The Difficulties of Definite Integration

James H. Davenport*
Department of Computer Science
University of Bath
Bath BA2 7AY England
J.H.Davenport@bath.ac.uk

September 13, 2003

1 Indefinite Integration and Differentiation

Indefinite integration is the inverse operation to differentiation, and, before we can understand what we mean by indefinite integration, we need to understand what we mean by differentiation.

1.1 What is differentiation?

1. An analytic operation:
   $f'(x_0) = \lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0}$.
2. An algebraic operation, satisfying $(a+b)' = a' + b'$, $(ab)' = a'b + b'a$, $x' = 1$.

We note the different ways in which the fact that we mean "differentiation with respect to x" is expressed in the two formulations.

1.1.1 Two interpretations of atan

Analytically, $\mathrm{atan}(x) = y : y = \tan(x)$ and $-\pi/2 < y < \pi/2$.
Algebraically, $\mathrm{atan}(f)' = \frac{f'}{1+f^2}$. Therefore only defined "up to a constant".
Analytically, c is a constant iff $c(x_1) = c(x_2) \forall x_1, x_2$, and, in this view, the Heaviside function is not a constant.
Algebraically, c is a constant iff $c' = 0$, and, in this view, the Heaviside function is a constant.

*The author is grateful to many people for their discussions of definite integration, notably Andrews Adams, Jacques Carrette, Tony Hearn and Daniel Lichtblau.

1.1.2 Comparing the two approaches

- The success of computer algebra is that one can model the first by the second.
- The problem of computer algebra is that the model, particularly when it comes to inverse differentiation and the handling of "up to a constant", is not perfect.

The first deals with functions (say $\mathbf{R} \mapsto \mathbf{R}$), the second with algebraic expressions.

1.2 Indefinite integration

- Solved for elementary functions (algebraic, exponential, trigonometric, hyperbolic and their inverses) [17, 18, 12, 20, 7];
- Solved for elementary and error functions [8];
- Solved for elementary functions and logarithmic integrals [9];

always subject to the underlying problem of deciding if constants are zero or not [16].

Many possible expressions for the answer.

1.2.1 General outline

Let $f \in C(x, \theta_1, \dots \theta_n)$, θ_i elementary.

1. Consider $f \in K(\theta_n)$.
2. Let F be the most general form of $\int f$ (Liouville's Theorem).
3. Split $f = F'$ into polynomial and rational function parts.
4. Polynomial part solve by equating coefficients of θ_n.
5. Rational part solve by integration by parts.

1.2.2 Most general form (Liouville)

Let $f = p(\theta_n) + \frac{q(\theta_n)}{r(\theta_n)}$. Then F is

$$\overline{p}(\theta_n) + \frac{\overline{q}(\theta_n)}{\overline{r}(\theta_n)} + \sum_i c_i \log v_i,$$

where $\overline{r}, v_i$ divide r and the c_i are constants.
It is then simple to deduce degree bounds etc.

But this theorem is hard to generalise, for example[1] $\int \operatorname{erf}(ax)\operatorname{erf}(bx)$ involves $\operatorname{erf}(\sqrt{a^2+b^2}x)$!

1.3 Examples of choice of integral

1.3.1 Which arctan?

$$\int \frac{1}{x^2-8x+17}\mathrm{d}x = \arctan(x-4) = \arctan\left(\frac{x-5}{x-3}\right) \tag{1}$$

$$[\arctan(x-4)]_2^4 = \arctan 2 \approx 1.107$$

$$\left[\arctan\left(\frac{x-5}{x-3}\right)\right]_2^4 = \frac{-\pi}{4} - \arctan 3 \approx -2.034$$

At $x=3$ the second form of the integrand went through the "branch cut at infinity" of arctan, so the two answers differ by

$$-\pi = \lim_{x\to 3^+} \arctan\frac{x-5}{x-3} - \lim_{x\to 3^-} \arctan\frac{x-5}{x-3}.$$

This particular case is resolved by [14], who show how to choose an expression in arctan without such "spurious" branch cuts. Nonetheless, this very simple example shows what can go wrong with an over-simplistic use of indefinite integration to do definite integration.

A further interesting phenomenon occurs when one tries to integrate one of these alternatives: $\arctan\left(\frac{x-5}{x-3}\right)$. Maple returns

$$(x-3)\arctan\left(\frac{x-5}{x-3}\right)+\ln\left(\frac{-2}{x-3}\right)-\frac{1}{2}\ln\left(1+\left(\frac{x-5}{x-3}\right)^2\right)-\arctan(x-4)+\frac{\pi}{4},$$

which at $x=2$ evaluates to a real number, but at $x=4$ evaluates to $\frac{1}{2}\ln 2 + i\pi$, so a naïve implementation of definite integration through indefinite integration would return a complex number after integrating a real function through a real range. In fact, $i\pi$ is precisely the jump discontinuity in the integral as we pass through $x=3$, but it could as well have had a real part also.

1.3.2 Logarithms and signs

It would be natural to assume[2] that

$$\int \frac{1}{1-x}\mathrm{d}x = -\log(1-x), \tag{2}$$

[1] More precisely: $\int \operatorname{erf}(ax)\operatorname{erf}(bx)\mathrm{d}x =$

$$x\operatorname{erf}(ax)\operatorname{erf}(bx)+\frac{e^{-a^2x^2}\operatorname{erf}(bx)}{a\sqrt{\pi}}+\frac{e^{-b^2x^2}\operatorname{erf}(ax)}{b\sqrt{\pi}}-\frac{b\operatorname{erf}\left(\sqrt{a^2+b^2}x\right)}{a\sqrt{\pi}\sqrt{a^2+b^2}}-\frac{a\operatorname{erf}\left(\sqrt{a^2+b^2}x\right)}{b\sqrt{\pi}\sqrt{a^2+b^2}}.$$

[2] I am grateful to Jacques Carrette for this example.

and similarly that

$$\int \frac{-1}{x-1} \mathrm{d}x = -\log(x-1). \tag{3}$$

The integrands in equations (2–3) are clearly equal in the normal algebra of $\mathbf{Q}(x)$. However, the integrals differ by

$$\begin{cases} -i\pi & x > 1 \\ i\pi & x < 1 \\ i\pi & \Im(x) > 0 \\ -i\pi & \Im(x) < 0 \end{cases}.$$

Again, we see that "up to a constant" has a more subtle meaning than we might expect, and what might seem a trivial choice has ramifications. In particular, the right-hand side of (2) has its branch cut immediately above the real axis (for $x > 1$).

In has been suggested [19] that the correct answer is

$$\int \frac{1}{1-x} \mathrm{d}x = -\log|1-x|. \tag{4}$$

This, of course, is only valid over $\mathbf{R}$, and differs from the answer in either (2) or (3) by $\pm i\pi$ as necessary, to make it real everywhere on $\mathbf{R}$.

1.4 The Fundamental Theorem of Calculus

There are many possible statements of this theorem, but we shall use the following one.

Theorem 1 *Given $f : [a,b] \mapsto \mathbf{R}$ ($[a,b]$ may be infinite) and a function F, if*

- *f is Riemann-integrable on $[a,b]$ and*
- *throughout $[a,b]$*[3]*, $F' = f$,*

then $\int_a^b f(x)\mathrm{d}x = F(b) - F(a)$.

It is well known that F has to be continuous, and differentiable.

The Appendix gives an exegesis of the Fundamental Theorem in terms of elementary formulae, and gives a context in which the Theorem is applicable in computer algebra.

1.5 Examples of ignoring integrability

$\int \frac{1}{x^2}\mathrm{d}x = \frac{-1}{x}$. "So"

$$\int_{-1}^{1} \frac{1}{x^2} \mathrm{d}x \stackrel{?}{=} \left[\frac{-1}{x}\right]_{-1}^{1} = -1 - 1 = -2.$$

[3]In fact, it suffices to have $f(a) = \lim_{h \to 0^+} \frac{F(a+h)-F(a)}{h}$ and $f(b) = \lim_{h \to 0^-} \frac{F(b+h)-F(b)}{h}$.

$\int \frac{1}{x^3}\mathrm{d}x = \frac{-1}{2x^2}$. "So"

$$\int_{-1}^{1} \frac{1}{x^3}\mathrm{d}x \stackrel{?}{=} \left[\frac{-1}{2x^2}\right]_{-1}^{1} = -\frac{1}{2} + \frac{1}{2} = 0.$$

The first is plain wrong[4], the second happens to yield the Cauchy principal value of the integral (without warning).

1.6 So we should always check for Riemann integrability?

In theory, certainly. But:

- Riemann integrability is essentially an *analytic* question, and we have historically tried to do computer *algebra* (not a *good* reason);
- I know of no general-purpose algorithm for doing so, and the definition is hard to make constructive;
- I know of no computer algebra system which does so in general.

Instead most computer algebra systems **implicitly** rely on what I have come to call

Hypothesis 1 (of closed form integrability) *If f has no singularities in $[a, b]$ and has a closed form integral, then f is Riemann-integrable in $[a, b]$.*

What one can prove is that, if the elementary formula corresponding to f has no *apparent* singularities in a closed interval, then f is Riemann-integrable in that interval (see Appendix). One might ask about, say, $\frac{\cos\left(\frac{1}{x}\right)}{x^2}$, whose integral is $-\sin(1/x)$, but on an interval not including 0, it is in fact Riemann-integrable, and at 0 it is undefined (indeed as we approach 0 it is unbounded in both directions).

1.7 A pseudo-algorithm for definite integration

We assume that the interval of integration $[a, b]$ is finite: if not, limiting processes are in order, but the modifications are trivial[5].

1. If f does not have an indefinite integral in closed form F, give up.
2. If f or F have singularities in (a, b), split the integral at each singularity[6] and recurse.

[4] We are integrating a strictly positive, indeed ≥ 1, function.

[5] Provided that the singularities are bounded. If they are not, then purely algebraic techniques seem doomed to failre, and we need to reason about a "generic" singularity.

[6] There will be problems here if there are infinitely many singularities, and theorem-proving probably has much to offer in this case.

3. If f has a singularity at a, consider
$$\lim_{c \to a+} \int_c^b f(x)\mathrm{d}x.$$
4. If f has a singularity at b, consider
$$\lim_{d \to b-} \int_a^d f(x)\mathrm{d}x.$$

1.8 Integration is dependent on limits?

Just as Riemann pointed out, but not in the same manner in practice.

- By assuming hypothesis 1, we are ignoring the problems of letting mesh size tend to zero with bounded functions such as $\sin \frac{1}{x}$.
- The treatment of singularities at end-points is consistent with that of Riemann integration.
- By treating internal singularities more carefully and combining limits, it is possible to compute Cauchy principal values *if requested*. In the case of the second example of section 1.5, we would write

$$\begin{aligned}
\int_{-1}^{1} \frac{1}{x^3}\mathrm{d}x &= \int_{-1}^{0} \frac{1}{x^3}\mathrm{d}x + \int_{0}^{1} \frac{1}{x^3}\mathrm{d}x \\
&= \lim_{\epsilon \to 0^+} \left(\int_{-1}^{-\epsilon} \frac{1}{x^3}\mathrm{d}x \right) + \lim_{\epsilon \to 0^+} \left(\int_{\epsilon}^{1} \frac{1}{x^3}\mathrm{d}x \right) \\
&= (-\infty) + \infty = \text{undefined}.
\end{aligned}$$

But, combining limits to get the Cauchy principal value,

$$\begin{aligned}
&= \lim_{\epsilon \to 0^+} \left(\int_{-1}^{-\epsilon} \frac{1}{x^3}\mathrm{d}x \int_{\epsilon}^{1} \frac{1}{x^3}\mathrm{d}x \right) \\
&= \lim_{\epsilon \to 0^+} \left(\frac{1}{2} + \frac{-1}{2\epsilon^2} - \frac{-1}{2\epsilon^2} - \frac{1}{2} \right) \\
&= \lim_{\epsilon \to 0^+} 0 = 0
\end{aligned}$$

- Definite integration through indefinite integration cannot be better than the (one-sided) “limit” software it uses.
- But this may need to be tuned to the problems that tend to come up this way. However, special-purpose limit ideas in integration are bad [15].

1.9 Integration depends on singularity detection?

Certainly so. The integrator

- **must** detect all singularities, including jump discontinuities, in both f and F.

- **should not** report apparent singularities, where in fact cancellation means that there is not a problem.

* $3x^2\sin\left(\frac{1}{x}\right)-x\cos\left(\frac{1}{x}\right)$ (derivative of $x^3\sin\left(\frac{1}{x}\right)$) does not have a singularity at $x=0$.

Strong interaction with the "limit" code.

Consider [15], following a MAA problem, the following results from Mathematica.

$$\begin{aligned}\int_0^\pi\int_0^\pi|\sin(x-y)|\mathrm{d}y\mathrm{d}x &= 2\pi \quad \text{(right)}\\ \int_0^{2\pi}\int_0^{2\pi}|\sin(x-y)|\mathrm{d}y\mathrm{d}x &= 4\pi \quad \text{(wrong)}\end{aligned}$$

In both integrals the singularity $x=y$ is recognised, but in the second, $x\equiv y+\pi \pmod{2\pi}$ is not recognised as a (broken) singularity line, and therefore the contributions from $(0,\pi)\times(\pi,2\pi)$ and $(\pi,2\pi)\times(0,\pi)$ were computed as if the integrand were $\sin(x-y)$, and therefore gave 0.

2 Other ways of computing definite integrals

We are all familiar with other ways of computing definite integrals, as in

$$\begin{aligned} & \left(\int_0^\infty \exp(-x^2/2)\mathrm{d}x\right)^2\\ = & \left(\int_0^\infty \exp(-x^2/2)\mathrm{d}x\right)\left(\int_0^\infty \exp(-y^2/2)\mathrm{d}y\right)\\ = & \int_0^\infty\int_0^\infty \exp(-(x^2+y^2)/2)\mathrm{d}x\mathrm{d}y\\ = & \int_0^\infty\int_0^{\frac{\pi}{2}} \exp(-r^2/2)r\mathrm{d}\theta\mathrm{d}r\\ = & \frac{\pi}{2}\int_0^\infty \exp(-r^2/2)r\mathrm{d}r\\ = & \frac{\pi}{2}\left[-\exp(-r^2/2)\right]_0^\infty = \frac{\pi}{2}\\ \Rightarrow & \qquad \int_0^\infty \exp(-x^2/2)\mathrm{d}x = \sqrt{\frac{\pi}{2}}.\end{aligned}$$

2.1 Integration tables?

These seem to be the only way of handling such a variety of *ad hoc* techniques at the moment, though one might hope that a suitable series of tactics for a theorem prover might work.

- Notoriously unreliable; up to 26% error rates [13].

- Unless there is an exact match[7], there is the problem of ensuring that any transformations are *analytically* valid.
- Need theorem-prover support for verification and application [2, 3].

2.2 Integration by convolution of Meijer G functions

To the best of my knowledge, this method is only implemented in Mathematica and Maple.

2.2.1 Definition of Meijer G functions

Let:

$$
\begin{aligned}
a_s &= [a_1, \ldots, a_m] \qquad & \Gamma(1 - a_s + y) &= \prod_{i=1}^{m} \Gamma(1 - a_i + y); \\
b_s &= [b_1, \ldots, b_n] \qquad & \Gamma(b_s - y) &= \prod_{i=1}^{n} \Gamma(b_i - y); \\
c_s &= [c_1, \ldots, c_p] \qquad & \Gamma(c_s - y) &= \prod_{i=1}^{p} \Gamma(c_i - y); \\
d_s &= [d_1, \ldots, d_q] \qquad & \Gamma(1 - d_s + y) &= \prod_{i=1}^{q} \Gamma(1 - d_i + y).
\end{aligned}
$$

$$
\mathrm{G}([a_s, b_s], [c_s, d_s], z) =
$$

$$
G_{pq}^{mn}\left(z \,\middle|\, \begin{array}{cccccc} a_1 & \ldots & a_m & a_{m+1} = c_1 & \ldots & a_{m+p} = c_p \\ b_1 & \ldots & b_n & b_{n+1} = d_1 & \ldots & b_{n+p} = d_q \end{array} \right) =
$$

$$
\frac{1}{2\pi i} \oint_L \frac{\Gamma(1 - a_s + y)\Gamma(c_s - y)}{\Gamma(b_s - y)\Gamma(1 - d_s + y)} z^y \mathrm{d}y.
$$

Where the integration path L is one of:

L_∞ : from $\infty + i\phi_1$ to $\infty + i\phi_2$ with $\phi_2 > \phi_1$;

$L_{-\infty}$: from $-\infty + i\phi_1$ to $-\infty + i\phi_2$ with $\phi_2 > \phi_1$;

$L_{\gamma + i\infty}$: from $\gamma - \infty$ to $\gamma + i\infty$.

In fact the Meijer G functions are precisely those functions which have an expansion of the form $G(z) = z^\alpha \sum f_n z^n$ where $\frac{A(f_{n+1})}{B(f_{n+1})} = \frac{C(f_n)}{D(f_n)}$: A, B, C, D polynomials.

[7] Of both integrand and range.

2.2.2 Integration of Meijer G functions

If the integral is of the form $\int_0^\infty z^k G_1(z)\mathrm{d}z$ or $\int_0^\infty z^k G_1(z)G_2(z)\mathrm{d}z$ (where G_1 and G_2 are MeijerG functions), then a method based on convolution of MeijerG functions can compute the integral, subject to appropriate (Often quite hard to check, and to do with convergence properties at infinity)[8] conditions.

Since the Heaviside function $H_t(x) = \begin{cases} 1 & x \geq t \\ 0 & x < t \end{cases}$ can be represented as a MeijerG function, we can compute

$$\int_t^\infty z^k G_1(z)\mathrm{d}z = \int_0^\infty z^k G_1(z) H_t(z)\mathrm{d}z.$$

Hence, if both integrals on the right converge,

$$\int_0^t z^k G_1(z)\mathrm{d}z = \int_0^\infty z^k G_1(z)\mathrm{d}z - \int_t^\infty z^k G_1(z)\mathrm{d}z.$$

2.2.3 Common functions are Meijer G functions

$$\begin{array}{rcl}
z & = & \mathrm{G}([[], [2]], [[1], [0]], z) \\
z \ln z & = & -\,\mathrm{G}([[], [2, 2]], [[1, 1], []], z) \\
\exp(z) & = & \mathrm{G}([[], []], [[0], []], -z) \\
z\exp(z) & = & \mathrm{G}([[], []], [[1], []], -z) \\
\mathrm{Ei}(z) & = & \mathrm{G}([[], [1]], [[0, 0], []], z) \\
(z+1)\mathrm{Ei}(z) - \exp(z) & = & -\,\mathrm{G}([[], [2]], [[0, 0], []], -z) \\
z\mathrm{Ei}(-z) - \exp(-z) & = & \mathrm{G}([[], [2]], [[1, 0], []], z) \\
z\mathrm{Ei}(z) & = & \mathrm{G}([[], [2]], [[1, 1], []], z) \\
-(2\gamma - 1 + \ln(-z))z + \dfrac{z^2}{2}\,{}_2F_3(1, 1; 2, 2, 3; -z) & = & \mathrm{G}([[], [2]], [[1, 1], [0]], z) \\
-\left(1 + \gamma + \ln(z)\right)z - 1/2\,\left(\gamma - 7/2 + \ln(z)\right)z^2 & = & \mathrm{G}([[], [3, 3]], [[1, 1], [0]], z) \\
-\sqrt{z} I_1(2\sqrt{z}) & = & \mathrm{G}([[], []], [[1], [0]], -z) \\
\text{(where } I \text{ and } J \text{ are Bessel's functions)} & & \\
\sqrt{z} J_1(2\sqrt{z}) & = & \mathrm{G}([[], []], [[1], [0]], z) \\
z J_1(2z) & = & \mathrm{G}([[], []], [[1], [0]], z^2)
\end{array}$$

So many standard functions can be expressed in terms of MeijerG functions.

3 Pragmatics

Various issues come up in implementing these algorithms, which may affect running time, intelligibility/usability of the answer, or even whether an answer (as opposed to an error) is returned.

[8]Private Communication from Jacques Carrette

3.1 To expand or not to expand?

$\int_0^\infty (e^{i/x}-1)\sin x\mathrm{d}x$ *must not* be expanded [15], as the apparent singularities cancel, at both 0 and ∞.
On the other hand, consider the integral:

$$\int_2^5 \frac{7x^{13}+10x^8+4x^7-7x^6-4x^3-4x^2+3x+3}{x^{14}-2x^8+2x^7-2x^4-4x^3-x^2+2x+1}\mathrm{d}x.$$

This could be represented as

$$\int_2^5 \frac{7x^{13}}{x^{14}-2x^8+2x^7-2x^4-4x^3-x^2+2x+1}\mathrm{d}x+ \int_2^5 \frac{10x^8}{x^{14}-2x^8+2x^7-2x^4-4x^3-x^2+2x+1}\mathrm{d}x\ldots$$

As it happens, Mathematica does not expand and gets [15]:

$$\frac{1}{2}\left(\ln\frac{6102576361}{15553}+\sqrt{2}\ln\left(\frac{1}{\frac{1}{2}-\frac{232482\sqrt{2}}{9764515}}-1\right)\right)$$

Maple does, and gives a six-page answer, with three-line numbers.
But in many examples, expansion does help.

3.2 Conditional Results (parameters)

[15] considers $\int_0^1\int_0^1 |x-y|^n\mathrm{d}x\mathrm{d}y$.

- Mathematica, by default, generates nonsense, e.g. 4.67 for $n=3.2$.
- Maple, by default, fails to integrate.
- Mathematica, if asked to generate conditions, produces $\frac{2}{2+3n+n^2}$ if $\Re(n)>-1$, otherwise unevaluated.
- Maple, if told[9] [21] that $n>-1$, also produces $\frac{2}{2+3n+n^2}$.

[15] also considers $\int_{-\infty}^\infty e^{-\frac{a(x-b)^2}{\sigma^2}}\mathrm{d}x$.

- Mathematica integrates this via the MeijerG method (see section 2.2), for which it is told $a>0$, $\sigma\in\mathbf{R}$.
- With condition generation, it gets $\sqrt{\frac{\pi}{a}}|\sigma|$ if $\sigma\neq 0$ and $\Re(b)<0$, otherwise undefined.
- Without condition generation, it gets an unconditional 0, since the two halves are integrated using contradictory assumptions on b — see section 3.3.

[9] It seems that, in Maple 8, trying to `assume(Re(n)>-1)` actually makes n real and >-1.

- Maple gets the equivalent[10]

$$\begin{cases} \sqrt{\pi}\frac{1}{\sqrt{\frac{a}{\sigma^2}}} & \operatorname{csgn}\left(a\overline{\sigma}^2\right) = 1 \\ \infty & \text{otherwise} \end{cases}.$$

However, condition generation is not a panacea, since it can produce very unwieldy results: in Mathematica $\int_a^b \ln(x)\mathrm{d}x$ can consume over half a page due to the perceived problem of $\Im(b) = \Im(a)$, whereas Maple simply produces the (correct[11]) $b \ln b - b - a \ln a + a$.

3.3 R or C?

This is a question that bedevils the whole of computer algebra practice, but is acute here.

- Systems do not clearly document when $\mathbf{R}$ is assumed, and when $\mathbf{C}$: in analysis the two are separate domains. $\mathbf{R} \subset \mathbf{C}$ but $\overline{\mathbf{R}} \not\subset \overline{\mathbf{C}}$ (where the bar indicates topological closure). See also section 1.3.2, where the "best" answer over $\mathbf{R}$ is not valid over $\mathbf{C}$.
- No system that I know has a clear internal design philosophy in this area.
- Contradictory assumptions on the sign of b got $\int_{-\infty}^{\infty} e^{-\frac{a(x-b)^2}{\sigma^2}}\mathrm{d}x$ wrong.

4 Conclusions

- Indefinite integration can be viewed as the inverse of differentiation, and this can be viewed as an algebraic process.
- Definite integration is inherently an analytic process, and one cannot pass from one to the other without taking care of the analysis,
- *ad hoc* treatments of the analysis tend to work only for the cases considered.
- A definite integrator is a test of a lot of the subroutines of an algebra system! Examples we have seen are singularity detection, branch cut detection and limit computation, but general branch-cut respecting simplification is also required [4, 6, 5].

[10] Maple's "csgn" function is a useful complex extension of the traditional "sign" function, defined as $\operatorname{csgn}(x+iy) = \begin{cases} 1 & x > 0 \\ -1 & x < 0 \\ 1 & x = 0,\, y > 0 \\ -1 & x = 0,\, y < 0 \\ 0 & x = y = 0 \end{cases}.$

[11] Whether by careful consideration of putative difficulties, or a cavalier disdain for them, only the authors can say.

4.1 Computer Algebra Systems

There are some lessions that the designers of computer algebra systems need to take to heart if their systems are to have powerful definite integrators.

- If the subroutines are not powerful enough, the integrator will not be. They need to be able to return three-valued logics: singularities/there are none/cannot decide.
- The system needs to be able to check many transformations for *analytic* validity.
- **R** versus **C** needs to be explicit.
- Condition generation needs to be at least an option, and possibly the default for conditions not of measure zero.

4.2 Theorem provers using computer algebra

Here we outline some of the things that theorem-prover designers need to be aware of if they are using computer algebra systems as "black boxes".

- Use of computer algebra systems as oracles to compute indefinite integrals is pretty safe, and a positive result should be checkable by the theorem prover.
- In the current state of the art, use as oracles to compute arbitrary definite integrals is a gamble. Numerical cross-checking is recommended (it's what we all do!).
- Specific classes of definite integrals may well be safe in specific systems, e.g. rational functions (provided [14] is implemented).

4.3 Theorem provers to help computer algebra

Here we mention some of the ways in which a fruitful collaboration could be of benefit.

- Computer algebra systems have great difficulties with infintely many objects, and functions with infinitely many singularities are no exception. The ability to reason about "a generic singularity" would be helpful.
- Some of the proofs, and particularly side-conditions, on definite integrals in tables, need to be checked mechanically.
- Neither party, to be best of my knowledge, has a really good theory of branch cuts (but see [11, 5]).

References

[1] Abramowitz, M. and Stegun, I. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. 10th Edition, *US Government Printing Office* (1964).

[2] Adams, A., Gottliebsen, H., Linton, S. and Martin, U. Automated Theorem proving in Support of Computer Algebra: Symbolic Definite Integration as a Case Study. In *Proceedings ISSAC '99* (1999), S. Dooley, ACM, N. York, and 1999, Eds., pp. 253–260.

[3] Adams, A., Gottliebsen, H., Linton, S. and Martin, U. VSDITLU: a verified symbolic definite integral table look-up. *Proc. CADE 16* (1999), 112–126.

[4] Beaumont, J.C., Bradford, R.J. and Davenport, J.H. Better Simplification of Elementary Functions Through Power Series. In *Proc ISSAC 2003* (2003) J.R. Sendra, Ed., ACM Press, 2003, pp. 30–36.

[5] Bradford, R., Corless, R., Davenport, J., Jeffrey, D. and Watt, S. Reasoning about the Elementary Functions of Complex Analysis. *Annals of Mathematics and Artificial Intelligence 36* (2002), 303–318.

[6] Bradford, R. and Davenport, J. Towards Better Simplification of Elementary Functions. In *Proceedings ISSAC 2002* (2002), T. Mora, Ed., pp. 15–22.

[7] Bronstein, M. *Integration of Elementary Functions*. Ph.D. thesis, Univ. California Berkeley, 1987.

[8] Cherry, G. Integration in Finite Terms with Special Functions: the Error Function. *J. Symbolic Comp. 1* (1985), 283–302.

[9] Cherry, G. Integration in Finite Terms with Special Functions: the Logarithmic Integral. *SIAM J. Computing 15* (1986), 1–21.

[10] Corless, R., Davenport, J., Jeffrey, D. and Watt, S. According to Abramowitz and Stegun. *SIGSAM Bulletin 2 34* (2000), 58–65.

[11] Corless, R. and Jeffrey, D. The Unwinding Number. *SIGSAM Bulletin 2 30* (1996), 28–35.

[12] Davenport, J. On the Integration of Algebraic Functions. *Springer Lecture Notes in Computer Science 102* (1981).

[13] Klerer, M. and Grossman, F. Error Rates in Tables of Indefinite Integrals. *Industrial Math. 18* (1968), 31–62.

[14] Lazard, D. and Rioboo, R. Integration of Rational Functions — Rational Computation of the Logarithmic Part. *J. Symbolic Comp. 9* (1990), 113–115.

[15] Lichtblau, D. Issues in symbolic definite integration. Lecture and handout at Applications of Computer Algebra 2003, Raleigh, N.C., 28 July 2003.

[16] Richardson, D. Some Unsolvable Problems Involving Elementary Functions of a Real Variable. *Journal of Symbolic Logic 33* (1968), 514–520.

[17] Risch, R. The Problem of Integration in Finite Terms. *Trans. A.M.S. 139* (1969), 167–189.

[18] Risch, R. The Solution of the Problem of Integration in Finite Terms. *Bulletin AMS 76* (1970), 605–608.

[19] Stoutemyer, D.R. Should computer algebra programs use $\ln x$ or $\ln|x|$ as their antiderivation of $1/x$. *DERIVE Newsletter 26* (Jun 1997), 3–6.

[20] Trager, B. *Integration of Algebraic Functions*. Ph.D. thesis, M.I.T. Dept. of EE&CS, 1984.

[21] Weibel, T. and Gonnet, G. An Assume Facility for CAS with a Sample Implementation for Maple. In *Proceedings DISCO '92* (1993).

A Riemann Integrability of elementary functions

A.1 Preliminaries

Notation 1 *Let $K = C(x)$ be a differential field of characteristic zero, with $x' = 1$ and $c' = 0 \forall c \in C$.*

C will normally be a (constructive sub-field of) **R**.

Definition 1 *Let L be a differential extension of K, and $M = K(\theta)$. We say that M is an* elementary *extension of L if the constants of M are the constants of L and one of the following three holds:*

(i) *θ is algebraic over L, in which case we may assume, without loss of generality, that the minimal polynomial of θ is monic;*

(ii) *θ is transcendental over L and $\exists \phi \in L$ with $\theta' = \phi'/\phi$, in which case we say that θ is a* logarithm *of ϕ;*

(iii) *θ is transcendental over L and $\exists \phi \in L$ with $\theta'/\theta = \phi'$, in which case we say that θ is an* exponential *of ϕ.*

We say that M is elementary *over K if there is a chain $K = L_0 \leq L_1 \leq \ldots \leq L_n = M$ and each L_i is an elementary extension of L_{i-1}. We say that ϕ is an* elementary formula *if it can be written in some M an elementary extension of K.*

Since trigonometric and hyperbolic functions and their inverses can be written in terms of exponentials and logarithms, this covers them as well. We will consider such functions as "first-class objects", and not try to rewite them in terms of (often complex) exponentials and logarithms. We should also note that logarotihms are only defined up to an additiive constant, and exponentials up to a (non=zero) multiplicative constant.

A.2 Formulae and Functions

In doing definite integration via indefinite integration, the algebraic side of the theory is concerned with *formulae*, whereas the analytic side of the theory is concerned with the (partial) *functions* $\mathbf{R} \to \mathbf{R}$ that they denote.

Notation 2 *Let greek letters denote formulae, and the corresponding latin letters the (partial) functions* $\mathbf{R} \to \mathbf{R}$ *that they denote so that* f *is the function denoted by* ϕ*, and* G *that denoted by* Γ*. We note that the map from formulae to functions is many–one, as in equation (1).*

We need to know precisely which function is denoted by a formula. For this we use the branch cuts of [10], which refines, in terms of behaviour on the branch cuts, those of [1], with one exception. Since we are concerned with functions $\mathbf{R} \to \mathbf{R}$ we will use $\sqrt[n]{}$ to denote the *real* n-th root function[12], viz. for n even, the positive n-th root of a positive number, $\sqrt[n]{0} = 0$, and undefined otherwise; whereas for n odd $\sqrt[n]{x} = y \in \mathbf{R}$ where $y^n = x$.

We can then regard elementary formulae as expression trees, which are to be evaluated eagerly[13] from the leaves to the root.

Definition 2 *We say that* ϕ *is* wholly real *on* $[a, b]$ *if and only if, for every* $x \in [a, b]$*, all nodes in the evaluation of* ϕ *at* x *give real values.*

The "wholly real" properties of the various elementary operations are given in table 1.

A.3 Applications to the Fundamental Theorem

Proposition 1 *If* ϕ *is an elementary formula which is wholly real on* $[a, b]$*, then the corresponding function* f *is continuous, and hence Riemann-integrable, on* $[a, b]$*.*

This follows from the fact that the composition of continuous functions is continuous, and, from the definition of "wholly real", no singularities intervene. The examples of section 1.5 are not counter-examples, since $\frac{1}{x^2}$ and $\frac{1}{x^3}$ are not wholly real on $[-1, 1]$ because of the case $x = 0$.

[12] Maple refers to this as the `surd` function.

[13] That is to say that every node is evaluated, even if we know that it is later to be multiplied by 0, as in evaluating $(x-3)\arctan\left(\frac{x-5}{x-3}\right)$.

Table 1: Properties of elementary operations

Operation(s)	Wholly real on
$+, -, \times$	$(-\infty,\infty)\times(-\infty,\infty)$
$/$	$(-\infty,\infty)\times(-\infty,0), (-\infty,\infty)\times(0,\infty)$
$\sqrt[n]{}$	$\begin{cases} [0,\infty) & n \text{ even} \\ (-\infty,\infty) & n \text{ odd} \end{cases}$
Other algebraic functions	depends on principal branch (no standard definition of this)
ln	$(0,\infty)$
exp, sin, cos, sinh, cosh, tanh, sech	$(-\infty,\infty)$
tan, sec	Any interval without a $(n+\frac{1}{2})\pi : n \in \mathbf{Z}$
cot, cosec	Any interval without a $n\pi : n \in \mathbf{Z}$
coth, cosech, arccsch	$(-\infty,0), (0,\infty)$
arcsin, arccos	$[-1,1]$
arctan, arccot, arcsinh	$(-\infty,\infty)$
arcsec, arccoth	$(-\infty,-1), (1,\infty)$
arccsch	$(1,\infty)$
arccosh	$[1,\infty)$
arctanh	$(-1,1)$
arcsech	$(0,1]$

Theorem 2 (Computer algebra's Fundamental Theorem of Calculus) *If ϕ and Φ are elementary formulae which are wholly real on $[a,b]$, with $\Phi' = \phi$, then F is an indefinite integral (in the analytic sense) of f, and $\int_a^b f(x)\mathrm{d}x = F(b) - F(a)$.*

Because Φ is wholly real, F is continuous. Since $\Phi' = \phi$, $F' = f$ wherever both are defined, since both algebraic differentiation of Φ and analytic differentiation of F obey the same rules, and therefore throughout $[a,b]$. By Proposition 1, f is Riemann-integrable on $[a,b]$. Therefore the conditions of the Fundamental Theorem of Calculus (Theorem 1) are verified, and the conclusion follows.

The example of section 1.3.1 fails to be a counter-example, since $1/(x-3)$ is not wholly real at $x = 3$, and hence the proposed integral $\arctan\left(\frac{x-5}{x-3}\right)$ is not totally real in $[2,4]$. We note that $\gamma = (x-3)\arctan\left(\frac{x-5}{x-3}\right)$ is, by our definition, not wholly real in $[2,4]$, even though the corresponding g can be made continuous in $[2,4]$ by defining $g(3) = 0$, as limits suggest. This precaution in our definition is necessary to avoid concealed jump discontinuities.

A.4 Higher Functions

The same argument applied to any liouvillian functions, i.e. those generated by the solution of first-order linear differential equations. However, in order to apply the result, one needs the equivalent entries to table A.2, and the limiting routines need to know the appropriate behaviours of these new functions.