On the Problem of Recognizing Zero

S. C. JOHNSON

Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey

ABSTRACT. The problem of recognizing when a complicated mathematical expression equals zero has great importance in symbolic mathematics. This paper gives two algorithms which can be applied to many such problems, and discusses a concrete example.

The algorithms are based on the recognition that many interesting functions (such as exponentiation) are eigenvectors of well studied transformations (such as differentiation).

KEY WORDS AND PHRASES: algebraic manipulation, zero recognition, canonical form

CR CATEGORIES: 5.39

1. *Introduction*

The problem of dealing with general mathematical expressions by computer has been intensively studied. Brown [1], Moses [2–4], Risch [5], and Richardson [6] have all described environments in which, with some success, one can do arithmetic operations and integration over certain classes of functions. On the other hand, Richardson [7] and Caviness [8] have shown that the problem of deciding whether a mathematical expression is identically zero is undecidable for certain other classes of functions.

The above authors begin by defining a class of expressions which represent mathematical functions, and describing arithmetic operations on these expressions. The development of algorithms using this approach is frequently complex, and the results are limited to the particular class of functions studied. The results are immediately applicable to computer algebra systems, however,

This paper describes two general mathematical algorithms for deciding when a sum of elements of a commutative ring R is zero, given appropriate conditions. They reduce the recognition of zero sums in R to the recognition of zero expressions in a subset of R. Because these algorithms are stated abstractly, their development is fairly simple. The actual application of these algorithms to computer algebra systems requires representing abstract rings and fields in a computer, in such a way that these algorithms can be effectively computed. We will thus assume that the rings and fields to be discussed in the following are well represented in concrete terms in the computer. In particular, the arithmetic operations and all functions considered will be considered effectively computable, and various other simple mathematical operations (such as cancellation and the distributive law) will be

This paper is a revised version of one with the same title that appears in "The Proceedings of the Second Symposium on Symbolic and Algebraic Manipulation, March 23–25, 1971," held by ACM-SIGSAM, which was prepared before the symposium.

assumed to be effectively applicable in the representation. A formal exposition of the required properties of the representation, as well as a comparison with the results of Brown [1] and Richardson [6], will await a later paper.

2. *The Algorithms*

This section is devoted to some mathematical background and the statements of the two algorithms which will be applied in the remaining sections. The reader may find it profitable to read this section in parallel with the next two sections.

Let K be a field whose structure is known completely for our purposes. In particular, we shall assume that we can recognize when an expression in K equals zero.

Let R be a commutative ring which contains K, and let φ be an effectively computable function $\varphi: R \to R$.

Definition: An element $u \in R$, $u \neq 0$, is an *eigenvector* of φ if there is an element $E(u) \in K$ with

$$\varphi(u) = E(u)u.$$

$E(u)$ is called the *eigenvalue* of u.

At this point, we need to make a practical convention; when we say that u is an eigenvector, we shall mean that we know not only u but also its eigenvalue. This is a crucial requirement, since the eigenvalue is in K, and we have assumed that the structure of K is well understood.

Throughout the following, we will deal with transformations φ which satisfy three axioms:

Axiom 1: $\varphi(u + v) = \varphi(u) + \varphi(v)$, for all $u, v \in R$.

Axiom 2: $\varphi(K) \subset K$, and $\varphi(1) = 0$.

Axiom 3: If u is an eigenvector, then u is invertible and u^{-1} is an eigenvector. If u and v are eigenvectors, then uv is. Moreover, we must be able to effectively compute the eigenvalues of u^{-1} and uv from their representations.

We now draw some simple consequences from the axioms, which can be quickly proved by the interested reader.

Axiom 1 implies that $\varphi(0) = 0$. Moreover, Axiom 2 implies that every nonzero element of K is an eigenvector. By Axiom 3, we thus deduce that φ applied to an eigenvector is either 0 (if and only if its eigenvalue is 0) or another eigenvector, whose eigenvalue can be effectively computed.

The set of all t in R with $\varphi(t) = 0$, the kernel of φ [denoted $\ker(\varphi)$], is of central importance in this work. Notice that every nonzero element of $\ker(\varphi)$ is an eigenvector with eigenvalue 0. Since 1 is in $\ker(\varphi)$, all the integers which are in K are also in $\ker(\varphi)$. In general, however, $\ker(\varphi)$ is not contained in K.

We now describe the first zero recognition algorithm:

ALGORITHM 1

Suppose we have some algorithm for deciding, given $t \in \ker(\varphi)$, if $t = 0$. Then, given a sum $S = \sum_{i=1}^{n} u_i$, with the u_i eigenvectors, $i = 1, \cdots, n$, we may decide if $S = 0$ as follows:

(1) If $n = 1$, then $J \neq 0$. Return.

(2) If $n > 1$, then compute

$$T = \varphi(u_n^{-1}S)$$

and test it for zero. (See details below).

(3) If $T \neq 0$ then $S \neq 0$. Return.

(4) If $T = 0$, then $u_n^{-1}S \in \ker(\varphi)$. By assumption, we can test if $u_n^{-1}S = 0$.

(5) If $u_n^{-1}S \neq 0$, then $S \neq 0$. Return.

(6) If $u_n^{-1}S = 0$, then $S = 0$. Return.

It remains to explain Step (2). We have:

$$\begin{aligned} u_n^{-1}S &= \sum_{i=1}^{n} u_n^{-1}u_i \\ &= 1 + \sum_{i=1}^{n-1} u_n^{-1}u_i . \end{aligned}$$

By Axiom 3, we may compute the eigenvalues $b_i = E(u_n^{-1}u_i)$, $i = 1, \cdots, n - 1$. Thus

$$T = \varphi(u_n^{-1}S) = \sum_{i=1}^{n-1} b_i(u_n^{-1}u_i).$$

If all the b_i are zero, then $T = 0$. Otherwise T is a sum of at most $n - 1$ eigenvectors with known nonzero eigenvalues, and we may apply Algorithm 1 recursively to decide whether $T = 0$.

We may in fact go even farther in this direction. Suppose, as above, that we can decide when elements of $\ker(\varphi)$ are zero. Let $A_1, \cdots, A_n \in R$ have the property that $\varphi(A_i)$ is an eigenvector, $i = 1, \cdots, n$. (Note that all eigenvectors have this property.) Then we may use the following algorithm to decide if $S = \sum_{i=1}^{n} A_i$ is zero.

ALGORITHM 2

Use Algorithm 1 to decide if $\varphi(S) = \sum_{i=1}^{n} \varphi(A_i)$ is zero. If $\varphi(S) \neq 0$ then $S \neq 0$. If $\varphi(S) = 0$, then S is in $\ker(\varphi)$, so we can decide if it is zero by assumption.

Clearly, this process may be extended to allow us to decide when a sum $\sum_{i=1}^{n} A_i$ is zero, provided only that each A_i can be transformed to a finite sum of eigenvectors by a finite number of applications of φ.

3. An Application

Let R be a field, and let φ be a *derivation* on R; that is, φ satisfies:

(1) addition law: $\varphi(a + b) = \varphi(a) + \varphi(b)$,

(2) multiplication law: $\varphi(ab) = a\varphi(b) + b\varphi(a)$.

The most common case of a derivation is when R is a field of infinitely differentiable functions, and φ is differentiation.

Let K be a subfield of R with $\varphi(K) \subset K$. Then we have the following result.

THEOREM. *The axioms of Section 2 hold for φ.*

PROOF. Axiom 1 is precisely the addition law. Axiom 2 follows because $\varphi(K) \subset K$ and

$$\varphi(1 \cdot 1) = 1 \cdot \varphi(1) + 1 \cdot \varphi(1)$$

or

$$\varphi(1) = \varphi(1) + \varphi(1)$$

so that

$$\varphi(1) = 0.$$

Let u and v be eigenvectors with eigenvalues $E(u)$ and $E(v)$. Then one can easily show that

$$\varphi(uv) = (E(u) + E(v))uv$$

and

$$\varphi(u^{-1}) = -E(u)u^{-1}.$$

Thus, uv and u^{-1} are eigenvectors, and we have

$$E(u^{-1}) = -E(u),$$

$$E(uv) = E(u) + E(v).$$

The eigenvectors of φ may include many interesting functions, for example
(1) rational functions,
(2) ae^b, with a and b rational functions,
(3) b^α, with b a rational function and α any rational number.
Algorithm 2 enables us to deal with functions A such that $\varphi(A)$ is an eigenvector; that is, indefinite integrals of eigenvectors. This set may include
(4) $\log b$, for b a rational function,
(5) $\arctan x$, $\arcsin x$,
(6) $\operatorname{erf}(x) = \int_0^x e^{-x^2} dx$.
The problem of recognizing zero expressions in $\ker(\varphi)$ is typically a hard theoretical problem, but provides little practical difficulty. Brown (1) and Moses (3) suggest practical approaches.

4. Another Application

Let K be a field of rational functions in one variable m. Let R be a field of functions of m such that R includes K. Define $\varphi: R \rightarrow R$ as

$$\varphi(u)(m) = u(m + 1) - u(m), \qquad u \in R;$$

that is, φ is the first difference function.

Definition: A function $u(m) \in R$ is *factential* if there is a rational function $a(m) \in K$ with

$$u(m + 1) = a(m)u(m).$$

For example, with K equal to the field of rational functions in m over the rational numbers, 2^m and $m!$ are factential, but $2^m + 1$ and $(m^2)!$ are not. The name "factential" describes the fact that both factorial and exponential functions are factential.

We leave to the reader the following simple proposition.

THEOREM. (1) *All rational functions are factential;*

(2) *factential functions are closed under the operations of taking inverses and multiplication;*

(3) *The nonzero factential functions are precisely the eigenvectors of φ; if $u(m + 1) = a(m)u(m)$, then $E(u) = a(m) - 1$;*

(4) *φ satisfies the three axioms.*

Thus Algorithm 1 can be used to allow us to tell when sums of factentials are identically zero, provided that we can tell when a function $u(m)$ in $\ker(\varphi)$ is identically 0.

Moreover, Algorithm 2 allows us to deal with functions

$$A(m) = \sum_{i=0}^{m-1} u(i),$$

where $u(m)$ is factential, since $\varphi(A) = u$.

5. *An Example*

We give a computational example of the use of these algorithms. Let K be the field of rational functions of x over the rational numbers, and let φ be differentiation with respect to x. Let R be a field of infinitely differentiable functions of x. Then, for eigenvectors u and v, and a rational number r, we have

(a) $E(r) = 0$,

(b) $E(uv) = E(u) + E(v)$,

(c) $E(u/v) = E(u) - E(v)$,

(d) $E(u^r) = rE(u)$.

Moreover, if $s(x)$ is a rational function,

(e) $E(s(x)) = s'(x)/s(x)$,

(f) $E(e^{s(x)}) = s'(x)$.

We shall investigate the expression

$$2xe^{3x}(3x + 1)^{\frac{1}{2}} - \int_0^x \frac{3te^{3t}}{(3t + 1)^{\frac{1}{2}}}\, dt - \int_0^x 2[e^t(3t + 1)^{\frac{1}{2}}]^3\, dt.$$

Set

$$A_1 = 2xe^{3x}(3x + 1)^{\frac{1}{2}},$$

$$A_2 = \int_0^x \frac{-3te^{3t}}{(3t + 1)^{\frac{1}{2}}}\, dt,$$

$$A_3 = \int_0^x - 2[e^t(3t + 1)^{\frac{1}{2}}]^3\, dt.$$

We wish to see if $A_1 + A_2 + A_3$ is zero. Applying algorithm 2, we examine $\varphi(A_1) + \varphi(A_2) + \varphi(A_3)$, a sum of eigenvectors, to see if it is zero. A_1 is already an eigenvector: we have

$$E(A_1) = E(2x) + E(e^{3x}) + \tfrac{1}{2}E(3x + 1)$$

$$= \frac{2}{2x} + 3 + \frac{3}{2(3x + 1)}$$

$$= \frac{18x^2 + 15x + 2}{2x(3x + 1)}.$$

Thus we write

$$U_1 = \varphi(A_1) = \frac{18x^2 + 15x + 2}{2x(3x + 1)} A_1,$$

$$U_2 = \varphi(A_2) = \frac{-3xe^{3x}}{(3x + 1)^{\frac{1}{2}}},$$

$$U_3 = \varphi(A_3) = -2[e^x(3x + 1)^{\frac{1}{2}}]^3.$$

U_1, U_2, and U_3 are eigenvectors; their eigenvalues are given by

$$\begin{aligned} E(U_1) &= E(18x^2 + 15x + 2) - E(2x) - E(3x + 1) + E(A_1) \\ &= \frac{9}{2}\left(\frac{36x^3 + 60x^2 + 27x + 4}{(3x + 1)(18x^2 + 15x + 2)}\right), \\ E(U_2) &= \frac{18x^2 + 9x + 2}{2x(3x + 1)}, \\ E(U_3) &= \frac{3}{2}\left(\frac{6x + 5}{3x + 1}\right). \end{aligned}$$

We now ask if $U_1 + U_2 + U_3$ is zero. Applying Algorithm 1, ask if $\varphi(U_1/U_3) + \varphi(U_2/U_3)$ is zero. We have

$$\begin{aligned} V_1 &= \varphi\left(\frac{U_1}{U_3}\right) = [E(U_1) - E(U_3)]\frac{U_1}{U_3} = \frac{3(1 - 3x)}{(18x^2 + 15x + 2)(3x + 1)} \cdot \frac{U_1}{U_3}, \\ V_2 &= \varphi\left(\frac{U_2}{U_3}\right) = \frac{1 - 3x}{x(3x + 1)} \cdot \frac{U_2}{U_3}. \end{aligned}$$

V_1 and V_2 are eigenvectors, and we have

$$\begin{aligned} E(V_1) &= E(3(1 - 3x)) - E(3x + 1) - E(18x^2 + 15x + 2) + E(U_1) - E(U_3) \\ &= -6(3x + 9)/(3x + 1)(3x - 1) \end{aligned}$$

and

$$E(V_2) = -6(3x + 9)/(3x + 1)(3x - 1).$$

Applying algorithm 1 to $V_1 + V_2$, we examine the eigenvector V_1/V_2. The eigenvalue of V_1/V_2 is

$$E(V_1/V_2) = E(V_1) - E(V_2) = 0.$$

Thus, $V_1/V_2 \in \ker(\varphi)$, so $1 + V_1/V_2$ is a constant. We must now appeal to a procedure which tells when elements of $\ker(\varphi)$ are zero. In this case, we could substitute $x = 1$; we have

$$\begin{aligned} A_1(1) &= 4e^3, \\ U_1(1) &= (\tfrac{35}{8} \cdot 4e^3) = \tfrac{35}{2}e^3, \\ U_2(1) &= -\tfrac{3}{2}e^3, \\ U_3(1) &= -2(e \cdot 2)^3 = -16e^3, \end{aligned}$$

and thus

$$V_1(1) = \frac{-6}{35\cdot 4}\frac{U_1(1)}{U_3(1)} = \frac{3}{64},$$

$$V_2(1) = \frac{-2}{4}\frac{U_2(1)}{U_3(1)} = \frac{-3}{64}.$$

Thus $1 + V_1(1)/V_2(1)$ is zero, so $V_1 + V_2$ is. Now, going backwards, we know that

$$1 + U_1/U_3 + U_2/U_3 \in \ker(\varphi).$$

We again ask if this is zero; reasoning as above we find

$$1 + \frac{U_1(1)}{U_3(1)} + \frac{U_2(1)}{U_3(1)} = 1 - \frac{35}{32} + \frac{3}{32} = 0.$$

Thus $U_1 + U_2 + U_3$ is zero, so $A_1 + A_2 + A_3 \in \ker(\varphi)$. Now our kernel examining algorithm could substitute $x = 0$, and recognize that $A_1(0) = A_2(0) = A_3(0) = 0$, so the original sum $A_1 + A_2 + A_3$ is zero.

Notice that, once we knew the eigenvalues, no simplification was necessary except in the rational function field. The only time that the actual representations of the eigenvectors were examined was when it was necessary to tell if an element of ker (φ) was in fact zero.

REFERENCES

1. Brown, W. S. Rational exponential expressions and a conjecture concerning π and e. *Amer. Math. Monthly 76* (1969), 28–34.
2. Moses, J. The integration of a class of special functions with the Risch algorithm. *SIGSAM Bull. 13* (Dec. 1969).
3. Moses, J. Algebraic simplification: A guide for the perplexed. *Proc. 2nd Symposium on Symbolic and Algebraic Manipulation*. ACM, New York, 1971, pp. 282–304; *CACM 14*, 8 (Aug. 1971), 527.
4. Moses, J. Symbolic integration: The stormy decade. *Proc. 2nd Symposium on Symbolic and Algebraic Manipulation*. ACM, New York, 1971. *CACM 14*, 8 (Aug. 1971), 548.
5. Risch, R. H. The problem of integration in finite terms. *Trans. Am. Math. Soc. 139* (1969), 167–189.
6. Richardson, D. Solution of the identity problem for integral exponential functions. *Zeitschr. Math. Logik 15* (1969), 333–340.
7. Richardson, D. Some unsolvable problems involving elementary functions of a real variable. *J. Symbolic Logic 33* (1968), 514–520.
8. Caviness, B. F. On canonical forms and simplification. *J. ACM 17* (1970), 385–396.