\,, \\ 1 \quad , & t > \alpha\,. \end{cases} \tag{6.17}$$

By comparing (6.7) with (6.17) obtain (6.6).

If $\tilde{a} = 1$, (6.10) holds. By virtue of $\gamma \geq \alpha$ and (6.15) obtain $\alpha_m \alpha = \gamma\alpha = \alpha$ and pause $0(.)$ in (6.10) degenerates into change $1'_\alpha$. Taking into account also (6.16) find

$$x_\gamma^{(s)}(t\alpha) = \begin{cases} x^{(s)}(t)\,, & t \leq \alpha\,, \\ 1 \quad , & t > \alpha\,. \end{cases} \tag{6.18}$$

Comparison of (6.7) and (6.18) also leads to (6.6).

Statement 2. Let $x^{(s)}(t)$ be a transient in an s-input disjunctor with arbitrary input actions given in the form of (6.1). Add to the disjunctor $(s + 1)$-st input with action $0'_\alpha$. The transient in the resulting $(s + 1)$-input disjunctor is defined by

$$x^{(s+1)}(t) = {}_\gamma x^{(s)}(t \vee \alpha) \tag{6.19}$$

where γ is an arbitrary constant satisfying condition $\gamma \leq \alpha$.

Proof of this statement follows that of Statement 1. As above, represent $(s + 1)$-input disjunctor by a two-stage circuit and obtain

$$x^{(s+1)}(t) = x^{(s)}(t) \vee 0'_\alpha = \begin{cases} 1 \quad , & t \leq \alpha\,, \\ x^{(s)}(t)\,, & t > \alpha\,. \end{cases} \tag{6.20}$$

Then consider the same three cases:

1) $\alpha < \alpha_1$, and also $\alpha < \alpha_2, \ldots, \alpha < \alpha_m$, whence

$$x^{(s)}(t \vee \alpha) = x^{(s)}(t)\,. \tag{6.21}$$

Now, if $a = 0$ then

$$_{\gamma}x^{(s)}(t \vee \alpha) = x^{(s)}(t \vee \alpha) = x^{(s)}(t)\,, \tag{6.22}$$

But it follows from (6.20) (with due regard for the fact that for $a = 0$, $\alpha < \alpha_1$, and $t \leq \alpha$, $x^{(s)}(t) = 1$ holds) that:

$$x^{(s+1)}(t) = x^{(s)}(t)\,, \tag{6.23}$$

By comparing (6.22) with (6.23) obtain the desired formula (6.19).

If $a = 1$, then

$$_{\gamma}x^{(s)}(t \vee \alpha) = 0(\gamma \vee \alpha, \alpha_1 \vee \alpha)\, x^{(s)}(t \vee \alpha)\,. \tag{6.24}$$

By condition, $\alpha_1 > \alpha \geq \gamma$, therefore $\gamma \vee \alpha = \alpha$, $\alpha_1 \vee \alpha = \alpha_1$, pause $0(.)$ turns into $0(\alpha, \alpha_1)$, and we obtain through (6.24) and (6.21)

$$_{\gamma}x^{(s)}(t \vee \alpha) = \begin{cases} 1 & , \quad t \leq \alpha\,, \\ x^{(s)}(t \vee \alpha)\,, & t > \alpha\,, \end{cases} = \begin{cases} 1 & , \quad t \leq \alpha\,, \\ x^{(s)}(t)\,, & t > \alpha\,. \end{cases} \tag{6.25}$$

Comparison (6.20) with (6.25) leads to (6.19).

2) $\alpha > \alpha_m$. In this case $\alpha > \alpha_1, \ldots, \alpha > \alpha_m$ therefore $\alpha_1 \vee \alpha = \ldots = \alpha_m \vee \alpha = \alpha$ and all the changes in process $x^{(s)}(t \vee \alpha)$ will superpose in one change in point α. Whence

$$x^{(s)}(t \vee \alpha) = \begin{cases} \bar{a} & , \quad t \leq \alpha\,, \\ \bar{\tilde{a}} = x^{(s)}(t)\,, & t > \alpha\,. \end{cases} \tag{6.26}$$

If $a = 0$, find through (6.26)

$$_{\gamma}x^{(s)}(t \vee \alpha) = x^{(s)}(t \vee \alpha) = \begin{cases} 1 & , \quad t \leq \alpha\,, \\ x^{(s)}(t)\,, & t > \alpha\,. \end{cases} \tag{6.27}$$

Comparison of (6.20) with (6.27) leads to (6.19).

If $a = 1$, (6.24) holds (by condition $\alpha > \alpha_1$, $\alpha \geq \gamma$). Therefore $\gamma \vee \alpha = \alpha_1 \vee \alpha$, pause $0(.)$ in (6.24) degenerates into change $0'_{\alpha}$ snd by (6.26) obtain

$$_{\gamma}x^{(s)}(t \vee \alpha) = \begin{cases} 1 & , \quad t \leq \alpha\,, \\ x^{(s)}(t \vee \alpha)\,, & t > \alpha\,. \end{cases} = \begin{cases} 1 & , \quad t \leq \alpha\,, \\ x^{(s)}(t)\,, & t > \alpha\,. \end{cases} \tag{6.28}$$

By comparing (6.20) with (6.28) see once more that (6.19) holds.

3) $\alpha_i \leq \alpha \leq \alpha_{i+1}$. Write

$$x^{(s)}(t \vee \alpha) = a(\alpha_1 \vee \alpha, \alpha_2 \vee \alpha)\, \bar{a}(-, \alpha_3 \vee \alpha) \ldots \tilde{a}(-, \alpha_m \vee \alpha)\,.$$

In this case

$$\alpha_1 \vee \alpha = \ldots = \alpha_i \vee \alpha = \alpha\,;\, \alpha_{i+1} \vee \alpha = \alpha_{i+1}\,, \ldots \alpha_m \vee \alpha = \alpha_m\,. \tag{6.29}$$

Consequently,

$$x^{(s)}(t \vee \alpha) = \begin{cases} \bar{a} & , \quad t \leq \alpha\,, \\ x^{(s)}(t)\,, & t \leq \alpha\,. \end{cases} \tag{6.30}$$

Now, if $a = 0$, then

$$_{\gamma}x^{(s)}(t \vee \alpha) = x^{(s)}(t \vee \alpha) = \begin{cases} 1 & , \quad t \leq \alpha\,, \\ x^{(s)}(t)\,, & t > \alpha\,. \end{cases} \tag{6.31}$$

Comparing (6.20) with (6.31) obtain (6.19).

If $a = 1$, (6.24) holds. Since, by condition, $\alpha_1 \leq \alpha$, $\gamma \leq \alpha$ we have that $\gamma \vee \alpha = \alpha_1 \vee \alpha = \alpha$, pause $0(.)$ in (6.24) degenerates in change $0'_\alpha$, and by using (6.30) obtain

$$_{\gamma}x^{(s)}(t \vee \alpha(= \begin{cases} 1 & , \quad t \leq \alpha \\ x^{(s)}(t \vee \alpha)\,, & t > \alpha \end{cases} \Bigg| = \begin{cases} \wedge, t \leq \alpha \\ x^{(s)}(t), t > \alpha\,. \end{cases} \tag{6.32}$$

Comparison of (6.20) and (6.32) leads once more to (6.19).

Statement 3. Let $x^{(s)}(t)$ be a transient in an s-input disjunctor with arbitrary input actions. Add to the disjunctor $(s + 1)$-st input with action $1'_\alpha 0'_\beta$. Then transient $x^{(s+1)}(t)$ in the resulting $(s + 1)$-input disjunctor is defined as a sequence of two sub-processes:

$$x^{(s+1)}(t) = x^{(s)}_{\gamma'}(t\alpha)_{\gamma_1} x^{(s)}(t \vee \beta)\,, \tag{6.33}$$

where γ_1, γ_2 are arbitrary constants satisfying conditions:

$$\alpha \leq \gamma_1\,, \qquad \gamma_2 \leq \beta \tag{6.34}$$

Proof. Having represented $(s + 1)$-input disjunctor with a two-stage circuit, one can write:

$$x^{(s+1)}(t) = x^{(s)}(t) \vee 1'_\alpha 0'_\beta\,. \tag{6.35}$$

Apply the decomposition method (see Section 3). Divide process $1'_\alpha 0'_\beta$ into $1'_\alpha$ and $0'_\beta$. Then partial responses of the $(s + 1)$-input disjunctor are found through Statement 1:

$$x_1^{(s+1}(t) = x^{(s)}(t) \vee 1'_\alpha = x^{(s)}_{\gamma'}(t\alpha)\,, \qquad \gamma_1 \geq \alpha \tag{6.36}$$

and Statement 2:

$$x_2^{(s+1)}(t) = x^{(s)}(t) \vee 0'_\beta = \gamma_2\, x^{(s)}(t \vee \beta)\,, \qquad \gamma_2 \leq \beta\,. \tag{6.37}$$

One sees from (6.36) that $x_1^{(s+1)}$ does not end later than time α, and from (6.37) that $x_2^{(s+0)}(t)$ does not begin earlier then time β. But $\alpha < \beta$ (otherwise given action $1'_\alpha 0'_\beta$ does not exist). Therefore the first partial response ends before the second begins. Hence, the transient in the $(s+1)$-input disjunctor may be represented as a sequence of its partial responses

$$x^{(s+1)}(t) = x_1^{(s+1)}(t)\, x_2^{(s+1)}(t)\,. \tag{6.38}$$

By substituting (6.38) into (6.36) and (6.37) obtain (6.33).

7. Conclusion

This procedure features great labour comsumption due, primarily, to the number of input combinations turning function implemented by the circuit into unit and, to an appreciably smaller extent, to the length of the circuit. The procedure enables one to analyse manually circuits having up to 5 through 7 inputs (the length of circuits is, practically, unlimited) and may be easily mechanized. It is also useful in qualitative studies of transients in logic circuits (failures, risks, hazards, etc.).

References

1. *Glushkov, V. M.*, Synthesis of Digital Automata, Fizmatgiz, Moscow, 1962 (in Russian).
2. *Huffman, D. A.*, The Design of Hazard-Free Switching Networks, J. Assoc. Comput. Machinery **4** (1957) 1.
3. *McCluskey, E.*, Transients in Combinatorial Logic Circuits, in Redundancy Techniques for Computing System, Sparten Book Co., 1962.
4. *Eichelberger, E. B.*, Hazard Detection in Combinational and Sequential Switching Circuits, JBHJ. Res. and Development **9** (1965) 2.
5. *Roginskii, V. N.*, Transients in relay devices (discrete automata), in Information Transmission Networks and their Automation, "Nauka", Moscow, 1965 (in Russian).
6. *McNaughton, R.*, A Theorem about Infinite Valued Sentential Logic, J. Symb. Logic **16** (1951) 1.
7. *Ginzburg, S. A.*, Continuous Mathematical Logic and Function Representation, "Energiya", Moscow, 1968 (in Russian).
8. *Levin, V. I.*, Infinite Valued Logic and Transients in Finite Automata, Avtomatika i Vychislitelnaya Tekhnika (1972) 6 (in Russian).
9. *Levin, V. I.*, Decomposition of Transients in Finite Automata, Avtomatika i Vychislitelnaya Tekhnika (1973) 3 (in Russian).
10. *Anger, S. H.*, Hazards and Delays in Asynchronous Szquential Swithcing Circuits, IRE Trans. on Circuits Theory **6**, 1 (1959).
11. *Moisil, Gr. C.*, Algebraic Theory of Discrete Automatic Devices, Bucuresti, Ed. Tehnica 1959 (Acad. Republicii Populare Romine. Stinta si tehnica) (in Roumanian).

MAGYAR
TUDOMANYOS AKADEMIA
KONYVTARA

NOTE TO CONTRIBUTORS

Two copies of the manuscripts (each duly completed by figures, tables and references) are to be sent either to

E. D. Teryaev coordinating editor

Department of Mechanics and Control Processes

Academy of Sciences of the USSR

Leninsky Prospect 14, Moscow V-71, USSR

or to

J. Kocsis coordinating editor

Department of Automation

Technical University

H—1111 Budapest XI., Goldman György tér 3, Hungary

The authors are requested to retain a third copy of the submitted typescript to be able to check the proofs against it.

The papers, preferably in English or Russian, should be typed double-spaced on one side of good-quality paper, with wide margins (c. 4—5 cm) should carry the title of the contribution, the author(s)' name, and the name of the country. At the end of the typescript the name of that author who manages the proof-reading should also be given.

An abstract of about 50 to 100 words should head the paper.

The authors are encouraged to use the following headings: Introduction, (outlining the problem), Methods and results, Discussion, Conclusions, References. The entire material should not exceed 15 pages including tables and references. The proper location of the tables and figures must be indicated on the margins. Mathematical notations should follow up-to-date usage.

The summary — possibly in Russian if the paper is in English and *vice versa* — should contain a brief account of the proposition and indications of the formulas used and figures shown in the paper. The summary is not supposed to exceed 10—15 per cent of the paper.

The authors will be sent sheet-proofs which they are to return by next mail to the sender Regional Editorial Board.

Authors are entitled to 100 reprints free of charge. Rejected manuscripts will be returned to the authors.

К СВЕДЕНИЮ АВТОРОВ

Рукописи в двух экземплярах (каждый из которы должен содержать рисунки, таблицы и литературу направляются

Е. Д. Теряеву — Научный секретарь журнала

Отделение механики и процессов управления

Академия Наук СССР

Ленинский Проспект, 14, Москва В-71, СССР

или

Я. Кочишу — Научный секретарь журнала

Кафедра Автоматизации
Будапештского Технического
Университета
H—1111 Будапешт XI, площадь Голдман Д., 3,
Венгрия

Авторам рекомендуется оставлять у себя копию всех представленных ими материалов для справок при корректуре.

Статьи, желательно на русском или английском языках, отпечатанные на бумаге хорошего качества, с промежутком в два интервала и широкими (4—5 см) полями должны содержать наименование статьи, фамилию автора (авторов), название страны. В конце статьи необходимо также указать фамилию автора, ответственного за корректуру гранок.

Статье должна предшествовать аннотация объемом до 50—100 слов.

Авторы при написании статьи должны придерживаться следующей формы: введение (постановка задачи), основное содержание и результаты, обсуждение, выводы и литература. Объем статьи не должен превышать 15 печатных страниц, включая таблицы и ссылки. Последовательность таблиц и рисунков должна быть отмечена на полях. Математические обозначения рекомендуется давать в соответствии с современными требованиями и традициями.

К статье обязательно должно быть приложено резюме-реферат. Резюме — на русском языке, если статья написана на английском, и наоборот — должно содержать краткое изложение текста статьи со ссылками на необходимые формулы и графики, имеющиеся в основном тексте. Объем разюме не должен превышать 10—15% объема статьи.

Авторам высылаются гранки статьи, которые они должны незамедлительно возвратить в Региональную секцию Редколлегии журнала.

Авторам обеспечивается бесплатно 100 оттисков их статей. Рукописи непринятых статей возвращаются авторам.

CONTENTS • СОДЕРЖАНИЕ

Pugachev, V. S.: Aleksandr Mikhailovich Letov (*Пугачев В. С.*: Александр Михайлович Летов) 239

Гоппа В. Д.: Универсальные алгебраические коды для симметричных каналов (*Goppa, V. D.*: Universal algebraic codes for symmetric channels) 241

Haber, R.—Keviczky, L.: Adaptive dual extremum control by finite order Volterra model (*Хабер Р.—Кевицки Л.*: Адаптивное дуальное экстремальное управление при помощи конечной модели Вольтерра) 247

Левин В. И.: Вычисление переходных процессов в комбинационных конечных автоматах (*Levin, V. I.*: Calculation of the transient processes in the combinational finite automata) 261

Kwiatek, B.—Zieba, R.: Homomorphisms of stochastic and nondeterminate automata (*Квиатек Б.—Зиеба Р.*: Гомоморфизм стохастических и недетерминированных автоматов) 277

Book review 285