

**МАТЕМАТИЧЕСКОЕ
МОДЕЛИРОВАНИЕ**

Министерство высшего и среднего специального образования
Латвийской ССР

Латвийский ордена Трудового Красного Знамени
государственный университет имени Петра Стучки

Вычислительный центр

МАТЕМАТИЧЕСКОЕ МОДЕЛИРОВАНИЕ

Межведомственный сборник научных трудов



Латвийский государственный университет им. П.Стучки
Рига 1978

В сборнике помещены работы, посвященные математическому моделированию интегральных схем. Круг вопросов, затронутых в работах, касается задач: создания комплексов программ анализа, исследования численных методов решения систем нелинейных трансцендентных и дифференциальных уравнений, создания эффективных численных методов моделирования компонентов интегральных схем, а также синтеза оптимальных фильтров.

Тематика сборника заинтересует широкий круг специалистов по прикладной математике и электронике, а также аспирантов и студентов, занимающихся машинным проектированием электронных схем.

РЕДАКЦИОННО-ИЗДАТЕЛЬСКОЕ УЧРЕЖДЕНИЕ:

Э.Я.Гринберг (отв.ред.), Л.З.Кашинельсон,
Б.С.Польский, Г.Ф.Страутманис

Печатается по решению редакционно-издательского совета
ЛГУ им.П.Стучки от 30 декабря 1977 года

© Латвийский государственный университет им.П.Стучки, 1978

М 30501-016y 175-78
М 812(II)-78

КОМПЛЕКС ПРОГРАММ ДЛЯ МОДЕЛИРОВАНИЯ ИНТЕГРАЛЬНЫХ СХЕМ (КОМПРОМИС)

Э.Н.Гринберг, Л.Э.Кацнельсон, И.Э.Озолинь,
А.А.Ванагс, А.А.Элисте, А.Н.Заворин, Е.С.Кельман,
Л.С.Похвалина, И.И.Кан-Каган
Вычислительный центр ЛГУ им.П.Стучки

На основе автоматизированной системы анализа модулей схем (АСАМС-Ф) [1-3] и дальнейших работ [5], [6] разработан комплекс программ для моделирования интегральных схем (КОМПРОМИС-2). КОМПРОМИС позволяет проводить для великих схем статический, динамический, статистический анализ, для линейных схем - частотный и шумовой анализ, расчет нулей и полюсов передаточной характеристики, а также статистический анализ для сквозного (нелинейного и линейного) анализа схем.

Возможности КОМПРОМИСа при проведении анализ нелинейных схем значительно расширены по сравнению с АСАМС-Ф.

1. В качестве модели транзистора используется модифицированная модель Эберса-Молла,^х для которой учитывается эффект Эрли и зависимость коэффициента усиления от тока базы;

2. Число транзисторов для проведения статического и динамического анализов увеличено до 40-60;

3. Широко используется модульный принцип для введения новых моделей транзисторов и диодов и построения системы;

4. Расширены возможности сервисных программ и алфавитно-цифрового языка.

При обнаружении ошибки в записи исходной информации система программ не прекращает свою работу, она печатает порядковый номер ошибки, характер ошибки и продолжает ввод остальной исходной информации, тем самым находит практически все ошибки в исходной информации. Система образует признак *ZBLKR*, значение которого зависит от правильности или характера ошибки в исходной информации.

^х Модель введена по предложению В.Маментова и П.Рожукална.

$ZBLKR=0$, если ошибки в исходной информации не обнаружены, в остальных случаях $ZBLKR \neq 0$. Признак $ZBLKR$ позволяет управлять работой системы в зависимости от характера ошибки.

В КОМПРОМИСе предусмотрена возможность многократного расчета схем при измененных номиналах без ввода информации о всей схеме, а также расширения схемы, т.е. добавления к введенной - новой схемы, которая имеет некоторые общие узлы с первоначальной схемой.

Для расчета статических характеристик используются различные модификации метода сопряженных градиентов, овражного, Флетчера-Пауэлла, Ньютона. Для расчета динамических характеристик используются различные методы неявного интегрирования. Широко используется техника работы с разреженными матрицами.

Ниже описываются дополнительные возможности КОМПРОМИСа при анализе схем по сравнению с АСАМС-Ф.

I. Статистический анализ схем:

Статистический анализ проводится в системе, позволяющей моделировать до 20 транзисторов.

Процесс статистического анализа схем включает:

- 1) многократное измерение номиналов схемы;
- 2) пересчет некоторых номиналов транзисторов;
- 3) статистическую обработку номиналов схемы: определение средних значений, дисперсий, третьих моментов, матрицы корреляций;
- 4) многократный анализ схемы при случайно выбранных номиналах;
- 5) статистическую обработку полученных результатов анализа.

Комплекс программ предназначен для выполнения всего процесса статистического анализа, за исключением сбора информации о номиналах схемы.

Для удобства пользователя комплекс программ реализован в виде трех отдельных частей, которые работают последовательно. В первой части, исходя из вольтамперных характеристик транзисторов и диодов, рассчитываются номиналы двухполосников модели, необходимые для проведения анализа схем. Предусмотрено несколько вариантов использования этой части. В результате работы подготавливается необходимая информация на перфокартах для 2-ой части. Главная задача второй части комплекса программ - подготовка информации для третьей - основной части комплекса. В этой части проводится статистическая обработка номиналов исследуемой схемы для получения средних значений дисперсий, третьих моментов, матрицы корреляции. При использовании комплексом эта часть работает дважды. Первый раз для получения матрицы корреляции, ее оценки и обработки, второй раз для получения усеченной матрицы корреляции и вывода на перфокарты средних значений, дисперсий, третьих моментов.

Третья часть комплекса проводит статистический анализ схемы, обрабатывая вычисленные параметры схемы. Исходной информацией для этой части являются информация о схеме, средних значениях, дисперсиях, третьих моментах, законах распределения, верхних и нижних границах номиналов, матрице корреляции, количестве бросаний (для метода Монте-Карло) и др.

Предусмотрена возможность прерывания счета с накоплением информации на магнитной ленте.

Комплекс позволяет проводить частотный и шумовой анализ схем.

Порядок работы комплекса при проведении сквозного анализа следующий: для нелинейной схемы рассчитывается рабочая точка, модель Эберса-Молла для транзистора и диода пересчитывается в малосигвальную модель, производится частотный и шумовой анализ схем, рассчитываются нули и

х) Программу обработки по методу моментов составили Н.А. Фонин, по методу Монте-Карло - А.Я. Бигельсон.

полосы. Затем, если это необходимо, вводятся изменения номиналов и анализ повторяется.

2. Анализ схем по малому сигналу

Анализ схемы по малому сигналу включает в себя: частотный анализ, шумовой анализ, поиск частоты единичного усиления и накопление значений функционалов для статистической обработки. При этом предполагается, что схема, полученная в результате анализа исходной схемы по большому сигналу, может рассматриваться как линейный трех- или четырехполюсник.

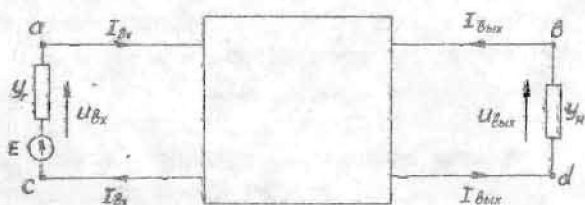


Рис. I. Четырехполюсник.

Через $I_{вх}$ и $I_{вых}$ обозначены комплексные амплитуды соответственно входного и выходного токов схемы, а через $U_{вх}$ и $U_{вых}$ обозначены комплексные амплитуды напряжений соответственно на входе и выходе схемы. Буквами a и c обозначены входные узлы схемы, а буквами b и d - выходные узлы схемы. Y_r - комплексная проводимость генератора на входе. Y_n - комплексная проводимость нагрузки на выходе. E - комплексная амплитуда ЭДС генератора. В схеме, подлежащей анализу по малому сигналу, могут присутствовать следующие элементы: сопротивления, индуктивности, емкости, а также диоды и транзисторы, заданные библиотечными моделями.

2.1. Частотный анализ.

При выполнении частотного анализа схемы, приведенной

на рис. I, могут быть вычислены следующие передаточные характеристики как функции от частоты:

- У-параметры, т.е. коэффициенты в уравнениях передачи четырехполюсника

$$I_{вх.} = Y_{11} U_{вх.} + Y_{12} U_{вых.}$$

$$I_{вых.} = Y_{21} U_{вх.} + Y_{22} U_{вых.}$$

- инвариантный коэффициент устойчивости K.

$$K = \frac{2 \cdot \operatorname{Re}(Y_{11}) \cdot \operatorname{Re}(Y_{22}) - \operatorname{Re}(Y_{12} \cdot Y_{21})}{\operatorname{Mod}(Y_{12} \cdot Y_{21})};$$

- коэффициент передачи напряжения генератора KE

$$KE = \frac{U_{вых.}}{E} = - \frac{Y_{21} \cdot Y_1}{(Y_{22} + Y_H)(Y_1 + Y_{11}) - Y_{12} \cdot Y_{21}};$$

- коэффициент передачи напряжения

$$KU = \frac{U_{вых.}}{U_{вх.}} = - \frac{Y_{21}}{Y_{22} + Y_H};$$

- входное комплексное сопротивление $Z_{вх.}$

$$Z_{вх.} = \frac{U_{вх.}}{I_{вх.}} = \frac{Y_{22} + Y_H}{Y_{11} \cdot (Y_{22} + Y_H) - Y_{12} \cdot Y_{21}};$$

- входное сопротивление для дифференциального сигнала $R_{вх.}$

$$R_{вх.} = \operatorname{Re}(Z_{вх.});$$

- входная емкость для дифференциального сигнала $C_{вх.}$

$$C_{вх.} = \frac{1}{\omega \cdot \operatorname{Im}(Z_{вх.})};$$

- выходное комплексное сопротивление $Z_{вых.}$

$$Z_{OUT} = \frac{U_{ВЫХ}}{I_{ВЫХ}} = \frac{Y_{11} + Y_1}{(Y_{11} + Y_1) \cdot Y_{22} - Y_{12} \cdot Y_{21}} ;$$

- выходное сопротивление для дифференциального сигнала R_{OUT}

$$R_{OUT} = Re(Z_{OUT}) ;$$

- выходная емкость для дифференциального сигнала C_{OUT}

$$C_{OUT} = \frac{1}{\omega \cdot Im(Z_{OUT})} ;$$

- входное сопротивление для синфазных сигналов $R_{IN\ SIN\ F}$

$$R_{IN\ SIN\ F} = Re(Z_{IN\ SIN\ F}) ;$$

$Z_{IN\ SIN\ F}$ определяется как входное комплексное сопротивление сх мы при закороченных между собой входах а и с ;

- входная емкость для синфазных сигналов $C_{IN\ SIN\ F}$

$$C_{IN\ SIN\ F} = \frac{1}{\omega \cdot Im(Z_{IN\ SIN\ F})} ;$$

- коэффициенты относительного ослабления синфазных сигналов $CMRR$ (раз.) и $CMRR$ (дб.).

$$CMRR (раз.) = 2 \frac{KU' + KU''}{KU' - KU''} ;$$

$$CMRR (дб.) = 20 \cdot \lg CMRR (раз.) ,$$

где

KU' - коэффициент усиления напряжения по входу α при

заземленном входе c (рис.2,а);

KU^* - коэффициент усиления напряжения по входу c при заземленном входе a (рис.2,б).

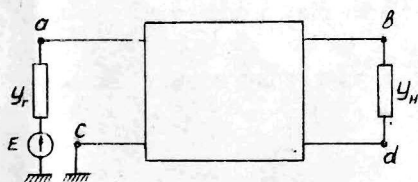


рис.2,а

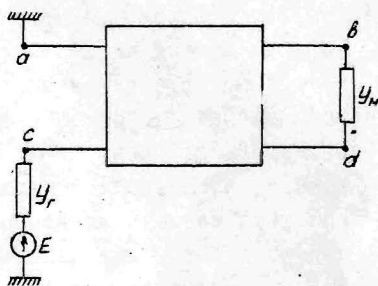


рис.2,б

Схемы для расчетов коэффициентов усиления.

2.2. Шумовой анализ.

При выполнении шумового анализа для каждого заданного значения температуры вычисляются следующие величины как функции от частоты:

- амплитуда суммарного выходного шума $V_{NOISE OUT}$

$$V_{NOISE OUT} = \sqrt{\sum_{L=1}^n |V_L|^2},$$

где V_L - выходное шумовое напряжение, создаваемое L -ым источником шума в схеме и n - общее количество источников шумов в схеме;

- эквивалентная амплитуда входного шума $V_{NOISE IN EQU}$ вычисляется с помощью коэффициента передачи напряжения

$$V_{NOISE IN EQU} = V_{NOISE OUT} / KU.$$

2.3. Накопление значений функционалов для статистической обработки.

Функционалами для статистической обработки могут

быть компоненты частотных и шумовых характеристик схемы, а также значение частоты единичного усиления и соответствующее ей значение фазочастотной характеристики.

2.4. Задание входной информации для анализа схемы по малому сигналу.

Для проведения анализа схемы по малому сигналу может потребоваться следующая информация:

- наименование схемы;
- номера узлов входа-выхода;
- описание нагрузки и генератора;
- перечень требуемых видов анализа;
- информация о частотах;
- информация о температурах;
- параметры для поиска частоты единичного усиления;
- информация о количестве, последовательности и форме вывода результатов на печать;
- перечень функционалов, значения которых накапливаются для статистической обработки.

2.4.1. Наименование схемы.

Наименование не является обязательной. Используется для вывода результатов.

2.4.2. Номера узлов входа-выхода схемы.

Узлы схемы, отмеченные на рис.1 буквами a и c , считаются входными узлами схемы, а узлы, отмеченные буквами b и d , - выходными. Номера узлов a, b, c, d должны удовлетворять следующим условиям:

$$a \neq b; a \neq c; a \neq d; b \neq c, b \neq d$$

и быть целыми неотрицательными числами. Номера узлов входа-выхода не могут иметь значения, превышающего максимальный номер узла в схеме. Допустим вариант, когда $c=0$ или $d=0$, а также $c=d$ и $c=d=0$.

Если $c=0$ или $c=d$, то для такой схемы не будут вычисляться величины $R_{IN(SNF)}$, $C_{IN(SNF)}$, $CMRR$ (раз)

и *CMRR* (дб.).

2.4.3. Описание нагрузки и элемента генератора.

Набор допустимых элементов в данном случае ограничен. Допустимыми элементами являются: сопротивление, емкость, индуктивность или двухполосник, состоящий из элементов, перечисленных трех типов.

2.4.4. Виды анализа.

Выполнение частотного анализа обеспечивается картой:
AC ANALYSIS.

Выполнение шумового анализа задается картой:

NOISE ANALYSIS.

В информации могут присутствовать обе карты вместе или только одна из них. Наличие любой из этих карт означает также вывод на печать результатов для заданного вида анализа.

Отсутствие карт видов анализа допустимо, а при задании информации для статистики даже необходимо.

2.4.5. Информация о частотах.

Обязательна для шумового, частотного и статистического анализов. Можно задать два типа частот: линейную и круговую. Для каждого типа частот возможны две формы задания: таблица и интервал.

2.4.6. Информация о температурах. Используется для выполнения шумового анализа и расчета соответствующих функционалов для статистического анализа.

2.4.7. Печать результатов.

Результаты частотного и шумового анализов выводятся на печать в виде таблиц и в смешанной форме: таблиц и графиков.

2.4.8. Функционалы для статистической обработки при сквозном анализе.

Для статистической обработки могут быть использованы любые десять из нижеперечисленных функционалов:

$Re(Y_{11}), Im(Y_{11}), Re(Y_{12}), Im(Y_{12}), Re(Y_{21}), Im(Y_{21}),$
 $Re(Y_{22}), Im(Y_{22}), K, Re(KE), Im(KE), Mod(KE),$
 $Arg(KE), Re(KU), Im(KU), Mod(KU), Arg(KU),$
 $Mod(Z_{IN}), Arg(Z_{IN}), R_{IN}, C_{IN}, Mod(Z_{OUT}),$
 $Arg(Z_{OUT}), R_{OUT}, C_{OUT}, R_{IN\ SINP}, C_{IN\ SINP},$
 $CMRR(\rho_{3}), CMRR(\delta), V_{NOISE\ OUT}, V_{NOISE\ IN\ EQU},$
 $F1, Arg(KU(F1)).$

Предусмотрено накопление значений функционалов для 31 частоты.

3. Расчет полюсов и нулей передаточной функции.

Пусть после анализа схемы по большому сигналу в ней выделены входное $U_{вх}$ и выходное $U_{вых}$ (рис.1) напряжения. Тогда при задании в исходной информации соответствующих входных и выходных узлов рассчитываются полюса и нули передаточной функции

$$T(\rho) = \frac{U_{вых}}{U_{вх}} = \frac{\Delta_1(\rho)}{\Delta_2(\rho)},$$

где $\Delta_2(\rho)$ - определитель системы линейных алгебраических уравнений, записанной по методу узловых потенциалов; $\Delta_1(\rho)$ - определитель, отличающийся от $\Delta_2(\rho)$ первым или двумя первыми столбцами, в зависимости от того, является ли схема трехполюсником или двухполюсником.

В схеме, подлежащей анализу на полюсы и нули, могут присутствовать те же элементы, что и в случае анализа по малому сигналу, а именно: сопротивления, индуктивности, емкости, а также диоды и транзисторы, заданные библиотечными моделями.

Для проведения расчета полюсов и нулей задается ин-

формация о схеме для анализа по большому сигналу и дополнительно к ней - информация о входных и выходных узлах. Последняя задается так же, как для анализа по малому сигналу.

Комплекс КОМПРОМИС выполнен на языке Фортран для ЭВМ ДЖИИ-400.

В настоящее время проводятся работы по расширению возможностей комплекса и перевода его на ЕС ЭВМ.

Литература

1. Аринь Э.И., Гринберг Э.Я., Кацнельсон Л.З. и др. Автоматизированная система анализа моделей схем на языке Фортран (АСАМС-Ф), ч.1. Общее описание и инструкция по использованию. Рига, ЛГУ им.П.Стучки, 1972. 68 с.
2. Аринь Э.И., Гринберг Э.Я., Кацнельсон Л.З. и др. Автоматизированная система анализа моделей схем на языке Фортран (АСАМС-Ф), ч.2. Описание алгоритмов и программ. Рига, ЛГУ им.П.Стучки, 1972, 184 с.
3. Гринберг Э.Я., Кацнельсон Л.З. и др. Автоматизированная система анализа моделей схем на языке Фортран (АСАМС-Ф), ч.3. Структура и программы системы. Рига, ЛГУ им.П.Стучки, 1972, 172 с.
4. Заворин А.Н. О некоторых методах решения разреженных систем узловых уравнений.-В настоящем сборнике, с.15-31.
5. Гринберг Э.Я., Кацнельсон Л.З., Заворин А.Н. и др. Система анализа электронных схем и линейной связи.-В кн.: Латв.мат.ежегодник, 14. Рига, Зинатне, 1974, с.25-32.
6. Гринберг Э.Я., Кацнельсон Л.З. и др. Моделирование интегральных схем на ЭВМ. - В кн.: Латв. мат. ежегодник, 20, Рига, Зинатне, 1977, с.36-48.

О НЕКОТОРЫХ МЕТОДАХ РЕШЕНИЯ РАЗРЕЖЕННЫХ СИСТЕМ УЗЛОВЫХ УРАВНЕНИЙ

А.Н.Заворин

Вычислительный центр ЛГУ им.П.Стучки

Расчет электрических цепей обычно бывает связан с необходимостью решения систем линейных алгебраических уравнений, записываемых по какому-либо из известных методов, например, по методу узловых потенциалов, контурных токов и т.п. Эти системы обладают целым рядом свойств, которые могут быть использованы при численном решении. В частности, если электрическая цепь имеет много узлов (контуров) невысоких степеней, т.е. таких, каждый из которых связан с небольшим числом других узлов (контуров), то система узловых (контурных) уравнений цепи оказывается разреженной. Разреженные системы уравнений часто встречаются при решении и других прикладных задач.

Определение. Матрица (и система уравнений с такой матрицей) порядка m называется разреженной, если число M ее ненулевых элементов удовлетворяет неравенству $M \ll m^2$.

Из этого определения видно, что оно имеет смысл только для систем достаточно высокого порядка.

Для решения разреженных систем на ЭВМ в принципе могут применяться те же численные методы, которые применяются для решения обычных систем линейных алгебраических уравнений и которые работают со строками и столбцами как с целыми. Однако такой подход в случае разреженных систем оказывается неэффективным, поскольку он предполагает выполнение арифметических операций в основном с нулевыми коэффициентами. Если при этом используется точный метод, то общее количество арифметических операций у него не может быть меньше числа операций, которого требует метод Гаусса, и сравнимого с m^3 [1]. Итерационные методы на каждой итерации требуют числа операций, сравнимого с m^2 . Таким образом, при возрастании порядка m в случае разре-

менных систем резко возрастают как ненужные в принципе затраты вычислительных операций, так и затраты памяти.

В связи с таким положением в последнее время были разработаны специальные алгоритмы, предназначенные для решения разреженных систем. Эти алгоритмы требуют хранения в памяти и выполнения арифметических операций только с ненулевыми элементами матриц.

В настоящей работе мы проанализируем некоторые из таких алгоритмов применительно к случаю разреженных систем узловых уравнений. При этом главное внимание уделим сравнению эффективности итерационных алгоритмов и алгоритмов точного решения.

I. Укажем сначала на некоторые свойства системы узловых уравнений, наиболее важные с точки зрения применимости численных методов.

Запись матрицы системы узловых уравнений для произвольной цепи осуществляется по следующему правилу. Все множества узлов цепи разбиваются на непересекающиеся подмножества, в которые входят либо отдельные узлы, либо группы узлов, если они связаны между собой источниками напряжения. Назовем эти подмножества независимыми узлами. Для каждого такого узла в матрице записывается отдельная строка. Принадлежащий ей диагональный элемент равен сумме всех (в общем случае комплексных) проводимостей, взятых со знаком плюс и связывающих независимый узел со всеми остальными узлами цепи. Недиагональный элемент с индексами i, k равен сумме всех проводимостей, взятых со знаком минус и связывающих независимые узлы i и k между собой.

Матрица $A = (a_{ik})_{i,k=1,2,\dots,m}$ системы узловых уравнений

$$AX = B, \quad (1)$$

отвечающей резисторной цепи без зависимых источников, обладает следующими свойствами:

- I.1. Если цепь не содержит сечений, целиком состоящих из источников тока и контуров, состоящих только из источников напряжения, то матрица системы положительно определена, а следовательно, и не вырождена.
- I.2. Матрица симметрична.
- I.3. Диагональные элементы матрицы положительны, а недиагональные неположительны.
- I.4. Для матрицы выполнены слабые признаки сумм по строкам и по столбцам, т.е. например, для всех строк

$$|a_{ii}| \geq \sum_{\substack{k=1 \\ k \neq i}}^m |a_{ik}|, \quad i=1,2,\dots,m$$

и по крайней мере для одного $i=n$ выполняется строгое неравенство

$$|a_{nn}| > \sum_{\substack{k=1 \\ k \neq n}}^m |a_{nk}|$$

- I.5. Если цепь содержит много независимых узлов невысоких степеней, то матрица разрежена.
- Свойством разреженности обладает и матрица системы узловых уравнений с комплексными операторными проводимостями, отвечающей схеме с реактивными элементами.

2. Рассмотрим вопрос о применимости для решения систем с указанными выше свойствами некоторых численных методов, а именно: методов Гаусса, простой итерации и Зейделя.

Обратимся сначала к методу Гаусса. Свойства системы узловых уравнений позволяют применить для нее решения схему единственного деления. Действительно, из свойства I.1, следует, что при реализации схемы все ведущие элементы отличны от нуля, а это гарантирует получение решения. Легко видеть, что схема единственного деления не нарушает свойства I.4, в процессе исключения переменных, поэтому эта схема эквивалентна схеме с выбором главного элемента

в столбце или строке.

Сходимость итерационного метода Зейделя обеспечивается свойствами I.1 и I.2 [3].

При использовании метода простой итерации из свойства I.1 сразу следует, что последовательные приближения не выйдут из некоторой окрестности решения. Это, однако, еще не означает строгой сходимости метода, которая тем не менее имеет место. Чтобы это обосновать, докажем следующую теорему, обобщающую теорему о сходимости метода простой итерации для системы с неразложимой матрицей [4].

Теорема. Пусть матрица A системы

$$AX = B \tag{1}$$

разложима, т.е. после перестановки одних и тех же строк и столбцов она может быть разбита на подматрицы в виде

$$A^* = \begin{pmatrix} A_{11} & 0 & 0 & \dots \\ A_{21} & A_{22} & 0 & \dots \\ A_{31} & A_{32} & A_{33} & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix} \tag{2}$$

Пусть, далее, все диагональные квадратные матрицы A_{ii} неразложимы, и в каждой из подматриц $(A_{11}), (A_{21}, A_{22}), (A_{31}, A_{32}, A_{33}), \dots$ матрицы A^* выполнен слабый признак сумм по строкам. Тогда метод простой итерации для системы (1) сходится.

Доказательство. Для простоты изложения допустим, что матрица A^* имеет вид

$$A^* = \begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix}, \quad A_{11} = (a_{ix}^*)_{i,x=1,2,\dots,p} \tag{3}$$

Поскольку матрица A приведена к виду (3) с помощью перестановок одинаковых строк и столбцов, то система уравнений



$$A^* X^n = B^* \quad (4)$$

совпадает с системой (I) с точностью до перенумерации неизвестных и уравнений. Поэтому сходимость метода простой итерации для системы (I) будет следовать из сходимости этого метода для системы (4). Последняя же имеет место в том и только в том случае, если корни уравнения

$$\begin{vmatrix} a_{11}^* \lambda & a_{12}^* & \dots & \dots \\ a_{21}^* & a_{22}^* \lambda & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{m1}^* & a_{m2}^* & \dots & a_{mm}^* \lambda \end{vmatrix} = 0, \quad (5)$$

составленного для матрицы A^* , по модулю меньше единицы [4].

По условию для матрицы A^* выполнен слабый признак суммы по строкам, поэтому уравнение (5) не может иметь корней, по модулю больших единицы.

Допустим теперь, что уравнение (5) имеет корень λ такой, что $|\lambda| = 1$. Тогда существует ненулевой вектор $Y = (y_1, y_2, \dots, y_m)$, удовлетворяющий однородной системе с матрицей, содержащейся в уравнении (5), и при этом

$$|a_{ii}^*| |y_i| = \left| \sum_{k=1}^m a_{ik}^* y_k \right|, \quad i = 1, 2, \dots, m \quad (6)$$

Обозначим

$$\alpha = \max_{1 \leq i \leq m} |y_i| = |y_s|, \quad 1 \leq s \leq m \quad (7)$$

Из слабого признака суммы по строкам и из (6) имеем

$$\sum_{k \neq s} |\alpha_{sk}^*| |y_s| \leq |\alpha_{ss}^*| |y_s| < \sum_{k \neq s} |\alpha_{sk}^*| |y_k| \quad (8)$$

или

$$\sum_{k \neq s} |\alpha_{sk}^*| (|y_s| - |y_k|) < 0, \quad (9)$$

откуда следует, что если только некоторый $\alpha_{sk}^* \neq 0$, то $|y_k| = |y_s| = \alpha$.

Предположим теперь, что $1 \leq s \leq p$, тогда $p > 1$ и найдется по крайней мере один такой номер $s_1 < p$, что $s_1 = s$ и $\alpha_{s_1 s_1}^* \neq 0$, ибо в противном случае матрица A_n разложима. Следовательно, $|y_{s_1}| = \alpha$ и неравенство (9) справедливо для строки с номером s_1 . Для нее проведем аналогичное рассмотрение и получим новый номер $s_2 \leq p$ такой, что $|y_{s_2}| = \alpha$. Продолжая процесс далее найдем, что все $|y_i| = \alpha$, $1 \leq i \leq p$, в противном случае матрица A_n оказалась бы разложимой. Тогда из неравенства (8) получаем $|\alpha_{s_1 s_1}^*| \leq \sum_{k \neq s_1} |\alpha_{s_1 k}^*|$, $s_1 = 1, 2, \dots, p$, что противоречит слабому признаку сумм для матрицы A_n .

Пусть, наконец, $p+1 \leq s \leq m$. В этом случае, так же как и в предыдущем, опираясь на неразложимость матрицы A_{22} и на неравенство (9), устанавливаем, что все $|y_i| = \alpha$, если $p+1 \leq i \leq m$. Для $1 \leq i \leq p$ равенство $|y_i| = \alpha$ будет иметь место в том случае, если отличны от нуля соответствующие a_{ik} , $p+1 \leq i \leq m$, входящие в матрицу A_{21} . Таким образом, снова приходим к противоречию со слабым признаком сумм по строкам для подматрицы (A_{21}, A_{22}) . Теорема доказана.

Примечание 1. Доказанная теорема справедлива для систем как с действительными, так и с комплексными коэффициентами.

Примечание 2. Аналогичная теорема может быть доказана и при выполнении слабого признака сумм по столбцам.

Итак, пусть система узловых уравнений отвечает схеме, содержащей омические сопротивления, емкости, индуктивнос-

ти, независимые источники тока и напряжения и не содержащей зависимые источники. Как уже указывалось выше, матрица такой системы удовлетворяет слабым признакам суммы по строкам и столбцам. Поэтому, если эта матрица неразложима, то сходимость метода простой итерации сразу следует из теоремы.

Если же матрица системы разложима, причем она приводится к виду (3) перестановкой неодинаковых строк и столбцов, то легко видеть, что существует перестановка одинаковых строк и столбцов, приводящая матрицу к тому же виду. Такая перестановка, сохраняющая симметрию матрицы, эквивалентна перенумерации неизвестных. Таким образом, матрица A_{ii} в (3) будет нулевой и разложимость будет указывать на то, что электрическая цепь состоит из двух по существу независимых схем, связанных между собой узлами с известным напряжением. Этим схемам матрицы A_{ii} и A_{jj} и для каждой из них выполнен слабый признак суммы (предполагается, что электрическая цепь не содержит сечений из источников тока). Следовательно, и в этом случае сходимость метода простой итерации имеет место.

Для линейных схем с зависимыми источниками матрица системы узловых уравнений может не удовлетворять слабому признаку суммы, причем это обстоятельство может зависеть не только от способа подключения двухполосников друг к другу, но и от величин операторных проводимостей. В частности, при анализе цепей с транзисторами, заменяемыми линейными моделями [5], содержащими зависимые источники, часто встречается следующая ситуация. При малых по модулю значениях комплексной частоты в матрице системы узловых уравнений признак суммы не выполняется. Однако при больших значениях частоты, когда операторные проводимости емкостей превышают проводимости зависимых источников, матрица системы удовлетворяет признаку суммы и сходимость метода простой итерации следует из доказанной теоремы.

3. Сравним теперь эффективность итерационных методов

и метода Гаусса, учитывая только количество арифметических операций, которые они требуют для решения системы, работая со строками и столбцами как с целыми. Метод Гаусса, реализованный по схеме единственного деления, требует $m/3(m^2 + 5m - 1)$ операций умножения и деления и $m/3(m-1)(m-1/2)$ операций сложения и вычитания [1].

Итеративные методы на каждом шаге требуют и тех и других операций по $m(m-1)$. Если взять отношения соответствующих величин для обоих видов операций, то при достаточно больших значениях m они будут сравнимы с коэффициентом κ , равным

$$\kappa = \frac{m}{3}, \quad (10)$$

Этот коэффициент определяет количество итераций, требующих выполнения того же числа операций, что и метод Гаусса. Вместе с тем он показывает, что при решении однотипных задач сравнительная эффективность итеративных методов растет с возрастанием m , если при этом не происходит значительного увеличения числа итераций, дающих решение задачи с заданной точностью.

4. Сравнение эффективности методов, реализованных с учетом разреженности матриц, основанное только на сравнении требуемых арифметических операций, имеет весьма малый практический смысл. Уменьшение объема арифметических операций является лишь одной стороной дела. Важно, чтобы при этом были приемлемыми также используемый объем памяти и объем, неизбежно появляющихся вспомогательных операций. В настоящем пункте мы рассмотрим некоторые алгоритмы и программы, предназначенные для решения больших разреженных систем и основанные на ранее отмеченных методах, а именно: символическое генерирование алгоритма Кроута [6], алгоритм по схеме единственного деления метода Гаусса, разработанный автором, программы методов простой итерации и Зейделя.

Алгоритм, предложенный в работе [6], не существуе

нован на известной модификации метода Гаусса. Матрица системы

$$AX = B \quad (I)$$

разлагается в произведение двух треугольных матриц L и U , после чего решение находится из двух систем

$$\begin{cases} LY = B \\ UX = Y \end{cases} \quad (II)$$

Элементы матриц L и U вычисляются по простым рекуррентным формулам.

Алгоритм реализован в виде программы *GNSO*, которая в свою очередь формирует программу *SOLVE*, осуществляющую непосредственное решение заданий системы. Основной исходной информацией для *GNSO* служит вспомогательная матрица \bar{A} , состоящая из нулей и единиц и вводимая в ЭВМ поразрядно (по битам). Элемент матрицы \bar{A} полагается равным единице в том и только в том случае, если соответствующий элемент матрицы A отличен от нуля. Основной информацией для *SOLVE* служат одномерные массивы ненулевых элементов матрицы A и столбца B .

Программа *GNSO* символически осуществляет разложение матрицы A на треугольные сомножители и решает системы (II), последовательно формируя ФОРТРАН-выражения для вычисления элементов матриц L и U , а затем и самого решения. Эта последовательность ФОРТРАН-выражений, не содержащая циклов и включающая только арифметические операции, и является программой *SOLVE*.

Работая с матрицей \bar{A} , программа *GNSO* составляет выражения только из ненулевых элементов матрицы A .

Таким образом, описанный выше подход обеспечивает экономию памяти за счет использования не самой матрицы A а списка ее ненулевых элементов и вспомогательной матрицы \bar{A} , занимающей сравнительно немного места, благодаря по-

разрядной записи элементов. Экономия в объеме вычислений достигается за счет выполнения арифметических операций только с ненулевыми элементами матрицы A . Программа *SOLVE* после трансляции может быть использована многократно, если необходимо решать несколько раз систему (I) с изменяющейся матрицей A , по неизменной \bar{A} .

Кной способ использования разреженности применен автором при разработке программы, реализующих схему единственного деления метода Гаусса для матриц с симметрично расположенными ненулевыми элементами. Весь процесс решения системы (I) разбит на два этапа. Первый этап осуществляется с помощью программы *TRBAN*, которая приводит исходную матрицу к треугольному виду. На втором этапе программа *PETPE* решает систему с треугольной матрицей A после необходимого преобразования столбца свободных членов.

Программы требуют задания следующей информации: одномерного массива AM ненулевых элементов матрицы A , взятых по столбцам; одномерного массива $JCTR$ номеров строк матрицы, которым принадлежат соответствующие элементы массива AM ; одномерного массива $JKON$, содержащего индексы элементов массива AM , которыми заканчиваются все столбцы.

Программа *TRBAN*, в отличие от *GN50*, непосредственно выполняет все арифметические операции, работая только с ненулевыми элементами матрицы A . Однако от программы *SOLVE* она отличается циклической структурой и наличием вспомогательных и, в частности, логических операций. Эти операции необходимы для настройки циклов на работу с элементами определенных столбцов матрицы, при определении принадлежности этих элементов определенным ее строкам, при вставке новых ненулевых элементов, появляющихся в ходе исключения по Гауссу. Поскольку число новых ненулевых элементов существенно образом зависит от расположения исходных элементов в матрице, в программе *TRBAN* предусмотрена операция перестановки строк и столбцов. Она обеспечивает исключение поддиагональных элементов в том из еще оставшихся столбцов,

в которых этих элементов содержится минимальное число.

Программа REIRE работает с информацией, подготовленной для нее программой TPEAN. Прежде чем решать систему с треугольной матрицей, REIRE выполняет арифметические операции с элементами столбца свободных членов в соответствии с тем, как происходило исключение в TPEAN. Все необходимые коэффициенты находятся на местах исключенных поддиагональных элементов матрицы.

Следует отметить особо два режима использования программ TPEAN и REIRE. Часто бывает нужно решать систему с неизменной матрицей, но с изменяющимися свободными членами. В этом случае выполняется одно обращение к TPEAN, после чего для каждого нового столбца свободных членов работает только REIRE.

Если же требуется решать систему с изменяющейся матрицей, но неизменной структурой расположения ненулевых элементов, то перед вторичным и последующими обращениями к TPEAN массив AM нужно заполнять в соответствии с содержанием массивов ЭСТР и ЭКОН, оставшихся после первого обращения. Это позволяет избежать повторения всех операций, необходимых для вставки новых ненулевых элементов и перестановки строк и столбцов.

Алгоритмы и программы методов простой итерации и Зейделя для разреженных систем используют ту же информацию, что и программы TPEAN и REIRE. Как и эти последние, они обеспечивают экономию машинного времени и памяти ЭВМ за счет хранения в памяти и выполнения арифметических операций только с ненулевыми элементами матрицы решаемой системы. При этом они требуют очень небольшого числа вспомогательных операций, необходимых для организации циклов и оценки погрешности очередного приближения.

5. Эффективность всех указанных выше алгоритмов и программ в сильной мере зависит от того, какая конкретно решается система. Однако они имеют ряд важных признаков, позволяющих сделать выводы об их эффективности весьма об-

щего характера.

Обратимся сначала к программам *GNSO* и *SOLVE*. Основным недостатком алгоритма, положенного в основу программы *GNSO*, является то, что он включает хотя и символическую, но тем не менее полную процедуру разложения матрицы на треугольные сомножители с учетом как ненулевых, так и нулевых элементов. Следовательно, общее количество операций при работе *GNSO* сравнимо с m^3 , так что сама по себе

GNSO не дает экономии машинного времени по сравнению с обычными программами, не учитывающими разреженности. В противоположность *GNSO* программа *SOLVE* включает в себя арифметические операции только с ненулевыми элементами матрицы, поэтому время ее работы для разреженной системы, например, пятидесятого порядка, может быть в 100 раз меньшим времени работы *GNSO*.

Таким образом, программы *GNSO* и *SOLVE* более эффективны, чем программы, не учитывающие разреженности, только в том случае, если необходимо много раз решать разреженную систему уравнений с изменяющимися коэффициентами, но с неизменной структурой расположения нулевых коэффициентов.

Недостатком этого алгоритма является также необходимость резервирования значительного объема памяти для программы *SOLVE* на случай, если количество ненулевых элементов матрицы окажется большим. В качестве примера можно привести такой крайний случай, как система с полной матрицей. Объем необходимой памяти при этом будет сравнимы с m^2 .

Программы *TPEAN* и *PETPE* лишены основных недостатков программы *GNSO* и *SOLVE*, поскольку они работают только с ненулевыми элементами матрицы. Однако из-за циклической структуры и других вспомогательных операций им требуется для решения больше машинного времени, чем программе *SOLVE*. Слабой стороной программы *TPEAN*, как и программы *GNSO*, является принципиальная возможность переполнения массива

ненулевых элементов в процессе решения системы. Размерность массива задается заранее и с расчетом на разреженность системы, а его длина в каждом конкретном случае определяется только после окончания процесса исключения. Если же действительная длина массива при решении некоторой системы превысит его размерность, то без изменения программы это решение становится невозможным.

Программы, основанные на итерационных методах, лишены всех вышеотмеченных недостатков. При работе с разреженными системами они обеспечивают максимальную экономию памяти. Выполняя арифметические операции только с ненулевыми элементами матриц, эти программы требуют незначительного числа вспомогательных операций, благодаря простоте своей циклической структуры. Однако при всех достоинствах программы итерационных методов обладают одним очень существенным недостатком, а именно: часто в условиях плохой сходимости они не обеспечивают решения задачи за практически приемлемое время, требуя для достижения заданной точности большого числа итераций.

Выше, в пункте 3 было отмечено, что при решении системы с полной матрицей эффективность итерационных методов, по сравнению с методом Гаусса, возрастает с увеличением порядка системы, если при этом не уменьшается скорость сходимости. Рассмотрим теперь этот вопрос для случая разреженных систем. За основу возьмем программу ТРЕАН и программу метода простой итерации.

Обозначим, как и ранее, через m порядок системы. Пусть в каждом столбце матрицы системы, имеющей симметрично расположенные ненулевые элементы, вне диагонали находится в среднем S ненулевых элементов, причем пусть $S \ll m$. Пусть, далее, при реализации схемы единственного деления метода Гаусса для этой системы новые ненулевые элементы не появляются.

Тогда для проведения матрицы к треугольному виду программе ТРЕАН требуется не более, чем $(m-1)(10S^2+14S+8)$ операций. Из них $(m-1)(2S^2+1)$ арифметических операций

необходимо для собственно исключения. Остальное приходится на долю вспомогательных арифметических операций с индексами, операторов присваивания и оператора /F арифметический.

Программа метода простой итерации при решении той же системы требует на каждом итерационном шаге не менее, чем $m(6s+13) - 3s(s+1)$ операций. $2(2m-s)(s+1)$ арифметических действий используется для получения следующего приближения и его оценки, а остальные операции являются вспомогательными.

Таким образом, коэффициент K для рассматриваемых программ при сделанных выше допущениях относительно системы, удовлетворяет неравенству

$$K \leq \frac{(m-1)(10s^2 + 14s + 8)}{m(6s+13) - 3s(s+1)}. \quad (12)$$

Полученная оценка существенно отличается от (10). Дело в том, что во многих классах задач, описываемых системами линейных уравнений, при увеличении порядка m среднее число ненулевых элементов в строке матрицы остается неизменным или меняется незначительно. Это, в частности, обычно имеет место в системах узловых уравнений. Но из формулы (12) следует, что в таком случае при увеличении m коэффициент K остается практически постоянным, а следовательно, не меняется и число итераций, требующее того же количества операций, что и программа ТРЕАН.

6. В заключение приведем некоторые данные, полученные при решении конкретных задач с помощью указанных выше программ.

Программы *GNSO* и *SOLVE* были использованы для решения системы уравнений порядка $m=57$ [6]. Потребность в многократном решении этой системы возникла при интегрировании системы нелинейных дифференциальных уравнений не явным методом. В ходе интегрирования на каждом шаге изменялись коэффициенты алгебраической системы, однако струк-

тура расположения ненулевых коэффициентов оставалась неизменной. Таких коэффициентов система содержала 281.

На решение этой задачи на IBM-7094 программа потребовала 7 секунд машинного времени, сформировав при этом 81,4 ФОРТРАН-выражений для программы SOLVE. В матрицах L и U число ненулевых элементов оказалось равным 455. После трансляции программа SOLVE уже непосредственно решала систему для задаваемых на каждом шаге интегрирования значений коэффициентов, затрачивая на каждый отдельный вариант 0,03 секунды. Для сравнения эта же система решалась с помощью программы, также основанной на методе разделения матрицы на треугольные множители, но не учитывавшей разреженности. При этом каждый вариант системы требовал 1,8 секунды машинного времени.

Программы TPEAN и RETPE, а также программы итерационных методов, учитывающие разреженность, были применены для решения контрольной системы узловых уравнений порядка $m = 50$ с 206 ненулевыми коэффициентами.

Программа TPEAN была испытана в двух режимах работы. В первом режиме, который может иметь место при первом решении системы, программа TPEAN, благодаря заведомо плохой нумерации неизвестных, должна была осуществить многочисленные операции по их перенумерации с целью уменьшения числа новых ненулевых элементов. Время работы программы составляло 3,6 секунды, а общее количество ненулевых элементов оказалось равным 316.

Во втором режиме матрица приводилась к треугольному виду без дополнительных операций по вставке новых ненулевых элементов и перенумерации неизвестных. Для этого была использована информация, полученная в первом режиме работы TPEAN. В этом случае затраты машинного времени составили 0,35 секунды.

Программе RETPE для решения системы с треугольной матрицей потребовалось в обоих режимах 0,05 секунды машинного времени.

шинного времени.

Приведенные выше данные, относящиеся к программе ТРЕАН при первом решении системы, указывают на сильную зависимость времени ее работы от нумерации неизвестных и характера расположения в матрице ненулевых элементов. Это время обычно удается значительно уменьшить, если еще на этапе подготовки системы к расчету занумеровать ее неизвестные и уравнения так, чтобы в матрице системы из всяких двух столбцов меньший номер имел столбец с меньшим числом ненулевых элементов.

При решении контрольной суммы узловых уравнений программами, основанными на методах простой итерации и Зейделя, процесс итерирования прекращался, когда относительная погрешность приближенного решения становилась меньше 10^{-6} . На однократное решение системы методом простой итерации потребовалось 235 итераций и 14 секунд, а методом Зейделя - 117 итераций и 6,6 секунд машинного времени. Программы РЕТРЕ и ТРЕАН во втором режиме потратили на ту же задачу 0,4 секунды. Таким образом, чтобы сравняться с ними, программам итеративных методов необходимо было обеспечить сходимость не более, чем за 7 шагов.

Для примера получим соответствующую оценку из формулы (12). В соответствии с условиями задачи положим $m=50$ в числителе $S \approx 5$, так как конечное число ненулевых элементов у программы ТРЕАН равно 316, в знаменателе $S \approx 3$. Тогда $K \leq 10$. Такого результата для итеративных методов без применения специальных способов ускорения сходимости можно достичь только за счет выбора начальных приближений, очень близких к решению. На практике же такой выбор удается осуществить весьма редко.

Литература

1. Клеев В.В., Коковкин-Шербак Н.И. О минимизации числа арифметических операций при решении линейных алгебраических систем уравнений. - Журнал вычислит. математики и мат. физики, 1965, т.5, № I, с.21-33.
2. Селу С., Рид м.Б. Линейные графы и электрические цепи. М., Высшая школа, 1971, 448с.
3. Фадеев Д.И., Фадеева В.Н. Вычислительные методы линейной алгебры. М.-Л., Физматгиз, 1963, 656с.
4. Коллатц Л. Функциональный анализ и вычислительная математика. М., Мир, 1969, 447с.
5. Idleman T.E. SLIC-A simulator for linear integrated circuits. - IEEE Journ. of solid-state circuits, 1971, v.6, No.4, p.188-203.
6. Gustavson F.G., etc. Symbolic generation of an optimal Crout algorithm for sparse systems of linear equations. - Journal of the ACM, 1970, v.17, No.1, p.87-109.

ВЫБОР ШАГА ПРИ РЕШЕНИИ ЖЕСТКИХ СИСТЕМ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ ЯВНЫМИ МЕТОДАМИ

А.Н.Заворин, Е.В.Мамонтов
Вычислительный центр ЛГУ им. П.Стучки,
Рижский НИИ микроприборов

Решение жестких систем обыкновенных дифференциальных уравнений явными методами требует большого числа шагов из-за ограничения на величину шага интегрирования h вида

$$h < \frac{\alpha}{|\lambda|}, \quad (I)$$

где λ - наибольшее по модулю мгновенное собственное значение системы, α - константа, зависящая от метода. Выполнение этого ограничения обеспечивает устойчивость метода при решении системы. Для метода Рунге-Кутты первого порядка (метода Эйлера) $\alpha=2$, для метода четвертого порядка $\alpha=2,78$. Поэтому большой практический интерес представляет способ увеличения константы α в выражении (I), предложенный в работе [1] для метода Эйлера. Это увеличение достигается за счет интегрирования системы парами шагов, длины которых выбираются специальным образом.

В настоящей работе этот способ изменения шага интегрирования анализируется для метода Эйлера и рассматривается применительно к методам Рунге-Кутты второго и третьего порядка, обладающих значительно более высокой точностью.

I. Метод Эйлера

Пусть система дифференциальных уравнений

$$X' = F(t, X) = AX, \quad X(t_0) = X_0, \quad (2)$$

где X' и X - векторы и A - квадратная матрица порядка m , имеет отрицательные, действительные собственные значения $\lambda_1 \ll \lambda_2 \ll \dots \ll \lambda_m < 0$, $\lambda_1 \ll \lambda_m$.

Система (2) решается методом Эйлера

$$Y_{n+1} = Y_n + hF(t_n, Y_n),$$

где Y_n - вектор приближенного решения в точке $t=t_n$, $Y_0 = X_0$. Шаг интегрирования меняется, причем каждый второй шаг полагается равным

$$h_2 = -\frac{1}{\lambda}, \quad (3)$$

где $\lambda = \lambda_1$ - наибольшее по модулю собственное значение системы (2) (предполагается, что t_0 достаточно велико и при интегрировании с шагом h_2 заданная точность выдерживается). Обозначим через h_1 величину шага, предшествующего h_2 . Тогда формула Эйлера имеет вид

$$Y_{n+2} = (1+h_1A)(1-\frac{1}{\lambda}A)Y_n \quad (4)$$

Предположим для простоты изложения, что матрица A приводится преобразованием подобия к диагональному виду

$$\text{diag}(\lambda_1, \lambda_2, \dots, \lambda_m) = T^{-1}AT$$

Полагая

$$Y_n = TZ_n \quad (5)$$

и умножая (4) слева на T^{-1} , получаем

$$Z_{n+2} = BZ_n, \quad (6)$$

где $B = \text{diag}\left\{\left(1+h_1\lambda_1\right)\left(1-\frac{\lambda_1}{\lambda}\right), \left(1+h_1\lambda_2\right)\left(1-\frac{\lambda_2}{\lambda}\right), \dots, \left(1+h_1\lambda_m\right)\left(1-\frac{\lambda_m}{\lambda}\right)\right\}$. Точное решение системы $X(t) \rightarrow 0$ при $t \rightarrow \infty$, поэтому к нулю должны стремиться также Y_n и Z_n при $n \rightarrow \infty$. Следовательно, величина h_1 должна быть таковой, чтобы выполнялись неравенства

$$\left| \left(1+h_1\lambda_i\right)\left(1-\frac{\lambda_i}{\lambda}\right) \right| < 1 \quad i=1, 2, \dots, m \quad (7)$$

При $i=1$ это неравенство выполняется при любом h_1 , поскольку его левая часть обращается в нуль. Из этих со-

образений и была выбрана величина h_2 .

Каждое из неравенств (7) накладывает свое ограничение на величину h_1 . Мы исследуем наихудший случай, имея в виду, что выполняется неравенство

$$\lambda < \lambda_i < 0. \quad (8)$$

Раскрывая модуль в выражении (7), получаем неравенство

$$1 - \frac{\lambda_i}{\lambda} + h_1 \lambda_i - h_1 \frac{\lambda_i^2}{\lambda} > -1$$

(второе неравенство выполняется при любом $h_1 > 0$ вследствие (8)). Положим теперь $k = \frac{\lambda_i}{\lambda}$, тогда из последнего неравенства получаем

$$h_1 < \frac{k-2}{\lambda k(1-k)}, \quad 0 < k < 1$$

Легко видеть, что правая часть последнего неравенства достигает минимума при $k = k^* = 2 - \sqrt{2} \approx 0,58579$ и это минимальное значение равно $-\sqrt{2}/(3\sqrt{2}-4)\lambda \approx -5,62843 \frac{1}{\lambda}$. Для всех других значений k ограничение на величину менее жесткое.

Получим теперь неравенство для средней величины шага интегрирования $h_{cp} = \frac{h_1 + h_2}{2}$

$$h_{cp} < \frac{k^2 - 2}{2\lambda k(1-k)} \quad (9)$$

В наихудшем случае, т.е. при $k = k^*$, имеем

$$h_{cp} < -\frac{2}{(2-\sqrt{2})\lambda} \approx -3,41421 \frac{1}{\lambda}$$

Полученные результаты отличаются в худшую сторону от результатов, приведенных в работе [1]. Однако и они указывают на существенное увеличение коэффициента α в неравенстве (1) при решении системы уравнений методом Эйлера парами шагов, каждый второй из которых при достаточно большом t полагается равным (3). Это увеличение особенно велико, если спектр собственных значений системы распадается на две группы, одна из которых содержит собственные значения, очень близкие к λ , а другая — очень малые по модулю собственные значения. Действительно, из (9)

следует, что правая часть этого неравенства стремится к бесконечности при $k \rightarrow 1$ и при $k \rightarrow 0$.

Сложнее обстоит дело, если группа больших по модулю собственных значений расположена в достаточно большом интервале. Пусть, например, все большие собственные значения системы (2) находятся в интервале $[-1, -k^*]$, малые - в интервале $[-0,001; -0,004]$ и среди собственных значений имеются $\lambda_1 = -1, \lambda_2 = -k^*, \lambda_m = -0,0001$. Пусть $[0,30000]$ - интервал интегрирования (он определяется обычно величиной, пропорциональной $-\frac{t}{\lambda_m}$).

Для устойчивого интегрирования этой системы обычным образом методу Эйлера потребуется не менее 15000 шагов. Поскольку система имеет собственное значение $\lambda_2 = -k^*$, то интегрирование парами шагов потребует шагов примерно в 1,7 раза меньше. Однако число шагов может быть сокращено, если решать эту систему не парами, а сериями по $z+1$ шагу, из которых $z > 1$ шагов выбираются равными (3).

Действительно, для сохранения устойчивости при интегрировании такой серией должны выполняться неравенства

$$|(1+h_i \lambda_i)(1 - \frac{\lambda_i}{\lambda})^z| < 1, \quad i=1, 2, \dots, m$$

Вводя снова переменную $k = \frac{\lambda_i}{\lambda}$, $0 < k \leq 1$ и раскрывая модуль, приходим к неравенству

$$h_i < -\frac{1+(1-k)^z}{\lambda k(1-k)^z}$$

Обозначим через h_{cp} средний шаг $h_{cp} = \frac{1}{z+1}(h_i - \frac{z}{\lambda})$. Для него получаем

$$h_{cp} < -\frac{1+(1+z)k(1-k)^z}{\lambda k(1+z)(1-k)^z}$$

Легко видеть, что правая часть этого неравенства неограниченно возрастает при каждом фиксированном $k > 0$ если $z \rightarrow \infty$.

Докажем, что точка минимума по k правой части смещается к нулю при $z \rightarrow \infty$. Для этого найдем ее произ-

водную и положим равной нулю. После несложных преобразований получим, что точками минимума правой части могут быть нули полинома

$$P(k) = -(1-k)^{z+1} + k(1+z) - 1$$

Производная этого полинома, равная

$$P'(k) = (z+1)[(1-k)^z + 1],$$

положительна при $0 < k < 1$ и $z \geq 1$. Вместе с этим $P(0) \cdot P(1) < 0$. Следовательно, полином $P(k)$ имеет на рассматриваемом интервале один нуль.

Пусть, далее, k^* - нуль полинома $P(k)$ для некоторого значения z . Тогда $P(k^*)$ для $z+1$, равный $k^*(z+1)$, положителен, что и доказывает наше утверждение.

Из проведенного выше рассмотрения можно сделать следующие практические рекомендации. Если при решении некоторой системы парами шагов устойчивость имеет место при значении коэффициента α в (I), близкого к $3,4$, то система имеет собственное значение, близкое к λk^* . В этом случае можно увеличить число шагов в серии z . При достаточно большом z коэффициент α возрастет из-за смещения к нулю наилучшей точки и из-за возрастания величины n_{cp} .

Легко видеть, что в приведенном выше примере увеличение z приводит к значительному увеличению коэффициента α . Например, при значениях z равных 2, 5 и 9 коэффициент α приблизительно равен соответственно 4, 57; 24, 9 и 200. Значение α при $z=9$ указано для случая, если система имеет собственное значение $\lambda_1 = -0,001$. Если же у системы в интервале $[-0,001; -0,0001]$ нет других собственных значений, кроме $\lambda_m = -0,0001$, то $\alpha \approx 1270$ при $z=9$. Таким образом, число шагов серии, наиболее подходящее для решения системы с двумя группами собственных значений, сильно различающихся по модулю, зависит от этого различия.

Рассмотрим теперь вопрос о влиянии неточности опре-

деления величины λ_1 , а с нею и величины шага h_2 , на коэффициент α в (I) при практическом интегрировании. Этот вопрос весьма важен, поскольку при практическом интегрировании нелинейных жестких систем обыкновенных дифференциальных уравнений трудно получить без больших вычислений достаточно точную оценку для λ_1 .

Обозначим через μ приближенное значение $\lambda = \lambda_1$, и предположим сначала, что $\mu < \lambda$, тогда, поступая совершенно аналогично тому, как и для точного λ , получим ограничение на средний шаг вида

$$h_{cp} < \frac{\rho^2 - 2}{2\mu\rho(1-\rho)},$$

где $\rho = \frac{\lambda_1}{\mu} < 1$, $\lambda < \lambda_1 < 0$. Вводя обозначение $\alpha = \frac{\lambda}{\lambda_1}$; ($\alpha < 1$) получаем

$$h_{cp} < \frac{\kappa^2 \alpha^2 - 2}{2\lambda\kappa(1-\alpha\kappa)},$$

где $\kappa = \frac{\lambda_1}{\lambda}$. Производная правой части неравенства положительна. Поэтому с уменьшением α величина h_{cp} уменьшается и стремится к $-\frac{1}{\lambda}$ при $\alpha \rightarrow 0$. Таким образом, при определенных значениях α коэффициент α в (I) меньше при интегрировании парами шагов, чем при интегрировании обычным образом. При $\kappa = \kappa^*$ это имеет место, если $\alpha < 2 - \sqrt{3}$.

Пусть теперь $\mu > \lambda$. Тогда раскрывая модуль в выражении (7) для $\lambda_1 = \lambda$, получаем два неравенства

$$h_{cp} < \frac{\rho^2}{2\lambda_1(1-\rho)}$$

$$h_1 > \frac{-2+\rho}{\lambda_1(1-\rho)}$$

Первое неравенство показывает, что при $\rho > 2$, т.е. при $\mu > \frac{\lambda_1}{2}$, коэффициент α метода меньше, чем у обычного метода Эйлера. Из второго неравенства следует, что

требование устойчивости при $p > 2$ накладывает на шаг интегрирования ограничение снизу, что вообще говоря недопустимо.

Таким образом, из проведенных рассмотрений следует, что при практическом интегрировании парами шагов приближение для λ необходимо находить достаточно точно, ибо в противном случае средняя величина шага может не превышать или даже быть меньше, чем величина шага у обычного метода Эйлера.

2. Метод Рунге-Кутты второго порядка

Из разностных уравнений (6) видно, что при решении парами шагов методом Эйлера каждый второй шаг h_2 выбирается так (3), чтобы разностное уравнение, соответствующее наибольшему по модулю собственному значению, имело нулевое решение. Эту идею можно попытаться распространить на метод Рунге-Кутты второго порядка

$$Y_{n+1}^* = Y_n + hF(t_n, Y_n) \quad (II)$$

$$Y_{n+1} = Y_n + \frac{h}{2} [F(t_n, Y_n) + F(t_{n+1}, Y_{n+1}^*)]$$

Применяя этот метод для решения системы (2), вводя снова шаги h_1 и h_2 и осуществляя преобразование (5) получим разностные уравнения вида (6). Диагональными элементами матрицы B будут выражения

$$(1 + h_1 \lambda_i + \frac{1}{2} h_1^2 \lambda_i^2)(1 + h_2 \lambda_i + \frac{1}{2} h_2^2 \lambda_i^2), \quad i = 1, 2, \dots, m \quad (I2)$$

Эти уравнения не имеют нулевого решения ни при каких значениях h_2 , так как полиномы в (I2) не имеют действительных корней. Таким образом, идея использованная для метода Эйлера, не подходит в чистом виде для метода (II).

Попытаемся однако найти другим способом значение h_2 , которое в сочетании с h_1 давало бы увеличение коэффици-

ента α в (I). При этом должно выполняться, в частности, неравенство

$$|(1+h_1\lambda + \frac{1}{2}h_1^2\lambda^2)(1+h_2\lambda + \frac{1}{2}h_2^2\lambda^2)| < 1,$$

где $\lambda = \lambda_i$. Обозначим h_1, λ , h_2, λ соответственно через x_1 и x_2 . Пусть $P(x) = 1 + x + \frac{x^2}{2}$, тогда последнее неравенство можно записать в виде

$$P(x_1) \cdot P(x_2) < 1,$$

так как $P(x) > 0$ при любом действительном x , если

$$P(x_1) < \frac{1}{P(x_2)}$$

После элементарных преобразований из этого неравенства получаем

$$|x_1 + 1| < \sqrt{\frac{x_2^2 + 2x_2 - 2}{x_2^2 + 2x_2 + 2}} \quad (13)$$

Выражение в правой части неравенства (13) определено на множестве $[-1 - \sqrt{3}, 0]$, где $x^2 + 2x - 2 < 0$. Обозначим $\frac{1}{2}(x_1 + x_2)$ через x_{cp} . Тогда для x_{cp} из (13) получим, в частности,

$$x_{cp} > \frac{1}{2}(x_2 - 1 - \sqrt{\frac{x_2^2 + 2x_2 - 2}{x_2^2 + 2x_2 + 2}})$$

Легко видеть, что правая часть этого неравенства достигает своего минимума на интервале $[-1 - \sqrt{3}, 0]$ в точке $x_2 = -2$. Этот минимум также равен -2 .

Из проведенных рассмотрений следует, что решение системы дифференциальных уравнений методом (II) парами шагов не дает никакого роста коэффициента α в (I) по сравнению с обычным интегрированием. Действительно, при использовании пар шагов наибольшее значение α равно 2, т.е. равно коэффициенту метода (II) для постоянного шага.

3. Метод Рунге-Кутты третьего порядка

Рассмотрим теперь характер изменения константы α при решении системы (2) парами шагов у метода Рунге-Кутты третьего порядка

$$Y_{n+1/2} = Y_n + \frac{1}{2} h F(t_n, Y_n)$$

$$Y_{n+1}^* = Y_n - h [F(t_n, Y_n) - 2F(t_n + \frac{h}{2}, Y_{n+1/2})] \quad (14)$$

$$Y_{n+1} = Y_n + \frac{h}{6} [F(t_n, Y_n) + 4F(t_n + \frac{h}{2}, Y_{n+1/2}) + F(t_n + h, Y_{n+1}^*)]$$

Легко видеть, что константа α этого метода для постоянного шага равна

$$\alpha = 2,5127 \quad (15)$$

При решении системы (2) с парой шагов h_1, h_2 после замены переменной (5) метод (14) эквивалентен системе разностных уравнений (6). Диагональные элементы матрицы B в этом случае равны

$$(1 + h_1 \lambda_i + \frac{h_1^2 \lambda_i^2}{2} + \frac{h_1^3 \lambda_i^3}{6})(1 + h_2 \lambda_i + \frac{h_2^2 \lambda_i^2}{2} + \frac{h_2^3 \lambda_i^3}{6}), \quad i=1, 2, \dots, m \quad (16)$$

Как и в случае метода Эйлера в качестве h_2 выберем значение, обращающее в нуль правую скобку в (16) при $i=1$. Таким образом разностное уравнение в (6), отвечающее наибольшему по модулю собственному значению системы (2) будет иметь нулевое решение. Легко видеть, что

$$h_2 = \frac{S}{\lambda} \approx -\frac{4,596}{\lambda}, \quad \lambda = \lambda_1 \quad (17)$$

Устойчивость метода (14) при решении системы (2) имеет место, если выполняются неравенства

$$|(1 + h_1 \lambda_i + \frac{h_1^2 \lambda_i^2}{2} + \frac{h_1^3 \lambda_i^3}{6})(1 + h_2 \lambda_i + \frac{h_2^2 \lambda_i^2}{2} + \frac{h_2^3 \lambda_i^3}{6})| < 1, \quad i=1, 2, \dots, m \quad (18)$$

Поскольку величина h_2 выбрана (17), то исследуем, какой может быть величина h_1 , чтобы выполнялись все неравенства (18). Для этого раскроем модуль в выражении (18), подставив в него (17). Предполагая, что $\lambda_i \in [\lambda, 0)$, получим одно неравенство

$$x^3 k^3 + 3x^2 k^2 + 6xk + 6 + \frac{6}{b} > 0, \quad (19)$$

где $x = \lambda h$, $k = \frac{\lambda_i}{\lambda}$, $b = \frac{s^3 k^3}{6} + \frac{s^2 k^2}{2} + sk + 1$

Если в неравенстве (19) задать определенное значение $k \in (0, 1)$, то это неравенство определит значение x , которое с точностью до знака представляет собой коэффициент a в неравенстве (1) для соответствующего значения λ_i . Наименьшее значение $x = -4,88657$ достигается при $k = k^* = 0,7975$. В этом наилучшем случае, когда в спектре собственных значений системы есть $\lambda_i = k^* \lambda$, на величину $h_{cp} = \frac{1}{2}(h_1 + h_2)$ накладывается ограничение

$$h_{cp} < -\frac{3,241}{\lambda} \quad (20)$$

Если же $\lambda_i \rightarrow \lambda$, или $\lambda_i \rightarrow 0$, то $h_{cp} \rightarrow \infty$.

Из выражений (10), (15), (20) следует, что метод (14) при решении жесткой устойчивой системы (2) каррами шагов имеет большую константу a , чем при использовании постоянной шага, но несколько меньшую, чем у метода Эйлера для пары шагов. Учитывая, однако, что при малых шагах интегрирования метод (14) обеспечивает более высокую точность, чем метод Эйлера, он может быть более предпочтительным, по крайней мере, на начальном отрезке интервала интегрирования.

Литература

1. Моренков И.И., Джурунтаев Д.З. Алгоритм выбора шага интегрирования системы дифференциальных уравнений при малом анализе электронных схем. — Микроэлектроника, 1974, вып. 3, № 3, с. 231-234.

ИССЛЕДОВАНИЕ МЕТОДОВ РЕШЕНИЯ ЗАДАЧИ
СТАТИЧЕСКОГО АНАЛИЗА ЭЛЕКТРОННЫХ
СХЕМ НА ТЕСТОВЫХ ФУНКЦИЯХ

Е.С.Кельман

Вычислительный центр ЛГУ им.П.Стучки

1. Введение

Известно, что задачу статического анализа (ЗСА) электрической цепи можно свести к нахождению такого минимума функции $f(x)$, что $f(\tilde{x})=0$. Причем, функция f имеет вид

$$f(x) = \sum_{i=1}^n f_i^2(x), \quad (1)$$

где $f, f_i: R^n \rightarrow R$. Эффективное применение к ЗСА описываемых в литературе методов невозможно без предварительной доработки их на тестовых функциях, тестовых и в практических ЗСА.

В настоящей работе описываются численные эксперименты с тремя методами — методом Ньютона для решения уравнений (МНУ), методом Флетчера-Пауэлла [4] (МФП) и методом сопряженных градиентов (МСГ) — на 10 тестовых функциях. Следуя условиям наших практических задач, мы руководствовались следующим.

1. Любое вычисление функции f занимает значительно больше времени, чем, например, вычисление норм нескольких векторов. Поэтому в методах использовались различные сравнения, чтобы уменьшить число вычислений f .

Вычисление всех f_i ($i=1, \dots, n$) при одном значении x будем также называть одним вычислением функции.

2. Нет программы, вычисляющей градиент σf или производные функций f_i по аналитическим формулам. В [5] по-

казано, что если функция имеет вид (I), то ошибка при вычислениях j -ой составляющей градиента по формуле

$$\frac{\partial f(x)}{\partial x^{(j)}} = 2 \sum f_i(x) \frac{f_i(x+\Delta_j) - f_i(x)}{\|\Delta_j\|} \quad (2)$$

на порядок меньше ошибки формулы

$$\frac{\partial f(x)}{\partial x^{(j)}} \approx \frac{f(x+\Delta_j) - f(x)}{\|\Delta_j\|} \quad (3)$$

Вектор-приращение Δ_j в этих формулах имеет лишь j -ую ненулевую координату. Эта координата положительна и полагалась в наших экспериментах равной

$$\|\Delta_j\| = \begin{cases} 10^{-6}, & \text{если } |x^{(j)}| \leq 1, \\ 10^{-6} \cdot |x^{(j)}|, & \text{если } |x^{(j)}| > 1. \end{cases} \quad (4)$$

Легко видеть, что на вычислении градиента по обеим формулам тратится n вычислений функции (считая, что f и все f_i в точке x известны).

Поскольку на каждом шаге ММ и МСГ требуется вычисление градиента и поиск минимума по направлению, естественно попытаться тратить мало вычислений функций на поиски по направлениям, не увеличив из-за этого существенно число шагов. Нужно, следовательно, сравнить варианты методов, отличающиеся поиском минимумов по направлениям.

В наших экспериментах считалось необходимым там, где это возможно, проводить сравнения на одинаковых участках траекторий минимизации (траекторий, соединяющих точки последовательных приближений). На этой основе выяснялись и устранялись причины медленного движения методов по траекториям. Использование же критериев [1, гл.5], без учета

траекторий - недостаточно, т.к. оценки методов получаются случайными, противоречивыми или неубедительными.

2. Методы

Маленькими латинскими буквами с индексами будем обозначать векторы-столбцы, верхним индексом T - транспонирование.

Метод Флетчера-Пауэлла применяется в форме [4], основанной на [3]. Основные операции, реализующие один шаг метода, следующие.

1. Вычисляем градиент g_i функции f в точке x_0 .
2. Полагаем $i=0$, $H_0=I$ - единичной матрице.
3. Вычисляем i -ое направление поиска минимума, как $S_i = -H_i g_i$.
4. Если не выполняется условие

$$(g_i, S_i) < 0, \quad (5)$$

обновляем метод; полагаем $x_0 = x_i$, $g_0 = g_i$ и переходим к оп.2. Если выполняется, вычисляем $i+1$ -ое приближение к точке минимума $x_{i+1} = x_i + \alpha_i S_i$. Вычисление α_i называется поиском минимума по направлению S_i и сводится к приближенному нахождению точки минимума функции

$$\varphi_i(\alpha) = f(x_i + \alpha S_i), \alpha > 0. \quad (6)$$

5. Вычисляем градиент g_{i+1} в точке x_{i+1} и полагаем $y_i = g_{i+1} - g_i$.
6. Матрицу, соответствующую точке x_{i+1} вычисляем, как

$$H_{i+1} = H_i + \frac{(\alpha_i S_i)(\alpha_i S_i)^T}{(\alpha_i S_i, y_i)} - \frac{(H_i y_i)(H_i y_i)^T}{(H_i y_i, y_i)}$$

7. Полагаем $i = i + 1$ и переходим к оп.3. Условие (5) является достаточным условием существования $\alpha_i > 0$ в оп.4. Это же условие проверяется в МСТ.

Опишем один шаг метода сопряженных градиентов.

1. Полагаем $i = 0$, вычисляем градиент g_0 функции f в точке x_0 . Полагаем $S_0 = -g_0$ и переходим к оп.5.

2. Вычисляем градиент g_i в точке x_i .

3. Вычисляем i -ое направление поиска минимума

$$S_i = -g_i + \beta_i S_{i-1}, \quad \text{где}$$

$$\beta_i = \begin{cases} \frac{(g_i, g_i)}{(g_{i-1}, g_{i-1})} & , \text{ если } i \neq kl, \\ 0 & , \text{ если } i = kl, \quad k = 1, 2, 3, \dots \end{cases} \quad (7)$$

По формуле (7) через каждые l шагов метод обновляется: направление поиска полагается равным антиградиенту.

4. Если не выполняется (5), полагаем $i = 0, x_0 = x_i, g_0 = g_i, S_0 = -g_0$.

5. Вычисляем x_{i+1} как в оп.4 ММ.

6. Полагаем $i = i + 1$ и переходим к оп.2. В литературе [6] принимается $l = n$ в (7), предлагаются другие формулы вычисления β_i . В наших экспериментах использовались только формула (7), в качестве l брались $l = n, l = n + 1, l = n + 3$ и др.

В программе ММ, учитывая, что матрицы H_i симметрические, резервируется место для половины (с диагональю) квадратной действительной матрицы размерности n и шести действительных векторов; в программе МСТ — для четырех действительных векторов.

На наш взгляд эффективность этих методов в решающей степени зависит от программы поиска минимума по направлению. Приведем ее основные операции.

1. Вычисляем величину пробного шага вдоль луча $\alpha > 0$.

2. Выделяем интервал, содержащий точку минимума α_i функции (6).
3. Сужаем этот интервал до величины ε_i .
4. На полученном интервале функция (6) интерполируется параболой.
5. В качестве α_i выбирается α , в котором значении функции (6) было наименьшим.

Отметим, что, уменьшив ε_i в оп.3, можно получить минимум по направлению с любой точностью. Однако это изменялось редко из-за большого числа вычислений функции. На оп.4 требуется одно вычисление функции, а в оп.1,2 - параметры подобраны так, чтобы, во-первых, получать первый минимум на луче $\alpha > 0$ (вообще говоря, локальный), а во-вторых, тратить на это обычно 2 + 4 вычисления функции.

Учитывая (I), поиск минимума \check{x} функции f можно свести к решению системы уравнений

$$F(x) = 0 \quad , \text{ где } F(x) = (f_1(x), f_2(x), \dots, f_n(x))^T.$$

Пусть вектор Δx мал, $F'(x)$ обозначает квадратную матрицу

$$F'(x) = \left(\left(\frac{\partial f_i(x)}{\partial x^{(j)}} \right) \right), \quad i, j = 1, 2, \dots, n. \quad (8)$$

Тогда выполняется

$$F(x + \Delta x) \approx F(x) + F'(x) \cdot \Delta x. \quad (9)$$

Предположим далее, что $x + \Delta x$ есть точка минимума. Тогда $F(x + \Delta x) = 0$, и из (9) получим

$$F'(x) \cdot \Delta x = -F(x). \quad (10)$$

Теперь сформулируем операции одного шага метода Ньютона. I. Полагаем $i = 0$

2. Вычисляем функцию F и матрицу F' в точке x_i .
3. Вычисляем Δx_i , решая систему линейных уравнений (10) при $x = x_i$.
4. Следующее приближение к точке минимума

$$x_{i+1} = x_i + \Delta x_i.$$

Полагаем $i = i + 1$ и переходим к оп.2.

Описанный таким образом метод Ньютона соответствует частному случаю метода наименьших квадратов Гаусса [1, с.235]. Широко известно преимущество этого частного случая - вычисление следующих приближений требует решения системы линейных уравнений, а не обращения матрицы.

Матрицы $F'(x_i)$ вычислялись по конечноразностной формуле. Векторы-приращения Δ_j выбирались (4). Чтобы вычислить столбец матрицы $F'(x_i)$, надо вычислить все функции f_i в точке $x_i + \Delta_j$ - это эквивалентно одному вычислению функции. Таким образом на одном шаге МНУ требуется $n+1$ вычислений функции - все на оп.2. В программе МНУ резервируется место для действительной матрицы $n \times (n+1)$ и трех действительных векторов.

Отметим, что хотя МНУ нашел точки минимума тестовых задач, применение его в практических ЭСА невозможно без дополнительных операций, предохраняющих от расходимости.

3. Тестовые функции

В настоящем разделе приводятся десять тестовых функций, заимствованных из [1], [2]. Римскими цифрами нумеруются начальные приближения. В этих источниках имеются некоторые характеристики тестовых функций: другие точки минимумов, линии уровней, седловые точки и др.

I. Quadratic function 1.

$$f = (x^{(1)})^2 + 100(x^{(2)} - 1)^2 + (x^{(3)} - 2)^2,$$

$$n = 3, \check{x} = (0; 1; 2)^T, f(\check{x}) = 0.$$

$$I x_0 = (3; 2; 1)^T, f(x_0) = 110, [2],$$

$$II x_0 = (-10; 10; -10)^T, f(x_0) = 8544, [2],$$

$$III x_0 = (100; 0; 0)^T, f(x_0) = 10104, [2].$$

2. Quadratic function 2.

$$f = (x^{(1)} + x^{(2)} - 2)^2 + 10^4(x^{(1)} - x^{(2)})^2,$$

$$n = 2, \check{x} = (1; 1)^T, f(\check{x}) = 0.$$

$$I x_0 = (100; 0)^T, f(x_0) \approx 0.10001 \cdot 10^8, [2],$$

$$II x_0 = (10; 10.001)^T, f(x_0) = 324.05, [2],$$

$$III x_0 = (-10; 10)^T, f(x_0) \approx 0.40000 \cdot 10^8, [2].$$

3. Quadratic function 3.

$$f = (x^{(1)} + 2x^{(2)} + 3x^{(3)})^2 + 100(x^{(2)} - 1)^2 + (x^{(3)} - 2)^2,$$

$$n = 3, \check{x} = (-8; 1; 2)^T, f(\check{x}) = 0.$$

$$I x_0 = (3; 2; 1)^T, f(x_0) = 201, [2],$$

$$II x_0 = (-10; 10; -10)^T, f(x_0) = 8644, [2],$$

$$III x_0 = (100; 0; 0)^T, f(x_0) = 10104, [2].$$

4. Rosenbrock's function.

$$f = 100 [x^{(2)} - (x^{(1)})^2]^2 + (1 - x^{(1)})^2,$$

$$n = 2, \check{x} = (1; 1)^T, f(\check{x}) = 0.$$

$$I x_0 = (-1.2; 1)^T, f(x_0) = 24.2, [2],$$

$$II x_0 = (-2.547; 1.489)^T, f(x_0) \approx 0.25108 \cdot 10^4, [1, c. 222].$$

5. Beales's function

$$f = \sum_{i=1}^3 \{c_i - x^{(1)} [1 - (x^{(2)})^i]\}^2,$$

$$c_1 = 1.5, c_2 = 2.25, c_3 = 2.625, n = 2, \check{x} = (3; 0.5)^T, f(\check{x}) = 0,$$

$$I x_0 = (0; 0)^T, f(x_0) \approx 14.203, [2],$$

$$II x_0 = (8; 0.8)^T, f(x_0) = 2.0427, [1, c. 222],$$

$$III x_0 = (2, 2.2)^T, f(x_0) = 0.52978, [1, c. 478] (\text{задача 35}).$$

6. Powell's function No.1.

$$f = (x^{(1)} + 10x^{(2)})^2 + 5(x^{(3)} - x^{(4)})^2 + (x^{(2)} - 2x^{(3)})^2 + 10(x^{(1)} - x^{(4)})^4,$$

$$n = 4, \check{x} = (0; 0; 0; 0)^T, f(\check{x}) = 0,$$

$$x_0 = (3; -1; 0; 1), f(x_0) = 245, [2].$$

7. Powell's function No.2.

$$f = [1 + (x^{(1)} - x^{(2)})^2]^{-1} + \sin\left(\frac{1}{2}\pi x_2 x_3\right) + \exp\left\{-\left(\frac{x^{(1)} + x^{(2)}}{x^{(2)}} - 2\right)^2\right\}.$$

В таком виде в точке $\check{x} = (1; 1; 1)^T$ функция имеет максимум и $f(\check{x}) = 3$. Мы запрограммировали функцию, преопределив f как $f = -f + 3$.

Тогда $f(\check{x}) = 0$, и при $x_0 = (0; 1; 2)^T$ имеем $f(x_0) = 1.5$ [2].

8. "Cube".

$$f = 100 [x^{(2)} - (x^{(1)})^2]^2 + (1 - x^{(1)})^2,$$

$$n = 2, \check{x} = (1; 1)^T, f(\check{x}) = 0,$$

$$I \check{x}_0 = (-1.2; 1)^T, f(x_0) = 749.04 [2],$$

$$II \check{x}_0 = (0.248; -3.082)^T, f(x_0) \approx 959.86 [1, c. 2.22].$$

9. Helical Valley.

$$f = 100 [(x^{(3)} - 10\theta)^2 + (r - 1)^2] + (x^{(2)})^2$$

$$\theta(x^{(1)}, x^{(2)}) = \begin{cases} \frac{1}{2\pi} \arctg\left(\frac{x^{(2)}}{x^{(1)}}\right), & x_1 > 0, \\ \frac{1}{2\pi} (\arctg\left(\frac{x^{(2)}}{x^{(1)}}\right) + \pi), & x_1 < 0, \end{cases}$$

, где

$$r(x^{(1)}, x^{(2)}) = [(x^{(1)})^2 + (x^{(2)})^2]^{\frac{1}{2}},$$

$$n = 3, \check{x} = (1; 0; 0)^T, f(\check{x}) = 0,$$

$$x_0 = (-1; 0; 0)^T, f(x_0) = 2500, [2].$$

10. Wood's function:

$$f = 100 [x^{(2)} - (x^{(1)})^2]^2 + (1 - x^{(1)})^2 + 90 [x^{(4)} - (x^{(3)})^2]^2 + (1 - x^{(3)})^2 + 10.1 \{ (x^{(2)} - 1)^2 + (x^{(4)} - 1)^2 \} + 12.8 (x^{(2)} - 1)(x^{(4)} - 1),$$

$$n=4, \bar{X} = (1; 1; 1; 1)^T, f(\bar{X}) = 0,$$

$$X_0 = (-3; -1; -3; -1)^T, f(X_0) = 19192, [2].$$

Функции *Beale*, *Pow2* и *Wood* при программировании не были представлены в виде (1), и поиск их минимумов методом Ньютона не осуществлялся. Градиенты для функций *Pow2* и *Wood* вычислялись по формуле (3), для остальных функций - по формуле (2).

4. Численные эксперименты

Эксперименты проводились на ЭВМ ДУИИ-400. На ней числа, представленные с обычной точностью, содержат 11 значащих десятичных цифр. Результаты экспериментов приводятся в приложении, в табл. 1-8. Использовались следующие условные обозначения. Целое число в скобках - количество шагов (или номер приближения к минимуму). Целое число без скобок - число вычислений функций. Действительное число, записанное с порядком, как константа языка ФОРТРАН - значение функции. Константой .0 обозначались значения функций, меньше машинного нуля (10^{-77}). Знак * использовался для шагов, на которых по всем координатам приближения достигалось пять точных значащих цифр. Точнее: отличие по каждой координате удовлетворяло условию аналогичному (4), только с множителем 10^{-5} .

Если, используя МП и МТ, при поиске минимума по i -му направлению полагать $\epsilon_i = 10^{-7}/15, n$, то норма ошибки поиска будет не больше 10^{-7} . При таком использовании методов (табл. 1, табл. 5) на каждый шаг тратилось 30 + 40 вычислений функций, главным образом на сужение. Среди вариантов МТ несколько лучшим оказался вариант с обновлением через $n+1$ шаг ($L = n+1$). Его результаты приводятся в табл. 5 + 7.

Сравнивая табл. 1 с табл. 6, табл. 3 с табл. 7, видно, что

МСТ по числу шагов "отстает" от МЭП. Оказывается, траектории минимизации методов близки и МСТ движется по ним "медленнее". Исключением является случай "Cube" - I в табл.3. В нем из-за неточного вычисления x_1 , направление S_1 , неудачно: приближение $x_2 \approx (0.11; 0.0045)^T$, тогда как при точном вычислении x_1 , приближение $x_2 \approx (0.81; 0.54)^T$ и значительно ближе к точке минимума $\bar{x} = (1; 1)^T$. В дополнение к табл.7 укажем, что МСТ при минимизации функции Pow 1 достиг трех точных значащих цифр за 60 шагов, 913 вычислений функций, $f = 0.56 \cdot 10^{-11}$. На функции Wood траектории МЭП и МСТ различны, и, хотя МСТ достигает двух точных значащих цифр на 19-ом шаге (290 вычислений функций, $f = 0.66 \cdot 10^{-3}$), четырех точных значащих цифр нет и на последнем 56-ом шаге (853 вычисления функций, $f = 0.32 \cdot 10^{-7}$).

Для примера движения по траектории минимизации выbran МЭП на функции Rosenbrock (табл.4). Точки I + 6 расположены вдоль "дна оврага" функции. В качестве точек No 0 приводятся начальные приближения. В качестве точки No 1 при минимизации с первого начального приближения было взято второе приближение к минимуму, при минимизации со второго начального приближения - приближение, ближайшее к точке No 1 первого приближения.

Точки No 2 - приближения, ближайшие к точке с координатами (-0.5; 0.25), точки No 3 - приближения, ближайшие к точке с координатами (0.; 0.), точки No 4 - приближения, ближайшие к точке с координатами (0.5; 0.25), точки No 5 - приближения. точки минимума с одной точной значащей цифрой, точки No 6 - с пятью точными значащими цифрами.

При поиске минимумов тестовых функций МЭП значительно превосходил МСТ. Для получения точек минимумов квадратичных функций с пятью точными значащими цифрами ему потребовалось 1-2 шага, результаты для неквадратичных функций приведены в табл.8.

5. Заключение

Все три метода - МФП, МСГ и МНУ - могут быть успешно использованы для поиска минимумов, когда производные вычисляются по конечноразностным формулам. На тестовых функциях МФП оказался несколько лучше МСГ, а МНУ значительно превосходил МФП.

Проведенные исследования позволили улучшить методы, применяемые в системе КОМПРОМИС.

Литература

1. Химмельблау Д. Прикладное нелинейное программирование. М., Мир, 1975, 540с.
2. Greenstadt J. A Quasy-Newton Method with No Derivatives. -Math.of Comp., 1972, v.6, No.117, p.145-166.
3. Fletcher R., Powell M.J.D. A Rapidly Convergent Descent Method for Minimisation. -Comp.J., 1963, 6, 2, p.163-168.
4. Кельман Е.С. Модифицированный алгоритм Флетчера-Пауэлла. Алгоритмы и программы. -Информационный бюллетень, ВНИИЦентр, 1973, вып.2, № 51, с.29.
5. Кацнельсон Л.З., Кельман Е.С. Анализ нелинейных электрических цепей по постоянному току методом Флетчера-Пауэлла. -Б кн.: Латв.мат.ежегодник, 14. Рига, Зинатне, 1974, с.36-50.
6. Поляк Б.Т. Метод сопряженных градиентов. -Труды Второй зимней школы по математическому программированию и смежным вопросам, 1969, вып.1, с.152-201.

Приложение. Таблицы I + 8.

Таблица I.

Число шагов до получения приближений с пятью точными значащими цифрами и значения функций от них при минимизации МП. При поиске минимума по i -му направлению $\epsilon_i = 10^{-7}/15,1$

Rosenbrock		Beale	
I	II	I	II
(18) .70-19	(23) .24-10	(8) .72-11	(13) .26-12

Beale	Pow 1	Pow 2	"Cube"
III			I
(7) .56-11	(29) .12-18	(9) .15-9	(13) .55-11

"Cube"	Helical Valley	Wood
II		
(15) .89-12	(20) .12-9	(42) .20-8

Таблица 2.

Число шагов, число вычислений и значения квадратичных функций от приближений с пятью точными значащими цифрами при минимизации МП. При поиске минимума по i -му направлению сужение не проводилось (ϵ_i достаточно большое число)

Quadratic 1			Quadratic 2			Quadratic 3		
I	II	III	I	II	III	I	II	III
(2) 15 .33-11	(3) 22 .65-11	(4) 30 .79-21	(4) 25 .37-15	(3) 26 .43-12	(2) 13 .22-5	(4) 28 .16-18	(4) 29 .32-17	(4) 28 .17-12

Таблица 3

Число вычислений и значения неквадратичных функций при минимизации МЭП. При поиске минимумов по направлениям сужение не проводилось (ε_1 достаточно большое число)

№ шага	Rosenbrock		Beale			Pow 1	Pow 2	"Cube"		Helical Valley	Wood
	I	II	I	II	III			I	II		
0	1 .24+2	1 .25+4	1 .14+2	1 .20+1	1 .53+0	1 .21+3	1 .15+1	1 .75+3	1 .96+3	1 .25+4	1 .19+5
5	34 .19+1	39 .52+1	29 .92-3	40 .21+0	28 .96-4	39 .14-2	32 .39-2	44 .54+0	33 .14+0	33 .11+2	43 .20+2
10	68 .32+0	97 .18+1	59 .53-22	120 .19-7	120 *152	96 .10-5	-	104 .10+0	66 .42-2	76 .15+1	80 .49+1
15	102 .57-7	146 .10+0	-	152 .53-22	179 *152	179 .16-7	-	153 .85-3	98 .66-5	119 .14-1	126 .35+1
20	*132 1.0-13	177 .12-1	-	-	241 *152	241 .27-13	-	*184 .86-12	-	*154 .58-15	177 .15+1
25	-	*208 .99-12	-	-	-	-	-	-	-	-	226 .79+0
Послед- ний шаг	(21) 138 .21-21	(27) 220 .0	(10) 59 :53-22	(15) 152 :53-22	(9) 52 :11-20	(23) 269 :84-18	(9) 64 :15-10	(21) 190 :84-18	(18) 116 :16-18	(21) 160 :57-20	(44) 436 :43-9

Примечание 1. У функции Pow 1 в последнем приближении x_{23} одна координата имеет четыре точные значащие цифры точки минимума, остальные - пять.

Примечание 2. У функции Wood приближения x_{40} - x_{44} мало различаются. В них четыре точные значащие цифры точки минимума у одной координаты, пять - у остальных.

Таблица 4

Минимизация функции Розенброка ММН. При поиске минимумов по направлениям сужение не проводилось (ε , достаточно большое число)

№ точ-ки	Число шагов, число вычислений и значения функции	Координата		Число шагов, число вычислений и значения функции	Координата	
		1	2		1	2
0	(0) 1 .24+ 2	-1.20	1.00	(0) 1 .25 + 4	-2.55	1.49
1	(2) 13 .38+ 1	-.93	.84	(6) 52 .43+ 1	-.91	.75
2	(4) 27 .25+ 1	-.52	.23	(8) 77 .27 + 1	-.51	.20
3	(6) 41 .14+ 1	-.10	-.031	(11) 110 .11 + 1	.057	-.043
4	(10) 68 .32+ 0	.51	.23	(13) 131 .28 + 0	.52	.24
5	(15) 102 .57+2	.96	.91	(21) 185 .17 - 2	.99	.97
6	(20) 132 1.00-13	1.00	1.00	(25) 208 .99 -12	1.00	1.00

Таблица 5

Число шагов до получения приближений с пятью точными значащими цифрами и значения функций от них при минимизации МСГ. Обновления через $n+1$ шаг. При поиске минимума по i -му направлению выполнялось условие $\epsilon_i = 10^{-1}/\|s_i\|$

Rosenbrock		Beale		Pow 2	"Cube"		Hel, V.
I	II	I	II		I	II	
(26)	(29)	(11)	(14)	(13)	(14)	(20)	(31)
.48-8	.17-9	.37-14	.17-9	.73-10	.41-13	.17-8	.72-15

Таблица 6

Число шагов, число вычислений и значения квадратичных функций от приближений с пятью точными значащими цифрами. МСГ с обновлением через $n+1$ шаг. При поиске минимума по i -му направлению выполнялось условие $\epsilon_i = 10^{-1}/\|s_i\|$

Quadratic 1			Quadratic 2			Quadratic 3		
I	II	III	I	II	III	I	II	III
(2) 34	(4) 67	(4) 70	(3) 60	(4) 62	(3) 45	(5) 77	(7) 98	(7) 117
.69-11	.34-17	.13-14	.44-11	.20-9	.10-17	.47-8	.23-15	.14-11

Таблица 7

Число шагов, число вычислений и значения функций от приближений с пятью точными значащими цифрами. МСГ с обновлением через $n+1$ шаг. При поиске минимума по i -му направлению выполнялось условие $\epsilon_i = 10^{-1}/\|s_i\|$

Rosenbrock		Beale		Pow2	"Cube"		Hel, V.
I	II	I	II		I	II	
(29) 324	(29) 340	(11) 121	(20) 255	(12) 147	(11) 121	(23) 261	(27) 348
.52-14	.20-8	.53-14	.58-9	.15-10	.23-9	.12-11	.13-12

Таблица 8.

Число вычислений и значения функций при поиске минимумов МНУ:

№ ша-га	Rosenbrock		Pow 1	"Cube"		Hel.V.	№ ша-га	Pow 1 продол-жение	Hel.V. про-долже-ние
	I	II		I	II				
0	1	1	1	1	1	1	8	41	33
	.24+ 2	.25+4	.21 +3	.75+ 3	.96+ 3	.25+ 4		37- 7	13- 4
1	4	4	6	4	4	5	10	51	41
	.23+ 4	.16 +5	.10 +2	.46+ 4	.72+ 4	.14+ 4		15- 9	13-27
2	7	7	11	7	7	9	12	61	-
	.53- 9	.71 -8	.63 +0	.27- 8	.63- 9	.28+ 4		57-12	
3	10	10	16	10	10	13	14	71	
	.13-20	.13-20	.39 -1	.12-19	.12-19	.59+ 3		23-14	
4	-	-	21	-	-	17	16	81	
			.25 -2			.29+ 3		95-17	
6			31			25	17	86	
			.96 -5			.65+ 1		65-18	

Примечание 1. Приближения x_1 для функций Rosenbrock и "Cube":

ROS-I $x_1 \approx (1.00; -3.84)$, ROS-II $x_1 \approx (1.00; -11.58)$
 "Cube" I $x_1 \approx (1.00; 7.78)$, "Cube"-II $x_1 \approx (1.00; .15)$.

Примечание 2. У функции Pow 1 в последнем приближении x_{17} одна координата имеет четыре точные значащие цифры точки минимума, остальные пять.

ИЗОЭКСТРЕМАЛЬНЫЕ ДРОБИ ЧЕБЫШЕВА ДЛЯ
ОПТИМАЛЬНОГО СИНТЕЗА ЭЛЕКТРИЧЕСКИХ ЦЕПЕЙ

Э.Я.Гринберг, Л.З.Кацнельсон, М.М.Неизвестный
Вычислительный центр ЛГУ им.П.Стучки

В статье описаны алгоритмы вычисления изоекстремальных дробей Чебышева, пригодных для решения задач оптимального синтеза электрических цепей.

Дробь Чебышева (типа $\frac{n}{m}$, $n \geq m$) — это несократимая рациональная дробь

$$F(y) = \frac{\alpha_0 y^n + \alpha_1 y^{n-1} + \dots + \alpha_n}{\beta_0 y^m + \beta_1 y^{m-1} + \dots + \beta_m}$$

с действительными коэффициентами, которая достигает своего наименьшего и наибольшего значений на отрезке $[-1, 1]$ попеременно $n+1$ раз, причем эти значения равны соответственно -1 и 1 .

Для задач синтеза используются изоекстремальные дроби Чебышева, т.е. дроби, у которых $m=2r$ и которые, являясь дробями Чебышева, принимают в промежутке $[h, \infty)$ ($h > 1$) максимально возможные значения.

Установлено [3] существование и единственность таких дробей, причем для них выполняется условие изоекстремальности, т.е. в промежутке $[h, \infty)$ существует $r+1$ точка

$$h = \beta_1 < \beta_3 < \dots < \beta_{2r+1},$$

в которых дробь принимает минимальное на $[h, \infty)$ значение

$$F(\beta_1) = F(\beta_3) = \dots = F(\beta_{2r+1}) = \min_{x \in [h, \infty)} F(x) \quad (I)$$

В [1], [2] приведены выражения дробей Чебышева через их полюсы, т.е. корни знаменателя. Изоекстремальные дроби имеют двойные действительные полюсы $\beta_2, \beta_4, \dots, \beta_{2r}$, чередующиеся с экстремальными точками

$$\beta_1 < \beta_2 < \beta_3 < \beta_4 < \dots < \beta_{2r} < \beta_{2r+1}$$

и выражения дробей Чебышева в этом случае имеют вид

$$F(x) = \cos \left[2(n-m) \arctg \sqrt{\frac{1-x}{1+x}} + 4 \sum_{i=1}^r \arctg \sqrt{\frac{(\beta_{2i+1})(1-x)}{(\beta_{2i}-1)(1+x)}} \right] \quad (2)$$

при $-1 \leq x \leq 1$,

$$F(y) = \frac{R^{2(n-m)} C_1^2 \dots C_r^2 S^{2(n-m)} D_1^2 \dots D_r^2}{2^{n+1} (x-\beta_2)^2 (x-\beta_4)^2 \dots (x-\beta_{2r})^2} \quad (3)$$

при $1 \leq x < \infty$. R, S, C_i, D_i — известные функции от $x, \beta_2, \beta_4, \dots, \beta_{2r}$.

Представление (3) вместе с (1) позволяет составить систему уравнений относительно $\beta_2, \beta_4, \dots, \beta_{2r}, \beta_{2r+1}$

$$F(\beta_{2i-1}) = F(\beta_{2i+1}) \quad (4)$$

$i=1, 2, \dots, r$. Величина $\beta_{2r+1} = h$ определена промежутком $[h, \infty)$.

Еще r уравнений получаются из условия обращения в нуль производной $F'(x)$ в точках $\beta_3, \beta_5, \dots, \beta_{2r+1}$. Чтобы выписать эти уравнения, нужно приравнять нулю производную $F'(x)$ соотношения (2) и сократить ее на отличный от нуля множитель. В результате получаются уравнения

$$n-m+4 \sum_{i=1}^r \frac{\mu_i}{(1-\mu_i^2) \beta_{2j-1} + \mu_i^2 + 1} = 0, \quad (5)$$

$$j=1, 2, \dots, r; \quad \mu_i = \sqrt{\frac{\beta_{2i+1}}{\beta_{2i}-1}}$$

Подстановка в (2) или (3) значений $\beta_2 < \beta_4 < \dots < \beta_{2r} < \beta_{2r+1}$, являющихся решением системы (4),

(5), дает при заданных n , m и h искомую изокстремальную дробь Чебышева, т.е. ее коэффициенты a_0, a_1, \dots, a_n , b_0, b_1, \dots, b_m .

Изокстремальные дроби Чебышева типа $\frac{n}{2r} (n-2r=0)$ или 1) - это дроби Золотарева. Дробь типа $\frac{n}{2r} (n-2r=2)$ можно получить путем преобразования дроби Золотарева. Для упомянутых дробей известны явные параметрические представления. Для изокстремальных дробей Чебышева типа $\frac{n}{2r}$ с $n \geq 2r+3$ хорошего способа построения нет. Предлагается следующий алгоритм алгебраического построения изокстремальных дробей Чебышева [4]:

1. Учет по возможности многих свойств искомой функции путем ее представления в рациональном виде с минимальным числом параметров.

2. Составление условий изокстремальности в виде системы \sum алгебраических уравнений между параметрами.

3. Преобразование системы \sum в более простой вид, например, $g(s, b) = 0$, где s, b - параметры, представление других параметров в виде рациональных функций от s и b . Нахождение границ параметров s, b .

4. Получение в наиболее удобном для расчета виде алгебраических соотношений

$$g(h, s, b) = 0, H = H(s, b),$$

где h - частота среза, H - гарантированное затухание в полосе зашумления.

Для выполнения условий 3. и 4. используются системы полиномиальных выкладок на ЭВМ.

На языке ФОРТРАН составлены программы для получения характеристик изокстремальных дробей обоими методами. Параметрами программ являются $a_{n, \max}$ - неравномерность затухания в полосе пропускания, h - частота среза и порядок дроби Чебышева n и m . Рассчитаны полные таблицы, которые позволяют синтезировать оптимальные RC фильтры с числом элементов, не превышающим 12 ($n + \frac{1}{2} m \leq 12$).

По заданным параметрам рассчитываются передаточная функция

$$T(p) T(-p) = \frac{1}{1 + \kappa F(y)}, \quad (y = -p^2, \kappa = \text{const})$$

нули и полюсы ее, добротности контуров, H - гарантированное затухание в полосе зашумления. Ограничение на число элементов в фильтра не существенно.

Ниже приводится выборка из полной таблицы рассчитанных передаточных функций.

$$a_{max} = 0,01$$

$$h = \Omega_x = 1,15$$

n	m	a _{min}	нули (Ω _z)	полюсы (ρ _j)		κ ₀ = √(δ ₀ /α ₀)
				Re ρ _j	Im ρ _j	
1	2	3	4	5	6	7
5	4	7,62110	1,090682	2,09157		1,56975
			1,402901	0,0325578	± 1,06812	
6	4	14,0075	1,085859	0,293475	± 1,20524	
			1,280084	0,0314275	± 1,05239	1,05780
				0,210501	± 1,08700	
				0,978337	± 0,740645	
8	4	26,3488	1,081321	0,126232	± 1,01860	0,345332
			1,191529	0,387783	± 0,865953	
				0,0256660	± 1,03422	
				0,664578	± 0,351773	
3	2	0,316057	1,131358	7,60305		7,55396
				0,0245459	± 1,12743	

I	2	3	4	5	6	7
4	2	I.20568	I.I09456	0.0283286 I.4485I	± I.I0I4I ± I.44349	4.II940
5	2	3.26965	I.099632	I.209I9 0.0326924 0.63729I	± I.08447 ± I.I9098	2.I4784
6	2	6.57835	I.093990	0.03489I0 0.8I7935 0.377337	± I.06986 ± 0.454065 ± I.083I9	I.I0I15
7	2	IO.5979	I.0903I0	0.7003II 0.258064 0.573726 0.0344959	± I.03606 ± 0.644667 ± I.05744	0.559886
8	2	I4.843I	I.0877I3	0.42I887 0.0324840 0.564547 0.I9I495	± 0.743574 ± I.04748 ± 0.272000 ± I.0I389	0.283335
7	6	24.8933	I.08I346 I.I86092 I.757I66	I.07457 0.0264370 0.I35596 0.50944I	± I.03687 ± I.03I86 ± 0.9230I8	0.2740I2
8	6	32.5789	I.079493 I.I56948 I.509090	0.I04438 0.33I367 0.0228I90 0.743469	± I.0I383 ± 0.920672 ± I.02930 ± 0.450I04	0.223908

$$\sigma_{max} = 0.01 \text{ дБ}$$

$$h = \Omega_s^* = 1.2$$

n	m	σ_{min}	нули (Ω_i)	полюсы (ρ_j)		$\kappa_0 = \sqrt{\frac{b_2}{a_2}}$
				Re ρ_j	Im ρ_j	
1	2	3	4	5	6	7
3	2	0.448803	1.164245	6.73764 0.0336439	± 1.15797	6.72035
4	2	1.82728	1.138739	0.0400243 1.38795	± 1.12462 ± 1.35617	3.67454
5	2	4.87125	1.127247	1.17326 0.0453177	± 1.10094	1.91891
6	2	9.21829	1.120641	0.631950 0.382701	± 1.14534 ± 1.95641	0.984177
7	2	14.0432	1.116334	0.0461076 0.801252	± 1.08068 ± 0.442389	0.500625
8	2	18.9183	1.113295	0.688218 0.265333	± 1.01910	0.500625
9	2	24.0432	1.110334	0.565962 0.0434656	± 0.632174 ± 1.06466	0.500625
10	2	29.9183	1.107295	0.418486 0.193455	± 0.732281 ± 1.00244	0.253418
11	2	35.8183	1.104295	0.556683 0.0394381	± 0.267674 ± 1.05255	0.253418
12	4	9.94479	1.117034 1.468571	1.84581 0.0431385	± 1.08130	1.23409
13	4	17.3403	1.111340 1.331861	0.348996 0.0398663	± 1.18935 ± 1.06131	0.835661
14	4	27.3403	1.105965 1.232063	0.242221 0.938523	± 1.07438 ± 0.677400	0.835661
15	4	30.6149	1.105965 1.232063	0.141266 0.390803	± 1.01158 ± 0.838638	0.274335
16	4	30.6149	1.105965 1.232063	0.643911 0.0309938	± 0.337083 ± 1.03919	0.274335

2	3	4	5	6	7	
7	6	28.6442	I.106055 I.226710 I.856324	0.992580 0.0324617 0.156367 0.525014	± 1.04307 ± 1.02609 ± 0.872545	0.190365
8	6	36.8726	I.103859 I.193462 I.584913	0.119681 0.347083 0.710266 0.0275861	± 1.00942 ± 0.889981 ± 0.415095 ± 1.03404	0.156438

$$\alpha_{max} = 0.01$$

$$h = \Omega_k^2 = 1.25$$

n	m	α_{min}	нули (Ω_i)	ПОЛЮСЫ (ρ_j)		$K_0 = \sqrt{\frac{\delta_2}{\alpha_0}}$
				Re ρ_j	Im ρ_j	
1	2	3	4	5	6	7
3	2	0.601001	I.195436	6.16907 0.0433052	± 1.18622	6.08246
4	2	2.53402	I.166783	0.0522264 I.34233	± 1.14531 ± 1.28561	3.33181
5	2	6.50347	I.153831	I.14456 0.0573423 0.629633	± 1.11410 ± 1.10936	I.74089
6	2	11.6500	I.146380	0.387937 0.0558516 0.787975	± 1.03595 ± 1.08883 ± 0.432856	0.893581
7	2	17.0618	I.141521	0.678455 0.271508 0.559927 0.0508108	± 1.00637 ± 0.622029 ± 1.07000	0.454675

I	2	3	4	5	6	7
8	2	22.4200	I.136096	0.415989 0.204146 0.550358 0.044889I	± 0.723181 ± 0.993934 ± 0.264138 ± 1.05631	0.230203
5	4	12.0258	I.142513 I.528200	I.68475 0.0528486 0.392037	± 1.09224 ± 1.17009	I.00637
6	4	20.0182	I.136054 I.379513	0.0472396 0.266514 0.908800	± 1.06826 ± 1.06207 ± 0.632074	0.683705
8	4	34.2191	I.129942 I.269928	0.152887 0.392908 0.628024 0.0354335	± 1.00554 ± 0.817869 ± 0.325742 ± 1.04320	0.225304
7	6	31.8104	I.130091 I.264623 I.945028	0.036114 0.172999 0.0376738 0.533446	± 1.02035 ± 1.04821 ± 0.833791	0.139870
8	6	40.4936	I.127594 I.227812 I.653383	0.131925 0.357682 0.0316481 0.685529	± 1.00525 ± 0.865730 ± 1.03797 ± 0.390325	0.115381

$$\alpha_{max} = 0.01$$

$$h = \Omega_x^2 = 1.30$$

n	m	α_{min}	ПОЛЮСЫ (ρ_j)			$K_0 = \sqrt{\frac{\beta_2}{\alpha_0}}$
			нули (Ω_i)	$Re \rho_j$	$Im \rho_j$	
1	2	3	4	5	6	7
3	2	0.77222I	I.225309	5.67882		5.57189
				0.0534654	$\pm I.21254$	
4	2	3.3008I	I.193834	0.0646024	$\pm I.16324$	3.05617
				I.30672	$\pm I.2269I$	
5	2	8.09596	I.179570	I.1208I		I.59787
				0.0684996	$\pm I.12474$	
				0.628902	$\pm I.08013$	
6	2	I3.866I	I.171359	0.392759	$\pm I.01967$	0.820457
				0.0642492	$\pm I.09520$	
				0.777022	± 0.424828	
7	2	I9.74II	I.166005	0.570308		0.417557
				0.276777	± 0.996347	
				0.555055	$\pm 0.61353I$	
				0.05668826	$\pm I.07416$	
8	2	25.5020	I.16223I	0.414075	$\pm 0.7156II$	0.211439
				0.208362	± 0.987286	
				0.545074	± 0.261144	
				0.0493120	$\pm I.05925$	
5	4	I3.8995	I.167257	I.57033		0.841778
			I.583524	0.0617329	$\pm I.10152$	
				0.426006	$\pm I.15030$	
6	4	22.3682	I.160110	0.0537332	$\pm I.07478$	0.573236
			I.424117	0.285824	$\pm I.05063$	
				0.885354	± 0.597628	
7	4	30.1289	I.156011	0.797102		0.339712
			I.349323	0.209650	$\pm I.01500$	
				0.562290	± 0.738017	
				0.0461163	$\pm I.05793$	

1	2	3	4	5	6	7
8	4	37.3660	I.153335 I.30560I	0.162253 0.39445I 0.615268 0.0392319	\pm I.0003I \pm 0.801245 \pm 0.3I66I3 \pm I.04353	0.18942I
7	6	34.5732	I.153537 I.300517 2.026428	0.894232 0.186728 0.042254I 0.538173	\pm I.01487 \pm I.05259 \pm 0.802908	0.106824
8	6	43.652I	I.150768 I.260519 I.71665I	0.142070 0.365292 0.668123 0.0352008	\pm I.00138 \pm 0.845915 \pm 0.371640 \pm I.04128	0.0883643

$$\sigma_{max} = I$$

$$h = \Omega_A^* = I.074 (I.073529...)$$

n	m	σ_{min}	нули (Ω_i)	Полосы (ρ_j)		$\kappa_0 = \sqrt{\frac{b_0}{a_0}}$
				Re ρ_j	Im ρ_j	
1	2	3	4	5	6	7
3	2	6.94.36	I.075937	1.01260 0.0534132	\pm I.01619	0.905779
4	2	11.4090	I.06110I	0.431774 0.0358170	\pm 0.651024 \pm I.00660	0.490085
5	2	15.5727	I.051524	0.403376 0.229777 I.00295	\pm 0.799527 \pm I.0066I	0.254572
6	2	19.4479	I.050755	0.393144 0.146697 0.0208118	\pm 0.352168 \pm 0.863419 \pm I.00129	0.130240

I	2	3	4	5	6	7
7	2	23.0955	I.048294	0.258313		0.0661336
				0.211538	± 0.546814	
				0.0992924	± 0.898622	
				0.0168823	± 1.00048	
8	2	26.5682	I.046554	0.211476	± 0.242206	0.0334381
				0.157690	± 0.664008	
				0.0733111	± 0.920701	
				0.0140757	± 1.00006	
9	2	29.9050	I.045257	0.189607		0.0168612
				0.171389	± 0.409058	
				0.121434	± 0.739867	
				0.0567055	± 0.935745	
				0.0118580	± 0.999828	
10	2	33.1342	I.044250	0.164229	± 0.184853	0.00848692
				0.139932	± 0.527055	
				0.0962187	± 0.791836	
				0.0454664	± 0.946508	
				0.0103479	± 0.999767	
11	2	36.2764	I.043446	0.149812		0.00426639
				0.140955	± 0.325711	
				0.115659	± 0.612989	
				0.0780801	± 0.829089	
				0.0373613	± 0.954637	
				0.00893084	± 0.999677	
12	2	39.3472	I.042787	0.133945	± 0.149537	0.00214265
				0.121077	± 0.434210	
				0.0968555	± 0.677301	
				0.0646737	± 0.856728	
				0.0375279	± 0.960828	
				0.00810085	± 0.999763	

I 12	42.3577	I.042239	0.123831		0.00107532
			0.118887	±0.270281	
			0.104510	±0.519001	
			0.0821212	±0.726615	
			0.0544795	±0.877844	
			0.0272466	±0.965613	
			0.00766186	±0.999994	
I 4	45.3177	I.041774	0.112988	±0.125577	0.000539370
			0.105393	±0.368279	
			0.0907937	±0.586262	
			0.0704436	±0.765223	
			0.0466059	±0.894411	
			0.0237007	±0.969686	
			0.00688287	±0.999987	

$$T(\rho) T(-\rho) = K_0^2 \frac{A_m(\rho) A_m(-\rho)}{B_n(\rho) B_n(-\rho)};$$

$$A_m(\rho) = (\rho^2 + \Omega_1^2) \dots (\rho^2 + \Omega_r^2); \quad m = 2r;$$

при $n = 2t$

$$B_n(\rho) = (\rho + p_1)(\rho + \bar{p}_1) \dots (\rho + p_t)(\rho + \bar{p}_t),$$

при $n = 2t + 1$

$$B_n(\rho) = (\rho + p_0)(\rho + p_1)(\rho + \bar{p}_1) \dots (\rho + p_t)(\rho + \bar{p}_t).$$

Литература

1. Чебышев П.А. Вопросы о наименьших величинах, связанных с приближенным представлением функций.- Полн. собр. соч., т.2, М., Изд-во АН СССР, 1947, с.151-238.
2. Марков А.А. Избранные труды по теории непрерывных дробей и теории функций, наименее уклоняющихся от нуля. М., Гостехиздат, 1948, 376с.
3. Кацнельсон Л.В. Об одном классе рациональных функций.- В кн.: Латв.мат.ежегодник, 6. Рига, Зинатне, 1969, с.107-114.
4. Гринберг Э.Я., Кацнельсон Л.В. Об алгебраическом построении изокстремальных дробей Чебышева.- В кн.: Латв.мат.ежегодник, 3. Рига, Зинатне, 1968, с.89-100.

МЕТОД ОПТИМАЛЬНОГО СИНТЕЗА ПОЛИНОМИАЛЬНЫХ ФИЛЬТРОВ С ОГРАНИЧЕНИЯМИ НА ДОБРОТНОСТЬ

М.М. Неизвестный

Вычислительный центр ЛГУ им. П. Стучки

Коэффициент передачи $H(p)$ четырехполюсника по э. д.с. определяется [1] как отношение входного возбуждения к выходному напряжению

$$H(p) = \frac{E}{U} \quad (1)$$

где E - входная операторная э.д.с., U - выходное операторное напряжение, p - комплексная частота.

Мы рассматриваем полиномиальные фильтры четного порядка $n=2k$. Поэтому $H(p)$ можно представить в виде

$$H(p) = \alpha \prod_{i=1}^k (p^2 + \alpha_i p + \beta_i) \quad (2)$$

с действительными $\alpha, \alpha_i, \beta_i, i \in [1:k]$. Если $H(p)$ имеет действительные корни, то такая запись может не быть единственной.

Поскольку нас не интересует фазочастотная зависимость, мы будем пользоваться амплитудной функцией $F(y)$ которая есть квадрат модуля коэффициента передачи [1].

$$F(y) = H(p)H(-p), \quad y = -p^2 \quad (3)$$

Ее можно представить в виде

$$F(y) = \alpha \prod_{i=1}^k (y^2 - c_{ii} y + c_{ii}), \quad (4)$$

где $\alpha = \alpha^2$, $c_{1i} = 2\beta_i - \alpha_i^2$, $c_{2i} = \beta_i^2$. Если раскрыть скобки, то амплитудная функция есть полином степени n от квадрата частоты $y = -p^2 = \omega^2$

$$F(y) = \alpha y^n + a_1 y^{n-1} + \dots + a_{n-1} y + a_n. \quad (5)$$

Для ее физической реализуемости необходимо и достаточно, чтобы при положительных значениях y , выполнялось условие $F(y) > 0$.

Задача оптимального синтеза фильтра низких частот может быть сформулирована следующим образом.

Среди всех физически реализуемых амплитудных функций $F(y)$ степени n с заданной величиной $\alpha' = \alpha / \min_{y \in [0,1]} F(y)$, где $\alpha > 0$ старший коэффициент $F(y)$, найти ту, для которой величина

$$\frac{\max_{y \in [0,1]} F(y)}{\min_{y \in [0,1]} F(y)} \quad (6)$$

принимает наименьшее значение, т.е. которая имеет наименьшую неравномерность в полосе пропускания.

Решение такой задачи дается полиномами Чебышева.

Мы будем рассматривать аналогичную задачу, но на более узком множестве полиномов, чем множество всех физически реализуемых амплитудных функций.

Пусть заданы функция $\Gamma(\xi, \zeta)$ и константа $q > 0$. Будем рассматривать те амплитудные функции, для которых найдется представление (4), удовлетворяющее условию

$$\begin{aligned} 0 &\leq \Gamma(c_{1i}, c_{2i}), \quad i \in [1:k], \\ \max_{i \in [1:k]} \Gamma(c_{1i}, c_{2i}) &\leq q \end{aligned} \quad (7)$$

Например, если

$$\Gamma(\xi, \zeta) = \frac{\sqrt{\zeta}}{2\sqrt{\zeta} - \xi},$$

то

$$q_i = \Gamma(c_{1i}, c_{2i}) = \frac{\sqrt{c_{2i}}}{2\sqrt{c_{2i}} - c_{1i}}$$

есть квадрат добротности i -го квадратичного множителя в представлении (4), а (7) означает, что задано ограничение на величину добротности.

В сформулированной выше задаче оптимального синтеза фильтра нижних частот минимизация величины

$$\frac{\max_{y \in [a,1]} F(y)}{\min_{y \in [a,1]} F(y)} \quad (6)$$

эквивалентна минимизации величины

$$\max_{y \in [a,1]} F(y) - \min_{y \in [a,1]} F(y). \quad (8)$$

Поэтому при решении этой задачи мы можем считать, что задан старший коэффициент a полинома $F(y)$ и минимизировать

$$\max_{y \in [a,1]} F(y).$$

При наличии ограничения (7), мы не можем таким образом упростить задачу, так как величины

$$\Gamma(c_{1i}, c_{2i}), \quad i \in [1:k]$$

могут изменяться при увеличении полинома на константу по

самым разным законам.

Поэтому введем нормировку следующим образом.

Положим

$$F_m = \min_{y \in [0,1]} F(y) = 1 \quad (9)$$

и рассмотрим задачу минимизации величины

$$\max_{y \in [0,1]} \{F(y) - 1\} \quad (10)$$

на множестве \tilde{H}_q полиномов степени $n = 2k$, удовлетворяющих условиям:

1) старший коэффициент $F(y)$ фиксирован и равен $d > 0$;

2) для всех $y \geq 0$

$$F(y) \geq 1;$$

3) существует представление полинома $F(y)$ в виде

(4)

$$F(y) = d \prod_{i=1}^k (y^2 - c_{1i}y + c_{2i}),$$

для которого выполняется ограничение (7)

$$0 \leq \min_{i \in [1:k]} \Gamma(c_{1i}, c_{2i})$$

$$\max_{i \in [1:k]} \Gamma(c_{1i}, c_{2i}) \leq q,$$

где $q > 0$ - заданное число.

Дальнейшее изложение поведем для случая

$$\Gamma(\xi, \xi) = \frac{\sqrt{\xi}}{2\sqrt{\xi} - \xi}, \quad (11)$$

т.е. будем считать, что задано ограничение на добротность, которая задается для множителей в (2) по формуле

$$Q_i = \frac{\sqrt{A}}{a_i}, \quad i \in [1:k] \quad (12)$$

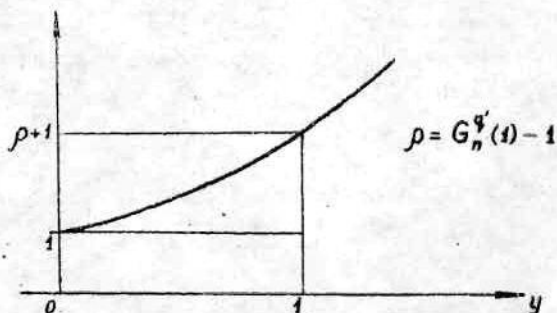
и, значит, для (4)

$$Q_i^2 = \frac{\sqrt{C_{2i}}}{2\sqrt{C_{2i} - C_{1i}}}$$

В этой ситуации множество полиномов, удовлетворяющих условиям 1)-3) не пусто: в нем содержится по крайней мере полином

$$G_n^{q'}(y) = (\sqrt{\alpha} y^n + \sqrt{\alpha} \frac{1-2q'}{q'} y + 1)^k, \quad n=2k, \quad (13)$$

где $0 < q' \leq \min(\frac{1}{2}, q)$, имеющий вид



Обозначим через H_q подмножество в пространстве R_{n+1} , состоящее из точек, координаты которых являются коэффициентами полиномов из \bar{H}_q . Тогда мы будем иметь задачу в минимизации функционала $\Phi: H_q \rightarrow R$

$$\Phi(\alpha) = \max_{y \in [0,1]} \{F(y) - 1\}, \quad (14)$$

где $\alpha = (\alpha, a_1, a_2, \dots, a_n) \in H_q \subset R_{n+1}$,

$$F(y) = \alpha y^n + a_1 y^{n-1} + \dots + a_{n-1} y + a_n. \quad (5)$$

Не представляет большой сложности установить корректность сформулированной задачи, т.е. доказать, что существует точка $a^* \in H_q$, для которой

$$\Phi(a^*) = \inf_{a \in H_q} \Phi(a) \quad (15)$$

Соответствующий полином является искомой амплитудной функцией.

Если не требовать выполнения условия 3) в определении множества \tilde{H}_q , то мы получим известное решение, задаваемое полиномом Чебышева

$$T_n^*(y) = \frac{\alpha}{2^{1/n-1}} [\cos n \arccos(2y-1) + 1] + 1. \quad (16)$$

Пусть q_r^* - максимальная добротность этого полинома.

Если теперь $q \geq (q_r^*)^2 = q_r^*$, то полином $T_n^*(y) \in \tilde{H}_q$ и является единственным решением поставленной задачи. Поэтому основной интерес представляет задача минимизации функционала Φ на множестве H_q , когда

$$0 < q < (q_r^*)^2 = q_r^*. \quad (17)$$

Отметим, что при

$$0 < q \leq \frac{1}{2}$$

В.В.Гришбергом построено явное единственное решение, задаваемое полиномами (13)

$$G_n^*(y) = (\sqrt{\alpha} y^2 + \sqrt{\alpha} \frac{1-2q}{q} y + 1)^k, \quad n=2k.$$

Наложим одну из основных теорем, используемых при

решении.

Пусть $F(y) \in \bar{H}_q$ и

$$0 \leq q_1 = \Gamma(c_{11}, c_{21}) \leq q_2 = \Gamma(c_{12}, c_{22}) \leq \dots \leq q_k = \Gamma(c_{1k}, c_{2k}) = q_{\max}$$

Отметим здесь, что если i -ый множитель в (2) имеет действительные корни, то соответствующая добротность

$$q_i \leq \frac{1}{2},$$

$$q_i \leq \frac{1}{4} < \frac{1}{2}.$$

Теорема. Если $q \leq q_r^*$, то для того, чтобы полином $F(y) \in \bar{H}_q$ был решением поставленной задачи аппроксимации, необходимо выполнение условия

$$q_{\max} = q \quad (18)$$

Доказательство здесь приводить не будем.

Из этой теоремы следует, что при $q \leq q_r^*$ найдется целое число r , $1 \leq r \leq k$, для которого

$$0 \leq q_1 \leq q_2 \leq \dots \leq q_{k-r} < q_{k-r+1} = q_{k-r+2} = \dots = q_k = q_{\max} = q. \quad (19)$$

Так при $q \leq \frac{1}{2}$ $r = k$, при $q = q_r^*$ $r = 1$.

Исходя из этого, расчет оптимального полинома можно организовать следующим образом: определяем число (например, просчитывая случаи $r = 1, 2, \dots, r_1$ до тех пор пока не будут выполняться условия 1)-3). Тогда

$$F(y) - 1 = x(B, y)(\alpha_j y^m + x_1 y^{m-1} + \dots + x_{m-1} y + x_m) - 1 = x(B, y)(x_1 y^{m-1} + \dots + x_m) - [1 - \alpha_j x(B, y) y^m],$$

(20)

где $B = (B_1, B_2, \dots, B_r)$, $B_i > 0$

для $i \in [1:r]$,

$$K(B, y) = \prod_{i=1}^r (y^2 - A_i y + B_i).$$

$$A_i = 2u\sqrt{B_i}, \quad u = 1 - \frac{1}{2q}, \quad q_i = \frac{\sqrt{B_i}}{2\sqrt{B_i} - A_i} = q,$$

$$i \in [1:r];$$

(21)

$$x = (x_1, \dots, x_m) \in R_m;$$

$$m = n - 2r.$$

Если известен вектор B , то минимизация $\Phi(\alpha)$ сводится к задаче наилучшего одностороннего (сверху) приближения функции $f(y) = 1 - \alpha K(B, y) y^m$ на отрезке $[0, 1]$ полиномами с весом

$$P(x, y) = K(B, y) (x_1 y^{m-1} + \dots + x_m).$$

Таким образом, решая последнюю задачу при каждом B , мы получим функцию

$$\varphi(B) = \min_{x \in R_m} \max_{y \in [0, 1]} \{P(x, y) - f(y)\}, \quad (22)$$

которая зависит от r переменных. Минимизируя функцию $\varphi(B)$ в допустимой области, мы приходим к решению исходной задачи.

К сожалению, функция $\varphi(B)$ может оказаться многоэкстремальной. Кроме того, нет доказательства единственности решения задачи.

Для решения задачи одностороннего приближения можно использовать аналог алгоритма Валле-Пуссена [2] для решения классической задачи наилучшего равномерного приближения функций.

В заключение отметим, что применять такой метод нам кажется возможным и в некоторых других случаях. Например, функцию $\Gamma(\xi, \zeta)$ можно выбрать такой, чтобы она задавала ограничения на добротность для полосового фильтра. Или вместо задания старшего коэффициента α полинома $F(y)$ можно задавать величину гарантированного затухания при некотором $y > 1$.

Литература

1. Лангин А.А. Расчет электрических фильтров. Рига, Зинатне, 1974, 183с.
2. Демьянов В.Ф., Малоземов В.Н. Введение в миниманс. М., Наука, 1972, 368с.

О СХОДИМОСТИ ИТЕРАЦИОННЫХ МЕТОДОВ
 ДЛЯ РЕШЕНИЯ СТАЦИОНАРНЫХ ЗАДАЧ
 ТЕОРИИ ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ

Б.Я.Мартузан, Б.С.Польский
 Вычислительный центр ЛГУ им.П.Стучки

§ I. Рассмотрим систему стационарных уравнений, описывающую распределение потенциала электрического поля и концентрации носителей тока в одномерной полупроводниковой структуре:

$$\Delta \varphi = n - p - N, \quad (1)$$

$$\Lambda_1(\varphi)n = R(p, n), \quad (2)$$

$$\Lambda_2(\varphi)p = R(p, n), \quad (3)$$

где:

$$\Delta \varphi = \frac{d^2 \varphi}{dx^2}, \quad (4)$$

$$\Lambda_1(\varphi)n = \frac{d}{dx} \left[D_n \left(\frac{dn}{dx} - n \frac{d\varphi}{dx} \right) \right], \quad (5)$$

$$\Lambda_2(\varphi)p = \frac{d}{dx} \left[D_p \left(\frac{dp}{dx} + p \frac{d\varphi}{dx} \right) \right], \quad (6)$$

$$R(p, n) = \frac{np - 1}{\tau_n(p+1) + \tau_p(n+1)}$$

φ - потенциал электрического поля, n - концентрация электронов, p - концентрация дырок, D_n , D_p - коэффициенты диффузии, N - результирующая концентрация примеси, τ_n , τ_p - времена жизни носителей. (Все переменные безразмерные.)

Дискретизация системы уравнений (I)-(3) приводит к следующей разностной схеме:

$$\Lambda^h \varphi = n - p - N \quad (7)$$

$$\Lambda_1^h(\varphi)n = R(p, n) \quad (8)$$

$$\Lambda_2^h(\varphi)p = R(p, n) \quad (9)$$

где $\varphi = (\varphi_0, \varphi_1, \dots, \varphi_i, \dots, \varphi_M)$, $n = (n_0, n_1, \dots, n_i, \dots, n_M)$, $p = (p_0, p_1, \dots, p_i, \dots, p_M)$

а Λ^h , Λ_1^h , Λ_2^h - некоторые трехточечные разностные операторы, аппроксимирующие Λ , Λ_1 и Λ_2 . В точках $i=0$ и $i=M$ ставятся краевые условия. Они общеизвестны [I], и мы для краткости их не выписываем. Разностная схема (7)-(9) является системой нелинейных алгебраических уравнений относительно неизвестных φ_i, n_i, p_i .

Рассмотрим два итерационных метода для ее решения:

$$I. \Lambda^h \varphi^{l+1} = n^l - p^l - N \quad (10)$$

$$\Lambda_1^h(\varphi^{l+1})n^{l+1} = R(p^l, n^l) + \frac{n^{l+1} - n^l}{\tau_1} \quad (11)$$

$$\Lambda_2^h(\varphi^{l+1})p^{l+1} = R(p^l, n^l) + \frac{p^{l+1} - p^l}{\tau_2} \quad (12)$$

$$2. \quad \Lambda^h \varphi^{l+1} = n^l - \rho^l - N + \frac{\varphi^{l+1} - \varphi^l}{\tau_2} \quad (13)$$

$$\Lambda_1^h(\varphi^{l+1})n^{l+1} = R(\rho^l, n^l) + \frac{n^{l+1} - n^l}{\tau_2} \quad (14)$$

$$\Lambda_2^h(\varphi^{l+1})\rho^{l+1} = R(\rho^l, n^l) + \frac{\rho^{l+1} - \rho^l}{\tau_2} \quad (15)$$

где: $l = 0, 1, 2, \dots$ номер итерации, τ_1 и τ_2 - итерационные параметры, φ^0 , n^0 , ρ^0 - заданные начальные приближения. Второй метод отличается от первого тем, что в правую часть уравнения Пуассона введен "нестационарный" член $(\varphi^{l+1} - \varphi^l)/\tau_2$. Реализация же алгоритмов, как в первом, так и во втором случае следующая: вначале по известным n^l и ρ^l из соотношений (10) или (13) определяется φ^{l+1} , затем из соотношений (11), (12) или (14), (15) по известным φ^{l+1} , n^l и ρ^l определяются n^{l+1} и ρ^{l+1} . Очевидно, что каждое из соотношений (10)-(15) является системой линейных алгебраических уравнений с трехдиагональной матрицей и эффективно решается методом прогонки [4].

Первый метод был предложен в работе [2]. В работе [3], на модельной задаче методом малых возмущений, было показано, что сходимость имеет место при условии:

$$\tau_1 \ll \frac{1}{DN} \quad (16)$$

D и N - максимальные значения коэффициента диффузии и концентрации легирующей примеси. Для сильнолегированных биполярных структур ($N \approx 10^{18}$) τ_1 очень мало, что приводит к низкой скорости сходимости метода и делает его практически непригодным для решения стационарных задач.

Второй метод был предложен в [6]. В этой работе численные расчеты для конкретной биполярной транзисторной структуры показали значительно более высокую эффективность метода [2] по сравнению с методом [1]. В настоящей работе на модельной задаче методом малых возмущений будут получены условия сходимости метода [2], а также будут даны рекомендации по выбору параметра τ_2 при решении реальных задач теории полупроводниковых приборов.

§ 2. Анализ сходимости итерационного процесса (I3) - (I5) в силу нелинейности исходной системы уравнений (7) - (9) крайне затруднителен. Поэтому мы сделаем следующие упрощающие предположения:

1. Полагаем $R(\rho, n) \equiv 0$.

2. Пренебрегаем неосновными носителями, в данном случае дырками, т.е. рассматриваем систему, состоящую из уравнений (1) и (2).

3. Считаем, что $D_n = D = \text{const}$, $N = \text{const} > 0$.

Тогда для простейшей разностной аппроксимации операторов Δ и Δ_1 система (7) - (8) принимает следующий вид:

$$\frac{\varphi_{i+1} - 2\varphi_i + \varphi_{i-1}}{h^2} = n_i - N \quad (I7)$$

$$n_{i+1} \left[\frac{D}{h^2} - \frac{D(\varphi_{i+1} - \varphi_i)}{2h^2} \right] - n_i \left[\frac{2D}{h^2} + \frac{D(\varphi_{i+1} - \varphi_i)}{2h^2} - \frac{D(\varphi_i - \varphi_{i-1})}{2h^2} \right] + n_{i-1} \left[\frac{D}{h^2} + \frac{D(\varphi_i - \varphi_{i-1})}{2h^2} \right] = 0, \quad (I8)$$

где h - шаг разностной сетки.

Метод 2 для этой разностной схемы запишется в виде:

$$\frac{\varphi_{i+1}^{l+1} - 2\varphi_i^{l+1} + \varphi_{i-1}^{l+1}}{h^2} = n_i^l - N + \frac{\varphi_{i+1}^{l+1} - \varphi_i^l}{\tau_2} \quad (I9)$$

$$n_{i-1}^{l+1} \left[\frac{D}{h^2} - \frac{D(\varphi_{i-1}^{l+1} - \varphi_i^{l+1})}{2h^2} \right] - n_i^{l+1} \left[\frac{2D}{h^2} + \frac{D(\varphi_{i+1}^{l+1} - \varphi_i^{l+1})}{2h^2} - \frac{D(\varphi_i^{l+1} - \varphi_{i-1}^{l+1})}{2h^2} \right] +$$

$$+ n_{i-1}^{l+1} \left[\frac{D}{h^2} + \frac{D(\varphi_i^{l+1} - \varphi_{i-1}^{l+1})}{2h^2} \right] = \frac{n_i^{l+1} - n_i^l}{\tau_2} \quad (20)$$

Очевидно, что сеточные функции $\varphi_i = c$ и $n_i = N$ ($i = 0, 1, 2, \dots, M$), где c - произвольная постоянная, являются решением системы (I7)-(I8). Будем искать решение системы (I9)-(20) в виде:

$$\varphi_i^{l+1} = c + Z_i^{l+1} \quad (21)$$

$$n_i^{l+1} = N + Y_i^{l+1} \quad (22)$$

Подставляя (21) и (22) в (I9)-(20), пренебрегая членами второго порядка, для поправок Z_i и Y_i получим следующую систему уравнений:

$$\frac{Z_{i+1}^{l+1} - 2Z_i^{l+1} + Z_{i-1}^{l+1}}{h^2} = Y_i^l + \frac{Z_i^{l+1} - Z_i^l}{\tau_2} \quad (23)$$

$$D \frac{Y_{i+1}^{l+1} - 2Y_i^{l+1} + Y_{i-1}^{l+1}}{h^2} = Y_i^l ND + \frac{Y_i^{l+1} - Y_i^l}{\tau_2} + ND \frac{Z_i^{l+1} - Z_i^l}{\tau_2} \quad (24)$$

$$Z_0^{l+1} = Z_N^{l+1} = Y_0^{l+1} = Y_N^{l+1} = 0 \quad (25)$$

Или:

$$\Lambda^h Z^{l+1} = Y^l + \frac{Z^{l+1} - Z^l}{\tau_2} \quad (26)$$

$$D\Lambda^h y^{l+1} = NDy^l + \frac{y^{l+1} - y^l}{T_2} + ND \frac{z^{l+1} - z^l}{T_2}; \quad (27)$$

где

$$y^{l+1} = (y_0^{l+1}, y_1^{l+1}, \dots, y_M^{l+1}), \quad z^{l+1} = (z_0^{l+1}, z_1^{l+1}, \dots, z_M^{l+1})$$

$$(\Lambda^h z^{l+1})_i = \frac{z_{i+1}^{l+1} - 2z_i^{l+1} + z_{i-1}^{l+1}}{h^2}, \quad i = 1, 2, \dots, M-1, \quad (28)$$

$$z_0^{l+1} = z_M^{l+1} = y_0^{l+1} = y_M^{l+1} = 0.$$

Лемма. Пусть

$$T_1 = \max\left(0, \frac{1+D}{ND} - \sqrt{\frac{4}{N^2D} + \frac{h^2}{ND}}\right), \quad (29)$$

$$T_2 = \frac{1+D}{ND} + \sqrt{\frac{4}{N^2D} + \frac{h^2}{ND}}.$$

Если

$$T_1 < T_2 \leq T_2, \quad (30)$$

то итерационный процесс (26)-(27) сходится.

Доказательство. Пусть u_k и λ_k ($k=1, 2, \dots, M-1$) собственные функции и собственные значения оператора Λ^h [4]. Представим z^{l+1} и y^{l+1} в виде:

$$z^{l+1} = \sum_{k=1}^{n-1} z_k^{l+1} u_k, \quad y^{l+1} = \sum_{k=1}^{n-1} y_k^{l+1} u_k. \quad (31)$$

Подставляя (31) в (26)-(27) и используя равенство

$$\Lambda^h u_k = -\lambda_k u_k \quad (32)$$

получим.

$$z_k^{l+1} = \frac{z_k^l}{1 + \tau_2 \lambda_k} - \frac{\tau_2 y_k^l}{1 + \tau_2 \lambda_k} \quad (33)$$

$$y_k^{l+1} = \frac{\tau_2 \lambda_k N D z_k^l}{(1 - \tau_2 \lambda_k)(1 + \tau_2 D \lambda_k)} + \frac{(1 + \tau_2 \lambda_k - \tau_2^2 N D \lambda_k) y_k^l}{(1 + \tau_2 \lambda_k)(1 + \tau_2 D \lambda_k)} \quad (34)$$

Запишем (33)-(34) в виде

$$v_k^{l+1} = A_k v_k^l \quad (35)$$

где

$$v_k^{l+1} = \begin{pmatrix} z_k^{l+1} \\ y_k^{l+1} \end{pmatrix}, \quad A_k = \begin{pmatrix} \frac{1}{1 + \tau_2 \lambda_k} & -\frac{\tau_2}{1 + \tau_2 \lambda_k} \\ \frac{\tau_2 \lambda_k N D}{(1 - \tau_2 \lambda_k)(1 + \tau_2 D \lambda_k)} & \frac{1 + \tau_2 \lambda_k - \tau_2^2 N D \lambda_k}{(1 + \tau_2 \lambda_k)(1 + \tau_2 D \lambda_k)} \end{pmatrix} \quad (36)$$

Если мы покажем, что собственные значения μ_k матрицы A_k по модулю меньше 1 для всех $k=1, 2, \dots, M-1$, то лемма будет доказана. Уравнение, определяющее μ_k имеет вид

$$\mu_k^2 - \frac{2+(1+D)\lambda_k T_2 - T_2^2 N D \lambda_k}{(1+T_2 \lambda_k)(1+T_2 D \lambda_k)} \mu_k + \frac{1}{(1+T_2 \lambda_k)(1+T_2 D \lambda_k)} = 0 \quad (37)$$

Выпишем дискриминант D_k этого уравнения:

$$D_k = \frac{T_2^2 N^2 D^2 \lambda_k^2 \left[T_2^2 - \frac{2(1+D)}{ND} T_2 + \frac{(1-D)^2}{N^2 D^2} - \frac{4}{ND \lambda_k} \right]}{4(1+T_2 \lambda_k)^2 (1+T_2 D \lambda_k)^2} \quad (38)$$

Покажем, что $D_k \leq 0$ для всех $k=1, 2, \dots, M-1$. Действительно, нетрудно видеть, что если:

$$\frac{1+D}{ND} - \sqrt{\frac{4}{N^2 D} + \frac{4}{ND \lambda_k}} \leq T_2 \leq \frac{1+D}{ND} + \sqrt{\frac{4}{N^2 D} + \frac{4}{ND \lambda_k}} \quad (39)$$

то

$$D_k^* = T_2^2 - \frac{2(1+D)}{ND} T_2 + \frac{(1-D)^2}{N^2 D^2} - \frac{4}{ND \lambda_k} \leq 0. \quad (40)$$

Поскольку [4]:

$$0 < \alpha \leq \lambda_k \leq \beta = \frac{4}{h^2}, \quad k=1, 2, \dots, M-1, \quad (41)$$

то в силу условия (39) $D_k \leq 0$ для всех $k=1, 2, \dots, M-1$.
По тогда

$$|(M_{21})_1| = |(M_{21})_2| = \frac{1}{\sqrt{(1+\tau_2\lambda_n)(1+\tau_2D\lambda_n)}} < 1 \quad (42)$$

для всех $k=1, 2, \dots, M-1$ и, тем самым, лемма доказана.

Замечание. Поскольку скорость сходимости итерационного процесса (35) определяется наибольшим по модулю собственным значением матрицы A_k , то из (42) следует, что максимальная скорость сходимости будет при $\tau_2 = \tau_2$.

§ 3. В качестве приложения, полученных выше результатов, приведем пример расчета диодной np структуры со следующими значениями параметров: длина всей структуры $L = 2$ мкм, глубина залегания np перехода $y_0 = 1$ мкм, времена жизни носителей $\tau_n = \tau_p = 10^{-8}$ сек. Распределение примеси задавалось по формуле:

$$N(x) = 10^{10} e^{-cx^2} - 10^{10} \text{ (см}^{-3}\text{)},$$

где c выбрана таким образом, что $N(y_0) = 0$. Коэффициенты диффузии зависели от концентрации легирующей примеси и электрического поля по эмпирической формуле, предложенной в [6]. На отрезке $[0, L]$ была построена неравномерная сетка, содержащая 25 узлов:

$$0,0(0,125) 0,5; 0,5(0,05) 1,20; 1,20(0,075) 1,5; 1,7; 2,0.$$

Координаты приведены в микронах.

Использовалась разностная аппроксимация, предложенная в [7]. Для этой разностной аппроксимации итерационный метод 2 имеет следующий вид:

$$\frac{1}{h_i^* h_{i+1}} \varphi_{i+1}^{l+1} - \left(\frac{1}{h_i^* h_{i+1}} + \frac{1}{h_i^* h_i} + \frac{1}{\tau_2} \right) \varphi_i^{l+1} + \frac{1}{h_i^* h_i} \varphi_{i-1}^{l+1} = \quad (43)$$

$$= n_i^l - p_i^l - N_i - \frac{\varphi_i^l}{\tau_2},$$

$$\begin{aligned} & \frac{(D_n)_{i+\frac{1}{2}}(\varphi_{i+1}^{L+1} - \varphi_i^{L+1})}{h_i^* h_{i+1}(\exp(\varphi_{i+1}^{L+1} - \varphi_i^{L+1}) - 1)} n_{i+1}^{L+1} - \left[\frac{(D_n)_{i+\frac{1}{2}}(\varphi_{i+1}^{L+1} - \varphi_i^{L+1})}{h_i^* h_{i+1}(1 - \exp(\varphi_{i+1}^{L+1} - \varphi_i^{L+1}))} + \right. \\ & \left. + \frac{(D_n)_{i-\frac{1}{2}}(\varphi_i^{L+1} - \varphi_{i-1}^{L+1})}{h_i^* h_i(\exp(\varphi_i^{L+1} - \varphi_{i-1}^{L+1}) - 1)} + \frac{1}{\tau_2} \right] n_i^{L+1} + \frac{(D_n)_{i-\frac{1}{2}}(\varphi_i^{L+1} - \varphi_{i-1}^{L+1})}{h_i^* h_i(1 - \exp(\varphi_i^{L+1} - \varphi_{i-1}^{L+1}))} n_{i-1}^{L+1} = \\ & = R(p_i^L, n_i^L) - \frac{n_i^L}{\tau_2}, \end{aligned}$$

(44)

$$\begin{aligned} & \frac{(D_p)_{i+\frac{1}{2}}(\varphi_{i+1}^{L+1} - \varphi_i^{L+1})}{h_i^* h_{i+1}(1 - \exp(\varphi_{i+1}^{L+1} - \varphi_i^{L+1}))} p_{i+1}^{L+1} - \left[\frac{(D_p)_{i+\frac{1}{2}}(\varphi_{i+1}^{L+1} - \varphi_i^{L+1})}{h_i^* h_{i+1}(\exp(\varphi_{i+1}^{L+1} - \varphi_i^{L+1}) - 1)} + \right. \\ & \left. + \frac{(D_p)_{i-\frac{1}{2}}(\varphi_i^{L+1} - \varphi_{i-1}^{L+1})}{h_i^* h_i(1 - \exp(\varphi_i^{L+1} - \varphi_{i-1}^{L+1}))} + \frac{1}{\tau_2} \right] p_i^{L+1} + \frac{(D_p)_{i-\frac{1}{2}}(\varphi_i^{L+1} - \varphi_{i-1}^{L+1})}{h_i^* h_i(\exp(\varphi_i^{L+1} - \varphi_{i-1}^{L+1}) - 1)} p_{i-1}^{L+1} = \\ & = R(p_i^L, n_i^L) - \frac{p_i^L}{\tau_2}, \end{aligned}$$

(45)

где

$$h_i = x_{i+1} - x_{i-1} \quad (x_i - \text{безразмерные})$$

$$h_i^* = \frac{-h_i + h_{i+1}}{2}, \quad D_{i+\frac{1}{2}} = \frac{D_i + D_{i+1}}{2}$$

На рис. 1 приведена зависимость величины

$$\delta = \max_i \frac{|\varphi_i^{L+1} - \varphi_i^L|}{|\varphi_i^L|}$$

от номера итерации для I-го и 2-го методов, при заданном прямом смещении $\psi = 0,7$ в. Начальные приближения в обоих случаях были выбраны следующим образом: n_i^0 и ρ_i^0 — были решениями системы уравнений:

$$\begin{aligned} n_i^0 - \rho_i^0 - N_i &= 0, \\ n_i^0 \rho_i^0 &= 1, \end{aligned}$$

а φ^0 — линейная функция. Непосредственно применять полученные выше результаты для определения τ_2 нельзя, поскольку D , N , h зависят от i . Мы определяли τ_2 равенством:

$$\begin{aligned} \tau_2 &= \min_{1 \leq i \leq n-1} \left[\frac{1 + (D_n)_{i-1/2}}{|N_i|(D_n)_{i-1/2}} + \sqrt{\frac{4}{N_i^2 (D_n)_{i-1/2}} + \frac{H_i^2}{|N_i|(D_n)_{i-1/2}}} \right] \approx \\ &\approx \min_{1 \leq i \leq n-1} \left[\frac{1}{|N_i|(D_n)_{i-1/2}} + \frac{H_i}{\sqrt{|N_i|(D_n)_{i-1/2}}} \right] \end{aligned} \quad (46)$$

где $H_i = \min(h_i, h_{i+1})$.

При вычислении τ_2 считали, что D_n не зависит от электрического поля. В данном примере величина τ_2 оказалась равной $0,398 \cdot 10^{-5}$. Заметим, что при $\tau_2 = 0,5 \cdot 10^{-5}$ наблюдалась неустойчивость вычислительного алгоритма. Для I-го метода τ_1 определялось следующим образом:

$$\tau_1 = \min_{1 \leq i \leq n-1} \frac{1}{|N_i|(D_n)_{i-1/2}} \quad (47)$$

и в нашем примере $\tau_1 = 0,798 \cdot 10^{-6}$. Сравнение кривых I и

2 показывает значительно большую эффективность метода 2 по сравнению с методом 1.

Программа, реализующая изложенный выше алгоритм написана на языке Фортран. Расчеты проводились на ЭВМ GE-415. Время, затрачиваемое на одну итерацию, составляло 0,28сек.

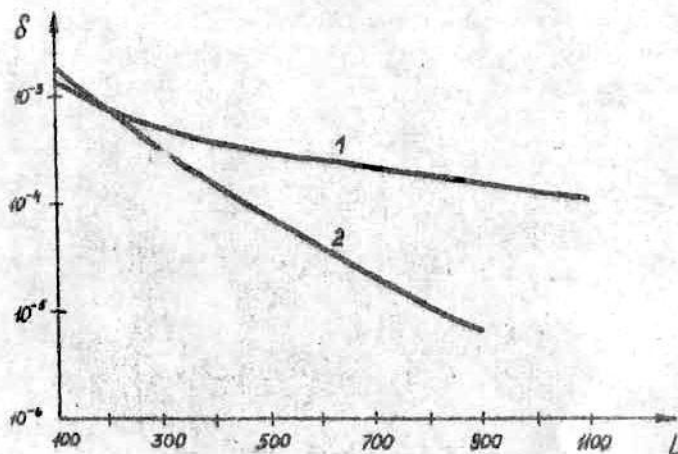


Рис. I. Зависимость относительной погрешности от номера итерации. 1 метод - кривая 1, 2 метод - кривая 2.

Литература

1. Arandjelovic V. Accurate numerical Steady-state Solution for a diffused one-dimensional junction diode.-Solid State Electr., v.13, 1970, p.855-871.
2. Reiser M. Computing methods in semiconductor problems.-Comp.Meth. in Appl.Scienc. and Eng., Int.Symp., Ver - sailles, 1973.
3. Reiser M. On the stability of finite difference schemes in transient semiconductor problems.-Comp.Meth.in Appl. Mech. and Engin., 1973, 2, p.65-68.
4. Самарский А. Введение в теорию разностных схем. М., Наука, 1971, 550с.
5. Мартуван Б.Я., Польский Э.С. Численное моделирование одномерной биполярной транзисторной структуры.-Изв.АН ЛатвССР, Сер.физ. и техн.наук, в печати.
6. Sanghey D.M., Thomas R.E. Carrier mobilities in silicon empirically related to doping and field.-Proc.IEEE (Lett.), v.55, 1967, p.2192-2193.
7. Sharfetter D.L., Gummel H.K. Large signal analysis of a silicon Read diode oscillator.-IEEE Trans.Electr.Dev., 1969, v.ED-16, p.64-77.

РАСЧЕТ ОДНОМЕРНОГО ДИФФУЗИОННОГО РАСПРЕДЕЛЕНИЯ БОРА И ФОСФОРА В КРЕМНИИ

Н.М.Иванов, В.И.Новиков
Рижский НИИ микроприборов

ВВЕДЕНИЕ

В практике машинного проектирования интегральных схем (ИС) и их компонентов все большее распространение получают модели, связывающие электрические характеристики приборов с параметрами материала и геометрическими размерами диффузионных областей. Во многих разработанных к настоящему времени комплексах программ проектирования ИС распределение концентраций примеси аппроксимируется с помощью функций, являющихся аналитическим решением уравнения диффузии в простейшей постановке [1, 2]. Такой подход может привести к значительным ошибкам по электрическим характеристикам приборов, поэтому представляют интерес модели, более точно учитывающие физические явления технологического процесса.

ПОСТАНОВКА ЗАДАЧИ

В настоящей работе рассматривается одномерная модель диффузионного процесса, учитывающая динамику роста стекловидной пленки на поверхности кремния. Предполагается, что толщина пленки w зависит от времени t по линейно-параболическому закону [3].

$$w = \sqrt{C_2 t + (C_1 + w_0)^2} - C_1, \quad (1)$$

где C_1, C_2 — коэффициенты, зависящие от температуры,
 w_0 — начальная толщина пленки.

При этом слой кремния толщиной $(w - w_0)\epsilon$ переходит в стекло. Коэффициент ϵ принимается равным отношению плотности стекла к плотности кремния.

Выберем за начало отсчета по глубине x границу раздела стекла и кремния. Тогда уравнения диффузии для концентраций атомов примесей $\bar{u}(x, t)$ в стекле и $u(x, t)$ в кремнии можно представить в виде

$$\frac{\partial \bar{u}}{\partial t} = -\frac{\partial \bar{J}}{\partial x}, \quad -W \leq x \leq 0, \quad (2)$$

$$\frac{\partial u}{\partial t} = -\frac{\partial J}{\partial x}, \quad 0 \leq x \leq x_{\infty}, \quad (3)$$

где

$$\bar{J} = -\bar{D} \frac{\partial \bar{u}}{\partial x} - W_1' \bar{u}, \quad (4)$$

$$J = -D \frac{\partial u}{\partial x} - \epsilon W_2' u \quad (5)$$

— плотности диффузионных потоков. Концентрационную зависимость коэффициента диффузии в кремнии выберем согласно [4]:

$$D = D_0 \left(1 + \frac{u^2}{c^2}\right) \left(1 + \frac{u}{\sqrt{u^2 + n_i^2}}\right).$$

Коэффициент диффузии в стекле \bar{D} считаем постоянным.

Первое граничное условие, связывающее уравнения (2), (3)

$$\bar{u}(0, t) = \eta u(0, t), \quad (6)$$

следует из предположения о сегрегации примеси на границе "стекло-кремний". Кроме того, для уравнения (3) выполняется

$$\frac{\partial}{\partial x} u(x_{\infty}, t) = 0, \quad (7)$$

так как толщина кремниевой пластины намного превышает глубину диффузии.

Диффузионный процесс можно условно разделить на два

этапа. В течение первого этапа на поверхность кремния подается газовая смесь, содержащая диффузанта и окислитель. При этом предполагается, что в образующемся стекле примесь распределяется равномерно, т.е. $\bar{u}(x, t) = C_0$ и с учетом (6)

$$u(0, t) = \frac{C_0}{2} \quad (8)$$

Таким образом можно ограничиться решением уравнения (3) с условиями (7), (8) и начальным условием $u(x, 0) = 0$.

На втором этапе подача диффузанта прекращается, хотя процесс образования стекла на поверхности кремния может продолжаться. В этом случае будем проводить совместное решение уравнений (2), (3) с условиями (6), (7) и условием непрерывности диффузионного потока

$$\bar{J}(0, t) = J(0, t). \quad (9)$$

Кроме того, для уравнения (2) потребуем выполнение условия

$$\frac{\partial}{\partial x} \bar{u}(-w, t) = 0, \quad (10)$$

в предположении, что с поверхности стекла примесь не испаряется. За начальное условие возьмем распределение примеси на предыдущем этапе диффузии.

КОНЕЧНО-РАЗНОСТНАЯ АППРОКСИМАЦИЯ

Для решения поставленной задачи будем использовать конечно-разностный метод.

Введем в области $-w \leq x \leq 0$, $0 \leq t \leq t_x$ прямоугольную сетку с постоянным шагом $h = x_i - x_{i-1}$. Переменный шаг по времени $\tau_j = t_j - t_{j-1}$ выберем из условия $w(t_j) - w(t_{j-1}) = h$, накладывая ограничение $\tau_j \leq \tau_{max}$.

В области $0 \leq x \leq x_{\infty}$, $0 \leq t \leq t_x$ переменный шаг по x выберем следующим образом:

$$h_i = x_{i+1} - x_i = \begin{cases} h_0, & 0 \leq i < i_1, \\ 2h_0, & i_1 \leq i < i_2, \\ 4h_0, & i_2 \leq i < n. \end{cases}$$

На рис. I представлено примерное расположение узлов сетки.

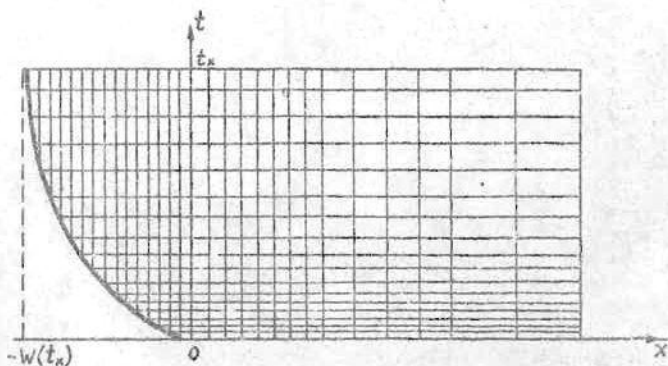


Рис. I. Расположение узлов сетки при $i_1 = 8$, $i_2 = 12$, $n = 14$.

Значения сеточной функции $v(x_i, t_j)$ будем, как обычно, обозначать через u_i^j . Заменяя разностями частные производные в уравнении (3), получаем

$$u_i^j - u_i^{j-1} = \frac{\tau_j}{h_i} (J_{i-1/2}^j - J_{i+1/2}^j), \quad (II)$$

где $h_i = \frac{1}{2} (h_i + h_{i-1})$. С помощью подстановки

$$u = z \exp\left(-\frac{\epsilon W_t x}{D}\right)$$

в выражение для плотности потока (5) находим

$$\frac{\partial z}{\partial x} = -\frac{\tau}{D} \exp\left(\frac{\epsilon W_t x}{D}\right). \quad (I2)$$

Предположим, что в интервале $[x_i, x_{i+1}]$ плотность потока и коэффициент диффузии постоянны: $J = J_{i+1/2}^j$, $D = D_{i+1/2}^{j-1}$. Тогда, интегрируя обе части равенства (12), получаем

$$z_{i+1}^j - z_i^j = \frac{J_{i+1/2}^j}{\varepsilon W'(t_j)} \left[\exp\left(\frac{\varepsilon W'(t_j) x_i}{D_{i+1/2}^{j-1}}\right) - \exp\left(\frac{\varepsilon W'(t_j) x_{i+1}}{D_{i+1/2}^{j-1}}\right) \right]$$

или

$$J_{i+1/2}^j = \varepsilon W'(t_j) \frac{u_i^j \exp\left(-\frac{\varepsilon W'(t_j) h_i}{D_{i+1/2}^{j-1}}\right) - u_{i+1}^j}{1 - \exp\left(-\frac{\varepsilon W'(t_j) h_i}{D_{i+1/2}^{j-1}}\right)} = \alpha_i^j u_i^j - \beta_i^j u_{i+1}^j,$$

где

$$\beta_i^j = \frac{\varepsilon W'(t_j)}{1 - \exp\left(-\frac{\varepsilon W'(t_j) h_i}{D_{i+1/2}^{j-1}}\right)}, \quad \alpha_i^j = \beta_i^j \exp\left(-\frac{\varepsilon W'(t_j) h_i}{D_{i+1/2}^{j-1}}\right).$$

Подставляя полученное выражение в (11) и учитывая, что

$$J_{i-1/2}^j = \alpha_{i-1}^j u_{i-1}^j - \beta_{i-1}^j u_i^j,$$

приходим к системе линейных уравнений

$$A_i^j u_{i-1}^j - c_i^j u_i^j + B_i^j u_{i+1}^j = -u_i^{j-1}, \quad (13)$$

где

$$A_i^j = \frac{\tau_j}{h_i} \alpha_{i-1}^j, \quad B_i^j = \frac{\tau_j}{h_i} \beta_i^j, \quad c_i^j = 1 + \frac{\tau_j}{h_i} (\alpha_i^j + \beta_{i+1}^j).$$

Аналогично можно найти коэффициенты системы уравнений, описывающей распределение концентрации в стекле:

$$\bar{A}_i \bar{u}_{i-1}^j - \bar{c}_i \bar{u}_i^j + \bar{B}_i \bar{u}_{i+1}^j = -\bar{u}_i^{j-1}. \quad (14)$$

Учитывая, что в этом случае нумерация узлов по пространству выбрана в обратном порядке, получаем

$$\bar{A}_j = \frac{\tau_j}{h} \frac{w'(t_j)}{1 - \exp(-\frac{w'(t_j)h}{\bar{D}})}, \quad \bar{B}_j = \bar{A}_j \exp(-\frac{w'(t_j)h}{\bar{D}}), \quad \bar{C}_j = 1 + \bar{A}_j + \bar{B}_j.$$

Метод получения разностных схем (I3), (I4) аналогичен методу приведенному в работе [5]. В применении к рассматриваемой задаче эта схема значительно экономичней по числу узлов сетки, чем обычная центрально-разностная схема, для устойчивости которой должны выполняться условия

$$h < \frac{2\bar{D}}{w'(0)}, \quad \max h_i < \frac{2D_0}{\varepsilon w'(0)}.$$

Для совместного решения систем линейных уравнений (I3), (I4) целесообразно использовать метод встречной прогонки [6].

РЕЗУЛЬТАТЫ РАСЧЕТА

На рис.2, в качестве примера, приводятся расчетные графики распределения бора и фосфора для двухстадийного диффузионного процесса, основные параметры которого сведены в таблицы I и 2.

Для каждого режима диффузии, характеризуемого определенной температурой (T), продолжительностью (t) и скоростью роста стекла, в таблицах указаны значения начальной и конечной толщины стекла (w₀, w_t), коэффициенты диффузии в стекле (D̄) и в кремнии (D). Приводятся также рассчитанные согласно [1] значения поверхностных сопротивлений (R_s), глубины p-n переходов (x_j), поверхностные концентрации (C_s). В таблице I показано изменение параметров диффузии бора при диффузии фосфора (см. режимы при 1050°C и 960°C).

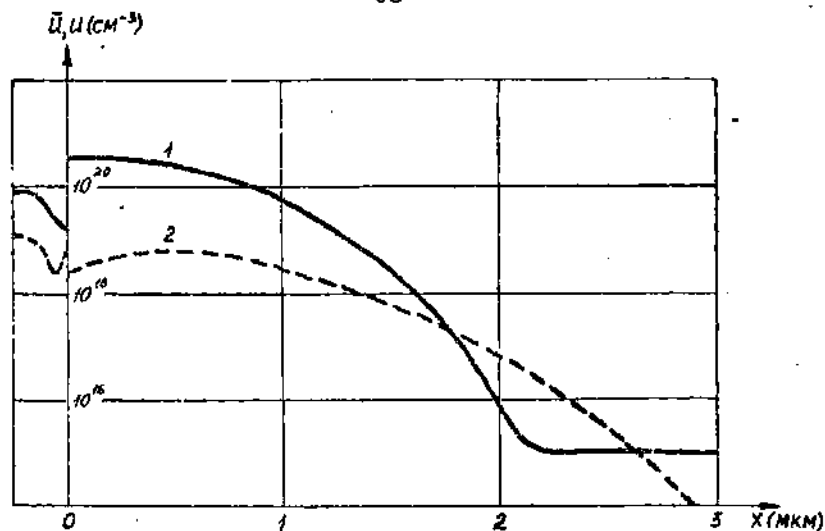


Рис.2. Распределение концентраций фосфора (1) и бора (2) в стекле и в кремнии.

Таблица I. Основные параметры диффузии бора

$T^{\circ}C$	\bar{D} МКМ ² /МИН	$\min D$ МКМ ² /МИН	$\max D$ МКМ ² /МИН	t МИН	W_0 МКМ	W_t МКМ	X_t МКМ	R_t ОМ/О	C_s СМ ⁻³
950	$5,15 \cdot 10^{-7}$	$1,03 \cdot 10^{-8}$	$8,20 \cdot 10^{-4}$	20	0	0,16	0,18	110	$3,6 \cdot 10^{20}$
			$8,20 \cdot 10^{-4}$	20	0,16	0,16	0,25	85	$3,3 \cdot 10^{20}$
			$7,04 \cdot 10^{-4}$	5	0,16	0,17	0,27	80	$2,8 \cdot 10^{20}$
1150	$5,50 \cdot 10^{-7}$	$2,50 \cdot 10^{-5}$	$3,63 \cdot 10^{-3}$	15	0	0,09	0,50	103	$1,5 \cdot 10^{19}$
			$3,08 \cdot 10^{-3}$	30	0,09	0,51	2,17	179	$1,6 \cdot 10^{18}$
			$3,01 \cdot 10^{-3}$	10	0,51	0,53	2,38	174	$4,6 \cdot 10^{18}$
1050	$4,35 \cdot 10^{-6}$	$1,36 \cdot 10^{-4}$	$2,52 \cdot 10^{-4}$	5	0,53	0,54	2,40	176	$3,8 \cdot 10^{18}$
			$2,46 \cdot 10^{-4}$	20	0,54	0,54	2,42	178	$3,1 \cdot 10^{18}$
960	$3,11 \cdot 10^{-9}$	$1,41 \cdot 10^{-5}$	$1,45 \cdot 10^{-5}$	20	0,54	0,57	2,17	183	$1,7 \cdot 10^{17}$
			$1,62 \cdot 10^{-5}$	10	0,57	0,57	2,42	182	$1,6 \cdot 10^{18}$

Таблица 2. Основные параметры диффузии фосфора

$T^{\circ}C$	\bar{D} МКМ ² /МИН	$\min D$ МКМ ² /МИН	$\max D$ МКМ ² /МИН	t МИН	W_0 МКМ	W_2 МКМ	X_j МКМ	R_s ОМ/□	C_s СМ ⁻³
1050	$4,81 \cdot 10^{-5}$	$9,63 \cdot 10^{-4}$	$7,61 \cdot 10^{-3}$	6	0	0,13	0,65	3,8	$5,0 \cdot 10^{20}$
			$7,61 \cdot 10^{-3}$	20	0,13	0,14	1,61	4,8	$4,6 \cdot 10^{20}$
960	$3,35 \cdot 10^{-6}$	$6,70 \cdot 10^{-5}$	$1,50 \cdot 10^{-3}$	20	0,14	0,25	1,72	4,8	$4,0 \cdot 10^{20}$
			$1,34 \cdot 10^{-3}$	10	0,2	0,26	1,75	4,8	$3,6 \cdot 10^{20}$

Результаты расчета глубин р-п переходов и поверхностных диффузионных сопротивлений дали хорошее совпадение с экспериментальными значениями этих параметров.

В вычислениях использовалась разностная сетка, содержащая около 100 узлов по пространственной координате, включая область стекла. Шаг по времени составлял в среднем 0,5 мин. Путем дальнейшего увеличения числа узлов сетки было установлено, что при изменении концентраций от 10^{21} см⁻³ до 10^{16} см⁻³ относительная погрешность изменяется от 5% до 10%. Расчеты проводились на ЭЦВМ GE-415. Затраты машинного времени для рассмотренного выше примера составили около 5 минут.

Разработанная программа успешно решает задачу описания физической структуры ИС, исходя из технологической информации. Это позволяет использовать ее в системе программы проектирования интегральных компонентов.

Литература

1. Казеннов Г.Г., Баталов Б.В., Руденко А.А., Авдеев Е.В., Чеботарев Е.В. Проектирование физической структуры транзисторов логических интегральных микросхем с помощью ЭММ. - Микроэлектроника, 1974, вып.7, с.100-114.
2. Ермак Е.В., Жаров М.М., Никишина В.И., Петросянец К.О., Русаков С.Г., Смолко Г.Г., Харин В.И. Комплекс программ для расчета электрических характеристик компонентов биполярных ИС.-Электронная промышленность, 1974, № 12, с.15-22.
3. Основы технологии кремниевых интегральных схем. Окисление. Диффузия. Эпитаксия. Пер. с англ. под ред. Р.Бургера и Р.Донована. М., Мир, 1969, 451 с.
4. Thai N.D. Anomalous diffusion in semiconductors - quantitative analysis.-Solid-State Electronics, v.13, 1970, p.165-172.
5. Sharfetter D.L., Gummel H.K. Large signal analysis of a silicon Read diode oscillator.-IEEE Trans, Electr.Dev. 1969, v.ED-16, p.64-77.
6. Самарский А.А. Введение в теорию разностных схем. М., Наука, 1971, 552 с.

О МОДИФИКАЦИЯХ МЕТОДА НЬЮТОНА ДЛЯ АНАЛИЗА ИНТЕГРАЛЬНЫХ СХЕМ

Л.Э.Кацнельсон, Е.С.Кельман
Вычислительный центр ЛГУ им.П.Стучки

Проблема выбора хорошего начального приближения для анализа статических характеристик интегральных схем любым численным методом очень актуальна.

В статье предлагаются модификации метода Ньютона, которые использовались в системе программ КОМПРОМИС и позволили просчитать целый ряд практически важных схем с большим усилением с произвольных начальных приближений.

В комплексе программ анализа интегральных схем (КОМПРОМИС) задача статического анализа сводится к решению системы нелинейных уравнений

$$f_j(x) = 0, \quad j = 1, \dots, n, \quad (1)$$

где $f_j(x): R^n \rightarrow R$, а компоненты $x^{(1)}, x^{(2)}, \dots, x^{(n)}$ вектора x — напряжения на открытых переходах транзисторов и диодов [2]. Решение системы (1) можно искать как такой минимум \bar{x} функции

$$f(x) = \sum_{j=1}^n f_j^2(x), \quad (2)$$

что $f(\bar{x}) = 0$. Для минимизации (2) применялся метод Флетчера-Пауэлла [1]. При этом оказалось целесообразным применить следующий прием.

Для каждого перехода транзисторов и диодов вычислялась константа — минимальное напряжение, при котором переход можно считать открытым (током через элемент DI перехода нельзя пренебречь). Если в процессе минимизации (2) некоторое $x^{(k)}$ — напряжение на открытом переходе — становилось значительно меньше соответствующей константы или напряжение на закрытом переходе становилось больше констан-

ты, то система (I) переформировывалась: менялись функции f_j и порядок n системы. Каждый закрывающийся переход уменьшал n на 1, открывающийся - на 1 увеличивал. И минимизация продолжалась далее с новыми функциями f_j в (2).

Этим приемом удалось значительно сократить время минимизации и просчитать схемы, которые невозможно было просчитать раньше.

Константы - минимальные напряжения, при которых переход открыт, положительны. С другой стороны, по физическому смыслу напряжения на открытых переходах меньше единицы. Таким образом

$$0 < x^{(k)} < 1, \quad k = 1, 2, \dots, n. \quad (3)$$

Неравенства (3) выделяют область, в которой целесообразно искать решение. При нарушении условия $x^{(k)} < 1$ функции f_j очень велики (возможны переполнения), а если $x^{(k)}$ меньше соответствующей константы, скорость сходимости резко уменьшается.

Метод Ньютона [I], примененный для решения (I) там, где сходилос, сходилос очень быстро. Часто, однако, при вычислении следующего $l+1$ -го приближения по формуле [I]

$$x_{l+1} = x_l + \Delta x_l \quad (4)$$

оказывалось, что компоненты вектора Δx_l велики (вплоть до $10^5 \div 10^7$ по модулю) и переставало выполняться условие (3).

Обозначим $g(x)$ градиент функции (2), верхним индексом T - транспонирование, через $F(x)$ - вектор $F(x) = (f_1(x), f_2(x), \dots, f_n(x))^T$, через $F'(x)$ - квадратную матрицу

$$F'(x) = (\psi_{jk}) = \left(\frac{\partial f_j(x)}{\partial x^{(k)}} \right), \quad j, k = 1, 2, \dots, n.$$

Известно, что

$$(g(x_i), \Delta x_i) < 0. \quad (5)$$

Это следует из того, что

$$\begin{aligned} g(x) &= \left(\frac{\partial f(x)}{\partial x^{(1)}}, \frac{\partial f(x)}{\partial x^{(2)}}, \dots, \frac{\partial f(x)}{\partial x^{(n)}} \right)^T = \\ &= \left(2 \sum f_j \frac{\partial f_j}{\partial x^{(1)}}, 2 \sum f_j \frac{\partial f_j}{\partial x^{(2)}}, \dots, 2 \sum f_j \frac{\partial f_j}{\partial x^{(n)}} \right)^T = 2(F'(x))^T \cdot F(x), \end{aligned}$$

Так как (из [1])

$$F'(x_i) \Delta x_i = -F(x_i),$$

получаем

$$\begin{aligned} (g(x_i), \Delta x_i) &= 2((F'(x_i))^T F(x_i), \Delta x_i) = \\ &= 2(F(x_i), F'(x_i) \Delta x_i) = -2(F(x_i), F(x_i)), \end{aligned}$$

что и доказывает (5).

Из (5) следует, что существует $\alpha_i > 0$ такое, что, положив

$$x_{i+1} = x_i + \alpha_i \Delta x_i, \quad (6)$$

имеем

$$f(x_{i+1}) < f(x_i). \quad (7)$$

Известные в литературе способы предлагают вычислять α_i как минимум функции f вдоль направления Δx_i , либо некоторым другим способом, чтобы выполнялось (7). При этом теряется преимущество метода Ньютона [1] - отсутствие поиска по направлению и может значительно увеличиться число шагов.

Мы применяли (6'), полагая $\alpha_i = 1$, когда все компоненты Δx_i по модулю не больше некоторой константы $\delta > 0$, либо, в противном случае, вычисляли α_i так, чтобы максимальная по модулю компонента $\alpha_i \Delta x_i$ была равна δ .

На разных схемах операционных усилителей изучалось влияние величины δ на скорость сходимости. Благодаря (3), удалось быстро подобрать δ . Лучшие результаты (в смысле надежной сходимости и малого числа вычислений функций (1)) получились при

$$\delta = 0,05.$$

Для получения решения модифицированному таким образом методу Ньютона требовалось в среднем в n раз меньше шагов, чем методу Флетчера-Пауэлла [1]. При этом на первых шагах, как правило, не выполнялось (7), α_i было немного меньше 1, но $f(x_{i+1})$ оставалось того же порядка, что $f(x_i)$.

В некоторых случаях дополнительное ускорение сходимости в 2-3 раза удалось получить следующим приемом. Если на m шагах подряд требовалось максимально изменять (в одну и ту же сторону) одну и ту же компоненту, то на m -том шаге δ_m полагалось равным $1,5^{m-1} \delta_1$.

Эта модификация метода Ньютона позволила просчитать большое число схем операционных усилителей с усилением до нескольких миллионов.

Литература

1. Кельман Е.С. Исследование методов решения задачи статического анализа электронных схем на тестовых функциях. В настоящем сборнике, с.42-57.
2. Гришберг Э.И., Мацнельсон Л.Э., Осолить И.Э., Ванато А.А., Зильте А.А., Кельман Е.С., Похвалина Л.С. Моделирование интегральных схем на ЭВМ. - В кн.: Латв.мат.ежегодник. 21, Рига, Зинатне, 1977, с.136-151.

ЧИСЛЕННАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА ОПТИМАЛЬНОГО СИНТЕЗА НИЗКОДОБРОТНЫХ ФИЛЬТРОВ

М.М.Неизвестный

Вычислительный центр ЛГУ им.П.Стучки

В статье рассмотрим некоторые результаты численной реализации метода оптимального синтеза фильтров с низкими добротностями, сформулированного в работе [1] сборника.

Изложим кратко эту задачу и схему ее решения.

Нам нужно среди полиномов F степени n с положительным старшим коэффициентом α , удовлетворяющих условиям:

$$1) F_m = \min_{y \in [0,1]} F(y) = 1,$$

$$2) 0 \leq q_j = \frac{\sqrt{C_{2j}}}{2\sqrt{C_{2j}-C_{1j}}} \leq q_j, \quad j \in [1:k]$$

где четное $n=2k$ и для простоты считаем, что

$$F(y) = \alpha \prod_{j=1}^k (y^2 - C_{1j}y + C_{2j}),$$

$$3) F(y) \geq 1 \quad \text{для всех } y \geq 0,$$

найти тот, будем называть его оптимальным, для которого величина

$$F_m = \max_{y \in [0,1]} F(y)$$

принимает минимальное значение.

Величины $q_j, j \in [1:k]$ будем называть добротами полинома F , а число q_j - величиной ограничения на доброты.

Если максимальная добротность q_j полинома Чебышева (преобразованного соответствующим образом) не превосходит величины ограничения q_j , то он является решением поставленной задачи. Поэтому предлагаемый метод используется при $q_j < q_j$. Если $q_j \leq \frac{1}{2}$, то решение также выписывается в

явной форме [1]. Итак, при $\frac{1}{2} < q < q_7$ будем искать решение в виде

$$F(y) - 1 = \left[\prod_{j=1}^r (y^2 - A_j y + B_j) \right] (\alpha y^m + x_1 y^{m-1} + \dots + x_m) - 1 = \\ = \left[\prod_{j=1}^r (y^2 - A_j y + B_j) \right] (x_1 y^{m-1} + \dots + x_m) - [1 - \alpha y^m \prod_{j=1}^r (y^2 - A_j y + B_j)],$$

где r - число множителей оптимального полинома, имеющих доброту, равную q ,

$$A_j = (2 - \frac{1}{q}) \sqrt{B_j}, \quad j \in [1:r],$$

$$m = n - 2r,$$

доброты полинома $x_1 y^{m-1} + \dots + x_m$ строго меньше q .
Определим в случае $r < k$ функционал $\varphi: R_r \rightarrow R$

$$\varphi(B) = \min_{\{x\}} \max_{y \in [0,1]} F(y) - 1,$$

где $\{x\}$ - это множество полиномов степени $m-1$,
 $B = (B_1, \dots, B_r)$ - вектор свободных членов. Для вычисления значения функционала φ в точке B необходимо решить задачу равномерного приближения сверху функции

$$1 - \alpha y^m \prod_{j=1}^r (y^2 - A_j y + B_j)$$

полиномами с лесом

$$\left[\prod_{j=1}^r (y^2 - A_j y + B_j) \right] (x_1 y^{m-1} + \dots + x_m).$$

При этом используется несколько видоизмененный алгоритм Ремеза, изложенный, например, в [2]. Функционал φ интересует нас только в положительном ортанте R_r^+ пространства R_r , так как величины $B_j, j \in [1:r]$ все положитель-
ные.

Для нахождения оптимального полинома F^* и соответствующей величины F_M^* остается проминимизировать функционал φ . При этом в точке минимума должны выполняться условия 2) и 3).

Случай $r=k$ в указанную схему не укладывается, так как теряет смысл определение функционала φ . Однако проведенные исследования делают правдоподобной гипотезу:

Случай $r=k$ имеет место только тогда, когда

$$q \leq \frac{\sqrt{4+\sqrt{2}}}{2(\sqrt{4+\sqrt{2}}-\sqrt{2})},$$

при этом если $q \geq \frac{1}{2}$, то

$$F^*(y) = \alpha \left[y^2 - \frac{2(2q-1)}{\sqrt{2}\sqrt{4q-1}} y + \frac{4q^2}{\sqrt{2}(4q-1)} \right]^k,$$

если $q \leq \frac{1}{2}$, то

$$F^*(y) = \alpha \left(y^2 - \frac{2q-1}{\sqrt{2}q} y + \frac{1}{\sqrt{2}} \right)^k.$$

Вторая часть сформулированной гипотезы установлена [1]. Во всех остальных случаях имеет место $r < k$ и для получения решений можно воспользоваться описанным методом.

При этом поскольку значение r априори неизвестно и нет метода, гарантирующего получение глобального минимума функционала φ , возникают определенные вычислительные трудности.

Для поиска решения нужно было бы провести вычисления для значений r равных 1, 2 и т.д. до первого такого значения $r=r_*$, когда выполняются все условия на искомый полином. Соответствующий полином и будет оптимальным.

Организуя вычислительную работу по указанному методу, мы использовали предположение о том, что оптимальный полином F^* непрерывно зависит от величины ограничения q . Установить такой факт представляется задачей, эквивалентной по сложности доказательству или опровержению теоремы единственности оптимального полинома, и во всей полноте этого сделать не удалось. В связи с этим предположением мы считаем, что при уменьшения величины q значение r может только возрастать и, значит, существуют числа

$$\frac{1}{2} = q^{(0)} < q^{(1)} = \frac{\sqrt{4+\sqrt{\alpha}}}{2(\sqrt{4+\sqrt{\alpha}} - \sqrt{\alpha})} < q^{(2)} < \dots < q^{(k-1)} < q^{(k)} = q_*$$

такие, что при

$$q \in [q^{(j)}, q^{(j+1)}], j \in [1:k-1]$$

$$r = k - j.$$

Поэтому, если задана величина $q = \bar{q}$ ограничения на доброту, при которой нужно определить оптимальный полином F^* , начинаем расчеты с значения $q = q_*$. Затем, уменьшая значение q с некоторым, вообще говоря переменным, шагом $h\bar{q}$ и решая задачу при каждом таком q , определяем с нужной точностью значения констант $q^{(j)}$, $j \in [1:k-1]$. После этого, узнав при каком j выполняется включение $\bar{q} \in (q^{(j)}, q^{(j+1)})$, получаем оптимальный полином для нужного значения $q = \bar{q}$. Вычислив при данном α константы $q^{(j)}$, $j \in [1:k-1]$, без больших вычислительных затрат, можем находить оптимальный полином для произвольных значений q .

Выполнение свойства 3), которому должен удовлетворять искомый полином по условию физической реализуемости, про-

вернем только для окончательного ответа. Проведенные расчеты показывают, что то условие всегда соблюдается, если выполнены все остальные требования.

Практически бывает удобно задавать не старший коэффициент α полинома F , а его значение $C > 1$ при некотором $y = h > 1$. Требуется при этом решить задачу, аналогичную изложенной выше. Нам не известно доказательство эквивалентности таких задач. Однако решения их в практически интересных случаях при соответствующем выборе параметров α, h, C совпадают. Для решения второй задачи используется точно такой же метод, как и для первой, только вместо функционала φ определяется функционал $\varphi: R_+ \rightarrow R$, который отличается выражением для $F(y) - 1$. В этом случае

$$F(y) - 1 = [(y-h) \prod_{j=1}^n (y^2 - A_j y + B_j)] (x, y^{m-1} + \dots + x_m) - [1 - C \prod_{j=1}^n \frac{y^2 - A_j y + B_j}{h^2 - A_j h + B_j}]$$

$$\varphi(B) = \min_{\{x\}} \max_{y \in [a, 1]} F(y) - 1.$$

Далее приводятся оптимальные полиномы для второй задачи, полученные при $n=6$, $h=2,25$, $C=3090$ и различных значениях величины ограничения на доброту q .

Чебышевский полином с такими параметрами

$$F(y) = 1 + \varepsilon^2 [\cos 3 \arccos(2y-1)]^2, \quad \varepsilon^2 = 0.11916978$$

имеет максимальную доброту $q_T > 40$ и значение $F_m = 1.12$.

Таблица 1

Доброты и значения r в зависимости от q

$q = q_3$	40	35	10	7.5	2.25	1.82	0.5
q_2	3.342	3.606	7.118	7.5	2.25	1.82	0.5
q_1	0.521	0.530	1.057	1.179	1.513	1.82	0.5
r	1	1	1	2	2	3	3

В таблице 1 приведены доброты и значения r полиномов, соответствующих различным $q < q_T$. Для приведенного примера

$$q^{(1)} = 1.82,$$

$$7.5 < q^{(2)} < 10,$$

$$40 < q^{(3)} = q_T.$$

Таблица 2

Зависимость F_M от q

q	q_T	40	35	10	7.5	2.25	1.82	0.5
F_M	1.12	1.14	1.20	1.97	2.32	8.6	9.4	49.5

В таблице 2 приведены значения соответствующих величин F_M (нервномерность в безразмерных величинах).

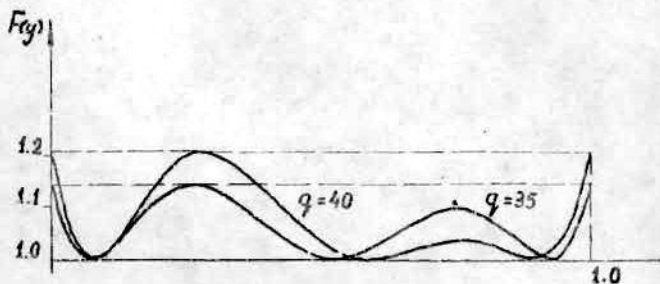


Рис. 1. Оптимальные полиномы при $q = 35$ и $q = 40$.

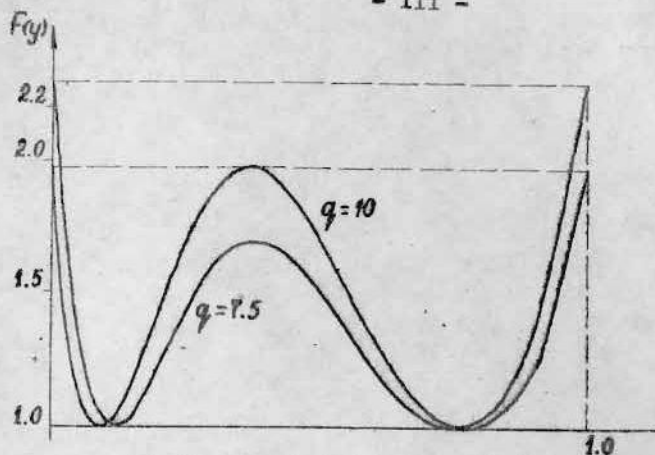


Рис.2. Оптимальные полиномы при $q=7,5$ и $q=10$.

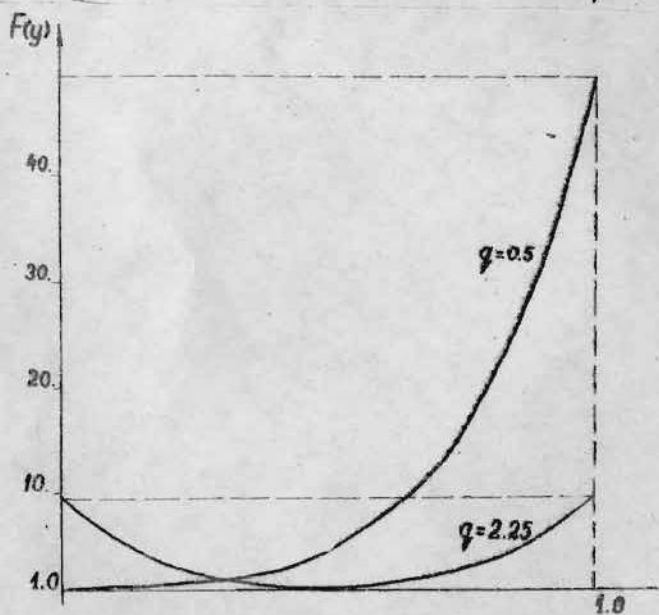


Рис.3. Оптимальные полиномы при $q=0,5$ и $q=2,25$.

На рисунках 1,2,3 изображены оптимальные полиномы при различных значениях q .

Литература

1. Неизвестный М.М. Метод оптимального синтеза полиномиальных фильтров с ограничениями на добротность. - В настоящем сборнике. с. 71 - 79.
2. Демьянов В.Ф., Малоземов В.Н. Введение в минимакс. М., Наука, 1972.

ПОЛИНОМИАЛЬНЫЕ ИНВАРИАНТЫ ГРАФА

А.О.Кац

Вычислительный центр ЛГУ им. П.Стучки

Настоящая работа посвящена изучению общих свойств полиномиальных инвариантов обыкновенных графов. Из-за ограниченного объема здесь приводятся только основные результаты. Полный текст статьи со всеми доказательствами и определениями на основных понятиях будет опубликован в Латвийском математическом ежегоднике.

Под полиномиальным инвариантом графа будем понимать полином от одной или нескольких переменных, связанный с графом, значения коэффициентов которого инвариантны относительно перенумерации вершин графа.

Пусть G — обыкновенный граф с вершинами, занумерованными от 1 до n . Рассмотрим следующий полиномиальный инвариант

$$V_G(\lambda, \delta, \alpha, \beta) = \text{Det}(\lambda E + \delta \Delta + \alpha A + \beta B). \quad (1)$$

Здесь $\lambda, \delta, \alpha, \beta$ — независимые переменные, A — матрица смежности графа G , B — матрица смежности дополнительного графа \bar{G} , E — единичная матрица, $\Delta = \{\delta_i\}_n$ — диагональная матрица, у которой элемент δ_i равен степени вершины i графа G .

Определители вида (1) встречаются при исследовании электрических цепей, динамических систем цепей Маркова, систем массового обслуживания и т.д. Непосредственно на языке теории графов некоторые частные случаи (1) уже изучены ранее.

Э.Я.Гринбергом [2] исследован определитель

$$Q_G(\lambda) = V_G(0, 1, -1, 0). \quad (2)$$

Л.М.Лихтенбаумом [4] - полином

$$Q_G(\lambda) = V_G(\lambda, 1, -1, 0), \quad (3)$$

А.К.Кельмансом [3] - полином

$$B_\lambda^n(G) = \frac{1}{\lambda} V_G(\lambda, 1, -1, 0), \quad (4)$$

автором [1] - полином

$$P_G(\lambda) = V_G(\lambda, 0, -1, 0). \quad (5)$$

Полиномиальные инварианты содержат различную информацию о графе. Так, известен геометрический смысл коэффициентов полиномов $P_G(\lambda)$ [1], $B_\lambda^n(G)$ [3], главных миноров Q_G [2]. Таким образом, связи (2)-(5) между полиномами устанавливают зависимости между инвариантами графа, например, такими как веса систем контуров и числе корневых лесов в случае орграфа, веса систем циклов и веса лесов и деревьев в случае неграфа в классе однородных графов.

Области применения полиномиальных инвариантов в задачах различения графов на изоморфизм различны. Однако, имеет место утверждение:

Над классом обыкновенных однородных графов среди полиномиальных инвариантов типа (I) нет инварианта сильнее характеристического полинома (5).

Приведенное утверждение является следствием соотношения между полиномиальным инвариантом (I) и характеристическим полиномом (5) для однородных степени p графов:

$$V_G(\lambda, \delta, \alpha, \beta) = \frac{\lambda + p\delta + p\alpha + (n-p-1)\beta}{\lambda + p\delta + p\alpha - (p+1)\beta} (-\alpha + \beta)^n P_G\left(\frac{\lambda + p\delta - \beta}{-\alpha + \beta}\right).$$

Далее рассмотрим класс всех обыкновенных графов. В этом классе полиномиальный инвариант (I) четырех переменных однозначно выражается двумя полиномиальными инвариантами двух переменных, а именно:

$$V_G(\lambda, \delta, \alpha, \beta) = \frac{\alpha}{\alpha - \beta} (\beta - \alpha)^n V_G\left(\frac{\lambda - \beta}{\beta - \alpha}, \frac{\delta}{\beta - \alpha}, -1, 0\right) - \frac{\beta}{\alpha - \beta} (\alpha - \beta)^n V_G\left(\frac{\lambda - \alpha}{\alpha - \beta}, \frac{\delta}{\alpha - \beta}, 0, -1\right). \quad (6)$$

Если же при сравнении графов не учитывать непосредственно степени вершин, т.е. рассматривать полиномиальные инварианты типа $V_G(\lambda, 0, \alpha, \beta)$, то из (6) очевидно следующее утверждение:

Над классом обыкновенных графов полиномиальный инвариант $V_G(\lambda, 0, \alpha, \beta)$ равносителен паре инвариантов - характеристическому полиному графа G и характеристическому полиному дополнительного графа \bar{G} .

Полином $B_\lambda^n(G)$ (4) учитывает непосредственно степени вершин графа ($\delta=1$), таким образом тройка полиномиальных инвариантов

$$P_G(\lambda), P_{\bar{G}}(\lambda), B_\lambda^n(G) \quad (7)$$

демонстрирует силу полиномиальных инвариантов типа (I) при установлении изоморфизма графов.

Нами вычислялись полиномы (7) для различных семейств неизоморфных графов. Результаты следующие:

Неизоморфные графы с совпадающими $P_G(\lambda)$ и $P_{\bar{G}}(\lambda)^x$ различаются полиномами $B_\lambda^n(G)$ и, наоборот, неизоморфные графы с совпадающими $B_\lambda^n(G)$ [3] различаются полиномами $P_G(\lambda)$ и(или) $P_{\bar{G}}(\lambda)$.

x Построение одного семейства таких графов дается в полном тексте статьи.

Литература

1. Беце А.О. (Кац А.О.). О коэффициентах характеристического полинома графа.-В кн.: Латв.мат.ежегодник, 3. Рига, Зинатне, 1968, с.75-80.
2. Гринберг Э.Я. Анализ марковских процессов. Алгоритмы вычислений.-В кн.: Новости медицинского приборостроения, М., 1971, вып.1, с.45-48.
3. Кельманс А.К. О свойствах характеристического многочлена графа.-В кн.: Кибернетику - на службу коммунизму. М.-Л., Энергия, 1967, с.24-41.
4. Лихтенбаум Л.М. Новые теоремы о графах.-Сибирский математический журнал, 1962, т.3, № 4, с.561-568.

О ЗАТЫЛОВАНИИ РЕЖУЩЕГО ИНСТРУМЕНТА

И.И.Кан-Каган
Вычислительный центр ЛГУ им.П.Стучки

Цель настоящей статьи — продемонстрировать строгую постановку и строгое же решение задачи о затыловании режущего инструмента. Делается это лишь для простейшего случая, а именно — для случая фасонной дисковой фрезы с передним углом $\gamma=0$ и с профилем, который не должен меняться при переточках^{*}.

Будем пользоваться двумя системами координат. В качестве первой выберем произвольную систему $Oxyz$ декартовых прямоугольных координат с осью Oz , направленной по оси фрезы. Вторая — система цилиндрических координат, начало которой совпадает с O и ось которой направлена по Oz . Эти системы будут использоваться параллельно без специальных оговорок.

Пусть в цилиндрических координатах уравнение задней поверхности зуба фрезы

$$\rho = \rho(z, \varphi)$$

Так как передний угол фрезы предположен равным 0 , то после всякой переточки (и до переточек) передняя грань зуба представляет собой плоскость, описываемую уравнением

$$\varphi = \varphi_0$$

Поэтому уравнения режущей кромки должны записываться в виде

$$\begin{cases} \varphi = \varphi_0 \\ \rho = \rho(z, \varphi_0) \end{cases} \quad (I)$$

* Если фасонная фреза предназначена для обработки, например, винтовой поверхности, то профиль ее, вообще говоря, зависит от диаметра и должен, следовательно, меняться по мере уменьшения диаметра в результате переточек.

С другой стороны, из предположения о неизменности профиля фрезы при переточках следует, что поворот плоскости передней грани зуба на угол φ должен вызывать одно и то же изменение величины ρ при всех z . Но тогда уравнение задней поверхности зуба фрезы должно иметь вид

$$\rho(z, \varphi) = f_1(z) + f_2(\varphi) \quad (2)$$

Здесь функция $f_1(z)$ известна. Она определяется профилем фрезы. Функция же $f_2(\varphi)$ должна быть найдена.

Эта последняя функция ищется, исходя из условия неизменности заднего угла фрезы при переточках. Задний угол представляется естественным измерять в плоскости, нормальной к режущей кромке. При этом угол может, вообще говоря, меняться от одной точки кромки к другой. В дальнейшем будет, однако, рассматриваться не задний угол, а его дополнение до 90° . Этот дополнительный угол обозначим $F(z)$. Функция $F(z)$ должна быть задана, как и профиль фрезы (функция $f_1(z)$).

В декартовой системе координат задняя поверхность зуба фрезы может быть задана уравнениями

$$\begin{aligned} x &= \rho(z, \varphi) \cos \varphi = [f_1(z) + f_2(\varphi)] \cos \varphi \\ y &= \rho(z, \varphi) \sin \varphi = [f_1(z) + f_2(\varphi)] \sin \varphi, \end{aligned} \quad (3)$$

в которых z, φ — суть независимые параметры.

Далее, декартовы координаты точки режущей кромки зуба с цилиндрическими координатами $(z_0, \varphi_0, \rho(z_0, \varphi_0))$ суть

$$\begin{cases} x = [f_1(z_0) + f_2(\varphi_0)] \cos \varphi_0 \\ y = [f_1(z_0) + f_2(\varphi_0)] \sin \varphi_0 \\ z = z_0. \end{cases} \quad (4)$$

Кроме того, из (3) следует

$$\left. \frac{\partial x}{\partial z} \right|_{\substack{z=z_0 \\ \varphi=\varphi_0}} = f_1'(z_0) \cos \varphi_0$$

$$\left. \frac{\partial y}{\partial z} \right|_{\substack{z=z_0 \\ \varphi=\varphi_0}} = f_1'(z_0) \sin \varphi_0 \quad (5)$$

$$\frac{\partial z}{\partial z} = 1$$

Полученные выражения - не что иное, как координаты вектора касательной в точке (4) к проходящей через эту точку режущей кромке, так как для всех точек режущей кромки величина φ постоянна. Отсюда уравнение в декартовых координатах нормальной плоскости, проведенной через точку (4) к режущей кромке, проходящей через эту точку

$$\{ [f_1(z_0) + f_2(\varphi_0)] \cos \varphi_0 - X \} f_1'(z_0) \cos \varphi_0 + \{ [f_1(z_0) + f_2(\varphi_0)] \sin \varphi_0 - Y \} \times \\ \times f_1'(z_0) \sin \varphi_0 + z_0 - Z = 0, \quad (6)$$

где X, Y, Z - текущие координаты точки плоскости.

Уравнение же в декартовых координатах плоскости $\varphi = \varphi_0$

$$X \sin \varphi_0 - Y \cos \varphi_0 = 0 \quad (7)$$

где X, Y - снова текущие координаты точки плоскости.

Легко видеть, что плоскости (6) и (7) пересекаются по прямой, уравнения которой

$$\begin{cases} Y = X \operatorname{tg} \varphi_0 \\ Z = z_0 + [f_1(z_0) + f_2(\varphi_0)] f_1'(z_0) - X \cdot \frac{f_1'(z_0)}{\cos \varphi_0} \end{cases} \quad (8)$$

Здесь X, Y, Z — текущие координаты точки прямой, причем X играет роль параметра.

Подставим теперь в (6) вместо X, Y, Z соответственно выражения для x, y из (3) и z . Получим после несложных преобразований

$$[f_1(z_0) + f_2(\varphi_0)] f_1'(z_0) - [f_1(z) + f_2(\varphi)] f_1'(z) \cos(\varphi - \varphi_0) + z_0 - z = 0. \quad (9)$$

Это уравнение устанавливает ту зависимость между параметрами φ и z , при наличии которой уравнения (3) превращаются в уравнения линии пересечения плоскости (6) с вадней поверхностью зуба.

Из (9) следует

$$\frac{d\varphi}{dz} = - \frac{f_1'(z) f_1'(z_0) \cos(\varphi - \varphi_0) + 1}{f_1'(z_0) \{ f_1'(\varphi) \cos(\varphi - \varphi_0) - [f_1(z) + f_2(\varphi)] \sin(\varphi - \varphi_0) \}}$$

$$\left. \frac{d\varphi}{dz} \right|_{\substack{z=z_0 \\ \varphi=\varphi_0}} = - \frac{[f_1'(z_0)]^2 + 1}{f_1'(z_0) f_1'(\varphi_0)}$$

(10)

Теперь, дифференцируя (3) по z , получаем

$$\left. \frac{dx}{dz} \right|_{\substack{z=z_0 \\ \varphi=\varphi_0}} = \left[f_1'(z_0) - \frac{[f_1'(z_0)]^2 + 1}{f_1'(z_0)} \right] \cos \varphi_0 + [f_1(z_0) + f_2(\varphi_0)] \cdot \frac{[f_1'(z_0)]^2 + 1}{f_1'(z_0) f_2'(\varphi_0)} \sin \varphi_0$$

$$\left. \frac{dy}{dz} \right|_{\substack{z=z_0 \\ \varphi=\varphi_0}} = \left[f_1'(z_0) - \frac{[f_1'(z_0)]^2 + 1}{f_1'(z_0)} \right] \sin \varphi_0 - [f_1(z_0) + f_2(\varphi_0)] \cdot \frac{[f_1'(z_0)]^2 + 1}{f_1'(z_0) f_2'(\varphi_0)} \cos \varphi_0$$

$$\left. \frac{dz}{dz} \right|_{\substack{z=z_0 \\ \varphi=\varphi_0}} = 1 \tag{II}$$

Выражениями (II) даются координаты вектора касательной в точке (4) к линии пересечения плоскости (6) с заданной поверхностью зуба.

Очевидно, угол между прямой (8) и вектором (II) должен быть равен $F(z)$.

Координаты направляющего вектора прямой (8) суть

$$\begin{aligned} \frac{dX}{dX} &= 1 \\ \frac{dY}{dX} &= \operatorname{tg} \varphi_0 \\ \frac{dZ}{dX} &= -\frac{f_1'(z_0)}{\cos \varphi_0} \end{aligned} \tag{I2}$$

Беря скалярное произведение векторов (II) и (I2) и дели его на произведение длин этих векторов, получаем

$$\begin{aligned} & \frac{1}{\cos \varphi_0} \cdot \frac{1}{f_1'(z_0)} + \frac{f_1(z_0)}{\cos \varphi_0} \\ & \sqrt{1 + \frac{1}{[f_1'(z_0)]^2} + [f_1(z_0) + f_2(\varphi_0)]^2 \left\{ \frac{[f_1'(z_0)]^2 + 1}{f_1'(z_0) f_2'(\varphi_0)} \right\}^2} \cdot \sqrt{1 + \operatorname{tg}^2 \varphi_0 + \frac{[f_1'(z_0)]^2}{\cos^2 \varphi_0}} = \\ & = \cos[F(z_0)]. \end{aligned}$$

Поскольку последнее равенство должно выполняться при любых z_0 и φ_0 , заменяем эти величины текущими координатами и получаем после несложных преобразований

$$\frac{f_2'(\varphi)}{\sqrt{[f_2'(\varphi)]^2 + [f_1(z) + f_2(\varphi)]^2 \{ [f_1'(z)]^2 + 1 \}}} = \cos [F(z)]. \quad (13)$$

Это и есть уравнение для отыскания $f_2(\varphi)$.

Существует ли при всех $f_1(z)$ и $F(z)$ решение такого уравнения, неизвестно, и представляется правильным искать в том или ином наперед заданном классе функций такую функцию $f_2(\varphi)$, чтобы отклонение левой части (13) от правой части было в каком-то смысле минимальным при изменении z и φ в некоторых заданных пределах.

Была предпринята попытка для выпуклой радиусной фрезы найти такую функцию среди полиномов. Однако требуемые вычисления ввиду своей трудоемкости оказались невыполнимыми даже на ЭВМ.

Таким образом, задача о затыловании инструмента оказывается далеко не простой даже в простейшем случае. Тем больших трудностей, естественно, ожидать в случаях, когда профиль инструмента должен меняться при переточках. Поскольку неизвестно, разрешимо ли уравнение (13), следует считать совершенно беспредметными рассуждения о теоретически точной затылованной поверхности. Вопрос о существовании таковой в настоящее время является открытым, а если и будет решен, то скорее всего отрицательно.

По-видимому, в качестве ближайшей цели можно выдвинуть поиск не слишком трудоемких методов приближенного решения уравнения (13).

УДК 681.3.06:621.396.6

КОМПЛЕКС ПРОГРАММ ДЛЯ МОДЕЛИРОВАНИЯ ИНТЕГРАЛЬНЫХ СХЕМ (КОМПРОМИС). Гринберг Э.Я., Кацнельсон Л.Э., Озолить И.Э., Ванег А.А., Зилите А.А., Заворин А.Н., Кельман Е.С., Похвалина Л.С., Кан-Каган И.И. Математическое моделирование, 1978.

Описывается комплекс программ для моделирования нелинейных и линейных интегральных схем. Комплекс позволяет проводить для нелинейных схем статический, динамический, статистический анализ, для линейных схем - частотный и шумовой анализ, расчет нулей и полюсов передаточной характеристики, а также статистический анализ для сквозного (нелинейного и линейного) анализа схем.

Библиография (5), рис.2.

УДК 518:512.8

О НЕКОТОРЫХ МЕТОДАХ РЕШЕНИЯ РАЗРЕЖЕННЫХ СИСТЕМ УЗЛОВЫХ УРАВНЕНИЙ. Заворин А.Н. Математическое моделирование, 1978.

В работе анализируется эффективность некоторых прямых и итерационных методов и основанных на них алгоритмов и программы при решении больших разреженных систем линейных алгебраических уравнений, описывающих электрические цепи.

Библиография (6).

УДК 518:517.91/.94

ВЫБОР ШАГА ПРИ РЕШЕНИИ ЖЕСТКИХ СИСТЕМ ДИФФЕРЕНЦИАЛЬНЫХ УРАВНЕНИЙ ЯВНЫМИ МЕТОДАМИ. Заворин А.Н., Мамонтов Е.В. Математическое моделирование, 1978.

Исследуется способ интегрирования устойчивых жестких систем дифференциальных уравнений методами Рунге-Кутты первого и третьего порядков с использованием серий шагов, длины которых выбираются специальным образом. Если спектр собственных значений некоторой системы распадается на две

группы (больших и малых по модулю), то интегрирование сериями шагов существенно повышает эффективность указанных методов.

Библиография (1).

УДК 518.12:621.372.061

ИССЛЕДОВАНИЕ МЕТОДОВ РЕШЕНИЯ ЗАДАЧИ СТАТИЧЕСКОГО АНАЛИЗА ЭЛЕКТРОННЫХ СХЕМ НА ТЕСТОВЫХ ФУНКЦИЯХ. Кельман В.С. Математическое моделирование, . 1978.

В работе описываются численные эксперименты с тремя методами - методом Ньютона для решения уравнений, методом Флетчера-Пауэлла, методом сопряженных градиентов - на десяти тестовых функциях. Проведено сравнение методов и показано, что все три метода могут быть успешно использованы для поиска минимумов, когда производные вычисляются по конечноразностной формуле. Проведенные исследования позволили улучшить методы, применяемые в системе анализа электронных схем (КОМПРОМИС).

Библиография (6), табл.8.

УДК 621.372.54

ИЗОЭКСТРЕМАЛЬНЫЕ ДРОБИ ЧЕБЫШЕВА ДЛЯ ОПТИМАЛЬНОГО СИНТЕЗА ЭЛЕКТРИЧЕСКИХ ЦЕПЕЙ. Гринберг Э.Я., Кацнельсон Л.З., Неизвестный М.М. Математическое моделирование, . 1978.

Приведено краткое описание алгоритмов расчета и выдержки из таблиц для изоэкстремальных дробей Чебышева.

Библиография 4.

УДК 518.6

МЕТОД ОПТИМАЛЬНОГО СИНТЕЗА ПОЛИНОМИАЛЬНЫХ ФИЛЬТРОВ С ОГРАНИЧЕНИЯМИ НА ДОБРОТНОСТЬ. Неизвестный М.М. Математическое моделирование, . 1978.

В статье предлагается один способ численного решения задачи синтеза полиномиальных электрических фильтров с низкими добротностями.

Решение сводится к минимизации некоторых функционалов. Глобальный минимум дает оптимальное решение.

Для случая небольших добротностей ($Q \leq \frac{\sqrt{\epsilon}}{2}$), при которых амплитудно-частотная характеристика является монотонной, приведено явное выражение соответствующих полиномов.

Библиография (2), рис.1.

УДК 621.382.3.001

О СХОДИМОСТИ ИТЕРАЦИОННЫХ МЕТОДОВ ДЛЯ РЕШЕНИЯ СТАЦИОНАРНЫХ ЗАДАЧ ТЕОРИИ ПОЛУПРОВОДНИКОВЫХ ПРИБОРОВ. Маргузан Б.Я., Польский Б.С. Математическое моделирование, 1978.

Предложен итерационный метод численного решения фундаментальной системы уравнений физики полупроводников. На модельной задаче исследована сходимость метода. Эффективность метода показана на примере расчета конкретной диодной структуры. Библиография (7), рис.1.

УДК 621.315.592.3

РАСЧЕТ ОДНОМЕРНОГО ДИФФУЗИОННОГО РАСПРЕДЕЛЕНИЯ БОРА И ФОСФОРА В КРЕМНИИ. Иванов Н.М., Новиков В.И. Математическое моделирование, 1978.

Теоретически рассмотрено диффузионное распределение бора и фосфора в кремнии с учетом диффузии в пленке двуокиси кремния. Уравнения диффузии записаны в движущейся системе координат, соответствующей линейно-параболическому закону роста окисла. Для уравнения, рассматриваемого в области кремния введена концентрационная зависимость коэффициента диффузии. Решение уравнений получено с помощью неявного конечно-разностного метода. Представлены результаты расчета концентраций примеси, сопротивлений диффузионных слоев, глубин R/L переходов. Библиография (6), рис.2.

УДК 518.12:621.372.061

О МОДИФИКАЦИЯХ МЕТОДА НЬЮТОНА ДЛЯ АНАЛИЗА ИНТЕГРАЛЬНЫХ СХЕМ. Кацнельсон Л.Э., Кельман Е.С. Математическое моделирование, 1978.

Предлагаются модификации метода Ньютона, позволившие просчитать ряд практически важных схем с большим усилением с произвольных начальных приближений.

Библиография (2).

УДК 518.6:621.372.54

ЧИСЛЕННАЯ РЕАЛИЗАЦИЯ АЛГОРИТМА ОПТИМАЛЬНОГО СИНТЕЗА НИЗКОДОБРОТНЫХ ФИЛЬТРОВ. Неизвестный М.М. Математическое моделирование, 1978.

Приведены некоторые результаты численных экспериментов по расчету оптимальных коэффициентов передачи полиномиальных фильтров с низкими добротностями. Предлагается способ организации соответствующих вычислений.

Библиография (2), рис. 3, табл.2.

УДК 519.1

ПОЛИНОМИАЛЬНЫЕ ИНВАРИАНТЫ ГРАФА. Кац А.О. Математическое моделирование, 1978.

В работе показано, что над классом обыкновенных однородных графов нет полиномиального инварианта рассматриваемого вида сильнее характеристического полинома графа. Над классом всех обыкновенных графов выделена тройка полиномов, демонстрирующая силу полиномиальных инвариантов.

Библиография (4).

УДК 621.0.013

О ЗАТЫЛОВАНИИ РЕЖУЩЕГО ИНСТРУМЕНТА. Кан-Каган И.И. Математическое моделирование, 1978.

В статье изложены строгая постановка и строгая решение задачи о затыловании режущего инструмента в простейшем случае.

СОДЕРЖАНИЕ

Гринберг Э.Я., Кацнельсон Л.З., Озолинь И.Э., Ванагс А.А., Зилите А.А., Заворин А.Н., Кельман Е.С., Похвалина Л.С., Кан-Каган И.И. Комплекс программ для моделирования интегральных схем (КОМПРОМИС)	3
Заворин А.Н. О некоторых методах решения разреженных систем узловых уравнений	14
Заворин А.Н., Мамонтов Е.В. Выбор шага при решении жестких систем дифференциальных уравнений явными методами ...	31
Кельман Е.С. Исследование методов решения задачи статического анализа электронных схем на тестовых функциях..	41
Гринберг Э.Я., Кацнельсон Л.З., Неизвестный М.М. Изокстремальные дроби Чебышева для оптимального синтеза электрических цепей	57
Неизвестный М.М. Метод оптимального синтеза полиномиальных фильтров с ограничениями на добротность	70
Мартузан Б.Я., Польский Б.С. О сходимости итерационных методов для решения стационарных задач теории полупроводниковых приборов	79
Иванов Н.М., Новиков В.И. Расчет одномерного диффузионного распределения бора и фосфора в кремнии	92
Кацнельсон Л.З., Кельман Е.С. О модификациях метода Ньютона для анализа интегральных схем	101
Неизвестный М.М. Численная реализация алгоритма оптимального синтеза низкодобротных фильтров	105
Кац А.О. Полиномиальные инварианты графа	113
Кан-Каган И.И. О затыловании режущего инструмента	117
Р е ф е р а т ы	123