SYMBOJIC MATHEMATICAL LABORATORY

by

WILLIAM ARTHUR MARTIN

B.S., Massachusetts Institute of Technology
1962

M.S., Massachusetts Institute of Technology
1962

SUBMITTED IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
January, 1967

Signature of Author     WILLIAM A. MARTIN
Department of Electrical Engineering, January 9, 1967

Certified by         Marvely
Thesis Supervisor

Accepted by         Ray
Chairman, Departmental Committee on Graduate Students
SYMBOLIC MATHEMATICAL LABORATORY

by

WILLIAM ARTHUR MARTIN

Submitted to the Department of Electrical Engineering on January 9, 1967 in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

ABSTRACT

A large computer program has been developed to aid applied mathematicians in the solution of problems in non-numerical analysis which involve tedious manipulations of mathematical expressions. The mathematician uses typed commands and a light pen to direct the computer in the application of mathematical transformations; the intermediate results are displayed in standard text-book format so that the system user can decide the next step in the problem solution.

Three problems selected from the literature have been solved to illustrate the use of the system. A detailed analysis of the problems of input, transformation, and display of mathematical expressions is also presented.

Thesis Supervisor: Marvin L. Minsky

Title: Professor of Electrical Engineering
ACKNOWLEDGMENTS

The work herein was supported in part by the M.I.T. Computation Center, and in part by Project MAC, an M.I.T. research program sponsored by the Advanced Research Projects Agency, Department of Defense, under Office of Naval Research Contract number Nonr-4102(01). Reproduction in whole or in part is permitted for any purpose of the United States Government.

The author expresses his gratitude to Marvin Minsky for his supervision of this thesis and to the directors of Project MAC for providing a sympathetic environment where the work could be carried out. I also wish to thank Herbert Teager and Oliver Selfridge for reading the manuscript and Tim Hart for his assistance when I first joined the Artificial Intelligence Group.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Chapter Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ABSTRACT</td>
<td>2</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENT</td>
<td>3</td>
</tr>
</tbody>
</table>

(Chapter Titles Are Followed by a Short Description of Their Contents)

## CHAPTER I. INTRODUCTION

The purpose of the thesis is stated to be the advancement of the computer programming technology necessary for a useful symbolic mathematical laboratory. Other relevant research work is then briefly reviewed for the purpose of establishing the strategy used in planning the work.

## CHAPTER II. A STEP-BY-STEP SOLUTION OF THREE PROBLEMS IN NON-NUMERICAL ANALYSIS

We describe how the programming system is used. The form of the typed commands to the system is specified, and a brief description is given of each of the available mathematical transformations. The step-by-step solution of three problems of current interest is then presented to demonstrate how the system is actually used.

## CHAPTER III. ADDITIONAL GENERAL FEATURES

Additional features are presented which would make the system described in Chapter II more convenient for the user. These features are divided into three classes: methods for adding new mathematical structures, algorithms, and notations; methods for handling error conditions and helping the user to understand the system; and methods for helping the user with the bookkeeping involved in the solution of his problem.
### TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>CHAPTER IV.</th>
<th>INTRODUCTION TO THE REMAINING CHAPTERS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The remaining chapters analyze in greater detail many aspects of the system presented in Chapter II and Chapter III.</td>
<td>82</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER V.</th>
<th>PROBLEM AREAS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A number of areas in mathematics should be investigated in order to enlarge the mathematical laboratory. They are classified according to the degree to which they appear to require new heuristic methods.</td>
<td>83</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER VI.</th>
<th>INTERNAL REPRESENTATION AND MANIPULATION OF MATHEMATICAL EXPRESSIONS</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>The first part of this chapter contains a description of the operations necessary for manipulating mathematical expressions, and some properties of the expressions which determine the flow of program control. Operations on the elementary functions are analyzed and then the need for additional notation for structures involving sets is discussed. The rest of the chapter contains a detailed description of each of the major LISP transformation routines used in the system. This description would be good preliminary reading for workers planning to construct a similar system.</td>
<td>85</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>CHAPTER VII.</th>
<th>HASH-CODING FUNCTIONS OF A COMPLEX VARIABLE</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A common operation in the routines described in Chapter VI is the comparison of symbolic mathematical expressions. This operation was accomplished by &quot;hash-coding&quot; the expressions using finite field arithmetic and then comparing the resulting hash code numbers. The scheme for generating the hash codes is explained in this chapter. This scheme is the most unexpected result of the thesis.</td>
<td>189</td>
</tr>
<tr>
<td>CHAPTER</td>
<td>TITLE</td>
<td>PAGE</td>
</tr>
<tr>
<td>---------</td>
<td>-------------------------------------------</td>
<td>------</td>
</tr>
<tr>
<td>VIII</td>
<td>THE LANGUAGE OF MATHEMATICAL EXPRESSIONS</td>
<td>201</td>
</tr>
<tr>
<td>XI</td>
<td>SYNTAX AND DISPLAY OF MATHEMATICAL EXPRESSIONS</td>
<td>203</td>
</tr>
<tr>
<td>X</td>
<td>LINEAR INPUT OF MATHEMATICAL EXPRESSIONS</td>
<td>225</td>
</tr>
<tr>
<td>XI</td>
<td>TWO-DIMENSIONAL INPUT OF MATHEMATICAL EXPRESSIONS</td>
<td>234</td>
</tr>
<tr>
<td>XII</td>
<td>THE SYSTEM PROGRAMS</td>
<td>238</td>
</tr>
</tbody>
</table>

A system which mathematicians will accept must have a convenient language. This chapter discusses some possibilities for such a language and introduces the discussion of several languages found in the following chapters.

This chapter describes a program for displaying mathematical expressions in standard textbook notation. A complete set of syntax rules for this complex textbook language is given. A table driven picture compiler which uses these syntax rules is then described.

The productions of an operator grammar, which contains the system commands as a subset of its terminal strings, are described. A parsing algorithm for this grammar is then described.

The problem of parsing a two-dimensional array of characters into a mathematical expression is discussed. Two methods of converting the two-dimensional array into a linear string are given. First, some restrictions proposed by Klerer which give the array a linear order are discussed. Then a method of reversing the generation procedure in the previous chapter is given. A method for handling context dependent relationships such as $a$ in $(a)$ is still needed.

The advantages of system programs are listed and the use and acceptance of the system programs described in the following two chapters is discussed.
### TABLE OF CONTENTS (continued)

<table>
<thead>
<tr>
<th>Chapter</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>XIII.</td>
<td>PDP-6 LISP INPUT-OUTPUT FOR THE DATAPHONE</td>
<td>241</td>
</tr>
<tr>
<td></td>
<td>The use of a routine which allows PDP-6</td>
<td></td>
</tr>
<tr>
<td></td>
<td>LISP to communicate with Project MAC time</td>
<td></td>
</tr>
<tr>
<td></td>
<td>sharing is described.</td>
<td></td>
</tr>
<tr>
<td>XIV.</td>
<td>PDP-6 LISP INPUT-OUTPUT FOR THE DISPLAY</td>
<td>246</td>
</tr>
<tr>
<td></td>
<td>The use and implementation of a display</td>
<td></td>
</tr>
<tr>
<td></td>
<td>language which has been embedded in a</td>
<td></td>
</tr>
<tr>
<td></td>
<td>version of PDP-6 LISP is described. This</td>
<td></td>
</tr>
<tr>
<td></td>
<td>language makes it possible to write LISP</td>
<td></td>
</tr>
<tr>
<td></td>
<td>functions which generate displays and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>analyze the motions of the light pen.</td>
<td></td>
</tr>
</tbody>
</table>

| Appendix | Listing of the LISP Programs               | 267  |

| Bibliography |                                             | 331  |
| Biographical Note |                                         | 336  |
Chapter I

INTRODUCTION

The purpose of this thesis is to advance the computer programming technology required for the real time solution of non-numerical problems in applied mathematics. Generally, the first use of computers in any area has been the mechanization of the routine parts of tasks in essentially the same manner as they were formerly done by hand. Later, when the computer resources are better understood, the methods of doing these tasks are altered to make better use of the computer. J. C. R. Licklider(2) has described the "man-computer symbiosis" which would result from the most economical utilization of the abilities of man and currently available machines in a problem solving situation. "Men will make approximate and fallible, but leading contributions, and they will define criteria and serve as evaluators, judging the contributions of the equipment and guiding the general line of thought. In addition, men will handle the very low probability situations when such situations arise. ...The information-processing equipment, for its part, will convert hypotheses into testable models and then test the models against data." Licklider stated further that the equipment will simulate, transform, interpolate, extrapolate, carry out routinizable clerical operations, and remember precise values and exact details.
This degree of man-machine cooperation has been and will continue to be approached by stages. This is because it is very difficult to determine the many small features necessary to make workable a system that differs greatly from those currently in use, and because mathematicians will not use a system which is not clearly more convenient than their tried and true aids, such as pencil and paper. Furthermore, considerable research will be required to determine just how much of mathematics is routinizable. One might envision four roles which the computer could take, each requiring increased responsibility.

The first step would be to help the mathematician overcome straightforward drudgery. This is often done by programming the simple iterative operations needed in straightforward problems in numerical analysis. A second goal would be to help the mathematician work in areas that he understands in principle, but where he does not have the technical facility, gained from long practice and accuracy, which frees him to focus on the creative aspects of the problem. Such a skill is required, for example, for the approximate solution of differential equations using series expansion techniques; the programming system constructed for this thesis operates at this level. A third goal would be to help the mathematician in areas where he is only a beginner--where he knows the names and purposes of the algorithms, but does not know their details. The computer must know and apply criteria which determine what methods should be used on a given example. Some current research on symbolic integration seems to be a step toward systems of this type. (8) Research in computer problem-solving tactics for the use of fixed resources on specific tasks would be useful for systems at this
level. A fourth goal would be the construction of an almost autonomous problem solver; the machine would help formulate the general strategy of the problem solution. It is at this level that any productive research in computer theorem-proving strategy would be most applicable to a working system.

Of course, to build a working system which represents a significant advance toward these goals requires a great deal of research in several areas, often too much for any one project. It is difficult to work much in advance of the general level of programming technology without having to do a lot of ground work, which only builds tools for the project at hand. To develop the system in this thesis, which realizes in the laboratory the second of the above goals, it was necessary to set up systems programs and a programming language environment for real-time interaction with large symbolic programs. It required new techniques for input and output of mathematical expressions. Mathematical expressions were represented as data in the computer, and we had to investigate algorithms needed to transform them. Specific problem areas where mathematicians could most benefit from such a system had to be defined. Working in so many areas, only some of the possibilities could be explored and a fixed strategy with respect to the trade-offs between programmer effort, program efficiency and usefulness, and the number of new ideas had to be followed.

The particular emphasis of the thesis can best be understood through a review of work developed concurrently with it. A bibliography of the work in the field has been prepared by Sammet (1), and no attempt will be made to mention every paper. In the sections to follow some research areas will be listed. Then the different research projects carried out
by other people will be discussed, grouped according to their area of emphasis. In conclusion, the choice of research for this thesis will be stated.

To better use the computer as an aid to the mathematician, there are several lines of research which can be pursued, each affording the possibility of doing something new without overreaching for currently unrealistic levels of overall performance. These are as follows:

1. The provision of computer capability for solving existing problems in mathematics. Special programs might be written for a particular problem. The programs would be simple, but the problem solution would be new.

2. The development of complex programming structures and languages suitable for more complex mathematical algorithms.

3. The development of input-output techniques to bring the mathematician closer to the computer.

4. The discovery of procedures for mechanizing the more routine mathematical transformations.

5. The development of methods for solving problems which are feasible only with computer implementation.

6. The extension of mathematical rigor to computer procedures and languages in the context of mathematical procedures.

There is naturally some difference of opinion as to the usefulness, significance, and feasibility of research in the different areas, and over the past five years more than fifteen people and projects have taken their stand.
Several programs have been written in the LISP 1.5 language, which was developed by the M.I.T. Artificial Intelligence Group. LISP 1.5 provides both numerical and list structure data and automatic storage allocation. In addition, the specification of algorithms as recursive functions is particularly suitable for operating on the elementary functions in analysis, which can be represented as a tree structure of subexpressions. To demonstrate the capabilities of LISP, Hart\(^{(3)}\) wrote a program to put the elementary functions in a simplified form. He had to define a canonical order for expressions representing the elementary functions and he found that not only were some simplifications very hard to program, but that the use of an expression determined its "simplest" form.

Carrying on this work, Russell and Wooldridge\(^{(4)}\) at Stanford, also interested in complicated LISP programs, added new features to cope with the inefficiencies and shortcomings of Hart's program. They introduced user options, such as expansion of products, and special representation for polynomials. They added routines for differentiation. They had ideas for discovering subexpressions such as \(\sin^2 x + \cos^2 x\) and methods to avoid simplifying identical subexpressions more than once, but they quit working and the program was not improved further, probably because each change required many small revisions. A new simplification program was written by Korevold.\(^{(5)}\) He reduced the amount of fixed program structure and implemented a way for adding simplification rules in the form of patterns containing a free variable. He had a complicated strategy for applying the rules, but the matching of the rule to the data only covered simple cases. At the same time, Fenichel\(^{(6)}\) at Harvard recognized the importance of an explicit body of rules; this
body of rules would provide program flexibility. In contrast to the others, Fenichel emphasized the formal power and structure of his program instead of showing how his implementation was an efficient model for the data and algorithms. He emphasized the idea of computation by local transformations.

There are two more examples of complex programs for mathematical expression transformation. While Hart was writing the first of the programs for algebraic simplification, Slagle\(^7\) wrote a program which solved freshman calculus integration problems. Although Slagle primarily wanted to show how heuristics and a judicious choice of subproblems could be used to solve examples in an area for which a general algorithm was not known, he also had to find properties of mathematical expressions which would be useful in selecting a trial transformation. Thus, he wrote a complicated routine for matching mathematical expressions against a pattern. Slagle's program demonstrated a working scheme for the heuristic selection of subgoals and provided food for thought about executive structures, but it misled some people to think that few integration problems could be solved algorithmically. Recently Moses has increased the range of examples which can be integrated with algorithms. Moses\(^8\) has continued the work on pattern matching routines.

Concurrent with this work was the work of those who took a skeptical view of the possibility of creating complex programs which were efficient enough for immediate use. The ALPAK system\(^9\), developed at Bell Telephone Laboratories, is a polynomial manipulation package which has been used to solve problems for the research workers there. It is so
much more efficient than the systems above, that it provides a definite incentive to produce optimum coding for common classes of problems. In addition, it provides mathematicians with a useful facility, but it is limited to problems which can be solved by polynomial manipulation. An excellent polynomial manipulation package has also been developed by Collins\(^{(10)}\) at I.B.M. In his paper, he gives a good discussion of the relative efficiency of different systems. The FORMAC language\(^{(11)}\), also developed at I.B.M., is intended for the efficient solution of problems requiring algebraic manipulation. FORMAC programs are similar to FORTRAN programs and require statements which are more familiar to programmers than to mathematicians, such as the declaration of symbolic variables and the direction of input-output. The language is not as flexible as the LISP systems, but it is faster. Another language is FORMULA ALGOL\(^{(12)}\), which is being developed at Carnegie Tech. It is a version of ALGOL which contains list structure and mathematical expression data types and a built-in pattern matching facility. It contains numerous design decisions, such as the use of an addition operator with only two arguments, which seem controversial.

During this same period, better man-machine communication was developed in certain systems for real-time interaction with programs for numerical analysis. One of the first programs was the JOSS system, developed at RAND in 1963.\(^{(13)}\) The fact that it was implemented on a small machine limited its capability to simple calculations, but it stands as an example of the attention to details of "human engineering" which makes a system pleasant to use. Collier\(^{(14)}\) demonstrated the real-time solution of numerical problems by methods which could not be used...
in a closed shop computer facility, since they required frequent man-
machine interaction. Culler introduced the graphical display of 127
points approximating a function to be transformed numerically. He
attempted to use special keys on a user console to simplify the input
of expressions and commands. The primitives of Culler's system are
very elementary; the user must first combine these operations into
programs which perform the transformation he needs.

On-line numerical analysis is also performed by MAP(15), which
operates at Project MAC. MAP contains error-checking and comments to
guide the user as he learns. It also establishes a set of conventions
for the addition of new subroutines, so that they can be used
conveniently. MATHLAB(16) is the first working program for on-line
non-numerical analysis. It is being developed at the MITRE Corporation
with computer support from Project MAC. The initial version contained
the Russell-Wooldridge simplification program. In addition to the
routines for expression expansion and differentiation, there were
routines for factoring and integration. There are many differences, but
with the exception of the factoring and integration routines, the initial
version of MATHLAB was essentially a subsystem of the system described
here.

There has been only a little work on the mathematics associated
with algebraic symbol manipulation. Several programs have been written
to accomplish a specific task which requires mathematical manipulation,
but these are rather simple from a programming point of view, and there
have not been enough of them to change the applied mathematician's
concept of the resources available to him. Therefore, they have not
resulted in much new mathematics.
To improve the efficiency of his polynomial manipulation package, Collins developed a new greatest common divisor algorithm for polynomials with integer coefficients. Moses has investigated algorithms for solving systems of polynomial equations and for integration. Manove (17) has mechanized the integration of rational functions, employing a program for the factorization of polynomials in several variables over the integers written by Bloom (17). Engelman (17) is expanding his MATHLAB through the inclusion of these routines as well as others that have been written for the manipulation of matrices, direct and inverse Laplace transforms, the solution of equations, and the typewriter display and editing of mathematical expressions. An interesting theoretical result was obtained by Richardson (18), who recently proved that the questions of whether (i) an elementary function is identically zero or (ii) is integrable in closed form, are recursively unsolvable.

The mathematical laboratory described in this thesis is very complex and explores a number of new ideas about hash coding, descriptive properties of mathematical expressions, and system organization. An examination is made of the types of programming structures needed in non-numerical analysis. Some very complicated routines for expression analysis and display are explained. On the other hand, no new mathematical algorithms are introduced and the programs are explained without mathematical formalism. As opposed to the other complex LISP systems, the system was constructed to be immediately useful as an aid to applied mathematicians in the solution of certain problems. As opposed to some of the other problem solving systems, it does not solve these problems in an economical manner. Instead, attention has been
concentrated on providing a rich enough notation for input and display and a great enough variety of mathematical transformations to achieve better man-computer symbiosis during the problem solving procedure. As a demonstration of this man-machine interaction, we will present "scripts" showing the solution of some non-trivial problems in applied mathematics.

The current hardware configuration for the mathematical laboratory is shown in Figure 1.

![Diagram showing hardware configuration](image)

The system is quite inefficient, in terms of the amount of the total computation time required by the slow transmission of information over the dataphone, the frequent disk accesses necessitated by the small 7094 core memory, and the fact that many of the routines run interpretively. These inefficiencies could be eliminated by introducing the entire system into the PDP-6, with the large core memory now available.
The configuration in Figure 1 was used initially as it was the only means then available for obtaining the minimum requirements for the system: a time-shared machine with a large memory connected to a display with fast light pen response. The use of a peripheral computer for input has the disadvantage that the input signals cannot be interpreted by the peripheral computer in terms of the whole data base, which is in the main machine. This limits the immediate feedback which can be used to simplify the user's input task. In addition, the use of two machines complicates the day-to-day routine required to develop a system. The justification for two machines would be the more economical assignment of the different types of computation to be done, or the need to have the user located at a terminal remote from the main machine.
Chapter II

A STEP BY STEP SOLUTION OF
THREE PROBLEMS IN NON-NUMERICAL ANALYSIS

The mathematical laboratory consists of a PDP-6 computer linked by
dataphone with the Project MAC time sharing system. When the laboratory
is in operation, versions of the LISP programming language are running
in both the PDP-6 computer and the Project MAC 7094 computer. The user
communicates with the PDP-6, which is used only for input-output, by
means of a teletype, a DEC Type 30 display, a light pen, and a Calcomp
plotter. The PDP-6 relays messages to and from the Project MAC time
sharing system, where transformation routines written in the LISP
language are applied to the mathematical expressions. Mathematical
expressions can be displayed and plotted in standard textbook format.
Input of expressions and commands is from the teletype, but subexpressions
of displayed expressions can be referenced during input. If the light
pen is pointed at the main connective of a subexpression, the
subexpression can now be referenced in the typed input string. The
exact format of the input commands and a brief description of the
available mathematical transformations will be presented in the next
section. This description is followed by the step-by-step solution of
three problems in applied mathematics. The details of the programs will
be presented in later chapters.
The Mathematical Operators

The commands typed at the PDP-6 are similar to Algol statements. The commands are typed and executed one at a time. More complex operations involving the definition and alteration of commands and the introduction of more pneumonics and man machine interaction will be described later. The commands consist of infix operators, functions, and variables. Functions and variables can be subscripted and any subexpression can be quoted. A sample command which requires most of the notation is:

\[ \delta'\text{El}'(X + Y) \times \text{'DRV('}T,1,\text{'DRV('}U,2,\text{El}')) + E2,20 + 'f(I,J)(X,Y))' \]

In words: the name El is assigned to the expression which is the sum of three terms. The first term is the product of \((X + Y)\) with the unevaluated first derivative with respect to lower case \(T\) of the second derivative with respect to \(U\) of the expression currently named El. The second term is the \(20^{th}\) subexpression of a displayed expression currently named E2; this subexpression has been indicated with the light pen. The final term is the square of a subscripted function of \(X\) and \(Y\). The notation may seem somewhat complex, but as will be seen, a complex notation is required to express in a compact way the many small steps required to solve a particular problem.
The infix operators are:

\[ A*B \]  \( B \) is given name \( A \). As such it is written on the disk. The value of \( * \) is \( B \).

\[ \lceil A,N \rceil \]  The \( N \)th subexpression of \( A \) is the value of \( \lceil \) Intensify the desired subexpression of \( A \) by pointing to its main connective with the light pen. Then type \( \lceil A \rceil \), and the computer will type \( N \). If the expression has no main connective, point to one of its arguments and type \( \lceil A \rceil \) instead of \( \lceil A \rceil \). Consider all minus signs to be unary.

\[ A=B \]  Equate \( A \) and \( B \)

\[ A+B \]  \( A \) plus \( B \)

\[ A-B \]  \( A \) minus \( B \)

\[ A*B \]  \( A \) times \( B \)

\[ A/B \]  \( A \) divided by \( B \) \( A/B^C \) is equivalent to \( A/(B^C) \)

\[ A*B \]  \( A \) to the power \( B \)

The functional, subscripting, and set notation is:

\[ A(C,D,E) \]  \( A \) is a function with arguments \( C,D, \) and \( E \).

\[ A[I,J] \ (C,D,E) \]  \( A_{I,J} \) is a function with arguments \( C,D, \) and \( E \).

\[ A[I,J] \]  \( A_{I,J} \) is a variable.

\[ (A,B,C) \]  This is a set with three elements. By convention \( (A) = A \).
Either an expression or a variable name can be quoted. A function name always stands for itself. Quoting a function name means that its arguments will be evaluated but that the function will not be evaluated. For example let \( F(X,Y) = X - Y \) be a function and let \( X \) and \( Y \) be names for \( A \); then

\[
F(X,Y) \text{ evaluates to } 0 \\
'F(X,Y) \text{ evaluates to } F(A,A) \\
F('X,Y) \text{ evaluates to } X - A \\
'(F(X,Y)) \text{ evaluates to } F(X,Y).
\]

Quoting a function or variable name does not quote its subscripts. Numbers are taken as quoted automatically.

Causes the letters which follow it to be lower case for purposes of display.

As in CTSS, there are two editing characters:

? Deletes all the characters of a command back to the initial ?.

" Deletes only the immediately preceding character.

"?" must be the first and last character of every command. ";" causes the current intensified subexpression to be raised one level. For example, if the \( A \) in \( A^B + C \) is intensified, then when ; is typed, \( A^B \) will be intensified.
Other available operators are:

**ALLSUMEXPAND**(EXP)  Applies SUMEXPAND to every summation in expression EXP.

**BRINGOVER**(EXP, X)  Subexpression X, which has been indicated with the light pen is brought to the other side of equation EXP.

**COLLECT**(EXP, SET)  Top level terms in EXP are collected on powers of the expressions in set SET.

**DEPENDENCE**(EXP)  Returns a set of the variable and function names in EXP.

**DELSUBST**(EXP, OLDDEL, NEWDEL)  \[ \frac{dx}{d \text{OLDDEL}} \rightarrow \frac{dx}{d \text{NEWDEL}} \] for each such subexpression in EXP.

**DRV**(X_{i}, N_{i}, ..., X_{n}, N_{n}; Y)  Differentiate Y N_{i} times with respect to X_{i}, for each i.

**DRVDO**(EXP, X)  All indicated derivatives with respect to X in EXP are carried out as far as possible.

**DREVFACTOR**(EXP, X, N)  \[ \frac{d^{N+M}}{dx^{N+M}} \rightarrow \frac{d^{N}}{dx^{N}} \left( \frac{d^{M}}{dx^{M}} \right) \] for each such subexpression in EXP.

**DRVZERO**(EXP, X)  All derivatives with respect to X in EXP are set equal to zero.
EVALUATE(\text{EXP},\text{SET})

\text{SET} is a set of equations; whenever the left side of one of these equations can be matched to a subexpression in \text{EXP}, the right hand side is substituted. The left sides must be variables or functions. A match occurs whenever a binding of the function variables and subscripts can be made.

EXCHANGE(\text{EXP})

If the top level connective of \text{EXP} is binary, its arguments are exchanged, right to left.

EXPAND(\text{EXP})

Multiplies out all expressions of the form \(a^b(b+c)\) in \text{EXP}. In addition,

\[
\frac{d}{dx}(a^b) \rightarrow \frac{da}{dx} + \frac{db}{dx}
\]

FACTOROUT(\text{EXP},\text{FACTOR},Y)

The factor \text{FACTOR} is factored from each term of \text{EXP}. The third argument \(Y\) is optional. If \(Y\) is present, the factor \text{FACTOR} is renamed \(Y\).

GROUP(\text{SET})

The set \text{SET} of terms which have been indicated by the light pen in \text{EXP} are grouped within the associated sum or product. The value of \text{GROUP} is the grouped set of terms.

LEFT(\text{EXP})

Returns the left argument of the main binary connective of \text{EXP}.

LIMIT(\text{EXP},X,N)

Determines the limiting value of \text{EXP} as \(X\) approaches \(N\).

MULTIPLYTHROUGH(\text{EXP},X)

Multiplies each top level term of \text{EXP} by \(X\).

NEWNAME()

Creates a name of the form \text{Fn}, where \(n\) is an integer.
NORMPOLY(EXP,X) Every sum in EXP is treated as a polynomial in X and a power of X is factored out so that the lowest power of X in the polynomial will be zero.

REPLACE(E,X,Y) Expression X replaces Y in the expression named E. Y is a term indicated with the light pen or a group of terms indicated with GROUP. If the light pen has been used to construct X, the resulting expression position is named HOLE. HOLE can then be used for the third argument. If X is equal to NIL, then the third argument is omitted from the expression named E.

RIGHT(EXP) Returns the right argument of the main binary connective of EXP.

SIMPENIFY(EXP) Simplifies expression EXP.

SOLVE(EXP,X) Solves equation EXP for variable X as far as possible.

SPLIT(EXP) Subparts of EXP are named and replaced by their names in EXP, so that EXP will contain less than 100 subexpressions.

SUBSTITUTE(EXP,X,Y) Substitute X for each occurrence of Y in EXP.

SUM(σ,EXP) \( \sum (a+b) \rightarrow \sum a + \sum b \)

SUMEXPAND(σ,EXP) Expands the finite summation EXP.

TERM(σ,EXP,N) Returns the Nth argument of the top level connective of EXP, or NIL if there is no Nth argument.

TRUNCATE(σ,EXP,VAR,N) Expands EXP up to power N in variable VAR.

SUM(I,N1,N2,Y) Sum expression Y for values of I from N1 to N2.
ITG(X,L1,L2,Y) Integrate Y with respect to X between limits L1 and L2.

Expressions which are assigned names are kept on the disk. The expression most recently computed always has the name LAST. When A*B is executed, if A is not "LAST" and is already the name of an expression, then this old value of A is given the name OLD. Thus, if A*A+2 is executed and then is found to be incorrect, the old value of A can be retrieved.

Operators used for input-output and disk storage are:

EDISPLAY(E) Displays the expression named E on the PDF-6 scope.

EPRINT(E) Prints out the internal form of the expression named E with PLS, PRD, EQN, and FMV in infix form; the other operators in prefix form.

EDELETE(E) Deletes expression named E from the disk.

This completes the description of the PDF-6 commands.

The Poincaré-Lighthill Procedure Applied to $\frac{\partial^2 x}{\partial t^2} + \omega^2 x = c k^3$

The Poincaré-Lighthill procedure is typical of a number of procedures used to find the first few terms in the asymptotic expansion of the function which is the solution to a mildly non-linear differential equation. The equation chosen here is that for a harmonic oscillator with a small forcing function. These solution procedures involve assuming a series expansion in powers of the small parameter $\varepsilon$ for one or more of the parameters and variables in the equation, substituting these series into the differential equation, and thus obtaining a series of relations between the coefficients of like powers of $\varepsilon$. Each of these equations
is then treated in turn by whatever methods seem appropriate. Thus it is in general necessary to see these equations before the next steps in the solution process can be determined.

When a typed command has been completed, the machine makes a response of acknowledgment. This standard response will be omitted in the dialogue to follow; only the typed commands and the displayed equations will be shown. A running discussion of the dialogue is included, and the displayed equations are shown as plotted by the CALCOMP plotter on the pages following their use. A photograph of the same equations displayed on the scope is shown at the end of the section. The reader should be aware that in the equation syntax used, more than one line of an expression can occur over a divide bar or within brackets. This is illustrated by equation Q16 in the last section.

Enter the differential equation.

#'E1<-'(BRV(3,2,X(3:T)) + OMEGA*2*X(3:T) = EP*X(3:T)+J)#

#EDISPLAY('E1)#

A new independent variable \( \tau \) is introduced in order to stretch the time. Type in expressions for series expansions for \( X \) and \( t \) in terms of functions of \( \tau \). That is, a solution for \( X(\tau) \) rather than for \( X(t) \) will be found. Since \( t \) depends on \( \tau \), equation E1 can be used to find an equation in derivatives of \( X(\tau) \). As a final step the inverse relationship \( \tau(t) \) will be found, so that \( X(\tau) \) will give \( X(t) \).

#'E2<-'(X(TAU) = SUM(1,0,INF,EP*1*X[1](TAU)))#

#EDISPLAY('E2)#

#'E3<-'(I(TAU) = TAU + SUM(1,INF,EP*J*:I(J)(TAU)))#

#EDISPLAY('E3)#
In order to substitute $\tau$ for $t$ it is necessary to apply the transformation
\[
\frac{d^2 x}{d t^2} \rightarrow \frac{d}{d \tau} \left( \frac{d x}{d \tau} \right)
\]
to equation E1.

Display E4 for comparison with the substituted result below.

Now substitute $\tau$ for $t$, and $X(\tau)$ for $X(t)$

Now substitute the series for $X(\tau)$ and perform the indicated differentiation with respect to $\tau$.

Now expand both sides to first order in $\varepsilon$.
\[ (E1) \quad \frac{d^2}{dt^2} \chi(t) + \omega^2 \cdot \chi(t) = \varepsilon \cdot \chi(t)^3 \]

\[ (E2) \quad \chi(\tau) = \sum_{I : \theta}^\alpha \varepsilon^I \cdot \chi_1(\tau) \]

\[ (E3) \quad \tau(\tau) = \tau + \sum_{j=1}^\alpha \varepsilon^j \cdot \tau_j(\tau) \]
The zero order terms form the harmonic oscillator equation; the solution can be written down by inspection as $A \cos \omega t$. Use the light pen to form an equation of the first order terms.

Bring the terms in $X_1(r)$ to the left side of the first order equation. Substitute for $X_0(r)$ and carry out the indicated differentiation.

It is now necessary to substitute an identity for $\cos^3 \omega t$ and to collect terms on $\sin \omega t$ and $\cos \omega t$.

Theoretical considerations require that the coefficients of $\cos \omega t$ and $\sin \omega t$ must be zero if there is to be a periodic solution for $X_1(r)$. From the coefficient of $\sin \omega t$ it is apparent that $t_1(r)$ must be some constant $C$. This constant is determined from the coefficient of $\cos \omega t$.
(E4) \[
\frac{a}{dt} \left[ \frac{a}{dt} X(t) \right] + \omega^2 \cdot X(t) = \varepsilon \cdot X(t)^3
\]

(E5) \[
\frac{\left\{ \frac{a}{d\tau} \left[ \frac{a}{d\tau} X(\tau) \right] \right\}}{\left\{ \sum_{i=1}^{n} \frac{a}{d\tau} X_i(\tau) \cdot \varepsilon_i + 1 \right\}} + \omega^2 \cdot X(\tau) = \varepsilon \cdot X(\tau)^3
\]

(E6) \[
\frac{\left\{ \sum_{i=1}^{n} \frac{a}{d\tau} X_i(\tau) \cdot \varepsilon_i \right\} \cdot \left\{ \sum_{i=1}^{n} \frac{a}{d\tau} X_i(\tau) \cdot \varepsilon_i \cdot \varepsilon_i + 1 \right\} \cdot \left\{ \sum_{i=1}^{n} \frac{a}{d\tau} X_i(\tau) \cdot \varepsilon_i \right\} \cdot \left\{ \sum_{i=1}^{n} \frac{a}{d\tau} X_i(\tau) \cdot \varepsilon_i + 1 \right\} \cdot \left\{ \sum_{i=1}^{n} \frac{a}{d\tau} X_i(\tau) \cdot \varepsilon_i \cdot \varepsilon_i + 1 \right\}}{\left\{ \sum_{i=1}^{n} \frac{a}{d\tau} X_i(\tau) \cdot \varepsilon_i + 1 \right\} \cdot \left\{ \sum_{i=1}^{n} \frac{a}{d\tau} X_i(\tau) \cdot \varepsilon_i \right\} \cdot \left\{ \sum_{i=1}^{n} \frac{a}{d\tau} X_i(\tau) \cdot \varepsilon_i + 1 \right\}} + \omega^2 \cdot \sum_{i=1}^{n} \varepsilon_i \cdot X_i(\tau) = \varepsilon \cdot \left[ \sum_{i=1}^{n} \varepsilon_i \cdot X_i(\tau) \right]^3
\[ (-2) \left\{ \frac{a^2_{\tau}}{a_{\tau}} x_{\tau}(r) \left[ \frac{a}{a_{\tau}} t_1(r) + \frac{a^2}{a_{\tau}^2} x_1(r) + (-1) \left[ \frac{a^2}{a_{\tau}^2} t_1(r) \right] \left[ \frac{a}{a_{\tau}} x_\tau(r) + x_1(r) \cdot \omega^2 \right] \right] \right\} + \frac{a^2_{\tau}}{a_{\tau}} x_\tau(r) = x_{\tau}(r)^3 \]
\[ (E9) \]
\[
\omega^2 \cdot \chi_1(\tau) + \frac{\partial^2}{\partial \tau^2} X_1(\tau) = (-2) \left[ \frac{\partial}{\partial \tau} t_1(\tau) \right] \cdot \cos(\omega \cdot \tau) \cdot \omega^2 \cdot A + (-1) \cdot \sin(\omega \cdot \tau) \cdot \omega \cdot A \cdot \left[ \frac{\partial^2}{\partial \tau^2} t_1(\tau) \right] + A^2 \cdot \cos(\omega \cdot \tau)^2
\]

\[ (E10) \]
\[
\frac{\partial^2}{\partial \tau^2} X_1(\tau) + X_1(\tau) \cdot \omega^2 = (-1) \cdot \omega \cdot A \cdot \left[ \frac{\partial^2}{\partial \tau^2} t_1(\tau) \right] \cdot \sin(\omega \cdot \tau) + \left[ (-2) \left[ \frac{\partial}{\partial \tau} t_1(\tau) \right] \cdot \omega^2 \cdot A + \frac{1}{4} A^3 \right] \cdot \cos(\omega \cdot \tau) + \frac{1}{4} \cdot \cos(3 \cdot \omega \cdot \tau) \cdot A^2
\]

\[ (E11) \]
\[
C = \frac{\frac{5}{4} \cdot A^2}{\omega^2}
\]
So \(X_0 = A \cos \tau\) and to first order \(t = \tau + \frac{3A^2}{8\omega^2} \tau\).

Thus to first order \(\tau = t(1 - \frac{3A^2}{8\omega^2})\) and \(X_0 \approx A \cos (1 - \frac{3A^2}{8\omega^2})t\).

One effect of the non-linear term is thus seen to be an alteration in the frequency of the zero order term.

In conclusion, note the large number of small steps necessary to solve this problem. These are the result of doing almost the entire solution in the machine. In the case of a small problem such as this, some of these steps could be done by hand, the object here is to illustrate the steps which would be required for a larger problem. This problem also illustrates how rather lengthy intermediate calculations can lead to some rather concise results. The perturbation of the frequency in the zero order function is found to have a simple expression.

**Plasma Accelerator Electrode Boundary Layers**

The second problem is a duplication of the work in the first three sections of Chapter Three of a 1963 Masters Thesis by J. S. Draper for the M.I.T. Department of Aeronautics and Astronautics. This thesis investigates the laminar compressible boundary layer on the electrode walls of a direct-current crossed field plasma accelerator under very special physical conditions. Many of the assumptions used under these conditions are set forth in the paper "Electrode Boundary Layers in

In summary, the entire solution procedure is as follows:

1. Write down 5 non-linear partial differential equations:
   - Momentum
   - State
   - Continuity
   - Energy
   - Electron mobility as a function of temperature.

These equations relate
   - $U$ stream velocity
   - $V$ lateral velocity
   - $t$ temperature
   - $\rho$ density
   - $\mu$ electron mobility
   - $P$ pressure

in terms of the independent variables $x$ and $y$. The constants are:
   - $j$ current
   - $B$ magnetic field
   - $C_p$ specific heat
   - $K$ compressibility
   - $\sigma$ conductivity
   - $R$ gas constant
2. The absence of variation in the $y$ direction in the free stream is used to find the momentum and state equations there. These two reduced equations are solved for $\frac{dP}{d\xi}$ which is eliminated from the five main equations, since $P$ is not a function of $y$.

3. The relation $H = C_p T + \frac{U^2}{2}$ is used to substitute derivatives of $H$ for those of $T$ in the energy equation which then becomes an enthalpy equation. $T$ is thus eliminated from this equation. Simplification of the resulting expression requires introduction of the momentum relation. This step is performed because the enthalpy equation has a term proportional to $[1 - \frac{1}{P_r}]$ where $P_r$ is $\frac{C_p U^2}{K}$ and can be approximated as 1, thus eliminating this term.

4. Next, a change of independent and dependent variables is made. The change of independent variables is such that it approximates a similarity transformation for low Mach number. These transformations change $x$ and $y$ to $\Theta$ and $\eta$. In addition $\frac{U}{U_m}$ is defined as $f'$, $\frac{C}{C_m}$ as $\Theta$, and $\frac{H}{H_m}$ as $g$. The momentum and enthalpy equations are transformed, using the continuity and state equations as side conditions. $\Theta$ is then changed to $M$.

There result two non-linear differential equations in $f$ and $g$ and their derivatives with respect to $\eta$ and $M$.

5. $f$ and $g$ are then approximated as $f'(\eta, M) = b(M)\eta + c(M)\eta^2$, $g(\eta, M) = e(M)\eta + f(M)\eta^2$. These approximations are substituted into the two non-linear differential equations. The equations are then integrated with respect to $\eta$ between the wall and the edge of the velocity boundary layer $\delta_u$ and the edge of the enthalpy boundary.
layer δe respectively. b(M_o) and c(M_o) are eliminated from the result by the relations $f'(\delta u, M_o) = 1$ and $g(\delta e, M_o) = 1$. There result two ordinary linear differential equations for the derivatives of $\delta u$, $c$, $\delta e$, and $f$ with respect to $M_o$.

6. Two more linear differential equations for $\delta u$, $c$, $\delta e$, and $f$ are generated by choosing the coefficients of the approximations in step 5 so as to satisfy the momentum and enthalpy equations produced in step 4 exactly at the extremal points $f''' = 0$ and $g' = 0$.

7. The four resulting linear differential equations are solved for the derivatives $\frac{\delta u}{\delta M_\infty}$, $\frac{\delta c}{\delta M_\infty}$, $\frac{\delta e}{\delta M_\infty}$, and $\frac{\delta f}{\delta M_\infty}$ by Gaussian reduction. These four expressions are then numerically integrated with a Runge-Kutta method.

This problem has several interesting features. It is a demonstration of the notation use by workers in this area. The algebraic expressions are of a size difficult to manipulate by hand, but within the capabilities of current machines. The final symbolic result is large; it is difficult to write the corresponding numerical integration program correctly when this result must be input by hand, but here it is developed in the machine and could then be transformed into the required numerical program. Note that the symbolic steps are needed in order to cast the problem in terms of the independent and dependent variables of interest. The problem is characterized by the application of simplifying side conditions and physical assumptions. As such, it involves a number of manipulations for the purpose of expression condensation. This will be apparent from the following step by step reproduction of the first three sections of Chapter III. These steps bring the described solution through the
application of the similarity transformation to the momentum equation.

Input the momentum equation:
\[ \frac{\partial}{\partial x} (\rho u^2) + \frac{\partial}{\partial y} (\rho v u) = \rho \frac{\partial u}{\partial x} + \rho \frac{\partial u}{\partial x} + \rho \frac{\partial u}{\partial y} \]

Input the energy equation:
\[ \frac{\partial}{\partial x} (\rho c_p) + \frac{\partial}{\partial y} (\rho v c_p) = \rho \frac{\partial u}{\partial x} + \rho \frac{\partial u}{\partial x} + \rho \frac{\partial u}{\partial y} + \rho \frac{\partial u}{\partial y} \]

The boundary layer solutions must match the free stream solution.
The free stream values are indicated by the subscript \( \infty \). At the free
stream, there is no variation in the boundary layer with respect to \( y \).
To save rewriting, define sets containing the variables to be subscripted.

Then in the free stream \( D_1 \) and \( D_2 \) become:

\[ D_5 \text{ SIMPLIFY} \left( \text{SUBSTITUTE} \left( \text{DRVZERO} \left( D_1, 'Y', D_4, D_3 \right) \right) \right) \]

\[ D_6 \text{ SIMPLIFY} \left( \text{SUBSTITUTE} \left( \text{DRVZERO} \left( D_2, 'Y', D_4, D_3 \right) \right) \right) \]

\#EDISPLAY('D5)\#
\#EDISPLAY('D6)\#

\( D_5 \) and \( D_6 \) will now be used to express some of the expressions in \( D_1 \) and
\( D_2 \) in terms of the free stream quantities. In order to substitute for
some of the terms in a sum, the terms must be grouped using the light
pen. This is somewhat inconvenient.

\#EDISPLAY(D1)\#
\[(D5)\] \[\rho_0 \cdot \left( \frac{\partial}{\partial x} u_0 \right) \cdot u_0 = -\frac{\partial p}{\partial x} + j \cdot B\]

\[(D6)\] \[\rho_0 \cdot C_p \cdot \left( \frac{\partial}{\partial x} T_0 \right) \cdot u_0 = u_0 \cdot \left( \frac{\partial p}{\partial x} \right) + \frac{j}{\sigma_0} \frac{\partial^2}{\partial x^2}\]

\[(D1)\] \[\rho \cdot \left[ u \cdot \left( \frac{\partial u}{\partial x} \right) + v \cdot \left( \frac{\partial u}{\partial y} \right) \right] = \frac{\partial}{\partial y} \left( \mu \cdot \left( \frac{\partial u}{\partial y} \right) \right) - \frac{\partial p}{\partial x} + j \cdot B\]
Now the last two terms in D1 are replaced by the left side of D5.

\[ D7\rightarrow\text{REPLACE ('D1,LEFT(D5), GROUP((&D1,35,:D1,38)))} \]

\[ \#\text{EDISPLAY('D7')} \]

\[ D8\rightarrow\text{LEFT(D2) = EXPAND(SUBSTITUTE(}

\[ \text{RIGHT(D2), RIGHT(SOLVE(D6,'(X,1,F'))), '(DRV(X,1,F'))} \]

\[ \#\text{EDISPLAY('D8')} \]

The last two terms in D8 are now put in a factored form.

\[ D8\rightarrow\text{REPLACE('D8,FACOTOROUT(GROUP((:D8,62,:D8,78)),':(3*2/SIGMA[INF]))}

\[ \text{HOLE})} \]

\[ \#\text{EDISPLAY('D8')} \]

In order to effect a cancellation, equation D8 will now be transformed by replacing the temperature, t, with the enthalpy, H, using the definition

\[ H = C_p t + U^2/2. \]

First \( H = C_p t + U^2/2 \) is solved for temperature \( t \).

\[ D9\rightarrow\text{SOLVE('(X,Y) = ((P)*T(X,Y) + U(X,Y)*2/2),'(T(X,Y))} \]

Now the substitution is made for \( t \) and \( t_m \).

\[ D10\rightarrow\text{EXPAND(SUBSTITUTE(D8,(DRV('X,1,RIGHT(D9)),SUBSTITUTE(DRV('X,1,RIGHT(D9)),

\[ D4,D5),DRV('Y,1,RIGHT(D9)), '(DRV(X,1,T), DRV(X,1,T[INF]),

\[ DRV(Y,1,T)))})} \]
\[ \begin{align*}
\text{(D7)} & \quad \rho \left( u \left( \frac{\partial u}{\partial x} \right) + u \left( \frac{\partial u}{\partial y} \right) \right) = \frac{\partial}{\partial y} \left( \frac{\partial u}{\partial y} \right) + \rho u \left( \frac{\partial u}{\partial y} \right) + \frac{\partial}{\partial x} \left( \frac{\partial u}{\partial x} \right) u_n \\
\text{(D8)} & \quad \rho \cdot c_v \left( u \left( \frac{\partial u}{\partial x} \right) + u \left( \frac{\partial u}{\partial y} \right) \right) = \frac{a}{\partial y} \left( \frac{\partial u}{\partial y} \right) \cdot k + \left( \frac{\partial u}{\partial y} \right) \cdot \mu + \rho \cdot c_v \cdot \left( \frac{\partial u}{\partial y} \right) \cdot a_x \cdot c_v \cdot \left( \frac{\partial u}{\partial y} \right) = \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial y^2} + \frac{j^2}{\sigma} \left( - \frac{u_{\infty} \cdot c_v}{\sigma} \right)
\end{align*} \]
Now, there are a number of tedious grouping steps.

First, factor RHO out of two of the terms on the left side of the equation.

Now factor RHO·U from the other two terms and bring them to the right side.

A machine matching operation would be better for the next factoring step which must be done twice when light-pen pointing is used.

The quantity $K/C_p\mu$ is factored out of two of the terms and set equal to $1/P_e$.

$C_p\mu/K$ is the Prandtl number $P_e$. It is close to 1, and setting it to 1 will effect a simplification if the identity

$$\mu \frac{d u}{d t} = \frac{d}{d t} [\mu \cdot U(X,Y) \frac{d U(X,Y)}{d Y} - U \cdot \frac{d [\mu d u]}{d Y}]$$

is also substituted. This simplification means that the heat conduction away from a point is just equal to the viscous dissipation at that point.
\[ \{ \Omega \} - U(x, y) \left[ \frac{\partial}{\partial y} U(x, y) \right] - U \cdot \text{RHO} + \left[ \frac{\partial}{\partial y} H(x, y) \right] \cdot U \cdot \text{RHO} - U(x, y) \left[ \frac{\partial}{\partial x} U(x, y) \right] - U \cdot \text{RHO} + \left[ \frac{\partial}{\partial x} H(x, y) \right] \cdot U \cdot \text{RHO} \]

\[ = -\frac{a}{a y} \frac{k \cdot U(x, y) \left[ \frac{\partial}{\partial y} U(x, y) \right]}{C_r} + \frac{a}{a y} \frac{k \left[ \frac{\partial}{\partial y} H(x, y) \right]}{C_r} + \mu \left[ \frac{\partial U}{\partial y} \right]^2 + U \left[ \frac{\partial}{\partial x} H(x, y) \right] \cdot \text{RHO} \]

\[ - U \cdot U(x, y) \left[ \frac{\partial}{\partial x} U(x, y) \right] \cdot \text{RHO} + \frac{1}{\sigma^2} - \frac{U \cdot j^2}{U_j \cdot \sigma} \]
\[ (011) \quad u(x,y) \left[ \frac{\partial}{\partial y} u(x,y) \right] \cdot \text{RHO} + \text{RHO} \cdot \left\{ \left[ \frac{\partial}{\partial x} H(x,y) \right] \cdot u + \left[ \frac{\partial}{\partial y} H(x,y) \right] \cdot u \right\} - u(x,y) \left[ \frac{\partial}{\partial x} u(x,y) \right] \cdot \text{RHO} \]

\[ = - \frac{\kappa \cdot u(x,y)}{c_p} \cdot \frac{\partial}{\partial y} u(x,y) + \frac{\kappa \cdot \left[ \frac{\partial}{\partial y} H(x,y) \right]}{c_p} + MU \cdot \left( \frac{\partial u}{\partial y} \right)^2 + \left[ \frac{\partial}{\partial x} H_{xx}(x,y) \right] \cdot \text{RHO}_{xx} \]

\[ - u \cdot u_{xx}(x,y) \left[ \frac{\partial}{\partial x} u_{xx}(x,y) \right] \cdot \text{RHO}_{xx} + \frac{L}{\sigma} - \frac{U \cdot u_{xx}^2}{U_{xx} \cdot \sigma_{xx}} \]
\[ \text{RHO} \left\{ \left[ \frac{\partial}{\partial X} H(X, Y) \right] \cdot U + \left[ \frac{\partial}{\partial Y} H(X, Y) \right] \cdot U \right\} \]

\[ = -\frac{a}{\partial Y} K \cdot U(X, Y) \left[ \frac{\partial}{\partial Y} U(X, Y) \right] \left( \frac{a}{\partial Y} \right) \frac{C_r}{C_p} + \frac{a}{\partial Y} K \cdot \left[ \frac{\partial}{\partial Y} H(X, Y) \right] \left( \frac{a}{\partial Y} \right) \frac{C_r}{C_p} + NU \left[ \frac{\partial}{\partial Y} U(X, Y) \right]^2 + U \left[ \frac{\partial}{\partial X} H^*(X, Y) \right] \cdot \text{RHO} \]

\[ - U \cdot U(X, Y) \left[ \frac{\partial}{\partial X} U^*(X, Y) \right] \cdot \text{RHO} \cdot U + \frac{U \cdot \sqrt{z}}{U^* \cdot \sqrt{z}} \cdot \text{RHO} \cdot U(X, Y) \left\{ - \left[ \frac{\partial}{\partial Y} U(X, Y) \right] \cdot U - \left[ \frac{\partial}{\partial X} U(X, Y) \right] \cdot U \right\} \]
\[ \text{011} \]

\[ \text{RHO} \cdot \left( \left[ \frac{a}{a x} H(x, y) \right] \cdot U + \left[ \frac{a}{a y} H(x, y) \right] \cdot U \right) \]

\[ \begin{align*}
= - \frac{a}{a y} \frac{U(x, y) \cdot \left[ \frac{a}{a y} U(x, y) \right]}{p_e} + \frac{a}{a y} \frac{\kappa \cdot \left[ \frac{a}{a y} H(x, y) \right]}{c_e} + \mu \cdot \left( \frac{a y}{c_e} \right)^2 + U \cdot \left[ \frac{a}{a y} H_e(x, y) \right] \cdot \text{RHO} \cdot U \\
- U \cdot U_e(x, y) \cdot \left[ \frac{a}{a x} U_e(x, y) \right] \cdot \text{RHO} + \frac{1}{\sigma} \cdot \frac{U \cdot \left[ \frac{a}{a x} U(x, y) \right]}{U_e \cdot \sigma_e} - \text{RHO} \cdot U(x, y) \cdot \left\{ - \left[ \frac{a}{a y} U(x, y) \right] \cdot U - \left[ \frac{a}{a x} U(x, y) \right] \cdot U \right\}
\end{align*} \]
Next, the substitution in D11 of the right side of equation D7 for its left side effects a nice cancellation.

\[ \phi \text{D11-SIMPLIFY(REPLACE('D11,-RIGHT(D7)),GROUP((ID11,117,ID11,122))))} \]

\[ \phi \text{D11-LEFT(D11) = SIMPLIFY(EXPAND(SUBSTITUTE(RIGHT(D11),'(U, U[INF]),'(U,(X,Y),}
\text{U[INF])(X,Y)))))} \]

\[ \phi \text{EDISPLAY('D11)} \]

The momentum equation D7 and the enthalpy equation D11 are now ready for the similarity transformation.

The transformations of independent variables are:

\[ \phi \text{'D12-'(XB(X) = ITG(X,0,X,P=U[INF]/(P[0]^2U[0])))} \]

\[ \phi \text{'D13-'(ETA(X,Y) = (U[INF]/U[0])*(U[0]/(2*nu[0]*XB(X))^((1/2)))*}
\text{ITG(Y,0,Y,RHO/RHO[0]))} \]

Next, to compute the required differentials.

First

\[ \frac{1}{dX} = \frac{d\delta}{dx} + \frac{\frac{d\theta}{dx}}{\frac{d\theta}{dx}} \cdot \frac{1}{dX} \]

\[ \phi \text{'D14-DRV('X,1,RIGHT(D13))'-(DEL(ETA))+DRV('X,1,RIGHT(D12)+'DEL(XB))} \]

\[ \phi \text{EDISPLAY('D14)} \]

D14 now contains the expression for \( \eta \) as a factor.

\[ \phi \text{'D14-REPLACE('D14,FACTOROUT('D14,'RIGHT(D13),'(ETA(X,Y)),''HOLE)} \]

\[ \phi \text{EDISPLAY('D14)} \]

\( \eta \) is now substituted for its definition and \((P \cdot U_m)/(P_0 \cdot U_0)\) is factored out.

\[ \phi \text{'D14-SIMPLIFY(FACCTOROUT(SUBSTITUTE(D14,DRV('X,1,RIGHT(D12)),}
\text{''(DRV(X,1,XB(X))),'P=U[INF]/(P[0]^2U[0]))))} \]

\[ \phi \text{EDISPLAY('D14)} \]
\[011\] \[ \frac{\partial}{\partial t} \left[ \frac{\partial}{\partial x} H(x, y) \right] \cdot U + \frac{\partial}{\partial v} H(x, y) \cdot U \]

\[ = \frac{\partial}{\partial v} \left[ \frac{\partial}{\partial v} H(x, y) \right] \cdot MU - U(x, v) \cdot \left[ \frac{\partial}{\partial v} M U \cdot \left( \frac{\partial}{\partial v} U \right) U \cdot \left[ \frac{\partial}{\partial v} H(x, y) \right] \cdot RHO \cdot U \cdot U(x, y) \cdot \left[ \frac{\partial}{\partial v} H_0(x, y) \right] \cdot RHO \cdot U \]

\[ + \frac{1}{\sigma} \frac{U \cdot U}{U_0 \cdot U} - RHO \cdot U(x, y) \cdot \left[ \frac{\partial}{\partial v} U(x, y) \right] \cdot U \cdot \left[ \frac{\partial}{\partial v} U(x, y) \right] \cdot U \]

\[011\] \[ \frac{\partial}{\partial t} \left[ \frac{\partial}{\partial x} H(x, y) \right] \cdot U + \frac{\partial}{\partial v} H(x, y) \cdot U \]

\[ = \frac{\partial}{\partial v} \left[ \frac{\partial}{\partial v} H(x, y) \right] \cdot MU + \frac{\partial}{\partial v} H(x, y) \cdot U \cdot \left[ \frac{\partial}{\partial v} H(x, y) \right] \cdot RHO \cdot \left[ \frac{\partial}{\partial v} H_0(x, y) \right] \cdot U \cdot \frac{1}{\sigma} - \frac{1}{\sigma} \frac{U}{U_0 \cdot U} \]
\[
\begin{align*}
\text{[014]} & \quad \frac{\frac{\partial}{\partial x} xB(x) \cdot NU_x \cdot 2 \cdot U_x \cdot \left[ \frac{U_x}{2 \cdot NU_x \cdot XB(x)} \right]^{-\frac{1}{2}} \cdot U_x \cdot \left( \int \frac{RHO \, dY}{RHO \, \delta Y} \right) \cdot \text{DEL}(\text{ETA})}{\left[ 2 \cdot NU_x \cdot XB(x) \right]^2 \cdot 2 \cdot U_x} \\
\text{[014]} & \quad \frac{\frac{1}{2} \cdot \text{ETA}(X, Y) \cdot \left[ \frac{\partial}{\partial x} xB(x) \right] \cdot \text{DEL}(\text{ETA})}{XB(x)} + \frac{P \cdot U_x \cdot \text{DEL}(XB)}{P_x \cdot U_x}
\end{align*}
\]
The differential for \(1/dY\) is
\[
\frac{d}{dX} \text{ and } \frac{d}{dY}
\]
Expressions D14 and D15 are the required differentials for \(\frac{d}{dX}\) and \(\frac{d}{dY}\) respectively.

Now to transform the momentum equation D7. First, substitute for the differentials of the independent variables and the new normalized dependent variable \(f\), defined by \(U/U_{\infty} = \frac{f'}{U_{\infty}}\).

Use the equation of state to recognize that
\[
\frac{d}{dY} (\rho T) = \frac{dP}{dn} = 0
\]
since \(P\) is taken independent of \(Y\). A substitution for \(P\) should therefore be made.

The machine responds with "EXPRESSION TOO LARGE"

Therefore, the left side of equation D17 will be treated first, the substitution for \(P\) will be deferred.
\[ (D14) \quad \frac{P \cdot U_\infty \left[ \frac{(-\frac{1}{2}) \cdot \eta(x, y) \cdot \nabla \eta}{x B(x)} + \nabla(x B) \right]}{P_\infty \cdot U_\infty} \]

\[ (D15) \quad \frac{\rho \cdot \left[ \frac{U_\infty}{2 \cdot n_\infty \cdot x B(x)} \right]^{\frac{1}{2}} \cdot U_\infty \cdot \nabla \eta}{\rho_{\infty} \cdot U_\infty} \]
An expression for $V$ will now be developed and substituted into D18. It was shown in Chapter II of the thesis that the continuity equation yields

$$V = -\frac{\rho_0}{\rho} \frac{1}{dx} \left(2u_0^2 \rho_0 \frac{du}{dx} \frac{1}{\rho} \right) \frac{1}{2} \left(\frac{\rho}{\rho_0} \right)^{\frac{1}{2}} \left(\frac{\rho_0}{\rho} \right)^{\frac{5}{2}}$$

Enter this expression.

`'D19'=-(\rho/\rho)*DRV(X,1,((2*u[0]*u[0]*u[0])*XB)+(1/2)*F(XB,ETA)))`  

Now substitute the new independent variables:

`'D20'=SUBSTITUTE(DELSUBST(D19,'(DEL(X))','D14),'(U[INF])(XB,ETA),
(U[INF],XB(ETA),ETA(X,Y))')`  

Now, substituting this expression for $V$ into the partially transformed left side of the momentum equation, D18, the entire expression is differentiated as far as possible, expanded and simplified.

`'D21'=EXPAND(DRVD(DRVD(SUBSTITUTE(D18,D20,'V'),'ETA'),'XB'))`  

Now some shorthand will be introduced to make D21 easier to read.

`'D22'=((F,U,F1,F2,F01,F11,U[INF]))`  

`'D23'=((F(XB,ETA),U(XB),DRV(ETA,1,F),DRV(ETA,2,F),DRV(XB,1,F),DRV(ETA,1,XB,1,F),U[INF])(XB))`  

`'D24'=SUBSTITUTE(D21,D22,D23)`  

`EIDISPLAY('D21')`  

It is now convenient to factor a large coefficient from each term and set it equal to 1 since it will also be factored from the other side and set to 1 there.

`'D25'=SIMPACTUOUT(D21,'(RHO*U[INF]*3*F/(2*F(0)*U[0]*XB)),1))`  

`EIDISPLAY('D25')`
Returning to the right side of D17.

\[ D_{25} \cdot \text{RIGHT}(D17) \]

\[ \text{EDISPLAY('D25')} \]

The deferred substitution for \( P \) is now made.

\[ D_{27} \cdot \text{REPLACE('D25',FACTORYOUT(,'D25',44,'((RHO\ast T)/(RHO[0] \ast T[0])), '(P/P[0])), 'HOLE)'} \]

The expression is now differentiated as far as possible, as was the left side.

\[ D_{28} \cdot \text{EXPAND(DRVD0(DRVD0(D27,'ETA'),'X8'))} \]

The simplifying notation is substituted.

\[ D_{29} \cdot \text{SUBSTITUTE(D28,D22,D23)} \]

\[ \text{EDISPLAY('D28')} \]

Finally the large factor is removed from each term as it was from the right side.

\[ D_{29} \cdot \text{SIMPCLUY(FACTORYOUT(D28,'(RHO\ast U[1\ast INF]+3\ast P)/(2\ast P[0] \ast U[0] \ast X8)),1))} \]

The two sides of the transformed momentum equation are now recomumed.

\[ D_{30} \cdot \text{SIMPCLUY(D24-D29) = 0} \]

\[ \text{EDISPLAY(D'30)} \]

Using some more relations developed earlier in the thesis, a final substitution will be made, from the equation of state at constant pressure.
\[ (021) \quad - \frac{U_e^4 \cdot F_2 \cdot \rho_0 \cdot F_0 \cdot P}{U_e \cdot P_e} + \left( -\frac{1}{2} \right) \frac{U_e^3 \cdot F_2 \cdot \rho_0 \cdot f \cdot P}{X_B \cdot U_e \cdot P_e} + \frac{F_1^2 \left( \frac{\partial}{\partial X_B} U_e \right) \cdot P \cdot U_e^4 \cdot \rho_0 \cdot F_1 \cdot \rho_0}{U_e \cdot P_e} + \frac{U_e^5 \cdot F_{11} \cdot P \cdot F_1 \cdot \rho_0}{U_e \cdot P_e} \]

\[ (024) \quad -2 \cdot F_2 \cdot F_0 \cdot X_B \cdot F_2 \cdot f - \frac{2 \cdot F_1^2 \left( \frac{\partial}{\partial X_B} U_e \right) \cdot X_B}{U_e} + 2 \cdot F_{11} \cdot F_1 \cdot X_B \]
\[ \begin{align*}
\text{RHO} \cdot 2^{-\frac{1}{2}} \cdot U_e^{-\frac{1}{2}} \cdot \text{NU}_e^{-\frac{1}{2}} \cdot \text{XB}^{-\frac{1}{2}} \cdot U_e(\text{XB}) \\
\text{RHO}_e \\
\left\{ \frac{a}{\text{d} \text{ETA}} \frac{\text{M} \text{U}_e \cdot t \cdot \text{RHO} \cdot 2^{-\frac{1}{2}} \cdot U_e^{-\frac{1}{2}} \cdot \text{NU}_e^{-\frac{1}{2}} \cdot \text{XB}^{-\frac{1}{2}} \cdot U_e(\text{XB})}{\text{t}_e \cdot \text{RHO}_e} \left[ \frac{\text{d}}{\text{d} \text{ETA}} U_e(\text{XB}) \right] \left[ \frac{\text{d}}{\text{d} \text{ETA}} f(\text{XB}, \text{ETA}) \right] \right\} \\
\text{RHO}_e \cdot P \cdot U_e(\text{XB})^{-2} \cdot \left\{ \left( -\frac{1}{2} \right) \cdot \text{ETA} \cdot \left[ \frac{\text{d}}{\text{d} \text{ETA}} U_e(\text{XB}) \right] \frac{\text{XB}}{\text{XB}} + \frac{\text{d}}{\text{d} \text{XB}} U_e(\text{XB}) \right\} + \frac{\text{d}}{\text{d} \text{XB}} U_e(\text{XB}) \\
\text{P}_e \cdot U_e
\end{align*} \]
\[(D28) \quad \frac{1}{2} P \cdot MU \cdot U^3 \left( \frac{\partial F}{\partial ETA} \right) \cdot RHO + \frac{\left( \frac{\partial}{\partial XB} U \right) \cdot U \cdot P \cdot RHO}{U \cdot P} \]

\[(D30) -2 \cdot F_2 \cdot F_0 \cdot XB = 2 \cdot F_1 \cdot \frac{\left( \frac{\partial}{\partial XB} U \right) \cdot XB}{U} + 2 \cdot F_{11} \cdot F_1 \cdot XB - \frac{2 \cdot \left( \frac{\partial}{\partial XB} U \right) \cdot RHO \cdot XB}{U \cdot RHO} - \frac{MU \cdot \left( \frac{\partial F}{\partial ETA} \right)}{RHO \cdot NU} = 0 \]

\[(D31) \quad -2 \cdot F_2 \cdot F_0 \cdot XB = 2 \cdot F_1 \cdot \frac{\left( \frac{\partial}{\partial XB} U \right) \cdot XB}{U} + 2 \cdot F_{11} \cdot F_1 \cdot XB = 0 \]

\[2 \cdot \left[ \frac{\left( \frac{1}{2} \cdot (\text{GAMMA}-1) \cdot M^2 - B \cdot \gamma + 1 \right) \cdot C + B \cdot \gamma + \left( \frac{1}{2} \right) \cdot M^2 \cdot (\text{GAMMA}-1) \cdot F_1^2}{U} \right] \cdot \frac{\left( \frac{\partial}{\partial XB} U \right) \cdot XB}{U} - \frac{MU \cdot \left( \frac{\partial F}{\partial ETA} \right)}{RHO \cdot NU} = 0 \]
\[ \frac{\rho_\infty}{\rho} = \frac{1}{t_{\infty}} \text{ but } \frac{1}{t_{\infty}} = \Theta \text{ so } \frac{\rho_\infty}{\rho} = \Theta, \text{ and from Chapter II,} \]

\[ \Theta = (\Theta_s - \Theta_u)g + \Theta_u - (\Theta_s - 1)f^2 \]

\[ \Theta_s = 1 + \frac{Y-1}{2} \cdot \frac{N}{\infty} \]

The final transformed equation is:

\[ 'D31' \text{SIMPLIFY(SUBSTITUTE(REPLACE(}'D30',FACTOROUT('D30,38,'}'(RHO[INF]/RHO),


'(1 + (GAMMA - 1)*M[INF]²/2),'(THETA[S])'))' \]

# EDISPLAY('D31)'

In conclusion, since physical approximations are involved in the expression condensation, close man-machine interaction is required. The greatest drawbacks to sufficient interaction are the input notation and the lack of sufficient facilities for abbreviation on output.

A Multiple Channel Queueing Problem

The third problem is taken from Interim Technical Report No. 13 of the M.I.T. Center for Operations Research entitled "A Class of Queueing Problems". This was a 1955 Doctor's Thesis by H.N. Garber. The queueing situation shown by the example in Figure 1 is treated.
Arrival times at the queueing complex are exponentially distributed with mean $\lambda$. Arrivals enter any channel which becomes vacant and progress through several phases of service with exponentially distributed service times with mean $k\mu$. There are $M$ channels and $k$ phases of service in the general case. It is desired to find the probability distribution of the number of units in the system. The solution can be carried a number of steps for the general case. A set of equations relating the probabilities of the states are written, where a state of the system is taken as a certain number $N$ in the system and a description of which phases of which channels are occupied. Each equation is then multiplied by the appropriate variables so that the summation of the equations will yield an equation with the generating function for the state probabilities as its left side. Unfortunately, the generating function also appears in a summation on the right side. The independent variables of the generating function are constrained to make each term of this summation equal to zero. This constraint is expressed in a change of independent variables. It is then shown that a summation over values of the new independent variables will yield the old generating function.
This generating function still involves the state probabilities for a partially full system as constants to be determined. There are a number of relations which can be used to determine these initial probabilities. As the third problem this evaluation will be carried out for the case of two channels with two phases of service.

The source of expression complexity is different in this problem. In the first problem complexity arose because a small parameter allowed an unsolvable problem to be split into a spectrum of solvable ones. In the second problem complexity was the result of the number of important phenomena in the physical process being described. Here, complexity is the result of the number of states in the process being described, all of which are very similar. In this problem there would be the best chance for reduction in complexity through proper notation.

This problem has been included so that the reader can better refine his intuitive notion of the types of mathematical operations and notation needed to solve problems in different areas of applied mathematics.

Summation expansion, function evaluation, limits, and some more grouping operations are introduced.

The function in the transformed variables is:

\[ \hat{Q}^{(Z_0)}(X_0, Y_0, Z_0, K_0) = (Z_0(2Z_0 + 1))^A[K_0] + (E[Z_0 + 1]^Q[X_0 + 1]^R_0 / K_0 + E[Z_0 + 1]^Q[2]^R_0 / K_0]) / (2^Z_0(2^Z_0 - \alpha^X_0, Y_0, Z_0, K_0)) \]

Where \( \hat{W} \), \( \alpha \), and \( A_\tau \) are defined by:

\[ \hat{W}(Z_0; K) = 1 + \text{THETA} - \text{THETA}^2 \times K \]
\[ Q_3 = (\text{ALPHA}(Q[1],Q[2])) = (\text{E}^{(-2\pi i q[1])}/K) + (\text{E}^{(-2\pi i q[2])}/K)/2 \]

\[ Q_4 = (A[R] = P(1,R+1,0) - (2\pi i (Z^k - 1)) \times P(1,R,0)) \]

\[ \text{EDISPLAY}('Q1') \]

It is useful to have the denominator of \( Q_1 \) contain only powers of \( z^k \). This is done by using the identity \((x-y)(x^{k-1} \ldots y^{k-1}) = (x^k - y^k)\). At present the system contains no operators for achieving this goal, so it must be done by brute force.

\[ \text{Q5 = LEFT(Q1) = } (\text{SUM}(J,0,K-1,\text{Z}(Z^k)) \times J \times \text{ALPHA}(Q[1],Q[2],(K-1-J))) \]

\[ \text{EDISPLAY}('Q5') \]

Now substitute 2 for the number of phases of service, \( k \). Then expand the summations.

\[ \text{Q6 = SIMPLIFY(ALLLSUMEXPAND(SIMPLIFY(SUBSTITUTE(Q5,2,K))))} \]

\[ \text{EDISPLAY}('Q6') \]

The substitution \( k = 2 \) is made in all the initial equations as well since they will be used several times.

\[ \text{Q1 = SIMPLIFY(SUBSTITUTE(Q1,2,K))} \]
\[ \text{Q2 = SIMPLIFY(SUBSTITUTE(Q2,2,K))} \]
\[ \text{Q3 = SIMPLIFY(SUBSTITUTE(Q3,2,K))} \]
\[ \text{Q4 = SIMPLIFY(SUBSTITUTE(Q4,2,K))} \]

Now write down an expression for the generating function as a summation of \( Q_6 \) over the transformed independent variables, the \( q \)'s.
\[
K(z, 0, 1, 0_2) = \frac{Z^{2k+1}}{\sum_{R=1}^{k} Z \cdot A_R \cdot e_k} + e_k
\]

\[
\{01\}
\]

\[
\{02\}
\]
\[ K(Z,q_1,q_2) = \frac{\frac{1}{2} \cdot Z^{2\cdot q_1 + 1} \cdot \left( \sum_{k=0}^{q_2} \left( Z \cdot W \left( Z^k \right) \right)^2 \cdot \text{ALPHA} \left( q_1, q_2 \right)^{q_1+1} \right) \cdot \sum_{k=0}^{q_2} Z^k \cdot A_k \cdot \left( e^{-q_1 \cdot 1 \cdot \text{e}^k - q_2 \cdot 1 \cdot \text{e}^k} \right)}{Z^k \cdot W \left( Z^k \right) \cdot \text{ALPHA} \left( q_1, q_2 \right)^k} \]

\[ K(Z,q_1,q_2) = \frac{\frac{1}{2} \cdot \left( \text{ALPHA} \left( q_1, q_2 \right) + Z \cdot W \left( Z^q \right) \right) \cdot Z^q \cdot \left( Z \cdot A_1 \cdot \left( e^{-q_1 \cdot 1 \cdot \text{e}^q + q_2 \cdot 1 \cdot \text{e}^q} \right) + Z^q \cdot A_q \cdot \left( e^{-q_1 \cdot 1 \cdot \text{e}^q + q_2 \cdot 1 \cdot \text{e}^q} \right) \right)}{Z^q \cdot W \left( Z^q \right) \cdot \text{ALPHA} \left( q_1, q_2 \right)^q} \]
\[ q7^* \cdot (w(\bar{z} \bar{2}, v[1], v[2]) = (1/4) \cdot \text{SUM}(q[1], 0, 1, \text{SUM}(q[2], 0, 1, ((v[1]/2)^*(-1)*q[1] + v[1]\bar{z}^2/(v[2]^2))\cdot((v[2]/2)^*(-1)*q[2] + v[2]\bar{z}^2/(v[2]^2)) + k(2,q[1],q[2])))) \]

Now expanding the summations, substituting the appropriate values of \( q6, q3, \) and \( k, \) and simplifying.

\[ q8^* \cdot \text{Simplify}(\text{Evaluate}(\text{AllSumExpand}(q7), (q6, q3))) \]

To evaluate \( p(1,1,0) \) and \( p(1,2,0) \) certain known conditions are next imposed. \( q8 \) is known to have a zero of order four in \( Z \) for all values of \( v_1 \) and \( v_2, \) so one would like to collect terms on \( Z^4. \) The simplest way to explore this would be to expand \( q8 \) and collect terms on \( Z^4. \) Unfortunately, this leads to roughly a sixteen-fold growth in expression size and to memory overflow.

Inspection of \( q8 \) shows that it might be rearranged while in factored form. As the first step, subexpressions which are polynomials in \( Z, v_1, \) or \( v_2, \) have factors of \( Z, v_1, \) or \( v_2 \) removed so that their lowest order term is of zero order in these.

\[ q8^* \cdot \text{Simplify}(\text{Nnorm}(\text{Nnorm}(\text{Nnorm}(q8, 'z'), (v[1])), (v[2]))) \]

Next the center two terms are combined.

\[ q8^* \cdot \text{Replace}(q8, \text{FactorOut}(\text{Group}((q81,85, 'q81,108)), (a[2]/w(2z2)), '\text{HOLE}) \]

#EDISPLAY('q82')
\[ H(Z^2, U_1, U_2) = \frac{1}{4} \left( \sum_{\sigma_1} \frac{U_1}{Z} \cdot \left( \frac{1}{U_1} + U_1 \right) \right) \cdot \left( \frac{U_2}{Z} \cdot \left( \frac{1}{U_2} + U_2 \right) \right) \cdot \left( \frac{U_3}{Z} \cdot \left( \frac{1}{U_3} + U_3 \right) \right) \cdot \left( \frac{Z \cdot \omega(Z^2)}{Z^2} \cdot \left( \frac{1}{\omega(Z^2)} - 1 \right) \right) \cdot \left( Z \cdot \omega(Z^2) \cdot \left( \frac{1}{\omega(Z^2)} - 1 \right) \right) \]
The resulting term is then expanded.
\[ \texttt{\texttt{\texttt{'Q83*REPLACE('Q82, EXPAND('Q82,90), 'HOLE))}}} \]
\[ \texttt{\texttt{\texttt{'EDISPLAY('Q83))}}} \]

Now the other two terms are combined.
\[ \texttt{\texttt{\texttt{'Q84*REPLACE('Q83, FACTOROUT(GROUP((('Q83,111,'Q83,30)),1/Q83!70), 'HOLE))}}} \]
\[ \texttt{\texttt{\texttt{'EDISPLAY('Q84))}}} \quad \text{EXPRESSION TOO LARGE} \]

\[ \texttt{\texttt{\texttt{Q84 will not display, so it is reduced in size by naming a subpart.}}} \]
\[ \texttt{\texttt{\texttt{'Q85*SPLIT(Q84))}}} \]
\[ \texttt{\texttt{\texttt{'EDISPLAY('Q85))}}} \]

Next the numerator of the larger term in Q85 is arranged on powers of \( z \).
\[ \texttt{\texttt{\texttt{'Q85*REPLACE('Q85, COLLECT(EXPAND(SUBSTITUTE(('Q85,31,F2,'F2)),F2)),F2), 'HOLE))}}} \]
\[ \texttt{\texttt{\texttt{'EDISPLAY('Q85))}}} \quad \text{EXPRESSION TOO LARGE} \]
\[ \texttt{\texttt{\texttt{'Q85*SPLIT(Q85))}}} \]
\[ \texttt{\texttt{\texttt{\texttt{\texttt{\texttt{\texttt{'EDISPLAY('Q85))}}}}} \}

Forming a term from the renamed parts of Q85 one has for the
coefficient of \( z^2 \) in Q87.
\[ \texttt{\texttt{\texttt{'Q85*SIMPPLY(F2 + F3 + F4 + '2*A1) + F5 + F6))}}} \]
\[ \texttt{\texttt{\texttt{'EDISPLAY('Q85))}}} \]

Looking at equation Q87, one can see that in a double power series
expansion in \( V_1 \) and \( V_2 \), only the coefficient of \( V_1^2V_2^2 \) does not have a
zero of order 4 at \( z = 0 \). Setting this coefficient equal to zero one
obtains by inspection of Q87:
\[ \frac{A_2}{2w(z^2)} + \frac{A_1 + z^2 \cdot w(z^2)A_2}{z(z^2 \cdot w^2(z^2) - 1)} = 0 \]
\begin{align*}
\text{(081)} & \quad H\left(Z^2, U_1, U_2\right) \\
& = \frac{1}{4} Z^2 \cdot U_1 \cdot U_2 \\
& \quad \cdot \left\{ \frac{\frac{1}{2} \cdot (Z+U_1) \cdot (Z+U_2) \left[ Z \cdot W\left(Z^2\right) + 1 \right] \cdot (Z \cdot A_1 + 2 \cdot Z \cdot A_2)}{Z^2 \cdot W\left(Z^2\right) - 1} + \frac{(Z+U_1) \cdot (-Z+U_2) \cdot A_2}{W\left(Z^2\right)} + \frac{(-Z+U_1) \cdot (Z+U_2) \cdot A_2}{W\left(Z^2\right)} \right\} \\
& \quad + \frac{\frac{1}{2} \cdot (-Z+U_1) \cdot (-Z+U_2) \left[ Z \cdot W\left(Z^2\right) - 1 \right] \cdot (-2 \cdot A_1 + 2 \cdot Z \cdot A_2)}{Z^2 \cdot W\left(Z^2\right) - 1} \\
& \quad + \frac{\frac{1}{2} \cdot (Z+U_1) \cdot (Z+U_2) \left[ Z \cdot W\left(Z^2\right) - 1 \right] \cdot (Z \cdot A_1 + 2 \cdot Z \cdot A_2)}{Z^2 \cdot W\left(Z^2\right) - 1} \right\}
\end{align*}
\[ H\left( z^2, u_1, u_2 \right) \]
\[ = \frac{1}{4} \cdot Z^2 \cdot U_1 \cdot U_2 \left\{ \frac{1}{2} \cdot (Z+U_1) \cdot (Z+U_2) \cdot \left[ Z \cdot \omega \left( z^2 \right) + 1 \right] \cdot (2 \cdot A_1 + 2 \cdot Z \cdot A_2) }{ \left[ Z^2 \cdot \omega \left( z^2 \right) - 1 \right] } + \frac{A_2 \cdot \left[ (Z+U_1) \cdot (-Z+U_2) + (-Z+U_1) \cdot (Z+U_2) \right] }{ \omega \left( z^2 \right) } \]
\[ + \frac{1}{2} \cdot (-Z+U_1) \cdot (-Z+U_2) \cdot \left[ Z \cdot \omega \left( z^2 \right) - 1 \right] \cdot (-2 \cdot A_1 + 2 \cdot Z \cdot A_2) }{ \left[ Z^2 \cdot \omega \left( z^2 \right) - 1 \right] } \]
\[ H\left( z^2, u_1, u_2 \right) = \frac{1}{2} z^2 \cdot u_1 \cdot u_2 \cdot \left\{ \frac{\frac{1}{2} (z+u_1) \cdot (z+u_2) \cdot \left[ z \cdot w \left( z^2 \right) + 1 \right] \cdot (2 \cdot A_1 + 2 \cdot z \cdot A_2)}{z^2 \cdot w \left( z^2 \right) - 1} + A_2 \cdot \left[ 2 \cdot u_2 \cdot u_1 - 2 \cdot z^2 \right] \right\} \frac{w \left( z^2 \right)}{w \left( z^2 \right) - 1} \]
Enter the equation into the machine.


Evaluating \( W, A_1 \) and \( A_2 \) one obtains at \( Z = 0 \)

\[ q^{10} \text{SIMPLIFY(SUBSTITUTE(EVALUATE(q^9, (q^6, q^2)), 0, 'Z'))} \]

\#EDISPLAY('Q10')\#

Recognizing that by definition \( P(1,3,0) = P(1,1,0) \)

this equation can be solved for \( P(1,2,0) \).

\[ \text{Q11=SIMPLIFY(SOLVE(SUBSTITUTE(Q10, ('P(1,0,0)', 'P(1,3,0)', 'P(1,2,0)')))} \]

\#EDISPLAY('Q11')\#

Q11 can be simplified somewhat.

\[ \text{Q11=LEFT(Q11) = (-SIMPPLY(FACTOROUT((Q11,11,'Q11,29)/ FACTOROUT((Q11,43,'Q11,49))))} \]

\#EDISPLAY('Q11')\#

Let \( u = z^2 \), the generating function for the unconditional state

probabilities \( G(u) = \sum_{n=0}^{\infty} u^n p(n) \) can now be written. There are special

case terms for \( n = 1, \) and \( n = 2. \) The other terms are found from \( H(u,1,1). \)

The most compact formula for \( H \) is Q86.

\[ \text{Q12=t(G(u)) = ('P(0,0,0)) + 2*W*('P(1,0,0)) + RIGHT(Q11)) + SIMPLIFY(SUBSTITUTE(EVALUATE(RIGHT(Q86), (Q2, Q4)), 'U^+(1/2), 1,1,'Z, V[1],V[2]))} \]

The original transition equations give \( P(1,1,0) = \Theta P(0,0,0), \) anticipating

that \( P(1,3,0) \) may also occur, substitute \( P(1,3,0) = P(1,1,0) + \Theta P(0,0,0). \)
\((Q87)\) 
\[ H\left( Z^2, U_1, U_2 \right) \]
\[ = \frac{1}{4} Z^2 \cdot U_1 \cdot U_2 \cdot \left\{ \frac{(F2 + F3 + F4 + 2 \cdot A_1 + F5 + F6) \cdot Z^2 + 2 \cdot U_1 \left( Z^2 \right) \cdot A_2 \cdot Z^2 + 2 \cdot A_1 \cdot U_1 \cdot U_2}{Z^2 \cdot W\left( Z^2 \right)^2 - 1} \right\} + \frac{A_2 \cdot \left( 2 \cdot U_1 \cdot U_2 \cdot 2 \cdot Z^2 \right)}{W\left( Z^2 \right)} \]

\((Q88)\) 
\[ 2 \cdot U_2 \cdot A_2 + 2 \cdot U_1 \cdot A_2 + 2 \cdot W\left( Z^2 \right) \cdot U_2 \cdot U_1 \cdot A_2 + 2 \cdot B_1 + 2 \cdot U_2 \cdot W\left( Z^2 \right) \cdot A_1 + 2 \cdot U_1 \cdot W\left( Z^2 \right) \cdot A_1 \]

\((Q10)\) 
\[ \frac{\frac{1}{3} \cdot P(1, 3, 0) - (2 \cdot \sigma + 1) \cdot P(1, 2, 0)}{(\theta + 1)} + \frac{1}{2} \cdot (2 \cdot \theta + 1) \cdot P(1, 1, 0) + (-\frac{1}{2}) \cdot P(1, 2, 0) = 0 \]

\[(Q11)\] 
\[ P(1, 2, 0) = -\frac{\frac{1}{3} \cdot P(1, 1, 0) + P(1, 1, 0) \cdot \theta + \frac{1}{2} \cdot P(1, 1, 0)}{(\theta + 1)} \]

\[ \left[ -\frac{\frac{1}{2} \cdot \frac{(-1)}{\theta} \cdot \frac{\theta}{(\theta + 1)} \right] \]
Then use Q11 to eliminate $P(1,1,0)$.

\[\text{#}'Q13\text{=}\text{SUBSTITUTE}(Q12, \text{RIGHT}(Q11), '(P(1,2,0)))'\]

\[\text{#}'Q14\text{=}\text{SIMPLIFY}(\text{SUBSTITUTE}(Q13, '(\text{THETA}^*P(0,0,0), \text{THETA}^*P(0,0,0)),
\text{'}(P(1,3,0), P(1,1,0))))'\]

Q14 contains only the unknown $P(0,0,0)$ which can be determined from
the condition $G(1) = 1$.

\[\text{#}'Q15\text{=}\text{SIMPLIFY}(\text{SUBSTITUTE}(\text{RIGHT}(Q14), 1,',U) = 1)''\]

The machine types out INDETERMINATE, indicating that $\neq 0$ has been
replaced by UNDEFINED.

\[\text{##DISPLAY}'(Q15)'\]

The operator LIMIT will be tried; this operator uses l'Hopital's
Rule. It is slow, and so it should not be used when substitution will
suffice.

\[\text{#}'Q16\text{=}\text{LIMIT}(\text{RIGHT}(Q14), ',U,1)''\]

\[\text{##DISPLAY}'(Q16)'\]

Q16 can be simplified by factoring and expansion.

\[\text{#}'Q17\text{=}\text{FACTOROUT}(\text{EXPAND}(Q16), '(1((3*\text{THETA} + 2)\text{*(1-2*\text{THETA}))})'\]

\[\text{#}'Q17\text{=}\text{LEFT}(Q17)\text{*EXPAND}(\text{RIGHT}(Q17))'\]

\[\text{##DISPLAY}'(Q17)'\]

Q17 is equal to 1 and can be solved for $P(0,0,0)$.

\[\text{#}'Q18\text{=}\text{SOLVE}(Q17=1, ',(P(0,0,0))'\]

\[\text{##DISPLAY}'(Q18)'\]

The expression for $P(0,0,0)$ shows that the SOLVE routine could be
improved. This expression is now substituted into Q14 in order to produce
\begin{align}
(011) \quad P(1,2,0) &= \frac{P(1,1,0) \cdot [8+2 \cdot e \cdot (e+1)+2]}{(3 \cdot e+2)} \\
(015) \quad P(0,0,0) + \frac{2 \cdot P(0,0,0) \cdot e \cdot [e+2 \cdot e \cdot (e+1)+2]}{(3 \cdot e+2)} + 2 \cdot e \cdot P(0,0,0) + \text{UNDEFINED} \cdot \frac{1}{4} &= 0
\end{align}
the final expression for the generating function H.

\$'Q19\cdot\text{Simplify}(\text{Evaluate}(Q14, Q18))'$

Taking a census of Q19 shows that it is probably too large to display without being split.

\$\text{Census}(Q18)\$

\$\text{Eprint}('\text{Last})\$

1438

In the thesis this expression was evaluated numerically to form a table of values. This problem would provide a basis for further work in automatic simplification.

The preceding problem solutions show that the current system can be used for work on existing problems. No one part of the system is particularly weak, but there are many interesting possibilities for improvement.
\[ p(0,0,0) + \frac{2 \cdot \left[ g + 2 \cdot (g+1) \cdot e + 2 \right] \cdot g \cdot p(0,0,0)}{(3 \cdot e + 2)} + 2 \cdot p(0,0,0) \cdot e \]

\[ \frac{16 \cdot \left[ e + 2 \cdot (e+1) \cdot g + 2 \right] \cdot p(0,0,0) \cdot e^2}{(3 \cdot g + 2)} + 16 \cdot p(0,0,0) \cdot e^2 - \frac{6 \cdot \left[ g + 2 \cdot (g+1) \cdot e + 2 \right] \cdot g \cdot p(0,0,0)}{(3 \cdot e + 2)} \]

\[ \frac{+4 \cdot p(0,0,0) \cdot e}{(-2 \cdot e + 1)} \]

\[ 6 \cdot e^2 \cdot p(0,0,0) + 7 \cdot e \cdot p(0,0,0) + 2 \cdot p(0,0,0) + 2 \cdot e^2 \cdot p(0,0,0) \]

\[ \frac{(3 \cdot g + 2) \cdot (-2 \cdot e + 1)}{(-2 \cdot e + 1)} \]

\[ p(0,0,0) = \begin{bmatrix} 2 \cdot e^3 & 2 & 7 \cdot e^3 & 6 \cdot e^2 \\ -6 \cdot e^2 - e + 2 & -6 \cdot e^2 - e + 2 & -6 \cdot e^2 - e + 2 & -6 \cdot e^2 - e + 2 \end{bmatrix}^{-1} \]
In the preceding chapter a system for manipulating mathematical expressions was introduced. In this chapter we discuss additional features which would make this system more convenient in a wider variety of situations. Many of these would let the user tailor the program to his particular needs.

Whenever a user is provided with means for expanding and modifying a program, the question of efficiency becomes an important one. Efficiency normally results from planning a program function as a unit; building the function up bit by bit usually leads to waste. To an extent, though, it may be possible to design the system to guide the user, without unduly restricting him, into avoiding many serious inefficiencies. A careful analysis of the users' needs is required in order to determine what should be built into the program and what can be added later. Some of the features which the user might want are discussed in this chapter. They are divided into three classes: new mathematical structures, algorithms, and notation; bookkeeping; and "systems" instructions to the user and corrections of his mistakes.

**Structures**

There are many frequently-used mathematical structures: sets, rings, groups, fields, vector spaces, matrices, and tensors, to name a few. It might be possible to define some of these structures partly in terms of
the others, but each structure will probably need its own unique data representation for really efficient computation. Some very good polynomial manipulation programs have been written (as noted in Chapter 1); Maurer (4) uses a special representation for finite groups, and Hearn (5) has devised a data representation for tensors. Housekeeping programs should be written to accept the wide variety of data forms which can be expected. Special mathematical structures can form an important part of an input language even if they are not manipulated, since they let the mathematician state his problem in concise and familiar terms. For example, the initial steps in the solution of a set of equations can sometimes be expressed concisely in vector notation. The machine can perform the operations on the individual components.

Notation

Special mathematical structures will require symbols to represent them. The user should be able to define new symbols, and new arrangements of symbols. Hand-written input would make it easy to define new symbols; they could also be assigned to blank buttons or keys. In the current system, a string of input characters such as $\Omega B\Omega A$ can be assigned an output symbol such as $\omega$. The system could be extended so that the user could define a new output symbol by choosing points to be displayed from a raster or by drawing a series of straight lines with the light pen. The user should also be able to make abbreviations such as $x'$ for $\frac{dx}{dt}$. New combinations of symbols such as summation, $\sum_{i=0}^{\infty} x^i$, could be described in the picture language given later, but a more interesting possibility would be to generalize from hand drawn examples.
Algorithms

The user will want to add new mathematical functions and transformations in several ways. Special properties will be important for a given operation. For example, a simplification routine needs to know that \( \sin^2 x + \cos^2 x = 1 \), while a differentiation routine needs to know that \( \frac{d \sin x}{dx} = \cos x \). A user may want to define a function only for certain arguments and later expand his definition. All the machinery of the "advice taker" programs, such as Teitlaman's PILOT(1), could be used. When a transformation reaches arguments for which it is undefined, it can ask the user for advice. For example, in the current system the power series expansion routine will query the user about the poles of expressions, or simply ask for the expansion of a subexpression. It would be better if the user could input a general method for handling such expressions. However, in that case he must define the conditions under which it applies, and this requires a language which deals with concepts such as "rational function". The problem of combining new information with the old is an important one and should be pursued in the future.

New transformations could also be defined as combinations of old ones. If the string of user commands is remembered by the machine, the user could repeat an operation by referring back to his earlier commands. The string of commands could be displayed, so that the user could edit it or add comments. He might assign a name to a substring of commands and give it arguments to produce a function. These command strings could form a programming language with the addition of conditional expressions. Furthermore, the commands could be treated as mathematical
expressions and should be legitimate data for the system, so that they
could be simplified or otherwise transformed. The development of a
command language will require the introduction of properties useful for
describing mathematical expressions. For example, a method might apply
only to expressions rational in sines and cosines. A specification of
this method would require a check to see if the expression had this
form. Such a language would allow a mathematician to describe his
problem solution to the machine in the most general way, thereby
increasing the number of problems which could be solved with the same
string of commands. In applying the same commands to other problems,
the machine could remember certain characteristics of the original problem
and ask the user for help when new problems require further generalization
or new methods.

Numerical Algorithms and Notation

A mathematical laboratory also needs a strong numerical capability.
Even when symbolic manipulation is an essential part, most practical (and
probably most theoretical) problems will require numerical methods
somewhere in their solutions. Sometimes the equations describing a
numerical method must be tailored to suit each problem. For example,
an iterative technique may have a starting procedure which depends on
the boundary conditions. This tailoring might be done with symbolic
routines. In other cases, the equations to be processed numerically
must be transformed symbolically before the numerical routines can be
applied. This was the case in the second problem solved in the previous
chapter.
Numerical evaluation can also be used in order to gain insight regarding exact or approximate symbolic solutions. For numerical work, mathematical structures must be represented as numerical arrays. Methods are needed for converting expressions from symbolic form to numerical form and the reverse. In addition, a method is needed to link with the large store of existing numerical routines. Finally, there must be input-output for numerical data in the form of tables and graphs. The MAP system and the Culler system provide many of the features needed, such as automatic scaling of graphs, formatting of tables, and the use of the light pen and parameter knobs for describing and altering numerical functions and constants.

Bookkeeping

Routine bookkeeping is often required for the solution of a large problem. The machine should assume this function in an inobtrusive manner. For example, the machine might keep a record of what expressions had been combined to form other expressions and display this information in the form of a graph. Touching some part of the graph with the light pen would cause the machine to display that part in greater detail. There might be a way for the user to name pages of expressions or to ask for all of the expressions with a certain property.

The machine could find missing relations needed to solve a set of equations, so that the user would not have to specify the arguments completely.
There would probably be many questions which the machine could answer about past inputs. Of course, the machine can also collect data of interest to system designers, such as the distribution of properties of expressions to which a given transformation is applied.

**User Errors**

Finally, the machine must help the user overcome his difficulties in using the system. It might be possible to develop a systematic method of rejecting inputs which are syntactically or semantically meaningless. The formal structure of mathematics makes this easier than it would be for some other real-time systems. The machine can use heuristics to determine what the user might have meant. Optional instructive comments like those in the MAP system would also be useful.

As can be seen from this chapter, the working mathematical laboratory will be a very complicated system. Development of the means necessary in order to create such a system will be one of the important steps in the next generation of programming research.
Chapter IV

INTRODUCTION TO THE REMAINING CHAPTERS

The remaining chapters describe the system in great detail and suggest possibilities for further work. These chapters are largely independent of one another and can be read in any order. Each chapter or section becomes more detailed toward the end. Therefore, one strategy would be to read the first part of each chapter and then read some of them to the end. Chapter VII, Chapter IX, and the first part of Chapter VI contain the most significant results.
Chapter V

PROBLEM AREAS

In this short chapter we mention a number of classical mathematical transformations for which it should be possible to write computer programs. Some have already been investigated by others, but all of them need more study. For some areas it seems possible to find an algorithm which will handle many of the expressions which arise in practice. Examples of these are differentiation, solution of equations by Gaussian reduction, and certain matrix and tensor operations. Even though an algorithm may in theory handle all expressions of a given class, it may in fact be limited by space or time considerations. Often, however, it can be improved through use of the mathematical properties of the transformation to be performed. For example, a sum with a positive integer exponent can always be expanded by repeated multiplication followed by collection of like terms. However, expansion of $\left(A_0 \cdots A_n\right)^K$ in this manner leads to $(n+1)^K$ terms, which combine to yield $\binom{n+K}{n}$ different terms. The $\binom{m}{n}$ terms can be found directly by the multinomial theorem; thus, in general, reducing the time and the maximum storage required. In the case of Gaussian reduction, savings can be made on the average through manipulation of zeros.
In other problem areas, algorithms can be used to transform only some examples, but the range of examples which can be transformed can probably be considerably extended with heuristic methods. Examples of these areas are limits, Laplace transforms, expansion in a Laurent series, integration of elementary functions in closed form, and solution of some differential equations.

Finally, in the case of non-linear differential equations, for example, it seems that only isolated special cases could be solved by the machine. It is, however, difficult to estimate a priori whether further research will lead to useful methods. As stated in the introduction, integration of the elementary functions in closed form has been proved recursively unsolvable. Such a proof not only affirms the futility of finding a general scheme but can also indicate which subclasses may be impossible. Perhaps more such proofs can be found.

In the absence of such a proof, another measure of problem difficulty is the extent to which the solution is a global function of the input mathematical expressions making up the problem statement. When the solution is not a global function, then local properties of the input may give clues to local properties of the solution, thus making it possible to guess at the form of the answer.

Mathematical transformations should be a good medium for the investigation of computer problem solving.
Chapter VI
INTERNAL REPRESENTATION AND MANIPULATION
OF MATHEMATICAL EXPRESSIONS

Introduction

The main body of this long chapter contains detailed descriptions of
the effects of the LISP routines which have been written to manipulate
mathematical expressions. This detailed description is introduced by a
general discussion of the problems and possibilities involved in writing
such routines. As the structures to be manipulated become more complex,
the methods for representing them and for describing their transformation
become, in the current usage, more a matter of choice. This is because
the complex structures are described with a very general syntax, the
ambiguities being resolved through context, and because there are fewer
well established conventions for representing these structures.
Therefore, the problem of manipulating the elementary functions will be
discussed in some detail before more general structures involving sets
are introduced.
The Elementary Functions

A. Introduction

The expressions for the elementary functions are defined recursively as follows. All floating point numbers, integers, and pairs of integers representing rational numbers are elementary functions. If $u$ and $v$ are elementary functions, then $u+v$, $u-v$, $u^v$, $u/v$, $\log u$, and $-u$ are elementary functions. The trigonometric functions may be represented explicitly. In addition, arbitrary functions will be allowed. These functions are assumed only to be differentiable. It is allowed to write $a+b+c$ and $a\cdot b\cdot c$; the program assumes association.

Normally, transformation of these expressions involves two steps. Certain properties of the source expression are investigated and then the results of this investigation control the flow of control as the output expression is generated, often by combining parts of the source expression. It is difficult to list the important properties of expressions without examining actual problems. Some useful properties are those used by Slagle's Integration program SAINT and the routines described at the end of this chapter.

B. Properties of Elementary Functions

The properties can be classified according to their complexity as follows. Those from SAINT are indicated by (S).

1. Tests involving no investigation of subparts of the expression.
   a. Is the expression an integer. (S)
   b. Is the expression the variable of integration. (S)
   c. Expression is 0.
   d. Expression is 1.
2. Tests which are applied to the expression and its arguments
   in a completely recursive manner; that is, the test which
   is applied to the expression involves applying the identical
   test to the arguments of the expression, unless a terminating
   condition is met.
   a. Depth: the maximum number of arguments within arguments. (§)
   b. Length: the number of subexpressions in the expression. (§)
   c. Algebraic function: the expression contains only the
      algebraic operators and constant exponents. (§)
   d. Dependent on a given variable. (§)

3. Tests which involve applying tests to the arguments of an
   expression which are a function of the main connective of
   the expression.
   a. Rational function: if the main connective is POWER,
      then the exponent must not be a rational function, but
      rather an integer. (§)
   b. Rational function of sines and cosines: if the main
      connective is a sine or cosine, then the argument must
      be the variable of integration. (§)
   c. Exponential: as in b., the expression contains a
      subexpression of the form e^y. (§)
   d. Integral with indefinite limits.
   e. Integral with respect to a given variable.
   f. Derivative operator applied to an arbitrary function.
   g. Lowest power of a variable in power series expansion
      of an expression.
b. $\text{SUM}$ with numerical lower limit.

1. $\text{SUM}$ with numerical or infinite upper limit.

4. Tests which involve the comparison of certain subparts of the expression.
   a. Exponent base: there is some constant, $b$, and variable, $x$, such that any subexpression which is an exponential is of the form $b^{nx}$, where $n$ is an integer. (S)
   b. Comparison of arguments of a product or sum for equality.
   c. All terms of a sum contain a common factor.

5. Non-recursive structures made up of parts each of which has certain properties.
   a. The expression contains a non-constant sum raised to an integer power. (S)
   b. The expression contains a non-constant sum as a factor of a product. (S)
   c. There is a subexpression of the form $c_3 + c_2v + c_1v^2$; $c_1 \neq 0$, $c_2 \neq 0$. (S)
   d. The expression is the product of two factors with certain properties. (S)
   e. The expression is the product of a factor with given properties and the expression $\cos^{2n+1}v$, where $v$ is a variable. (S)
   f. The expression is an exponential with constant exponent.
g. Is expression of the form $x^n$; $n$ an integer.

6. Some transformations require arbitrarily complex computations on the parts of the expression in order to establish the flow of control.
   a. Comparison of ranges of several summations and grouping together of like segments of the ranges.
   b. Generations of the lower order terms of a multinomial expansion.

C. **Methods for Testing Properties of Elementary Functions**

There are several mechanisms which can be used to determine if a given expression lies in one of these classes. The best method depends on which of the six named classes is involved as well as the associated constructive operations and the nature of the expressions under consideration.

Many transformations can be applied in turn to larger and larger subparts of the expression. Others are applied to certain parts of the expression as a whole. While those of the most complex type must make use of intermediate results as well as the initial expression.

It would be unlikely that any one scheme would be the most efficient for every combination of analysis and construction. The various transformations should be handled by methods appropriate to their complexity.

The efficiency of a transformation must be taken as an average over the set of expressions to which it is applied. Very often a transformation is applicable to only some of the expressions to which it is applied. Since a great deal of time can be spent in attempting to apply irrelevant transformations, the efficiency of a transformation might
be evaluated in two ways: the average time required for application to
a random expression, and the average time required for application to an
expression to which it is appropriate. As expressions are transformed,
more information about them will become known; it might be possible to
pass this information along so as to improve the efficiency of further
transformations. Furthermore, the physical situation which the
expressions represent may give clues to their properties. Preserving
such information involves trading space for time.

The expressions being considered form a tree structure; the
main connective is considered to be at the top, with the branches
hanging down. Some methods typically come down from the top, others
come up from the bottom; some first go one way and then the other.
Class 1 tests are applied at the top.

Class 2 tests can be applied going either way. For instance,
in the case of depth, one can come down the tree incrementing the depth
count at each node and then take the maximum of the counts at the ends
of the branches, or one can come up the tree, incrementing the maximum
branch count at each node. The number of incrementing and binary
maximum operations is the same in either case.

The situation is more interesting for class 3 tests. Since
the test to be applied to some of the subexpressions depends on the
main connective above, all possible tests must be applied if one comes
up from the bottom. Coming down from the top one can bring the
information about the main connective along and apply only the needed
test. If several tests are applied at once, then most of the tests will
probably be needed on a given subexpression and not so much is lost.
An advantage of applying all tests to a subexpression is that identical subexpressions can be treated in the same manner since they are not distinguished by context. If several tests can be computed at once by bit manipulations, then the extra cost would be less. It seems that the great loss is probably that many of the tests, if applied from the top, would fail or succeed before all of the subexpressions had been tested, so that a saving would result from some subexpressions not being tested at all. (This seems to be a weakness of K. Korsvold's program).

The tests in class 4 involve the comparison of expressions for equivalence. Because of the defining relations, equivalent expressions can be written in several forms. For instance, the factors of sums and products can be written in any order. Some existing schemes impose some additional constraints on the way expressions are written so that there will be a canonical form. Another possibility is to compute a hash code number for each expression which is invariant under the alternate ways of writing the expression. The method of doing this and the merits of the scheme are discussed in the next chapter.

The exponent base operation is the first test which requires comparison of subexpressions located at arbitrary nodes of the expression tree. A program which goes up or down the expression tree must have some way to communicate the information at one node to itself when it is processing another node.

The tests in class 6 are those which do not fit into one of the other classes. Their existence indicates that although programs which move up and down the expression tree are useful in a large number of cases, there will be tests for which this type of executive structure
is very inefficient. Programs which carry out complex algorithms involving many forms of intermediate results will be needed. To introduce concepts which will be useful in such programs, operations for constructing elementary functions will now be considered.

D. Constructing Elementary Functions

All constructions which form elementary functions from other elementary functions must do so either by substitution or by combining the input functions with the elementary operators. Useful intermediate forms are a string of terms which are to be factors in a sum or product. Keeping such a string removes the need to test each time before adding new terms.

E. Specification of Algorithms for Manipulating Elementary Functions

No matter what level or flexibility a language for algorithm specification is to have, to be convenient and efficient it must reflect the structure of the algorithms to be specified. Based on the preceding discussion the following types of operations are needed. One begins with a basic program sequence that goes down and up the expression tree. When this sequence is associated with a specific set of rules, it is given a name by which it can be called. There are rules for the downward pass and the upward pass. The rules for the downward pass consist of a conditional specification of what is to be applied to the arguments, and a method of reserving information for lower levels and also for any following operation. The conditional is a match containing the operator and functions or exact match on the arguments. The arguments and groups of them can be given mnemonic names to be used in this match. There is a way to name all arguments satisfying some match. Matched arguments can
be given special names to be used in the rest of the specifications. There is a way to break out immediately with the answer on either the upward or the downward pass. The upward pass consists of another match and a construct. The construct employs a pattern but the operators in the pattern which have variable arguments are simplified if possible. The construct pattern can contain functions. FLIP(III-II) might be useful in such a language.

**Representation of Sets**

A set might be represented by giving its name and some of its properties, or by giving a general member and directions on how to specialize this general member to any particular member, or by giving all of the members. Members of ordered sets can be indexed to form arrays. More than one of these representations could be used simultaneously. Typical sets are an indexed summation of terms, or an indexed set of equations. In collecting terms one might want to combine several summations of the form \( \sum_{i=m}^{n} a_i x^i \) to form a new set of summations, no two of which contain the same power of \( x \). This amounts to combining members of different sets; that is, the coefficients, which members have the same value of a property, that is power of \( x \). The summation index makes it possible to carry out this operation without explicitly testing each member of each set. This is a very general form in which many algorithms dealing with sets can be expressed, but greater computer efficiency and clarity of expression might result from using special concepts.
As another example of the use of a set concept, consider the placement of the coefficients of a polynomial into an array, so that the Routh test can be applied. A series of terms is generated by combining these coefficients. The generation of these terms is best specified as operations on an array of expressions. The requirement that the resulting terms all have the same sign then leads to relations between their parameters.

Two further examples are similar. Any particular term in a multinomial expansion can be written down directly in terms of the base and the exponent. This is done by forming all combinations of powers of the variables which fulfill certain restrictions. Computation with integers is used to determine the flow of control as the powers are formed. A similar but more complex case is the formation of terms in the power series of the multidimensional transformation of a non-linear differential equation as described by Van Trees. (V2)

As a final example, consider the multiple channel, multiple phase queueing system described by Garber and discussed in Chapter II. The basic equations describe a system such as the one shown below:

```
<p>| | | | | | | | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
```

Channel 1

Channel 2

Channel 3
A system of \( n \) channels, each with \( k \) phases of service, is described by the total number \( n \) of units in the system and the state \( r_j \) of each channel. The arrival rate is \( \lambda \) and the service rate is \( \mu \). The equations are

\[
(\text{n+1}) p(n; r_1, r_2, \ldots, r_n, 0, \ldots, 0) = 
\sum_{j=1}^{n} p(n; r_1 + \delta_{j,1}, r_2 + \delta_{j,2}, \ldots, r_n + \delta_{j,n}, 0, 0, \ldots, 0) 
+ \sum_{j=n+1}^{n} p(n+1; r_1, \ldots, r_n, \delta_{j,n+1}, \delta_{j,n+2}, \ldots, \delta_{j,n})
\]

whenever \( n \leq m \) and subject to:

\[
k u p(n; r_1, \ldots, r_{i-1}, k+1, r_{i+1}, \ldots, r_n, 0, \ldots, 0) = 
\frac{\lambda}{n-k+1} p(n-1; r_1, \ldots, r_{i}, 0, r_{i+1}, \ldots, r_n, 0, 0, \ldots, 0) 
+ \delta_{n,n-k+1} k u p(n+1; r_1, \ldots, r_{i-1}, 1, r_{i+1}, \ldots, r_n)
\]

whenever \( r_i = k \). While if \( n > m \)

\[
(1+\theta) p(n; r_1, r_2, \ldots, r_m) = \frac{1}{m} \sum_{j=1}^{n} p(n; r_1 + \delta_{j,1}, r_2 + \delta_{j,2}, \ldots, r_n + \delta_{j,n}) 
+ \theta p(n+1; r_1, r_2, \ldots, r_m)
\]

subject to

\[
p(n; r_1, \ldots, r_{i-1}, k+1, r_{i+1}, \ldots, r_m) \equiv p(n+1; r_1, \ldots, r_{i-1}, 1, r_{i+1}, \ldots, r_m)
\]
when \( r_{l} \leq k \). It is understood that although only those equations for which the full channels have the lowest index have been written, there is in fact an equation for each permutation.

Note that this representation involves the use of several external conditionals, some written in English, as well as the internal conditional \( d_{ij} \). In addition, it contains the inductive "..." notation and the English statement about permutations.

Some of the external English conditional statements arise because the \( d_{ij} \) does not provide enough generality as an internal conditional. The "..." notation arises in part because there is no standard way to mention a particular subscript of a quantity which is indexed by a number of subscripts which is itself a variable. It is for this same reason that the English statement about the permutations is made. The syntax used has great flexibility, but does not contain concepts appropriate to this statement. If the expression is represented internally in the computer in this form, it will be difficult to write algorithms to transform it.

An alternate form of notation would involve introducing a general internal conditional, functional arguments, and some specific concepts about sets. The function \textsc{setsum} \([\text{set}, \text{function}]\) has as value the sum of the quantities produced by applying the function \text{function} to each element of set \text{set}. The function \text{setsubst} \([x, y, z]\) substitutes element \( x \) for element \( y \) in set \( z \). The concept of bound variables is also used. The current element to which \text{setsum} is applying \text{function} is bound to the name, element. In this manner, the expression can be written:
\[ \text{EQN} (N, N, K, \text{RSET}, \emptyset) \equiv \]
\[ \text{If } N \leq M \text{ then } (M+\emptyset) \cdot P[N, \text{RSET}] = \text{SETSUM} (\text{RSET}, \emptyset) \]
\[ \text{if } \text{element} = 0 \text{ then } P[N+1, \text{SETSUBST} (1, \text{element}, \text{RSET})] \]
\[ \text{or if } \text{element} = k \text{ then } \frac{\lambda}{M-N+1} \cdot P[N-1, \text{SETSUBST} (0, \text{element}, \text{RSET})] \]
\[ + \text{ if } N=M \text{ then } k \cdot P[N+1, \text{SETSUBST} (1, \text{element}, \text{RSET})] \]
\[ \text{else } 0 \]
\[ \text{else } P[N, \text{SETSUBST} (\text{element} + 1, \text{element}, \text{RSET})] \]
\[ \text{else } (1+\emptyset) \cdot P[N, \text{RSET}] = \text{SETSUM} (\text{RSET}, \emptyset) \]
\[ \text{if } \text{element} = k \text{ then } \frac{1}{m} \cdot P[N+1, \text{SETSUBST} (1, \text{element}, \text{RSET})] \]
\[ + \emptyset \cdot P[N-1, \text{RSET}] \]
\[ \text{else } \frac{1}{m} \cdot P[N, \text{SETSUBST} (\text{element} + 1, \text{element}, \text{RSET})] + \emptyset \cdot P[N-1, \text{RSET}] \]

There are several alternative ways in which this expression could be written. The choice would depend largely on the operations to be performed. The typical procedure is to apply a number of steps and then to generate a special case from the general expression.

In closing, another area where the problem solution goes from the general to the particular is the area of vector spaces and matrices.

Methods of Specification

The amount of information about the input to be saved in the internal representation depends on the range of tasks which the machine is to perform in the man-machine interaction. One possible task is the arrangement of statements input in any order into an order which allows their sequential evaluation. In some cases missing steps or limits could be pointed out. This requires some representation of what statements belong together as well as properties of the statements such as the
information needed to evaluate them. With respect to evaluation, the
more problem oriented, rather than procedure oriented, the man-machine
discourse becomes, the more necessary it is to represent objects by a
set of properties which would allow generation of the object by an
appropriate routine. This is true, for example, of sets representing
permutations and combinations subject to certain restrictions.

Such properties are also needed if the machine is to perform
difficult problem solving tasks. Often, for example, the factors of a
polynomial can be found only if information about them is known from
physical considerations. This information must be stored as a property
of the polynomial.

The Experimental Routines

The mathematical expression transformation routines will now be
described in detail. Conclusions about the types of program
organization which were needed for these routines have already been
presented in the first part of the chapter. This detailed description
will be useful to those who want to write similar programs. The
transformation routines are written as a hierarchy of LISP functions.
Some of the functions perform operations which are common to many of
the transformations. These functions comprise a low level programming
language. In addition, there are many functions which perform system
and bookkeeping operations. To see the purpose of each of the function
types, the processing of a sample command will now be traced through the
system. A knowledge of the LISP programming language will be assumed for
the remainder of the chapter.
Suppose the user were to type the command:

\[
\_\_E := \text{Simplify}('X, 1, 'X^2 + '(X+1)^2))\]

the PDP-6 sends this command, as a list of characters, to the 7094. There, the LISP function \texttt{APARSE}, described in Chapter X, is used to parse the string of characters into a LISP expression. The above command becomes:

\[
(\texttt{EAASSIGN} (\texttt{EQUOTE} E) (\texttt{SIMP} \texttt{LIFY} (\texttt{DRV} (\texttt{EQUOTE} X) 1) (\texttt{PLS} (\texttt{PWR} (\texttt{EQUOTE} X) 2) (\texttt{PWR} (\texttt{EQUOTE} (\texttt{PLS} X 1) ) 2))))
\]

This expression is then evaluated in the LISP system. All mathematical expressions are kept on the disk. The function \texttt{EAASSIGN} is one of several \texttt{system functions}. It will write the mathematical expression which is the value of its second argument on the disk as a file with name \texttt{E FORM}. The first name \texttt{E} has been pseudo-quoted with \texttt{EQUOTE}. \texttt{EQUOTE} transforms typed input expressions into internal form. It is one of several \texttt{functions for transforming expressions into equivalent forms for input or output}. Next, consider the evaluation of the arguments of \texttt{DRV}, the derivative function. \texttt{DRV} will take the first derivative with respect to \texttt{X} of its third argument. Evaluating the third argument, \texttt{X} and \texttt{2} are given to the function \texttt{PWR}, which is one of a group of \texttt{functions for performing the algebraic operations}. \texttt{PWR}([X, 2]) \rightarrow (\texttt{PWR} X 2 \texttt{NIL}) Similarly \texttt{EQUOTE} [(\texttt{PLS} X 1)] \rightarrow (\texttt{PLS} X 1 \texttt{NIL}), \texttt{PWR} ([\texttt{PLS} X 1 \texttt{NIL}], 2) \rightarrow (\texttt{PWR} (\texttt{PLS} X 1 \texttt{NIL}) 2 \texttt{NIL}), and the final result is (\texttt{PLS}((\texttt{PWR} X 2 \texttt{NIL}) (\texttt{PWR} (\texttt{PLS} X 1 \texttt{NIL}) 2 \texttt{NIL}) \texttt{NIL}). The functions \texttt{PLS} and \texttt{PWR} will perform certain simplifications, such as \texttt{PLS} [X, (\texttt{PLS} A B \texttt{NIL})] \rightarrow (\texttt{PLS} X (\texttt{PLS} A B \texttt{NIL}) \texttt{NIL}). This reduces the growth of expressions because of simple redundancies. Next, the
derivative function DRV is called. DRV issues a call to

$$\text{DIFF}(x, \text{PWR}(\text{PLS} X 2 \text{NIL}) (\text{PWR}(\text{PLS} X 1 \text{NIL}) 2 \text{NIL}) \text{NIL}) \text{NIL}).$$

The function

DIFF is one of several large routines for transforming the elementary
functions. It would take up too much space if it stayed in core
memory at all times. Therefore, DIFF is defined as a call to the

system function GETFILE. GETFILE[DIFF, (X EXP)] first reads in the
disk file DIFF LISP, which contains the DIFF routines. Then GETFILE

evaluates DIFF[X, EXP]. Finally, GETFILE reads the disk file DIFF
ERASE, which removes the DIFF routines from core memory. As stated
above, DIFF is given the expression:

$$(\text{PLS} \text{PWR} X 2 \text{NIL}) \text{PWR}(\text{PLS} X 1 \text{NIL}) 2 \text{NIL} \text{NIL}).$$

Before differentiating this expression, DIFF gives it to MKP, the

function for putting the variables in all subexpressions on their

property lists. MKP returns $$(\text{PLS}\text{PWR} X 2 (X) \text{PWR}(\text{PLS} X 1 (X)) 2$

$$(X) (X)).$$ Now DIFF can look on the property list of an expression

in order to see if it depends on the variable of differentiation. If

it does not, DIFF will immediately return the answer, 0. For the

example above, DIFF returns:

$$(\text{PLS} \text{PRD} 2 X \text{NIL}) \text{PRD} 2(\text{PLS} X 1 \text{NIL})(\text{PLS} 1 0 \text{NIL}) \text{NIL} \text{NIL}).$$

This result is given to SIMPLIFY. This routine is used so often that it

has been chosen to reside permanently in core memory. Therefore, no
call to GETFILE is needed this time. SIMPLIFY finds that

$$(\text{PLS} 1 0 \text{NIL}) \rightarrow 1,$$ then that $$(\text{PRD} 2 (\text{PLS} X 1 \text{NIL}) 1 \text{NIL}) \rightarrow$

$$(\text{PRD} 2 (\text{PLS} X 1 \text{NIL}) \text{NIL}).$$ Since this later expression is the argument

of a PLS, it is expanded to $$(\text{PLS} \text{PRD} 2 X \text{NIL} 2 \text{NIL}).$$ Thus, at the
top level SIMPLIFY finds:
(PLS (PRD 2 X NIL) (PRD 2 X NIL) 2 NIL), which can be simplified by collecting terms. To do this, the routine must discover that two of the terms have the same literal factor. To compare these factors, SIMPLIFY uses the functions for computing the hash code number of a symbolic expression. If the hash code numbers are the same, SIMPLIFY assumes that the factors are equivalent. The final answer, (PLS 2 (PRD 4 X NIL) NIL) is now written on the disk by EASSIGN.

In addition to the five classes of functions above, there are also a few general purpose functions. The functions in each of these six classes will now be explained. Then the functions which comprise each of the transformation routines will be described. These descriptions use several special terms.

The expressions are represented in LISP in operator prefix form. The individual operators are:

(Note: most of the operators can take any number of arguments in the obvious manner.)

1. (PLS A B C NIL) ≡ A+B+C
2. (PRD A B C NIL) ≡ A·B·C
3. (FRT I J NIL) ≡ \( \frac{I}{J} \)
4. (PWR A B NIL) ≡ A^B
5. (DRV A B C D E NIL) ≡ \( \frac{d^{B+D}}{d^{A+D}} \)
6. (ITG D A B C NIL) = \( \int_A^B \frac{C}{dD} \)
7. (SUM A B C D NIL) ≡ \( \sum_{A=B}^C D \)
8. \((\text{EVL} \ A \ B \ C \ D \ E \ \text{NIL}) \equiv E \int_{A=B}^{C=D}\)

9. \((\text{NAM} \ A \ B \ C \ \text{NIL}) \equiv C_{A,B}\)

10. \((\text{F} \ A \ B \ \text{NIL}) \equiv F(A,B)\)

11. \((\text{NAM} \ A \ B \ (F \ C \ D \ \text{NIL}) \ \text{NIL}) \equiv F_{A,B}(C,D)\)

12. \((\text{FTL} \ A \ \text{NIL}) \equiv A'\)

13. \((\text{ABS} \ A \ \text{NIL}) \equiv |A|\)

14. \((\text{DEL} \ A \ \text{NIL}) \equiv \frac{d}{dA}\)

15. \(\text{INF} \equiv \infty\)

16. \(\text{IDF} \equiv \text{indefinite}\)

17. \(\text{PI} \equiv \pi\)

18. \(*E \equiv e\)

19. \(*I \equiv i\)

20. \(\text{NIL} \equiv \text{undefined}\)

21. \((\text{XST} \ I \ B \ C \ A \ \text{NIL}) \equiv \)

22. \((\text{DIS} \ I \ B \ C \ A \ \text{NIL}) \equiv \)

23. \((\text{CND} \ (A \ B \ \text{NIL}) (C \ D \ \text{NIL}) \ \text{NIL}) \equiv \frac{A+B}{C+D}\)

The last element of each list is the property list of the expression. An empty property list is represented by NIL. When the operator is removed from an expression, the argument list remains. For example, the argument list of \((\text{PLS} \ A \ B \ C \ \text{NIL})\) is \((A \ B \ C \ \text{NIL})\). The body of an operator is the last argument in the list. For example, the body of \((\text{NAM} \ A \ B \ C \ \text{NIL})\) is \(C\). The expression \((F \ A \ B \ \text{NIL})\) is called an arbitrary function of \(A\) and \(B\). An expression such as

\((\text{PLS} \ (\text{PRD} \ B \ C \ \text{NIL}) \ (\text{PRD} \ C \ (\text{PWR} \ D \ E \ \text{NIL}) \ \text{NIL}) \ \text{NIL})\)
forms a tree of subexpression as shown below:

When a function is said to go **down the expression tree**, it processes the nodes of this tree in the normal order for LISP.

**Functions Which Put the Dependence Property on the Property Lists of All the Subexpressions**

As was stated in the introduction, it is sometimes useful to put the variables in a subexpression on its property list. This list of variables will be called the **dependence list** of the subexpression. Subscripted function names are distinguished from variables as can be seen in the following examples:

\[
(F \times Y \text{ NIL}) \rightarrow (F \times Y (X Y))
\]

\[
(\text{PLS A (F X Y NIL) NIL}) \rightarrow (\text{PLS A (F X Y (X Y)) (F A X Y)})
\]

\[
(\text{PLS A (NAM 1 (F X Y NIL) NIL) NIL}) \rightarrow (\text{PLS A (NAM 1 (F X Y (X Y)) (NAM 1 F) X Y)})
\]

The dependence lists are computed by **MKP[X]**. **MKP** goes down the expression tree to the atoms, which are left unchanged. Coming back up the expression tree, each argument list is given to **MKP1**. **MKP1** \([\text{ARGLIST, EXP, PLIST}]\) goes down the argument list making a dependence list by taking the union of the variables on which each argument depends with the accumulated result, **PLIST**. There are five cases necessary to
find the dependence of an argument. If the argument is a number, its
dependence is NIL. For an atomic argument, the result is a list of
the atom. If the argument is an arbitrary function, the result is the
function name added to the argument property list. All other expressions
except NAM bodies yield their property lists. As shown in the examples,
NAM bodies yield the entire expression with the property list removed
by RPLST and with arbitrary functions replaced by their names, combined
with the dependencies of the NAM body function arguments.

Functions Which Compute the Hash Code Number of an Expression

Chapter VII presents a scheme for assigning integers to expressions
in such a way that equivalent expressions will be assigned the same
integer. These integers are called hash code numbers because the
infinite set of mathematical expressions is mapped into a finite set
of integers by the assignment scheme. As explained in Chapter VII, if
\( a^{x+2} \) is to have the same hash code number as \( a^x \cdot a^2 \), then a special hash
code scheme must be used for exponents. For want of a better solution,
exponents of exponents are treated with the base scheme. The current
depth of exponentiation module 2 is assigned to the variable SIMPLEVEL.
SIMPLEVEL is 0 for the base arithmetic, which was called F arithmetic in
Chapter VII. SIMPLEVEL is 1 for the exponent or S arithmetic. The F
and S arithmetic is performed by the functions FPLS, FPRD, FDVD, FPWR,
SPLS, SPRD, SDVD, and SPWR. All of these functions except FDVD are
coded in LAP for greater efficiency; the functions perform the following
computations:

(Note: A and B are arguments and p is the finite field prime.)

**FPLS:** Answer = \((A+B) \mod p\)

**FFRD:** Answer = \((A \cdot B) \mod p\)

**FPWR:** First a check is made to see if the base is equal to the integer assigned to e. If so, special code is entered so that \(e^{n/2}\) will often hash into \(-1^n\). In any case, a check must be made for a negative exponent which indicates that the finite field square root of the base is to be taken. Finally, the base is multiplied by itself in order to create the series base, base², base⁴, base⁶ \(\cdots\) base²ⁿ. The answer is the product of all the terms in this sequence for which the corresponding bit of the exponent is a 1. The flow chart is:
SPLS: Either argument can be positive or negative.

\[
\begin{align*}
A > 0 & \quad B > 0 \quad \text{Answer} = (A+B) \mod (p-1) \\
A > 0 & \quad B < 0 \quad \text{Answer} = (-2A+B) \mod (p-1) \\
A < 0 & \quad B > 0 \quad \text{Answer} = (A-2B) \mod (p-1) \\
A < 0 & \quad B < 0 \quad \text{If } A+B \text{ is odd: Answer} = (A+B) \mod (p-1) \\
& \quad \text{Otherwise: Answer} = (\lfloor A+B \rfloor/2) \mod (p-1)
\end{align*}
\]

SPRD: Either argument can be positive or negative.

\[
\begin{align*}
A > 0 & \quad B > 0 \quad \text{Answer} = (A-B) \mod (p-1) \\
A > 0 & \quad B < 0 \quad \text{If } A-B \text{ is even: Answer} = \lfloor (A-B) \mod (p-1) \rfloor/2 \\
& \quad \text{Otherwise: Answer} = (A-B) \mod (p-1) \\
A < 0 & \quad B > 0 \quad \text{If } A-B \text{ is even: Answer} = \lfloor (A-B) \mod (p-1) \rfloor/2 \\
& \quad \text{Otherwise: Answer} = (A-B) \mod (p-1) \\
A < 0 & \quad B < 0 \quad \text{Answer} = -\phi\left\lfloor \frac{(A-B) \mod (p-1)}{2} \right\rfloor \mod (p-1) \\
& \quad \phi = 2n+2 \quad \text{where } p = 16n+13
\end{align*}
\]

SPWR: If the base is negative, then it is multiplied by \( (2n+2) \mod (p-1) \).

It is then raised to the exponent in the same manner as for FPWR.

First, the series base, base\(^2\), ..., base\(^{2^n}\) is formed, where each term is the square of the preceding one, taken mod \( p-1 \).

Then the answer is formed by multiplying together mod \( p-1 \) the terms \( k \) for which the \( k \)th digit of the exponent is a one.
SDVD: A is the numerator and B is the denominator.

\[ \begin{align*}
A > 0 & \quad B > 0 \quad A \text{ odd} \rightarrow B \text{ odd} \rightarrow \text{Answer} = (A \cdot B^{-1}) \mod (p-1) \\
& \quad T \rightarrow \text{Answer} = [-A \cdot (B/2)^{-1}] \mod (p-1) \\
& \quad T \rightarrow B \text{ odd} \rightarrow \text{Answer} = (A \cdot B^{-1}) \mod (p-1) \\
& \quad T \rightarrow \text{Answer} = \text{SDVD}(A/2, B/2) \\
A > 0 & \quad B < 0 \quad \text{Answer} = (2A \cdot B^{-1}) \mod (p-1) \\
A < 0 & \quad B > 0 \quad A \text{ odd} \rightarrow \text{Answer} = (A \cdot B^{-1}) \mod (p-1) \\
& \quad T \rightarrow \text{Answer} = (|A/2| \cdot B^{-1}) \mod (p-1) \\
A < 0 & \quad B < 0 \quad \text{Answer} = (|A| \cdot |B|^{-1}) \mod (p-1)
\end{align*} \]

Where \( B^{-1} \) is computed as \( \text{SPWR}(B, 8n+3) \), since \( 8n+3 \) is \( \pi(p-1) \), the number of integers less than \( p \) and relatively prime to \( p-1 \).

FDVD is computed as \( \text{FPFD}(A, \text{FPWR}[B, p-2]) \). Frequently, one wants to perform an operation with either the S or the F functions depending on the SIMPLELEVEL. For this reason, the A functions have been defined. These functions are APLS, APFD, ADVD, APWR, and ANEG. Their definitions are similar. For example:

\[ \begin{align*}
\text{APLS}[A,B] \equiv (\text{SIMPLELEVEL} = 0) & \rightarrow \text{FPLS}[A,B] \\
T \rightarrow \text{SPLS}[A,B]
\end{align*} \]


This completes the explanation of the algebraic operations used in the hash code scheme. There are several functions which use these operations to compute the hash code number of an arbitrary expression.

The hash code number of any expression, \( X \), is computed by the function ECODER[X, SIMPLELEVEL]. ECODER goes down the expression tree until
it reaches the individual variables and constants. It finds the hash code numbers for these and then computes the hash code number of each subexpression from the hash code numbers of its arguments. In going down the expression tree, the SIMPLELEVEL is changed whenever the exponent of a WR is entered or whenever the main connective of a subexpression is flagged with FCODE, indicating that it is a transcendental function. The code numbers of the individual atoms and numbers are computed by the function GETC[N]. GETC gives a floating point number to MKCD. In the case of a positive integer, it is only necessary to take APLS[integer, 0], or ANEG[-integer] for a negative integer. If N is a function, the answer is ADVD of GETC applied to the numerator and denominator. The remaining possible argument for GETC is a literal atom. GETC looks on the property list of the atom under the indicator CODE. If the atom has been coded before, or if it is *E, *I, or PI, its code number will be found. Otherwise, a random number is generated by RANNO and taken modulo p-1 by SPLS[random number, 0]. This code number is put on the property list of the atom and also returned as the answer. The floating point number hashing function, MKCD, hashes the number in three steps. First, the absolute value of the number is taken. Then the characteristic and fraction are separated, and the fraction is either multiplied or divided by $2^{\text{characteristic}}$, depending on the sign of the characteristic. Finally, if the floating point number was negative, ANEG of the answer is taken.

Returning to ECODE, the method of computing the code number of an expression from the code numbers of its arguments, depends on the operator of the expression. As stated above, if the operator symbol has
the property FCODE, this represents a function of one variable; the
functions in this class which are presently coded are: PSIN, FCOS,
PTAN, FSEC, FCSC, FCCT, PSINH, FCOSH, PTANH, FCOTH, FSECH, and FCSCH.
These functions will be discussed individually later. They all compute
a hash number from the hash number of their argument. If the operator
symbol has the property NCODE, this property flags a function of one
argument which is a list of the hash code numbers of the arguments of
the operator. When a new operator is introduced, a function for computing
its hash code can be invented and added to the system with NCODE. ECODE
applies this function to a list of the argument hash numbers. ECODE
calls APLS, APRD, and APWR for the operators PLS, PRD, and PWR
respectively. In the case of PLS and PRD, ECODE uses ECODEL to apply
the binary operators APLS and APRD to the indefinite number of arguments
of PLS and PRD. All other operators are handled by a call to FGPN.
This function produces a result which depends on the hash numbers of
the arguments and their order, but which has no invariant properties.
Taking the hash number of the operator name as the initial accumulated
answer, the accumulated answer is added in F to the hash number of the
next argument, and this sum is raised in F to the arbitrary power
125786059 to form the new accumulated answer. This process is continued
until all of the arguments have been used. This completes the discussion
of ECODE. In conclusion, the formulas for the transcendental functions
are:

\[
\begin{align*}
\text{PSIN} & = \frac{1}{2i} \left( e^{ix} - e^{-ix} \right) \\
\end{align*}
\]
\[
\text{FCOS} \quad \frac{1}{2} (e^{ix} + e^{-ix})
\]

\[
\text{FTAN} \quad u = e^{2ix} \quad \text{Answer} = 1 \quad \frac{1}{1+u}
\]

\[
\text{FSIN} \quad 1/\text{FCOS}
\]

\[
\text{FCSC} \quad 1/\text{FSIN}
\]

\[
\text{FCOT} \quad 1/\text{FTAN}
\]

\[
\text{FSINH} \quad -1 \cdot \text{FSIN}(ix)
\]

\[
\text{FCOSH} \quad \text{FCOS}(ix)
\]

\[
\text{FTANH} \quad \text{FCOSH}
\]

\[
\text{FCOTH} \quad 1/\text{FTANH}
\]

\[
\text{FSCH} \quad 1/\text{FCOSH}
\]

\[
\text{FCSCH} \quad 1/\text{FSINH}
\]

The operations in transcendental functions are performed in S arithmetic.

**Functions for Algebraic Operations**

The functions which perform algebraic operations can be divided into seven classes. Typically, the functions within a given class are defined for the same arguments, such as integers, rational numbers, or general expressions. Each class may also contain a few helping functions.

A. **Integer Functions**

Two functions accept integer arguments:

\[
\text{ABS}[N] = |N|
\]

\[
\text{FTEXP}[N] = N!
\]

B. **Rational Number Functions**

The arguments for these functions can be either rational numbers, represented by (FRT M N NIL), integers, or floating point
numbers. Since the functions for addition, multiplication, and division take either fractions or LISP numbers as input, they first send their arguments to DENUM for conversion to a standard form:

\[
\text{DENUM}[A \ B] = (\text{numerator of } A, \text{denominator of } A, \text{numerator of } B, \text{denominator of } B)
\]

Using this list, the three functions form the following quotients.

\[
\begin{align*}
\text{PLSNUM:} &\quad \frac{\text{NUM1} \cdot \text{DENOM2} + \text{DENOM1} \cdot \text{NUM2}}{\text{DENOM1} \cdot \text{DENOM2}} \\
\text{PRDNUM:} &\quad \frac{\text{NUM1} \cdot \text{NUM2}}{\text{DENOM1} \cdot \text{DENOM2}} \\
\text{DVDNUM:} &\quad \frac{\text{NUM1} \cdot \text{DENOM2}}{\text{DENOM1} \cdot \text{NUM2}}
\end{align*}
\]

These results are then given to REDNUM to be reduced to lowest terms. If one of the original arguments was a floating point number, then of course, the answer will be a floating point number. If the denominator is zero, then REDNUM returns \(\text{INF}\), representing \(\infty\). Otherwise, REDNUM reduces the fraction to lowest terms by finding the g.c.d. of the numerator and denominator with GCD. The final member of this class is the exponentiation function PARNUM. The exponent must be an integer. The special case \(0^{-N}\) results in \(\text{INF}\). Any other negative exponent results in the base being inverted with DVDNUM, and the exponent made positive. The remaining possibilities have a non-zero base with a positive exponent. If the base is negative, the sign of the answer is positive or negative depending on whether the exponent is even or odd.
C. Functions for Arbitrary Expressions

These functions combine expressions under the elementary operations. The expressions are not simplified, but simple redundancies are eliminated. EPLS[X, Y] and EPRD[X, Y] both call EFUN with appropriate functional arguments. EFUN performs the same operations for the additive and multiplicative group. EPRD must make the check for a zero argument, since a·0=0 has no counterpart in the additive group. For INF·0 EPRD returns NIL and prints INDETERMINATE. EFUN[X, Y, W, NUMFUN, Z, NUMP] has six arguments which have the following uses and values:

<table>
<thead>
<tr>
<th>Name</th>
<th>Description</th>
<th>Value for EPLS</th>
<th>Value for EPRD</th>
</tr>
</thead>
<tbody>
<tr>
<td>X</td>
<td>first expression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Y</td>
<td>second expression</td>
<td></td>
<td></td>
</tr>
<tr>
<td>W</td>
<td>operation name</td>
<td>PLS</td>
<td>PRD</td>
</tr>
<tr>
<td>NUMFUN</td>
<td>rational number function</td>
<td>PLSUM</td>
<td>PRDSUM</td>
</tr>
<tr>
<td>Z</td>
<td>group identity</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>NUMP</td>
<td>group identity predicate</td>
<td>EZEROP</td>
<td>EONEP</td>
</tr>
</tbody>
</table>

If either argument is undefined, EFUN just returns (W X Y NIL). If either argument is the identity, EFUN returns the other argument. If either argument is infinity, the answer is infinity. If none of these conditions hold, then each of the two expressions is examined by EFUNL. EFUNL constructs a number, U, and a term list, V. U is initially the identity and V is empty. If X is a number, it is combined with U. If X has the operator W, then its terms are placed on V, except that if the first term is a number, it is combined with U. Otherwise, X itself is placed on V. Y is treated in the same manner. EFUN then looks at this result. There are three possible cases. If the term list is empty,
EPUN returns the number. If the number is the identity and there is only one term on the term list, this term is the result. Otherwise, the number is placed on the term list if it is not the identity, and the operator name W is added to form the answer.

The next function in this group is EPWR[X, Y]. There are quite a number of possible outcomes. If the exponent is one, the answer is the base. If either argument is undefined, the answer is (PWR X Y NIL). A zero exponent yields an answer of 1 unless the base is infinity. If the base is a fraction, then it is inverted if necessary, so that the numerator will be larger than the denominator. When the base is inverted, the sign of the exponent is changed. This canonical ordering of exponentiated fractions allows the current simplification program to reduce \((a^{-\frac{1}{2}})^{\frac{3}{2}}\) to 2. If the base is a number and the exponent is an integer, then the result is computed by PWRNUM. If this is not so, but the base is a positive LISP number and the exponent is also a LISP number, then the result is computed by EXPT. Otherwise, the result is (PWR X Y NIL).

EVDU[X, Y] is defined as EPRD[X, EPWR[X, -1]] and ENEG[X] is defined as EPRD[X, -1]. EEQS[X, Y] is (EQN X Y NIL) and ESET[X, Y] is (ESET X, Y NIL). EMIN[X] returns the smallest number on argument list X. It compares the arguments with the function LSSP described in the section on predicates.

ENLOG[X], the natural logarithm of X, can have several outcomes. If X is one, the answer is zero. If X is zero, the answer is -\(\infty\). If X is \(\infty\), the answer is \(\infty\); and if X is \(e^A\), the answer is A. Otherwise, the answer is (NLLOG X NIL).
There is one more function in this group; \( EDRV[x_1, n_1, \ldots, x_n, n_n, \ldots, x, n_n], Y \) expresses the \( n_i \)th derivative of \( Y \) with respect to \( x_i \) for each \( i \). If \( Y \) does not have main operator \( DRV \), then the answer is \( (DRV \ldots x, n_1, \ldots, Y, NIL) \). Otherwise, \( EDRV \) gives the two arguments to \( EDRV1 \). \( EDRV1 \) gives \( Y, x_1, \) and \( n_1 \) to \( EDRV2 \), for each \( i \). \( EDRV2 \) goes through the list of differentiation variables on \( Y \) and adds \( x_i \).

D. Predicates

There are eight predicates.

- \( EEMUMBERP[X] \) is \( X \) a LISP number, or \( (FRT N M NIL) \)?
- \( ENUMBERP[X] \) is \( X \) a LISP number, \( INF \), or \( (FRT N M NIL) \)?
- \( EZERO[P[X] \) is \( X \) a zero LISP number?
- \( EONEP[X] \) is \( X \) a LISP number equal to 1.0 or \( (FRT N N NIL) \)?
- \( EMINUSP[X] \) is \( X \) a negative LISP number or \( (FRT -N M NIL) \)?
- \( EODDP[X] \) is \( X \) an odd positive LISP number?
- \( GRTP[X, Y] \) \( Y = Y \rightarrow NIL \)
  \( X = -\rightarrow T \)
  \( Y = -\rightarrow NIL \)

Otherwise, there is an error unless \( X \) and \( Y \) are LISP numbers, these are compared with \( GREATERP \).

\( LSSP[X, Y] = GRTP[Y, X] \)
E. Functions Which Combine Expressions with Hash Code Numbers
On the Front

There are four functions which combine lists of the form
(hash number • expression) under addition, multiplication, and
exponentiation. These functions are used by the simplification program.
DPLS, DPRD, and DPWR all call DPUN with appropriate functional arguments.
DPWR first checks to see if the hash of the base is zero. In this case
the base and exponent are combined with EPWR and the result is prefaced
with its hash code number. This special check is to avoid zero raised
to a negative power.

DPUN[X Y AFUN EFUN] is a function of the two arguments and
the proper A and E functions for the given operation. DPUN combines
the hash code numbers with the A function and the expressions with the
E function. U is the new hash number and V the new expression. If
the expression is NIL, then it is undefined and the hash of NIL is
computed instead of U. The hash scheme breaks down for undefined
operations. Next a check is made to see if U is 1, 0, or -1. In this
case the expression is taken as equivalent to the hash number. (Note:)
A hash number of -1 means that APLS[1, N] = 0). If none of these cases
apply, the result is (U-V).

F. Functions Which Make It Possible to Evaluate the Expressions
As LISP Functions

The operators have been given functional definitions so that
they can be evaluated.
PLS[X, Y] = EPLS[X, Y]
PRD[X, Y] = EPRD[X, Y]
NEG[X] = ENEG[X]
Dvd[X, Y] = EDVD[X]
EQM[X, Y] = EEQM[X]

NAM[X, Y] is an FEXPR. It evaluates its arguments and from the resulting list, (X₁ ... Xₙ), forms (NAM X₁ ... Xₙ NIL).

FUN[X, Y] is also an FEXPR. The first argument is the name of an arbitrary function. FUN evaluates its arguments and adds NIL to the end of the resulting list.

DRV[X, Y] is also an FEXPR. It evaluates its arguments and gives the resulting list to DRV1 which carries out the indicated differentiation. DRV1 gives each indicated variable of differentiation to DRV2, which calls DIFF in order to differentiate Y the required number of times.

G. Remaining Function

There is one more function which is needed to implement the convention that a set of one element and that element can be used interchangeably in commands. SETTER[X] returns X if its main connective is ESET, and (ESET X NIL) otherwise.
Functions for Input-Output

There are five operations performed to facilitate input-output.

1. Transformation of expressions from internal form before they are printed on the typewriter.

2. Transformation of typed input expressions into internal form.

3. Retrieval of the \( N \)th subexpression of an expression which has been pointed at with the light pen.

4. Retrieval of an expression with the \( N \)th subexpression removed.

5. Conversion of an expression retrieved by operation 3. or 4. into internal form.

Conversion of expressions for display is the subject of Chapter IX.

Expressions can be printed in a modified internal form with EPRINT. EPRINT(X) reads X FORM from the disk and removes the property lists from all subexpressions with RPLST. It gives the resulting expression to EPRINT1. EPRINT1 goes through the expression and gives all PLS, PRD, PWR, and EQN subexpressions which have more than one argument, to EPRINT2. EPRINT converts the expressions to internal operator form. For example, (PLS A B C) becomes (A PLS B PLS C).

If an input command contains a quoted expression such as '(A + B)', this will be parsed into (EQUOTE (PLS A B)). EQUOTE is an FEXPR which returns its argument unevaluated, but transformed by EQUOTEL and EQUOTE2. EQUOTEL puts a null property list on every subexpression. It also transforms (NEG X) into (PRD -1 X NIL) and (DWD A B) into (PRD A (PWR B -1 NIL) NIL). EQUOTE2 removes redundant levels of PLS and
PRD expressions. For example,

(PLS A (PLS C D NIL) NIL) becomes (PLS A B C NIL).

GETSUB\[X, N]\] returns the \(N^{th}\) subexpression of \(X\); the subexpressions are numbered in the order they would normally be evaluated by LISP.

The subexpression is located by GETSUB1. GETSUB1 returns NIL if the count does not reach \(N\), so it must list the answer to distinguish the answer NIL from failure. Therefore, GETSUB takes CAR of the result of GETSUB1. \(N\) is a free variable for GETSUB1; GETSUB1 decrements \(N\) at each new subexpression. When \(N\) is zero, GETSUB1 returns a list of that subexpression as the answer. If \(N\) is not yet zero and GETSUB1 reaches an atom or subexpression with operator ATOM, it returns NIL to indicate failure. If an expression has several arguments, GETSUB2 applies GETSUB1 to each argument in sequence, until \(N\) becomes zero or the arguments are exhausted. GETSUB is used to find a subexpression which is being pointed at by the light pen as described in Chapter IX. It may also be necessary to use the expression which contains this subexpression, perhaps replacing the subexpression with another one. Therefore, when GETSUB1 finds the \(N^{th}\) subexpression, the list (SUBMARK M NIL) is put on the property list of the subexpression by altering the list structure with REPLACA. More than one subexpression may be marked by successive calls to GETSUB. The SUBMARK's will have increasing integers \(M\) until the constant INPNUM is reset to 1. This mechanism is used by the command REPLACE which is described in the discussion of commands.

The function REPLACESUB will replace each subexpression which has a SUBMARK structure on its property list, by that structure.
When a subexpression of expression E has been indicated with the light pen, then the structure E PFORM, which is produced by the picture compiler, is given to GETSUB. GETSUB finds the subexpression as described above. The PFORM subexpression is then transformed into internal form by APOFF. APOFF makes the following transforms on all applicable subexpressions:

1. \((\text{DIV} A B \text{NIL}) \rightarrow (\text{PRD} A (\text{PWR} B -1 \text{NIL}) \text{NIL})\)
2. \((\text{NEG} A \text{NIL}) \rightarrow (\text{PRD} -1 A \text{NIL})\)
3. \((\text{FUNCTION} F B C \text{NIL}) \rightarrow (F B C \text{NIL})\)
4. \((\text{FUNCTION} (\text{NAM} I F \text{NIL}) B C \text{NIL}) \rightarrow (\text{NAM} I (F B C \text{NIL}) \text{NIL})\)
5. \((\text{DRV} 1 1 2 3 (F X Y \text{NIL}) \text{NIL}) \rightarrow (\text{DRV} X 1 Y 3 (F X Y \text{NIL}) \text{NIL})\)
6. \((\text{PAREN} A \text{NIL}) \rightarrow A\)
7. \((\text{ATOM} A \text{NIL}) \rightarrow A\)

APOFF1 and APOFF 2 perform transformation 5. APOFF3 performs transformation 4.

System Functions

There are fourteen specialized functions which communicate with the disk and the teletype. There are five types of disk files:

1. Files containing LISP functions which comprise a given mathematical transformation. These files have second name LISP.
2. Files containing LISP statements which delete from core the functions in a given transformation file. These have second name ERASE.
3. Files containing a single mathematical expression. These have second name FORM.

4. Files containing an intermediate form of a single displayed expression. These files are generated by the picture compiler to facilitate light pen reference to subexpressions. They have second name PFORM.

5. Files which contain scope instructions to display an expression. These files have second name DFORM.

There are functions which read, write and delete these files.

There is not enough space in core memory for all of the LISP functions which execute the commands. Therefore, the functions for a given operation, such as differentiation, are brought into core only while that step is being performed. In the case of differentiation, for example, the top level function DIFF(Y, X) is temporarily defined as a call to GETFILE[DIFF, (Y X)]. GETFILE reads file DIFF LISP, executes DIFF(Y, X), and then reads file DIFF ERASE which contains LISP commands to remove the differentiation routines from core and redefine DIFF as a call to GETFILE.

All mathematical expressions which have been defined in the system are written on the disk with second name FORM by a call to EASSIGN[Name, Expression]. The input command E←X becomes EASSIGN[E, X]. EASSIGN also deletes any old files with name E DFORM or E PFORM. An expression can be deleted by a call to EDELETE[E] which deletes E FORM, as well as E PFORM and E DFORM. The FEXPR FORM[E] has as value the expression in file E FORM. Unquoted names in input commands are given to FORM.
When a light pen reference is made, the function PFORDER[Name, N] is called. This function reads file Name PFORM and uses GETSUB to find the N\textsuperscript{th} subexpression. If an operation such as REPLACE is called, then the expression Name PFORM must be saved, so that a new subexpression can be substituted for the old one. To do this the call to PFORDER is replaced by a call to GETINSUB and the required PFORM is read into core and saved on the LISP constant INPFORM. GETINSUB then uses the expression on INPFORM instead of reading from the disk. When GETSUB locates the N\textsuperscript{th} subexpression, it alters the list structure of INPFORM to mark the place.

EDISPLAY[E] will type the display instructions for E. These instructions are preceded by a dollar sign, so that the PDP-6 will interpret them as a command, instead of typing them on the teletype. EDISPLAY first checks to see if there is a file E PFORM, indicating that the picture compiler has already been used to compute the displayed form of expression E. In this case EDISPLAY chains to the CTSS system command PRINT, printing file E DFORM. Otherwise, it is necessary to chain first to the file LISPP SAVED, which contains the picture compiler. This file reads E FORM and creates the corresponding PFORM and DFORM files. The DFORM file originally has name LISP LSPOUT, but this is changed to E DFORM with the CTSS RENAME command.

When a command is typed at the PDP-6, the string of symbols is given to the LISP function COMMAND. This function calls APARSE to parse the symbols into a LISP expression. It then evaluates this expression for its effect and returns an answer of NIL to indicate that the evaluation is complete.
A method of questioning the user has been provided with the function

\texttt{QUESTION(X, Y)}. \(X\) is a list of expressions and \(Y\) is a question to be
typed out, which may contain some of the expressions in list \(X\).
\texttt{QUESTION} first types the word \texttt{QUESTION}. Then, it generates a name for
each expression in list \(X\) and writes the expression on the disk with
this name. It also displays each expression. Next, it replaces the
expressions in question \(Y\) with their names. The question is then
typed. The user can make one of four responses. If he types \texttt{GIVEUP},
\texttt{QUESTION} calls \texttt{ERROR}. If he types \texttt{ANSWERR X}, \texttt{LISP} expression \(X\) is
returned as the answer. He may also type \texttt{ANSWER X Y}, in this case
\(X Y\) is taken as a pair for \texttt{EVALQUOTE} and the result is returned as the
answer. A special case check is made for \texttt{ANSWER COMMAND (X)}. In this
case string \(X\) is parsed with \texttt{APARSE} and evaluated to produce the
answer. If the user does not type \texttt{GIVEUP}, \texttt{ANSWER}, or \texttt{ANSWERR}, then
his response is taken as a pair for \texttt{EVALQUOTE} and the result is typed
out. \texttt{QUESTION} then types, "THIS QUESTION TO YOU IS PENDING", repeats
the question and awaits a further user response.

There are three general purpose file reading functions.

\texttt{READF(X Y)} opens file \(X Y\), reads one list, then closes the file.

\texttt{FILESEND(X Y)} reads one list from file \(X Y\) and prints the members of
this list one at a time. \texttt{WHOLEFILEREAD(X Y)} returns a list of all the
lists in file \(X Y\).
General Purpose Functions

There are a number of functions which have proven useful in transforming the mathematical expression representations used in this thesis.

APND[X Y] forms an argument list from the argument lists X and Y. For example, APND[(A B NIL) (C D E NIL)] is (A B C D E NIL).

ARGLIST[X] If the next to last element of list X is an arbitrary function with arguments, or a subscripted arbitrary function with arguments, then the list of arguments is returned. Otherwise, NIL is returned.

DELETE[X Y] removes the first occurrence of Y from list X.

DEPEND[Y X] is true if expression Y depends on variable X.

Either Y is an atom and X is equal to Y, or X is a member of the property list of Y. If X is a subscripted variable or function, then it will appear on the property list of Y in a special format, without arguments and with a null property list. In this case, X must be transformed into this format before it can be compared against the property list of Y.

FLOAT[X] X must be a LISP number or the form (FRT N M NIL). If X is a LISP number, it is floated by adding it to floating zero. Otherwise, the numerator and denominator are floated in this manner and then the quotient is taken.

LAST[X] returns the last member of list X.

MAP[X, FN] is like MAPLIST except that it applies FN to CAR of the list X, and it does not apply FN to the last member of list X.
MATCH[EXP, PATT, CONST] is a very simple matching function. The expression EXP is matched against the pattern PATT in a two step process. If a match occurs, then the expression CONST is evaluated against a special ALIST, MALIST which is used during the matching step. In the matching step, MATCH2 compares EXP and PATT element by element, all elements must be the same until the ends of both expressions are reached or until the atom DOLLAR appears as an element on PATT. If this match succeeds, then the second part of the match is attempted. If PATT has a function on its property list, then the value of this function applied to EXP must be T. If both of these conditions are met, the match succeeds.

MKCLEANP[X] sets to NIL all subexpression property lists of expression X.

MKFLS[X] checks argument list X for three conditions:

(NIL) → 0

(A NIL) → A

(A ··· B NIL) → (PLS A ··· B NIL).

MKFRD[X] checks argument list X for three conditions:

(NIL) → 0

(A NIL) → A

(A ··· B NIL) → (PLS A ··· B NIL).

NILON[X] adds the element NIL to the right end of list X.

NTOL[X] returns the next to last element of list X.

NTOLSUBST[X, Y] substitutes Y for the next to last element of X.

NUMBER[X] is the same as LENGTH of X.
PUT[X Y Z] puts X on atom Y with indicator Z.

RANNOO[] returns a random integer which is formed by multiplying the previous random integer by 311. The previous random integer is kept as the value of the constant RANDOMNUMBER. This random number generator is based on a note by Martin Greenberger.

REMPROPS[LIST PROP] removes property PROP from each element of list LIST.

RPLST[X] removes the property list from each subexpression of expression X.

SIMP32[FN Y] evaluates (FN Y).

STRIKE[X Y] removes from list X any element for which the predicate Y is true.

SUBSTA[CONST PATT EXP] applies MATCH to each subexpression of EXP. SUBSTA starts at the ends of the branches of the expression tree and works toward the top.

TWOOFF[X] removes the last two elements from list X.

UNIONLIST[X Y] adds to list Y any members of list X which are not members of Y.

Transformations

The implementation of each of the transformations in Chapter II will now be described. Conclusions about these transformations have already been presented in the chapter introduction. The most informative are SIMPLIFY and TRUNCATE. The use of light pen references is explained in the description of REPLACE.
ALLSUMEXPAND

ALLSUMEXPAND[x] uses SUBST to apply SUMEXPAND to each SUM

subexpression in x. Since SUBST comes up from the bottom,

\[ \sum_{i=0}^{l} \sum_{i=0}^{l} x^i \]

would be expanded to

\[ \sum_{j=0}^{n} x^j + \sum_{j=0}^{1} x^j. \]
BRINGOVER

BRINGOVER(NAME, TERM) brings term TERM to the other side of the equation with name NAME. The term is indicated with the light pen.

BRINGOVER is not very clever; the term must be one or more factors in a sum or product. In order to compute the answer, BRINGOVER needs both the term and the equation with a hole where the term was removed.

Normally, the function PFORREAD which evaluates a light pen reference returns the subexpression referenced; the rest of the expression is lost. To save the rest of the expression BRINGOVER has been defined as an PEXPR. It has two arguments, the first will evaluate to the name of the equation and the second to the term. First, BRINGOVER finds the equation name and reads the PFORM with that name from the disk. The constant INFORM is assigned the PFORM as value. Now, before the second argument is evaluated the function name PFORREAD is replaced with GETINSUB. Instead of reading the PFORM from the disk as PFORREAD would do, GETINSUB uses INFORM. The hole where the light pen reference was found is marked by altering the list structure of the subexpression property list. If the term is constructed from several light pen references, each will be marked. Having evaluated its second argument to obtain the term, BRINGOVER next calls REPLACESUB to replace the marked subexpressions in INFORM with the lists (SUBMARK N NIL). N numbers the marked subexpressions. INFORM is converted to standard internal form by APOFF and SUBMARK expressions with a number other than 1 are deleted from it by BRING01.
BRINGOVER now has the required term and equation with the hole marked by (SUBMARK 1 NIL). It first checks to make sure that it was really given an equation. If not, it just replaces the term. Otherwise, it checks for each of the following legal possibilities.

Let $X$ be the term to be moved.

\[
X = Y \rightarrow 1 = XY^{-1}
\]
\[
Y = X \rightarrow X^{-1}Y = 1
\]
\[
AX = Y \rightarrow A = YX^{-1}
\]
\[
A + X = Y \rightarrow A = Y - X
\]
\[
Y = AX \rightarrow X^{-1}Y = A
\]
\[
Y = A + X \rightarrow Y - X = A
\]

In each of these cases, it performs the required operation to produce the answer. Otherwise, the answer is produced by replacing the term in the hole.
COLLECT

COLLECT[EXP, SET] collects terms in the top level of expression EXP in powers of the variables in set SET. If the expression is an equation, it will collect terms on each side. When possible, COLLECT will combine terms involving summation. For example,

\[ ax^2 + \sum_{i=1}^{10} bx^i + d \cdot \sum_{j=5}^{15} cx^{j+2} \]

will be collected on powers of x to yield:

\[ bx + (\sigma + b)x^2 + \sum_{i=3}^{6} bx^i + \sum_{i=7}^{10} (b + d \cdot c)x^i + \sum_{i=11}^{17} d \cdot cx^i \]

This transformation is straightforward enough, but since there may be several variables on which to collect terms, and several factors within each term, as well as many intersecting ranges of summation, an inefficient algorithm could be very slow. Therefore, the data is rearranged into a special form which makes it easy to sequence through the collection variables, term factors, and summation ranges. Thus, the LISP functions are specialized to this routine. The complex data structures made the routine more difficult to debug.

The flow of control through the functions which comprise this routine is shown in Figure 1. The key data structure is shown in Figure 2.
Figure 1

Flow of Program Control in COLLECT

- COLLECT
  - CLCT3
    - Simplify input expression.
  - CLCT31
    - If expression is an equation, do one side at a time.
  - CLCT32
    - Executive function.
  - CLCT5
    - Each term is a product of factors.
  - CLCT4
    - Convert each factor to special format.
  - CLCT8
    - Transform the expression into the special format.
  - CLCT7
    - Collect the terms on each of the variables.
  - CLCT9
    - Convert the answer back to standard internal form.
  - CLCT10
    - Collect terms on one variable.
The top level function \texttt{COLLECT(\texttt{EXP, SET})} simply calls \texttt{CLCT3(\texttt{EXP, SET})}. \texttt{CLCT3} changes \texttt{SET} into a list of variables by removing the operator and property list; if a set of one element has been represented by the element itself, \texttt{CLCT3} just lists the element. \texttt{CLCT3} also simplifies \texttt{EXP}. \texttt{CLCT3} gives \texttt{EXP} and the list of variables \texttt{Y} to \texttt{CLCT31}. \texttt{CLCT31} checks \texttt{EXP} to see if its main connective is \texttt{EQN}. If it is, then the two sides are collected individually, otherwise, the entire expression is collected. The part to be collected is converted to a \texttt{PLS} term list and given to \texttt{CLCT32}. \texttt{CLCT32} issues the function calls necessary to complete the collection of terms. First, the \texttt{PLS} term list \texttt{X} is converted by \texttt{CLCT8}. The conversion is shown by the example in Figure 2. \texttt{CLCT8} takes the \texttt{PLS} terms one at a time; each of the \texttt{PLS} terms is converted to a \texttt{PRD} term list by \texttt{CLCT5} and then given to \texttt{CLCT4}. The \texttt{PLS} term itself is added to the front of the answer produced by \texttt{CLCT4} to form the list structure shown in Figure 1. To produce its result, \texttt{CLCT4} checks each of the \texttt{PRD} terms to see if it has \texttt{FWR} or \texttt{SUM} for a main connective. If the main connective is \texttt{FWR}, a term of the form:

\begin{quote}
(hash of base, exponent, remaining \texttt{PRD} terms, preceding \texttt{PRD} terms)
\end{quote}

is constructed. If the main connective is \texttt{SUM}, several slightly more complex terms may be constructed. First, \texttt{\#1\#} is substituted for the running variable in the body of the \texttt{SUM} and it is converted to a \texttt{PRD} term list. The \texttt{SUM} body is then examined by \texttt{CLCT41}. If a term of the \texttt{SUM} body is a \texttt{FWR} and the base does not depend on \texttt{\#1\#} but the exponent does, then the exponent is evaluated by \texttt{CLCT411}. \texttt{CLCT411} returns \texttt{NIL} unless the exponent is of the form \texttt{\#1\#+N} or \texttt{\#N+N}, \texttt{N} an integer, in which case it returns \texttt{N}. If \texttt{CLCT411} returns \texttt{NIL}, then the factor can-
not be collected and the machine asks the user if it should continue, ignoring this factor. If CLCT411 returns an integer, the terms in the SUM body PRD list before and after the term in question are concatenated together and made into a PRD. If the integer returned is not 0, that is we have \((\text{base} \times I^a)\), then a change of variables is made in the SUM so that we have \((\text{base} \times I^a)\) by substituting \(I^a-N\) for \(I^a\). A term of the form:

\[
(\text{hash of base, } \text{(lower limit } + \ N\text{, upper limit } + \ N\text{), remaining PRD terms, preceding PRD terms including residue SUM body})
\]

is then generated. If CLCT41 finds a term in the SUM body PRD list which does not depend on the running variable \(I^a\), then if the term is a FWR a term of the form:

\[
(\text{hash of base, exponent, remaining PRD terms, preceding PRD terms including SUM with term in question factored out and running index changed to } \times J^a)
\]

is generated. If the term is not a FWR, a term of the form:

\[
(\text{hash of term, 1, remaining PRD terms, preceding PRD terms including SUM with term in question factored out and running index changed to } \times J^a)
\]

is generated. That is, the term is factored out of the SUM. If a term cannot be generated, then CLCT41 moves on to the next term in the SUM body PRD list. Returning to CLCT4, if it does not find the main connective to be FWR or SUM, then it forms a term:

\[
(\text{hash of term, 1, remaining terms in PRD, preceding terms in PRD})
\]

When all terms have been examined, CLCT4 returns to CLCT8.
After the expression is transformed by CLCT8, the list of variables to be collected on is transformed by CLCT5. CLCT5 generates a structure of the form: \(((\text{hash of variable 1}) (\text{hash of variable 2}) \ldots (\text{hash of variable n}) 0),\) which will be called the collection list.

The transformed expression and list of variables are given to CLCT7. The purpose of CLCT7 is just to give the terms of the top level PLS one at a time to CLCT10, which puts them on the collection list in the appropriate place. The PRD list of each of the terms in the top level PLS has been converted by CLCT4 into a list of terms each of which begins with a hash number of a variable or expression and contains the information needed to rewrite the PRD with powers of that variable or expression factored out. CLCT10 goes completely through these PRD terms for each member of the collection list. If none of the PRD terms has a base which matches one of the collection variables, then the PLS term which these PRD terms make up is put at the end of the collection list. If more than one of the terms match, then the user is asked if it is OK to collect on the first variable or expression which matched. A mechanism is provided which lets him answer this question once for all. When the first match is found, the matching PRD list form and collection list entry are given to CLCT11. CLCT11 removes the hash number from the front of the collection list entry and then puts it back after the entry has had the new term added to it by CLCT11.

After the hash number has been removed from the front, the collection list is a list of the form

\[(X_1 Y_1 X_2 Y_2 \ldots X_n Y_n)\] where

\(X_i\) is either an integer exponent or a list \((X_1, X_2)\) giving a range of
integer exponents, and Yi is the corresponding residue after factoring. Since part of a range Xi may match part of the range of the PRD list term, the remaining pieces must then be matched. CLCT1111 tries all the current pieces against the current range Xi, adding the residue to Yi for each piece that matches and splitting Xi if part of it matches. CLCT11 then proceeds to X(i+1) if any pieces of the range of the PRD list term remain. When Xn has been reached, all remaining pieces are put at the end by CLCT1111.

Once collection has been finished the collection list must be converted back to standard form. This is accomplished by CLCT9, which gives each term of the list to CLCT91 after removing the hash code number. CLCT91 generates a top level term for each pair Xi, Yi; a PWR if Xi is a single variable and a SUM if Xi is a range.
(A (PRD B C D NIL) (SUM J 0 10 (PRD E (PWR G *I* NIL) NIL) NIL) NIL)

is the term list to be converted to special format. There are three terms illustrating a single element, a product of factors, and a summation. The special format is:

<table>
<thead>
<tr>
<th>Term</th>
<th>Decomposition of term into factors</th>
</tr>
</thead>
<tbody>
<tr>
<td>First term</td>
<td></td>
</tr>
<tr>
<td>((A NIL)</td>
<td>(#A 1 (NIL) NIL))</td>
</tr>
<tr>
<td>Second term</td>
<td></td>
</tr>
<tr>
<td>((B C D NIL)</td>
<td>(#B 1 (C D NIL) (NIL))</td>
</tr>
<tr>
<td></td>
<td>(#C 1 (D NIL) (D NIL))</td>
</tr>
<tr>
<td></td>
<td>(#D 1 (NIL (C B NIL)))</td>
</tr>
<tr>
<td>Third term</td>
<td></td>
</tr>
<tr>
<td>((SUM <em>I</em> 0 10</td>
<td>(#D E (PWR G <em>I</em> NIL) NIL) NIL)</td>
</tr>
<tr>
<td></td>
<td>(#E 1 ((PWR G <em>I</em> NIL) NIL) NIL)</td>
</tr>
<tr>
<td></td>
<td>(#G (0 10) (NIL) (E NIL)) )</td>
</tr>
</tbody>
</table>

Hash code of factor base

Range of exponent

Following and preceding cofactors

Figure 2

Conversion of terms for more efficient collection.
DEPENDENCE

DEPENDENCE[EXP] returns a set of the variables and arbitrary
function names in EXP. It does this by making a set out of the
property list of EXP, after MKP has been used to find the dependencies.

If EXP is a number, DEPENDENCE returns NIL.
DELSUBST

DELSUBST[EXP, OLDEL, MEMDELFREE] uses SUBSTA to search for subexpressions representing the first derivative with respect to the OLDEL variable. For example, OLDEL might represent $\frac{d}{dx}$ and SUBSTA might find $\frac{d}{dx}(a+b)$. Then if MEMDELFREE represented, for example, $\frac{d}{dy} + e \frac{d}{dz}$, SUBSTA would use DELSUBST1 to make the substitution

$\frac{d}{dx}(a+b) \rightarrow e \frac{d}{dy}(a+b) + e \frac{d}{dx}(a+b)$. This is one of the few uses of the DEL operator, (DEL X NIL) ≡ $\frac{d}{dx}$. 
The operator DRV was described in the section on algebraic functions. There it was explained that DRV[x, 2, y, 1, z] would use DIFF to differentiate Z, first with respect to Y, and then twice with respect to X. The function DIFF[EXP, VAR] will now be described; DIFF differentiates EXP once with respect to VAR. The routine uses MKP to compute the dependence of subexpressions. Then, as DIFF comes down from the top, it can immediately replace a constant subexpression by zero. This method is certainly faster than simplifying a completely differentiated constant. However, it would be faster to work up from the bottom without the call to MKP. Since, MKP can be used in several routines, the method used here requires less program. It is very easy to specify derivatives with patterns and so they have been used here for the transcendental functions. The length of this routine is primarily the result of checks to reduce generation of redundant expressions.

DIFF differentiates an expression with respect to a given variable. The top level function, DIFF(y, x), gives the FORM y and variable x to DIFF2. DIFF2 makes one pass down to the ends of the branches of the expression tree, rewriting the expression as it goes. The transformation performed on each subexpression is a function of its main connective, the dependence of the subexpression and each of its arguments on the variable of expansion, and in some cases, the main connective of one of the arguments. However, the derivative of the expression depends on the derivatives of its arguments only in the sense that if some of the derivatives of the arguments are zero, the
result is collapsed to avoid an explosive growth of redundancies. The various cases will now be discussed one at a time, in the order in which the expression is tested.

The derivative of a constant is zero. At the end of a branch, the expression must be the variable of expansion with derivative 1. If the operator has property DIFF, this is a pattern into which single argument of the function is substituted for \(*Y\)*, and which is then multiplied by the derivative of the argument. For example, SIN has the pattern (COS \(*Y\) NIL), so that (PRD (COS X NIL) 1 NIL) would be formed as the derivative of (SIN X NIL). An arbitrary function initiates a call to DIFF8. DIFF8 forms a PLS term list. A term of the PLS is formed for each dependent argument of the function. DIFF8 forms the derivative of the function with respect to this dependent argument and multiplies this result by the derivative of the argument. The derivative of the body of a SUM or XST is taken, and all other arguments of these operators are left unchanged. PLS generates a call to DIFF3. DIFF3 forms a PLS term list of the derivatives of all the dependent arguments of the PLS. PRD initiates a call to DIFF4; DIFF4 also forms the term list of PLS. Each dependent term of the PRD is differentiated and multiplied by its cofactor to form a term of the PLS. PMR changes \(a^b\) into \(a^b(b^{-1} \cdot \log a)\cdot b\). The integral sign is removed from an indefinite integral with respect to the variable of differentiation.
(A convention here). Otherwise, $\int_{a}^{b} f(y)\,dy$ becomes

$$b \cdot (-f(a)) + b f(b) + \int_{a}^{b} f(x)\,dx$$

where the third term is replaced by zero if $f$ is independent of $x$. In the case of DRV, if the DRV body is an arbitrary function or a subscripted arbitrary function, then DIFFS is used to apply the chain rule as above. Otherwise, EDRV is used to add the new variable to the DRV list. Both arguments of EQN are differentiated, as are all the arguments of ESET. A DRV structure is set up around TNS. The second argument of all pairs in a CND is differentiated by DIFFS. For NAM, a check is first made to see if the expression is equal to the variable of differentiation. If this is not the case, the expression can be dependent only if it is a subscripted function. The expression is thus given to DIFFS for expansion by the chain rule. Any other permanent main connective does not have a derivative defined and will cause a call to ERROR.
DRVDO

DRVDO[EXP, VAR] carries out all indicated derivatives with respect to VAR in EXP. It uses DIFFDO1 to search out all DRV subexpressions. Each of these is given to DIFFDO2, the innermost ones first. DIFFDO2 examines each variable of differentiation on the DRV subexpression to see if any of them are VAR. If so, DIFFDO2 gives the DRV body, VAR, and the number of times, N, the body is to be differentiated with respect to VAR, to DIFFDO3. DIFFDO3 then issues N calls to DIFF. If there are variables of differentiation other than VAR, then these remain in a DRV subexpression.
DRYFACTOR

DRYFACTOR[EXP DRYFREE NFREE] uses SUBSTA to locate DRV subexpressions which indicate differentiation with respect to DRYFREE. It applies DRYFACTOR1 to each of these. DRYFACTOR1 searches down the list of variables of differentiation for DRYFREE. Then if the number of indicated derivatives with respect to DRYFREE is greater than NFREE, DRYFACTOR2 goes to the end of the argument list to form the new DRV subexpression. The resulting transformation is:

\[(DRV \ldots DRYFREE, (NFREE + M) \ldots Y NIL) \Rightarrow \]

\[(DRV \ldots DRYFREE, M \ldots (DRV DRYFREE NFREE Y NIL) NIL)\]
DRVZERO

DRVZERO(Exp, Var) uses SUBSTA to substitute zero for each DRV subexpression in Exp which indicates differentiation with respect to Var.
EVALUATE

EVALUATE[EXP, SET] searches expression EXP for variables and functions which can be evaluated using the equations in set SET. Only equations whose left hand side has one of the following forms are acceptable.

1. F
2. Fi, • • • ,j
3. F(X, • • • ,Y), matches F as well
4. Fi, • • • ,j (X, • • • ,Y), matches Fi, • • • ,j as well

The subscripts and function arguments can be either numbers or literals; if literals, they are taken as free variables to be bound. EVALUATE4 checks the equations and throws out unacceptable ones.

Next, EVALUATE1 searches EXP for instances of variables or functions which can be evaluated with the equations. Suppose, for example, that EXP was a+x and SET contained the equations x=y and y=b, then the search by EVALUATE1 would locate x and substitute y, yielding a+y. But further evaluation can now be done with the second equation. Therefore, whenever an evaluation is made, the variable MARKFREE is set so that EVALUATE1 will make another pass through EXP. EVALUATE1 gives EVALUATE2 EXP and one equation at a time. EVALUATE2 goes through EXP, calling EVALUATE5 to match each subexpression with the left side of the equation.

EVALUATE5 checks for each of the four cases given above. It uses EVALUATE6 to bind subscripts or arguments and substitute the values in the right side of the equation. Of course, for a match to occur the number of variables or subscripts must be equal.
EXCHANGE

EXCHANGE[EXP] will interchange the arguments of the main connective of EXP, if the main connective is binary.
EXPAND

EXPAND(EXP) goes through the expression EXP applying the distributive law. It expands sums raised to a positive integer power and distributes the derivative or integral of a sum. For historical reasons, the functions are named MULT. EXPAND calls MULT2(EXP). A special data form has been used, the resulting gain in efficiency is probably not worth the special programming required. If EXP is an equation, MULT2 handles each side separately; otherwise, the entire expression is handled. MULT2 gives the parts to be handled to MULT3, which returns a PLS term list at the top level. MULT2 makes this list of terms into a PLS and simplifies it. MULT3 recurs down the expression tree if the expression has PLS, PRD, DRV, ITG, or FWR for its main connective. All of the arguments of PLS, PRD, and FWR, and the main expressions of DRV and ITG are treated in turn by MULT3. Expressions with any other main connective are listed and returned as the value of MULT3. Coming back up the expression tree each function receives as arguments and returns for a value, a list of terms which are arguments of a PLS. Thus, at the end of each branch of the expression tree is a list with a single term. These lists are combined in a manner appropriate for the main connective which the function represents; MULT3 gives PLS to MULT4, PRD to MULT5, DRV and ITG to MULT6, and FWR to MULT7. In the case of DRV and ITG, MULT3 uses MULT31 to get a list of the arguments other than the main expression and passes this list along to MULT6 as well.
MULT4 receives a list whose members are lists of PLS terms, with a NIL at the right end. Since MULT4 represents PLS, it just appends these lists together to form its answer.

MULT5 also receives a list whose members are lists of PLS terms with a NIL at the right end. It has a second argument which is initially NIL and is used to build up the answer. This second argument is a list of PLS terms. MULT5 uses MULT51 to multiply each member of its first argument by the partial answer represented by the second argument. MULT51 multiplies each term of one of its arguments by all the terms of the other. The resulting list of terms forms the answer for MULT51.

To do this, MULT51 uses MULT511 to multiply each member of a list of terms by a single term. Each of these products is simplified to prevent expression growth.

MULT6 makes an integral or derivative subexpression out of each member of the list given it by replacing the member with the structure:

("first part of old list" "member" NIL).

MULT7 handles exponents; it distinguishes three cases.

If the list of terms which comprise the base argument has a single member, this member is raised to the exponent by EPWR. If the exponent is a positive integer, then the base is multiplied out, using MULT71. Otherwise, a single term is formed which represents the terms in the base as a PLS subexpression raised to the exponent. MULT71 multiplies the base by itself by using MULT51. It builds up the answer as its first argument. It decrements the exponent, which is the third argument, to count the required number of multiplications.
FACTOROUT

FACTOROUT[EXP, X, Y] is defined as an FEXPR so that it can have an optional third argument. The function will factor X from each term of expression EXP. If the third argument Y is present, Y will be substituted for the term factored out. For example, FACTOROUT[A + X, X, Y] yields \( Y \cdot \left( \frac{A}{X} + 1 \right) \), while FACTOROUT[A + X, X] yields \( X \cdot \left( \frac{A}{X} + 1 \right) \).

FACTOROUT supplies FACTOROUT1 with X for a third argument if Y is not given. FACTOROUT1 will do each side of an equation separately. To factor X out, FACTOROUT1 just uses MAPL to divide each term by X and simplify it.
GROUP

GROUP is used for the light pen selection of certain arguments of
a commutative operator which are to be handled as a group. For example,
one might select the a and b from the expression a+c+b+c. These
expressions might be replaced by d to yield d+c+c. The first
argument of GROUP will evaluate to the name, X, of an expression. The
second argument will evaluate to a set of terms or factors in a
subexpression of X with a commutative operator. To determine what
operator has been grouped, GROUP uses GROUP1 to search through the
entire expression for an expression which has been marked by GETSUB.
This marked expression will be an argument of the operator in
question. In order to do this, GROUP must save X PFORM for GROUP1.
It does this in the standard way. The first argument is evaluated to
find the name of the expression. Then, the constant INPFORM is set
equal to the corresponding PFORM. Next, the second argument is
evaluated to find the set of grouped terms; during the evaluation
GETSUB will mark INPFORM. INPFORM can then be given to GROUP1.

If INPFORM is not NIL when GROUP is entered, then some operation,
such as REPLACE, contains this call to GROUP in one of its arguments.
This function will need the marked INPFORM generated by GROUP.
Otherwise, INPFORM is set to NIL in order to regain free storage.
LEFT

LEFT[EXP] returns the first argument of EXP.
LIMIT

\( \text{LIMIT[EXP, X, N]} \) substitutes \( N \) for \( X \) in \( \text{EXP} \). However, if this substitution in a subexpression such as \( a \cdot b \cdot c \) should result in \( a \cdot 0 \cdot \infty \), then \( \text{LIMIT} \) will apply L'Hopital's Rule and evaluate \( a \cdot \frac{\text{LIMIT[b', X, N]}}{\text{LIMIT[1/c]', X, N]} \). \( \text{LIMIT} \) simplifies \( \text{EXP} \) and calls \( \text{LIMITI[EXP, X, N]} \). \( \text{LIMITI} \) goes down the expression tree until it reaches the end of a branch, where it substitutes \( N \) if the branch end is \( X \), or until it reaches a \( \text{PRD} \) subexpression. The factors of the \( \text{PRD} \) are then sorted on the program variables \( W, N, D \), depending on whether their limit is non-zero finite, zero, or infinity. When all the factors have been sorted, \( \text{LIMITI} \) checks for the cases:

- \( W \cdot 1 \cdot 1 \rightarrow W \)
- \( W \cdot 0 \cdot 1 \rightarrow 0 \)
- \( W \cdot 1 \cdot \infty \rightarrow \infty \)
- \( W \cdot 0 \cdot \infty \rightarrow W \cdot \frac{\text{LIMIT[N']}}{\text{LIMIT[D']}} \)
MULTIPLYTHROUGH

MULTIPLYTHROUGH[EXP, X] will multiply an expression or each side of an equation by X. If the expression or equation side to be multiplied by X is a PLS, MULTIPLYTHROUGH will multiply each term by X. In all other cases, it will just multiply EXP by X. The function uses EPRD for multiplication.
NEWNAME uses the LISP character handling functions to generate a new atom of the form \( F_n \), where \( n \) is an integer. The integer \( n \) is kept as the value of the constant \( NAMENUMBER \). \( NAMENUMBER \) is incremented each time \( NEWNAME \) is called.
NORMPOLY

NORMPOLY[EXP, X] uses SUBSTA to search for all PLS subexpressions of EXP. Then NORMPOLY2 factors a power of X out of each term in the subexpression so that it will be a polynomial whose lowest power in X is zero. For example,

\[ ax + bx^2 \rightarrow x(a + bx) \]

\[ \frac{a}{x} + bx^2 \rightarrow \frac{1}{x}(a + bx^3) \]

Since SUBSTA comes up from the bottom,

\[ ax + c(dx + ex^2) \rightarrow x(a + c(d + ex)) \]

NORMPOLY1 finds the power of X to be factored out by taking the minimum of the lowest power of X in each term. The lowest power of X in each term is computed by NORMPOLY4. If the term is a PFD, NORMPOLY4 uses NORMPOLY3 to add the lowest powers of each factor. The lowest power of X is of course 1, and the lowest power of X to a number is the number. All other expressions are assigned a lowest power of zero.
REPLACE

REPLACE[X, Y, Z] replaces the subexpression Z, which has been indicated with the light pen in the expression named X, by the expression Y. If the evaluation of Y contains a light pen reference to X, and Y is to be replaced at this same spot, then Z can be 'HOLE.

Furthermore, if Y is 'NIL, then Z will just be deleted from X.

Obviously, REPLACE must obtain the expression named X, with the subexpression Z marked by a light pen reference. To do this REPLACE uses a method common to all functions which need as an argument the expression in which a light pen reference has been made. REPLACE is defined as an FEXPR. It evaluates its first argument to get the name X of the expression. Then it reads from the disk the expression X PFORM. X PFORM is generated from X by the picture compiler, and it has the same subexpression structure as the displayed form of X.

The constant INFORM is assigned X PFORM as value. The light pen reference count INFRNUM is set to zero. Next, REPLACE evaluates its second argument with the function PFORMREAD, which would read X PFORM from the disk in order to return a subexpression indicated by the light pen, replaced by GETINSUB, which will use INFORM instead. Each light pen reference is marked in INFORM by GETSUB, which is called by GETINSUB. GETSUB alters the list structure to mark it and uses INFRNUM to count the subexpressions. When the second argument is evaluated, REPLACE has the expression Y and on INFORM the expression named X with the references marked in it. If Z is HOLE, then REPLACE just substitutes Y into X at a subexpression reference. If Z is not HOLE, then Z contains the light pen reference indicating where Y must be
substituted. A new XFFORM is read in and Z is evaluated.

Now X has the proper subexpressions marked. X is converted from
a PFORM to internal form by APOFF; the Nth marked subexpression is
replaced by (SUBMARK N NIL). It remains to substitute Y. It is
assumed that if there is more than one SUBMARK expression, they are
all arguments of the same operator. REPLACE uses SUBST to locate
this operator. REPLACE2 checks each argument of an operator to see if
it is a SUBMARK. When an operator containing a SUBMARK argument is
found, REPLACE1 replaces the first SUBMARK with Y and then REPLACE3
deletes the remaining SUBMARKS. If Y is NIL, then all SUBMARKS are
deleted.
RIGHT

RIGHT[EXP] returns the second argument of EXP.
SIMPLIFY

SIMPLIFY starts at the ends of the branches of the expression tree and works toward the top, performing the following simplifications:

\[ 0 \cdot a \rightarrow 0 \]
\[ 1 \cdot a \rightarrow a \]
\[ 0 + a \rightarrow a \]
\[ a + (b+c) \rightarrow a + b + c \]
\[ a \cdot (b \cdot c) \rightarrow a \cdot b \cdot c \]
\[ a^n \cdot a^m \rightarrow a^{n+m} \]
\[ i\cdot a + j\cdot a \rightarrow (i+j) \cdot a \]
\[ i\cdot a + j\cdot(a+b) \rightarrow (i+j) \cdot a + j\cdot b \]
\[ (a \cdot b)^n \rightarrow a^n \cdot b^n \]
\[ \int -i\cdot a \rightarrow -i \int i\cdot a \]
\[ \frac{d}{dx} -i\cdot a \rightarrow -1 \cdot \frac{d}{dx} i\cdot a \]

where \( i \) and \( j \) are rational numbers. Furthermore, all indicated arithmetic operations involving integers, floating point numbers, and fractions are carried out. The above operations require the comparison of expressions for equality. This is accomplished by hash coding the expressions and comparing the resulting hash code numbers. Additional simplification is also achieved when an expression which is identically equal to zero or one is replaced by its hash code number.
Although the simplification program is not long, the use of the hash code scheme and several other schemes to improve efficiency make it tedious to read. Therefore, the sources of difficulty in each scheme will now be discussed.

Generation of a hash code number for a complex argument is moderately expensive. Therefore, it would be desirable to generate each hash code only once. However, the storage of a hash code for each subexpression would take up a significant fraction of the available storage. A decision was made to save the hash code numbers only at the current level of simplification. The hash code number of a subexpression is computed from the hash code numbers of its arguments. The hash code numbers of the arguments are then discarded. In the expression \( d \times (b+c) \), it would be incorrect to compute the hash code of \((b+c)\) and discard those of \(b\) and \(c\), since the next operation is also addition and requires the comparison of \(d\) against \(b\) and \(c\) in an effort to collect terms. To solve this problem, the next higher operator is made available to the program. This is normally adequate, but in the expression \(e \cdot (b-b+d^2)\), the next higher operator from \(d^2\) is \(\times\); therefore, the hash code number of \(d\) will be discarded and the hash code of \(d^2\) computed. Then, when \(b-b+d^2\) simplifies to \(d^2\), the hash code number of \(d\) must be recomputed for comparison with \(e\). This is hopefully not a frequent situation.

If an expression is identical to zero or one, then it is probable that its hash code number will be zero or one. Therefore, whenever a new subexpression is formed, its hash code should be checked. If it is zero or one, the subexpression is taken as the same. The D functions
described earlier perform this operation. The hash code number for each subexpression is kept on the front of it. The D functions combine the hash code numbers and the expressions, perform the identity checks, and then put the resulting hash code number on the front of the resulting expression.

A further complication in the hash code comparison is that the numbers must also be compared with the appropriate group inverse. For example, the code of \((c \cdot (a-b))\) is the negative of the code for \(c \cdot (b-a)\). But the expression \(c \cdot (b-a) + c \cdot (a-b) + d\) should be simplified to \(d\). Similarly, since \(\cosh^2 x - \sinh^2 x = 1\), \((\cosh x - \sinh x) \cdot (\cosh x + \sinh x) \cdot d\) should be simplified to \(d\). In fact, the only frequent comparison of multiplicative inverses would be in expressions such as \((\frac{1}{2})^{+15}\), \(2^{-\frac{15}{2}}\).

Since it is expensive to compute the multiplicative inverse, this has been solved by representing exponentiated fractions with the larger integer as numerator. Thus, \((\frac{1}{2})^{+15}\) becomes \(2^{-\frac{15}{2}}\).

Finally, two types of arithmetic are used in the hash code scheme; the base arithmetic and the exponent arithmetic. This was explained in the discussion of the functions which compute the hash code numbers. For want of a better solution, exponents of exponents are coded in the base arithmetic. Therefore, as SIMPLIFY goes through the expression it must keep track of the current level of exponentiation modulo 2.

In this sense, the addition of exponents of like factors in a product is not the dual of the addition of coefficients of like terms in a sum.

A mechanism has been added to insure that subexpressions represented by the same list structure will be simplified only once. After a subexpression is simplified, the list structure is altered so that other
references to the subexpression will encounter the simplified form.
The expression is marked "already simplified". The mark also gives the
level of exponentiation for which the hash code number of the simplified
expression has been computed. If the current reference has the other
level, a new hash code number must be computed.

Combining terms or factors in a sum or product is a form of sorting
operation. However, only an exhaustive search is used here. In order
to speed up the search, the terms are combined in a special format.
Whenever the operator at the level above is PLS and the current level
is PLS, or the operator above is FLD or FWR and the current operator
is FLD or FWR, then the sort format is passed up without conversion to
standard internal form. During the combination of terms a numerical
term is kept separate, for any number can be combined with another
number, but only non-numerical terms with the same base can be combined.
Non-numerical terms are compared by the hash code numbers of their
bases. Unfortunately, it may result that the combination of two non-
numerical terms results in a numerical term. For example, $\frac{3}{5}$ is
considered non-numerical, but $2^{\frac{1}{5}} \cdot 2^{\frac{2}{5}}$ results in 2. Numerical terms
generated in this manner are combined with the number which is kept
apart when the sort format is converted to standard internal form.
Conversely, in the expression $(2a)^{\frac{1}{5}} \cdot b$, the number 2 must be considered
a non-numerical term after it is exponentiated.

The functions which carry out these operations can now be
described. The top level function, SIMPLIFY, calls SIMP2.
SIMP2[SX, Y, SIMPLEVEL] is a function of the expression to be simplified,
the operator of which this expression is an argument, and the
simplification level, that is, whether F or E arithmetic should be used in computing the hash code number of the expression. Going down the expression tree, SIMP2 must look for four conditions. The end of a branch of the expression tree is reached whenever an FRT operator or an atom is encountered. This expression is given to SIMP3, the function which handles the upward pass. If a PWR operator is encountered, then the SIMPLEVEL is logically OR'ed with 1 when SIMP2 is applied to the exponent. There is a similar step for the arguments of transcendental functions. It is also possible that the expression will have the form: ((hash-expression)·N), where N is the simplevel. SIMP2 checks for this structure by testing if CDR of the structure is atomic. This structure is set up whenever an expression is simplified, thus, if two subexpressions are in fact the same list structure, then when the first one has been simplified, SIMP2 will find the above structure when it reaches the second. It is possible that the two expressions do not occur at the same simplevel, therefore, the simplevel N at which the hash number was computed is part of the structure. If N is equal to the current simplevel, CAR of the structure is returned as the answer by SIMP2. Otherwise, the hash number is first replaced by one calculated for the current simplevel. In all other cases, SIMP2 goes on down the expression tree without a change in simplevel.

Coming back up the expression tree, SIMP3 checks for nine possibilities. SIMP3 has three arguments X, Y, and Z, the current operator, the expression term list, and the operator of the expression of which the current expression is an argument. If the current operator is flagged with FCODE, then it is a transcendental function. No
simplification is done on the expression. The hash number is computed from the hash number of the argument by the $F_{0}$CODE function. If the resulting hash number is zero or one, then the expression is replaced with the hash number in the result by SIMP33. The result consists of the expression with the hash number on the front of it. If the current operator is flagged with NSIMP, then the list of arguments is given to the NSIMP function and the list of corresponding argument hash code numbers is given to the NCODE function. Since the input argument list contains the arguments with the hash code number of each on the front of it, the hash numbers and arguments are split into separate lists by the functions SIMP34 and SIMP31. The results of the NCODE and NSIMP functions are given to SIMP33 as above. The operators PLS and PRD cause SIMP3 to call SIMP5 with arguments appropriate to addition or multiplication. The arguments of SIMP5 and their possible initial values and purposes are:

- **X**: term list of expression to be simplified
- **Y**: operator above
- **SU**: 0-0 or 1-1, appropriate identity and hash code
- **V**: NIL, list of collected terms
- **PM**: DPLS or DPRD, appropriate group operator
- **PNZ**: DPRD or DFWR, appropriate group operator
- **R**: PLS or PRD, appropriate group operator

SIMP5 removes each term from list X and takes one of six courses of action. The term may be:
1. A number. (M-N) The number is added to SU with FNW.

2. An expression having the same main connective as the current level and left in special sorting format: ((sort format).(M-N))
(M-N) is combined with SU by FNW. (sort format) is combined with V by SIMP5, using FNW, FNZ, and R.

3. An expression having the same main connective as the current level, but in standard form. The term is converted to sort format by SIMP5 and then given back to SIMP5, where it satisfies case 2.

4. A PWR and R is PRD. The term is converted to single term sort format by SIMP5 and given back to SIMP5 where it satisfies case 6.

5. A PRD with a numerical coefficient when R is FLS. The term is converted to single term sort format by SIMP5 and given back to SIMP5 where it satisfies case 6.

6. Anything else. The term may be in single term sort format already if the level below is PWR or if the term has been converted by SIMP5. Otherwise, it is listed with (1·1) to make a single sort term with a coefficient of one.

A more detailed explanation of the data formats will now be given along with a description of the functions auxiliary to SIMP5. Collecting terms in a product or sum is carried out using a special format. A list of the terms is formed; each member of the list has the form:

((code·base) (code·exponent))

Terms to be added to the list are put into this form; in the case of addition, the exponent is the numerical coefficient of the term. The code number of the base of a new term is then compared with the code number of each of the old terms on the list. If a match is found, the
exponents are added. Otherwise, the new term is put at the end of the list. Thus, functions are needed to convert expressions into sorting format and from sorting format into normal form, and to add new terms to the sorting list. The function SIMP51 adds a list of new terms in sorting format to the sorting list. It gives each term to SIMP52.

SIMP52 makes the comparison of bases described above. SIMP53 transforms terms from sort format into normal form. SIMP53 has as arguments the sort expression, the group operator, and the exponent operator. The base and exponent from the sort list are given to the exponent operator which constructs a term in normal form with its hash number on the front. If this term is a number, it is combined with SU in SORT5. If this is the last term on the sort list, then SIMP53 returns a list of the hash number, the converted term, and NIL. Otherwise, SIMP53 combines the hash number with that on the front of the list for the rest of the converted terms and adds the new term to this list, putting the new hash number on after it. SIMP55 is used to convert a single term into sort format. Given (code PWR base exponent NIL) and 1, SIMP55 creates ( (((code base) (code exponent))) (1 1)) which is the sort structure with one term, as it is passed up to SIMP5. In the case of PLS, the simelevel is not raised for the calculation of the exponent hash code number, as it is for PRD. SIMP54[X, Y, Z] converts a PLS or PRD expression X, to a sort list. Y is the group identity, and Z is T if X is a PRD expression. The conversion is straightforward, if a term has an exponent in the group sense, then it is used, otherwise 1 is used. The structure shown for SIMP55 is created. If the expression has a numerical coefficient, it is used rather than the identity. Returning
to SIMP3, when all terms have been sorted, a result is passed up which depends both on the results of the sort and the level above. If there is only a numerical term, this is passed up. If the current operator is the same as the operator above, the sort format is passed up.

Otherwise, the sort format is converted to standard form by SIMP53. If the expression part should reduce to a PLS when the current operation is PRD and the numerical part is not 1 and the operation above is PLS, then the distributive law is applied, the number is multiplied through the expression with a MAPL and the result is converted to sort format by SIMP54. Otherwise the number is just combined with the expression to produce the result to be passed up.

Returning to SIMP3, FWR produces a call to SIMP7. If the base is a PRD or FWR, then it may already be in sort format; if not, it is converted to sort format by a call to SIMP54 or SIMP55. If the base is not a FWR or PRD, the base and exponent are given to DFWR to produce the result. Otherwise, the sort format is given to SIMP71, which multiplies the exponent times each of the exponents on the sort format and simplifies each result with SIMP5. The number is exponentiated with DFWR. SIMP7 then checks to see how the result should be passed up. If the operation above is PRD or FWR, then the sort format is passed up.

Since the result of DFWR applied to the number and the exponent may be an expression rather than a number, it is necessary to add the result to the sort list if it is not a number and take a dummy number of 1.

If the above level is not PRD or FWR, then the sort format is converted to standard form by SIMP53 and then passed up.
Returning to SIMP3, an operation NIL indicates that the expression is an atom or FRT, and a hash code is generated by GENC and put on the front of the atom or FRT. Otherwise the expression is treated as an arbitrary function. The hash is computed by FGIN, and the hash numbers are removed from the arguments of the expression by SIMP31 before the operator and hash number are added to the front.

DRV and I7G are sent to SIMP6, which brings out a minus sign.

All expressions above the lowest level on the expression tree are sent from SIMP3 to SIMP6 where the list structure is altered as explained above.
SOLVE

SOLVE[EXP, X] applies the obvious transformations in order to solve an equation for a specified variable or expression, X. When all possibilities fail, it returns the equation in its partially solved state. SOLVE first checks to see if the expression is an equation and transfers to ERROR if it is not. Next, SOLVE substitutes the atom "SOLVE" for the expression X. This allows the use of MKP to compute dependence on X, no matter what the form of X. After the substitution, SOLVE1 calls EXPAND to multiply out both sides. SOLVE3 then sorts both sides into terms which depend on "SOLVE" and those which do not. SOLVE1 moves the dependent terms to the left side and the others to the right. The resulting terms are given to SOLVE2. The terms which depend on "SOLVE" are the first argument, those terms which do not are the second argument, and "SOLVE" is the third argument. SOLVE2 first checks to see if the dependent expression consists of only the variable itself, in which case the solution is accomplished; otherwise, it examines the main connective in order to apply a simplifying transformation.

If the connective is PL$S$ it tries to factor a dependent term out of the left side, leaving only an independent term as a cofactor. The right side can then be divided by the cofactor. This is done as follows. SOLVE51 looks at the first term of the PL$S$, if it is a PL$D$, it checks to see if only one of the factors is dependent and returns this one. If more than one is dependent, SOLVE51 fails by returning NIL. If the first term of the PL$S$ is not a PL$D$, SOLVE51 returns the entire term. SOLVE5 then tries to factor this dependent term out of
each of the PLS terms, leaving an independent cofactor. It uses SOLVE32 to examine each term. SOLVE32 checks to see if the term to be factored out is the whole term, or if the term to be factored out is a factor of the whole term. In the case of a PRD it calls SOLVE32I to examine each factor. If SOLVE35 is successful, the left side is divided by the cofactor. Otherwise, the current state of the equation is the answer.

If the left side connective is PRD, SOLVE4 tries to remove exactly one dependent factor and divide the right side by the independent cofactor.

If the left side connective is PWR, one of two transformations is done. If the exponent is independent, the right hand side is symbolically taken to the converse of the exponent. Otherwise, the log of both sides is taken, the left side is given to SOLVE5 which multiplies it out and simplifies it and then removes any independent terms of the top level PLS to the right side.

If the main connective is NLOG, SIN, COS, TAN, ASIN, ACOS, or ATAN, then an inverse of the right side is taken. SOLVE2I checks to see if the left and right sides have the same one of these connectives, in which case it can just be removed from both sides.
SPLIT

SPLIT[EXP] makes two passes through the expression. First, SPLIT1 replaces the property list of each subexpression with a count of the number of subexpressions which that subexpression contains. Then, SPLIT2 uses these counts to reduce the total number of subexpressions in EXP to less than 100. It does this by naming parts of EXP and substituting these names for the parts. The named subexpressions are placed on the disk with EASS.

SPLIT1 uses several functions to count the subexpressions. It goes down to the bottom of the expression tree, and SPLIT3 computes the count on the way back up. To do this SPLIT3 adds the number of subexpressions in each of the arguments. It adds 1 to this sum in order to count the current subexpression. It adds 1 by considering the operator to be an argument. The number of subexpressions in a given argument is computed by SPLIT4. SPLIT4 counts an atom as 1; other arguments have their count in place of a property list.

SPLIT2 goes down the expression tree to prune it. If it reaches an atom, it leaves it alone, as it would save nothing to rename an atom. Also, if the expression size, \( N \), is smaller than the allotted size \( N \), it needs no pruning. Otherwise, the arguments are pruned in reverse order to their size. SPLIT6 finds the size \( n \) of the largest argument. If renaming only part of this argument will bring the total count below \( N \), then SPLIT2 is applied to this argument with an allotment equal to the difference between \( N \) and the sum of the count for the other arguments. Otherwise, this argument has its count reduced to
zero by SPLIT5 and SPLIT8, indicating that it is to be renamed. Then
the R is subtracted from M, and SPLIT6 finds the next largest argument.
SUBSTITUTE

SUBSTITUTE[X Y Z] substitutes each member of set Y for the corresponding member of set Z in the expression X. It checks to see if Y and Z are sets and gives the arguments to SUBST1. The property lists are set to NIL so that expressions can be matched with EQUAL. SUBST1 gives the members of sets Y and Z to SUBST2, a pair at a time. SUBST2 substitutes the member of Y for each instance of the member of Z, with the following exceptions. a₁ does not contain an instance of a. SUBST2 uses SUBST3 to avoid testing the a in a₁ against the member of Z. Furthermore, SUBST2 will substitute:

A for B in (B X Y NIL) and
A₁ for B in (B X Y NIL).
If the argument of SUMEACH is of the form $\sum a + \cdots + b$, it will return $\sum a + \cdots + \sum b$. 
SUMEXPAND

If the argument of SUMEXPAND is of the form \[ \sum_{i=n}^{m} x \], where \( m \) and \( n \) are integers, then SUMEXPAND will return the expanded summation.
TERM

\textsc{term}(x, n) \text{ returns the } n^{th} \text{ argument of } x \text{ or NIL if } x \text{ has less than } n \text{ arguments.}
TRUNCATE

TRUNCATE[(Y, X, N)] finds terms to the \(N^{th}\) power in the expansion about the origin of \(Y\) as a function of \(X\). \(Y\) must have a finite number of poles at the origin. Since \(Y\) is allowed to have poles at the point of expansion, it cannot be expanded by just evaluating successive derivatives at the origin. Furthermore, differentiation can often lead to very large expressions. Therefore, the method used here is to expand the smallest subexpressions in a power series, and then to combine these series to form the expansion of successively larger subexpressions. The initial expansions must contain enough terms so that the final combined result will be valid to power \(N\). The required number of terms is calculated by making two passes through the expression. During the first pass, the minimum power of \(X\) in each subexpression is placed on the subexpression's property list. Then, the second pass uses these minimum powers as it goes down the expression tree in order to calculate the maximum required power of each subexpression expansion. The expansion of the individual atoms in a power series is simple, but many functions were needed to efficiently compute the expansion of each operator from the expansions of its arguments. The program is further complicated by the expansion of indexed summations. In only some cases can the program find the power of the variable of expansion as an increasing function of the index of summation. If this can be done and the summation has a finite lower limit, then the lowest power can be found. The second pass then carries out the expansion by repeated substitution. Arbitrary functions are expanded by differentiation. The arguments are still expanded in the
normal manner before the differentiation, unless they also require
differentiation. In this case, the differentiation is performed only
once.

This routine is so large that it will not all fit in memory at
one time. It has been split into two parts. The first part finds the
minimum powers, associates a minimum and maximum power with each atom
or constant subexpression and expands summations. The second part,
which is compiled in the file TRNK9 LISP, combines the subexpression
series in order to compute the final result. The most complicated part
is the direct generation of given terms in a multinomial expansion.
Each of the functions will now be described in detail. TRNK first
defines MIN to have the BODER function EMIN. It then defines MIN to
have an NSIMP function which leaves the expression unchanged unless all
of the arguments are numerical, in which case EMIN is used to find the
minimum argument as a result. TRNK next calls TRNK11 to explore the
expression, storing information on the expression's property lists.
The expression is then passed to TRNK122 which chains to the second
half of the program TRNK9.

TRNK11 passes each side of an EQN or all of any other expression
to TRNK111, which applies TRNK7, then TRNK8, and finally TRNK21.

A dependent expression is one which contains the variable of
expansion. TRNK7 applies TRNK71 to all dependent PLS or PRD
subexpressions in the expressions. TRNK71 changes the PLS's or PRD's
into a series of binary ones, such that any dependent argument is in
a binary one, but the independent arguments are lumped together to
produce one of the forms.

The independent terms are collected on the last argument of TRNK71. The function TRNK72 is used to avoid a redundant level if no independent terms have been collected.

The next step is to go down the expression tree in order to find the lowest power of the variable of expansion in each subexpression. The function TRNK8 goes down the tree to the individual variables and constant subexpressions. Coming back up the tree, the function TRNK82 dispatches control to the several functions which compute the lowest power of an expression from the lowest power of its arguments. The progress of TRNK8 down the expression tree is interrupted when a SUM subexpression is encountered. The lowest power of the variable of expansion may be expressed in terms of the variable of summation. TRNK8 calls TRNK83 to explore the expression. If the lower limit of summation is a number and the upper limit is a number or infinity, then the SUM body, B, expansion variable, X, and summation variable, i, are given to TRNK831. There are several cases where an expression for the lowest power can be found. If B is independent of X, then B is a constant and the lowest power is 0. If B is independent of i, then the lowest power can be found in the normal way, using TRNK8. Otherwise, the method used depends on
the main connective of B. For PRD it is the sum of the lowest power of the arguments. For PLS it is the minimum lowest power of the arguments. FWR is handled only if it is of the form $f(x)^{g(i)}$; in this case the lowest power in $f(x)$ is multiplied by $g(i)$. If it is RAN, the lowest power is 0. Otherwise, the user is asked to return an expression for the lowest power as a function of $i$. The expression for the lowest power as a function of $i$ which results from the above procedure is next checked to see that it is only a function of $i$ and in addition an increasing one.

If these conditions are met, then the lowest power is found by plugging in the lower limit of summation and the exponent expression is saved by putting it on the property list of the SUM subexpression under the indicator SUM. Once again, if the exponent expression does not meet these conditions, the user is asked for the lowest power.

Coming back up the expression tree, TRNS82 dispatches control to the following functions. PLS and PRD are computed directly by taking the minimum and the sum respectively of the lowest powers of the arguments. FWR sends control to TRNS83. Only two cases can be handled; a zero exponent reduces the whole expression to 1, and an integer exponent can be multiplied by the lowest power of the base. In other cases the user is asked for the lowest power. DRV sends control to TRNS86. This function checks to see if differentiation is performed with respect to the variable of expansion. If so, the user must be asked for the lowest power. Otherwise, the lowest power is the same as that of the DRV body. IYG sends control to TRNS87. Here again, the user must be asked if the limits are dependent and otherwise the lowest power is the same as that of the IYG body. If the user is asked
for the lowest power in DRV or ITG, then he must later be asked for the expansion. This is flagged by changing DRV and ITG to ASKDRV and ASKITG. All other connectives send control to TRNK88. They are treated as arbitrary functions which must be expanded by differentiation. Consequently, if any of the arguments have poles at zero, the user is asked, otherwise, the lowest power is taken as zero.

TRNK21 puts the highest needed power on the property list of the whole expression and gives the expression to TRNK2 which goes down the expression tree. When TRNK2 reaches an atom, it returns an ATOM structure with a minimum power of 1 or 0 depending on whether or not the atom is the variable of expansion. If the subexpression is independent, the maximum power is put on its property list and the downward recursion of TRNK2 is terminated. Otherwise, the maximum power needed for each of the arguments of the subexpression depends on its main connective. For PROD, the maximum power of each argument is the maximum power for the FRD minus the minimum power for the other argument. For PWR, it is now assumed that the exponent is an integer, the maximum power for the base is the difference between the maximum power for the PWR and (exponent-1)·(minimum power for base). SUM initiates a call to TRNK24. TRNK24 forms a PLS through expanding the SUM by substituting increasing integers starting with the lower limit of summation into the SUM body until the SUM is fully expanded or until the evaluation of the exponent function under the same substitution shows that the desired maximum power has been reached. The resulting PLS expression is then given to the top level function TRNK11 which applies both the minimum and maximum power passes. For all other main connectives, the maximum power for the arguments is that for the subexpression.
After TRNK2 has finished, the expression is given to TRNK122. This function creates a file LISPT LISP which contains the doublet TRNK123 (expression, max power, expansion variable), and then chains to the SAVED file LISPT. When control comes back, the answer is in file LISPTR LISP and this answer is read in and defined as a form to complete the operation of TRNK. The operation of file LISPT SAVED will now be described.

The top level function of LISPT SAVED, TRNK123, writes the file LISPTR LISP containing the result of applying TRNK12 to the expression generated by the main system. TRNK12 operates on each side of an EQN or the whole expression otherwise. TRNK9 is applied by TRNK12 to the expression and the resulting list of coefficients is given to TRNK121 which generates a FLS containing the terms of the desired expansion. TRNK121 just recurs down the list of coefficients using EPWR and EPRD, since the first coefficient is by definition that of the highest required power and no coefficients are omitted.

For purposes of efficiency TRNK9 uses a special data form, coming up the expression tree each function receives a list of coefficients, from the highest required to the start of all zero coefficients, with none omitted in between. TRNK9 goes down the expression tree to the individual variables and constant expressions, where it sets up the coefficient lists. It removes the ATOM level. TRNK31 is used to generate leading zeros for the coefficient lists. Downward recursion goes only through the body of ITG and DRV. For ASKITG and ASKDRV the user is asked for the expansion. For PWR, two cases are distinguished. If the exponent times the lowest power of the base is greater than the highest power required, then the coefficient list is NIL. Otherwise, TRNK9 is applied only to the base. In the case of NAM or an arbitrary function, a check is made to see if
any of the arguments have poles. If they do, then an effort is made to remove them by applying TRNK9 to the arguments and then converting the result back to normal form. The third argument of TRNK9 is set as a flag, for the NAM or arbitrary function subexpression must be expanded by differentiation, and this need not be repeated if the same situation is encountered at a level within the subexpression. If the arguments do not have poles at zero, the differentiation step is applied to the subexpression once the atom structure has been removed from it by ATOMOFF.

Coming back up the expression tree, the coefficient lists are given to functions associated with the main connective at the particular node. PLS sends control to TRNK1. TRNK1 adds corresponding coefficients with EPLS until one of the two coefficient lists is exhausted, the remaining coefficients of the other list are then taken as is. PRD sends control to TRNK5. If either argument list is NIL, all coefficients are zero on that list so the result is NIL. Otherwise, the arguments are passed to TRNK2222 in a special form best illustrated:

\[(0 \ 0 \ 0 \ a_0 \ a_1 \ a_2)\]

\[(b_3 \ b_2 \ b_1 \ b_0)\]

The number of leading zeros in the first list is one less than the number of coefficients in the second list, and the coefficients in the first list have been reversed. TRNK2222 removes the zeros from the first list, one by one. Before removing each zero it sends both lists to TRNK3, which generates one coefficient in the answer. This coefficient is a sum of terms generated by multiplying by means of EPRD corresponding terms in the above lists. This TRNK2222 process is terminated by a final call to TRNK3 once all zeros have been removed.
For PWR, there are two possible calls. If the exponent is positive, SRFWR is called and TRNKL is then used to remove trailing 0's from the coefficient list which SRFWR produces. SRFWR takes four arguments, the coefficient list, the maximum power on this list, the maximum power needed in the result, and the exponent to which the series represented by the coefficient list is to be raised. If the coefficient list contains coefficients of negative powers, the SRFWR2 is called; otherwise, SRFWR3 is called. The operation of SRFWR3 will be described first. The purpose of this function is to generate the coefficients of the desired terms in the multinomial expansion, without generating the entire expansion. For each term, SRFWR3 initiates a call to SRFWR1. The desired coefficient could be written as a sum of all products of the coefficients of the base series which have the form

\[ \frac{i_0}{a_{0}} \frac{i_1}{a_{1}} \frac{i_2}{a_{2}} \cdots \frac{i_n}{a_{n}} \]

where \( i_1 + i_2 + \cdots + i_n = N \), the exponent, and \( 1 \cdot i_1 + 2 \cdot i_2 + \cdots + n \cdot i_n = P \), the desired power. The desired coefficient is, however, in fact, written with the above sum factored on the \( a_j \)'s, taken in the dictionary order.

The coefficient is built up in this manner by an exchange of calls by SRFWR1 and SRFWR11. SRFWR1 has as arguments the needed power, \( P \), the number of coefficients yet to be used \( M \), the power \( R \), which the next coefficient in the dictionary order can be used to generate, the power, \( W \), of this coefficient in the base series, the minimum power, \( S2 \), in the base series, and the list, \( A \), of remaining coefficients. SRFWR1 recurs on the argument \( R \). \( R \) is started off at the minimum value which will allow
generation of a term, it is then increased by 1 until \( R \geq M \), in which case more the \( N \) total base coefficients would be used, or until 
\[ P*(R+1)^V < S2*(P*(R+1) \) which states that if the remaining \( M-(R+1) \) powers were all used on the lowest coefficient \( S2 \), the total power generated would still be greater than that required. The terms created for each value of \( R \) are added with ZFLS. Each of these terms is created by a call to SRPWR1. SRPWR1 checks to see if the coefficient list, \( A \), has only one more member. In this case it creates the desired term \((\text{coef})^{R/R!}\); otherwise, it constructs this term, but multiplies it by the term created by giving the remaining coefficient list to SRPWR1, along with the remaining number of powers needed, \( P-RV \), the remaining coefficients available, \( M-R \), and the lowest value of \( R \) which you can start with next, \( \text{MAX}[0, P-2R+M*(2-V)] \). The expression \( P-2R+M*(2-V) \) is best viewed as \( (P-RV)-(M-R^*(V-2)) \); \( P-RV \), as given above, is the number of powers still needed, \( (M-R)*(V-2) \) is the number of powers which would be produced if none of the next coefficient were used. The difference is the number of powers which must be used on the next coefficient, for using a coefficient on the next, rather than the one after next, will increase the generated power by one.

SRPWR2 has six arguments. \( N \) is the maximum power needed, \( X \) is the list of positive coefficients, \( Y \) is the list of negative coefficients in ascending order. \( S2 \) is the lowest negative coefficient and \( S1 \) is the highest positive coefficient. \( S \) is the exponent to which the base series is raised. SRPWR2 uses four program variables: \( J, U, M, \) and \( P \). The result is built up on \( U \) and is a sum of terms for each coefficient. The current coefficient is \( J \), which is indexed from the lowest which can be
made up to \( N \), the highest needed. For each value of \( J \), a coefficient is built up which is a sum of terms. Each of these terms is of course a product of negative and positive coefficients such that the sum of the powers is \( J \), and the number of coefficients is \( S \). The terms can be classified by the number, \( M \), of negative coefficients and the number, \( P \), of negative powers which they contain. SRPWR generates all possible values of \( M \) and \( P \) which will produce the given \( J \); it then calls SRPWR1 for each pair of values of \( M \) and \( P \), once to produce the negative part of the coefficient, and once to produce the positive part. The positive and negative parts are then multiplied together with EPRD. The values of \( M \) and \( P \) are generated as follows. \( M \) is started off at 0 if \( J \) is positive, or so that \( M \times J < S2 \) if \( J \) is negative. \( M \) is then incremented until \( M > S1 \times (S-M) - J \). The meaning of this relation is as follows. Since the highest value of negative base coefficient is -1, \( M \) negative coefficients must create at least \( M \) negative powers. The maximum number of positive powers which can be generated with the remaining \( S-M \) coefficients is \( S1 \times (S-M) \), but \( J \) of these are needed for the result, so that \( S1 \times (S-M) - J \) can be matched against negative ones and this number must be greater than or equal to \( M \). Now, for each feasible number of negative coefficients, terms for corresponding possible negative powers are generated. The number of negative powers is given by \( P \), which is started off at the minimum of \(-M \times S2\), the most negative you can make, and \( S1 \times (S-M) - J \), the most negative you can absorb with extra positive powers. \( P \) is then decremented until it is less than \( M \) or until it is less than \(-J \), which means that if \( J \) is negative, not enough negative powers to produce it will be generated. Each completed coefficient is
multiplied by $S'$. Returning to TRNK9, if the exponent of FWR is negative, then a call to TRNK63 is given. TRNK63 inverts the series created by SRFWR as for a positive exponent by a call to TRNK6. TRNK6 first checks for a null series which would indicate division by zero is called for; in this case it notifies the user. Otherwise, it sets up TRNK61, which does the inversion. TRNK61 builds up the answer, term by term, on its first argument $W$. It recurs on its last argument $N$ which is a term counter, incremented from 1 up to the number of coefficients on the list $U$, of the series to be inverted. For each value of $N$ a term computed by TRNK62 is added to $W$. This term is $\frac{1}{a_0}$ times the sum of the product of corresponding terms of $W$ and $U$.

Returning once again to TRNK9, ITG and DRV send control to TRNK93. Since only the case where the integral or derivative do not involve the variable of expansion is handled, TRNK93 goes through the coefficient list putting the integral or derivative structure and arguments on each coefficient. The only remaining cases for TRNK9 are NAM and an arbitrary function; they are handled in the same manner. The arguments are treated as explained on the downward pass and then the expression in standard form is passed to TRNK95 for expansion by differentiation. It is at this point that the check is made to see if the subexpression under consideration is contained in a larger one which must also be expanded by differentiation, in which case the expression is returned undifferentiated by TRNK95. Otherwise, TRNK95 calls TRNK13 to expand the expression by differentiation. TRNK13 uses DIFF2 to repeatedly differentiate the expression, which it keeps on program variable $U$. Before each differentiation a term of the
expansion is set up by substituting zero for the variable of expansion in the expression \( U \), using TRNK131. TRNK131 substitutes zero for the variable of expansion, except that if it comes to a DRV subexpression it creates an EVL level.

This completes the explanation of TRNK.
Chapter VII

HASH CODING FUNCTIONS OF A COMPLEX VARIABLE

Introduction

Several of the routines described in the preceding chapter require the comparison of algebraic expressions for equivalence. The comparison is made by hash coding the expressions and comparing the resulting hash code numbers. The formulas required for the hash coding scheme and the functions which apply them were described in the last chapter. The mathematical considerations underlying these formulas will now be described.

The elementary functions of a complex variable are those which can be expressed by the following recursive scheme. Any complex constant or variable will be called an expression; if $u$ and $v$ are expressions, then so are $u + v$, $u - v$, $u^v$, $e^v$, $-u$ and $\frac{1}{u}$. The trigonometric and hyperbolic functions may be expressed explicitly. Because of the defining relations of the complex field and the trigonometric identities, there are infinitely many expressions for any given function. Two expressions will be said to be equivalent if they represent the same function.

Existing schemes for expression comparison use the defining relations along with some additional conditions to put each expression in a canonical form. If the canonical forms of two expressions are identical, they must represent the same function. This method has certain drawbacks. First, putting the expression in a canonical form requires the comparison of many subparts of the expression with each other. In particular, the commutative
law requires that the terms in sums and products be sorted. Second, discovery that two expressions are equivalent requires a comparison of every subpart of one with the corresponding subpart of the other. Third, the problem of reducing all equivalent expressions to one canonical form is recursively unsolvable (7) and the existing schemes fail in many cases.

This chapter explores a probabilistic approach. Suppose \( F(z) \neq G(z) \) (\( F \) and \( G \) are elementary functions), then \( F(z) - G(z) = 0 \) has, at most, a countable number of solutions, while the complex numbers are uncountable. Therefore, the probability that \( F(z) - G(z) = 0 \) for a point \( z \) chosen at random is 0. Thus, it would be possible to test for equivalence of expressions by comparing their values at a randomly selected point. It is possible to get some approximation to this fact with the finite arithmetic of a computer.

One method would be to substitute a random floating point number for each occurrence of each distinct variable and then evaluate the resulting expression using floating point arithmetic. This method is limited by overflow and roundoff error. For example, if \( x \) is a floating point number chosen at random from a flat distribution, then with probability one half \( x^2 \) is larger or smaller than all the floating point numbers; it does not appear possible to find a rule for mapping \( x^2 \) back into the floating point numbers such that the code numbers of equivalent expressions will be very nearly the same. This overflow is difficult to avoid by restricting the initial choice of floating point numbers since expressions of the form \( u^v \) are allowed. Furthermore, if two expressions, \( x \) and \( y \) are of different
orders of magnitude, then, because of roundoff error, \( x + y \) may evaluate to either \( x \) or \( y \). This is a particularly serious disadvantage since it is likely that an expression will be compared with subparts of itself. The same problems arise with a floating point approximation to the complex numbers.

Another strategy, which we investigate here, is to use a finite field, instead of the infinite field of real numbers.

**Finite Fields and the Exponent Arithmetic**

a. **Finite Fields**

The use of floating point numbers in the code number scheme is limited because the sum or product of two floating point numbers is not necessarily a floating point number. This problem is avoided if a finite field, \( F \), is used, since the field can be chosen small enough so that every element can be represented by a computer number. The task is to choose \( F \) such that expressions which are equivalent in the complex numbers are also equivalent in it. That is, we need a homomorphism from the complex numbers onto \( F \). We now develop a field which meets this requirement in many, but not all, cases.

An abelian group \( G \) is a set of elements with an operation \( \times \) and an identity element \( e \) such that:

1. \( a \in G, \ b \in G \) then \( a \times b \in G \)
2. \( a \in G \) then \( ae = ea = a \)
3. \( a \in G \) then \( a^{-1} \in G \) \( \Rightarrow aa^{-1} = a^{-1}a = e \)
4. \( a \in G, \ b \in G \) then \( ab = ba \)

A finite field \( F \) is a finite set of elements with an operation \( + \) under
which the elements of \( F \) form a group with identity 0 (the additive group),
and an operation \( \cdot \) under which the elements of \( F' = F - 0 \) form a group
with identity 1 (the multiplicative group). In addition the relations
\[
a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{and} \quad a \cdot 0 = 0
\]
hold.

If \( m \) and \( n \) are integers and \( p \) is a prime integer then \( c = (m + n) \mod p \)
means that \( c \) equals the remainder of \( (m + n) / n \). Multiplication \( \mod p \)
is defined similarly. It can be verified that the integers less than a prime
form a finite field under the operations addition and multiplication \( \mod p \).
The additive inverse of 1, \( -1 \) is seen to be \( p - 1 \) since \( p - 1 + 1 = 0 \mod p \).

b. The element \( i \)

In the complex field there is an element \( i \) such that \( i^2 = -1 \), so
such an element is also required in \( F \). To see how this restricts the
choice of \( p \) one needs the fact found in the references that the multiplicative
group \( F' \) of a finite field is cyclic. This means that there is an element \( \alpha \)
(called a generator) in \( F' \) such that every element in \( F' \) is some power of \( \alpha \).
In fact, \( F' \) can be written \( 1, \alpha, \alpha^2, \ldots, \alpha^{p-2} \) and \( \alpha^{p-1} = 1 \).
Since \( p \) is a prime it is odd and so \( \frac{p-1}{2} \) is an integer. \( \alpha^{(p-1)/2} = -1 \) since \( \alpha^{(p-1)/2} \neq 1 \)
and \( \{\alpha^{(p-1)/2}\}^2 = 1 \). If \( \frac{p-1}{2} \) is even then \( r = \frac{p-1}{4} \) is an integer such that if
\( i = \alpha^r, \ i^2 = -1 \). Note that either \( \alpha^r \) or \( \alpha^{3r} \) can be chosen as \( i \) and the other
becomes \( -i \). We have thus shown that \( F \) will have an element \( i \) if and only
if \( p \) is of the form \( 4q + 1 \).
c. The Exponent Arithmetic

In the complex numbers, one might have to test for the equivalence of two expressions such as $u^{v+1}$ and $u^v \cdot u$, where the exponent arithmetic is also performed in the complex numbers. However, since the multiplicative group is a cyclic group with one less element than $F$, $v + 1$ must be computed mod $(p - 1)$. Since an isomorphism does not exist between the additive and multiplicative groups of $F$, the exponent operations cannot be performed in it. This failure of the finite field evaluation to be recursive in the exponent direction is a serious limitation. Furthermore, since $p - 1$ is not a prime the exponent operations will not form a field.

Fortunately, many expressions encountered in analysis have rather simple exponents and so much can be saved by evaluating the exponents in the exponent arithmetic which we now define.

Let the basic elements of $E$ be the integers less than $p - 1$. Addition and multiplication are mod $(p - 1)$. It is easy to see that all elements have an additive inverse. No even integers have a multiplicative inverse, however, for this would imply:

$$2 \cdot s \cdot (2 \cdot s)^{-1} = 4 \cdot q \cdot m + 1$$

or

$$2 \cdot (s \cdot (2 \cdot s)^{-1} - 2 \cdot q \cdot m) = 1$$

which is a contradiction since 1 has no divisors in the integers. If we take $q$ prime, then the odd integers other than $q$ have a multiplicative inverse as a consequence of the Euler theorem (see Ref., Albert, p. 47):
Let \( \phi(m) \) be the number of integers \( g \) such that \( 0 < g \leq m \) and \( g \) is prime to \( m \). Then
\[
\phi(m) = 1 \mod m
\]
for every \( a \) relatively prime to \( m \).

The failure of the even integers to have a multiplicative inverse means that \( u^{1/2} + 1/2 \) will not evaluate to \( u \). We therefore adjoin to the basic elements of \( E \) the element \( \xi \), where \( 2 \xi = 1 \). Closing \( E \) under multiplication and addition would require that all the elements of the form \( b\xi^i (b \text{ an integer less than } p - 1 \text{ and } b \text{ odd}) \) be in \( E \). However, many cases can be covered if we allow only elements \( a \) or \( b\xi \).

d. **Square Roots in \( F \)**

\( u^\xi \) should evaluate to be the square root of \( u \) in \( F \), however, only one half the elements in \( F \) have a square root, these are the even powers of the generator, \( \alpha \), of \( F' \). Since \((\alpha^n)^2 \) is even for even \( n \) and any \( r \) only \( 1/4 \) of the elements could have a square root computable by raising the element to some power, that is by assigning some integer to \( \xi \). However a method of finding the roots of this smaller set can be found, as will be shown next, i.e., there exists a \( \xi \) such that if \( u = \alpha^{4n} \), then \( u^\xi = \alpha^{2n} \).

If \( p \) is of the form \( 4q + 3 \), then since \( 4n(q + 1) = 2n \mod (4q + 2) \) for \( n < q \) one finds \((\alpha^{4n})^{q+1} = \alpha^{2n} \), \( q + 1 \) is therefore the proper value for \( \xi \).

The requirement that \( p \) be of the form \( 4q + 3 \) is unfortunately in conflict with the earlier requirement that \( p \) be of the form \( 4q + 1 \). By choosing \( p = 8q' + 5 = 4 (2q' + 1) + 1 \) one obtains by a similar argument the square root of \( 1/4 \) of the elements with \( \xi = q' + 1 \).
e. Trigonometric Identities

Define in the usual manner:

\[
\sin \theta = \frac{e^{i\theta} - e^{-i\theta}}{2i} \quad \text{sinh} \, \theta = \frac{e^{\theta} - e^{-\theta}}{2}
\]

\[
\cos \theta = \frac{e^{i\theta} + e^{-i\theta}}{2} \quad \cosh \, \theta = \frac{e^{\theta} + e^{-\theta}}{2}
\]

where \( e \) is an element of \( F \) yet to be chosen. Note that \( i^2 = -1 \) in the E arithmetic but this does not arise in taking sums and products of the above functions. For instance:

\[
\sin^2 \theta + \cos^2 \theta = \left( \frac{e^{i\theta} - e^{-i\theta}}{2i} \right)^2 + \left( \frac{e^{i\theta} + e^{-i\theta}}{2} \right)^2
\]

\[
= \frac{e^{2i\theta} - 2 + e^{-2i\theta}}{4} + \frac{e^{2i\theta} + 2 + e^{-2i\theta}}{4}
\]

\[
= -\frac{2 - 1}{4} + \frac{2 + 1}{4}
\]

\[
= 1
\]

If \( \sin (\theta) \) is to equal \( \cos \left( \frac{\pi}{2} - \theta \right) \), then it is necessary that \( e^{i\pi} = -1 \) and \( e^{i\pi/2} = i \). If \( e \) is to have a square root it must be an even power of \( \varpi \); taking \( e = \varpi^{2n} \) one obtains for \( p = 8q^2 + 5 \):

\[
e^{i\varpi} = -1
\]

\[
\varpi = \sqrt[2n]{4q^2 + 2}
\]
2πn = (4q' + 2) \mod (8q' + 4)

nπ = (2q' + 1) \mod (4q' + 2)

Equation (1)

For any choice of \( n \), \( e \) is determined, providing equation (1) can be solved for the element \( \pi \). Some reflection will show that trigonometric calculations may involve roots of \( e \) greater than 2. Suppose \( n \) is chosen odd, then one square root of \( e \), \( \alpha^2 \), has no square root, nor does \(-\alpha^2\), the other square root of \( e \). That \(-\alpha^2\) does not have a square root is a consequence of the choice of -1 as a square. \(-\alpha^{2m} = -1 \cdot \alpha^{2m} = \text{(square)} \) (nonsquare) and a square times a nonsquare must be a nonsquare. From this one can see that if \( e \) is to have a 2\( m \)th root, it must be chosen of the form \( \alpha^{2m} \).

Note that the choice of \( n \) divides the elements of \( F^\prime \) between the roots and powers of \( e \).

Returning to the solution of equation (1) for the case \( n = 1 \), we see that no satisfactory solution is possible. For (1) implies that \( i\pi = 2^m(2q' + 1) \) + \( 2q' + 1 \) must be taken prime in order to reduce the number of elements which map into the same element in the \( E \) arithmetical. Therefore, \( 2q' + 1 \) must divide either \( i \) or \( \pi \), but this is not acceptable. Suppose \( 2q' + 1 \) divides \( \pi \), then in the \( E \) arithmetical 4\( \pi = 0 \). Since either 4\( \pi \) or 4i is a highly probable element, this relation is unacceptable. Once again, a patch can be inserted so that \( e^{im/2} \) will equal i in the most frequent cases.

When an element is exponentiated in the \( F \) arithmetical, special checks are made. If the base is \( e \) and \( i\pi \) divides the exponent, then \(-1)^{\text{exp}/i\pi} \) is computed. \( i\pi \) is chosen small enough so that small multiples of it will be less than the prime \( p \).
Summary of the Requirements

1. $p$ of the form $8q' + 5$.
2. $2q' + 1$ prime

Machine Realization -- Finding a Prime

The requirements of section II are:

1. $p = 8q' + 5$
2. $2q' + 2$ prime

Another requirement is:

3. $p$ less than $1/2$ the largest machine integer.

This allows addition without overflow. The multiplicative inverse of an element $a$ in $F$ is found by noting that $a^{-1} = a^{p-2}$. To raise an element $a$ to any power we begin by multiplying it by itself, creating the numbers $b_1 = a^2$, we then express the power as a binary number and add up the appropriate $b_1$. This leads to the requirement:

4. $p - 2$ should be expressible as a few powers of 2.

To find a prime for the 7094 the following procedure was followed:

1. Beginning with $n = 2^{29}$ test if $p = 16n + 13$ is prime by dividing by every odd number up to $\sqrt{p}$.
2. Test if $4n + 3$ is prime.
3. Find a generator $\alpha$, of $F'$. (If $a$ is not a generator then $a^2 = 1$ or $a^4 = 1$ or $a^7 = 1$ or $a^{27} = 1$.

Raise $a$ to these four powers by the scheme above.)

Almost $1/2$ of the elements are generators so one is quickly found.
4. Compute \( i = x^{4n + 3} \)

This procedure resulted in:

\[ p = 8, 589, 949, 373 \]
\[ x = 13, 560, 097 \]
\[ i = 5, 525, 736, 173 \]

Some Example Problems

Ten trigonometry identity problems were selected from a textbook (8).

The scheme produced the same hash number for both sides in all but the last. The identities are:

1. \( \sin x \tan x + \cos x = \sec x \)
2. \( (\sin x \cot x + \cos x)/\cot x = 2 \sin x \)
3. \( \csc^2 x + \cot^2 x + 1 = 2/\sin^2 x \)
4. \( \cos x \cot x + \sin x = \csc x \)
5. \( (1 - \sin x)(\sec x + \tan x) = \cos x \)
6. \( \sin x/(1 - \cos x) = \tan x/(\sec x - 1) \)
7. \( \csc^4 x - \cot^4 x = \csc^2 x + \cot^2 x \)
8. \( (\sin x/(\sec x + 1)) + (\sin x/(\sec x - 1)) = 2 \cot x \)
9. \( \cos^6 x + \sin^6 x = 1 - 3 \sin^2 x \cos^2 x \)
10. \( \sqrt{(\sec x - 1)/(\sec x + 1)} = (1 - \cos x)/\sin x \)

The probability of Error

Estimation of the probability of error is difficult. The average probability of error for certain subsets of expressions will differ from that for all expressions. No statistics are available on the expressions which will be encountered in practice.
It is possible that two expressions which represent the same function will receive different code numbers because some exponent operation does not preserve the equivalence. Study of section II should make clear under what circumstances this will happen.

In the simplification program described in the Chapter VI, expressions with the same code number are considered equivalent. Therefore, an accidental match of non-equivalent expressions is very serious. If we could show that the operations in the E and F arithmetic mapped their sets of elements uniformly back onto themselves, then the probability of a match between two expressions selected at random from the set of all expressions would be $1/p$. Unfortunately, this is not the case. In the F arithmetic the operations of multiplication and addition and their inverses do satisfy this criterion. Looking at the cyclic group $F^*$, however, one sees that raising any element in $F$ except a multiple of $w$ to all powers will produce either $1/4, 1/2, or all the elements of $F$. Thus exponentiation tends to map the elements into the 4th powers of a generator and so increase the probability of random match.

The same bunching occurs in the E arithmetic. The distribution of elements after $n$ operations can only be found using a rather complicated two dimensional convolution. The distributions after one operation shown in Figure 1 indicate that for moderately complicated exponents the probability of error should remain in control.
Distribution of elements in E after one operation. Within each classification, the elements are ordered in the normal manner.
Chapter VIII
THE LANGUAGE OF MATHEMATICAL EXPRESSIONS

The introduction stated that man-computer symbiosis can bring about an effective problem solving team by assigning to man and computer those parts of a problem which they are best suited to solve. Each is much better than the other at certain jobs; therefore, the two together can do the problem better than either alone, if the cost of breaking the problem up is not too high. The task assignment may result in either parallel or serial operation of man and computer. In a parallel operation, each has continuous access to all important data generated by the other, while in a serial operation, data exchanges occur only at significant points in the problem. Sometimes, a task in a serial operation could be performed by either man or computer; the one which has performed the previous step often has the relevant information for the task. The savings in sending the task to the best partner must be great enough to cover the cost of a possible data exchange. In this case, the cost of data exchange is critical to the balance between man and machine. This would not be so if most tasks could only be performed by one or the other.

The solved problems presented in Chapter II require some serial operation with much data exchange. It is, therefore, important to investigate the possibilities for man-machine communication in non-numerical analysis. A language consists of a series of statements following certain syntactic rules which the communicants map into their respective models of the world.
Often, the language can be divided into a central part, following a rather small set of well defined rules, together with a series of succeedingly complex exceptions to these rules and methods for continued discourse. In the case at hand, the central part is the set of mathematical expressions used in analysis. These expressions will be investigated in this chapter and the following two chapters.

The available input-output devices for the language are the typewriter and the display. Either a one-dimensional or a two-dimensional mathematical expression syntax can be used with the typewriter or the display. The use of a one dimensional syntax for typewriter input is discussed in chapter X. Chapter IX describes a two dimensional syntax for display output. A two dimensional syntax for the typewriter is discussed by Klerer, however, he has not stated any useful general ideas. The problems involved in constructing a two dimensional input syntax for the display will be considered in Chapter XI.
Chapter IX

SYNTAX AND DISPLAY OF MATHEMATICAL EXPRESSIONS

Introduction

This chapter describes a LISP program which computes the pictorial representation of a mathematical expression from a LISP S-expression representation of the expression. Since the language of mathematics contains a number of special cases, the computation will be introduced with a summary of the results produced by each part of the program and the general principles followed.

Because of the possible use of many transformations and identities, a mathematical expression can be represented in many equivalent forms. For instance \(a(b+c)\) could be represented as \(ab + ac\). In certain contexts the picture of one of these forms may be easier to read than the picture of the others. For instance \(a - x\) might be preferred to \(a + (-1 \cdot x)\).

On the other hand, the use of several equivalent internal forms complicates the writing of routines needed to transform mathematical expressions. Furthermore, the most convenient internal form might be difficult to read. In the first pass of the picture compiler the internal form is transformed into an equivalent form which is easy to read. In addition, parentheses are added. The LISP expression resulting from this first pass has the form of a tree of subexpressions. Each node of the tree is a mathematical operation and the branches from that node are the arguments of that operation. At the end of the branches are the individual variables and
constants. The form of this internal tree is preserved throughout the remaining steps of the picture compilation. For example, the expression \( a = c \cdot d + e \) has the tree:

![Tree Diagram]

Mathematical operations are represented in "pictures" by symbols placed in special arrangements. A classification of the representations will be given later. The situation is complicated by the fact that sometimes the same symbols and positional notations are used in different combinations to represent different mathematical operations. The second pass of the picture compiler rewrites the internal tree in terms of these common symbols and positions. The symbols and positions are expressed in terms of a set of special forms. These forms are shown in Figure 1 through Figure II. The compiler starts at the base of the internal tree and works out to the ends of the branches, rewriting the tree at each node.

This rewriting introduces additional nodes into the internal tree, but these new nodes are marked so that the old structure is preserved. The display resulting from a node and its branches is called a picture part. Sometimes the form of a picture part depends on the dimensions of its arguments. In this case the second and third pass must be applied to the arguments before the second pass can be applied to this part.

In the third pass each picture part is inscribed in a rectangle. The compiler starts at the ends of the branches of the internal tree and works toward the base. The compiler first inscribes each individual constant and variable in a rectangle, defined by its width, height, and
depth. A dimensioned rectangle is then chosen in turn for each larger picture part; i.e. operator with arguments. The dimensions of each rectangle depend on its own operator and also on the dimensions of those rectangles associated with the arguments of the picture part. In addition, we compute the relative positions of the arguments of each picture part with respect to the lower left hand corner of its circumscribed rectangle.

Once the entire expression has been dimensioned, it can be positioned on the oscilloscope face. The final pass of the compiler then sends to the display the name, size, and position of each symbol. These symbols are interspersed with non-displaying left and right pseudo-parentheses. These pseudo-parentheses group the picture symbols into a tree of sub-expressions identical with the internal tree structure obtained from pass one of the picture compiler. Thus, a light-pen reference to a picture symbol can be identified with the smallest subexpression in the picture tree which contains it, and with the corresponding LISP subexpression in the internal tree. This makes it possible easily to designate mathematically meaningful picture-parts as arguments for other operations.

Transformations to Facilitate Semantic Interpretation

A method is needed for referencing meaningful subparts of the LISP source expression by pointing with a light pen to symbols in the displayed picture expression. Thus, before the picture is compiled, the source expression is transformed using mathematical relations into a form which makes it easy to establish a one-to-one mapping between subparts of the picture expression and mathematically meaningful subparts of the LISP
source expression. Then, in addition to the picture symbols, the
compiler sends the display a series of left and right pseudo-parentheses
which parse the picture symbols into a tree identical with the tree formed
by the transformed source expression. When the light pen is pointed at
a picture symbol, the smallest picture sub-expression which contains it
is intensified and the corresponding LISP expression may then be referenced.
The source expression is transformed in order to bring it into a better
pictorial form, for instance, the order of the product is reversed so that
\( (\int_{a}^{b} c\, dx) \, d \) can be written \( d \cdot \int_{a}^{b} c\, dx \).

Although the picture compiler is powerful enough to handle most all
of the notations which arise in mathematics, simple compiler rules have
only been written for the mathematical expressions listed below, which
arise in the solution of non-linear differential equations. These expressions
are represented by LISP S-expressions. In brief, S-expressions are
defined recursively as follows: Any number or string of alphanumeric
characters is an S-expression. One or more S-expressions separated by
spaces and surrounded by parentheses is an S-expression. Individual numbers
and alphanumeric strings are called atoms. Individual variables and
constants in the mathematical expressions are represented by atoms.
The mathematical operators are represented by S-expressions consisting
of the operator name, its arguments in some given order, and the atom NIL.
NIL denotes the null S-expression; it is replaced by temporary results
during the transformation of expressions and represents the property list
of the mathematical operator. The individual mathematical operators
currently in the program are:

(Note: most of the operators can take any number of arguments, in the obvious manner.)

1. \((\text{PLS} \ A \ B \ C \ \text{NIL}) \equiv A + B + C\)
2. \((\text{PRD} \ A \ B \ C \ \text{NIL}) \equiv A \cdot B \cdot C\)
3. \((\text{FRT} \ I \ J \ \text{NIL}) \equiv \frac{1}{J}\)
4. \((\text{PWR} \ A \ B \ \text{NIL}) \equiv A^B\)
5. \((\text{DRV} \ A \ B \ C \ D \ E \ \text{NIL}) \equiv \frac{d}{dA} \frac{B+D}{dAB \cdot dC^D} \ E\)
6. \((\text{ITG} \ D \ A \ B \ C \ \text{NIL}) \equiv \frac{B}{A} \ C_d D\)
7. \((\text{SUM} \ A \ B \ C \ D \ \text{NIL}) \equiv \sum_{A = B}^{C} D\)
8. \((\text{EYL} \ A \ B \ C \ D \ E \ \text{NIL}) \equiv E \left| \frac{C}{A} \right|_{\text{B}}^{\text{D}}\)
9. \((\text{NAM} \ A \ B \ C \ \text{NIL}) \equiv C_{\ A, \ B}\)
10. \((\text{F} \ A \ B \ \text{NIL}) \equiv F(A, B)\)
11. \((\text{NAM} \ A \ B \ (F \ C \ D \ \text{NIL}) \ \text{NIL}) \equiv F_{A, \ B}(C, D)\)
12. \((\text{FTL} \ A \ \text{NIL}) \equiv A!\)
13. \((\text{ABS} \ A \ \text{NIL}) \equiv |A|\)

Where A, B, C, D and E are arbitrary expressions, I and J are integers, and F is any atomic symbol not recognized as an operator.

-\(X\) is represented by \((\text{PRD} \ -1 \ X \ \text{NIL})\) and \(\frac{1}{X}\) is represented by \((\text{PWR} \ X \ -1 \ \text{NIL})\).

If one of the factors of a sum or product is a number, it is often the left-most argument.
The first compiler pass performs the following transformations:

\[(\text{PLS A NIL}) \rightarrow \text{A}\]
\[(\text{PRD A NIL}) \rightarrow \text{A}\]
\[(\text{PLS N A B NIL}) \rightarrow (\text{PLS A B N NIL})\]
\[(\text{DRV S 1 (F S NIL) NIL}) \rightarrow (\text{DRV 1 1 (F S NIL) NIL})\]
\[-N \rightarrow (\text{NEG N NIL})\]
\[(\text{PRD -1 A NIL}) \rightarrow (\text{NEG A NIL})\]
\[(\text{PRD -2 A NIL}) \rightarrow (\text{NEG (PRD 2 A NIL) NIL})\]
\[(\text{PRD A (SUM A B C D NIL) E NIL}) \rightarrow (\text{PRD A E (SUM A B C D NIL) NIL})\]
\[(\text{PRD A (DEL B NIL) C NIL}) \rightarrow (\text{PRD A C (DEL B NIL) NIL})\]
\[(\text{PRD A (PWR B -1 NIL) NIL}) \rightarrow (\text{DVD A B NIL})\]
\[(\text{PRD A (PWR B (PRD -1 C NIL) NIL) NIL}) \rightarrow (\text{DVD A (PWR B C NIL) NIL})\]
\[(\text{F A B NIL}) \rightarrow (\text{FUNCTION F A B NIL})\]
\[(\text{PRD A (1 TG A B C D NIL) E NIL}) \rightarrow (\text{PRD A E (ITG A B C D NIL) E NIL})\]
\[(\text{FRT 1 -2 NIL}) \rightarrow (\text{NEG (FRT 1 2 NIL) NIL})\]
\[X \rightarrow (\text{ATOM X NIL})\]
\[(\text{NAM X Y (F A B NIL) NIL}) \rightarrow (\text{FUNCTION (NAM X Y F NIL) A B NIL})\]

Where \(N\) is a number, \(X\) is a number or variable symbol, and \(S\) is a non-atomic subexpression.

In addition to these transformations, levels of parentheses are inserted. For instance \((\text{PRD A (PLS B C NIL) NIL}) \rightarrow (\text{PRD A (PAREN (PLS B C NIL) NIL) NIL})\). No parentheses are used around the arguments of \text{FUNCTION}, \text{PAREN}, \text{ABS}, \text{EVL} or \text{NAM}. A size change always removes the need for parentheses. In addition, the integration and summation
symbols stand in place of parentheses for operators on their left and a transcendental function is parenthesized if it is raised to a power. Otherwise, parentheses are placed around an argument of an operator if the main operator of the argument has lower precedence. Precedence is determined from the following list: PAREN, NAM, =, FUNCTION, FLT, PWR, FRT, PRD, =, DVD, ITG, DRV, NEG, =, PLS, =, SUM. Operators separated by = have equal precedence.

The expression resulting from these transformations is saved as the internal tree for the later analysis of light pen references. For example, the expression \( Y - X^2 + \frac{A}{B} \) would be represented by the S-expression

\[
\text{[EQN (ATOM Y NIL) (PLS (PWR (ATOM X NIL) (ATOM 2 NIL) NIL) (ATOM A NIL) (ATOM B NIL) NIL) NIL]NIL)NIL).}
\]

It has the tree structure

```
         =
        /|
       / \
      /   \
     Y    +
    / |   |
   /  |  |
  /   |  |
 /    |  |
/a  /  |
/  /  |
/   /  |
/    /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
/     /  |
```

**Transformations to Facilitate Syntactic Construction**

To a very large extent, mathematical notation becomes, in the mind of the individual, a model of the abstract concepts he is manipulating. For instance, consider the operation of canceling terms from both sides of an equation or from the numerator and denominator of an indicated division, or the operation of bringing a term to the other side of an equation,
or the operation of matrix multiplication. Therefore, the notation must have properties which make it easy to visualize.

Lack of generality is thus introduced in the course of making mathematical notation easier to read. Different mathematical operators should be represented by different sorts of notation and the mathematical operator with the highest precedence should be represented by the most compact form. For example, compare the equivalent Boolean expressions

\[ \{ ((A \lor B) \land C) \lor (D \lor E) \} \land A + B + C + D \lor E. \]

Furthermore, when an operation is applied to complex arguments, it should be possible to visualize it in the same manner as when it is applied to simple ones. For instance, the introduction of parentheses or other such fences allows \((A + B)C\) to be visualized in the same manner as \(DC\), with the operation expressed as concatenation. Sometimes a change of size is used to aid in visualizing a complex argument as a single unit. As a final example, consider the choice of \(f\) to represent the derivative when it is to be thought of as an operation of high precedence, but the choice of \(\frac{df}{dt}\) for the derivative when its properties as a ratio are to be used.

Although it is difficult to make strict classifications, it is useful to distinguish seven distinct pictorial forms used in mathematical notation. As stated above, the more compact forms are often used for operations of higher precedence with the exception that fences and size changes permit the use of complex arguments in the simpler forms; size changes allow complex forms to have high precedence.
The forms in rough order of decreasing precedence, with some examples of each are:

1. supersub \[ U_1' \quad : \quad X^2 \quad : \quad _2 T_3 \]
2. concatenation with variable size fence \[ fn(x, y); \quad X \quad ; \quad \underbrace{a}_b; \quad (a) \]
symbols and separation symbols
3. concatenation \[ ZK; \quad X Y Z \]
4. binding symbol \[ \frac{a+b}{c} ; \int _a ^b xdx \quad ; \quad \sum _{x-3} ^6 X \]
5. hybrid \[ \frac{d^3f}{d xdy} \]
6. concatenation with infix symbols \[ x ^{a y}; a + b + c ; \quad A \land B ; A \cdot B \cdot C \]
7. title \[ (E1) \quad X \]

Since the same mechanisms, such as concatenation, are used in more than one of these seven forms, we will discuss the simpler ones first.

a. Concatenation

A rewrite rule may be associated with any operator. For example, if concatenation is used to represent multiplication then there might be a rule:

\[ (PRD \ A \ B \ NIL) \rightarrow \ (CONCAT \ A \ B \ NIL) \]

The computation is simplified if concatenation is considered to be a binary operator. It is therefore necessary to have recursive rules such as:

\[ (PRD \ X_1 \ X_2 \ \ldots \ \ldots \ X_n \ NIL) \rightarrow \ (CONCAT \ X_1 \ (CONCAT \ X_2 \ \ldots \ (CONCAT \ X_n-1 \ X_n \ NIL) \ \ldots \ NIL)NIL). \]
This rule, however, makes a multiple level structure out of a single level structure and thus destroys the form of the internal tree which is needed to parse the symbols for the display. To remove this difficulty an indicator is put on the property list of the CONCAT form to show that it is not to be considered a node in the internal tree when the symbols are sent to the display by pass four. A pseudo-form, DELIMIT, is introduced to correspond to the old PRD operator. The shape of DELIMIT is just the shape of its argument. The rule above then becomes:

\[(\text{PRD} \ X_1 \ X_2 \ --- \ X_n \ \text{NIL}) \rightarrow (\text{DELIMIT}(\text{CONCAT} \ X_1 \ \text{CONCAT} \ X_2 \ --- \\
\text{CONCAT} \ X_{n-1} \ X_n \ \text{UNDELIMIT}) \ --- \ \text{UNDELIMIT}) \ \text{UNDELIMIT})\text{NIL}\]

b. Concatenation with variable size fence symbols and separation symbols.

In the example fn(X, Y), the size of the comma does not depend on the size of the arguments X and Y so it can be handled by the form CONCAT. On the other hand, the size of the parentheses depends on the dimensions of the enclosed expression and so it is necessary to introduce a parenthesis form. The rule for functions is:

\[(\text{FUNCTION} \ X_1 \ X_2 \ --- \ X_n \ \text{NIL}) \rightarrow (\text{DELIMIT}(\text{CONCAT} \ X_1(\text{PAREN}\text{CONCAT} \\
X_2(\text{CONCAT}\text{ATOM} \ , \ \text{UNDELIMIT}) \ --- \ \text{CONCAT}\text{ATOM} \ , \ \text{UNDELIMIT})) \\
X_n \ \text{UNDELIMIT}) \ \text{UNDELIMIT}) \ \text{UNDELIMIT}) \ \text{UNDELIMIT}) \ \text{UNDELIMIT}) \ \text{UNDELIMIT}) \ \text{NIL})\.

The size of the parentheses is chosen during the dimension pass. They are considered to be symbols introduced by the PAREN form, rather than arguments of the form. The dimension pass also tells the PAREN operator the depth of the parentheses contained in its argument. The PAREN
operator then chooses parentheses, brackets, or braces accordingly. For example, the compiler would produce \[ \left\{ a \cdot \left[ c \cdot (b+d) \right]^2 \right\}^3. \] In the case of transcendental functions, the parentheses are omitted if the argument is a single variable, a product of two variables, or a division.

c. **Binding symbol**

Often a size change is associated with a binding symbol. One might want to make the size change recursive; however, it takes only a few size changes to make the range of sizes too large. The current system uses characters which differ in size by a factor of two and in this case more than one size change is unsatisfactory. Fortunately, expressions occurring in practice would rarely require more than one size change if the recursive rule were to be used. The second pass carries the size with it as it goes out the branches of the internal tree; the rewrite rule for the mathematical operator of any node of the tree may specify that certain branches from that node are to be rewritten using the smaller size. In order to present the rewrite rules, the smaller size has been indicated by a 1 on the rewritten expressions property list. In actual fact, the size is saved on a pushdown list where it can be retrieved by the third pass. An example of a rule involving a size change is the rule for integrals where the expressions for the limits are rewritten using the smaller size:

\[
(\text{ITG } X1 \; \text{(ATOM A NIL)} \; \text{(ATOM B NIL)} \; X2 \; \text{NIL}) \rightarrow
(\text{ITG} \; \text{(CONCAT} \; \text{(ATOM D (UNDELIMIT}) \; X1 \; \text{(UNDELIMIT}) \; )
\text{ATOM A (I) } \text{(ATOM B (I) } \text{X2 NIL}).
\]
d. **Supersub**

This notation is used for exponentiation and for subscripting. The placement of subscripts is a property of the particular variable or function name. The subscripts can be placed at any of the four corners, so that there are 15 possible spatial arrangements. To avoid using 15 distinct but similar operators, null arguments have been introduced. They are treated as picture parts having zero dimensions. When the rewriting pass discovers a NAM operator it picks up a subscript-placement-list associated with the variable or function name in question. This is a list of the form

\[(X_1 \ X_2 \ldots \ X_n),\]

where each \(X_i\) is one of the symbols \(NE, NW, SE, SW\). The \(i\)th argument of NAM is then placed at corner \(X_i\). Two or more arguments at the same corner are separated by commas. Placement of arguments must start at the NE corner and proceed counter-clockwise. When the subscript-placement-list is exhausted the remaining arguments are placed at the SE corner. For example if \(H\) has the subscript-placement-list \((NW\ SW\ SW)\) then:

\[
\text{(NAM X1 X2 X3 X4 (ATOM H NIL)} \ NIL\rightarrow\ \text{(SUPERSUB NIL X1)}
\]

\[
\text{(CONCAT X2 (CONCAT (ATOM , (UNDELIMIT)) X3 (UNDELIMIT))}
\]

\[
\text{(l UNDELIMIT) X4 (ATOM H NIL) NIL).}
\]

If a subscripted variable name is raised to a power the exponent takes the NE position if there are no NE subscripts. The rewrite rule uses SUPERSUB and the operator PWRUP which preserves the original tree by removing the exponent from the SUPERSUB level during the dimension pass and using it for its second argument.
For example:

\[(PWR \ (NAM \ X1 \ X2 \ NIL) \ X3 \ NIL) \rightarrow (PWRUP \ (SUPERSUB \ X3 \ NIL \ NIL \ X1 \ X2 \ NIL) \ NIL \ NIL).\]

Note that this combination of levels must be handled carefully if one desires to use a scheme where duplicate subexpressions are rewritten and dimensioned only once.

c. Concatenation with infix symbols

This form has two possible spatial representations. If it is too wide for one line, then some of the arguments can be placed on additional lines. In order to choose the correct form the arguments must be rewritten and dimensioned and the available line width must be known. The arguments and symbols are concatenated horizontally until a line is filled. Successive lines are then concatenated vertically.

In the case \(A + - B\) the connecting symbol \(+\) can be omitted. Similarly one might want to write \(A \times B\), but \(A(B+C)\), omitting the \(\times\) when the parentheses remove the ambiguity. On the other hand, when the expression is continued on the next line one would still write:

\(A \times (B+C)\) but \(AB\). This problem is resolved by associating with each connecting symbol lists of those operators which cause it to be suppressed.

There is a list of those operators which suppress it from the left, and a list of those operators which suppress it from the right. In addition there is a list of operators which replace the connecting symbol when a new line is started. Before an argument is rewritten and dimensioned its main operator is checked against the above lists so that the connecting symbol may be omitted if necessary.
f. **Title**

The title operator takes as arguments an expression and its name. The name is rewritten, dimensioned, and positioned on the left edge of the display; the expression must then fit into the remaining space.

g. **Hybrid**

Some operators are simply combinations of the others.

In conclusion, rewriting is controlled by the available space, the type of operator, and the size and type of its arguments. The form is dependent on the dimensions only if line width is applied as a constraint.

**Dimensions**

The dimension pass starts at the ends of the branches of the internal tree and works towards the base. At the ends of the branches are the individual variable names. These are either the names of special characters such as $\&, or a sequence of characters to be displayed by the character generator. The names of special characters are marked with a flag and have their dimensions associated with them. The width of sequences of generated characters must be computed using their size and the fact that in the SAL language used for display $\$ and $\& are mapped into character generator case shifts and $/ is used as a quote character, and also the fact that some generated characters are non-spacing.

Five dimension functions must be written for each form. There are functions for the width, height, depth, symbols, and arguments. The compiler uses these functions to put the dimensions of each picture part on its property list. The three functions for the width, height, and depth
have a list of the dimensioned arguments of the form as input. They
use helping functions to retrieve the dimensions from the property lists
of the arguments. The position of an argument or symbol is specified
by the coordinates of the lower left hand corner of the rectangle containing
it. The function for the arguments has as value an S-expression of the
form \((\Delta X_1 \ \Delta Y_1 \ ... \Delta X_n \ \Delta Y_n)\) where \(\Delta X_i\), \(\Delta Y_i\) is the position of the \(i^{th}\)
argument with respect to the lower left hand corner of the circumscribed
rectangle. The function for the symbols has as output an S-expression
of the form \((\Delta X_1 \ \Delta Y_1 \ \text{name}_1 \ \text{size}_1 \ \text{width}_1 \ ... \ \Delta X_n \ \Delta Y_n \ \text{name}_n \ \text{size}_n \ \text{width}_n)\)
where the \(i^{th}\) symbol which is added to the display by this form has name
\(\text{name}_i\), size \(\text{size}_i\), width \(\text{width}_i\), and relative position \(\Delta X_i \ \Delta Y_i\). The functions
for the symbols and arguments have as input the width, height, and depth,
as well as a list of the dimensioned arguments. The computation of the
dimensions and relative positions by independent functions breaks down
when symbols such as parentheses must be chosen, as it is not desirable
to repeat this choice for each of the five functions. To solve this problem
a mechanism is set up by which the width function, which is executed first,
can communicate its choice to the others.

Each of the forms currently in the system is pictured in the following
figures. Dotted lines have been drawn to indicate how the arguments and
center lines are placed. It might be well to note that the human eye is
sensitive to even the smallest misalignments.
Figure 1: CONCAT

Figure 2: SUPSUB

Figure 3: EVL
Figure 4: SUM

Figure 5: PAREN

Figure 6: LVCONCAT

Figure 7: TITLE
Figure 8: VCONCAT

Figure 9: DVD

Figure 10: ITG

Figure 11: ABS
Output to the Display

The LISP picture compiler program runs in Project MAC time sharing. After an expression has been dimensioned, the compiler outputs a description of the picture to be displayed over the dataphone to PDP-6 LISP. The picture is displayed on the PDP-6 scope using the SAL and MACROSAL languages. Each expression picture is set up as a single MACROSAL object with the picture tree marked by left and right pseudo-parentheses.

The compiler starts at the base of the tree and goes out the branches, taking them in left to right order. The compiler sends each symbol to the display as it is encountered. If two successive symbols are not adjacent, then the compiler links them with a non-displaying relative vector. Whenever the compiler encounters a new node of the internal tree it follows the last symbol sent with a left pseudo-parenthesis. After the compiler has finished this node and all branches extending from it the corresponding right pseudo-parenthesis is sent. This information is also sent to the disk so that the picture can be reconstructed without being recompiled.
Example

The following figure is a photograph of two displayed expressions.

The source expression for E1 is:

(EQN(PWR(NAM O(R OMEGA NIL)NIL)(A NIL))
(PLS(PRD(PRD 2 6N(PWR PI -1 NIL)NIL))
(PRD(PLS PI OMEGA NIL) (PWR(PLS 4A 1 NIL)-1 NIL)NIL)NIL)
(PRD(PRD 2 6A OMEGA (PWR PI -1 NIL) (PWR(PLS PI OMEGA NIL) (PRD -1 6A
NIL)NIL)NIL)
(ITG (T O PY (PRD(NAM I(THETA 4T NIL)NIL))
(NAM(PLS 6A -1 NIL) (M(PRD OMEGA 4T NIL)NIL)NIL)NIL)NIL)

The source expression for E2 is:

(O(SUM 6V(PRD 2N(PWR 2 -1 NIL)NIL))
(PLS 6N -2 NIL) (PRD(ABS(NAM(PLS 6V 1 NIL)B NIL)NIL))
(PRD(PLS(PLS(PLS 6N(PRND -1 6V NIL)NIL))
(PLS(NAM(PLS 6N(PRND -1 6V NIL)-1 NIL)P NIL)
(PRND -1NAM(PLS 6N(PRND -1 6V NIL)-2 NIL)P NIL)NIL)NIL)
(NIL) (NAM(PLS 6N(PRND -1 6V NIL)-2 NIL)P NIL)NIL)
(PLS PI OMEGA NIL) (NAM(PLS 6V 1 NIL)
THETA NIL) (PWR(PLS 6V 1 NIL)-2 NIL)
(PWR(PLS 6N(PRND -1 6V NIL)NIL)-1 NIL)NIL)NIL)NIL)NIL)
\[ P_w(w) = \frac{2 \cdot 1}{\pi} \cdot \frac{\log w}{(a+1)} + \frac{2 \cdot a \cdot w}{\pi \cdot (\log w)} \cdot \int_{-\infty}^{\infty} \theta(t) \cdot M \cdot L \cdot t \cdot dt \]

\[ 0 \left( \sum B_i \cdot \frac{[(n-v) \cdot (P_r - P_s) + P_s]}{(v+1) \cdot (n-v)} \right) \cdot \log(v+2) \cdot \theta \]
Conclusion

One measure of program complexity is the extent and predictability of the combinations of independent parts of the input data which must be made in order to determine the flow of program control. In this respect the display of mathematical expressions seems to lie midway in difficulty between the display of lines of English text and the display of arbitrary graphs. In the display of English text, the data is organized in a string, the only global property required is a character count. Hyphenation decisions are based on the word at the end of the line. On the other hand, the display of graphs seems to require the simultaneous positioning of several nodes, and might require a character algebra or an iterative approach. In the display of mathematical expressions, decisions can be made at each node of the expression tree, based on information collected from above and below. The central importance of this tree structure makes LISP a convenient language for the program.
Chapter X
LINEAR INPUT OF MATHEMATICAL EXPRESSIONS

R.W. Floyd describes an algorithm for parsing a language with a precedence grammar. Floyd defines an operator grammar as one with no productions containing adjacent non-terminal symbols. A precedence grammar is then an operator grammar in which given two ordered terminal symbols, one either appears before, at the same time, or after the other, in all series of productions leading to a string in the language. The productions shown in Figure 1 will, with three exceptions, produce the desired input strings as a subset of all the strings which they produce. With the exception of N, which stands for any digit, and L, which stands for any letter of the alphabet, the letters are the non-terminal symbols. The initial non-terminal symbol is V. The productions 10, 12, 29 and 42 all represent classes of productions formed by taking any number of S's or P's, separated by commas.

To verify that these productions form a precedence grammar, the table shown in Figure 2 is formed. For any two terminal symbols T₁ and T₂, T₁ is >, <, or ≥ T₂, if it appears before, after, or at the same time as T₂ in a series of productions leading to a string of terminal characters, and in which T₁ and T₂ are at some point in the series, either adjacent or separated only by non-terminal characters. Floyd gives a systematic way of checking a proposed set of productions for these relations. In Figure 2 it can be seen that only one of these relations holds for each ordered pair of

225
terminal symbols, and so the productions do form a precedence grammar and a string formed by these productions can be parsed using Floyd's algorithm.

These productions produce strings which differ from the strings to be input in three ways, and so the input strings must be preprocessed. In the input strings, the unary minus is written the same as the binary one, the preprocessor recognizes it by the immediately preceding left parenthesis and changes it to theta, in order to conform with productions 20 and 43. Similarly, the quotes, ", which precede a left parenthesis are changed to double quotes, ", to agree with productions 3, 7, 42, 43 and 24. Finally the string !B, N typed at the PDP-6 is changed to BIN by the PDP-6 as it picks up the subexpression address integer N from displayed expression B.

To the parsing algorithm must be added a description of how meaning is assigned to the parsed expression. As the input string is parsed, the non-terminal symbol on the left side of each production discovered is represented by the LISP expression which is the meaning of that production. For example X + Y would be identified as an instance of production 13 and changed to (PLS X Y). The steps in parsing X + C * D would be:

\[
\begin{align*}
X &+ C * D \\
X &+ (PRD C D) \\
(PLS X (PRD C D)) 
\end{align*}
\]

The function which forms the LISP expression associated with each production is found on the property list of the left most terminal symbol of the right side of the production. Sometimes, the same terminal symbol is left-most in several productions, but the production in which the symbol lies
can then be determined from the other terminal symbols in the production or the total number of symbols in it. Productions 33 through 41 are tested for separately, so that a function will not have to be placed on the property list of each letter and digit. The method of assigning meaning to these productions is also different, for they must be transformed into LISP atoms and numbers. Corresponding to the meaning assignments are the rules:

\[
\begin{align*}
L & \rightarrow \text{(ATOM L)} \\
N & \rightarrow \text{(ATOM N)} \\
L(\text{ATOM...}) & \rightarrow \text{(ATOM L...)} \\
N(\text{ATOM...}) & \rightarrow \text{(ATOM N...).}
\end{align*}
\]

Then, a post-processing function is applied to the final parsed expression and the subexpressions starting with ATOM are replaced by the LISP atom or number made from the characters in the subexpression. The remaining meaning assignments are shown in Figure 3.

A flow chart of Floyd's parsing algorithm is shown in Figure 4. This chapter will be concluded with a description of the LISP program to perform this algorithm. The top level function APARSE accepts the list of input symbols, which is in reverse order, reverses it, and gives it to APARSE1 for the preprocessing step. APARSE then parses this preprocessed string, using the PROG variables U, V, P, R, and S for temporary results. U is the list of symbols to be parsed. The answer is built up on P. NIL is used for the left and right end terminal symbol. In brief, terminal symbols on P are combined into phrases, i.e. non-terminal symbols, and these
phrases are combined with other terminal symbols into larger phrases
till there is only one phrase, the result, between the two terminal NIL's.
The first step is to put the right terminal NIL on P. Then, the next symbol,
R, is taken from the input list U. When U is empty, the left terminal NIL
is taken. The next step is to look down the list of phrases and terminal
symbols on P, calling the terminal symbol under consideration S, until one
is found with higher precedence than R. This means that S belongs to a
non-terminal phrase which came down after the phrase containing R. Control
is sent to B, where all terminal symbols in this phrase which have equal
precedence with S, and the intervening phrases are determined. These
symbols and phrases are placed on V, where they are combined into a phrase
by a function associated with S. This phrase replaces its parts on P and
the comparison of symbols on P with R is resumed. If S is not greater than
R, then a check is made to see if R is NIL, in which case the process is
complete. If R is not NIL, then R is added to P and control goes to A,
where a new input symbol R is taken and the old R becomes S, since it is now
at the front of P. When this process is finished, the final phrase is operated
on by APARSE2 which combines the ATOM phrases into LISP atoms and
numbers. Description of APARSE2 will be deferred until the functions which
create the phrases have been described. As stated above, a function is
associated with each terminal symbol. Numbers, POINT, and literals use
the function PFN5. If the number or literal has no phrase immediately to
its right, then the structure (ATOM N) is created by PFN5. Otherwise
the phrase will be of this form and the new symbol is added to it.

(N (ATOM X ...) ) becomes (ATOM N X ...). Most of the other
operators use the function BPFN which creates a binary prefix structure
from a binary infix one. \((X, \text{op}, Y)\) becomes \([\text{op} X \ Y]\). The exceptions
are \((X \ \text{NEG} \ Y)\) which becomes \([\text{PLS} \ X \ (\text{PRD} -1 \ Y)]\), \((X \ \text{LBK} \ Y, Z, \ldots \ \text{RBK})\)
which becomes \([\text{NAM} \ Y \ Z \ldots X]\), and LP. For LP there are three cases,
\((\text{LP} \ \text{THETA} \ X)\) becomes \([\text{PRD} -1 \ X]\). \((\text{ATOM} \ldots) \ \text{LP} \ X, Y \ldots \text{RP})\)
becomes \((\text{ATOM} \ldots) \ X, Y \ldots\). Finally, the subscripted function case,
\((\text{NAM} \ldots \ F) \ \text{LP} \ X, Y \ldots \text{RP})\) becomes \([\text{NAM} \ldots \ F \ X \ Y \ldots]\). The
function PFIN is used to remove the commas in case two, and the function
PFN2 is used to remove the commas in case three, placing the function
structure at the end as the NAM body.

In conclusion, APARSE2 will be described. APARSE2 looks for the
structures \((\text{ATOM} \ A \ldots)\) which it changes into a LISP number if the first
character \(A\) is an integer or a LISP atom if \(A\) is a literal. Literal atoms
are formed by PFIN which uses the LISP character handling functions.
Numbers are formed by PFN4, the number is built up as the first member
of the list of digits, the second member being combined with it until all
digits have been used or until POINT is reached, indicating that the number
is floating point. The digits to the right of POINT are combined into a
floating point fraction by PFN4I and then added to the integer part.
1. \( V \rightarrow C \)
2. \( V \rightarrow R \rightarrow C \)
3. \( C \rightarrow 'a'(S) \)
4. \( C \rightarrow S \)
5. \( S \rightarrow P \rightarrow P \)
6. \( S \rightarrow P \)
7. \( P \rightarrow 'a'(D) \)
8. \( P \rightarrow D \)
9. \( D \rightarrow E \)
10. \( D \rightarrow \langle S, S \rangle \) etc.
11. \( D \rightarrow X(S) \)
12. \( D \rightarrow X \langle S, S \rangle \) etc.
13. \( E \rightarrow E + F \)
14. \( E \rightarrow E - F \)
15. \( E \rightarrow F \)
16. \( F \rightarrow G \)
17. \( F \rightarrow F \Theta G \)
18. \( F \rightarrow F/G \)
19. \( G \rightarrow H \)
20. \( G \rightarrow \langle QH \rangle \)
21. \( H \rightarrow H + I \)
22. \( H \rightarrow I \)
23. \( I \rightarrow X \)
24. \( I \rightarrow 'a'(D) \)
25. \( I \rightarrow (D) \)
26. \( I \rightarrow Y \)
27. \( I \rightarrow R \rightarrow Y \)
28. \( X \rightarrow R[F] \)
29. \( X \rightarrow R[F, \Theta] \) etc.
30. \( X \rightarrow R \)
31. \( R \rightarrow 'a'B \)
32. \( R \rightarrow B \)
33. \( Y \rightarrow .Y \)
34. \( Y \rightarrow N \)
35. \( Y \rightarrow NY \)
36. \( B \rightarrow L \)
37. \( B \rightarrow LA \)
38. \( A \rightarrow L \)
39. \( A \rightarrow N \)
40. \( A \rightarrow LA \)
41. \( A \rightarrow NA \)
42. \( I \rightarrow 'a'(S) \) etc.
43. \( I \rightarrow 'a'(QH) \)

**Figure 1**

Productions for the Input Grammar
**Figure 2**

Precedence Table for the Grammar
Fig. 3

Assignment of Meaning to the Productions
Figure 4

Floyd's Parsing Algorithm
Chapter XI

TWO DIMENSIONAL INPUT OF MATHEMATICAL EXPRESSIONS

In the preceding chapter, a mathematical expression represented in list structure form was changed into a text book style expression through the application of a series of rules. First, the order of arguments was rearranged and prefix operators became infix operators. Then, the expression was expressed in terms of two dimensional forms. Finally, the exact size and position of these forms and their corresponding symbols was determined. In this chapter, a method by which this process could be reversed will be suggested. Conditions under which the size and position of individual symbols can be used to parse an array of symbols will be given.

Klerer (4) has developed a special typewriter which allows the carriage to be spaced up and down as well as left and right. Characters typed on this machine are placed in a two dimensional array in core memory. Klerer has an algorithm for translating this array into a linear string. It would also be possible to generate characters on the display face. Teitleman (1) has demonstrated a program which recognizes individual hand printed characters. This program could be used to recognize a string of printed characters by assuming that all overlapping strokes form a single character. The program would remain several strokes behind the user, assuming that no character has more than say five strokes and that all of the strokes for a given character are made in sequence. Characters consisting of disconnected parts will be parsed together in the parsing program.
When the character is recognized, it will be replaced by a standard character in one of several standard sizes and moved into the nearest cell of an array on the display face. The user can erase and correct any false characters. The order in which the characters are generated can be saved as an heuristic aid in parsing. The characters are positioned in a matrix on the display face for two reasons. Positions such as $A^B$ must be interpreted either as $AB$ or $A^B$, the user can correct a false interpretation before parsing begins; furthermore, if characters are allowed to have any coordinates on the scope face, then the parsing program is given redundant information. The parsing program would have to consider many character positions as equivalent. This means that it does not have the possibility of terminating a string by finding no character in a given position.

The first step in parsing is to transform the two dimensional array of characters into a one dimensional string. To do this, information about the relative positions of the characters must be used. Since adjacent characters are often grouped together, it is efficient to sort the characters into a two dimensional array. Klerer uses an array with one cell for every possible typewriter position. Another possibility is to use a list structure array, with the coordinates of each character as its properties. This would take less storage if Klerer's array is nearly empty. To get the characters into a one dimensional string, it is necessary to map groups of characters with a given vertical relationship into a given horizontal relationship.

For instance, the characters comprising the numerator of a division will occur before the characters comprising the denominator. Klerer identifies
these groups of characters in expressions containing subscript, superscript, and binding symbol forms by imposing some constraints. First, the symbols which lie in a given area with respect to a binding symbol are taken to comprise an argument of that symbol. For example, the symbols in the numerator of a division must lie over the division bar. Second, the allowed position of input symbols is constrained so that a top to bottom, left to right scan will encounter the symbols in a canonical order, so that every binding symbol can be found before the symbols in its arguments. Finally, superscripts and subscripts must occur in a given position with respect to the base, and occur on the right only, so that the base will be scanned first by the left to right scan.

These are not very severe constraints; but it would be possible to relax them so that the case \( \frac{a + b + c}{d} \) could be handled. To do this, note that the forms can be ordered:

CONCAT
SUPSUB
LVCONCAT
VCONCAT
BINDING SYMBOL
TITLE

In addition, the binding symbols can be partially ordered. For example, if two division bars occupy the same horizontal space, the longer dominates the smaller. Once this is accomplished the one dimensional string is formed. First, all symbols related by CONCAT are found and replaced by a string with
the corresponding position and dimensions. The position and dimensions are found from the definition of the CONCAT form in the last chapter. Next, all SUPSUB's are formed in the same way. If no SUPSUB's are found then all LVCONCAT's are formed. However, if some SUPSUB's are found, then CONCAT's are considered again. In this manner, each higher level form is considered. When binding symbols are considered, the form for each lowest ranking member in each group of ordered binding symbols is constructed. This process is continued until all the symbols have been converted into one linear string. Any form which can be defined for the previous chapter can be treated in this manner. Of course size must be used to distinguish A from B. Context dependent forms, however, cannot be handled. For example, in \((a)\) the a and b would not be recognized to be in the VCONCAT relationship unless the vertical separation was near that specified in the last chapter. However, this relationship is implied by the large parentheses.

As stated, the above parsing scheme could be quite inefficient, since some forms might be considered several times before they were combined. It should be possible, however, to consider the expression in a systematic way, following the tree structure in most cases.

Once the expression has been converted to a linear string, it can be parsed by the algorithm given earlier for the typewriter.

Robert Anderson at Harvard is working on both a formal definition of the two dimensional syntax and a parsing scheme. He has independently suggested ideas similar to those in this chapter which he hopes to carry much farther.
Chapter XII
THE SYSTEM PROGRAMS

Three system programs were constructed as a basis for the mathematical laboratory: Time Sharing LISP, the PDP-6 LISP dataphone link, and the PDP-6 LISP display language. The development of these systems required an additional expenditure of programming effort and their use in the finished system results in some loss in both space and speed. Therefore, the description of these systems in the following chapters will be introduced by a discussion of the purpose they serve in the research and some of the factors which guided the many decisions, large and small, required for their construction.

In brief, the purpose of programming research is to investigate the structure and behavior of complex systems. Legitimate results of this research are working programs which demonstrate the realization of certain levels of technology in a given area, or program concepts which have been found useful for system description or for programming certain processes.

In practice, the development of working programs has often hinged largely on whether sufficient high calibre effort, properly organized, could be expended, rather than on the discovery of major ideas. System programs allow workers in related areas to help each other by working together to develop a system which will be used by all. Working alone results not only in the duplication of effort in writing programs to perform certain basic
functions, but results in a myriad of programs which are completely incompatible because of numerous trivial differences. Voluntary collaboration may be initiated by writing a system program and offering it to the public. If this is done, then it is necessary to give the public what it wants. This is borne out by the acceptance of these systems to date. The potential users of time-sharing LISP had goals similar to the writer. The system had only one objectionable new feature, the method of writing files, while providing many new conveniences. Several people contributed additional improvements and after two years Moses and Fenichel rewrote it completely to form an excellent system which was used in this thesis. Three people have used the dataphone program; they felt the LISP system took up too much space in the PDP-6, so they expended considerable effort to provide a new environment for the program and then never did anything with it. They never got around to improving the reliability of the dataphone program, which would have benefitted the author. Occasional malfunction of the dataphone program reduced acceptance by potential casual users. Acceptance of the display language at Project MAC among the more recreational programmers was retarded when a version of PDP-6 LISP with many small improvements superceded the version in which the display language was originally imbedded. Not enough care was taken in clearly delineating the steps required to move the embedded language to a new version. Also, some of the conventions, although useful, were too controversial. Considerable interest has been shown by owners of other PDP-6 machines, who would like some display facility. Also, the
embedding of I-O routines in LISP has become more popular.

System programs can also be an aid to the individual programmer. Important research programs can take a long time to construct. This leads programmers to plan far more than they can accomplish, to do work which has an incorrect emphasis in the light of the state of the art when the work is finished, and to lose time moving from one machine to another. If a programmer builds a system program as he builds his basic routines, then he is less likely to make the careless decisions which leave him inflexible with respect to the above difficulties. The important thing is not to undertake much more work because you are building a system, but to do what you do in a form suitable for a public system.

System programs are also useful in the discovery and communication of important concepts. In his recent thesis, Teitelman describes the PILOT system, which lets the programmer alter the effect of his procedures by inserting methods for handling special cases which arise as the program is used. In handling each case separately, the programmer must be careful, or he will neglect to search for a general concept which will appropriately describe all of the cases and which can be communicated to others. A programming language is a good method of communicating the many useful ideas needed to treat a new problem area.
Chapter XIII

PDP-6 LISP INPUT-OUTPUT FOR THE DATAPHONE

This chapter describes the PDP-6 LISP interface with the routine which sends and receives characters over the dataphone.

The dataphone operates in sequence break mode, requesting service each millisecond. It must be turned on and off with the LISP function dp; however, when it is running, input and output to it are handled in the same manner as to the teletype, paper tape reader, and paper tape punch. The dataphone is controlled by the control characters:

A Write on dataphone.
C Do not write on dataphone.
D Read from dataphone.
E Do not read from dataphone.

Control characters can be typed from the teletype or executed with the function cchar. Since CTSS truncates lines longer than 72 characters, a carriage return is inserted whenever 72 characters have been output to the dataphone since the last carriage return. If the user sends the ASCII character % through the dataphone program, it will be sent as the CTSS quit character, similarly ASCII & is interpreted as the CTSS interrupt character. There are several LISP functions written in machine language which are useful for dataphone programming:

dp (x) If x is NIL the dataphone is turned off;
otherwise it is turned on and initialized.

241
dpd (x)  
The value of dpd (x) is a flag for  
characters in the dataphone input buffer.  
If x is NIL, dpd returns NIL if there are  
no characters; otherwise, it waits until  
some arrive. When characters are  
received this function returns a list of  
the mode of the characters and the number  
of characters. The mode is a number  
determined by the sender. Characters  
sent from CTSS in the same manner as to  
a normal console are assigned mode 1.

ttyd { }  
Returns the number of characters in the  
teletype input buffer or NIL if there are  
none.

cchar (x)  
Executes the lower case ASCII character  
x as a control character.

To establish communication with CTSS, turn on the dataphone by  
executing dp (T). Then execute cts ( ) and type S followed by a space.  
The teletype can now be used as a CTSS console. Several useful features  
of the LISP functions cts, dwrite, and dpread are described below.  
LISP expressions for these functions are at the end of the chapter.
dpread { }

This function waits until there is input from the dataphone. It then prints the characters on the teletype one by one. If it receives $, it does not print this, but instead evaluates the S-expression which follows.

dpwrite (x;y)  

dpwrite takes two forms of input. If y is NIL dpwrite assumes that x is a list of atoms and sends over the dataphone the characters in the PNAMEs of these atoms, with a space between each PNAME. Otherwise it sends x as an S-expression.

css { }

ctss allows the user to operate the PDP-6 as a CTSS console. There are two modes; local and send. ctss is initially in local mode. In this mode it accepts S-expressions for eval but watches for the single atom S-expressions S, STOP, and L. If it finds S it goes into send mode. If it finds STOP it terminates returning NIL. If it finds L it sends over the dataphone the two S-expressions which are typed next,
but stays in local mode.

In send mode it sends characters one
by one while watching for @ and #.
@ returns it to local mode. If it sees
# it sends the evaluation of the S-
expression which follows and then types
#.
A System for Display Language Construction

The system provides a language macrosal for the generation of picture parts called objects. An object can be any combination of points, lines and characters. An object is generated by calling the function macrosal [NAME; DESCRIPTION] described in the second section. The most common way to describe an object is to establish a set point. The set point is established utilizing PARAMETER, LOCY, and LOCX statements. The exact format of these statements is discussed below. An object description is terminated with a STOP statement. If NAME is T, the current description will be appended to the description of the last object generated. The first example, disp, generates a large object in this manner.

The user communicates with a display through a light pen. As the light pen sweeps across the screen, its trajectory can be used in many different ways. For example, it may be used to determine a point, a set of points, or a line. Or if a subpart of the display has been defined as an object, the trajectory may be simply interpreted as a pointer to this object or a point on the object. The LISP functions embedded in the display language facilitate acquiring the data needed to make these different levels of interpretation.

One problem in utilizing the light pen is to determine when it is near the screen and not just being moved into place. This is solved by using the
width of the field of view of the pen as measured by a tracking cross. This width decreases as the pen approaches the screen and a center dot is displayed whenever the field of view is less than a certain prescribable limit. This pen distance is available to LISP.

The approach of the pen to an object is considered significant. Just how close the pen must come before being noticed is a program variable. Its most recent position within this distance is recorded; in addition, so are the last 5 such positions, each at least a prescribed distance from the preceding one. This represents a crude way of gradually forgetting the details of the past. It is also possible to get the current coordinates of the pen, obtain a list of all objects currently seen by the pen, to report when the pen sees an object with a name other than a given name, or to require an object to move with the pen.

The example function `sketch` is a LISP function using several of these features. The function uses a subroutine to display five different light buttons. If the light pen is held near one of these buttons, the tracking cross will be centered about the point where the pen is seen. The program interprets pointing at these buttons to mean 1) draw a line, 2) move a line, 3) delete a line, 4) suppress the cross and 5) return control to the teletype. To draw lines touch the first button, a new line will then be drawn whenever the pen leaves the screen and then returns. This process is terminated whenever the pen returns near one of the light buttons. Additional LISP functions could be written to expand `sketch` into a program similar to SKETCH PAD.
Often one wants to communicate to the display certain basic forms which are to be used in constructing larger units. These inputs might be a set of letters which are to be parsed into a sentence; or they might be a set of circuit elements or music symbols. In SKETCH PAD this is done by indicating the type of form and then moving the light pen so that the parameters of the form, the end points of a line segment for instance, can be abstracted from the trajectory. Alternatively, one could abstract both the type and the parameters from the trajectory. The example program argus uses a method developed by Teitelman which enables the user to teach the machine to replace a single line by a known form. A line is a single movement (however complex) on the surface of the display without lifting the pen. The LISP data structure is convenient for storing properties of the forms to be recognized.

The parsing of very large displayed expressions, such as LISP S-expressions for example, can be difficult for people. Furthermore, there may be alternative parsings. People can be aided by intensifying, upon request, grammatical subexpressions or sub-objects containing referenced segments or by providing additional displays meaningfully related to the first display. These might be rotated views of an object or shaded objects. Further development of the system is needed in this area.

It is the task of the programmer to organize a program and data base in such a way that the most needed inputs to a machine will have short representations and the most needed computations will be efficient. The combinatorial aspects are such that this must be done through a series of
levels of concepts. An important point is that it is not possible to complete
an entire level at a time. The most useful concepts at a given level only
become clear with the exploration of higher levels. The exploration of
higher levels without intermediate concepts is, however, almost impossibly
tedious. It is important in an experimental situation to have a system where
one can make changes to any desired depth and provide for the irregular
growth and reorganization of the data base.

In the present display system, experimentation will probably indicate
that new statements for macrosal are needed, or that certain objects occur
so often that more programming effort could well be spent in generating
them efficiently. To provide for these possibilities macrosal has been
programmed as a syntactic extension of the scope assembly language, sal.
Sal is a LISP function which creates objects from lists of octal numbers. It
is described in detail in part II. Provision has also been made for the
addition of machine language subroutines which alter the objects as they
are displayed. Furthermore, the system is organized so that no statements
need be made about features of the display language which are not needed.

By embedding these display facilities in LISP one makes available a
wealth of mechanisms which have proved useful in the analysis and generation
of language and in the development of systems which can be increased
incrementally in complexity.
Implementation

The use of LISP in this system has two distinct disadvantages. First, it is not possible to interrupt the LISP system at any point in time and immediately employ its full power. It may be in the midst of garbage collection. Garbage collection with the current version of PDP-6 LISP requires a noticable time. Second, the data types are too limited. It is not convenient to set up the type of list structure used in SKETCH PAD, but this can be approximated. A serious problem is the inability to set aside blocks of registers to contain display instructions and to store information about light pen actions.

To get around these problems a fixed buffer of 2048 words has been set aside for description of the display. All communication between LISP and the display goes through this buffer. This buffer contains two kinds of data structures; display lists and headers. One header is associated with each display list, which is a list of half word commands for the display. The headers build down from the top of the buffer and the display lists build up from the bottom.

During display an interrupt routine cycles through a dispatch table. Dispatches can occur to a pen track routine, a routine which displays the contents of the display buffer, a line drawing routine, and a routine which terminates the display.

The pen track routine displays a cross as was described earlier. In a crude effort to give the routine enough display time each cycle, it is called between display of each object in the display buffer. Use of the clock would
be better.

The format of the headers in the display buffer is as follows:

<table>
<thead>
<tr>
<th>NAME</th>
<th>Pointer to an atom</th>
</tr>
</thead>
<tbody>
<tr>
<td>A = ON-OFF bit</td>
<td>(display list length) pointer to start of display list - 1</td>
</tr>
<tr>
<td>B = Tracking Cross jump bit</td>
<td>Pen hit distance Pen hit count</td>
</tr>
<tr>
<td>C = Move with pen bit</td>
<td>Subr address A B C</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Most recent Y</th>
<th>Most recent X</th>
</tr>
</thead>
<tbody>
<tr>
<td>Y</td>
<td>X</td>
</tr>
<tr>
<td>Y</td>
<td>X</td>
</tr>
<tr>
<td>Y</td>
<td>X</td>
</tr>
<tr>
<td>Y</td>
<td>X</td>
</tr>
<tr>
<td>Y</td>
<td>X</td>
</tr>
</tbody>
</table>

Figure 1 Header Format

Each display list has a name which is kept as the first word of its header. When an object is referenced by a LiSP function, the headers are searched for one with the name mentioned.

The interrupt program cycles through the headers. It picks up a pointer for a BLKO instruction from the right half of the second word of each header. This BLKO is terminated by a STOP instruction at the end of
each display list. If the subr address is not 0, then the subr at this
address will be executed when the STOP is reached.

If the light pen is seen during display of some display list, control
goes immediately to an interrupt program. Several conditional branchings
can occur within this program. The interrupt program first re-displays
the display list as a check against light pen noise. If the pen is seen a
second time the pen cross movable bit is checked. If this bit is a 1 the
pen tracking cross is centered about the point where the light pen was seen
and the rest of the display list is displayed. Otherwise, a check is made
to see if the center of the pen cross is within a specified minimum distance
of the point seen. If not, the display list is continued with the light pen
reenabled. If the pen is close enough, the coordinates of the point seen
are stored in the fifth word of the header. The last five words of the header
contain the coordinates of points seen by the pen in the past. Each of these
history points is at least a specified distance from the preceding one.
This distance is in the left half of the third word of the header. When a
new point is seen, it is added to the history points if it is far enough away
from the one most recently stored or if there are none. The number of
history points is kept in the right half of the third word of the header. This
number can be set to zero by LISP. Whenever there are more than five
history points the oldest one is lost. After the point seen has been
appropriately stored, a check is made to see if the object should move with
the pen. If bit C is set, a pointer to the display list is transmitted to the
pen track routine. The display list is then finished.
Display lists are put into the buffer by the sal language. If a line is to be drawn with the pen, a set point for the line is created with the sal language. Its display list is then incremented by the line drawing routine. This incrementing is terminated when the pen leaves the screen.

**Display Functions**

makrosal \([X; Y]\) makrosal interprets its arguments and then calls sal on the result. X is the first argument for sal. Y is a description of a display list for sal. Y is a list of lists, each of which is a macro. The first word of each macro is an atom which has under the property MACROSAL a function of one argument. makrosal gets this function and applies it to the remainder of the macro list. The result of this function is a list for sal beginning and ending in mode 1. This list is prefaced by two integers which give respectively the mode that the display will end up in and the mode in which it must begin, if it is to interpret this list correctly.

makrosal appends the successive macro expansions, using the prefacing integers to create the proper linking of modes.
The following macro word formats are in the system:

1. (PARAMETER penenable scale intensity)
   Normally the first word in a set point, this statement creates a parameter word. If penenable, scale, or intensity is NIL, the corresponding field of the parameter word is not enabled.

2. (LOCY n)
   Creates a non-displaying Y point word which sets the scope Y coordinate to n. This is normally the second statement in a set point.

3. (LOCX n)
   Creates a non-displaying X point word which sets the scope X coordinate to N. This is normally the third and last statement in a set point.

4. (STOP)
   Creates a parameter word with stop enabled. This is normally the last statement in a display list.
5. (ISTOP)

Creates a parameter word with stop enabled and a flag that the previous word has a breakout bit. The use of ISTOP is explained in the description of SAL which follows.

6. (LOCYD n)

Creates a displaying Y point word which sets the display Y coordinates to n and displays a point.

7. (LOCXD n)

Creates a displaying X point word which sets the display X coordinate to n and displays a point.

8. (LINE \( X_1 \ Y_1 \ldots X_n \ Y_n \))

Creates a sequence of non-displaying line segments from vector words. The display starts at the last point displayed or set by LOCX, LOCY. Each increment has size \( X = X_{i+1} - X_i \), \( Y = Y_{i+1} - Y_i \).

9. (LIRED \( X_1 \ Y_1 \ldots X_n \ Y_n \))

Creates a line like LINE, but displays it.
10. \( \text{LONGLINE } X_1 \ Y_1 \ X_2 \ Y_2 \)  
   
   Creates a non-displaying vector 
   continue word. A line will be drawn from 
the current display coordinates to the edge 
of the display. The slope of the line is 
\( (Y_2 - Y_1)/(X_2 - X_1) \).

11. \( \text{LONGLINED } X_1 \ Y_1 \ X_2 \ Y_2 \)  
   
   Creates a line like \text{LONGLINE}, but 
displays it.

12. \( \text{CHAR } X_1 \ldots X_n \)  
   
   Creates a display of character made 
words which are the PNAME's of the atoms 
\( X_1 \). No spaces are inserted between the 
PNAME's.

\( \text{sal } [X;Y;Z] \)  

X is the name of the display list to be 
created or T. If X is T, then this list is 
appended to the last one created. If the 
previous list ends in \text{STOP} (30000_{\text{g}}), the 
\text{STOP} is removed. If it ends in \text{ISTOP} 
\( (4030000_{\text{g}}) \), it not only removes this but 
zeros the breakout bit \( (4000000_{\text{g}}) \) in the 
previous half word. \( Y \) is a description of 
the list to be created. There are two forms 
for the elements of \( Y \) corresponding to two
modes for the assembly function \texttt{sal}

The function is initially in mode 1.

In mode 1 \texttt{sal} removes lists of atoms, two at a time, from \texttt{Y}. Each atom is a number or is bound to one on the dotted pair list \texttt{Z}. For each pair of lists \texttt{sal} forms one display half word instruction by shifting the numbers on the second list the number of places specified by the corresponding number on the first list. When it encounters \texttt{NIL} on \texttt{Y}, it goes into mode 2. In mode 2 \texttt{sal} takes numbers or non-numerical atoms one at a time from \texttt{Y}. Each number is a display half word. If a non-numerical atom is encountered, \texttt{sal} looks at the previous number to see if it puts the display into increment or character mode. If in character mode, \texttt{sal} assembles the PNAME of the atom as characters. If in increment mode, \texttt{sal} gets a list of full words off the atom's property list with the indicator SCHAR. It assembles these as increment mode half words. Consecutive non-numerical atoms are assembled together in the same mode.
When \texttt{sal} finds NIL, it returns to mode 1.
Mode 1 is more flexible and mode 2 is
more economical. NIL is not a legal
object name.

**More Functions**

In the description below S stands for the name of an object with a set
point. W stands for the name of an object with or without a set point.

- \texttt{sx [S]} \hspace{1cm} Returns the current x coordinate of S.
- \texttt{sy [S]} \hspace{1cm} Returns the current y coordinate of S.
- \texttt{smv [S]} \hspace{1cm} Makes object S follow the light pen whenever
  S sees it. Returns S.
- \texttt{sunmv [S]} \hspace{1cm} Negates smv. Returns S.
- \texttt{sxyinc [S;X;Y]} \hspace{1cm} Increments the set point coordinates of S
  by X, Y. Returns S.
- \texttt{sc1r []} \hspace{1cm} Clears the display buffer. Returns NIL.
- \texttt{sdlt [W]} \hspace{1cm} Deletes object W from the display buffer
  and returns W.
- \texttt{ptrk [X,Y]} \hspace{1cm} Starts the pen tracking cross at X, Y and
  returns NIL.
- \texttt{puntrk []} \hspace{1cm} Stops the display of the pen tracking cross.
  Returns NIL.
- \texttt{px []} \hspace{1cm} Returns the x coordinate of the pen.
- \texttt{py []} \hspace{1cm} Returns the y coordinate of the pen.
ptchp [ ]
Returns NIL if the pen is far from the screen.

srn[ ]
Starts the display build up in the buffer.
Display will continue until sunrn[ ] is called.

sunrn [ ]
Stops the display and returns NIL.

pldw [X;Y]
To use this command, use macrosal to set up a set point at the current pen position followed by a zero length relative line and terminated with ISTOP. Then call pldw.
Increments of length at least X and scale Y will be added to this object until the pen leaves the screen. Returns NIL.

phclr [W]
Clears the pen approach history points of object W and returns W.

ph [W]
Returns a list of the pen approach history points for object W.

phcl [W;X]
Sets to X the minimum distance between pen approach history points for object W.
X is an integer between 0 and 1024.

phc2 [X]
Sets to X the minimum distance between pen approach history points which is assumed for newly created objects.
### phc 3 \([X] \)
Sets to \(X\) the maximum distance on the screen at which the pen can see an object. \(X\) is an integer between 0 and 40.

### sint \([W;Y] \)
Sets to \(Y\) the intensity of object \(W\). Returns \(W\). \(W\) is an integer between 0 and 7.

### scopy \([W1, W2] \)
Copies object \(W1\) and names the copy \(W2\). Returns \(W1\)

### plsh \([\ ]\)
Returns a list of all objects currently being seen by the pen.

### sscl \([W;Y] \)
Sets the scale of object \(W\) to \(Y\). Returns \(W\). \(Y\) is an integer between 0 and 3.

### plhw \([W] \)
Returns the name of the first object other than \(W\) seen by the pen.

### ptcchw \([\ ]\)
Waits until the pen is near the screen and then returns NIL.

### puntchw \([\ ]\)
Waits until the pen is not near the screen and then returns NIL.

### pjp \([W] \)
If the tracking cross is not following the pen and \(pjp \[W]\) has been executed, then when the pen sees object \(W\), the tracking cross will be centered about the point of \(W\) seen. Returns \(W\).

### punjp \([W] \)
Negates \(pjp\) and returns NIL.
pch [ ]
Returns the name of the most recent object to be seen by the pen or NIL if none have been seen since pohc was executed.

pohc [ ]
Sets the last object seen by the pen to NIL.

sline [S]
Returns NIL.

sline [S]
Returns a list of the coordinates of the end points of the line segments which make up object S.

ssubr [W, Y]
Causes the subroutine beginning at location Y to be executed each time object W is displayed.

sloc [W]
Intensifies a subexpression of object W which is indicated by the light pen. Subexpressions are marked in the object by pseudo parentheses which are not displayed. A pseudo left paren is indicated by the parameter word 600001y and a pseudo right paren by the parameter word 600002y.

sadd [W]
Counts the number of pseudo left parens to the intensified subpart of W.
(PRINDEF DISP MV EXPOND)
(DEFLIST (LAMBDA (N) (PROG (U V)(MACROSAL QUOTE DISP)
 (QUOTE (PARAMETER 1 2 3) (LOCY I) (LOCX I) (LINED 0 0 0)
 (ISTOP) ) ) (SETQ U (EXPAND N NIL) ) A (COND ( (NULL U) (RETURN NIL) )
 (MACROSAL T (LIST (COND ( (EQUAL (CAR U) (QUOTE U) ) (QUOTE (LINED
 0 0 0 4) ) ) (EQUAL (CAR U) (QUOTE D) ) (QUOTE (LINED 0 0 0 7777777774) )
 (EQUAL (CAR U) (QUOTE L) ) (QUOTE (LINED 0 0 7777777774 0) )
 (T (QUOTE (LINED 0 0 4 0) ) ) ) (QUOTE (ISTOP) ) ) ) (SETQ U (CDR U)
 (GO A) ) ) ) ) ) ) ) ) ) ) ) EXPR)

(DEFLIST (MV (LAMBDA (X) (COND ( (EQ X (QUOTE U) ) (QUOTE D) )
 ( (EQ X (QUOTE D) ) (QUOTE U) ) ( (EQ X (QUOTE L) ) (QUOTE R) )
 (T (QUOTE L) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) ) EXPR)

(PRINDEF EXPAND)
(DEFLIST (EXPAND (LAMBDA (N EXP) (PROG (A B C D E ) (SETQ EXP
 (QUOTE (U R D R) ) ) START (COND ( (ZEROP N) (RETURN EXP) ) ) (SETQ N
 (PLUS N -1) ) ) LOOP (SETQ A (CAR EXP) ) (SETQ EXP (CDR EXP) ) SETOB
 (CAR EXP) ) (SETQ EXP (CDR EXP) ) (SETQ C (CAR EXP) ) (SETQ E
 (NCONC E (LIST B A (MV B) A A B C B (MV C) B C C (MV B) C B D) ) )
 (COND (EXP (GO LOOP) ) ) (SETQ EXP E) (SETQ E NIL) (GO START) ) ) ) )
 ) ) ) ) ) ) ) ) EXPR)
(PRINDEF SKETCH)

(DEFLIST ( (SKETCH (LAMBDA (N X) (PROG (U W) (COND (X (GO A) ) ) (PJP

(BUTTON (QUOTE S) 100 1700) ) (PJP (BUTTON (QUOTE D) 300 1700) )

(PJP (BUTTON (QUOTE M) 500 1700) ) (BUTTON (QUOTE P) 700 1700)

(PJP (BUTTON (QUOTE R) 1100 1700) ) A (POHC) (SETQ W (PLHW NIL) )

B (COND ( (MEMBER W (QUOTE (D M) ) ) (GO C) ) ( (EQUAL W (QUOTE S) )

(GO G) ) ( (EQUAL W (QUOTE R) ) (GO H) ) ( (EQUAL W (QUOTE P) ) (GO I) )

( (EQUAL U (QUOTE M) ) (GO E) ) ( (EQUAL U (QUOTE D) ) (SDLT W) )

(GO A) C (SDLT U) (SETQ U W) (BUTTONON W) (SETQ W (PLHW W) ) (GO B)

E (SMY W) (PUNCHW) (SUNMY W) (GO A) G (SDLT U) (BUTTONON W)

(SETQ U W) D (PUNCHW) (PITCHW) (COND ( (GREATERP (PY) 1600) (GO A) )

(LINEDRAW (GENSYM) N I) (GO D) H (SDLT U) (RETURN NIL) I (SDLT U)

(SETQ U W) (BUTTONON W) (PUNTRK) (GO A) ) ) ) EXPR)

(PRINDEF LINEDRAW BUTTONON BUTTONONI)

(DEFLIST ( (LINEDRAW (LAMBDA (NAME QUALITY SCALE) (PROG NIL(PITCHW)

(SAL NAME (APPEND (QUOTE ( (1) (1) (1 4) ) ) (CONS (LIST 3414 SCALE)

(CONS (LIST 0) (APPEND (CONS (LIST PLUS 220000 (PY) ) ) (CONS

(QUOTE 0) ) (LIST (LIST PLUS 114000 (PX) ) ) ) ) (QUOTE ( (1) (1) 0 0 0 0

(1) (3000) ) ) ) ) ) ) ) ) ) NIL) (PLDW QUALITY SCALE) (RETURN NAME) ) ) ) EXPR)

(DEFLIST { (BUTTON (LAMBDA (NAME X Y) SXING (SAL NAME (QUOTE ( (
(DEFLIST ( 1 BUTTONON (LAMBDA (NAME) (BUTTONON1 NAME (SX NAME) (SY NAME) ) ) ) ) ) EXPR)

(DEFLIST ( BUTTONON1 (LAMBDA (NAME X Y) $XYINC (SAL NAME (QUOTE ( (0) (0) (34114) (0) (220000) (0) (114800) (0) (200010) (0) (0) 04000 (0) (200210) (0) (644000) (0) (30000) ) NIL X Y ) ) ) EXPR)

(30000) ) NIL X Y ) ) ) EXPR)
(PRINDEF ARGUS MAKSTRING1 MAKSTRING2)

DEFLIST ( (ARGUS (LAMBDA NIL (PROG (U V R Q MERGELIST NQ) (SETQ U (LINEDRAW (QUOTE ARGUS) 2 1)) (PUNCTCHW) (SETQ V (MAKSTRING (SLINE U) )) (SETQ R (MAP2 V ATRIESES (QUOTE GETC) )) (MAP2 AWEIGHTS R (FUNCTION (LAMBDA (WEIGHT LIST) (MAP LIST (FUNCTION (LAMBDA (CANDIDATE) (SETQ MERGELIST (MERGE3 MERGELIST) )))))))) (COND ( (NULL MERGELIST) (GO A)) (SETQ NQ (CDAR MERGELIST)) (SETQ Q (CAAR MERGELIST)) (MAP MERGELIST (FUNCTION MERGE3)) A (SDLT U) (MACROSAL (QUOTE LAST)) (LIST (QUOTE (PARAMETER 1 3 3)) (LIST (QUOTE LOCY) (AYMIN) (LIST (QUOTE LOCX) AXMIN)) (LIST (COND ( (GET Q (QUOTE SCHAR)) (QUOTE SCHAR)) (T (QUOTE CHAR))) (COND ( (NULL Q) (QUOTE ?)) (T Q)) (QUOTE (STOP))) (COND ( (NOT (EQUAL AMODE (QUOTE TRAIN))) (RETURN Q)) (PRINT R) (SETQ U (READ)) (COND ( (NOT (EQUAL Q U)) (GO B)) (CSE TQ AWEIGHTS (MAP2 AWEIGHTS R (FUNCTION (LAMBDA (X Y) (COND ( (MEMBER U Y) (PLUS X 1)) (T X))))))))) B (CSETQ ATRIESES (MAP2 V ATRIESES (QUOTE PUTC))) (RETURN U) )))))) EXPRL)

(DEFLIST ( (MAKSTRING (LAMBDA (LINE)) (PROG (U V MIN MAX V1 V2) (MAKSTRING2 (EVERYOTHER LINE)) (CSETQ AXMIN MIN (MAKSTRING2 (EVERYOTHER (CDR LINE)) )) (CSETQ AYMIN MIN (RETURN U)) )))))) EXPRL)

(DEFLIST ( (MAKSTRING1 (LAMBDA (X)) (COND ( (GREATER P (CAR X) V2) (LIST NIL (MAKSTRING3 X NIL)) (T (LIST (MAKSTRING3 X NIL) NIL)) )))))) EXPR)

(DEFLIST ( (MAKSTRING2 (LAMBDA (X)) (PROG NIL (SETQ MIN (CAR X)) (SETQ MAX (CAR X)) (MAP X (QUOTE MINMAX)) (SETQ (QUOTIENT (PLUS MAX (MINUS MIN)) 3) ) (SETQ V1 (PLUS MIN V1)) (SETQ V2 (PLUS V1 (QUOTIENT V 6))) (SETQ U (APPEND (MAKSTRING1 X) U)) (SETQ V1 (PLUS V1 V)) (SETQ V2 (PLUS V2 V1)) (SETQ U (APPEND (MAKSTRING1 X) U)) (RETURN NIL)) )))))) EXPRL)

(DEFLIST ( (MAKSTRING3 (LAMBDA (X LEFT)) (COND ( (NULL X) ) ( (OR (AND LEFT (GREATERP (CAR X) V2)) (AND NOT LEFT) (GREATERP V1 (CAR X)))) (PLUS 1 (MAKSTRING3 (CDR X) (NULL LEFT)) )))) (T (MAKSTRING3 (CDR X) LEFT)) )))))) EXPRL)
(PRINDEF MAP MAP2 SUBI EVERYOTHER MINMAX MERGE1
(MERGES GETC PUTC)
(DEFLIST (MAP (LAMBDA (X FN) (COND (NULL X) NIL) (T (CONS (FN (CAR X)) (MAP (CDR X) FN))))) EXPR)

(DEFLIST (MAP2 (LAMBDA (X Y GN) (COND (NULL X) NIL) (T (CONS (FN (CAR X) (CAR Y)) (MAP2 (CDR X) (CDR Y) FN))) EXPR)) EXPR)

(DEFLIST (SUBI (LAMBDA (X) (PLUS 68719476735 X)) EXPR)

(DEFLIST (EVERYOTHER (LAMBDA (X) (COND (NULL (CDR X)) (LIST (CAR X))) (T (CONS (CAR X) (EVERYOTHER (CDR X)))))) EXPR)

(DEFLIST (MINMAX (LAMBDA (X) (COND (GREATERP X MAX) (SETQ MAX X)) (GREATERP MIN X) (SETQ MIN X) (T NIL))) EXPR)

(DEFLIST (MERGE1 (LAMBDA (LIST) (COND (GREATERP (CDR LIST) NO) (PROG2 (SETQ NO (CDR LIST)) (SETQ) (CAR LIST)) (T (NIL))))) EXPR)

(DEFLIST (MERGE3 (LAMBDA (MERGELIST) (COND (NULL MERGELIST) (LIST (CONS CANDIDATE WEIGHT)) (EQUAL CANDIDATE (CAAR MERGELIST)) (CONS (CONS CANDIDATE (PLUS WEIGHT (CDAR MERGELIST)) (CDR MERGELIST)))) (T (CONS (CAR MERGELIST) (MERGE3 (CDR MERGELIST))))))) EXPR)

(DEFLIST (GETC (LAMBDA (PSTRING TREE) (COND (NULL PSTRING) TREE) (NULL TREE) (COND (ZEROP PSTRING) (LIST (LIST U)) (T (CONS NIL (PUTC (SUBI PSTRING) NIL))) (ZEROP PSTRING) (COND (MEMBER U (CAR TREE) TREE) (T (RPLACA TREE (CONS U (CAR TREE))))) (T (CONS (CAR TREE) (PUTC (SUBI PSTRING) (CDR TREE))) (LIST U))))) EXPR)
APPENDIX A.

LISTING OF THE LISP PROGRAMS

267
(OFFLIST((
(EDN RPFN)
(PLS RPFN)
(PWR RPFN)
(PRD RPFN)
(PFROMHEAD RPFN)
(EASSIGN RPFN)
((NEG ((LAMBDA X) (LIST (QUOTE PLS) (CAR X) (LIST (QUOTE PRD) (1 CADR X))))))
(ODN RPFN)
(LP LAMBDA X) (COND
 ((EQUAL CADR X) (QUOTE (HETA!)) (LIST (QUOTE PRD) (1 CADR X))))
 ((EQUAL CAR X) (QUOTE ATOM)) (CONSI (CAR X) (PFN1 (CDR X))))
 ((EQ (CAR X) (QUOTE EQQUOTE)) (CONSI (QUOTE FUNI) (CONS (CAR X) (PFN1 (CDR X))))
 ((EQUAL CADR X) (QUOTE LP))
 (APPEND (REVERSE (CDR (REVERSE (CAR X)))) (LIST
 (LAMBDA U) (COND
 ((EQ (CAR (LAST (CAR X))) (QUOTE EQUOTE)) (CONSI (QUOTE FUNIU))
 (IT U)) (CONS (LAST (CAR X)) (PFN1 (CDR X))))
 (CDR X)) (CONS (QUOTE ESET) (PFN1 (CDR X))))
 (T (CADR X))))
 (LAMBDA X) (CONSI (QUOTE NAM) (PFN2 (CDR X) (CAR X))))
 (POINT PFN5)
 (DEQUOTE (LAMBDA X)
 (COND
 ((EQ (ADDR X) (QUOTE THETA!)) (LIST (QUOTE EQUOTE)
 (LIST (QUOTE PRD) (1 CADDR X))))
 ((EQ (ADDR X) (QUOTE CMA)) (LIST (QUOTE EQUOTE)
 (CONSI (QUOTE ESET) (PFN1 (CDR X))))
 (T (LIST (QUOTE FQUOTE) (CADR X))))
 (T)
 (QUOTE (LAMBDA X) X))
 (QUOTE PFN1)
 DEFINE(1
 (LAMBDA X) (COND
 ((NULL (CDR X)) (CAR X))
 (T (LAST (CDR X))))
 (PFN1 (LAMBDA X) (LIST (CADR X) (CAR X) (CADDR X))))
 DEFINE(1
 (APARSE3 (LAMBDA X) (COND
 ((NUMBERP X)X)
 ((ATOM X) (LIST (QUOTE FORMX))
 ((EQ (CAR X) (QUOTE EQUOTE)) X)
 (T (CONSI CAR X) (MAPLIST (CDR X) (FUNCTION
 (LAMBDA U) (APARSE3 (CAR U))))))))
 )
 )
(OFFINF1
 (APARSE (LAMBDAX) (GETFILE (QUOTE APARSE) (LIST X)))
)
REMPROPS1 (PFN81 PFNSET PGRT PEG APARSE1 APARSE3
 APARSE2 PFN1 PFN2 PFN4 PFN5 PFN6 BPN1) EXPR1
REMPROPS1 (RP PLS EQU NEG PRO POINT BAR RBK PWR
 PFORMREAD EQUOTE) PGTR1
REMPROPS1 (LP CMA THETA LBK DEQUOTE) EQ1
REMPROPS1 (pun PLS PWR PPD PFORMREAD EASSIGN NEG
 OVD LD LBK POINT DEQUOTE EQUOTE) PP1)
(IPWRNUM (LAMBDA (X Y) (COND
   ((AND (ZEROP X) (MINUSP Y)) (QUOTE INF))
   ((MINUSP Y) (IPWRNUM (DULONGNUM X) (MINUS Y)))
   (EQUAL (CAR X) (QUOTE PRINT)) (DULONGNUM (IPWRNUM
      (CDR X) Y) (PWRNUM (CADDR X) Y))) (MINUSP X) (COND
      ((EQUALP Y) (MINUS (EXPT (MINUS X) Y))) (T (EXPT X Y))))))
(NOEXP (LAMBDA (X) (AN) (NUMBERP X) (ONEP (REMAINDER X 2))))
(QUESFNUM (LAMBDA (X Y) (PROG (U V W) (SETQ V X) (SETQ Y Y) (PRINT (QUOTE QUESTION)) (COND (NULL VI (GO A))) (SETQ U (INTERN (GET-FNSYM) (QUIT NAME))) (FAMILY (U CAR V)) (EDISPLAY
   U) (SETQ W (SUBST U (CAR V) W)) (SETQ V (CDR V)) (GO B)) A (PRINT W) C IS
   ETQ U (ERSETQ (RDFLX)) (COND (NULL U (GO C)) (EQUAL CAR U) (QUOTE G
   VUEP)) (ERROR (QUOTE (RETURNING UPWARD))) (EQUAL CAR U) (QUOTE ANSWER)
   (GO D)) (EQUAL CAR U (QUOTE ANSWER)) (GO E)) (SETQ V (ERSETQ IR
   DFLX)) (COND (NULL VI (GO C)) (ERSETQ (PRINT (EXECUTE (CAR V) (CAR
   V)))) (PRINT (QUOTE [THIS QUESTION TO YOU IS PENDING])) (PRINT W) (GO C)
   (SETQ V (ERSETQ (RDFLX)) (COND (NULL VI (GO C)) (RETURN (CAR V)) (E
   (SETQ U (ERSETQ IRDFLX)) (COND (NULL U (GO C)) (SETQ V (ERSETQ (RDF
   LX)) (COND (NULL VI (GO C)) (SETQ W (ERSETQ IEXECUTE (CAR U) (CAR V
   V)) (COND (NULL W) (GO C)) (RETURN W)))))
   (NUMBER (LAMBDA (X) (LENGTH X)))
)) DEFINE
(IFRAC (LAMBDA (X) (COND
   (NULL X) NIL)
   (ATOM X) (PRINT-STRUCT X)
   (ATOM (CAR X) (CONS (CAR X) (IFRAC (CDR X))))
   (T (ERROR X))))
))
SPECIAL(N)
COMMON(DOLLAR)
COMMON(RANDOMNUMBER)
COMPILE(WHOLEFILEREAD FILESEND
PAGEUPDATE EDISPLAY ARGLIST APOFF1 APOFF2 READF APOFF
EDDELETE GETSUM GETSUM1 GETSUM2
EPRINT EMIN LESS GRTP EMINUSP
PWRNUM GQDP QUESTION NUMBER I)
COMMON(Y)
COMPILE((MAPL STRIKE))
UNCOMMON(Y)
COMPILE((RANDOMNUMBER))
SPECIAL(W)
COMPILE((EFUN EFUN))
UNSPECIA TO ((ENLOG ENEG EDVD DEPEND EDRV EDRV) EDRV)
INT FLOAT EPWR APND PLSNUM PRDNUM DIVNUM DIVNUM GCD REGNUM
EENUMBERP EZEROP EZENEP FLEXP UNIONLIST
MKP ID MKPLS FSIN FCOS FTAN FSEC FCSC FCOT FSINH FCOSH
FTANH FCOTH FSECCH FCSCCH FDVD GETC FGFN
MKD RAND ECODE SIMPAN ANEG APAD APSR APWR AQVD
DPLS DFUN DPAD DFWR ECODE1 EPLS EPAD RPLST EQUOTE EQUOTE
MKP MKP I LAST)
COMMON(IY W NUMFUN NUMP))
COMPILE(MAPL EFUN EFUN))
UNCOMMON(IY W NUMFUN NUMP))
COMPILE(MKPLS MKPRD EPRD EDVD ENEG EPLS GRTP EMINUSP APND
LSSP FtLExP EPWR MKP MKP1 MKP2 MKP3 MKPS DEPEND
EZEROP LAST NTOL EDNEP PRDNUM PLSNUM DVNUM
DONUM REGNUM GCD ENUMBERP UNIONLIST FLOAT EDRV EDRV1 EDRV2))
STOP
DEFINE(!
  (DELSUBST LAMBDA (EXP OLDDEL NEWDELFREE) (SUBSTA
  (QUOTE)DELSUBSTi EXP NEWDELFREE))
  (LIST (QUOTE DRV) (CADR OLDDEL) (QUOTE DOLLAR) NIL
  EXP)))
  (DELSUBST) (LAMBDA (WFREE DEL) (SUBSTA
  (QUOTE LIST (QUOTE DRV) (CADR EXP) (NTOL WFREE) NIL))
  (QUOTE (DEL DOLLAR NIL)) DEL)))
)
COMPILE(DIFF DIFF3 DIFF4 DIFF6 DIFF8)
STOP
*EOF*
DEFINE
((ORVDOILAMBDA(X Y) (DIFFUO1 X Y))
 (DIFFDO1LAMBDA(X Y) (COND (ATOM X)
  X) ((EQUAL (CAR X) (QUOTE DRV)) (DIFFUO2 (CDR X) (DIFFUO1 (NTOL X) Y) Y
  NIL)) (T (CONS (CAR X) (NAPL (CDR X) (FUNCTION (LAMBDA (U) (DIFFUO1 U Y)
  1)))))) (DIFFDO2 (LAMBDA (X Y Z W) (COND (NULL (CDR X)) (COND (NULL
  W) (T (CONS (QUOTE DRV) (APPEND W (LIST Y NIL)))) ((EQUAL (CAR X) Z)
  (DIFFDO2 (CDR X) (DIFFUO3 ICADR X) Y Z) W) (T (DIFFDO2 (CDR X) Y
  Z (CONS (CAR X) (CONS (CDR X) 1)))))) (DIFFUO3 (LAMBDA (N Y Z) (COND (ZEROP N) Y) (T (DIFFUO3 (SUB1 I) (NKP (DIFF Y Z) Z))))))
DEFINE((
  EVALUATE(LAMBDA X SETI(PROG1U MARKFREE V)
  SETQ UI(EVALUATE6(SPTF TR SET))
  (SFT Q V X)
  A I(SETQ MARKFREE NIL)
  (SFT Q V) (EVALUATE6 V U) (*
  (COND (MARKFREE (GO X)))
  (RETURN V))
  (EVALUATE1(LAMBDA X Y) (COND
    (NULL (CDR X)) X)
    (TI(EVALUATE1(EVALUATE2 X (CAR Y)) (CDR Y)))))
  (EVALUATE2(LAMBDA X Y) (PROG (U)
    (COND (SETQ U (EVALUATE5 X (CDR Y) (CADR Y)))
      (HUG1 (SETQ MARKFREE (RETURN X)))
    (ATOM X) (RETURN X))
  (TI(ReturnX (CONS (CAR X) (MAP1 CDR X) (FUNCTION LAMBDA (V)
    (EVALUATE2 V Y)))))
  (EVALUATE1(LAMBDA (EXP FUN FORM) (COND
    (NULL FORM) NIL)
    (ATOM FUN) (COND ((EQ EXP FUN) FORM) (T NIL))
    (EQUICAR FUN) (QUOTE NAME) (COND
      (EQUICAR EXP) (QUOTE NAME) ((EVALUATE6 (INTO EXP), INTOL FUN)
      (EVALUATE6 (TWOFF EXP), (TWOFF FUN) FORM))
    (T NIL))
    ((EQUICAR FUN) EXP) FORM)
    ((EQ1 CAR FUN) (CAR EXP) ((EVALUATE6 (LASTOFF EXP) (LASTOFF FUN) FORM))
    (T NIL))
    (LASTOFF (LAMBDA X (COND
      (NULL (CDR X)) X)
    (TI (CONS (CAR X) (LASTOFF (CDR X)))))
    (EVALUATE6 (LAMBDA (EXP FUN FORM) (COND
      (NULL EXP) (COND (NULL EXP FORM) (T NIL))
    (NULL EXP NIL))
    (EQUICAR EXP (CAR FUN) (EVALUATE6 (CDR EXP) (CDR FUN) FORM))
    (NUMBERP CAR FUN NIL)
    (T (EVALUATE6 (CDR EXP) (CDR FUN) (SUBST CAR EXP (CAR FUN) FORM)))
    (EVALUATE6 (LAMBDA X (COND
      (NULL (CDR X)) X)
    (AND CAR (QUOTE EQU))
    (IS (ATOM (CADAR X)) (EQUICADAR X) (QUOTE NAME))
    (NULL (GET (CADAR X)) (QUOTE PERM)))
    (CONTINUE (CAR X) (EVALUATE6 (CDR X)))
    (TI (EVALUATE6 (CDR X)))
  )))
DEFINE ((EXPAND-LAMBDA(Y)(MULT2 Y))) (MULT31 (LAMBDA (Y) (COND (NIL (CDR Y)) NIL) (T (CONS ICAR Y) (MULT31 (CDR Y)))))) (MULT2 (LAMBDA (Y) (COND (EQUAL (CAR Y) (QUOTE EQ)) (COND (QUOTE EDQ) (LIST (QUOTE EDQ)) (MULT2 (CDR Y)) (MULT2 (CADR Y)) NIL) (T (CDR (SIMP2 (MKPLS (APPEND (MULT3 Y) (LIST NIL)) NIL))) (MULT2 (LAMBDA (Y) (SELECT ICAR Y) (QUOTE PLS)) (MULT4 (MAPL (LAMBDA (Y) (FUNCTION MULT2) (QUOTE PRD)) (MULT5 (MAPL (LAMBDA Y (FUNCTION MULT3) NIL))) (QUOTE DRV) (MULT1 (MULT31 Y) (MULT3 (NTOL Y)))))) (QUOTE IFG) (MULT6 (MULT31 Y) (MULT7 (MULT3 (CADR Y)) (MULT2 (ICADR Y)))) (LIST Y)))) (MULT4 (LAMBDA (Y) (COND (NIL (CDR Y)) (CAR Y) (T (MULT4 (CONS (APPEND ICAR Y) (CADR Y)) (CDR Y)))))) (MULT5 (LAMBDA (Y) (COND (NIL (CDR Y)) (V Y) (T (MULT5 (CDR Y) (MULT31 (CAR Y)) V NIL)))) (MULT5 (LAMBDA (Y) (COND (NULL (CDR Y)) (MULT51 (ICAR Y) V (Z1))))) (MULT51 (LAMBDA (X Y) (COND (NULL Y) (CONS (EPRD X (CAR Y)) (MULT51 X (CDR Y))))))) (MULT6 (LAMBDA (X Y) (COND (NULL Y) (CONS (APPEND X (CONS (CAR Y) (LIST NIL))) (MULT6 (X (CDR Y))))))) (MULT7 (LAMBDA (X Y) (COND (NULL (CDR X)) (LIST (EPWR (CAR X)) Y)) (AND (NUMBERP Y) (NOT (MINUSP Y)))) (MULT1 (LIST 1 Y) Y) (T (LIST (EPWR (MKPLS (APPEND ND X (LIST NIL)) Y 1))) (MULT7 (LAMBDA (X Y N) (COND (ZEROP N) Y) (T (MULT71 (MULT51 X Y NIL) X (SUB1 N)))))))
DEFINE
{(FACTOROUT1 LAMBDA (X FACTOR NAME) COND
(EQ (CAR X) QUOTE EQN) (LIST (CAR X))
(FACTOROUT1 (CADR X) FACTOR NAME))
(FACTOROUT1 (CADR X) FACTOR NAME NIL))}

(EZEROP X X)

(EQ (CAR X) QUOTE PL)) (EPRD NAME (MKPL SIMPL (CDR X))
FUNCTION LAMBDA (U (SIMPLIFY (COPY (EDVD U FACTOR))))) ))) )))

((EPRD NAME IS SIMPLIFY (COPY (EDVD X FACTOR))))) ))))))

NIL

(FACTOROUT1 LAMBDA (X Y) EVAL CONS
QUOTE FACTOROUT1 (COND
((NULL (CDR X)) (CONS (CAR X) CONS (CADR X) (CDR X)))
(T X))
Y))
IFEXPR)
DFFLIST

GROUP(LAMBDAARGS ALIST(PROG1UFREE X FLAG)
(COND(CDR ARGS)(ERROR'QUOTE(WRONG NUMBER OF ARGUMENTS FOR
GROUP)))
(COND(INPFORM(SETO FLAG T))
(CSETQ [NPNUM 0])
(CSETQ INPFORM(READ(EVAL(CAR(CLADDR ARGS))ALIST)(QUOTE PFORM)))
(SETO UFREE(EVAL(SUBST(QUOTE GETINSUB)(QUOTE PFORMREAD)(CAR ARGS))
ALIST))
(SETO W (INPFORM)
(COND(NULL FLAG)(CSETQ INPFORM NIL))
(RETURN(CONSGROUP(CDR W)(CAR L)(CDR UFREE)))))))
(FEXPR)

DEFINE(

GROUP(LAMBDA(X OP)(PROGIU)(RETURN(COND
NULL(CDR X)NIL)
(ATOM(CAR X))(GROUP(CDR XOP))
(EQ(CAR X)(LAST(CAR X))(QUOTE SUBMARK))OP)
(SETO U(GROUP(CDR X)(CAAR X))(U)
(T(GROUP(CDR X)(OP)))))
))
DEFINE
| LIMIT(LAMBDA(EXPR X N)(SIMPLIFY(LIMIT2(SIMPLIFY(EXPR) X N) X))){}
| LIMIT(LAMBDA(EXPR X N)COND
| (EQL EXP X N)
| (ATOM EXP)EXP)
| (TISELECT(CAR EXP)
| (QUOTE PRD)(LAMBDA(EXPR)(CAR EXP) X N))
| (MAPL EXP(FUNCT)(LAMBDA(U)(LIMIT1(U X N)))))
| LIMIT2(LAMBDA(EXPR X N)PROG(U V W N D)
| (SETO N 1)
| (SETO D 1)
| (SETO W 1)
| (SETO U 1)
| (COND((NULL(CDR U))GO A))
| (SETO V(SIMPLIFY(LIMIT1(CAR U) X M)))
| (COND((EZRUP V)SETO N(EPRD(CAR U) N)))
| (EQ V(QUOTE INF)(SETU D(EPRD(CAR U) D))
| (TISETO W(EPRD V W)))))
| (SETO U(CDR U))
| (GO B)
| ARETURN
| (COND((EONEP D)CON(EONEP N)(T O))
| (EONEP N QUOTE INF))
| (T(EPRD W LIMIT1(EDV0IDIFF N X X (DIFF(SIMPLIFY(EDV0 1(COPY D)) X X M) N))))
| )
| )
| )
LOAD(SIMP1)
EVALREAD(SIMP1 COMP NIL)
LOAD(SIMP2)
EVALREAD(SIMP2 COMP NIL)
LOAD(BASIC1)
EVALREAD(BASIC1 COMP NIL)
LOAD(BASIC2)
EVALREAD(BASIC2 COMP NIL)
LOAD(FIELD CHAIN SETUP)
CSET(INLIST NIL)
NBASIC LISP

DEFINITION
(EASSILAMBDA(NAME EXPRESSION)(PROG2
 (EASS NAME EXPRESSION) EXPRESSION))
(EASS LAMBDA EXP)(PROG NIL
 (FILEDELETE NAME (QUOTE FORM))
 (FILEDELETE NAME (QUOTE DFSPAN))
 (COND (AND (NOT (EQ NAME (QUOTE LAST))
 (NOT FILEGONE NAME (QUOTE FORM)))
 (FILEWRITE (QUOTE OLD) (QUOTE FORM))
 (READ NAME (QUOTE FORM)))
 (FILEWRITE NAME (QUOTE FORM) EXP)
 (RETURN NAME))
(EFEQ LAMBDA X Y)(LIST (QUOTE EQ) X Y NIL)
(EXCHANGE LAMBDA XI COND
 (COND (CDODDR X Y)
 (COND (LIC CAR (CAADR X) (CAADR X) (CAADDR X))))
 (MATCH (LAMBDA (EXP PATT CON)) (PROG) (ALIST))
 (SETQ MALIST (ALIST))
 (COND (MATCH2 EXP PATT) (RETURN (VAL CONST MALIST)))
 (RETURN EXP))
(MATCH LAMBDA (EXP PATT))
 (AND (MATCH2 EXP PATT)
 (OR (NULL (LAST PATT)) (APPLYCAR (LAST PATT))
 (LIST EXP) (MALIST))))
 (SUBSTAILAMBDA (CONST PATT EXP) (SUBSTA EXP))
 (SUBSTAILAMBDA XI COND
 (ATOM XI (MATCH X PATT CON))
 (MATCH CONS CAR XI (MAPCAR XI (FUNCTION SUBSTA))))
 (LIST EXP)
 (GETFILE LAMBDA (OP ARG5) (PROG1)
 (EVALREAD OPQUOTE LIST NIL)
 (SETQ UICOND)
 (GET OPQUOTE FEXPR1) (EVALCONS OP ARG5) (ALIST)
 (TAPPY OP ARGCONS (ALIST))
 (EVALREAD OPQUOTE ERASE (NIL))
 (CSETQ INLIST NIL)
 (RETURN U))
 (DRVZERO LAMBDA (EXP XFREE) (SUBSTA O
 (QUOTE (DRV DOLLAR (LAMBDA (U) (WITHRESPCTTO U XFREE)) EXP))
 (TERM LAMBDA XI COND
 (NULL CDR X) NIL)
 (LESF PP N) (CAR X)
 (TERM CAR X) (SUBST N) (COND)
 (CENSUS LAMBDA XI COND
 (ATOM XI))
 (NULL CDR X) (ADD1 (CENSUS CDR X))
 (LEFT LAMBDA XI CDR X))
 (DEPENDENCE LAMBDA XI COND
 (EENUMERATE X NIL)
 (ATOM X) (ESET X)
 (T (QUOTE ESET) (NIL) (LAST (MICX X))))
))
(RIGHT(LAMDA(X)(CAR X)))
(NEWNAM(LAMDA())PRG)
(FPANUPF)
(PACK1QUOTE F)
(NEWNAM1(EXPLOOD(CSETO NAMENUMF1ADD1 NAMENUMBER)))
(RETURN1INTERNKMAMI))
(NEWNAME1(LAMDA(X)ICOND)
(NEWNAM1(NULL XNIL))
(T1PRG2(PACK1CAR X))
(NEWNAM1(CDR X)))
(MATCH2(LAMDA(EXP PATT)(COND)
(NEWNAM1(CDR PATT))(NULL(CDR EXP)))
(NEWNAM1(CDR PATT)(QUOTE DOLLAR)T)
(IATOM EXP)NIL)
(FNULL(CDR EXP)NIL)
(EQUAL CAR EXP(CAR PATT)MATCH2(CDR EXP)
(CDR PATT))
(T NIL))
(WITHFSPSPECTID(LAMDA(EXP X)(MEMBER X(TWOFF EXP))))
(INTSUBST1(LAMDA(X Y)ICOND)
((NULL(CDR X))CONS YDCDR X))
((CONS1CAR X)(INTSUBST(CDR X)Y))))
(REPLACE1SUB1(LAMDA(X)(SUPSTA
(QUOTE(CAR)LAST EXP))))
(QUOTE(DOLLAR)(LAMDA(U)EDICAAR(LAST UI)
(QUOTE SUBMARK))))
(GETSUBS(LAMDA(X)(COND)
((NULL(CDR X))(REPLACE X)(LIST)(QUOTE SUBMARK)CSET0 INPNUM
ADD1 INPNUM)NIL))
((GETSUB3ICDR X)1))
(TWOFF(LAMDA(X)(COND)
((NULL(CDR X))NIL)
((CONS1CAR X)(TWOFFICDR X)1)))))
)
(CSET1NAMENUMBER 1)
DEFNP1((
(PUT1LAMDA(X Y Z)(QUOTE LIST
(X Y Z)NIL)))))
DEFNP(((
(ALISTLAMDA(X Y Z))
(NAME(LAMDA(X Y)CONSQUOTE NAME)(NILON(EVLIS X Y))))
(DRV(LAMDA(X Y)(DRV)EVLIS X Y))(VAL(LAST X Y)))
(FUN(LAMDA(X Y)NIL)(EVLIS X Y))
(FSET1LAMDA(X Y)(QUOTE SET1)(NILON(EVLIS X Y)NIL))
(PEXP1)
(CSET0(INPNUM 1)
(EPFNP1((
(DRVFACTOR(LAMDA(X Y Z)(GETFLIQUOTE DRVFACTOR
(LIST X Y Z))))
(EFLSUBST1(LAMDA(X Y Z)(GETFLIQUOTE MFLSUBST1)(LIST X Y Z))))
(SUMCAL(HLAMDA(X Y Z)(GETFLIQUOTE SUMCALH)(LIST X Y Z))
(MULTIPLYTHROUGH(LAMDA(X Y)(GETFLIQUOTE MULTIPLYTHROUGH
(LIST X Y)))))
(IDIFF(LAMDA(X Y Z)(GETFLIQUOTE IDIFF)(LIST X Y)))
(IDRV(LAMDA(X Y Z)(GETFLIQUOTE DRVD)(LIST X Y))))
(IEXPAND(LAMDA(X Y Z)(GETFLIQUOTE EXPAND)(LIST X Y)))))
ICOLLLECT(LAMDA(X Y Z)(GETFLIQUOTE COLLECT)(LIST X Y))))
\text{REPROP5(\text{NORMPOLY1 NORMPOLY2 NORMPOLY3 NORMPOLY4})EXPR1}
\text{DEFINE11}
\text{NORMPOLY1 LAMBDA(X Y1) GETFILE(IOLTE NORMPOLY1) (LST X Y1)1}}
\text{11}
DEFINITION
(NORMPOLY)(LAMBDA(X XFREE)(ISUBST
(QUOTE(NORMPOLY2 EXP IN NORMPOLY1(CDR EXP)(NORMPOLY4(CADR EXP) ))))
(QUOTE(PLS DOLLAR NIL) Y))
(NORMPOLY5)(LAMBDA(X N)(COND
(NULL(CDR X) N))
((NORMPOLY1(CDR X) MIN(NORMPOLY4(CAR X) N)) N)
(NORMPOLY6)(LAMBDA(X)(COND
((EQUAL X XFREE) NIL)
((EDICAR X)(QUOTE PWR)))(COND
((ATOM X) NIL)
((EDICAR X)(QUOTE PWR)))(COND
((EDICAR X)(QUOTE XFREE))(EQUALP(CADR X)(NIL))(NIL)
(CADR X))
(T X))
((EDICAR X)(QUOTE PRO))(NORMPOLY3(CDR X))
(T X) (T X))
(NORMPOLY5)(LAMBDA(A X)(COND
(TQUOTE)(NORMPOLY1(CAR X))
(NORMPOLY4)(LAMBDA(A X)(COND
((EDICAR X)(QUOTE N)))(NORMPOLY2(LAMBDA1 EXP N))
((EDICAR X)(QUOTE N)))(NORMPOLY5(CADR EXP)(NORMPOLY4(CADR EXP) ))))
(MULTIPLYTHROUGH EXP (EPWR XFREE(MINUS N) N) )
))
PICOM2 LISP
COMPILE (LATGHD2 NAMDS EUPY PICTIRE W H D A S W G HG DG
N LAST MKSYM MKCONCAT PUTP GESP MAPL PASSI HIEGER HHER1 PAREN
PAREN1 DIN_DOWN DIN_UP1 DIN_UP2 DIN_UP3 DIN_UP4 DIN_UP HALF)
DEFLIST((
  (ALIST LAMBDA (X Y))
) IFEXPR)

SPECIAL((SIZE V V2 XNC YNC SXNC SYNC BYNC BXNC ATOM WIDTH XPOS YPOS
  XREL YREL UI UZ WI)
  COMMON((SCOPEDIM))
  COMMON((FN PARENKIND DOLLAR BLANK FNS FRA ORDERLIST SLASHCOMMA))
  COMPILE((EDPY3 PARENNEXT PARENFEATER PAREN2 PAREN1
  PASS17 PASS16 PAREN2 PASS15 SPALER DIMUP6 MKWRUP LINER1 ATOM1
  NAMD1 NAMD2 NAMD3 NAMD4 PASS13 PASS14 DELIMIT DELIMIT1 DELIMIT2 EDPY2
  DIMDOWN1 DIMUP5 MKCLEANP PASS11 PASS12 SCOPEDIMSET ARGLIST
  EVLD1 NTOL MKUNDELIMIT LINER SIZE DRV01 DRV02 DRV03 NAMD1
  NAMD2 EPLS EDPY1 SIGNX SIGNY AB71))
DEFLIST
(REPLACE LAMBDA (X) (PROW FREE W)
(COND ((CADDR ARGS) (ERROR (QUOTE (LONG NUMBER OF ARGUMENTS FOR REPLACE)))
(CSETQ (INPNUM 0)
(Csetq INPFORM (EVAL (CAR ARGS)) (QUOTE PFORH))
(Csetq XFREE (EVAL (SUBST (QUOTE GETSUB) (QUOTE PFORH) (CADDR ARGS) ALIST))
(Csetq equal (CADDR ARGS) (QUOTE (EQUAL (QUOTE (EQUOTE HOLE))))
(Csetq (INPNUM 0)
(Csetq INPFORM (EVAL (CAR ARGS)) (QUOTE PFORH))
(EVAL (SUBST (QUOTE GETSUB) (QUOTE PFORH) (CADDR ARGS) ALIST)
(A1) (setq wink (REPLACE SUB INPFORM))
(Csetq INPFORM NIL)
(RETURN (SUBST (QUOTE (REPLACE) EXPI)
(QUOTE (DOLLAR (REPLACE 2) (SIMPPLY W))))))
(REXPR)
DEFINE
(REPLACE LAMBDA (X) (COND
((EQ CAR X) (QUOTE SUBMARK))
(COND (XFREE (CONS XFREE (REPLACE 1 (CDR X))))
(REPLACE 1 (CDR X)))
(REPLACE 1 (LAMBDA X) (COND
((NULL (CDR X)) X)
(EQ (CAR X) (QUOTE SUBMARK)) (REPLACE 2 (CDR X))
(CONS CAR X) (REPLACE 2 (CDR X)))
(REPLACE 2 (LAMBDA X) (COND
((ATOM X) NIL)
((NULL (CDR X)) NIL)
((EQ (CAR X) (QUOTE SUBMARK)) IT)
(LET (REPLACE 2 (CDR X)))))
)
SETUP LISP

FLAGI(PLS PRO SUH TMS DVD FRT EWN STT
NEG PWR DRY ITG UNN INS NAX CND
GRT LSS GRC LSE XST LHD ANN NOT ORR
EVL ABS FIT INF SIN COS TAN COT SEC
CSC SINH COSH TANH COT + ASIN ACOS ATAN ACOT
NEXP NLOG1PERM)

CSET(RANDOMNUMBER 17054832)
DEFLIST(I(SIN PI(COS FCOS)(TAN FTAN)
(SEC FSEC)(CSC FCSC)
(COT FCOT)(SINH FSIHNH)(COSH FCOSH)(TANH FTANH)
(COTH FCOTH)(SECH FSECH)(CSCH FCSCCH)(FCODE)
DEFLIST(I(1 5525736173) E 83647203441
PI 921361)CODEI
DEFLIST(I(1(SIN(COS ** NIL))
1COS(PWR -1(SIN ** NIL)(NIL))
1TAN(PWR(SEC ** NIL))2 NIL)
1COT(PWR -1(PWR(CSC ** NIL))2 NIL)
1SEC(PWR -1(PWR -1(SEC ** NIL))2 NIL)
1CSC(PWR -1(COT ** NIL)(CSC ** NIL))2 NIL)
1SINH(COSH ** NIL)
1COSH(SINH ** NIL)
1TANH(1PWR -1PWR(CSC ** NIL))2 NIL)
1COTH(PWR -1PWR(CSH ** NIL)(NIL))2 NIL)
1SECH(PWR -1SECH ** NIL))2 NIL)
1CSCH(PWR -1CSCH ** NIL))2 NIL)
1ASIN(PWR(PLS 1)PWR -1(PWR ** 2))2 NIL)
1ATAN(PWR -1FRT -1FRT 2 NIL))2 NIL)
1ACOS(PWR -1PWR -1(PWR ** 2))2 NIL)
1ACOT(PWR -1FRT -1FRT 2 NIL))2 NIL)
1NLOG(PWR ** -1 NIL))2 NIL)
) DIFF)
Y1 SIMPEG(IN(OL Y1))1111
(CONS(F9FN(GTC X1)) CONS X(S1KP31 Y1)) 111
(SIMPREG(LAMBDA(X) (SPRU(CANSICAGE -1 SIMPLEL) -1)) X) 11
(SIMP(EQUAL)(LAMBDA(X Y) ICOND
((EQUAL X Y)) 1
((EQUAL(NEG X)Y)=1
(T NIL)))))
COMMON(FLNM FSMW FNZ AFUN ELE FUN FNM FUN NURF UN FNM GARG)
SPECIAL(ISU SIMP LEVEL SX1)
COMPILE(SIMP31 SIMP33 SIMP32 SIMPLIFY SIMP2
SIMP5 SIMP51 SIMP52
SIMP341)
SIMP2  COMP
COMPILE((SIMP3 SIMP4 SIMP5 SIMP7 SIMP1 SIMP6 SIMP8
SIMP9 SIMP57 SIMP8 SIMPEQUAL SIMPAEG))
(DEFINITION
 (SPLIT1(LAMBDA(X)(MKCLEANP(SPLIT(T2(SPLIT1 X)100))))))
 (SPLIT1(LAMBDA(X)1(COND
 ((ATOM X))
 (TISPLIT1(CONS(CAR X)(MAPLICDR Y)(FUNCTION SPLIT1))(0))))
 (SPLIT3(LAMBDA(A X)1(COND
 ((NULL(CDR X))1(LIST M))
 (TICONS(CAR X)(SPLIT1(CDR X))(PLUS(SPLIT4(CAR X)1(IN))))))
 (SPLIT7(LAMBDA(A X)1(COND
 ((ATOM X))
 ((TOSPLIT1(LAST X))(EASS(NAMESPACE)(MKCLEANP X))))
 (NOT1(GREATERP (LAST X)N)X)
 (TIPROG(U M R))
 (SETO U X)
 (SETO M(LAST U))
 (SETO M1(SPLIT6 U))
 (COND (LESSP R 2)(RETURN
 (EASS(NAMESPACE)(MKCLEANP X)))))
 (SETO M(DIFFERENCE M R))
 (COND (LESSP M N)(RETURN(SPLIT7 U (DIFFERENCE N M))))
 (SETO U(SPLIT7 U R))
 (GO A))))))
 (SPLIT4(LAMBDA(X)1(COND
 ((ATOM X))
 (T(LAST X))
 (T(NULL(CDR X))(SPLIT4(CAR X))))
 (T(MAX(SPLIT4(CAR X))(SPLIT6(CDR X))))))
 (SPLIT7(LAMBDA(A X)1(COND
 ((NULL(CDR U))U)
 (EQUAL(SPLIT4(CAR U))(CONS(SPLIT7Z(CAR U))X)
 (SPLIT7(CDR U)1 M))))
 (TICONS(SPLIT7Z(CAR U)(LAST(CAR U)))(SPLIT7(CDR U)R M)))))
 (SPLIT5(LAMBDA(A U)1(COND
 ((NULL(CDR U))U)
 (EQUAL(R(SPLIT4(CAR U)))(CONS(SPLIT8(CAR U)(CDR U)))
 (TICONS(CAR U)(SPLIT5(CDR U)R))))))
 (SPLIT8(LAMBDA(U)1(COND
 ((NULL(CDR U))1(LIST 0))
 (TICONS(CAR U)(SPLIT8(CDR U)))))))
 )
DEFINE
  (SUMFRESH (LAMBDA (X XFRE2) (COND
      \AND (EQ \CAR XFRE2) (QUOTE SUM))
      (IF (CAR (INTOL XFRE2)) (QUOTE PLS))
      (CON (QUOTE PLS) (MAPL (\CDR (INTOL XFRE2))
        (FUNCTION (LAMBDA (U) (INTOLSUBST XFRE2 U)))))
    T XFRE2))
)
DEFINE
(SUMEXPANDILAMBDAX)(COND
((AND(EQ(CARX)QUOTESUM)(NUMBERP(CADRX)))(NUMBERP(CADDRX))))
(MKPLS(SUMEXPANDICADRXX)(CADRX))
(CADDRX)(CADDRICDRXX))
(TX))
(SUMEXPANDILAMBDAINZXX)(COND
((GREATERPNIZXX)(QUOTFNIL))
((EQUALNIZXILISTSUBSTNILXXNIL))
(T(CONS(SUBSTNILXX)(SUMEXPANDI(ADDI1NIL
NIXX)))))
)}
DEFINE
I((TRNK23 (LAMDA I Y N X) (PRG NIL)
(FILEWRITE QUOTE LISPTR) QUOTE LISP1 (TRNK12 Y N X))
(assignment)
(PRINT QUOTE (QUOTE (PLSOWRE))) (RETURN NIL) :I
I (LAMBDA (Y I) 'UNO (NULL (CDR Y)) I I (CAR Y) I (EAND (CDR Y))) (I T NIL)))) I (TRNK12 I LAMDA (Y I N X) (COND I (EQUAL I CAR Y) I QUOTE EQ1 (L I ST I CAR Y) I (TRNK12 (CAR Y) N X)) (TRNK12 (CADDR Y) N X) NIL) I IT MKPLS (TRNK12 (TRNK9 Y X NIL) NIL) X)) I I TRNK63 I LAMDA (Y U V S R)
(实习)
I (REVERSE (TRNK10) (REVERSE (SRPWRI Y R))
(MINUS S1)) I I TRNK62 (LAMDA (X Y) (COND I (NULL X) I (LIST NIL) I)) I (O R (ZERO I CAR Y) I I (ZERO I CAR Y) I I (TRNK62 (CDR X) I I (COND (I (CARE (CAR Y) I (CAR Y) (TRNK62 (CDR X) (CDR Y))) I)) (TRNK6) I LAMDA (W)) I (cond I (GREATERP (LENGTH U) W)) I (TRNK61 I CONS (EPROD W IMPLS TRNK62 W U)) I (I X V (AUX1 N1)) I I) I I TRNK62 (LAMDA (X Y) (PROG U V) I (COND (NULL X) I RETURN (TRNK61 I LAMDA V I (WEVEN V)
I (CDR U))) I (GETL (LAMDA I Y X) I (COND I (NULL Y) I NIL) I (EQUAL I CAR Y I X) I (I CAR Y V)) I I I (RETURN (TRNK61 I LAMDA X I X)) I I TRNK62 (LAMDA (X Y) I (COND I (OR
I (NULL X) (NULL Y) NIL) I I (REVERSE (TRNK2222 (TRNK5 (SUB1 (LENGTH Y)) I I)
(REVERSE X) Y))) I I I (TRNK2222 (LAMDA I X Y) I (COND I (NOT (ZERO I CAR X))) I I)
I (LIST MKPLS (TRNK3 X Y))) I I IT CONS I MKPLS (TRNK3 X Y) I I I I I TRNK2222 (C D R X))) I I)
I I I (TWOFF (LAMDA Y) I (COND I (NULL (CDR Y) Y)) NIL) I (COND I (C D R Y)) I)
I I I (TWOFF (CDR Y) Y)) (TRNK2 (LAMDA I X Y) I (COND I (NULL X) I Y) I (NULL
I X) I) I I I I I I I I (OR (EQUAL I CAR Y) I (CAR Y) I I (TRNK5 (LAMDA I X C)
I (COND (GREATERP (LENGTH U) W) I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I I
))))
(FILESEEK(QUOTE LISPRT)(QUOTE RETURN))
(SETO U(READI))
(FILEENDRQ(QUOTE LISPRT)(QUOTE RETURN))
(RETURN U1))
1)
SPECIAL((X N))
COMPILE((
EAND TRNK12 TRNK63 TRN:62 TRNK61 TRNK6 GETL TRNK5
TRNK2222 TWOFF TRNK1 TRNK51 TRNK3 SRPWR SRPWR1 SRPWR2
SRPWR3 SRPWR11 TAKE TRNK10 TRNK13 TRNK131 TRNK121 TRNK9
ATOMOFF TRNK91 TRNK92 TRNK93 TRK94 TRNK95 TRNK96 TRNK961))
UNSPECIAL((X N))
STOP
BIBLIOGRAPHY

J. Sammet has prepared an exhaustive bibliography. Therefore, only a few references appropriate to each chapter will be given here.

Sammet's bibliography is listed as the first reference.

Chapter 1


Chapter II


2. M.I.T. Mathematics Dept., Course 18.655, Notes and References.


Chapter III


Chapter V


Chapter VI


Chapter VII


6. Ankeny, N.C. M.I.T. Mathematics Department, discussion.


Chapter IX


Chapter X


Chapter XI


Chapter XIV


Biographical Note on the Author

William A. Martin was born in Oklahoma City on May 18, 1938, and graduated from Northwest Classen Highschool in 1956. He entered M.I.T. as a National Merit Scholar and received a B.S. and M.S. in Electrical Engineering in 1962. During this period he was elected to Tau Beta Pi, Eta Kappa Nu, and Sigma Xi. He was a member of the Beta Theta Pi social fraternity and captain of the wrestling team. He was a teaching assistant in the Electronic Circuits and Signals Laboratory and a Research Assistant in the Center for Operations Research.

Mr. Martin has held summer positions with electronics firms in England, Germany, and Switzerland and has been employed by the MITRE Corporation, Information International Incorporated, and IBM.

He published: