A DIFFERENTIAL-EQUATIONS APPROACH TO FUNCTIONAL EQUIVALENCE

by

John Shackell

University of Kent at Canterbury

1. Introduction

When computations involving transcendental functions are performed, for example in symbolic integration or calculation of limits, the problem arises as to how to recognise whether an expression is functionally equivalent to zero. This question is fundamentally connected with problems of algebraic simplification; see [3], [15], [8], [4]. As an example, consider the expression

$$e^{\sin^2 x(\log x+1)}e^{\sinh^2 x^2}-xe^{-\cos^2 x}e^{(\frac{1}{4}e^{2x^2}+1)}e^{-(\log x\cos^2 x-\frac{1}{4}e^{-2x^2})}.$$

This can be shown to be equivalent to zero by a calculation that is relatively straightforward but sufficiently involved to indicate that it is by no means trivial to decide such matters in general.

For some classes of expression, zero-equivalence is well known to be undecidable. If R_2 denotes the class generated by the constants 1 and π and the variable x using addition, multiplication and functional composition with sine and absolute value functions, then zero equivalence is recursively undecidable in R_2. This was shown by Caviness [5] using the methods of Richardson [18] and the undecidability of Hilbert's Tenth Problem, subsequently demonstrated by Matijacevic [14]. Richardson had earlier given a similar result for a somewhat larger class [18].

In order to decide zero-equivalence in a function class, one would in particular have to be able to do this for constant functions. This special case is related to some old, and clearly very difficult, questions in transcendental number theory; see [6], [7], [16], [12], [3]. It might well be undecidable for many function fields. In fact almost all the work that has been done on zero-equivalence in fields of transcendental functions assumes the existence of an oracle for deciding whether a constant is zero; this paper is no exception.

One of the main approaches to functional equivalence is to determine the algebraic dependencies between the various (apparently) transcendental functions appearing in the expression concerned. In this way, using the Risch Theorem, it is possible to handle fields built using arithmetic operations and functional composition with trigonometric, exponential and logarithmic functions. If instead of the Risch Theorem, one uses its generalisation due to Rothstein and Caviness, then one can include other functions defined by integrals, for example the error function; for further details see [19], [6], [16].

To seek algebraic dependencies between functions is to ask whether there exists a polynomial in them which is functionally equivalent to zero. For some purposes, for example in symbolic integration, this may indeed be what we want to know. However, to determine whether an expression is zero-equivalent we do not need to know whether *there exist* polynomials in the basic functions which are zero-equivalent but merely whether this is true of the expression given. The removal of the existential quantifier should and does make the problem easier (c.f. the distinction between deterministic and non-deterministic algorithms in the theory of NP completeness [9]). The methods outlined in this paper work directly with the given expression, which is regarded as a polynomial in a top-level basic function with coefficients in a function field containing the other basic functions. The top-level function is defined by a differential equation over the coefficient field. The techniques are entirely elementary and involve differentiation, substitution and calculation of gcds. Modulo the almost inevitable assumptions regarding constants, we are able to decide zero-equivalence in fields built using arithmetic operations and functional composition with functions defined as solutions of algebraic differential equations. In this paper we treat only first-order, first-degree equations. Higher orders and degrees will be considered in a later paper [21].

The author would like to thank colleagues at the University of Kent at Canterbury, and in particular Chris Woodcock, for a number of helpful comments while the research for this paper was in progress.

2. Field Extensions

For any field K, let $K[x]$ denote the ring of polynomials over K in one indeterminate and let $K(x)$ denote its quotient field, the field of rational functions. Let Q be a subfield of the field of complex numbers (for example Q might be the field of rational numbers), and let $F_0 = Q(x)$. Let $x_0 \in Q$ and for $i = 1,\ldots,n$, let F_i denote the simple extension of the field F_{i-1} by the solution of the differential equation

in which space?

$$X\left(\frac{a_{-2}}{X^2}+a_1+a_2X^2+\cdots\right)=1$$

no term $1/x$!

$\frac{dy}{dx} = \frac{1}{x}$ solve around $X = X_0 \neq 0$ Can we always write $y = \sum_{i=0}^{\infty} y_i (X - X_0)^i$? probably not... ? maybe with Laurent series?

$$\frac{dy}{dx} = Y_i(y) \quad ,$$

satisfying the initial condition $y(x_0) = y_i$; here Y_i is a rational function, with coefficients belonging to F_{i-1} and analytic at x_0, whose denominator does not vanish under the substitutions $y = y_i$, $x = x_0$. Under these conditions, the differential equation and initial condition together uniquely determine the solution $y(x)$ in a complex neighbourhood of x_0 and moreover it is analytic there; see [2] Ch.V for further details. Also each F_i is a differential field.

Let p be any element of F_n. We give a method for determining whether p is functionally equivalent to zero, assuming the existence of an oracle that determines whether a constant element of F_n is zero. We use induction on n, and take the ground case, $n = 0$, as given. We are also of course, taking x_0, the point of analyticity, as given. In practice the determination of such a point is an important matter and is likely to often be a troublesome one. In particular, deciding whether a given x_0 is a point of analyticity involves knowledge of when a constant is zero. Furthermore, taking a new x_0 may introduce new and complicated constant expressions.

Write $F = F_{n-1}$, $t_0 = y_n$ and write $Y_n(y)$ as a quotient of two polynomials, $Y_n = N/D$. Let $t(x)$ be the solution of the differential equation

$$\frac{dy}{dx} = \frac{N(y)}{D(y)} \quad , \qquad (1)$$

satisfying $t(x_0) = t_0$. After clearing denominators, we can reduce the question of whether $p \equiv 0$ to that of whether $P(t(x)) \equiv 0$, where $P(\lambda) \in F[\lambda]$ and is determined by p; here we are regarding λ as an indeterminate and writing $\equiv$ to denote functional equivalence.

3. The Star Derivation

For any $Q \in F[\lambda]$, we define

$$Q^* = \frac{\partial Q}{\partial x} D + \frac{\partial Q}{\partial \lambda} N \quad ; \qquad (2)$$

here, of course, $\partial Q/\partial x$ denotes the polynomial obtained by differentiating the coefficients of Q using the standard differentiation of the field F. It is easy to check that the operation * is a derivation on $F[\lambda]$.

Theorem 1

Let F_{alg} denote the algebraic closure of F. Let P be an element of $F[\lambda]$ which is square free and let G denote the gcd of P and P^*. If G contains a factor $\lambda - h$, where $h \in F_{alg}$, then $y = h$ satisfies (1). Conversely if $y = h$ satisfies (1) and $P(h) \equiv 0$, then $\lambda - h$ divides G .

Proof of Theorem 1

Let $\lambda - h$ be a factor of G, with $h \in F_{alg}$. Then $P(\lambda)$ may be written in the form

$$P(\lambda) = (\lambda - h) S \quad , \qquad (3)$$

with $S \in F_{alg}[\lambda]$. A priori h need not be differentiable at x_0. However it will be well-defined and analytic in a cut neighbourhood of x_0 and this suffices. The conclusion of the theorem then forces h to be analytic at x_0 also. From (3), we have

$$P^*(\lambda) = (\lambda - h) S^* + (N - Dh') S \quad . \qquad (4)$$

Since P is square free, $S(h) \not\equiv 0$. However, $P^*(h) \equiv 0$ and hence $N(h) - D(h)h' \equiv 0$; in other words h satisfies the differential equation (1).

Conversely if $P(h) \equiv 0$, then P may be written in the form (3). If h also satisfies (1) then (4) implies that $P^*(h) \equiv 0$. Hence $\lambda - h$ divides P^* and so it also divides G.

This completes the proof of Theorem 1.

4. The Algorithm

To determine whether $P(t(x)) \equiv 0$, where t satisfies (1) and the initial condition $t(x_0) = t_0$, we proceed as follows:

(i) Replace $t(x)$ by an indeterminate λ. Check that $P(\lambda)$ is not the zero polynomial (by induction we are able to do this). Then compute gcd $(P, \partial P/\partial \lambda)$ in the ring $F[\lambda]$ and divide P by the result in order to make P square free.

(ii) Compute the primitive part of the gcd of P and P^* in $F[\lambda]$. Multiply up by any quotients introduced in taking the gcd to give a polynomial $g(\lambda)$. Note that the gcd computation will require application of the algorithm recursively to each remainder in the sequence.

(iii) If g is independent of λ or if g fails to vanish under the substitutions $x = x_0$, $\lambda = t_0$ then $P(t(x)) \not\equiv 0$.

(iv) Otherwise substitute $x = x_0$ in g leaving λ as an indeterminate. If this reduces g to the zero polynomial, keep differentiating g partially with respect to x and substituting $x = x_0$ until for some k, $\partial^k g/\partial x^k$ does not reduce to the zero polynomial when $x = x_0$ (such a k exists because g is analytic at x_0 and not the zero polynomial). Now check whether g vanishes under the substitutions $x = x_0$, $\lambda = t_0$. If so then $P(t(x))) \equiv 0$, otherwise $P(t(x)) \not\equiv 0$.

We are of course assuming, by induction, that we are able to make any decisions as to whether elements of F are zero that are necessary for the gcd calculation. Also our assumptions concerning constants allow us to determine whether g and its derivatives vanish under substitutions. Given these, it is not difficult to see that the algorithim does what is claimed. For if g is independent of λ then, by Theorem 1, P cannot contain a factor $\lambda - h$ where h satisfies (1). On the other hand if g does depend on λ, it may be factored over F_{alg} in the form

$$g(\lambda) = A(x) \prod_{i=1}^{r} (\lambda - h_i) \quad . \qquad (5)$$

Let k be the least integer such that $A^{(k)}(x_0) \neq 0$. Then the first $(k-1)$ partial derivaties of g with respect to x all reduce to zero under the substitution $x = x_0$. If we write $f(\lambda) = \prod_{i=1}^{r} (\lambda - h_i)$, then by Leibniz Theorem,

$$\frac{\partial^k g}{\partial x^k} = A^k f + kA^{(k-1)}\frac{\partial f}{\partial x} + \ldots + A\frac{\partial^k f}{\partial x^k}. \qquad (6)$$

When we substitute $x = x_0$, all terms on the right of (6) except the first vanish. The first term does not vanish and so this k must be the value in (iv). Now $f(\lambda)$ vanishes under the substitutions $x = x_0, \lambda = t_0$ precisely when, for some i, $h_i(x_0) = t_0$, i.e. precisely when P contains a factor $\lambda - t$. This suffices to justify our algorithm.

An alternative to (iv) is to divide g by its leading coefficient to make it monic. The substitutions $x = x_0, \lambda = t_0$ may then lead to indeterminate forms and we need to evaluate the limit as $x \to x_0$. The procedure given in (iv) is equivalent to evaluating these by L'Hôpital's rule. However this is not necessarily the best technique; see [20], [10], for example.

5. A Worked Example

Consider the expression

$$F(x) = (e^{\log x})^2 + e^{\log x}(\log x - 2x) - x\log x + x^2.$$

We write μ for $\log x$, so μ is defined by $\mu' = 1/x$, $\mu(1) = 0$. Then we replace e^{μ} by λ, so $\lambda' = \lambda\mu' = \lambda/x$ and $\lambda(1) = 1$. $F(x)$ is then represented by

$$P(\lambda) = \lambda^2 + \lambda(\mu - 2x) - \mu x + x^2.$$

Calculation of $\gcd(P, \partial P/\partial\lambda)$ shows that P is square free. Then

$$P^* = 2\lambda^2 + \lambda(\mu - 4x + 1) - \mu x + 2x^2 - x.$$

Division of P by P^* gives remainder $\lambda(-\mu+1) + \mu x - x$. The coefficients of this polynomial have a common factor $-\mu+1$. A simple application of the algorithm shows that $-\mu+1$ is not equivalent to zero, and so we may divide the coefficients by it, giving $\lambda - x$ for the primitive part of the remainder. This turns out to divide P and so the primitive part of $\gcd(P, P^*)$ is equal to $\lambda - x$.

Our next step is to substitute $x=1$, $\lambda=1$ into $\lambda - x$ and observe that it vanishes. However it does not vanish if we only substitute for x and so $F(x) \equiv 0$.

6. Conclusion

The remaining problem is the one concerning constants. If one restricts attention to extensions via trigonometric, exponential and logarithmic functions then zero-equivalence of constants can be decided if one is prepared to assume the Schanuel conjecture. This may be stated as follows:

Let Q denote the field of rational numbers and let $Q(a,b,c,\ldots)$ denote its extension by $a,b,c,\ldots$ Suppose that $\alpha_1, \ldots, \alpha_n$ are complex numbers linearly independent over Q. Then the transcendence degree of the field $Q(\alpha_1, \ldots, \alpha_n, e^{\alpha_1}, \ldots e^{\alpha_n})$ is at least n .

It should be pointed out that proof of this lies well beyond the current state of development of transcendental number theory, but it is at least a plausible and long-standing conjecture; see [12], [1], [7], [3]. However, one can hardly make a similar conjecture concerning extensions via arbitrary solutions of first-order differential equations (let alone higher orders) and so the question is very much open as to what, in practice, one should do about constants at this level of generality.

Methods of approximating solutions of differential equations are well established (see [2] Ch. VII, for example) and this opens the possibility of using interval analysis to attack the constants problem; in this connection see [8]. Of course interval analysis can only ever give a semi-decision procedure. In other words it may establish definitively that a constant is not zero but if successive approximation intervals continue to straddle zero, one can only assert that the constant is possibly zero (or maybe probably zero where some meaning is assigned to word "probably"). Nonetheless, it may well be that interval analysis is the best tool available.

For further discussion concerning constants see [15], [8], [13], [17], [11], [3].

7. References

[1] A. Baker, *'Transcendental Number Theory'*, Cambridge University Press (1975).

[2] G. Birkoff and G.C. Rota, *'Ordinary Differential Equations'*, Ginn & Co. (1962).

[3] W.S. Brown, *'Rational Exponential Expressions and a Conjecture Concerning π and e'*, Amer. Math. Monthly **76** (1969), 28-34.

[4] B. Buchberger and R. Loos, *'Algebraic Simplification'*, in 'Computer Algebra: Symbolic and Algebraic Computation', B. Buchberger, G.E. Collins & R. Loos (eds.), 2nd edition, Springer-Verlag, Wien/New York (1983), 11-43.

[5] B.F. Caviness, *'On Canonical Forms and Simplification'*, J.A.C.M. **17/2** (1970), 385-396.

[6] B.F. Caviness, *'Methods for Symbolic Computation with Transcendantal Functions'*, in 'Proc. Conf. on Symbolic Computational Methods and Applications, St. Maximin (France)', A. Visconti (ed.) (1977), 16-43.

[7] B.F. Caviness and M.J. Prelle, *'A Note on Algebraic Independence of Logarithmic and Exponential Constants'*, SIGSAM Bull. **12/2** (1978), 18-20.

[8] J.P. Fitch, *'On Algebraic Simplification'*, Comput. J. **17/1** (1973), 23-27.

[9] M.R. Garey and D.S. Johnson, *'Computers and Intractability: A Guide to the Theory of NP-Completeness'*, Freeman & Co., New York (1979).

[10] K.O. Geddes and G.H. Gonnet, *'A New Algorithm for Computing Symbolic Limits using Generalised Hierarchial Series'*, Proceedings ISSAC 88.

[11] S.C. Johnson, *'On the Problem of Recognizing Zero'*, J.A.C.M. **18/4** (1971), 559-565.

[12] S. Lang, *'Transcendental Numbers and Diophantine Approximation'*, Bull. Amer. Math. Soc. 77/5 (1971), 635-677.

[13] W.A. Martin, *'Determining the Equivalence of Algebraic Expressions by Hash Coding'*, J.A.C.M. **18/4** (1971), 549-558.

[14] Yu.V. Matijacevic, *'Enumerable Sets are Diophantine'*, Sov. Math. Dokl. **11** (1970), 453-458.

[15] J. Moses, *'Algebraic Simplification: A Guide for the Perplexed'*, C.A.C.M. **14/8** (1971), 527-537.

[16] A.C. Norman, *'Computing in Transcendental Extensions'*, in 'Computer Algebra: Symbolic and Algebraic Computation', B. Buchberger, G.E. Collins & R. Loos (eds.), 2nd edition, Springer-Verlag, Wien/New York (1983), 169-172.

[17] A. Oldehoeft, *'Analysis of Constructed Mathematical Responses by Numeric Tests for Equivalence'*, Proc. ACM 24th Nat. Conf. (1969), 117-124.

[18] D. Richardson, *'Some Undecidable Problems Involving Elementary Functions of a Real Variable'*, J. Symbolic Logic **33** (1968), 514-520.

[19] M. Rothstein and B.F. Caviness, *'A Structure Theorem for Exponential and Primitive Functions'*, SIAM J. Comput. **8/3** (1979), 357-367.

[20] J.R. Shackell, *'Growth Estimates for Exp-Log Functions'*, preprint 1987.

[21] J.R. Shackell, *'Zero-equivalence in function fields defined by algebraic differential equations'*, preprint 1989.