Mixed Transcendental and Algebraic Extensions for the Risch-Norman Algorithm

Stefan T. Boettner

July 20, 2010

Elementary Functions

Definition

Let F be a field of functions. Then f is elementary over F if one of the following holds:

- $f = e^g$ for some $g \in F$
- $f = \log g$ for some $g \in F$
- f is algebraic over F

Elementary Extensions

Definition

Let F be a field of functions and K an extension of F. Then K is called an elementary extension if there exist $t_1, \ldots, t_n \in K$ such that $K = F(t_1, \ldots, t_n)$ and t_k is elementary over $F(t_1, \ldots, t_{k-1})$ for each $k \in \{1, \ldots, n\}$.

Definition

An elementary function is a function of any elementary extension of $\mathbb{C}(x)$.

Remark

Trigonometric functions and their inverses are elementary functions because they can be expressed using complex exponentials or logarithms.

Integration in Finite Terms

Definition

An (indefinite) integral is elementary if it is contained in some elementary extension of the function field that contains the integrand.

Problem

Determine in finitely many steps whether a given (indefinite) integral is elementary, and if so compute it.

The Risch Algorithm

The problem of integration in finite terms is solved by the Risch algorithm. However, due to its complexity, no complete implementation is known to exist.

The Parallel Risch or Risch-Norman algorithm is a simpler heuristic method for computing integrals. While it cannot prove that an integral is non-elementary, it has the advantage that it can be generalized to a much larger class of functions.

Differential Fields

Definition

Let R be a ring (field). A map $D : R \to R$ such that

- $D(a+b) = Da + Db$ and
- $D(ab) = aDb + bDa$

is called a derivation on R. The pair (R, D) is called a differential ring (field).

If the derivation D is clear from the context, we will just refer to R as a differential ring/field. The set

$$C(R) = \{c \in R | Dc = 0\}$$

is called the ring (field) of constants of R.

Differential Field Examples

Example

Let R be any ring. Define D by $Dx = 0$ for all $x \in R$. Then (R, D) is a differential ring, and every element of R is a constant.

Example

Let $F = \mathbb{C}(x)$ be the field of rational functions. The map D defined by the usual derivative

$$Df = \frac{df}{dx}$$

is a derivation.

Differentiation Rules

All the usual differentiation rules follow from the definition of differential rings/fields:

Theorem

Let (R, D) be a differential ring (field). Then

1. $D(ca) = cDa$ *for* $a \in R$ *and* $c \in C(R)$.
2. *If* R *is a field, then*

$$D\frac{a}{b} = \frac{bDa - aDb}{b^2}$$

   for any $a, b \in R$, $b \neq 0$.
3. $D(a^n) = na^{n-1}Da$ *for any* $a \in R$, $a \neq 0$ *and integer* $n > 0$ *(any integer* n *if* R *is a field).*

Differential Extension

Definition

Let (F, D) and (K, Δ) be differential fields such that K is a field extension of F. If $\Delta f = Df$ for all $f \in F$, then (K, Δ) is a differential extension of (F, D).

Example

Let $F = \mathbb{C}$ with $Dx = 0$ for all $x \in F$. Let $K = \mathbb{C}(x)$ be the field of rational functions with $\Delta = d/dx$ the usual derivative. Then (K, Δ) is a differential extension of (F, D).

Transcendental Differential Extension

If t is transcendental over F, then we can freely choose Dt in $F(t)$:

Theorem

Let (F, D) be a differential field, and let t be transcendental over F. Then for any $w \in F(t)$ there exists a unique derivation Δ on $F(t)$ such that $\Delta t = w$ and $(F(t), \Delta)$ is a differential extension of (F, D).

Transcendental Differential Extension

Example

Let (F, D) be a differential field, $f \in F$ and t transcendental over F. Extend D to $\Delta\colon F(t) \to F(t)$ by defining $\Delta t = (Df)t$. Then t models e^f.

Example

Let (F, D) be a differential field, $f \in F$ and t transcendental over F. Extend D to $\Delta\colon F(t) \to F(t)$ by defining $\Delta t = (Df)/f$. Then t models $\log f$.

Tower of Extensions

The Risch and basic Risch-Norman algorithms work with a tower of extensions of the rational function field:

$$K \subset K(x) \subset K(x, \theta_1) \subset \ldots \subset K(x, \theta_1, \ldots, \theta_n)$$

Here K is a computable field of constants (usually $\mathbb{Q}$) and θ_k is

- an exponential over $K(x, \theta_1, \ldots, \theta_{k-1})$,
- a logarithm over $K(x, \theta_1, \ldots, \theta_{k-1})$ or
- algebraic over $K(x, \theta_1, \ldots, \theta_{k-1})$ (full Risch algorithm only).

The recursive nature of the Risch algorithm relies on this tower structure, while the Risch-Norman method handles all extensions in parallel. It uses $K(x, \theta_1, \ldots, \theta_n)$ directly without intermediate field extensions.

Liouville's Theorem

Liouville's theorem gives us a strong hint how a closed form expression of an integral looks if it exists:

Theorem

Let (F, D) be a differential field, $C = C(F)$, and $f \in F$. If there exist an elementary extension K of F and $g \in K$ such that $Dg = f$, then there are $v_0 \in K$, $\lambda_1, \ldots, \lambda_n \in \bar{C}$ and $v_1, \ldots, v_n \in K(\lambda_1, \ldots, \lambda_n)$ such that

$$f = Dv_0 + \sum_{i=1}^{n} \lambda_i \frac{Dv_i}{v_i}.$$

Liouville's Theorem

According to Liouville's theorem, we may expect an integral, if it exists in closed form, to have the form

$$\int f\,dx = v_0 + \sum_{i=1}^{n} \lambda_i \log v_i.$$

with $v_0, \ldots, v_n$ functions from the same field as f, possibly extended by some algebraic constants, and $\lambda_1, \ldots, \lambda_n$ constants.

It consists of a *rational part* (v_0) and a *logarithmic part* (the sum of the $\lambda_i \log v_i$).

Basic Idea of the Risch-Norman Method

If the denominator of v_0 and the polynomials $v_1, \ldots, v_n$ in

$$f = Dv_0 + \sum_{i=1}^{n} \lambda_i \frac{Dv_i}{v_i},$$

can be predicted, then it is possible to determine the numerator of v_0, which is written as the sum of "all possible" monomials with unknown coefficients, and $\lambda_1, \ldots, \lambda_n$ by writing the equation over the common denominator and solving a system of linear equations.

Prediction of Denominator and Logarithms

Let q be the denominator of the integrand and

$$q = \prod_{j=1}^{m} q_j^{\nu_j}$$

its factorization into irreducibles. Define

$$\nu_j^* = \begin{cases} \nu_j & \text{if } q_j \text{ is an exponential} \\ \nu_j - 1 & \text{otherwise} \end{cases}$$

Then

$$q_0 = \prod_{j=1}^{m} q_j^{\nu^*}$$

is the prediction for the denominator of the integral. Moreover, every q_j may give rise to a term $\log q_j$.

A Complete Example

Consider the integral

$$\int \frac{x}{(a+e^{x^2})^2}\,dx$$

where a is a parameter.

We have to construct the field $K(x,\theta)$ where

- $K = \mathbb{Q}(a)$ is the field of constants,
- $Dx = 1$ and
- $D\theta = 2x\theta$.

A Complete Example

The integral is now written as

$$\int \frac{x}{(a+\theta)^2} dx.$$

Unless $a = 0$, the factor $a + \theta$ is not an exponential, so we make the ansatz

$$\int \frac{x}{(a+\theta)^2} dx = \frac{u(x,\theta)}{a+\theta} + \lambda \log(a+\theta)$$

where $u(x, \theta)$ is a polynomial in x and θ.

A Complete Example

Now let

$$u(x, \theta) = \sum_{i=0}^{n} \sum_{j=0}^{m} \beta_{ij} x^i \theta^j$$

with the coefficients β_{ij} to be determined. The *degree bounds* n and m are customarily taken one higher than the highest occuring power in the integrand, so $n = 2$ and $m = 1$ in this case, but no rigorous bound is known.

Alternatively, one may bound the total degree:

$$u(x, \theta) = \sum_{i+j \leq n} \beta_{ij} x^i \theta^j$$

A Complete Example

We now have to determine the coefficients β_{ij} and λ such that

$$\int \frac{x}{(a+\theta)^2} dx = \frac{u(x,\theta)}{a+\theta} + \lambda \log(a+\theta).$$

Apply D on both sides to obtain

$$\frac{x}{(a+\theta)^2} = \frac{(a+\theta)Du(x,\theta) - 2x\theta u(x,\theta)}{(a+\theta)^2} + \lambda \frac{2x\theta}{a+\theta}.$$

Multiplying this with the common denominator $(a+\theta)^2$ yields

$$x = (a+\theta)Du(x,\theta) - 2x\theta u(x,\theta) + 2\lambda x\theta(a+\theta).$$

A Complete Example

Substitute the ansatz with the undetermined coefficients for u:

$$\begin{aligned} x &= (2a\beta_{21} - 2\beta_{20})x^3\theta + (2a\beta_{11} - 2\beta_{10})x^2\theta + (2\beta_{21} + 2\lambda)x\theta^2 + \\ &\quad (2a\beta_{21} + 2a\beta_{01} + 2\beta_{20} - 2\beta_{00} + 2a\lambda)x\theta + 2a\beta_{20}x + \\ &\quad \beta_{11}\theta^2 + (a\beta_{11} + \beta_{10})\theta + a\beta_{10}, \end{aligned}$$

Now by comparing the coefficients of the monomials $x^s\theta^t$ we obtain a system of linear equations in the unknowns β_{ij} and λ.

A Complete Example

In matrix notation, we get this system of equations:

$$\begin{pmatrix} 2 & 0 & 0 & 0 & 0 & 0 & 2 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 2a & 0 & 0 & -2 & 0 & 0 & 0 \\ 0 & 2a & 0 & 0 & -2 & 0 & 0 \\ 2a & 0 & 2a & 2 & 0 & -2 & 2a \\ 0 & a & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2a & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & a & 0 & 0 \end{pmatrix} \begin{pmatrix} \beta_{21} \\ \beta_{11} \\ \beta_{01} \\ \beta_{20} \\ \beta_{10} \\ \beta_{00} \\ \lambda \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}.$$

A Complete Example

It has the solution

$$\beta_{21} = \frac{1}{2a^2}, \quad \beta_{11} = 0, \quad \beta_{01} = -\frac{1}{2a^2},$$
$$\beta_{20} = \frac{1}{2a}, \quad \beta_{10} = 0, \quad \beta_{00} = 0,$$
$$\lambda = -\frac{1}{2a^2}$$

corresponding to

$$\int \frac{x}{(a+e^{x^2})^2}\,dx = \frac{\frac{1}{2a^2}(x^2-1)e^{x^2} + \frac{1}{2a}x^2}{a+e^{x^2}} - \frac{1}{2a^2}\log(a+e^{x^2}).$$

Simultaneous Extensions

Instead with a tower of extensions, Risch-Norman works with a single simultaneous extension $K(x, \theta_1, \ldots, \theta_n)$.

Thus there is no reason to require that

$$D\theta_k \in K(x, \theta_1, \ldots, \theta_k),$$

but one may allow

$$D\theta_k \in K(x, \theta_1, \ldots, \theta_n).$$

Functions Satisfying a System of ODEs

We may therefore have

$$\begin{aligned} D\theta_1 &= f_1(x, \theta_1, \ldots, \theta_n) \\ &\vdots \\ D\theta_n &= f_n(x, \theta_1, \ldots, \theta_n) \end{aligned}$$

where $f_1, \ldots, f_n$ are rational functions. This can be used to model functions that satisfy a system of (possibly nonlinear) ODEs, or an ODE of higher order.

Many special functions satisfy such a system of ODEs.

Example: Airy Functions

The Airy functions are the solutions of the differential equation

$$y'' - xy = 0.$$

Therefore Airy functions may be modeled by the extension $K(x, \theta_1, \theta_2)$ if we define

$$\begin{aligned} Dx &= 1 \\ D\theta_1 &= \theta_2 \\ D\theta_2 &= x\theta_1 \end{aligned}$$

Trigonometric Functions

The basic trigonometric functions $\sin x$ and $\cos x$ satisfy the system of equations

$$\begin{aligned} \sin' x &= \cos x \\ \cos' x &= -\sin x, \end{aligned}$$

so it should be possible to model them using the extension $\mathbb{Q}(x, \theta_1, \theta_2)$ with

$$\begin{aligned} Dx &= 1 \\ D\theta_1 &= \theta_2 \\ D\theta_2 &= -\theta_1. \end{aligned}$$

Trigonometric Functions

Indeed, the method will compute simple integrals such as

$$\int \sin x \, dx = -\cos x.$$

But it will fail to compute

$$\int \sin^2 x \, dx = -\frac{1}{2} \sin x \cos x + \frac{x}{2}.$$

So what's wrong?

Trigonometric Functions

To analyze how the algorithm fails, take the derivative of

$$\int \sin^2 x \, dx = -\frac{1}{2} \sin x \cos x + \frac{x}{2}$$

and write the equation using expressions in the field $\mathbb{Q}(x, \theta_1, \theta_2)$:

$$\theta_1^2 = \frac{1}{2}\theta_1^2 - \frac{1}{2}\theta_2^2 + \frac{1}{2}$$

Evidently, the coefficients of the monomials are all different between the two sides of the equation! Why is this?

Trigonometric Functions and Algebraic Relations

The sine and cosine functions satisfy the relation

$$\sin^2 x + \cos^2 x = 1$$

which was not taken into account. In $\mathbb{Q}(x, \theta_1, \theta_2)$, $\theta_1^2 + \theta_2^2$ is not considered the same as 1!

This makes sense since we could as well have

$$\begin{aligned} \theta_1 &= \lambda \sin x \\ \theta_2 &= \lambda \cos x \end{aligned}$$

for any constant λ the way we defined θ_1 and θ_2.

Differential Ideals and Quotients

Definition

Let (R, D) be a differential ring and $I \subset R$ an ideal. Then I is called a differential ideal if $DI \subset I$.

Lemma

Let (R, D) be a differential ring and $I \subset R$ a differential ideal. Let $\pi : R \rightarrow R/I$ be the canonical projection. Then D induces a derivation D^* on R/I such that $D^* \circ \pi = \pi \circ D$, where $\pi : R \rightarrow R/I$ is the canonical projection.

Trigonometric Functions and Algebraic Relations

Instead of a field, construct the ring $R = \mathbb{Q}[x, \theta_1, \theta_2]$ with the derivation D defined as before. Then the ideal

$$I = \langle \theta_1^2 + \theta_2^2 - 1 \rangle \subset R$$

is a differential ideal:

Let $f \in I$. Then $f = g \cdot (\theta_1^2 + \theta_2^2 - 1)$ for some $g \in R$, and

$$Df = Dg \cdot (\theta_1^2 + \theta_2^2 - 1) + g \cdot \underbrace{(2\theta_1\theta_2 - 2\theta_1\theta_2)}_{=0} \in I.$$

Trigonometric Functions and Algebraic Relations

Therefore the quotient R/I is a differential ring with an induced derivation D^*.

We may now consider the field of fractions

$$K = \left\{ \frac{u}{v} \middle| u, v \in R/I \right\}$$

as the field to work over.

Trigonometric Functions and Algebraic Relations

It is easier to work with regular polynomials than with elements of quotient fields. Since elements of the quotient field can be represented by polynomials, we can in fact continue to work with polynomials, but we need to bring them to a unique normal form.

For example, we could impose the restriction that $\cos x$ does not appear with an exponent greater than one. Then any factor $\cos^2 x$ is rewritten as $1 - \sin^2 x$.

Gröbner Bases

In general, the tool to reduce elements to a normal form modulo an ideal I are Gröbner bases.

In a univariate polynomial ring, any ideal is principal, i.e. generated by a single polynomial v. A normal form of an arbitrary polynomial f is found by dividing f by v. The remainder is the normal form.

Gröbner bases generalize this procedure for multivariate polynomial rings in which ideals in general are not principal.

The Ideal of Algebraic Relations

The algebraic relations among the generators $x, \theta_1, \ldots, \theta_n$ form a differential ideal I. Any element of I is to be considered zero by the algorithm.

In addition to describing the function field by a set of differential equations, perform the following preparations before running the algorithm:

- Construct the ideal I of algebraic relations by specifying a set of generators.
- Compute a Gröbner basis of I so that later polynomials can be reduced to a normal form with that basis.

The Modified Algorithm

- Express the integrand as an element of $K(x, \theta_1, \ldots, \theta_n)$ pretending there are no algebraic relations.
- Make the same ansatz as before:

$$\int f\,dx = v_0 + \sum_{i=1}^{n} \lambda_i \log v_i.$$

- After applying the derivation D and multiplying the equation by the common denominator, reduce both sides of the equation to a normal form using a Gröbner basis of the ideal of algebraic relations.
- Obtain a system of linear equations by comparing the coefficients of monomials as before.

Algebraic Integrands

With these extensions, the algorithm is also able to compute certain integrals involving algebraic functions. Consider for example the integral

$$\int x\sqrt{1+x}dx.$$

The field $\mathbb{Q}(x, y)$ with y algebraic over $\mathbb{Q}(x)$ is constructed as the field of fractions of the quotient field of $R = \mathbb{Q}[x, y]/\langle y^2 - x - 1\rangle$. Moreover we define

$$Dy = \frac{1}{2y}.$$

Internally, the algorithm works with polynomials in x and y and reduces them modulo $I = \langle y^2 - x - 1\rangle$.

More Failures

While the reduction modulo the ideal of algebraic relations fixes many, there are still integrals that would fail to be computed, such as:

$$\begin{aligned}
\int \frac{1}{\sin x} dx &= \frac{1}{2}\log(\cos x - 1) - \frac{1}{2}\log(\cos x + 1) \\
\int \frac{\sqrt{1+x}}{x} dx &= 2\sqrt{1+x} - \log(\sqrt{1+x} + 1) + \\
&\quad \log(\sqrt{1+x} - 1) \\
\int \sqrt{e^x + 1}\, dx &= -x + 2\sqrt{e^x + 1} + 2\log\left(\sqrt{e^x + 1} - 1\right)
\end{aligned}$$

Obviously the reason is that the logarithms appear unexpectedly. We call such logarithmic terms *spurious*.

Spurious Logarithmic Terms

There are three common sources of spurious logarithmic terms:

- As a result of nonunique factorizations, e.g.

$$\sin^2 x = (1 + \cos x)(1 - \cos x)$$

- From polynomials which are units modulo the ideal of algebraic relations, e.g.

$$y \pm x \;\text{ where }\; y = \sqrt{1 + x^2}$$

- From factors of Dy that depend only on y, e.g.

$$y = 1 + \tan^2 x \;\Rightarrow\; Dy = 2y \cdot \tan x$$

Denominator Terms

There may be other factors besides exponentials that should not have their exponent decreased when the denominator is predicted:

$$\begin{aligned}\int \frac{1}{\sqrt{1+x^2}-1}dx &= \frac{-x}{\sqrt{1+x^2}-1} - \log\left(x-\sqrt{1+x^2}\right)\\ \int \frac{1}{\sin x+1}dx &= \frac{-\cos x}{\sin x+1}\end{aligned}$$

The reason is that they have roots of order higher than one despite being irreducible.

References

Manuel Bronstein.
Symbolic Integration I, Transcendental Functions, volume 1.
Springer-Verlag, 1997.

Bruno Buchberger.
A Theoretical Basis for the Reduction of Polynomials to Canonical Forms.
SIGSAM Bulletin, 10(3):19–29, 1976.

Keith O. Geddes and L. Yohanes Stefanus.
On the Risch-Norman integration method and its implementation in MAPLE.
In *ISSAC '89: Proceedings of the ACM-SIGSAM 1989 international symposium on Symbolic and algebraic computation*, pages 212–217, New York, NY, USA, 1989. ACM.