Open access • Journal Article • DOI:10.1145/321574.321591

On Canonical Forms and Simplification — Source link

B. F. Caviness

Institutions: Carnegie Mellon University

Published on: 01 Apr 1970 - Journal of the ACM (ACM)

Topics: Canonical form, Weyr canonical form, Canonical normal form, Canonical coordinates and Conjecture

On Canonical Forms and Simplification

by

Bobby Forrester Caviness

Department of Computer Science
Carnegie-Mellon University
Pittsburgh, Pennsylvania
May 1968

Submitted to the Carnegie-Mellon University
in partial fulfillment of the requirements
for the degree of Doctor of Philosophy

This work was supported in part by the Advanced Research Projects Agency of the Office of the Secretary of Defense (SD-146) and is monitored by the Air Force Office of Scientific Research. Distribution of this document is unlimited.

ACKNOWLEDGMENTS

I am most grateful to Dr. Alan J. Perlis, my advisor, for his guidance and encouragement during the course of the work on this dissertation.

To Dr. Henry S. Leonard, Jr., I am grateful for help with the sections on radical expressions.

My wife, Jane, has not only provided me with encouragement and understanding, but she has contributed to this thesis in many material ways. Among other things, she typed the rough drafts and helped with the programming.

Mrs. Janet Delaney provided frequent assistance with a sometimes difficult Formula Algol system, and Miss Carol Miller did an expert job of typing the final draft.

I am also most appreciative for financial support received from the National Science Foundation in the form of a graduate fellowship.

TABLE OF CONTENTS

Chapter I

Introduction

<u>Formula manipulation</u> is the process of carrying out operations and transformations on mathematical expressions or formulae. An expression is a string of symbols such as

$$(1) \qquad x^2 + e^{e^3+2} * x + 1.$$

(This is not actually a string but a two-dimensional figure. We shall not make a distinction between strings and figures except when necessary, and hence we will use such figures to represent strings.) With each expression is associated a function in a natural way. Thus by expression we mean such a string of symbols and by function, the natural function associated with such an expression.

Simplification

In formula manipulation process expressions with unnecessarily complicated structures are invariably generated. For example, most differentiation algorithms, when applied to the expression (1), will produce an expression similar to

$$(2) \qquad 2 * x^1 + (0 * e^3 + 0) * (e^{e^3} + 2) * x + (e^{e^3} + 2) * 1 + 0$$

instead of the functionally equivalent and structurally simpler expression

$$(3) \qquad 2 * x + e^{e^3+2}.$$

Such behavior seems to be characteristic of most formula manipulation algorithms. The process of reducing expressions like (2) to a simpler equivalent form like (3) is called simplification. Simplification is also taken to embrace other kinds of transformations such as finding common denominators for rational expressions, factoring, lexicographical ordering of subexpressions appearing in sums and products, etc.

The importance of keeping expressions in simplified form is threefold. First of all, simplified expressions require less memory space. For example, expression (2) typically requires about thirty storage cells whereas (3) only takes nine. Secondly, the processing of simplified formulae is faster and simpler. The processing is simpler in the sense that simplified formulae usually possess nice features which allow for cleaner and more precise algorithm design. Thirdly, functionally equivalent expressions are easier to identify when they are in simplified form. Indeed, simplification is of such a nature that almost no formula manipulation program can do without simplification capabilities.

Given the central role of simplification, it is hardly surprising to find that many algorithms for performing simplification have been reported in the literature. See Sammet's bibliography ([19] and [20]) for an extension listing of these. The usual form of attack of these algorithms has been to take a set of simplifying transformations that apply in obvious local cases and to try to weld these into a stable and coherent global schema for simplifying a class of expressions. The need for simplification and the kind of simplification transforms needed seem obvious in simple cases. However, as the expressions and algorithms increase in complexity the answers are no longer so obvious.

Fenichel [8] and Tobey, Bobrow, and Zilles [21] discuss the problems of the simplification algorithms in some detail. The main conclusion to be drawn from their discussion is that simplification only has meaning in a local context. For instance Fenichel points out that

$$\csc^2(x) - \cot(x) * \csc(x)$$

is easier to integrate than its structurally simpler equivalent

$$\frac{1}{1+\cos(x)} .$$

Thus in the context of integration the former expression is the simpler whereas in other cases the latter is more appropriate.

Recently, Richardson ([16] and [17]) provided some theoretical evidence of simplification problems when he proved that for sufficiently rich classes of expressions it is impossible to always determine when expressions are identically zero. Hence such simplification transforms as

$$x + 0 \rightarrow x$$

cannot always be applied since the O cannot always be identified.

Motivation and overview

The motivation for this work comes from these two sources. First of all we wanted to study the problems of simplification. But in order to guide our work on simplification it seemed desirable to study further the unsolvability angle. Thus in Chapter II we study Richardson's theorem and proof in detail. From Richardson's proof and from studies on the unsolvability of Hilbert's tenth problem, we draw some conclusions about sharpenings of Richardson's theorem.

With the limitations of these negative results in mind, we study in Chapter III the structure of some classes of expressions and prove the existence of canonical forms for these classes. The concepts of canonical and normal forms as developed in Chapter III preserve most of the important concepts of simplification. On the other hand, these concepts are global concepts that can be formalized and hence are appropriate for a careful study whereas the concept of simplification lacks these properties.

Then in Chapter IV, we discuss the implementation of the algorithms developed in Chapter III. The algorithms are implemented using Formula Algol ([10], [14], and [15]). The Formula Algol programs are included as appendix II with some output from actual runs.

Literature review

Other than the simplification algorithms mentioned previously, there is only a small literature that has any relevance to this study. First are the papers on simplification programs. Our approach is quite different from the approach of the these papers. Their approach is completely programmatic. Our primary aim is to study the structure of the classes of expressions in a formal way and only secondly to produce programs.

Otherwise there are two papers that have a particular relevance here. The first is the work of Richardson. The other is by W. S. Brown [3] and is very similar to some of our work in Chapter III. Both of these papers will be discussed later.

Then there are two other papers that should be mentioned here. P. J. Brown [2] has written an interesting paper in which he studies the existence of canonical forms in a more general setting than we have. Given a syntax for a set of expressions (or language) and a set of equivalence preserving transformations on the language, he investigates the properties that the set of transformations must have in order that unique canonical forms exist. The main purpose of the canonical form is to prove equivalence between expressions.

Since a number of unsolvable equivalence problems such as the word problems for semi-groups and groups, and the equivalence problem for mathematical expressions can be phrased as particular examples of the general kinds of calculi with which Brown works, one suspects that his methods cannot be too powerful. Even when his results are applicable, their application requires considerable ingenuity as he points out. As an application of his results he outlines a proof procedure for elementary trigonometric identities.

A paper by G. Rousseau [18] attacks some similar problems in a somewhat different realm. He proves the existence of an effective procedure for deciding whether or not functions contained in a subclass of the primitive recursive functions are identically zero. (This problem is recursively undecidable for the class of all primitive recursive functions.) Specifically let $\mathcal{E}$ be the set of functions obtainable from the initial functions Z (zero function), S (successor function), $x + y$, $x \cdot y$, x^y, $\max(x,y)$, $\min(x,y)$, $c \dot{-} x$, $x \dot{-} c (c = 1,2,\ldots)$ and the projection functions $U_i^n(x_1,\ldots,x_n)$, $i = 1,\ldots,n$; $n = 1,2,\ldots$ by substitution and the formation of bounded sums and products. There exists an effective procedure for deciding whether or not a given element of $\mathcal{E}$ is identically zero. This result does not seem relevant to our work since $\mathcal{E}$ does not contain the kinds of expressions with which formula manipulation is concerned.

Notation

Theorems and lemmas are numbered by chapter. Capital script letters are used to denote classes of expressions, capital printed letters to denote expressions within a class, small letters from the end of the alphabet are used as variables in expressions. The well-formed expressions of a particular class are indicated by Backus-Naur Form [1] grammars. Γ is the field of rational numbers, Ω the field of complex numbers. Ψ's are used to represent extension fields of Γ. J denotes the ring of integers.

Chapter II

Limitations on the Existence of Canonical Forms by Undecidability Results

In this chapter we study the negative side of the problem in order to obtain some guidance in the search for classes of expressions that possess canonical forms. First we give some definitions.

To be given a <u>class of expressions</u> $\mathcal{E}$ means to be given rules, such as a Backus-Naur Form (BNF) grammar, for determining the well-formed expressions in the class. The expressions must be formed from a finite set of atomic symbols, a subset of which must be designated as <u>variables</u>. Any member of $\mathcal{E}$ not containing a variable is called an <u>$\mathcal{E}$-constant</u> or <u>constant $\mathcal{E}$-expression</u>. The constants will usually form some well-known structure such as the ring of integers or field of rationals. Expressions are interpreted as functions over the domain $\mathcal{D}$ of constants.

If E_1 and E_2 are members of an expression class $\mathcal{E}$, E_1 is said to be <u>identical</u> to E_2 if E_1 and E_2 are the same string of atomic symbols. This relation is denoted by $E_1 = E_2$. E_1 and E_2 are said to be <u>functionally equivalent</u>, or simply equivalent, if for all assignments of values in $\mathcal{D}$ to their variables for which they are defined, they are equal. This realtion is denoted by $E_1 \equiv E_2$. Of course $E_1 = E_2$ implies $E_1 \equiv E_2$.

Examples: Let J be the ring of integers and Γ the field of rationals. Let $E_1 = x + 3$ and $E_2 = (x - x + 1)*(x + 3)$. Then $E_1 \equiv E_2$ over both J and Γ but $E_1 \neq E_2$. Similarly if $E_1 = \sin(x * \pi)$ and $E_2 = 0$ then $E_1 \equiv E_2$ over J, $E_1 \not\equiv E_2$ over Γ and $E_1 \neq E_2$.

Richardson's Theorem

Let $\mathfrak{R}$ consist of the class of expressions generated by

(i) the rational numbers and the real numbers π and $\log 2$,

(ii) the variable x,

(iii) the operations of addition, multiplication, and composition,

(iv) the sin, exp, and absolute value function.

(In the text we use informal definitions such as the above for expression classes. BNF definitions for most classes can be found in Appendix I.)

Theorem 1: If E is an expression in $\mathfrak{R}$ the predicate

$$'E \equiv 0'$$

is recursively undecidable. This decision problem will be referred to as Richardson's decision problem.

In order to derive some further results that follow from the proof of this theorem, the proof is included here. For the proof, a variation of the class $\mathfrak{R}$ is needed. Let $\mathfrak{R}_1$ be the class of expressions generated by

(i) the rational numbers and the real numbers π and $\log 2$,

(ii) the variables $x_1, x_2, \ldots, x_n$,

(iii) the operations of addition, multiplication and composition,

(iv) the sin and exp functions.

$\mathfrak{R}_1$ differs from $\mathfrak{R}$ in that is contains an arbitrary number of variables and does not contain the absolute value function. We shall show that for $G(x_1)$ in $\mathfrak{R}_1$ ($G(x)$ is a member of $\mathfrak{R}$ since it is an expression of only one variable) that the predicate 'there exist a real number a such that $G(a) < 0$' is recursively undeciable. This predicate will be referred to as the decision problem for $\mathfrak{R}_1$. Now suppose Richardson's decision problem is decidable and $G(x_1)$ in $\mathfrak{R}_1$. Consider $F(x) = |G(x)| - G(x)$. $F(x)$ is in $\mathfrak{R}$ and $F(x) \not\equiv 0$ if and only if there exists a constant b in $\mathfrak{R}$ such that $G(b) < 0$. But as we shall show we cannot decide if a real b exists such that $G(b) < 0$. But since the constants of $\mathfrak{R}$ (and $\mathfrak{R}_1$) are dense in the reals and all expressions are continuous functions, there exist a real b if and only if there exist a b in $\mathfrak{R}$ such that $G(b) < 0$. Thus if we can decide if $F(x) \equiv 0$, then we can decide if there exist such a real b, i.e., solve the decision problem for $\mathfrak{R}_1$. Except for the proof of the undecidability of the decision problem for $\mathfrak{R}_1$, the proof is complete.

The proof of the undecidability of the decision problem for $\mathcal{R}_1$ will be presented as a series of lemmas. The starting point is a result from a paper by Davis, Putnam, and Robinson [7].

Theorem 2: There exists a set of polynomials with integer coefficients

$$= \{P_i(y_1,y_2,\ldots,y_n,z_1,z_2,\ldots,z_n), i = 1,2,\ldots\}$$

such that the predicate

'there exist integers $a_1,a_2,\ldots,a_n$ such that

$$P_i(a_1,a_2,\ldots,a_n,2^{a_1},2^{a_2},\ldots,2^{a_n}) = 0\text{'}$$

is recursively undecidable.

Now consider

Lemma 1: For every $F(x_1,\ldots,x_n)$ in $\mathcal{R}_1$ there exists a $G(x_1,x_2,\ldots,x_n)$ in $\mathcal{R}_1$ such that

(i) $G(x_1,x_2,\ldots,x_n) > 1$ for all $x_1,x_2,\ldots,x_n$.

(ii) $G(x_1,x_2,\ldots,x_n) > F(x_1 + \Delta x_1,x_2 + \Delta x_2,\ldots,x_n + \Delta x_n)$ for all x_i and for $|\Delta x_i| < 1$, $i = 1,2,\ldots,n$.

G is called a dominating function for F.

Proof: The proof is by induction on the number of operators and primitive functions making up F.

If $F = c$, a constant, choose $G = |c| + 2$.

If $F = x_i$, choose $G = x_i^2 + 2$, $i = 1,2,\ldots,n$.

If $F = F_1 + F_2$ and G_1 and G_2 dominate F_1 and F_2 respectively choose $G = G_1 + G_2$.

If $F = F_1 * F_2$, choose $G = G_1 * G_2$.

If $F = \exp(F_1)$, choose $G = \exp(G_1)$.

If $F = \sin(F_1)$, choose $G = 2$.

In all cases G dominates F.

Q.E.D.

Using theorem 2 and lemma 1 we prove

Lemma 2: For each $P_i(y_1,y_2,\ldots,y_n,z_1,z_2,\ldots,z_n)$ in $\mathcal{R}$ there exists $F_i(x_1,x_2,\ldots,x_n)$ in $\mathcal{R}_1$ such that

(i) there exist integers $a_1,a_2,\ldots,a_n$ such that $P_i(a_1,a_2,\ldots,a_n,2^{a_1},2^{a_2},\ldots,2^{a_n}) = 0$ if and only if

(ii) there exist real numbers $b_1,b_2,\ldots,b_n$ such that $F_i(b_1,b_2,\ldots,b_n) = -1$ if and only if

(iii) there exist real numbers $b_1,b_2,\ldots,b_n$ such that $F_i(b_1,b_2,\ldots,b_n) < 0$.

Proof: Observe that

$Dx_jP_i^2(x_1,x_2,\ldots,x_n,2^{x_1},2^{x_2},\ldots,2^{x_n})$ is in $\mathcal{R}_1$ for each $j = 1,2,\ldots,n$, where $Dx_jP_j^2(x_1,x_2,\ldots,x_n)$ is the partial derivative of P_i^2 with respect to x_j. Let $K_j(x_1,x_2,\ldots,x_n)$ be the dominating function for $Dx_jP_i^2$. Define F_i by

$$F_i(x_1,x_2,\ldots,x_n) = (n+1)^2[\sum_{j=1}^{n}(\sin^2\pi x_j) * K_j^2(x_1,x_2,\ldots,x_n) + $$
$$+ P_i^2(x_1,x_2,\ldots,x_n,2^{x_1},2^{x_2},\ldots,2^{x_n})] - 1.$$

Obviously (i) implies (ii) implies (iii). It is only necessary to show that (iii) implies (i). Choose a_i to be the smallest integer such that $|a_i - b_i| \leq \frac{1}{2}$. We shall show that $P_i^2(a_i,a_2,\ldots,a_n) < 1$ which implies that $P_i(a_1,a_2,\ldots,a_n) = 0$ since P_i maps integers into integers. From $F_i(b_1,b_2,\ldots,b_n) < 0$ we have

$$\sum_{j=1}^{n}(\sin^2\pi b_j) * K_j^2(b_1,b_2,\ldots,b_n) + P_i^2(b_1,b_2,\ldots,b_n,2^{b_1},2^{b_2},\ldots,2^{b_n})$$
$$< \frac{1}{(n+1)^2}.$$

Hence $|\sin \pi b_j| \cdot K_j(b_1,b_2,\ldots,b_n) < \frac{1}{(n+1)^2} < \frac{1}{n+1}, j = 1,2,\ldots,n,$ and $P_i^2(b_1,b_2,\ldots,b_n,2^{b_1},2^{b_2},\ldots,2^{b_n}) < \frac{1}{(n+1)^2}$. By the n-dimensional mean value theorem

$$P_i^2(a_1,a_2,\ldots,a_n,2^{a_1},2^{a_2},\ldots,2^{a_n}) \leq P_i^2(b_1,b_2,\ldots,b_n) +$$
$$\sum_{j=1}^{n}|b_j - a_j| * Dx_jP_i^2(c_1,c_2,\ldots,c_n)$$

where c_j is between a_j and b_j. From the definition of K_j

$$P_i^2(a_1,a_2,\ldots,a_n,2^{a_1},a^{a_2},\ldots,2^{a_n}) < P_i^2(b_1,b_2,\ldots,b_n) +$$

$$= \sum_{j=1}^{n} |b_j - a_j| * K_j(b_1,b_2,\ldots,b_n).$$

The proof will be finished if we show that

$|b_j - a_j| \leq |\sin \pi b_j|$ for then $P_i^2(a_1,a_2,\ldots,a_n,2^{a_1},2^{a_2},\ldots,2^{a_n})$ will be less than a sum of $n + 1$ terms each of which is less than $\frac{1}{n+1}$.

We may assume without loss of generality that b_j is in $[0,\frac{1}{2}]$. Then $a_j = 0$. On $[0,\frac{1}{2}]$, $f(x) = \sin \pi x - x$ has a non-positive second derivative and hence is concave. Thus it takes its minimum at one of the end points. But $f(0) = 0$ and $f(\frac{1}{2}) = \frac{1}{2}$.

Q.E.D.

<u>Corollary</u>: For F in $\mathfrak{R}'$ the predicate 'there exist real numbers $b_1,b_2,\ldots,b_n$ such that $F(b_1,b_2,\ldots,b_n) < 0$' is recursively undecidable.

The next lemma will enable us to obtain the above corollary for expressions with only one variable. Note that we have not yet composed the primitive functions of $\mathfrak{R}_1$.

<u>Lemma 3</u>: Let $h(x) = x \sin x$ and $g(x) = x \sin x^3$. Then for any real numbers $a_1,a_2,\ldots,a_n$ and any $0 < \epsilon < 1$ there

exist $b > 0$ such that

$$|h(b) - a_1| < \epsilon$$
$$|h(g(b)) - a_2| < \epsilon$$
$$\cdots$$
$$|h(g(\ldots g(b)\ldots)) - a_n| < \epsilon.$$

Proof: The proof is by induction on n. Richardson first shows that for any two real numbers a_1 and a_2 there exists $b > 0$ such that $|h(b) - x_1| < \epsilon$ and $g(b) = x_2$. Let $c = |a_1| + |a_2| + \frac{2\pi}{\epsilon} + \epsilon + 1$. Pick b_2 in $(c, c + 2\pi)$ such that $h(b_2) = a_1$. Now pick b_1 so that

(1) $(b_2 - b_1)(b_2 + 1) < \epsilon$ and

(2) $b_2^3 - b_1^3 > 2\pi$.

This will be possible if

(3) $b_2 - \frac{\epsilon}{b_2+1} < \sqrt[3]{b_2^3 - 2\pi}$.

(3) can be proved by using the fact that if $f(x)$ is a monic polynomial and $a > h + 1$, where h is the absolute value of the negative coefficient of largest absolute value, then $f(a) > 0$. Applying this fact to the following polynomial in b_2 we obtain

$$b_2^4 + (2 - \frac{2\pi}{3\epsilon})b_2^3 + (1 - \epsilon - \frac{2\pi}{\epsilon})b_2^2 - (\epsilon + \frac{2\pi}{\epsilon})b_2 + \epsilon^2 - \frac{2\pi}{\epsilon} > 0.$$

This implies

$3b_2^2 \cdot \epsilon(b_2 + 1)^2 - 3b_2 \cdot \epsilon^2(b_2 + 1) + \epsilon^3 > 2\pi(b_2 + 1)^3$. Thus

$$b_2^3 - 3b_2^2(\tfrac{\epsilon}{b_2+1}) + 3b_2(\tfrac{\epsilon}{b_2+1})^2 - (\tfrac{\epsilon}{b_2+1})^3 < b_2^3 - 2\pi$$

for whence (3) follows. (2) implies that there exists b in (b_1, b_2) such that $g(b) = a_2$. Now

$$\begin{aligned}
|h(b) - h(b_2)| &\leq |b_2 \sin b_2 - b_2 \sin b| + |b_2 \sin b - b \sin b| \\
&\leq b_2|\sin b_2 - \sin b| + b_2 - b \\
&\leq b_2(b_2 - b) + b_2 - b \\
&< \epsilon \quad \text{by (1)}.
\end{aligned}$$

Hence b has the desired properties and the lemma is true for $n = 1$. Now suppose that it is true for $n = k$. Then there exists b' such that

$$|h(b') - a_2| < \epsilon$$

$$|h(g(b')) - a_3| < \epsilon$$

$$\cdots$$

$$|h(g(\ldots(g(b'))\ldots)) - a_{k+1}| < \epsilon.$$

By the preceding analysis there exists $b > 0$ such that $|h(b) - a_1| < \epsilon$ and $g(b) = b'$.

Q.E.D.

<u>Corollary</u>: For $G(x_1)$ in $\mathcal{R}_1$ the predicate 'there exist a real number a such that $G(a) < 0$' is recursively undecidable. (Note that $G(x)$ is in $\mathcal{R}$.)

Proof: Consider $G_i(x_1) = F_i(h(x_1), h(g(x_1)), \ldots, h(g(\ldots(g(x_1))\ldots)))$. If there exists an a such that $G(a) < 0$ then there exist $b_1, b_2, \ldots, b_n$ such that $F_i(b_1, b_2, \ldots, b_n) < 0$. Conversely if there exist $b_1, b_2, \ldots, b_n$ such that $F(b_1, b_2, \ldots, b_n) < 0$ then by lemma 2 there exist $b_1, b_2, \ldots, b_n$ such that $F(b_1, b_2, \ldots, b_n) = -1$. From the continuity of $G(x_1)$ and lemma 3 it follows that there exists an a such that $G(a) < 0$.

Q.E.D.

This corollary gives the undecidability of the decision problem for $\mathfrak{R}_1$ and hence completes the proof of theorem 1.

As a result of theorem 1, any class of expressions containing $\mathfrak{R}$ will not possess a canonical form.

Implications of the undecidability of Hilbert's tenth problem

Now we show that if Hilbert's tenth problem is undecidable then Richardson's result holds for a proper subset of $\mathfrak{R}$. Hilbert's tenth problem refers to one of the problems that David Hilbert [9] listed in a famous presentation in 1900. The problem is the one of deciding if an arbitrary polynomial (arbitrary with respect to degree and number of variables) with integral coefficients has integral roots. The problem is still unresolved but the evidence to date suggests that the problem is recursively undecidable. For this evidence see [4], [5], [6], and [7].

Let $\mathfrak{R}_2$ be the class of expressions generated by

(i) the rational numbers and the real number π,

(ii) the variable x,

(iii) the operations of addition, multiplication and composition,

(iv) the sin and absolute value functions.

Note that $\mathcal{R}_2$ is a proper subclass of $\mathcal{R}$ since it does not contain log 2 and the exp function.

Theorem 3: If Hilbert's tenth problem is recursively undecidable then for $E(x)$ in R_2 the predicate

$$'E(x) \equiv 0'$$

is recursively undecidable.

The proof is almost identical to the proof of theorem 1. Corresponding to the Davis, Putnam, Robinson theorem is the unsolvability of Hilbert's tenth problem. Thus we have by assumption that if $P(x_1,x_2,\ldots,x_n)$ is a polynomial with integral coefficients the predicate 'there exist integers $a_1,a_2,\ldots,a_n$ such that $P(a_1,a_2,\ldots,a_n) = 0$' is recursively undecidable. Then we have

Lemma 2': For each polynomial $P(x_1,x_2,\ldots,x_n)$ with integral coefficients there exists $F(x_1,x_2,\ldots,x_n)$ in $\mathcal{R}_3$ (where $\mathcal{R}_3$ is to $\mathcal{R}_2$ as $\mathcal{R}_1$ is to $\mathcal{R}$) such that

(i) there exist integers $a_1,a_2,\ldots,a_n$ such that

$P(a_1,a_2,\ldots,a_n) = 0$ if and only if

(ii) there exist real numbers $b_1,b_2,\ldots,b_n$ such that $F(b_1,b_2,\ldots,b_n) = -1$ if and only if

(iii) there exist real numbers $b_1,b_2,\ldots,b_n$ such that $F(b_1,b_2,\ldots,b_n) < 0$.

The proof is exactly like the proof of lemma 2 except that

$$F(x_1,x_2,\ldots,x_n) = (n+1)^2 [\sum_{j=1}^{n} (\sin^2 \pi x_j) * K_j^2(x_1,x_2,\ldots,x_n) + P^2(x_1,x_2,\ldots,x_n)] - 1$$

and hence does not involve the constant log 2 and the exp function as does the F of lemma 2. The remainder of the proof of theorem 3 is exactly like the proof of theorem 1.

Now consider the class of expressions $\mathfrak{R}_4$ generated by

(i) the rationals and π,

(ii) the variables $x_1,x_2,\ldots,x_n$,

(iii) the operations of addition, multiplication and restricted composition,

(iv) the functions sin and absolute value.

By restricted composition we mean that the primitive functions may not be rested. See appendix I for the BNF definition of $\mathfrak{R}_4$. Then we have

Theorem 4: If Hilbert's tenth problem is recursively undecidable, then for $E(x_1,x_2,\ldots,x_n)$ in $\mathfrak{R}_4$ the predicate '$E(x_1,x_2,\ldots,x_n) \equiv 0$'

is recursively undecidable.

<u>Proof</u>: Note that $F(x_1,x_2,\ldots,x_n)$ in lemma 2' is a member of $\mathcal{R}_4$ and hence lemma 2' holds for F in $\mathcal{R}_4$. Thus for $F(x_1,x_2,\ldots,x_n)$ in $\mathcal{R}_4$ we cannot decide if there exist real numbers $b_1,b_2,\ldots,b_n$ such that $F(b_1,b_2,\ldots,b_n) < 0$. Thus consider $E(x_1,x_2,\ldots,x_n) = |F(x_1,x_2,\ldots,x_n)| - F(x_1,x_2,\ldots,x_n)$ which is a member of $\mathcal{R}_4$. $E(x_1,x_2,\ldots,x_n) \not\equiv 0$ if and only if there exist real $b_1,b_2,\ldots,b_n$ such that $F(x_1,x_2,\ldots,x_n) < 0$.

Q.E.D.

This ends our discussion of undecidability results. These results will be used as a guide in the next chapter.

Chapter III

Canonical and Normal Forms

In this chapter more precise definitions of normal and canonical forms are given. Canonical forms are shown to exist for classes of exponential expressions that include most of $\mathcal{R}_3$ and $\mathcal{R}_4$. The existence of canonical forms implies, among other things, that functional equivalence is decidable. These results are then compared with some similar work of W. S. Brown and Richardson. Then we turn our attention to the radical expressions, i.e., rational roots of polynomials, and discuss the representation problem for these expressions. Some problems of algorithmic efficiency are also considered. The chapter concludes with a section that relates the work on exponential and radical expressions.

A _form_ is a generalized expression. For example $r_o + r_1 * x + \ldots + r_n * x^n$, where the r_i's are rational numbers, is a generalized polynomial expression, i.e., a polynomial form. A particular expression E is said to be in the form F or an instance of F if E matches F. This relation is denoted by E == F. Thus x and $1 + 3 * x^2$ match the above polynomial form but $(2 + 5 * x) * (x^2 + x^3)$ does not. An _f-normal form_ for a class of expressions $\mathcal{E}$ is a mapping f from $\mathcal{E}$ into $\mathcal{E}$ that satisfies for all E in $\mathcal{E}$ the following two properties:

(i) $f(E) \equiv E$

and

(ii) there exists a form F such that $f(E) == F$.

An <u>f-canonical form</u> is an f-normal form with the additional uniqueness property that for all E_1, E_2 in $\mathcal{E}$ such that $E_1 \equiv E_2$, $f(E_1) = f(E_2)$. If the particular f is clear from context or if we are speaking of an arbitrary f we shall frequently drop the prefix and simply use canonical (normal) form. If E is an expression such that $f(E) = E$ then E is said to be <u>in</u> (f-) <u>canonical</u> (normal) <u>form</u>. A class of expressions is called a <u>canonical</u> (normal) <u>class</u> or is said to <u>possess a canonical</u> (normal) <u>form</u> if there exists a canonical form for it. Further we adopt, without loss of generality, the convention that for all canonical forms f, if $E \equiv 0$ then $f(E) = 0$.

Frequently it is necessary to know that a total ordering can be imposed on a class of expressions. Usually this can be done in several different ways, but note that it can always be done by a lexicographical ordering scheme. In fact a well-ordering may be imposed on a class in this manner.

One further preliminary matter--a canonical (normal) form is not necessarily a computable function. For our purposes we need computable canonical (normal) forms. Rigorous proofs of the computability of the canonical (normal) forms will not be presented, but algorithms will be described for computing the canonical (normal) forms. Further Formula Algol programs that carry out these tasks are given in the appendices.

Examples

Let $\mathcal{P}$ be the class of polynomials generated by

(i) the rationals, the real number π, and the complex number i,

(ii) the variables $x_1, x_2, \ldots, x_n$,

(iii) the operations of addition, substraction and multiplication.

A rational number is in canonical form if it is an integer or if it is in the form p/q when p and q are integers, $q > 1$, and $\gcd(p,q) = 1$. A polynomial constant is in canonical form if it is an instance of the form

$$c_o + c_1 * \pi + \ldots + c_k * \pi^k$$

where the c_i are instances of $r_1 + i * r_2$, r_1 and r_2 being non-zero canonical rationals. A polynomial is in canonical form if it is an instance of the form

$$F = P_o(x_1, \ldots, x_{n-1}) + P_1(x_1, \ldots, x_{n-1}) * x_n + \ldots + P_k(x_1, \ldots, x_{n-1}) * x_n^k$$

where the $P_i(x_1, \ldots, x_{n-1})$ are non-zero canonical polynomials containing at most the variables $x_1, x_2, \ldots, x_{n-1}$.

It is well-known that there exists a function f mapping $\mathcal{P}$ into $\mathcal{P}$ such that $f(P) == F$ for all P in $\mathcal{P}$ and furthermore that f satisfies the uniqueness condition. There are other canonical forms for the class of polynomials; for example, the function that maps each polynomial into factored form.

<u>Sufficient conditions for the existence of a canonical form</u>

Given a class $\mathcal{E}$, closed under multiplication, a subclass $\mathcal{E}_2$ is linearly independent over a subclass $\mathcal{E}_1$ if for A_i in $\mathcal{E}_1$, X_i in $\mathcal{E}_2$, $i = 1,2,\ldots,n$, $A_1 * X_1 +\cdots+A_n * X_n \equiv 0$ implies that $A_1 \equiv A_2 \equiv\cdots\equiv A_n \equiv 0$. The following theorem establishes sufficient conditions for a class of expressions to have a canonical form.

<u>Theorem 1</u>: Let $\mathcal{E}$ be a class of expressions closed under multiplication. Suppose $\mathcal{E}_1$ and $\mathcal{E}_2$ are subclasses of $\mathcal{E}$ with the following two properties:

(i) $\mathcal{E}_1$ and $\mathcal{E}_2$ possess canonical forms f_1 and f_2 respectively.

(ii) All canonical members of $\mathcal{E}_2$ are linearly independent over $\mathcal{E}_1$.

Let $\mathcal{E}_2 : \mathcal{E}_1$ denote the set of all expressions $A_1 * X_1 + A_2 * X_2 +\cdots+ A_n * X_n$, $n = 1,2,\ldots,$ where

(i) A_i, X_i are in f_1-and f_2-canonical form respectively.

(ii) $A_i \neq 0$ for all $i = 1,2,\ldots,n$.

(iii) $X_i < X_j$ if $i < j$ where $<$ is any total ordering on $\mathcal{E}_2$.

If f is a mapping from $\mathcal{E}$ into $\mathcal{E}_2 : \mathcal{E}_1 \cup \{0\}$ such that $f(E) \equiv E$, then f is canonical.

<u>Proof</u>: The form for f is obviously $F_1^1 * F_1^2 + F_2^1 * F_2^2 + \cdots + F_n^1 * F_n^2$ where F^1 and F^2 are the forms for f_1 and f_2 respectively. To finish the proof it is only necessary to show that $E_1 \equiv E_2$ implies that $f(E_1) = f(E_2)$. Suppose $f(E_1) = A_1 * X_1 + A_2 * X_2 + \cdots + A_n * X_n$ and $f(E_2) = B_1 * Y_1 + B_2 * Y_2 + \cdots + B_m * Y_m$. Let $\{Z_1, Z_2, \ldots, Z_k\}$ be the distinct members of $\mathcal{E}_2$ occurring among the X_i and Y_i. Also assume that the Z_i's are in ascending order. Then

$$f(E_1) - f(E_2) \equiv C_1 * Z_1 + C_2 * Z_2 + \cdots + C_k * Z_k \equiv 0$$

$$\text{where } C_\ell = \begin{cases} \text{(i)} \quad A_i - B_j \text{ for some } i \text{ and } j \text{ if } Z_\ell \text{ appears in both } f(E_1) \text{ and } f(E_2). \\ \text{(ii)} \quad A_i \text{ if } Z_\ell \text{ appears in } f(E_1) \text{ but not in } f(E_2). \\ \text{(iii)} \quad -B_j \text{ otherwise.} \end{cases}$$

$f(E_1) - f(E_2) \equiv 0$ if and only if $C_\ell \equiv 0$, $\ell = 1,2,\ldots,k$. Thus $C_\ell \neq A_i$ and $C_\ell \neq -B_j$ since $A_i \neq 0 \neq B_j$. Hence $C_\ell = A_i - B_j \equiv 0$ which implies that $A_i = B_j$ since A_i and B_j are in canonical form. Thus $n = m = k$; $X_i = Y_i = Z_i$ and $A_i = B_i$, $i = 1,2,\ldots,n$. Hence $f(E_1) = f(E_2)$ and f is canonical.

Q.E.D.

This theorem is almost self evident. Its main purpose is to point out the main technique that is used to obtain new canonical forms, i.e., mapping classes of expressions into subclasses which are linear manifolds whose coefficient and basis sets are already known to possess canonical forms.

Now we are ready to find canonical forms for particular classes of expressions. We start by considering subclasses of $\mathcal{R}_4$ and $\mathcal{R}_2$ since our undecidabilility results of the last chapter imply that canonical forms cannot exist for the entire classes.

<u>Canonical form for first order exponential expressions</u>

In this section canonical forms for variations of the class $\mathcal{R}_4$ are presented. First we need some preliminary definitions and results. Let $C = \{(a_1,a_2,\ldots,a_n)\}$ be a set of n-tuples. For $n = 1$, C is called a <u>cascade set</u> if it contains infinitely many points. Now suppose cascade sets have been defined for $n < k$ and let $n = k$. Then the set

C of k-tuples is a cascade set if the set $C' = \{(x_1,x_2,\ldots,x_{k-1})$: there exist x_k such that $(x_1,x_2,\ldots,x_k)$ in $C\}$ is a cascade set and for each point $(x_1,\ldots,x_{k-1})$ in C' there exist infinitely many x_k such that $(x_1,\ldots,x_{k-1},x_k)$ is in C.

<u>Lemma 1</u>: Let $P(x_1,x_2,\ldots,x_n)$ be a polynomial in n variables over Ω and $C = \{(a_1,a_2,\ldots,a_n)\}$ be a cascade subset of Ω. If $P(x_1,x_2,\ldots,x_n) \equiv 0$ on C then $P(x_1,x_2,\ldots,x_n) \equiv 0$ on Ω.

<u>Proof</u>: The proof is by induction on n. For $n = 1$, $P(x_1)$ is a polynomial of one variable which has infinitely many zeros. Hence it must be the 0 polynomial. Now suppose the lemma holds for $n < k$ and $P(x_1,x_2,\ldots,x_k)$ is zero on the cascade set $C = \{(a_1,a_2,\ldots,a_k)\}$. Consider $P(x_1,x_2,\ldots,x_k)$ as a polynomial in the variable x_k, i.e., $P(x_1,x_2,\ldots,x_k) = P_0(x_1,\ldots,x_{k-1}) + P_1(x_1,\ldots,x_{k-1}) * x_k + \ldots + P_j(x_1,\ldots,x_{k-1}) * x_k^j$. Let $(a_1^0,a_2^0,\ldots,a_{k-1}^0)$ be an arbitrary point of C'. $P(a_1^0,a_2^0,\ldots,a_{k-1}^0,x_k)$ is a polynomial of one variable which is zero at infinitely many points and is hence identically zero. Thus $P_0(a_1^0,\ldots,a_{k-1}^0) = P_1(a_1^0,\ldots,a_{k-1}^0) = P_j(a_1^0,\ldots,a_{k-1}^0) = 0$. Thus each $P_i(x_1,x_2,\ldots,x_{k-1})$, $i = 0,1,\ldots,j$, is zero on the cascade set C' and hence by the induction hypothesis is identically zero on Ω. Hence $P(x_1,x_2,\ldots,x_k)$ is identically zero on Ω.

Q.E.D.

We also need a number theoretic result known as Lindemann's theorem (cf. [13], p. 117).

Theorem 2: Suppose $a_1, a_2, \ldots, a_k$ are distinct algebraic numbers. Then the set $\{e^{a_1}, e^{a_2}, \ldots, e^{a_k}\}$ is linearly independent over the algebraic numbers.

Now consider the class FOE of first-order exponentials generated by

(i) the rationals and the complex constant i,

(ii) the variables $x_1, x_2, \ldots, x_n$,

(iii) the operations of addition, multiplication, and restricted composition,

(iv) the exp function.

The class FOE contains as a subclass the class $\mathcal{P}$ of n variable polynomials over the field of complex rationals. These polynomials have a canonical form.

Theorem 3: Let $S_1(x_1, x_2, \ldots, x_n), S_2(x_1, x_2, \ldots, x_n), \ldots, S_k(x_1, x_2, \ldots, x_n)$ be distinct canonical members of $\mathcal{P}$. Then the set $\{\exp(S_1), \exp(S_2), \ldots, \exp(S_k)\}$ is linearly independent over $\mathcal{P}$.

Proof: Suppose

$$E(x_1, x_2, \ldots, x_n) = P_1(x_1, x_2, \ldots, x_n) * \exp(S_1) + \cdots$$
$$+ P_k(x_1, x_2, \ldots, x_n) * \exp(S_k) \equiv 0$$

where $P_i(x_1,x_2,\ldots,x_n)$ is in $\mathcal{Q}$, $i = 1,2,\ldots,k$. Suppose $(a_1,a_2,\ldots,a_n)$ is an arbitrary n-tuple of FOE constants. Then $E(a_1,a_2,\ldots,a_n) \equiv 0$ implies by Lindemann's theorem that either

(i) $P_i(a_1,a_2,\ldots,a_n) \equiv 0$ for all $i = 1,2,\ldots,k$

or

(ii) there exist $1 \leq i < j \leq k$ such that

$$S_i(a_1,a_2,\ldots,a_n) \equiv S_j(a_1,a_2,\ldots,a_n).$$

As a_n ranges over all the FOE constants either (i) or (ii) holds for infinitely many values of a_n. But this holds for arbitrary values of a_{n-1}, and hence as a_{n-1} ranges over the FOE constants there exist infinitely many values of a_{n-1} such that for each such value there exist infinitely many values of a_n such that for each such value either (i) or (ii) holds. Continuing in this fashion we see that there exist infinitely many values of a_1 such that for each value there exist infinitely many values of a_2 such that for each value

. . .

there exist infinitely many values of a_n such that for each value either (i) or (ii) holds. The set of all such n-tuples is a cascade set C. (ii) cannot hold on C for if it did then by lemma 1 $S_i \equiv S_j$ for all FOE constants, and hence $S_i = S_j$ since they are canonical. But by hypothesis S_i

S_j are distinct. Thus (i) must hold on C which implies by lemma 1 that the $P_i, i = 1,2,\ldots,k$, are functionally equivalent to 0.

Q.E.D.

Corollary 1: There exists a canonical form f for the first-order exponentials that maps each FOE into the form $P_1 * \exp(S_1) + P_2 * \exp(S_2) + \ldots + P_k * \exp(S_k)$ where the P_i are non-zero canonical members of $\mathcal{P}$ and the S_i are canonical members of $\mathcal{Q}$ with the property that, $S_i < S_j$ if $i < j$. Of course, if $E \equiv 0$, $f(E) = 0$.

Proof: Each FOE can be straight forwardly mapped into such an equivalent form by applications of the transformations $\exp(E_1) * \exp(E_2) \rightarrow \exp(E_1 + E_2)$ and $\exp(0) \rightarrow 1$. The fact that such a mapping is canonical follows from theorems 1 and 3.

Corollary 2: There exists a normal form for the class generated by

(i) the rationals and i,

(ii) the variables $x_1, x_2, \ldots, x_n$,

(iii) the operations of addition, substraction, multiplication, and restricted composition,

(iv) the exp, sin, cos, and tan functions.

Proof: By applying the following transformations

$$\sin(x) \rightarrow \frac{\exp(i*x) - \exp(-i*x)}{2i}$$

$$\cos(x) \rightarrow \frac{\exp(i*x) + \exp(-i*x)}{2}$$

and

$$\tan(x) \rightarrow \frac{\sin(x)}{\cos(x)}$$

each expression can be transformed into an equivalent expression which is a quotient of FOE's. Thus the normal form f maps each expression into an instance of the form P/Q where P and Q are the canonical form for the FOE's.

Q.E.D.

f is not canonical because P_1/Q_1 and P_2/Q_2 may be instances of P/Q such that $P_1/Q_1 \equiv P_2/Q_2$ but $P_1/Q_1 \not\equiv P_2/Q_2$. However there is a straight forward test for functional equivalence since $P_1/Q_1 \equiv P_2/Q_2$ if and only if $P_1 * Q_2 \equiv P_2 * Q_1$, $Q_1 \not\equiv 0 \not\equiv Q_2$. $P_1 * Q_2$ and $P_2 * Q_1$ are FOE's and hence are equivalent if and only if their canonical forms are the same. f would be canonical if the division operator were not included in the class.

This class of expressions contains all the primitives of $\mathcal{R}_4$ except the constant π and the absolute value function. Hence we have a canonical form for a large subclass of $\mathcal{R}_4$.

This shows, in a certain sense, that our results are fairly sharp.

Exponential expressions

In the FOE expressions composition is limited, i.e., the exp function cannot be nested. In this section a generalization of Lindemann's theorem is conjectured and the conjecture is used to obtain a canonical form for the class of exponential expressions. The exponential expressions are generated by

(i) the rationals and i,

(ii) the variable x,

(iii) the operations of addition, multiplication and composition,

(iv) the exp function.

The order of an exponential expression is the maximum number of nestings of the exp function. For example, all polynomials are of order 0, all FOE's are of order ≤ 1, and

$$\exp(\exp(\exp(x + 2) + 3 * i)) + \exp(x^2 + 5) + x^{10} + 1$$

is of order 3.

For the exponentials of order ≤ 1, theorem 3 gives a canonical form. Each order 1 expression is mapped into an expression of the form

$$(1) \quad P_1(x) * \exp(S_1(x)) + P_2(x) * \exp(S_2(x)) + \cdots + P_k(x) * \exp(S_k(*))$$

where the P_i are non-zero canonical polynomials and the S_i are distinct canonical polynomials, in ascending order.

Now define the mapping f on the exponential expressions as follows. If E has order ≤ 1, then $f(E)$ is the equivalent expression of the form (1). If f has been defined for expressions of order $\leq n - 1$ and E has order n, $f(E)$ is the equivalent expression of the form

$$(2) \quad P_1(x) * \exp(E_1(x)) + P_2(x) * \exp(E_2(x)) + \ldots + P_k(x) * \exp(E_k(x))$$

where the P_i are non-zero canonical polynomials and the $E_i(x)$ are f-form exponentials of order at most $n - 1$ with the property that $E_i < E_j$ if $i < j$.

<u>Conjecture</u>: Suppose $E_1, E_2, \ldots, E_k$ are distinct f-form exponential constants. Then the set of constants $\{\exp(E_1), \exp(E_2), \ldots, \exp(E_k)\}$ is linearly independent over the rationals.

If $E_1, E_2, \ldots, E_k$ are 0 order constants then the conjecture is a special case of Lindemann's theorem. However, the proof of the conjecture, if true, appears to be beyond current boundaries of number theoretic research since little seems to be known about such specific numbers as e^e. The conjecture implies that e^e is transcendental.

Assuming this conjecture, we can show that f is a canonical form for the exponential functions.

<u>Theorem 4</u>: If the above conjecture is true, then f is a canonical form for the exponential expressions.

<u>Proof</u>: It is only necessary to show that $f(E_1) \neq f(E_2)$ implies that $E_1 \not\equiv E_2$. We do this by showing that $f(E_1) - f(E_2) \not\equiv 0$. It is clear from the definition of f that $f(f(E_1) - f(E_2)) \neq 0$ if $f(E_1) \neq f(E_2)$. So it is sufficient to show that if E is of the form (2) and $E \neq 0$, then $E \not\equiv 0$. The proof is by induction on n, the order of E. By theorem 3 the result holds when $n = 1$. Assume the result holds for all expressions of form (2) with order less than n. Let

$$E(x) = P_1(x) * \exp(E_1(x)) + P_2(x) * \exp(E_2(x)) + \ldots$$
$$+ P_k(x) * \exp(E_k(x))$$

be an expression of order n. Assume $E(x) \equiv 0$. Consider any finite closed real interval I. For each rational r in I, $E(r) \equiv 0$. By the conjecture this implies that either

(i) $P_i(r) \equiv 0$ for all $i = 1,2,\ldots,k$

or

(ii) there exist $1 \leq i < j \leq k$ such that $E_i(r) \equiv E_j(r)$.

Since (i) or (ii) holds for every r in I then either (i) or (ii) holds for infinitely many r in I. Since each exponential expression is an entire analytic function and an analytic function is completely determined by its values at an infinite number of points on a closed interval, we have that either

(i) $P_i(x) \equiv 0$ for all $i = 1,2,\ldots,k$

or

(ii) $S_i(x) \equiv S_j(x)$.

But (i) does not hold by the definition of $E(x)$ and (ii) does not hold by the induction hypothesis. Thus we must conclude that $E(x) \not\equiv 0$.

Q.E.D.

Corollary 3 is an analogue of corollary 2.

Corollary 3: If the generalization of Lindemann's theorem is true then there exists a canonical form for the class generated by

(i) the rationals and i,

(ii) the variable x,

(iii) the operations of addition, subtraction, multiplication, division, and composition,

(iv) the exp, sin, cos, and tan functions.

Proof: The proof is the same as the proof for corollary 2 except that the canonical form for the exponential expressions is used instead of the canonical form for FOE. This corollary is very similar to a normal form theorem proved by W. S. Brown. Brown considers the class $\mathcal{B}$ of expressions generated by

(i) the rationals, π and i,

(ii) the variables $x_1, x_2, \ldots, x_n$ (denoted collectively by x),

(iii) the operations of addition, subtraction, multiplication, division and composition,

(iv) the exponential function.

He conjectures that if $E_1, \ldots, E_k$ are non-zero expressions in $\mathcal{B}$ such that the set $\{E_1, \ldots, E_k, i\pi\}$ is linearly independent over the rationals, then the set $\{\exp(E_1), \ldots, \exp(E_k), x, \pi\}$ is algebraically independent over the rationals. Then using this conjecture, he shows that there exists a normal form f for the class $\mathcal{B}$ that maps each expression into an equivalent expression of the form

$$\frac{g(\exp(E_1), \ldots, \exp(E_n), x, \pi, \omega_m)}{h(\exp(E_1), \ldots, \exp(E_n), x, \pi)}$$

where

(i) g and h are relatively prime polynomials over the rationals,

(ii) the degree of g in $\omega_m = \exp(i \cdot \pi/2m)$ is less than the degree of the minimal polynomial for the root of unity ω_m,

(iii) $E_1, E_2, \ldots, E_n$ are distinct normalized expressions,

(iv) the set $\{E_1, E_2, \ldots, E_n, i\pi\}$ is linearly independent over the rationals.

Further f is shown to have the property that $E \equiv 0$ if and only if $f(E) = 0$.

This is a very nice result in that the class $\mathfrak{B}$ contains all the primitives of $\mathfrak{R}_2$ except the absolute value function.[1]

Richardson has also proved a theorem that is somewhat similar to theorem 4 and Brown's result. He considers the class of expressions $\mathfrak{R}$ generated by

(i) the rationals and π,

(ii) the variables $x_1, x_2, \ldots, x_n$,

(iii) the operations of addition, subtraction, multiplication, division and composition,

(iv) the functions exp, sin, cos, and $\log|x|$.

He shows that if one assumes a decision procedure for deciding whether or not $\mathfrak{R}$ constants are equivalent to 0 and a procedure for deciding if $\mathfrak{R}$ functions are completely defined on an arbitrary interval, then a decision procedure can be

---

[1]Our results and Brown's results were obtained independently.

given to decide if an arbitrary $\mathfrak{R}$ expression is functionally equivalent to 0. He does not use a canonical form approach but by differentiation, multiplication and division finds a set of expressions $E_1,\ldots,E_k$ for which the predicates '$E_i \equiv 0$' are decidable and such that the original expression $E \equiv 0$ if and only if $E_i \equiv 0$ for all $i = 1,2,\ldots,k$.

Radical expressions

Now we turn our attention to a somewhat different class of expressions, the radical expressions. Radical expressions are rational roots of polynomials and rational expressions. The radical expressions are formed from

(i) the rationals,

(ii) the variable x,

(iii) the operations of addition, subtraction, multiplication, division, and rational exponentiation.

Rational exponentiation is the operation of raising expressions to rational powers. This operation may not be nested, i.e., expressions like $((x^2 + 2x)^{1/2} + 5)^{2/3}$ are not radical expressions as the expressions are defined here.

The radical expressions are to be interpreted as algebraic functions. In particular, this means that for each expression $E(x)$ there must exist an irreducible polynomial $P(y,x)$ such that $P(E(x),x) \equiv 0$. Thus expressions such as $(x^2)^{1/2}$ are to be interpreted as being functionally equivalent to either x or $-x$ depending on the branch of the square root function that is used. $(x^2)^{1/2} \not\equiv |x|$ for $|x|$ does not satisfy an irreducible polynomial and hence is not an algebraic function. In general, the single-valued branches are not analytic on the whole real line, and hence their domains must be restricted in a suitable manner.

If $R(x)$ is a rational expression, i.e., a member of the field $\Gamma(x)$ then $[R(x)]^{1/m}$, m a positive integer, is taken to be any fixed root of the polynomial equation $y^m - R(x) = 0$. In order to obtain a representation for a radical expression we shall determine from the expression a normal algebraic extension field of $\Gamma(x)$ to which the expression belongs. Given such a field we shall employ standard representations for the elements of such fields. For example,

$$(1) \qquad \frac{2^{1/4} + 3^{1/2} \cdot [-(x^2 + 1)]^{1/3}}{[x^2 + 1]^{1/2}}$$

is a member of the field $\Gamma(x)(\xi_{12}, 2^{1/4}, [x^2 + 1]^{1/6})$ where ξ_{12} is a primitive 12-th root of unity. Each element of this field may be uniquely represented in the form

$$(2) \quad \alpha_0 + \alpha_1[x^2 + 1]^{1/6} + \alpha_2[x^2 + 1]^{2/6} + \alpha_3[x^2 + 1]^{3/6} + \alpha_4[x^2 + 1]^{4/6} + \alpha_5[x^2 + 1]^{5/6}$$

where $\alpha_i (i = 0,1,\ldots,5)$ is in $\Gamma(x)(\xi_{12}, 2^{1/4})$. Each element of this field may be uniquely represented in the form $\beta_o + \beta_1[2]^{1/4} + \beta_2[2]^{2/4} + \beta_3[2]^{3/4}$ where $\beta_i (i = 0,1,2,3)$ is in $\Gamma(\xi_{12})$ in which field each element may be uniquely represented in the form $\gamma_o + \gamma_1\xi_{12} + \gamma_2\xi_{12}^2 + \gamma_3\xi_{12}^3$ with $\gamma_i (i = 0,1,2,3)$ in $\Gamma(x)$ where expressions have unique representations. In particular (1) may be written in the form (2) and is thus

$$(\frac{1}{x^2 + 1})\ (2^{1/4})[x^2 + 1]^{3/6} + [(\frac{1}{x^2 + 1})\xi_{12} + (\frac{1}{x^2 + 1})\ \xi_{12}^3][x^2 + 1]^{5/6}$$

Note that in a radical expression, a root such as $[R(x)]^{1/m}$ must be interpreted consistently wherever it appears in the expression. In the above expression $[R(x)]^{1/6}$ is taken to be the root of $y^6 - R(x)$ such that $[1]^{1/6} = 1$. Thus to be consistent $[-1]^{1/6} = \xi_{12}$ if ξ_{12} is the primitive 12-th root of unity $\exp(\pi \cdot i/6)$.

In general we shall be able to find, given any finite number of radical expressions, a common field to which they belong. This field will have the property that it can be constructed in a finite number of steps from the field $\Gamma(x)$ of rational expressions and the elements of the field will have unique representations. The basic procedure is based on the following well-known facts. A field is <u>explicitly given</u> if its elements can be uniquely represented with a finite alphabet and if the operations of addition, subtraction, multiplication, and division can be carried out in a finite number of steps. The field of rational numbers can be given explicitly. If the field Ψ is given explicitly, then every simple transcendental extension $\Psi(x)$ and every simple algebraic extension $\Psi(\Theta)$, with given irreducible defining equation $P(\Theta) = 0$, is explicitly given.

In the algebraic case the field elements may be uniquely represented in the form

$$\alpha_o + \alpha_1 \cdot \Theta + \alpha_2 \cdot \Theta^2 + \ldots + \alpha_{n-1} \cdot \Theta^{n-1}$$

where n is the degree of the defining equation for Θ. The operations are carried out as with polynomials over Ψ with the exception that the final result is reduced modulo $P(\Theta)$. In the transcendental case the elements of $\Psi(x)$ are simply rational expressions over Ψ in x and they have a unique representation as we have seen. Furthermore, if polynomials can be factored in Ψ in a finite number of steps then they can be factored in $\Psi(x)$ and $\Psi(\Theta)$ in a finite number of steps. Polynomials over Γ, the field of rational numbers, can be factored in a finite number of steps.

For further discussion along these lines see section 42 of van der Waerden [23] and Johnson [11].

Any radical expression can be straight-forwardly transformed into an equivalent expression which is a quotient of radical polynomials where each member of Γ appearing in the polynomials is an integer. A <u>radical polynomial</u> is a radical expression in which all powers are positive and which does not contain the division operator. For example the radical expression

$$3/2 * [\frac{2/5x^3 + x}{2x^4 + 9/2}]^{9/5} + [-x^{13} + 17/4x^9]^{11/3}$$

is equivalent to

$$\frac{3*[4x^3+10x]^{9/5}*[4]^{11/3}+[-4x^{13}+17x^9]^{11/3}*2*[20x^4+45]^{9/5}}{2*[20x^4+45]^{9/5}*[4]^{11/3}}$$

which is a quotient of radical polynomials. Radical polynomials are sums of products of polynomials and roots of polynomials in $J[x]$ where J is the ring of rational integers. A polynomial $P(x) = a_n \cdot x^n + a_{n-1} \cdot x^{n-1} + \ldots + a_1 \cdot x + a_o$ in $J[x]$ is said to be <u>primitive</u> if $P(x)$ is irreducible and

(i) $a_n > 0$

and

(ii) $\gcd(a_n, \ldots, a_1, a_o) = 1$ if $n > 0$ or else a_o is a prime in J.

Given any quotient of radical polynomials suppose $P_1(x)$, $P_2(x),\dots,P_n(x)$ are the elements of $J[x]$ which appear under a radical. In our example above $n = 4$ and 4, $4x^3 + 10x$, $20x^4 + 45$, and $-4x^{13} + 17x^9$ are the elements of $J[x]$ appearing under a radical. Each of the polynomials may be factored in a finite number of steps into products of -1 and powers of primitive polynomials.

The polynomials of our example factor into the products of primitive polynomials and -1 as follows:

$$4 = 2^2$$
$$4x^3 + 10x = 2 * x * (2x^2 + 5)$$
$$20x^4 + 45 = 5 * (4x^4 + 9)$$
$$-4x^{13} + 17x^9 = (-1) * x * (4x^4 + 17).$$

Now we want to determine the radical degree of each primitive polynomial. If $P(x)$ is a primitive polynomial appearing in a radical expression $E(x)$ and $P(x)$ appears in the expression raised to the rational powers $p_1/q_1, p_2/q_2, \dots, p_s/q_s$ where p_i, q_i are relatively prime integers, then $m = \mathrm{lcm}(q_1, q_2, \dots, q_s)$ is the <u>radical degree</u> of $P(x)$ in $E(x)$. Now let $P_1(x), P_2(x), \dots, P_n(x)$ be the primitive polynomials appearing in a quotient $E(x)$ of radical polynomials. If the $P_i(x)$ have the radical degrees $m_1, m_2, \dots, m_n$ respectively and if m_o is the radical degree of (-1) in $E(x)$, then we claim that

$E(x)$ belongs to $\Gamma(x)(\xi_m, [P_1(x)]^{1/m_1}, \ldots, [P_n(x)]^{1/m_n})$ where $m = \mathrm{lcm}(2m_o, m_1, \ldots, m_n)$. This is obviously so because $\Gamma(x)(\xi_m, [P_1(x)]^{1/m_1}, \ldots, [P_n(x)]^{1/m_n})$ consists by definition of rational combinations of ξ_m, $[P_i(x)]^{1/m_i}$ $(i = 1,2,\ldots,n)$ over $\Gamma(x)$ and $E(x)$ is a rational combination of $(-1)^{1/m_o}$ and $[P_i(x)]^{1/m_i}$. But $(-1)^{1/m_o}$ is an m-th root of unity since m is a multiple of $2m_o$. We assert that $\Gamma(x)(\xi_m, [P_1(x)]^{1/m_1}, \ldots, [P_n(x)]^{1/m_n})$ can be explicitly given. This is so because Γ can be explicitly given and hence $\Gamma(x)$ since it is a simple transcendental extension of Γ. The cyclotomic polynomial $\Phi_m(x)$ can always be explicitly constructed. For this construction see, for instance, Section 36 of van der Waerden. Hence $\Gamma(x)(\xi_m)$ can be given explicitly. For any $i (i = 1,2,\ldots,n)$ if $\Gamma(x)(\xi_m, [P_1(x)]^{1/m_1}, \ldots, [P_{i-1}(x)]^{1/m_{i-1}})$ is given explicitly then $\Gamma(x)(\xi_m, [P_1(x)]^{1/m_1}, \ldots, [P_i(x)]^{1/m_i})$ can be given explicitly. In order that Ψ_i (let Ψ_i denote $\Gamma(x)(\xi_m, [P_1(x)]^{1/m_1}, \ldots, [P_i(x)]^{1/m_i})$, $i = 0,1,\ldots,n$) be given explicitly it is only necessary to determine the minimal polynomial of $[P_i(x)]^{1/m_i}$ over Ψ_{i-1}. But the minimal polynomial is a factor of $y^{m_i} - P_i(x)$. Since Ψ_{i-1} is given explicitly we can factor $y^{m_i} - P_i(x)$ into irreducible factors and hence determine the minimal polynomial for $[P_i(x)]^{1/m_i}$.

With this knowledge we can now say

<u>Theorem 5</u>: Let $E(x)$ be any radical expression. The predicate '$E(x) \equiv 0$' is solvable.

<u>Proof</u>: It is only necessary to explicitly construct an extension field Ψ_n of Γ to which $E(x)$ belongs and then to find the unique representation of $E(x)$ in Ψ_n. We have already indicated how to construct Ψ_n. It is only necessary to indicate how the representation for $E(x)$ in Ψ_n can be found. $E(x)$ can be written as a quotient of radical polynomials. The radical polynomials can be considered as polynomials in the $n + 1$ variables ξ_m, $[P_1(x)]^{1/m_1}$, $\ldots, [P_n(x)]^{1/m_n}$ over $\Gamma(x)$ and can be put in the canonical form for such polynomials. Now using the defining equation for Ψ_1 over Ψ_o all the coefficients involving only ξ_m can be reduced to their unique representation in Ψ_1. Now if all the coefficients involving only ξ_m, $[P_1(x)]^{1/m_1}, \ldots,$ $[P_i(x)]^{1/m_{i-1}}$ have been put in their unique form in Ψ_{i-1}, then the coefficients involving $[P_i(x)]^{1/m_i}$ may be put in unique form in Ψ_i by using the defining equation for $[P_i(x)]^{1/m_i}$ over Ψ_{i-1}. Continuing in this fashion, we eventually obtain a representation for $E(x)$ as a quotient of two elements of Ψ_n. If the denominator is not 0 then

the division can be carried out (see section 19 of van der Waerden) to obtain a unique representation for $E(x)$ in Ψ_n. If the denominator is 0 then the expression is undefined. Then $E(x) \equiv 0$ if and only if the representation of $E(x)$ in Ψ_n is 0.

Q.E.D.

Theoretically speaking we have already said all that needs to be said. However, the above described algorithms are highly impractical. One of the primary reasons is that the factorization algorithms for polynomials over algebraic extensions of the rationals are not practical. The algorithms usually generate a large set of possible factors and then divide the original polynomial by each member of the set to see if any one is a factor. Only after trying all members of the generated set does one ascertain that an irreducible polynomial is in fact irreducible. Thus, the algorithms are particularly inefficient when dealing with irreducible polynomials. In many cases the polynomials $y^{m_i} - P_i(x)$ are irreducible and if we could determine this a priori without calling on the factorization algorithm, one of the main sources of inefficiency in the algorithm could be eliminated. It is the purpose of the next section to investigate this problem.

Irreducibility Considerations

The results of this section draw heavily on the material on pure equations as given in Tschebotaröw [22]. So let us begin by quoting the results from Tschebotaröw that we shall need. If m is an integer ξ_m denotes a primitive m-th root of unity, Ψ is a number field, i.e., a subfield of the complex numbers.

Theorem 6 (Tschebotaröw, p. 291): Let a be an element of Ψ and $m = 2^{\mu_o} q_1^{\mu_1}, \ldots, q_s^{\mu_s}$ be the prime decomposition of the positive integer m. The polynomial $y^m - a$ is irreducible over Ψ if and only if the polynomials

$$x^{2^{\mu_o}} - a,\ x^{q_1^{\mu_1}} - a, \ldots, x^{q_s^{\mu_s}} - a$$

are irreducible over Ψ.

Theorem 7 (Tschebotaröw, p. 291): Let q be an odd prime, μ a natural number, and a an element of the field Ψ. The pure polynomial $x^{q^\mu} - a$ is reducible over Ψ if and only if $x^q - a$ is reducible over Ψ, i.e., if and only if a is the q-th power of an element in Ψ.

Theorem 8 (Tschebotaröw, p. 293): Let a be an element of Ψ and μ an integer ≥ 2. The polynomial $x^{2^\mu} - a$ is reducible over Ψ if and only if the polynomial $x^4 - a$

is reducible over Ψ, i.e., if there is a b in Ψ such that either $b^2 = a$ or $-4b^4 = a$.

Lemma 2 (Tschebotaröw, p. 310, problem 6): Let m, m' be two positive integers such that $\gcd(m,m') = 1$. Then the cyclotomic polynomial $\Phi_m(x)$ is irreducible over $\Gamma(\xi_{m'})$.

Proof: $\xi_m \cdot \xi_{m'}$ is a primitive $(m \cdot m')$-th root of unity. Hence $\Gamma(\xi_{m'})(\xi_m)$ must have degree at least $\varphi(m \cdot m') = \varphi(m') \cdot \varphi(m)$ (φ is Euler's φ-function), which is true if and only if $\Phi_m(x)$ is irreducible over $\Gamma(\xi_{m'})$.

Q.E.D.

Following Tschebotaröw we say that a field Ψ has the E_m property if the cyclotomic polynomial $\Phi_m(x)$ is irreducible over Ψ. Γ has the E_m property for every positive integer m. Now we prove the following:

Lemma 3: Let m be an odd positive integer, $x^m - a$ an irreducible polynomial over Γ. If m' is a positive integer such that $\gcd(m,m') = 1$, then $x^m - a$ is irreducible over $\Gamma(\xi_{m'})$.

Proof: By contradiction. Suppose $x^m - a$ is reducible over $\Gamma(\xi_{m'})$. If $m = q_1^{\mu_1}, \ldots, q_s^{\mu_s}$ is the prime decomposition of m, then by the above theorems from Tschebotaröw $\Gamma(\xi_{m'})$ must contain a root of $x^{q_i} - a$ for at least one integer i in

[1,s]. Since $\Gamma(\xi_{m'})$ is a normal extension of Γ, if it contains one root of $x^{q_i} - a$, it must contain them all. The quotient of any two different roots of $x^{q_i} - a$ is a primitive q_i-th root of unity. Thus $\Gamma(\xi_{m'})$ contains a q_i-th root of unity which contradicts the previous lemma which states that $\Phi_{q_i}(x)$ is irreducible over $\Gamma(\xi_{m'})$.

Q.E.D.

In order to prove the next lemma we need the following theorem from Tschebotaröw, p. 299.

<u>Theorem 9</u>: Let q be an odd prime, k a natural number and let Ψ have the property E_{q^k}. If the polynomial $x^{q^k} - a$ is irreducible over Ψ, then it is still irreducible over $\Psi(\xi_{q^k})$.

Now we prove a generalization of Tschebotaröw's remark following the theorem.

<u>Lemma 4</u>: Let m be an odd positive integer, n any positive integer. If $x^m - a$ is irreducible over Γ then it is irreducible over $\Gamma(\xi_n)$.

<u>Proof</u>: Let $m = p_1^{\mu_1}, \ldots, p_s^{\mu_s}$ $(\mu_i > 0,\ i = 1,2,\ldots,s)$ be the prime decomposition of m and $n = 2^{\nu_o} p_1^{\nu_1} \ldots p_s^{\nu_s} \cdot p_{s+1}^{\nu_{s+1}}, \ldots, p_{s+t}^{\nu_{s+t}}$ where $\nu_i \geq 0$ for $i = 0,1,\ldots,s$ and $\nu_i > 0$ for $i = s+1,\ldots,s+t$. Let $n' = 2^{\nu_o} p_{s+1}^{\nu_{s+1}} \ldots p_{s+t}^{\nu_{s+t}}$ and

$m' = p_1^{\mu_1+\nu_1} \cdots p_s^{\mu_s+\nu_s}$. $x^m - a$ is irreducible over Γ if and only if $x^{p_i^{\mu_i}} - a$ is irreducible for $i = 1,2,\ldots,s$. $x^{p_i^{\mu_i}} - a$ is irreducible over $\Gamma(\xi_{n'})$ by the previous lemma since $\gcd(p_i^{\mu_i}, n') = 1$. Now let m'' be the positive integer such that $m' = p_i^{\mu_i+\nu_i} \cdot m''$. Then the previous lemma implies that $x^{p_i^{\mu_i}} - a$ is irreducible over $\Gamma(\xi_{n'}, \xi_{m''})$ since $\gcd(p_i^{\mu_i}, m'') = 1$. Further $\Gamma(\xi_{n'}, \xi_{m''})$ has the E_ℓ property where $\ell = p_i^{\mu_i+\nu_i}$. Hence by the immediately preceeding theorem from Tschebotaröw $x^\ell - a$, and hence $x^{p_i^{\mu_i}} - a$, are irreducible over $\Gamma(\xi_{n'}, \xi_{m''}, \xi_\ell) = \Gamma(\xi_n)$. Thus $x^m - a$ must be irreducible over $\Gamma(\xi_n)$.

Q.E.D.

Now we give our first general irreducibility result.

<u>Theorem 10</u>: Let ℓ be a positive integer, m an odd positive integer and $p_1, p_2, \ldots, p_n$ distinct positive prime integers. Then $x^m - p_n$ is irreducible over $\Gamma(\xi_\ell, [p_1]^{1/m}, \ldots, [p_{n-1}]^{1/m})$.

<u>Proof</u>: We shall assume $m|\ell$ and then show at the end of the proof that this assumption is not actually necessary. The proof is by induction on n. For $n = 1$ the theorem follows from the preceeding lemma. Let Ψ_i, $i = 0,1,\ldots,n$ denote the field $\Gamma(\xi_\ell, [p_1]^{1/m}, \ldots, [p_i]^{1/m})$.

Suppose the theorem is true for all $k < n$. Let $m = q_1^{\mu_1} \ldots q_s^{\mu_s}$ be the prime decomposition of m. By Tschebotaröw $x^m - p_n$ is reducible if and only if for some integer i in $[1,s]$, $x^{q_i} - p_n$ is reducible. This is true if and only if there exist in Ψ_{n-1} a root of the equation $x^{q_i} = p_n$. So suppose there exist γ in Ψ_{n-1} such that $\gamma^{q_i} = p_n$. Then by the induction hypothesis each element of Ψ_{n-1} and in particular γ may be represented uniquely in the form

$$(1) \quad \gamma = \alpha_0 + \alpha_1 [p_{n-1}]^{1/m} + \ldots + \alpha_{m-1} [p_{n-1}]^{(m-1)/m}$$

where α_i is in Ψ_{n-2}, $i = 0,1,\ldots,m-1$. Now consider an element σ_{n-1} of the Galois group of $\Psi_{n-1} : \Psi_{n-2}$ that has the property

$$\sigma_{n-1}([p_{n-1}]^{1/m}) = \xi_m [p_{n-1}]^{1/m}.$$

There exists such an element since $y^m - p_{n-1}$ is by the induction hypothesis irreducible over Ψ_{n-2}, and hence the Galois group of $\Psi_{n-1} : \Psi_{n-2}$ is transitive. Applying σ_{n-1} to (1) we obtain

$$(2) \quad \xi_m^{j_{n-1}} \gamma = \alpha_0 + \alpha_1 \xi_m [p_{n-1}]^{1/m} + \alpha_2 \xi_m^2 [p_{n-1}]^{2/m} + \ldots$$

$$+ \alpha_{m-1} \xi_m^{m-1} [p_{n-1}]^{(m-1)/m}$$

where $j_{n-1} = k_{n-1} \cdot (m/q_i)$, $0 \leq k_{n-1} < q_i$.

This implies that

$$(3)\quad \gamma = \alpha_o \xi_m^{-j_{n-1}} + \alpha_1 \xi_m^{1-j_{n-1}} [p_{n-1}]^{1/m} + \ldots + \alpha_{m-1} \xi_m^{m-1-j_{n-1}} [p_{n-1}]^{(m-1)/m}.$$

Thus from (1) and (3) and the linear independence of the $[p_{n-1}]^{j/m}$ $(j = 0,1,\ldots,m-1)$ over Ψ_{n-2}, we have that

$$(4)\quad \begin{cases} \alpha_o = \alpha_o \xi_m^{-j_{n-1}} \\ \alpha_1 = \alpha_1 \xi_m^{1-j_{n-1}} \\ \cdots \\ \alpha_k = \alpha_k \xi_m^{k-j_{n-1}} \\ \cdots \\ \alpha_{m-1} = \alpha_{m-1} \xi_m^{m-1-j_{n-1}} \end{cases}$$

This implies that $\alpha_k = 0$ unless $k = j_{n-1}$. Thus

$$(5)\quad \gamma = \alpha_{j_{n-1}} [p_{n-1}]^{(j_{n-1})/m} .$$

$\alpha_{j_{n-1}}$ may be written uniquely in the form

$$\alpha_{j_{n-1}} = \beta_o + \beta_1 [p_{n-2}]^{1/m} + \ldots + \beta_{m-1} [p_{n-2}]^{(m-1)/m}$$

where $\beta_i (i = 0,1,\ldots,m)$ is in Ψ_{n-3}. Thus

$$(6) \quad \gamma = (\beta_o + \beta_1 [p_{n-2}]^{1/m} + \ldots + \beta_{m-1}[p_{n-2}]^{(m-1)/m}) [p_{n-1}]^{(j_{n-1})/m}.$$

Consider σ_{n-2} the element of the Galois group of $\Psi_{n-2}: \Psi_{n-3}$ with the property that $\sigma_{n-2}([p_{n-2}]^{1/m}) = \xi_m [p_{n-2}]^{1/m}$. σ_{n-2} can be extended in the usual way to an element of the Galois group of $\Psi_{n-1}: \Psi_{n-3}$. Applying σ_{n-2} to (6) we obtain

$$(7) \quad \xi_m^{j_{n-2}} \gamma = (\beta_o + \beta_1 \xi_m [p_{n-2}]^{1/m} + \ldots + \beta_{m-1}\xi_m^{m-1}[p_{n-2}]^{(m-1)/m}) \cdot [p_{n-1}]^{(j_{n-1})/m}$$

where $j_{n-2} = k_{n-2}(m/q_i)$, $0 \leq k_{n-2} < q_i$. By analysis similar to the above we can see that $\gamma = \beta_{j_{n-2}} [p_{n-2}]^{(j_{n-2})/m} \cdot [p_{n-1}]^{(j_{n-1})/m}$. Continuing in this fashion we obtain, after $n - 1$ steps,

$$(8) \quad \gamma = \omega [p_1]^{j_1/m} [p_2]^{j_2/m} \ldots [p_{n-1}]^{(j_{n-1})/m}$$

where ω is in $\Gamma(\xi_m)$ and $j_\ell = k_\ell (m/q_i)$ $(\ell = 1,2,\ldots,n - 1)$ where k_ℓ is an integer in $[0,q_i)$. Thus

$$(9)\qquad p_1^{k_1}\cdot p_2^{k_2}\cdots p_{n-1}^{k_{n-1}}\cdot \omega^{q_i} - p_n = 0.$$

By Eisenstein's criterion this polynomial in ω is irreducible over Γ and hence by lemma 4 is irreducible over $\Gamma(\xi_\ell)$. But this is a contradiction since ω, an element of $\Gamma(\xi_\ell)$, is a root of the equation. Thus we must conclude that $x^{q_i} - p_n$ is irreducible over $\Gamma(\xi_\ell)$. Now if $m \nmid \ell$, then by the above analysis $y^m - p_n$ is irreducible over $\Gamma(\xi_{\ell\cdot m}, [p_1]^{1/m}, \ldots, [p_{n-1}]^{1/m})$ and hence is irreducible over Ψ_n since $\Gamma(\xi_{\ell\cdot m}, [p_1]^{1/m}, \ldots, [p_{n-1}]^{1/m}) \supset \Psi_n$.

Q.E.D.

Now for the second general irreducibility theorem.

<u>Theorem 11</u>: Let m be a positive integer, $P_1(x), P_2(x), \ldots, P_n(x)$ be distinct, irreducible, primitive polynomials over $J[x]$ with degree $P_i(x) > 0$ $(i = 1,2,\ldots,n)$. Then $P(y) = y^m - P_n(x)$ is irreducible over $\Omega(x)([P_1(x)]^{1/m}, \ldots, [P_{n-1}(x)]^{1/m})$ where $[P_i(x)]^{1/m}$ is any fixed root of $y^m - P_i(x) = 0$.

<u>Proof</u>: The proof is by induction on n. Let $m = 2^{\mu_0} q_1^{\mu_1} \ldots q_s^{\mu_s}$ be the prime decomposition of m. Consider the case

when $n = 1$. Then by Tschebotaröw[1], $P(y)$ is irreducible if and only if

(1) $y^2 - P_1(x)$ and

(2) $y^4 - P_1(x)$ and

(3) $y^{q_i} - P_1(x)$ $(i = 1,2,\ldots,s)$

are irreducible. Suppose at least one of (1), (2), and (3) is reducible. Then by Tschebotaröw there must exist $B(x)$ in $\Omega(x)$ such that either

$$(4)\begin{cases} B^2(x) = P_n(x) \text{ or} \\ -4B^4(x) = P_n(x) \text{ or} \\ B^{q_i}(x) = P_n(x). \end{cases}$$

Clearly we may assume $B(x)$ is in $\Omega[x]$. Since degree $P_1(x) > 0$, degree $B(x) > 0$. But then $P_1(x)$ must have multiple zeroes in Ω. But since $P_1(x)$ is irreducible over Γ this is not possible. Thus we must conclude that $y^m - P_1(x)$ is irreducible over $\Omega(x)$. Let Ψ_i denote the field $\Omega(x)([P_1(x)]^{1/m},\ldots,[P_i(x)]^{1/m})$, $i = 0,1,\ldots,n$, and assume that the theorem is true for all $k < n$. For $y^m - P_n(x)$ to be reducible over Ψ_{n-1} there must exist $B(x)$ in Ψ_{n-1} such that at least one of the equations in (4) holds.

---

[1]Strictly speaking, Tschebotaröw's theorems do not apply to $\Omega(x)$ since $\Omega(x)$ is not a number field. But almost certainly it is sufficient in Tschebotaröw to know only that the characteristic of the field is 0. So we shall use Tschebotaröw's theorems with such an assumption.

Again we will show that this assumption leads to a contradiction. First assume that $B^{q_i}(x) = P_n(x)$. By the induction hypothesis each element of Ψ_{n-1}, and in particular $B(x)$, may be uniquely represented in the following way.

$$(5) \quad B(x) = \alpha_o + \alpha_1 [P_{n-1}(x)]^{1/m} + \ldots + [B_{n-1}(x)]^{(m-1)/m}$$

where α_i $(i = 0,1,\ldots,m-1)$ is in Ψ_{n-2}. Consider the element σ_{n-1} of the Galois group of $\Psi_{n-1} : \Psi_{n-2}$ that maps $[P_{n-1}(x)]^{1/m}$ onto $\xi_m [P_{n-1}(x)]^{1/m}$. There is such an element in the group since by the induction hypothesis $y^m - P_{n-1}(x)$ is irreducible over Ψ_{n-2}, and hence the Galois group of $\Psi_{n-1} : \Psi_{n-2}$ is transitive. Applying σ_{n-1} to (5) we obtain

$$\xi_m^{j_{n-1}} \cdot B(x) = \alpha_o + \alpha_1 \xi_m [P_{n-1}(x)]^{1/m} + \ldots + \alpha_{m-1} \xi_m^{m-1} [P_{n-1}(x)]^{(m-1)/m}$$

where $j_{n-1} = k_{n-1} \cdot (m/q_i)$, $0 \leq k_{n-1} < q_i$. Thus

$$(6) \quad B(x) = \alpha_o \cdot \xi_m^{-j_{n-1}} + \alpha_1 \xi_m^{1-j_{n-1}} [P_{n-1}(x)]^{1/m} + \ldots$$
$$+ \alpha_{m-1} \xi_m^{m-1-j_{n-1}} [P_{n-1}(x)]^{(m-1)/m}.$$

From the linear independence of the $[P_{n-1}(x)]^{1/m}$, $j = 0,1,\ldots,m-1$, and from equations (5) and (6) we obtain

$$\begin{cases} \alpha_o = \alpha_o \xi_m^{-j_{n-1}} \\ \alpha_1 = \alpha_1 \xi_m^{1-j_{n-1}} \\ \cdot\ \cdot\ \cdot \\ \alpha_k = \alpha_k \xi_m^{k-j_{n-1}} \\ \cdot\ \cdot\ \cdot \\ \alpha_{m-1} = \alpha_{m-1} \xi_m^{m-1-j_{n-1}} \end{cases}$$

Thus $\alpha_k = 0$ unless $k = j_{n-1}$. Hence

$$(7) \quad B(x) = \alpha_{j_{n-1}} [P_{n-1}(x)]^{(j_{n-1})/m}.$$

$\alpha_{j_{n-1}}$ may be expressed as follows:

$$\alpha_{j_{n-1}} = \beta_o + \beta_1 [P_{n-2}(x)]^{1/m} + \ldots + \beta_{m-1} [P_{n-2}(x)]^{(m-1)/m}$$

where β_i $(i = 0, \ldots, m - 1)$ is in Ψ_{n-3}. Consider the element σ_{n-2} of the Galois group of $\Psi_{n-2}: \Psi_{n-3}$ that maps $[P_{n-2}(x)]^{1/m}$ onto $\xi_m [P_{n-2}(x)]^{1/m}$. σ_{n-2} may be extended in the usual way to an element of the Galois group of $\Psi_{n-1}: \Psi_{n-3}$. Applying σ_{n-2} to (7) we obtain

$$\xi_m^{j_{n-2}} \cdot B(x) = (\beta_o + \beta_1 \xi_m [P_{n-2}(x)]^{1/m} + \ldots$$
$$+ \beta_{m-1} \xi_m^{m-1} [P_{n-2}(x)]^{(m-1)/m}) [P_{n-1}(x)]^{(j_{n-1})/m}.$$

where $j_{n-2} = k_{n-2}(m/q_i)$, $0 \leq k_{n-2} < q_i$. By the independence of the $[P_{n-2}(x)]^{1/m}$ $(j = 0,1 \ldots m-1)$ over Ψ_{n-3} we obtain

$$B(x) = \beta_{j_{n-2}}[P_{n-2}(x)]^{(j_{n-2})/m}[P_{n-1}(x)]^{(j_{n-1})/m}.$$

Continuing in this fashion we eventually obtain

$$(8) \quad B(x) = \gamma[P_1(x)]^{j_1/m}[P_2(x)]^{j_2/m} \ldots [P_{n-1}(x)]^{(j_{n-1})/m}$$

where $j_\ell = k_\ell(m/q_i)$, $0 \leq k_\ell < q_i$ and γ is in $\Omega(x)$. Hence

$$(9) \quad P_1^{k_1}(x) \ldots P_{n-1}^{k_{n-1}}(x)\gamma^{q_i} - P_n(x) = 0.$$

We want to show that this is a contradiction by showing that no element γ of $\Omega(x)$ can satisfy (9). Suppose there exist relatively prime $\beta_1(x)$ and $\beta_2(x)$ in $\Omega[x]$ such that

$$\beta_1^{q_i}(x)/\beta_2^{q_i}(x) = P_n(x)/(P_1^{k_1}(x) \ldots P_{n-1}^{k_{n-1}}(x)).$$

Then

$$P_n(x) = (P_1^{k_1}(x) \ldots P_{n-1}^{k_{n-1}}(x) \cdot \beta_1^{q_i}(x))/\beta_2^{q_i}(x)$$

which implies that $\beta_1(x)$ must have degree 0 for otherwise $P_n(x)$ would have multiple roots. Letting $\beta_3(x) = \beta_2^{q_i}(x)/\beta_1^{q_i}$ we have

$$\beta_3(x)P_n(x) = P_1^{k_1}(x) \ldots P_{n-1}^{k_{n-1}}(x)$$

which implies that $\beta_3(x)$ is in $\Gamma(x)$. But this implies that $P_1(x),\ldots,P_n(x)$ are not distinct irreducible elements of $\Gamma(x)$ which is contrary to the hypothesis. Thus, there is no $\beta(x)$ in $\Omega(x)$ satisfying (9) and hence $y^{q_i} - P_n(x)$ is irreducible over Ψ_{n-1}. Now consider $y^4 - p_n(x)$. We can prove exactly as above that there does not exist $B(x)$ in Ψ_{n-1} such that $B^2(x) = P_n(x)$. This also demonstrates that $y^2 - P_n(x)$ must be irreducible. Now suppose there exists $B(x)$ in Ψ_{n-1} such that $-4B^4(x) = P_n(x)$. Once again suppose $B(x)$ is represented by (5). This time applying σ_{n-1} we obtain

$$\xi_m^{j_{n-1}} \cdot B(x) = \alpha_0 + \alpha_1 \cdot \xi_m[P_{n-1}(x)]^{1/m} + \ldots$$
$$+ \alpha_{m-1}\xi_m^{m-1}[P_{n-1}(x)]^{(m-1)/m}$$

where $j_{n-1} = k_{n-1}(m/4)$, $0 \le k_{n-1} < 4$. In fact, we may carry out corresponding analysis to obtain (8) where $q_i = 4$. Then we obtain

$$(10) \quad 4P_1^{k_1}(x),\ldots,P_{n-1}^{k_{n-1}}(x)\gamma^4 + P_n(x) = 0$$

where γ is in $\Omega(x)$. But the same argument that shows that no element of $\Omega(x)$ can satisfy (9) also shows that no element of $\Omega(x)$ can satisfy (10). Thus we must conclude that such a $B(x)$ does not exist and that $y^4 - P_n(x)$ is irreducible over Ψ_{n-1}.

Q.E.D.

Corollary 10: Let $m_1, m_2, \ldots, m_n$ be positive integers, $P_1(x), P_2(x), \ldots, P_n(x)$ be distance, irreducible, primitive polynomials over $J[x]$ with degree $P_i(x) > 0$ $(i = 1,2,\ldots,n)$. Then $P(y) = y^{m_n} - P_n(x)$ is irreducible over $\Psi = \Omega(x)([P_1(x)]^{1/m_1}, \ldots, [P_{n-1}(x)]^{1/m_{n-1}})$.

Proof: Let $m = \text{lcm}(m_1, m_2, \ldots, m_n)$. Then $y^m - P_n(x)$ is irreducible over $\Psi' = \Omega(x)([p_1(x)]^{1/m}, \ldots, [P_{n-1}(x)]^{1/m})$ by the preceeding theorem. Thus $y^m - P_n(x)$ is irreducible over Ψ since $\Psi' \supset \Psi$. Let $m = 2^{\mu_o} q_1^{\mu_1}, \ldots, q_s^{\mu_s}$ be the prime decomposition of m. The irreducibility of $y^m - P_n(x)$ implies that

$$y^2 - P_n(x)$$
$$y^4 - P_n(x)$$

and

$$y^{q_i} - P_n(x), \; i = 1,2,\ldots,s,$$

are irreducible over Ψ. But since m is a multiple of m_n this implies that $y^{m_n} - P_n(x)$ is irreducible over Ψ.

Q.E.D.

Theorems 10 and 11 establish conditions under which irreducibility can be determined a priori. The exact role of these theorems in the normal form algorithm will be discussed in the next chapter.

Combinations of exponential and radical expressions

The proof of theorem 5 actually establishes a normal form for the radical expressions, i.e., every radical expression E can be mapped into an equivalent expression of the form

$$G(\xi_m, [P_1(x)]^{1/m_1}, \ldots, [P_n(x)]^{1/m_n})$$

where G is a polynomial over $\Gamma(x)$ and $P_1(x), \ldots, P_n(x)$ are the primitive polynomials appearing in E. This representation is unique within the particular extension field determined by E. But we may have $E_1 \equiv E_2$, and E_1 and E_2 determine different extension fields of $\Gamma(x)$ and hence have different representations. For example consider $E_1 = -(2)^{1/2}$ and $E_2 = (-1)^{1/2} * (-2)^{1/4} * (-2)^{1/4}$. E_1 determines the field $\Psi_1 = \Gamma(x)(2^{1/2})$ and E_2 determines $\Psi_2 = \Gamma(x)(\xi_8)$. Now $E_1 \equiv E_2$ but the representation of E_1 in Ψ_1 is $-(2)^{1/2}$ whereas the representation of E_2 in Ψ_2 is $\xi_8^3 - \xi_8$ where ξ_8 is taken to be the particular root $\exp(i * \pi/4)$. On the other hand, given any finite subset $\mathcal{E}$ of the radical expressions, an extension field Ψ may be determined such that for all E in $\mathcal{E}$, E in Ψ. Hence if $E_1 \equiv E_2$ then the representations of E_1 and E_2 in Ψ will be identical. Ψ is determined as follows. For each E in $\mathcal{E}$ determine the primitive polynomials appearing in E. Let $P_1(x), P_2(x), \ldots, P_n(x)$ be a listing of the distinct primitive polynomials appearing in the expressions of $\mathcal{E}$. Determine the radical degree of each $P_i(x)$ as follows. Let $p_1/q_1, p_2/q_2, \ldots, p_n/q_n$ be the radical powers to which $P_i(x)$ occurs in the expressions of $\mathcal{E}$. Then $m_i = \mathrm{lcm}(q_1, \ldots, q_n)$ is the radical degree of $P_i(x)$ in $\mathcal{E}$. Then each E in $\mathcal{E}$ will belong to the field $\Psi =$ $\Gamma(x)(\xi_m, [P_1(x)]^{1/m_1}, \ldots, [P_n(x)]^{1/m_n})$ where $m = \mathrm{lcm}(2m_o, m_1, \ldots, m_n)$, m_o = radical degree of (-1) in $\mathcal{E}$. Thus for each finite subset of the radical expressions we have a canonical form. With these observations a normal form f can be obtained for certain combinations of exponential and radical expressions. The normal form will have the property that $E \equiv 0$ if and only if $f(E) = 0$. The form will be analogous to the form of corollary 1 with the polynomials replaced by radical expressions.

Consider the class C generated by

(i) the rationals,

(ii) the radical expressions,

(iii) the operations of addition, substraction, multiplication, and restricted composition,

(iv) the exp function.

<u>Theorem 12</u>: Let $\mathcal{E}$ be a finite subset of the radical expressions and $R_1(x), R_2(x), \ldots, R_n(x)$ be distinct canonical

members of $\mathcal{E}$. The set $\{\exp(R_1(x)),\ldots,\exp(R_n(x))\}$ is linearly independent over $\mathcal{E}$.

Proof: Let $P_1(x),\ldots,P_k(x)$ be the primitive polynomials appearing in $\mathcal{E}$. Let a be the largest real zero of $P_1(x),\ldots,P_k(x)$. Then all the members of $\mathcal{E}$ are analytic for real $x > a$. Let I be a closed interval in this half line. Then consider

$$(1) \quad E_1(x) * \exp(R_1(x)) + \ldots + E_n(x) * \exp(R_n(x))$$

where the E_i are members of $\mathcal{E}$. Suppose (1) is functionally equivalent to 0. Then by Lindemann's theorem, for infinitely many rationals r in I either

(i) $E_i(r) \equiv 0$ for all $i = 1,2,\ldots,n$,

or

(ii) there exist $1 \leq i < j \leq n$ such that $R_i(r) \equiv R_j(r)$.

(ii) implies that $R_i(x) \equiv R_j(x)$ which is not possible since R_i and R_j are distinct canonical members of $\mathcal{E}$. Thus (i) must hold which implies that $E_1(x) \equiv \ldots \equiv E_n(x) \equiv 0$.

Q.E.D.

Corollary 5: There exists a normal form f for the class C that maps each expression into the form

(1) $E_1(x) * \exp(R_1(x)) + \ldots + E_n(x) * \exp(R_n(x))$

where the $E_i(x), R_i(x)$ belong to a canonical subclass $\mathcal{E}$ of the radical expressions. The $E_i(x)$ are non-zero and $R_i(x) < R_j(x)$ if $i < j$. Further $f(E) = 0$ if $E \equiv 0$.

<u>Proof</u>: Each member of $\mathcal{C}$ can be straight forwardly mapped into an expression of the form

(2) $E_1'(x) * \exp(R_1'(x)) + \ldots + E_k'(x) * \exp(R_k'(x))$.

Let $\mathcal{E} = \{E_1'(x), \ldots, E_k'(x), R_1'(x), \ldots, R_k'(x)\}$. Replace each $E_i'(x)$ and $R_i'(x)$ in (2) by its canonical form in $\mathcal{E}$. Then if necessary, rearrange (2) by combining equal exponentials deleting 0 coefficients, and ordering the $R_i(x)$'s. Then (2) will be transformed from the form (2) into the form (1) or 0.

Q.E.D.

This concludes chapter III.

Chapter IV

Implementations

In order to complete our study of canonical forms two of the algorithms of chapter III have been implemented in Formula Algol (FA). EXPCAN is an implementation of the canonical form for exponential expressions that was given in theorem 4 of chapter III. RADCAN is a routine that transforms radical expressions into normal form. The actual programs and some sample runs are given in appendix II.

Supporting Routines

A number of basic supporting routines are necessary for EXPCAN and RADCAN. The more important ones are mentioned here. Since one of the primary purposes of the canonical form routines is to reduce the test for functional equivalence to a test for string identity, all arithmetic calculations must be exact. Hence, a set of routines for performing arbitrary precision integer arithmetic is provided. In addition routines for transforming polynomials into canonical form and for carrying out polynomial arithmetic are provided.

RADCAN requires additional supporting routines of a more special nature. NEWTON is a procedure that computes interpolation polynomials by Newton's method of finite differences. Routines are also provided for carrying out arithmetic in $\Gamma(\xi_n)$ and for inverting matrices over $\Gamma(\xi_n)$.

EXPCAN

EXPCAN takes any exponential expression E and transforms it into canonical form. The organization of EXPCAN is particularly simple in that it is a generalization of a routine for transforming polynomials in several variables into canonical forms. EXPCAN works as follows. If E does not involve the exp function, then E is a polynomial and is transformed into canonical form. Otherwise E is written as a polynomial in the 'variable' exp, i.e., E is transformed into the form

$$(1) \quad P_o(x) * \exp(E_o(x)) + \ldots + P_n(x) * \exp(E_n(x))$$

where the $P_i(x)$'s are canonical polynomials not involving exp. Then EXPCAN is called recursively to canonicalize $E_o(x),\ldots,E_n(x)$. (1) is then rearranged so that the canonicalized $E_o(x),\ldots,E_n(x)$ are in ascending order, exponentials with identical arguments are combined, and terms with 0 coefficients are deleted.

RADCAN

RADCAN, the routine for transforming radical expressions into normal form, is much more interesting than EXPCAN. The main difficulty posed by RADCAN is that the classical constructive methods for factoring polynomials over Γ and extensions of Γ are strangled by combinatorial problems. By considering the special cases with which RADCAN is concerned, many of the problems can be eliminated. But even so the remaining combinatorial problems are of such a magnitude that RADCAN is feasible for only a few expressions.

An overview of the organization of RADCAN is given in figure 1. To facilitate the explanation, let us consider the particular problem of transforming the radical expression $\sqrt[4]{8} + 10\sqrt{6}\sqrt[3]{x^2+1}$ into normal form. In the initialization step 1, RADX $\leftarrow \sqrt[4]{8} + 10\sqrt{6}\sqrt[3]{x^2+1}$. In step 2 the polynomials 8, 6 and $x^2 + 1$ are put on the list of expressions that appear under a radical. In step 3 each of the expressions must be factored into irreducible factors over Γ and their radical degrees determined. For the constants this is just a matter of finding their prime decomposition. To factor the polynomials, Kronecker's method is used. Given a polynomial $P(x)$, Kronecker's algorithm successively looks for factors of degree $1,2,\ldots,[n/2]$ where n is the degree of $P(x)$. To search for i-th degree factors $P(x)$ is evaluated at $i + 1$ integers $a_0,a_1,\ldots,a_i$. For each $(i + 1)$-tuple $(b_0,b_1,\ldots,b_i)$ of integers such that $b_i | P(a_i)$, an interpolation polynomial $Q(x)$ is generated such that $Q(a_i) = b_i$. Then if $Q(x) | P(x)$, $Q(x)$ is a factor. The rationale here is that $P(x) = Q(x) * S(x)$ if and only if $P(a_i) = Q(a_i) * S(a_i)$.

ENTER

↓

I. RADX ← Radical expression under consideration.

↓

II. Place each distinct canonical subexpression of RADX that appears under a radical on unfactored polynomial list (UFP).

↓

III. Factor each expression on UFP into primitive polynomials and determine the radical degree of each primitive polynomial.

↓

IV. Determine m from the radical degrees of (-1) and the primitive polynomials. Replace roots of (-1) by m-th roots of unity. Replace radical subexpressions by products of radical roots of primitive polynomials.

↓

V. Order primitive polynomials and for $k = 1,2,\ldots,N$, determine the factors of $y^{m_k} - P_k(x)$ over $\Gamma(x)(\xi_m, [P_1(x)]^{1/m_1}, \ldots, [P_{k-1}(x)]^{1/m_{k-1}})$ where $P_k(x)$ is the k-th primitive polynomial with radical degree m_k.

↓

VI. Consider RADX as a polynomial in the $N + 2$ variables x, ξ_m and roots of the primitive polynomials. Transform this polynomial into canonical form and then reduce expressions in ξ_m and roots of the primitive polynomials modulo their corresponding defining equations.

↓

END

FIGURE 1: FLOW DIAGRAM FOR RADCAN

Kronecker's method is guaranteed to find all the factors of $P(x)$ but it generates a large number of $Q(x)$'s which are not divisors. The number of $Q(x)$'s generated is usually much larger that 4^i. Johnson [12] has given some methods for eliminating some of the $Q(x)$'s. We have used only a couple of Johnson's simplest tricks because of severe code space limitations in the current FA system. Hence our current factorization program, KRONECKER, cannot reasonably handle polynomials of degree > 5.

In our example, the primitive polynomials are 2, 3 and $x^2 + 1$ with radical degrees 4, 2 and 3 respectively. Seven $Q(x)$'s are generated by Kronecker to determine that $x^2 + 1$ is irreducible.

In step 4 we compute $m = \text{lcm}(2 * m_o, m_1, \ldots, m_n)$ where m_o = radical degree of -1 and m_k = radical degree of the primitive polynomial $P_k(x)$. For our example $m = 12$. Now we replace the original subexpressions occurring under radicals by the corresponding products of radical roots of primitive polynomials. Thus RADX $\leftarrow [2^{1/4}]^3 + [2^{1/4}]^2 * 3^{1/2} * (x^2 + 1)^{1/3}$.

In step V we must factor $y^4 - 2$, $y^2 - 3$ and $y^3 - (x^2 + 1)$ over various extensions of $\Gamma(x)(\xi_{12})$. Corollary 4 of chapter III tells us <u>a priori</u> that $y^3 - (x^2 + 1)$ will be irreducible. In general, we know that all polynomials $P(y)$ whose constant term in y is of degree > 0 in x will be irreducible. So we only have to consider $P(y)$'s with constant terms which are prime integers. However, the combinatorial problems involved in factoring over algebraic extensions of the rationals are considerably more staggering than those of factoring over the rationals only. For example, van der Waerden (section 42) gives a method for factoring over extensions of the rationals. When this method is applied to the simple case of factoring $y^2 - 2$ over $\Gamma(\xi_{12})$, it leads to a polynomial of degree 72 that must be factored over Γ! Kronecker's algorithm, even with all of Johnson's improvements, is not practical for polynomials of such a large degree. However, some further improvements are possible. From theorem 10, chapter III, it follows that if the radical degrees of p_k and $p_1,\ldots,p_{k-1}$ are odd then $y^{m_k} - p_k$ is irreducible over $\Psi = \Gamma(\xi_m, [p_1]^{1/m_1},\ldots,[p_{k-1}]^{1/m_{k-1}})$. Hence the factorization algorithm does not have to be used in these cases. But we are still not able to handle simple cases such as the one above.

However, with some additional assumptions about the reducibility of our particular pure equations and by using an algorithm developed by Johnson [11] for factoring polynomials over extensions of Γ, the problem can be made more tractable. First of all we assume a generalization of theorem 10. Let $m, m_1,\ldots,m_{k-1}$ be any positive integers, m_k an odd positive integer. Then we assume that $y^{m_k} - p_k$ is irreducible over $\Gamma(\xi_m, [p_1]^{1/m_1}, \ldots, [p_{k-1}]^{1/m_{k-1}})$. Note that lemma 4 establishes this result for $k = 1$. Thus we only consider constants whose radical degrees are even. In this event we assume that $y^{m_k} - p_k$ factors over Ψ only if it factors over $\Gamma(\xi_m)$ and furthermore that it factors in the same way over both fields. With these assumptions we only need consider the reducibility of equations of the form $y^{m_k} - p_k$ over $\Gamma(\xi_m)$ where m is even. From Tschebotaröw's theorems and the above assumptions, such polynomials are reducible if and only (i) if there exists b_1 in $\Gamma(\xi_m)$ such that $b_1^2 = p_k$ or (ii) if $4|m$ and if there exists b_2 in $\Gamma(\xi_m)$ such that $-4b_2^4 = p_k$.

In order to determine the existence of such b_1 and b_2 we use a special case of Johnson's algorithm for factoring over normal extensions of the rationals. There are two combinatorial difficulties with this algorithm. First of all since the algorithm is a generalization of Kronecker's method for factoring polynomials over Γ, the b_i's are generated from i-tuples of integers in a fashion similar to the way the $Q(x)$'s are generated in Kronecker's original algorithm. As in the original algorithm, a large number of spurious i-tuples are generated, but here no methods are known for reducing this number. Secondly, for each radical expression two matrices over $\Gamma(\xi_m)$ of dimension $2^n \times 2^n$ and $3^n \times 3^n$, $n = \varphi(m)$, must be inverted. Since the elements of the matrices belong to $\Gamma(\xi_m)$ the inversions must be carried out symbolically which is significantly slower than numerical inversion. Because of time considerations we probably cannot handle expressions where $\varphi(m) > 4$. In our actual programs we cannot handle expressions with $\varphi(m) > 3$ because of storage limitiations.

For our example matrices of dimension 16×16 and 81×81 must be inverted. Over $\Gamma(\xi_{12})$ $y^2 - 2$ is irreducible and $y^2 - 3 = [y - (-\xi_{12}^3 + 2\xi_{12})] * [y + (-\xi_{12}^3 + 2\xi_{12})]$.

In step VI RADX $\leftarrow (-\xi_{12}^3 + 2\xi_{12}) * [2^{1/4}]^2 * (x^2 + 1)^{1/3} + [2^{1/4}]^3$.

Thus we have considered two algorithms. EXPCAN is very simple and practical. On the other hand, RADCAN is so encumbered by combinatorial difficulties that it cannot be considered as a practical routine.

Chapter V

Conclusion

The purpose of this dissertation has been to study the representations of formula expressions in a way that would give meaning to the so-called simplification problem. As a step toward this goal we have defined the concepts of canonical and normal forms as alternatives to the controversial and ill-defined concept of simplified form.

Then in this sense we have shown, following Richardson, that canonical forms do not exist for some very simple classes of expressions. On the other hand, we have shown that rather large subclasses of these classes possess canonical forms. This implies a certain sharpness to both our undecidability and canonical form results. However, to obtain desirable canonical form results, strong number theoretic conjectures had to be assumed both by us and by Brown. This fact lends a certain importance to theorem 3 of chapter III. For theorem 3 obtains a canonical form for a subclass of $\mathfrak{R}_4$ without resorting to any conjectures.

Then the class of radical expressions has been studied and a normal form algorithm derived for this class. With the observation that this form is canonical for finite subclasses, a normal form is obtained for a class that allows limited kinds of combinations of exponential and radical expressions.

In chapter IV the implementation difficulties for the algorithms are discussed. EXPCAN, the canonical routine for exponential expressions, was seen to be simple and hence is a practical tool. RADCAN, on the other hand, is seen to be a completely impractical algorithm. This fact makes the results on radical expressions of theoretical interest only. RADCAN does raise some interesting questions about the reducibility of certain kinds of pure equations. Some of these questions have been answered by theorems 10 and 11 of chapter III.

On the other hand, we have not been able to find any results concerning the exponential constants that would allow us to circumvent the assumption of a n-th order form of Lindemann's theorem in obtaining theorem 4 of chapter III. This problem is perhaps the hardest one raised by this thesis. A proof of our conjecture would certainly be an excellent result and a difficult task. But there may be other ways of considering the problem that would not require such strong number theoretic results. Further we have said nothing about many interesting classes of expressions--particularly ones containing the log function. Such results would probably not be too difficult to obtain, although they do not seem to have an immediate correspondence to results concerning exponentials as one might expect. Also, we have not considered classes containing special operations of interest such as integration and limits. Richardson has obtained an undecidability result for integration but it would be interesting to study the positive side of the picture.

Further results about the reducibility of the pure equations that arise from the radical expressions would be desirable. It might be possible to completely characterize the ways in which the equations reduce and hence to do away with the need for factoring algorithms in RADCAN. This would then make RADCAN a practical algorithm. Such results are not only interesting for their applications to RADCAN but would be interesting number theoretic results in their own right.

This thesis has probably raised more questions than it has answered. But hopefully the questions raised are the correct ones, the answers provided are useful ones and the approach a fruitful one.

Bibliography

1. J. W. Backus, 'The Syntax and Semantics of the Proposed International Algebraic Language of the Zurich ACM-GAMM Conference', Proceedings of the International Conference on Information Processing, Paris, 1959.

2. P. J. Brown, 'Canonical Forms and Artificial Languages', U. of North Carolina Computation Center Mathematical Notes No. 19, July 5, 1963.

3. W. S. Brown, 'Rational Exponential Expressions and a Conjecture Concerning π and e', Bell Telephone Laboratories, Inc., Murray Hill, N.J.

4. Martin Davis, 'Diophantine Equations and Recursively Enumerable Sets', Automata Theory, E. R. Caioniello (ed.), Academic Press, 1966, 146-152.

5. ____________, 'Extensions and Corollaries of Recent Work on Hilbert's Tenth Problem', Illinois J. Math. 7 (1963), 246-250.

6. ____________ and Hilary Putnam, 'Reductions of Hilbert's Tenth Problem', J. Symbolic Logic 23 (1958), 183-187.

7. ____________, ____________, and Julia Robinson, 'The Decision Problem for Exponential Diophantine Equations', Ann. of Math. 74 (1961), 425-436.

8. R. R. Fenichel, 'An On-line System for Algebraic Manipulation', Ph.D. Thesis, Harvard University, 1966.

9. David Hilbert, 'Mathematische Problème', _Bull_. _Amer_. _Math_. _Soc_. _8_, (1901-2), 437-479.

10. R. Iturriaga, 'Contributions to Mechanical Mathematics', Ph.D. Thesis, Carnegie Institute of Technology, 1967.

11. S. C. Johnson, 'A Factoring Algorithm for Polynomials over an Arbitrary Galois Extension of the Rationals', Bell Telephone Laboratories, Inc., Murray Hill, N. J.

12. ______________, 'Tricks for Improving Kronecker's Polynomial Factoring Algorithm', Bell Telephone Laboratories, Inc., Murray Hill, N.J.

13. Ivan Niven, _Irrational Numbers_, The Carus Mathematical Monographs, No. 11, The Mathematical Association of America, 1956.

14. A. J. Perlis and R. Iturriaga, 'An Extension to Algol for Manipulating Formulae , _Comm_. _A_._C_._M_. _7_ (1964), 127-130.

15. ______________, ______________, and T. A. Standish, A Definition of Formula Algol', Computation Center, Carnegie Institute of Technology, March, 1966.

16. Daniel Richardson, 'Some Unsolvable Problems Involving Functions of a Real Variable', _Notices of the Amer_. _Math_. _Soc_. _13_ (1966), 135.

17. ____________________, Untitled and undated mimeographed document received in private communication April 4, 1966.

18. G. Rousseau, 'A Decidable Class of Number Theoretic Equations', J. London Math. Soc. 41 (1966), 737-741.

19. Jean E. Sammet, 'An Annotated Descriptor-Based Bibliography on the Use of Computers for Non-numerical Mathematics', Computing Reviews 7 (1966), B1-B31.

20. ________________, Modification No. 1 to 'An Annotated Descriptor-Based Bibliography on the Use of Computers for Doing Non-numerical Mathematics', Appendix to SICSAM Bulletin No. 5, December, 1966.

21. R. G. Tobey, R. J. Bobrow and S. N. Zilles, 'Automatic Simplification in FORMAC', Proceedings of Fall Joint Computer Conference, November 1965.

22. N. Tschebotaröw, Grundzuge der Galois'schen Theorie, P. Noordhoff N.V., 1950.

23. B. L. van der Waerden, Modern Algebra, Vol. I, Frederick Ungar Publishing Company, New York, 1953.

Appendix I

Backus-Naur Form Definitions for Classes of Expressions

<u>General Definitions</u>

<non-zero digit>::= 1|2|3|4|5|6|7|8|9

<digit>::= <non-zero digit> | 0

<non-zero integer>::= <non-zero digit> | <non-zero integer><digit>

<integer>::= <non-zero integer> | 0

<rational>::= <integer> | <non-zero integer>/<non-zero integer>

<single variable>::= x

<multiple variable>::= $x_1|x_2|\cdots|x_n$

<u>Definition of the class $\mathcal{R}$</u>

<$\mathcal{R}$ primary>::= <rational>|π|log 2|<single variable>|(<$\mathcal{R}$>)

<$\mathcal{R}$ term>::= <$\mathcal{R}$ primary> | <$\mathcal{R}$ term>*<$\mathcal{R}$ primary>

<simple $\mathcal{R}$>::= <$\mathcal{R}$ term> | <simple $\mathcal{R}$>+<$\mathcal{R}$ term>

<$\mathcal{R}$>::= <simple $\mathcal{R}$> | sin(<$\mathcal{R}$>) | exp(<$\mathcal{R}$>) | abs(<$\mathcal{R}$>)

Note: abs(<$\mathcal{R}$>) is also denoted |<$\mathcal{R}$>| where "|" is an absolute value bar.

<u>Definition of the class $\mathcal{R}_4$</u>

<argument primary>::= <rational>|π|<multiple variable>|(<argument>)

<argument term>::= <argument primary> | <argument term>*
<argument primary>

<argument>::= <argument term> | <argument>+<argument term>

<$\mathcal{R}_4$ primary>::= <argument primary> | (<$\mathcal{R}_4$>)

<$\mathcal{R}_4$ term>::= <$\mathcal{R}_4$ primary> | <$\mathcal{R}_4$ term>*<$\mathcal{R}_4$ primary>

<simple $\mathcal{R}_4$>::= <$\mathcal{R}_4$ term> | <simple $\mathcal{R}_4$>+<$\mathcal{R}_4$ term>

<$\mathcal{R}_4$>::= <simple $\mathcal{R}_4$> | sin(<argument>) | abs(<argument>)

<u>Definition of the FOE class</u>

<FOE a.p.>::= <rational> |i| <multiple variable> | (<FOE a.>)

<FOE a.t.>::= <FOE a.p.> | <FOE a.t.>*<FOE a.p.>

<FOE a.>::= <FOE a.t.> | <FOE a.>+<FOE a.t.>

<FOE primary>::= <FOE a.p.> | (<FOE>)

<FOE term>::= <FOE primary> | <FOE term>*<FOE primary>

<simple FOE>::= <FOE term> | <simple FOE>+<FOE term>

<FOE>::= <simple FOE> | exp(<FOE a.>)

<u>Definition of radical expressions</u>

<base primary>::= <rational> | <single variable> | (<base>)

<base term>::= <base primary> | <base term>*<base primary>|
<base term>/<base primary>

<base>::= <base term> | <base>+<base term> | <base>-<base term>

<radical primary>::= <base primary> | (<radical>)

<radical factor>::= <radical primary> | <base>↑<rational>

<radical term>::= <radical factor> | <radical term>*
<radical factor> | <radical term>/<radical factor>

<radical>::= <radical term> | <simple radical>+<radical term> |
<simple radical>-<radical term>

<u>Definition of the class C</u>

<C primary>::= <radical> | (<C>)

<C term>::= <C primary> | <C term>*<C primary> | <C term>/<C primary>

<simple C>::= <C term> | <simple C>+<C term> | <simple C>-<C term>

<C>::= <simple C> | exp(<radical>)

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY *(Corporate author)* Carnegie-Mellon University Department of Computer Science Pittsburgh, Pennsylvania 15213	2a. REPORT SECURITY CLASSIFICATION UNCL 2b. GROUP

3. REPORT TITLE

ON CANONICAL FORMS AND SIMPLIFICATION

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*

5. AUTHOR(S) *(First name, middle initial, last name)*

Bobby Forrester Caviness

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
May 20, 1968	85	23

8a. CONTRACT OR GRANT NO. SD-146 ARPA b. PROJECT NO. 9718 c. 6154501R d. 681304	9a. ORIGINATOR'S REPORT NUMBER(S) 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

--Distribution of this document is unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Air Force Office of Scientific Research 1400 Wilson Boulevard (SRI) Arlington, Virginia 22209

13. ABSTRACT

The motivation for this work comes from these two sources. First of all we wanted to study the problems of simplication. But in order to guide our work on simplification it seemed desirable to study further the unsolvability angle. Thus in Chapter II we study Richardson's theorem and proof in detail. From Richardson's proof and from studies on the unsolvability of Hilbert's tenth problem, we draw some conclusions about sharpenings of Richardson's theorem.

With the limitations of these negative results in mind, we study in Chapter III the structure of some classes of expressions and prove the existence of canonical forms for these classes. The concepts of canonical and normal forms as developed in Chapter III preserve most of the important concepts of simplification. On the other hand, these concepts are global concepts that can be formalized and hence are appropriate for a careful study whereas the concept of simplification lacks these properties.

Then in Chapter IV, we discuss the implementation of the algorithms developed in Chapter III. The algorithms are implemented using Formula Algol. The Formula Algol programs are included as Appendix II with some output from actual runs.

DD FORM 1 NOV 65 1473

Security Classification

Security Classification

14. KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT

Security Classification