Integration of Algebraic Functions: A Simple Heuristic for Finding the Logarithmic Part

Manuel Kauers*
RISC-Linz
Johannes Kepler Universität
A-4040 Linz, Austria
mkauers@risc.uni-linz.ac.at

ABSTRACT

A new method is proposed for finding the logarithmic part of an integral over an algebraic function. The method uses Gröbner bases and is easy to implement. It does not have the feature of finding a closed form of an integral whenever there is one. But it very often does, as we will show by a comparison with the built-in integrators of some computer algebra systems.

Categories and Subject Descriptors

I.1.2 [**Computing Methodologies**]: Symbolic and Algebraic Manipulation—*Algorithms*

General Terms

Algorithms

Keywords

Symbolic Integration, Algebraic Functions

1. INTRODUCTION

In 1979, Norman and Davenport wrote [10]

> Ten years ago an integration program could be judged by the proportion of some set of (known integrable) examples it could cope with, whereas now failure to solve an easy problem is seen as a bug.

They characterize here the development from early heuristic approaches to solid algebraic algorithms for integration. Indeed, there is now a complete algorithmic solution to the problem of indefinite integration of elementary functions, "complete" in the sense that there is an algorithm which is guaranteed to find a closed form of an integral in terms of elementary functions whenever there is one [11, 12, 1, 5]. In this situation, why should we still care about incomplete heuristic approaches?

Parts of the complete integration algorithm are rather complicated. Implementors need a thorough mathematical understanding of the underlying theory and a lot of time to create a complete implementation of the complete algorithm. As a matter of fact, all of today's computer algebra systems only contain a partial implementation of the complete algorithm. Difficult parts of the algorithm that are only needed for certain types of integrals are often considered not worth the effort of implementing and are therefore left out. A simple heuristic may be an attractive way of filling these gaps. And even if a complete code is desired, a fast heuristic could be executed as a preprocessor before the complete code is entered. This may enhance the overall performance.

These reasons were stated by Bronstein [6, 5] in favor of parallel integration, an incomplete integration alternative to the Risch algorithm for integration of transcendental elementary functions. An incomplete (yet very successful) integrator of his based on this approach requires less than 100 lines of Maple code [4].

In this paper, we propose an incomplete method for finding the logarithmic part of an algebraic integral. As implementing our method requires no more than about ten lines of code in Maple or Mathematica, it might be interesting in situations where a full implementation of Davenport's [8] or Trager's [13] algorithm, which both involve complicated and time-consuming algebraic computations, is not adequate.

We use Gröbner bases to compute potential contributions to the logarithmic part. Very often, we can find the complete logarithmic part in this way. At least some components of the logarithmic part can usually be found, and some unintegrated part of the integrand may be left. We cannot give any proofs as to how often our method succeeds, but instead we measure its usefulness by comparing it with the integrators of Axiom, Maple, and Mathematica. We do believe that the integrators of these systems would benefit from including our method.

*Supported by the Austrian science foundation FWF, grants P19462-N18 and SFB F1305.

ISSAC'08, July 20–23, 2008, Hagenberg, Austria.

2. TRAGER'S ALGORITHM

Let k be a field, $\mathbb{Q} \subseteq k$. Let $m \in k[x,y]$ with $d := \deg_y m \geq 2$ irreducible over $\bar{k}(x)$. We consider the differential field $K := k(x)[y]/\langle m\rangle$ with the derivation D defined via $Dc = 0$ $(c \in k)$, $Dx = 1$. Elements of K whose minimal polynomial has a leading coefficient with respect to y that is free of x are called integral. We may assume that y itself is integral in K (otherwise, choose a different generator). For polynomials p we will write $\bar{p}$ for its residue class modulo an ideal that will be clear from the context. We will sloppily write $p(x, y, \gamma)$ for the polynomial obtained from p by substituting γ for the last indeterminate, and in similar situations. All these definitions and conventions will be used throughout the paper.

Suppose an integrand $f \in K$ is given. If the integral $\int f$ is elementary, then, according to Liouville's theorem, there exist $g \in K$ and $\gamma_1, \dots, \gamma_r \in \bar{k}$ and $p_1, \dots, p_r \in K(\gamma_1, \dots, \gamma_r)$ such that

$$f = Dg + \gamma_1 \frac{Dp_1}{p_1} + \cdots + \gamma_r \frac{Dp_r}{p_r},$$

and hence $\int f = g + \gamma_1 \log(p_1) + \cdots + \gamma_r \log(p_r)$. Here, g is called the algebraic part and the sum of the $\gamma_i \log(p_i)$ is called the logarithmic part of the integral.

Following Trager [13], g, the γ_i and the p_i may be computed as follows.

1. Choose a point $x_0 \in k$ where f has no pole or branch point, and perform the change of variables $x = 1/(x' - x_0)$, $dx = -1/x'^2\, dx'$. We call the new integrand again f and rename x' back to x.
2. Compute an integral basis $\omega = (\omega_1, \dots, \omega_d)$ of K. This is a basis which generates the subring $\mathcal{O}_{k[x]} \subseteq K$ of all integral elements as a $k[x]$-module.
3. Write
$$f = \frac{a_1\omega_1 + \cdots + a_d\omega_d}{b}$$
for $a_1, \dots, a_d, b \in k[x]$ such that $\gcd(a_1, \dots, a_d, b) = 1$. Then, using a generalization of Hermite reduction, compute $g, h \in K$ such that h has simple poles only and
$$f = Dg + h.$$
4. If the algebraic function $x^2 h(1/x)$ has a pole at the origin, return "not integrable"; the integral is not elementary in this case.
5. Write
$$h = \frac{u_1 + \cdots + u_d y^{d-1}}{v}$$
for $u_1, \dots, u_d, v \in k[x]$ such that $\gcd(u_1, \dots, u_d, v) = 1$. Compute the splitting field of
$$\operatorname{res}_x(\operatorname{res}_y(u_1 + \cdots + u_d y^{d-1} - tDv, m), v) \in k[t],$$
say $\gamma_1, \dots, \gamma_r \in \bar{k}$ generate this field as a vector space over $\mathbb{Q}$.
6. For each γ_i, construct an ideal $\mathfrak{a}_i \trianglelefteq \mathcal{O}_{k[x]}$ that encodes the finite places and multiplicities of the singularities a potential logand p_i with coefficient γ_i has to have.
7. For each $\mathfrak{a}_i$, determine whether there is a positive integer n_i such that $\mathfrak{a}_i^{n_i}$ is a principal ideal. If so, say $\mathfrak{a}_i^{n_i} = \langle p_i \rangle$ for some $p_i \in \mathcal{O}_{k[x]}$, then $\frac{\gamma_i}{n_i} \log(p_i)$ is the desired contribution to the logarithmic part of the integral. If no n_i exists, then return "not integrable"; the integral is not elementary in this case.
8. After undoing the substitution of Step 1, return
$$g + \frac{\gamma_1}{n_1} \log(p_1) + \cdots + \frac{\gamma_r}{n_r} \log(p_r).$$

The computation of an integral basis in Step 2 is easy if y is a radical over $k(x)$, i.e., if $m = ay^d - b$ for some $a, b \in k[x]$ [3]. Otherwise, the computation is not trivial, but algorithms for this purpose are known [13, 14]. The integral basis is needed for the Hermite reduction and also in Step 7. For a modified Hermite reduction [2, 3], however, an integral basis is not needed in the first place. The standard basis $1, y, \dots, y^{d-1}$ can be used instead as a first "approximation" to the integral basis. The modified Hermite reduction then returns, in addition to g and h, a refined basis $\omega_1, \dots, \omega_d$, which is just as close to an integral basis as was necessary for finding the desired g and h.

For details about steps 5, 6, and 7, we refer to Trager's thesis [13]. These are parts that are often only partially implemented (if at all), and for these we will give a simple alternative below.

3. CZICHOWSKI'S OBSERVATION

For the integration of rational functions $f \in k(x)$, Czichowski [7] observed that the logarithmic part can be read off directly from a certain Gröbner basis. Let $f = u/v$ with $u, v \in k[x]$ such that $\gcd(u, v) = \gcd(v, Dv) = 1$ and $\deg u < \deg v$, and consider the Gröbner basis

$$G = \{g_0, g_1, \dots, g_n\} \subseteq k[x, t]$$

of the ideal $\langle v, u - tDv \rangle \trianglelefteq k[x, t]$ with respect to the lexicographic order eliminating x. Denote by $c_i := \operatorname{cont}_x(g_i) \in k[t]$ and $p_i := \operatorname{pp}_x(g_i) \in k[x, t]$ the contents and the primitive parts of the g_i with respect to x $(i = 0, \dots, n)$, and suppose the g_i are sorted according to ascending leading terms. Then:

- $p_0 = c_n = 1$
- c_0 is the square free part of the Rothstein-Trager resultant $\operatorname{res}_x(v, u - tDv) \in k[t]$.
- $c_i \mid c_{i-1}$ $(i = 1, \dots, n)$
- for $q_i := c_{i-1}/c_i \in k[t]$ $(i = 1, \dots, n)$ we have
$$\int \frac{u}{v} = \sum_{i=1}^{n} \sum_{\gamma : q_i(\gamma) = 0} \gamma \log(p_i(x, \gamma)).$$

Example 1. *For $u = x^3 + 9x^2 - 18x + 9$ and $v = x^4 - 17x^2 - 18$ we have*

$$G = \{(2t-1)(8t^2 - 9), (2t-1)(x - 4t), 7x^2 - 152t^2 + 45\}$$

and therefore

$$\begin{aligned}
&\int \frac{x^3 + 9x^2 - 18x + 9}{x^4 - 17x^2 - 18} dx \\
&= \sum_{\gamma : 8\gamma^2 - 9 = 0} \gamma \log(x - 4\gamma) + \sum_{\gamma : 2\gamma - 1 = 0} \gamma \log(7x^2 - 152\gamma^2 + 45) \\
&= \frac{3}{\sqrt{8}} \log(x - 3\sqrt{2}) - \frac{3}{\sqrt{8}} \log(x + 3\sqrt{2}) + \frac{1}{2} \log(7x^2 + 7)
\end{aligned}$$

The following facts are immediate consequences of Czichowski's observation.

- If c_0 is irreducible then $G = \{c_0, p_1\}$ with
$$p_1(x, \gamma) = \gcd(u - \gamma Dv, v)$$
where γ is a root of c_0.

- More generally, if q is an irreducible factor of c_0, then the Gröbner basis of $\langle q, u - tDv, v\rangle \in k[x, t]$ with respect to an order eliminating x will have the form $\{q, p\}$ with $p(x, \gamma) = \gcd(u - \gamma Dv, v)$ where γ is a root of q.

- Consequently, if $q^{(1)}, \ldots, q^{(m)} \in k[t]$ are all the irreducible factors of c_0 and if $p^{(i)} \in k[x, t]$ is the corresponding element in the Gröbner basis of $\langle q^{(i)}, u - tDv, v\rangle$ $(i = 1, \ldots, m)$, then
$$\int \frac{u}{v} = \sum_{i=1}^{m} \sum_{\gamma:q^{(i)}(\gamma)=0} \gamma \log(p^{(i)}(x, \gamma)).$$

4. THE ALGEBRAIC CASE

Our goal is to extract the logarithmic part of an algebraic function integral from a Gröbner basis, similar as Czichowski does it for a rational function integral. Suppose, to this end, that for a given $f \in K = k(x)[y]/\langle m\rangle$ we have executed Trager's algorithm up to, and including, step 4 (cf. Section 2). The remaining integrand is then of the form u/v for some $u \in k[x, y]$ and $v \in k[x]$. We may assume that u/v has at least a double root at infinity.

In the rational case, we can identify the ideals $\langle q, v, u - tDv\rangle = \langle q, p\rangle \trianglelefteq k[x, t]$ for irreducible $q \in k[t]$ with principal ideals $\langle \bar{p}\rangle \trianglelefteq k[x, t]/\langle q\rangle$. These ideals give rise to a contribution $\sum_{\gamma:q(\gamma)=0} \gamma \log(p(x, \gamma))$ to the logarithmic part. Likewise, if in the algebraic case we can find some $p \in k[x, y]$ with $\langle q, m, v, u - tDv\rangle = \langle q, m, p\rangle \trianglelefteq k[x, y, t]$, then this ideal can be identified with the principal ideal $\langle \bar{p}\rangle \trianglelefteq k[x, y, t]/\langle q, m\rangle$. It gives rise to a contribution $\sum_{\gamma:q(\gamma)=0} \gamma \log(p(x, y, \gamma))$ to the logarithmic part of the integral.

Example 2. *For $u = y$ and $v = x^4 + 1$ with $m = y^2 - (x^2 + 1)$ and using $q = 128t^4 + 16t^2 + 1$ we find*
$$\{128t^4 + 16t^2 + 1, y^2 - 16t^2 - 2, x - 32t^3y\}$$
as the Gröbner basis of $\langle q, m, v, u - tDv\rangle$ with respect to lexicographic order $x > y > t$. It is easily checked that in fact
$$\langle v, u - tDv, m\rangle = \langle q, m, x - 32t^3y\rangle$$
(both ideals have the same Gröbner basis). Indeed,
$$\int \frac{\sqrt{x^2+1}}{x^4+1} dx = \sum_{\gamma:128\gamma^4+16\gamma^2+1=0} \gamma \log(x - 32\gamma^3 \sqrt{x^2+1}).$$

The Gröbner basis of $\langle m, v, u - tDv\rangle \trianglelefteq k[x, y, t]$ with respect to an order eliminating x and y will contain one univariate polynomial in t. For an irreducible polynomial $q \in k[t]$, the ideal $\langle q, m, v, u - tDv\rangle$ will be nontrivial iff q is a divisor of this polynomial. This restricts q to finitely many candidates that can be considered one after the other.

Example 3. *For $u = \mathrm{i}x + y$ and $v = x^4 + 1$ with $m = y^2 - (x^2 + 1)$ we find*
$$\begin{aligned}&\{(16t^2 - 8t + (2 - \mathrm{i}))(16t^2 + 8t + (2 + \mathrm{i})),\\ &128\mathrm{i}t^3 + 16t^2 - 16\mathrm{i}t + 7y^2 - 1,\\ &192yt^3 - 80\mathrm{i}yt^2 + 4yt + 14x - 9\mathrm{i}y\}\end{aligned}$$
as the Gröbner basis of $\langle m, v, u - tDv\rangle \trianglelefteq k[x, y, t]$. For $q = 16t^2 - 8t + (2 - \mathrm{i})$, we have
$$\langle q, v, u - tDv\rangle = \langle q, m, 2x + (1 - \mathrm{i})(4t - 1)y\rangle,$$
while for $q = 16t^2 + 8t + (2 + \mathrm{i})$, we have
$$\langle q, v, u - tDv\rangle = \langle q, m, 2x + (1 + \mathrm{i})(4t + 1)y\rangle.$$
Indeed,
$$\begin{aligned}&\int \frac{\mathrm{i}x + \sqrt{x^2+1}}{x^4+1} dx\\ &= \sum_{\gamma:16\gamma^2-8\gamma+(2-\mathrm{i})=0} \gamma \log(2x + (1 - \mathrm{i})(4\gamma - 1)\sqrt{x^2+1})\\ &\quad + \sum_{\gamma:16\gamma^2+8\gamma+(2+\mathrm{i})=0} \gamma \log(2x + (1 + \mathrm{i})(4\gamma + 1)\sqrt{x^2+1}).\end{aligned}$$

For a given ideal $\langle q, m, v, u - tDv\rangle \trianglelefteq k[x, y, t]$, there may or may not exist a $p \in k[x, y, t]$ such that $\langle q, m, v, u - tDv\rangle = \langle q, m, p\rangle$. It is not trivial to decide whether such a "principal generator" p exists, and even if it is known that there is one it is not obvious how to find it. But it turns out that often a principal generator p will belong to the Gröbner basis of $\langle q, m, v, u - tDv\rangle$ with respect to a block order $[x, y] > [t]$ which orders the block $[x, y]$ by a degree order, say degrevlex. We therefore suggest to compute this Gröbner basis and consider its elements as candidates for p. For each candidate p, we can simply check whether the ideal $\langle q, m, p\rangle$ has the same Gröbner basis, and if so, we have found a contribution to the logarithmic part. We know of no convincing algebraic justification of this heuristic, but we can assert that it does succeed in many cases (see the next section).

In order to give rise to a contribution to the logarithmic part of an integral, it is not necessary that $\mathfrak{a} := \langle \bar{v}, \bar{u} - tD\bar{v}\rangle \trianglelefteq k[x, y, t]/\langle q, m\rangle$ itself is a principal ideal. It suffices that some power of it is. For if $n \in \mathbb{N}$ and $p \in k[x, y, t]$ are such that $\mathfrak{a}^n = \langle \bar{p}\rangle$, this would give a contribution $\sum_{\gamma:q(\gamma)=0} \frac{\gamma}{n} \log(p(x, y, \gamma))$ to the logarithmic part. Therefore, if we fail to find a p with $\langle q, m, v, u - tDv\rangle = \langle q, m, p\rangle$ we check $\langle q, m\rangle + \langle v, u - tDv\rangle^2$, $\langle q, m\rangle + \langle v, u - tDv\rangle^3$, and so on. One of the main difficulties in the construction of a complete integration procedure is finding a bound on the exponent n above which it can be asserted that the integral is not elementary. But as our approach is heuristic anyway, we need not bother about finding a rigorous bound but choose some fixed number, say 12, and give up if we exceed this power without having found anything. (If the divisor is over the rationals, the choice of 12 is in fact a rigorous bound, as pointed out by one of the referees.)

Example 4. *For $u = y$ and $v = x^3 + 1$ with $m = y^2 - (x^2 + 1)$ we find*
$$\begin{aligned}&\{(9t^2 - 2)(9t^2 + 1), y + 9yt^2 - 27t^3 - 3t,\\ &x + 3ty - 1, y^2 + 3yt - 9t^2 - 2\}\end{aligned}$$
as the Gröbner basis of $\langle m, v, u - tDv\rangle$. The Gröbner basis of $\langle 9t^2 + 1, m, v, u - tDv\rangle$ is
$$\{9t^2 + 1, x + 3ty - 1, y^2 + 3ty - 1\},$$
and it turns out that
$$\langle 9t^2 + 1, m, v, u - tDv\rangle = \langle 9t^2 + 1, m, x + 3ty - 1\rangle.$$

The Gröbner basis of $\langle 9t^2-2, m, v, u-tDv\rangle$ is

$$\{9t^2-2, y-3t, x+1\},$$

but

$$\langle 9t^2-2, m, v, u-tDv\rangle \neq \langle 9t^2-2, m, y-3t\rangle \quad \text{and}$$
$$\langle 9t^2-2, m, v, u-tDv\rangle \neq \langle 9t^2-2, m, x+1\rangle,$$

so there seems to be no principal generator for $n=1$. The Gröbner basis for $\langle 9t^2-2, m\rangle + \langle v, u-tDv\rangle^2$ is

$$\{9t^2-2, x+3ty-1, y^2-6ty+2\},$$

and it turns out that

$$\langle 9t^2-2, m\rangle + \langle v, u-tDv\rangle^2 = \langle 9t^2-2, m, x+3ty-1\rangle.$$

Indeed,

$$\begin{aligned}\int \frac{\sqrt{x^2+1}}{x^3+1}dx &= \sum_{\gamma:9\gamma^2+1=0} \gamma \log(x-1+3\gamma\sqrt{x^2+1})\\ &\quad + \sum_{\gamma:9\gamma^2-2=0} \frac{\gamma}{2}\log(x-1+3\gamma\sqrt{x^2+1}).\end{aligned}$$

We have implicitly assumed so far that the integral closure $\mathcal{O}_{k[x]}$ of $k[x]$ in K is just $k[x,y]$, but it may actually be more that this. Just using $k[x,y]$ then in a computation may cause some contributions to the logarithmic part to be overlooked.

Example 5. *Consider the integral*

$$\int \frac{1}{\big(4x+4-\sqrt[3]{x^2(x+1)}\big)\sqrt[3]{x^2(x+1)}}dx.$$

Let

$$\begin{aligned}u &= 4(x+1)x^2 + x^2y + 16(x+1)y^2,\\ v &= x^2(x+1)(7x+8)(9x+8)\end{aligned}$$

and $m = y^3 - x^2(x+1)$. Then the integrand can be written $u(x, \sqrt[3]{x^2(x+1)})/v(x)$.

Note that despite of the double factor x in v, the integrand has simple poles only. It turns out that no principal generators can be found in the Gröbner bases of ideals $\langle q, m\rangle + \langle v, u-tDv\rangle^n \trianglelefteq k[x,y,t]$ for $n=1,\dots,30$ and any irreducible $q \in k[t]$.

By adjoining new formal elements to $k[x,y]$ and stipulating appropriate relations between them, the whole integral closure $\mathcal{O}_{k[x]}$ can be taken into account during the computation.

Example 6. *Continuing Example 5, an integral basis for $K = k(x)[y]/\langle m\rangle$ is given by $(1, y, y^2/x)$. Introducing a new indeterminate z referring to y^2/x, the integrand may be written r/s where*

$$\begin{aligned}r &= 4(x+1)x + xy + 16(x+1)z\\ s &= x(x+1)(7x+8)(9x+8).\end{aligned}$$

Let $\mathfrak{m} = \langle y^3 - x^2(x+1), z^3 - x(x+1)^2, zx-y^2\rangle \trianglelefteq k[x,y,z,t]$. Then $\mathfrak{m} + \langle s, r-tDs\rangle \cap k[t] = \langle (4t-3)(4t+3)\rangle$ and we have

$$\begin{aligned}&\langle 4t-3\rangle + \mathfrak{m} + \langle s, r-tDs\rangle^3\\ &\quad = \langle 4t-3\rangle + \mathfrak{m} + \langle 8+9x+6y+12z\rangle,\\ &\langle 4t+3\rangle + \mathfrak{m} + \langle s, r-tDs\rangle^3\\ &\quad = \langle 4t+3\rangle + \mathfrak{m} + \langle 8+7x+6y-12z\rangle.\end{aligned}$$

The principal generators were found in the Gröbner basis of the ideals on the left with respect to a block order $[x,y,z] > [t]$ ordering $[x,y,z]$ by degrevlex. Indeed,

$$\begin{aligned}&\int \frac{1}{\big(4x+4-\sqrt[3]{x^2(x+1)}\big)\sqrt[3]{x^2(x+1)}}dx\\ &= \tfrac{1}{4}\log\big(8+9x+6\sqrt[3]{x^2(x+1)}+12\sqrt[3]{x(x+1)^2}\big)\\ &\quad - \tfrac{1}{4}\log\big(8+7x+6\sqrt[3]{x^2(x+1)}-12\sqrt[3]{x(x+1)^2}\big).\end{aligned}$$

When y is not a radical over $k(x)$ the computation of an integral basis is a difficult and expensive task. Lazy Hermite reduction [2, 3] does not require an integral basis as input but begins the computation with $(1, y, \dots, y^{d-1})$ and refines this basis as much as necessary to complete the reduction. The approximation to the integral basis thus obtained may be used in search of the logarithmic part, as described above. No full integral basis is necessary to proceed this way, but of course the finer the approximation the more likely a logarithmic part will be found.

Additional formal elements can also be used if the integrand involves several different algebraic functions: instead of computing with a primitive element, each algebraic function arising in the integrand may be represented by an individual indeterminate.

Example 7. *For*

$$\begin{aligned}u &= (x-1)(4x+1)y + (x+1)(4x-3)z - 2(2x-1)yz,\\ v &= (x-1)(x+1)(4x-5)\end{aligned}$$

with $\mathfrak{m} = \langle y^2-(x+1), z^2-(x-1)\rangle$ we find first

$$\mathfrak{m} + \langle v, u-tDv\rangle \cap k[t] = \langle t^2(t-1)(t+3)\rangle$$

and then

$$\begin{aligned}&\langle t-1\rangle + \mathfrak{m} + \langle v, u-tDv\rangle^2 = \langle t-1\rangle + \mathfrak{m}\\ &\quad + \langle 16-20x-(9-12x)y+(13+4x)z-12yz\rangle\\ &\langle t+3\rangle + \mathfrak{m} + \langle v, u-tDv\rangle^2 = \langle t+3\rangle + \mathfrak{m} + \langle 4+3y-z\rangle.\end{aligned}$$

Indeed,

$$\begin{aligned}&\int \Big((x+1)(4x-3)\sqrt{x-1} + (x-1)(4x+1)\sqrt{x+1}\\ &\quad - 2(2x-1)\sqrt{x-1}\sqrt{x+1}\Big)\Big/(x-1)(x+1)(4x-5)\,dx\\ &= \tfrac{1}{2}\log(16-20x-(9-12x)\sqrt{x+1}\\ &\qquad + (13+4x)\sqrt{x-1} - 12\sqrt{x+1}\sqrt{x-1})\\ &\quad + \tfrac{1}{2}\log(4+3\sqrt{x+1}-\sqrt{x-1}).\end{aligned}$$

Alternatively, we may rephrase the integrand in terms of a primitive element. Let q be an indeterminate representing $\sqrt{x+1}+\sqrt{x-1}$ and $m = q^4 - 4xq^2 + 4$. Then the integrand may be written as $r/s \in k(x)[q]/\langle m\rangle$ with

$$\begin{aligned}r &= (2x-1)q^3 - 2(2x-1)q^2 + 2(x-2)q + 4x(2x-1),\\ s &= 4(x-1)(x+1)(4x-5).\end{aligned}$$

We find $\langle m, s, r-tDs\rangle \cap k[t] = \langle t^2(2t-1)(2t+3)\rangle$ and

$$\begin{aligned}\langle 2t-1, m, s, r-tDs\rangle &= \langle 2t-1, m, q^3-2q^2-q+2\rangle\\ \langle 2t+3, m, s, r-tDs\rangle &= \langle 2t+3, m, q+2\rangle.\end{aligned}$$

Substituting $\sqrt{x+1}+\sqrt{x-1}$ *for* q *leads to the alternative closed form*

$$\begin{aligned}&\log(2-4x-(3-4x)\sqrt{x+1}\\&\quad+(1+4x)\sqrt{x-1}-4\sqrt{x+1}\sqrt{x-1})\\&+\log(\sqrt{x+1}+\sqrt{x-1}+2).\end{aligned}$$

Summarizing, our procedure for finding the logarithmic part of an integral over an algebraic function u/v is as follows.

```
G := GröbnerBasis({v, u − tDv} ∪ M); int = 0
for all irreducible factors q of min G do
    A := {1}
    for n from 1 to 12 do
        A := GröbnerBasis((A · G) ∪ {q} ∪ M)
        for all p in A do
            if A = GröbnerBasis({q, p} ∪ M) then
                int := int + Σ_{γ:q(γ)=0} (γ/n) log(p(x, y, γ))
                next q
return int
```

Notational remarks.

- GröbnerBasis is meant to compute a Gröbner basis in $k[x,y,t]$ with respect to a block order $[x,y]>[t]$ using the degrevlex order for breaking ties in the block $[x,y]$.
- M is meant to contain the relations among the generators, typically $M=\{m\}$, but there may be more relations (and more indeterminates) in the case of nontrivial integral closures or multiple algebraic functions in the integrand.
- $\min G$ is meant to refer to the element of G with the lowest leading term. This is the unique element of G involving t but not x or y.
- If the polynomial $\min G$ has repeated factors, it is understood that the outer loop takes into account the multiplicities by repeating the body of a repeated factor an according number of times (cf. Section 6.3).
- $A\cdot G$ is meant to refer to the ideal product, i.e., if $A=\{a_1,\dots,a_r\}$ and $G=\{g_1,\dots,g_t\}$ then $A\cdot G:=\{a_ig_j : i=1,\dots,r,\ j=1,\dots,t\}$.
- The statement "next q" is meant to break the two inner loops and proceed with the next iteration of the outermost loop.

It is an easy matter to implement this procedure in an actual computer algebra system. In the appendix, we give a sample code for Mathematica.

5. EXPERIMENTS

Though Norman and Davenport may find it antiquated, we will judge our method by the proportion of some set of (known integrable) examples it can cope with. We compare its performance to the built-in integrators of Maple 10, Mathematica 5.2, and Axiom 3.9, knowing that any comparison of this sort is unfair to some extent.

We chose four different algebraic functions and randomly generated 1000 logarithmic expressions for each, which we differentiated to obtain candidate integrands in $\mathbb{Q}(x,y)$. A typical integrand from our collection is

$$\begin{aligned}&\Big(-214632x^7-90880x^5-39020x^4-238623x^3\\&\quad+4160x+7378+(-128496x^4-143517x^3+1300x\\&\quad+12614)\sqrt{x^4+1}\Big)\Big/\\&\Big(17424x^8+7040x^6+6138x^5+40414x^4\\&\quad+7040x^2+6138x+22990+(9504x^5+17061x^4\\&\quad+2200x^2+10494x+17666)\sqrt{x^4+1}\Big).\end{aligned}$$

The integrands were generated such as to admit a closed form in terms of a linear combination of three logarithms with logands of total degree 1 with respect to x and y. Typically, two of the logarithms involve constants from a quadratic extension of $\mathbb{Q}$. For example, the integral over the algebraic function above admits the closed form

$$\begin{aligned}&\tfrac{1}{44}(13+\mathrm{i}\sqrt{359})\log((9-\mathrm{i}\sqrt{359})x+11-33\sqrt{x^4+1})\\&+\tfrac{1}{44}(13-\mathrm{i}\sqrt{359})\log((9+\mathrm{i}\sqrt{359})x+11-33\sqrt{x^4+1})\\&-\tfrac{27}{4}\log(5-16\sqrt{x^4+1}).\end{aligned}$$

If an integrator did not deliver a logarithmic part in this form, but with higher degree logands, or with arctan expressions, we also accepted these as correct answers. The integrands were chosen at random, but simple rational functions in x produced occasionally by the test case generator were discarded.

We have implemented an integrator for algebraic functions in Mathematica that executes steps 1–4 of Trager's algorithm (as described in Section 2), followed by an implementation of our procedure for finding the logarithmic part. Singular was used as Gröbner basis engine. Our code for finding the logarithmic part is given in the appendix. We then presented the test case integrands to Maple (Mpl), Mathematica (Mma), Axiom (Ax) and our procedure (P). Each integrator got 180 seconds per integrand to perform the integration. The results are summarized below.

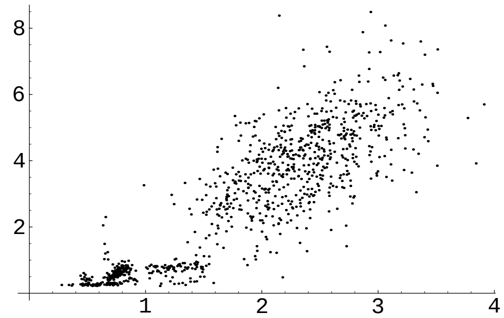
5.1 Integrals involving $\sqrt{x^2+1}$

	P	Ax	Mpl	Mma
Success	100.0%	100.0%	100.0%/40.5%	11.6%
Timeout	—	—	— /59.5%	—
Failure	—	—	— / —	88.4%

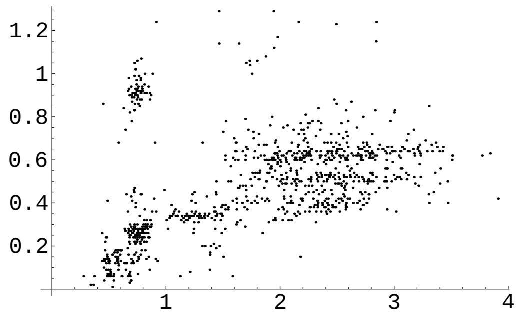
Maple's standard integrator does not appear to search for elementary closed forms of algebraic integrals by default. The first percentage shown for Maple refers to the default integration procedure, the second percentage refers to the case where an elementary closed form was explicitely requested.

The integrals that Mathematica could not do were returned unevaluated. For the other three integrators a runtime comparison might be interesting. One might think that iterated computation of Gröbner basis in P will make this integrator much slower than his competitors. But this is not the case.

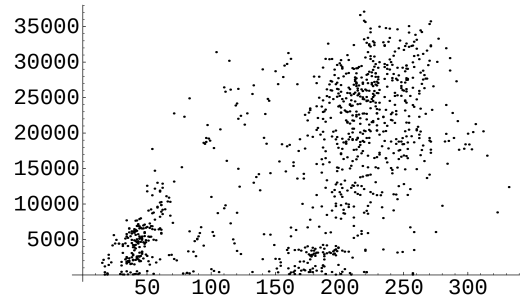
In the following diagram, a dot at the point (x,y) represents an integral that was evaluated by P in x seconds and by Ax in y seconds. We see that the runtime is somewhat correlated, with an advantage for Ax on fast examples and a slight advantage for P on slower examples. The center is at $(1.8, 2.8)$.



The corresponding figure for P vs. Mpl is shown next.



Mpl is much faster than P, but it should be noted that the results produced by Mpl are a whole lot messier than those of P or Ax. In the next figure, a point (x, y) corresponds to an integral for which P produced a logarithmic part of length x and Mpl produced a logarithmic part of length y. (For the "length" of a result, we simply counted the number of characters in Mathematica and Maple syntax, respectively.)



5.2 Integrals involving $\sqrt{x^3+1}$

	P	Ax	Mpl	Mma
Success	99.6%	24.3%	.2%/22.0%	2.2%
Timeout	.3%	6.6%	— /78.0%	.1%
Failure	.1%	69.1%	99.8%/ —	97.7%

The Failure line covers different behaviors. For Ax, a failure was counted when the integrator aborted with an error (e.g. "implementation incomplete"). Mpl never failed in this way, but always returned non-elementary closed forms involving elliptic functions. Mma returned the integrals unevaluated. Ways how P can fail are discussed in the next section.

The percentage of timeouts and failures increases more rapidly for Ax, Mpl and Mma than for P. Runtime comparisons for the cases in which at least two integrators succeed lead to less expressive diagrams, which are therefore not shown here.

5.3 Integrals involving $\sqrt{x^4+1}$

	P	Ax	Mpl	Mma
Success	99.5%	12.3%	1.8%/24.7%	1.8%
Timeout	.1%	16.4%	— /75.3%	.3%
Failure	.4%	71.3%	98.2%/ —	97.9%

Concerning the Failure line, the same remarks apply as for 5.2. The main difference to the previous case is the increase in timeouts for Ax. Also P slows down for this case: the runtime median here (counting successful cases as well as timeouts and failures) is 68.0 sec, compared to 16.0 sec in the previous case.

5.4 Integrals involving $\sqrt[3]{x^2+1}$

	P	Ax	Mpl	Mma
Success	67.9%	4.2%	— /18.1%	2.4%
Timeout	29.2%	95.8%	— /81.9%	—
Failure	2.9%	—	100.0%/—	97.6%

Failure here means for both Mpl and Mma that the integral was returned unevaluated.

Although the success rate of P drops considerably in this case, it is still superior to Ax, Mpl, and Mma. The substitution carried out in the first step of the integrator might lead to an algebraic function field with a nontrivial integral basis. The new indeterminates introduced in this case may be responsible for slowing down the Gröbner basis computations and raising the timeout percentage for P compared to the previous test cases.

6. REASONS FOR FAILURE

Our procedure will never detect that an algebraic function cannot be integrated in terms of elementary functions. Even if an elementary integral exists, there is no guarantee that our procedure will find it. It can fail for the following reasons.

6.1 The Principal Generator is Hidden

We search the principal generator p of an ideal among the elements of a Gröbner basis. Though the principal generator can often be found there, this is not always the case.

Example 8. *In Example 4 we found that the Gröbner basis of* $\mathfrak{a} := \langle 9t^2-2, m, v, u-tDv\rangle$ *is*

$$\{9t^2-2, y-3t, x+1\},$$

and since

$$\langle 9t^2-2, m, v, u-tDv\rangle \neq \langle 9t^2-2, m, y-3t\rangle \quad \textit{and}$$
$$\langle 9t^2-2, m, v, u-tDv\rangle \neq \langle 9t^2-2, m, x+1\rangle,$$

we concluded that there is no $p \in \mathfrak{a}$ *such that* $\mathfrak{a} = \langle 9t^2-2, m, p\rangle$. *However,* $p = x+3ty-1 \in \mathfrak{a}$ *and we do have* $\mathfrak{a} = \langle 9t^2-2, m, p\rangle$.

It does not harm if we overlook a principal generator in some ideal $\mathfrak{a}$ if we find a generator in one of its powers $\mathfrak{a}^n$. In the example above, the principal generator is found in the next step, as a member of the Gröbner basis for

$$\langle 9t^2-2, m\rangle + \langle v, u-tDv\rangle^2$$

(cf. Ex. 4). In our experiments, we have never observed a failure because the principal generator could not be found in any of the powers of an ideal.

6.2 The Power Bound is Exceeded

An obvious source of failures is that the fixed bound on the powers of $\langle v, u - tDv\rangle$ that are inspected is exceeded without that a principal generator was found. Of course, there could still be some principal ideal beyond the bound. In our experiments, we have used the bound 12 but this bound was never reached.

6.3 The Ideal is not Radical

For rational function integrands $u/v \in k(x)$, Czichowski has shown that the ideal $\langle v, u - tDv\rangle$ is always radical, in particular the polynomial in its Gröbner basis which is free of x will always be square free. For algebraic function integrands, this need no longer be true. Branch places over roots of the denominator may cause multiple factors.

In many of these cases, the correct result is obtained when the contributed logarithms are multiplied with the multiplicity of the corresponding factors of the univariate polynomial in t.

EXAMPLE 9. *For* $u = 1 + y$ *and* $v = (x^2 + 1)(x + 1)$ *with* $m = y^2 - (x^2 + 1)$ *we find*

$$\begin{aligned}\{&(4t^2 + 4t - 1)(8t^2 - 4t + 1)^2,\\ &-2560t^5 - 640t^4 + 1776t^3 - 1136t^2 + 306t + 9y - 41,\\ &1792t^5 + 576t^4 - 1184t^3 + 744t^2 - 160t + 9x + 19\}\end{aligned}$$

as the Gröbner basis of $\langle m, v, u - tDv\rangle \trianglelefteq k[x, y, t]$.

For $q = 4t^2 + 4t - 1$ *we find*

$$\langle q, m\rangle + \langle v, u - tDv\rangle^2 = \langle q, m, 1 - x - (1 + 2t)y\rangle.$$

For $q = 8t^2 - 4t + 1$ *we find*

$$\langle q, m\rangle + \langle v, u - tDv\rangle^2 = \langle q, m, 4t - 1 + x\rangle.$$

As $8t^2 - 4t + 1$ *is a double factor, we count the contribution from the latter ideal twice. Indeed,*

$$\begin{aligned}\int \frac{x + \sqrt{x^2 + 1}}{(x^2 + 1)(x + 1)}dx &= 2 \sum_{\gamma:8\gamma^2-4\gamma+1=0} \frac{\gamma}{2}\log(4\gamma - 1 + x)\\ &+ \sum_{\gamma:4\gamma^2+4\gamma-1=0} \frac{\gamma}{2}\log(1 - x - (1 + 2\gamma)\sqrt{x^2 + 1}).\end{aligned}$$

Handling multiple factors this way might, however, not be correct. All the failures reported in Section 5 for our integrator are of this kind. The logarithmic expression *int* returned by the integrator for an integrand u/v may then be viewed as a partial closed form, leaving $\frac{u}{v} - D(int)$ as unintegrated remainder. It usually pays off to apply the integrator to this remainder once again, for it may well succeed in integrating it in a second attempt.

EXAMPLE 10. *For* $u = 2x^3 + 6x^2 - 7x - 7 - (x - 1)(3x + 1)y$ *and* $v = (x^2 - 1)x(x^2 - x - 1)$ *with* $m = y^2 - (x + 1)$ *we find that the Gröbner basis of* $\langle m, v, u - tDv\rangle$ *contains the polynomial* $(t - 3)(t - 2)^2t(t + 6)(t + 8)$. *We have*

$$\begin{aligned}\langle t - 3, m, v, u - tDv\rangle &= \langle t - 3, m, x - 1\rangle,\\ \langle t - 2, m, v, u - tDv\rangle &= \langle t - 2, m, 1 + x + xy\rangle,\\ \langle t + 6, m, v, u - tDv\rangle &= \langle t + 6, m, y - 1\rangle,\\ \langle t + 8, m, v, u - tDv\rangle &= \langle t + 8, m, y + 1\rangle.\end{aligned}$$

However,

$$\begin{aligned}\frac{u}{v} - D\Big(&3\log(x - 1) - 6\log(y - 1)\\ &- 8\log(1 + y) + 4\log(1 + x + xy)\Big)\\ &= -\frac{2x^2 + x - 1 - (x + 2)y}{(x + 1)(x^2 - x - 1)} \neq 0.\end{aligned}$$

Applying the procedure to this nonzero remainder gives the result $-2\log(x + y)$. *Indeed,*

$$\begin{aligned}&\int \frac{2x^3 + 6x^2 - 7x - 7 - (x - 1)(3x + 1)\sqrt{x + 1}}{(x^2 - 1)x(x^2 - x - 1)}dx\\ &= 3\log(x - 1) - 6\log(\sqrt{x + 1} - 1) - 8\log(1 + \sqrt{x + 1})\\ &\quad + 4\log(1 + x + x\sqrt{x + 1}) - 2\log(x + \sqrt{x + 1}).\end{aligned}$$

In the examples in Section 5.2 and 5.3, when the integrator is applied twice, there are no failures any more. In the examples in Section 5.4, when the integrator is applied twice the failure rate drops from 2.9% to .1%. When it is applied once more there are no failures. In 26 of the 29 observed failures in 5.4, the "unintegrable" remainder was just a rational function in x, in 22 of these cases it was just a constant multiple of $1/x$.

7. CONCLUSION

We have described a procedure for finding the logarithmic part of an integral over an algebraic function. Our procedure is simple and efficient, and, although there is no guarantee, it finds the correct results in a great many cases. It has come off well in a comparison with the built-in integrators of Axiom, Maple, and Mathematica.

We do not see how our procedure could be turned into a complete algorithm. In particular, we have no convincing argument that would justify our observation that principal generators often show up in the Gröbner bases we compute. Obviously, any step towards an explanation of this phenomenon would also be a step towards a new complete algorithm, and therefore be highly interesting.

8. REFERENCES

[1] Manuel Bronstein. On the integration of elementary functions. *Journal of Symbolic Computation*, 9:117–173, 1990.

[2] Manuel Bronstein. The lazy Hermite reduction. Technical Report 3562, INRIA, 1998.

[3] Manuel Bronstein. Symbolic integration tutorial. ISSAC'98, 1998.

[4] Manuel Bronstein. Poorman's integrator. Maple package, available at http://www-sop.inria.fr/cafe/Manuel.Bronstein/pmint/, 2005.

[5] Manuel Bronstein. *Symbolic Integration I*, volume 1 of *Algorithms and Computation in Mathematics*. Springer, 2nd edition, 2005.

[6] Manuel Bronstein. Structure theorems for parallel integration. *Journal of Symbolic Computation*, 42(7):757–769, 2007.

[7] Günter Czichowski. A note on Gröbner bases and integration of rational functions. *Journal of Symbolic Computation*, 20(2):163–167, 1995.

[8] James H. Davenport. *On the Integration of Algebraic Functions*. Lecture Notes in Computer Science 102. Springer, 1981.

[9] Manuel Kauers and Viktor Levandovskyy. An interface between Singular and Mathematica. Technical Report 2006-29, SFB F013, Johannes Kepler Universität, 2006.

[10] A.C. Norman and J.H. Davenport. Symbolic integration – the dust settles? In *LNCS 72*, pages 398–407. Springer, 1979. Proceedings of ISSAC'79.

[11] Robert H. Risch. The problem of integration in finite terms. *Transactions of the American Mathematical Society*, 139:167–189, 1969.

[12] Robert H. Risch. The solution of the problem of integration in finite terms. *Bulletin of the American Mathematical Society*, 79:605–608, 1970.

[13] Barry M. Trager. *On the Integration of Algebraic Functions*. PhD thesis, MIT, 1984.

[14] Mark van Hoeij. An algorithm for computing an integral basis in an algebraic function field. *Journal of Symbolic Computation*, 18(4):353–363, 1994.

Appendix: Mathematica Code

The following ten lines form a Mathematica 5.2 implementation of the procedure described in this paper. Gröbner basis computations are done with Singular, accessed via an interface package [9]. The function to be integrated is num/den, where den is a squarefree polynomial in the integration variable. The list v contains the formal variables (typically, x and y, but perhaps additional indeterminates for encoding a nontrivial integral basis). The first element of v is taken as the integration variable. The list u is an ideal basis for the relations among the variables in v (typically, the minimal polynomial of y). An electronic version of this code can be found on the author's homepage.

```
LogarithmicPart[num_, den_, u_List, v_List] := Module[ {G, t, factors, f, r, i, F, p},
  G = SingularGroebner[Join[{den, num - t D[den, First[v]]}, u], v, {t}, MonomialOrder->"dp"];
  factors = DeleteCases[Rest[FactorList[First[G]]], {t, _}];
  F[p_, g_] := (t/#2*Log[#1])&@@PrincipalPower[g, Append[u, p], v, {t}];
  Plus@@Apply[r[f[#1], f[#2 F[#1, G]]]&, factors, {1}] /. {t -> #, f -> Function, r -> RootSum}];

PrincipalDivisor[gb_, u_, v__] := Module[{G = SingularGroebner[Join[gb, u], v, MonomialOrder->"dp"]},
  First[Append[Select[G, (SingularGroebner[Append[u, #], v, MonomialOrder->"dp"]===G)&, 1], 1]]];

PrincipalPower[gb_, u_, v__List, bound_Integer:12] := Module[ {id = gb, p, n = 1},
  While[ n < bound && (p = PrincipalDivisor[id, u, v])===1, n++;
    id = SingularGroebner[Join[SingularTimes[id, gb, v], u], v, MonomialOrder->"dp"]]; Return[{p, n}]]];
```

Example.

```
In[1]:= LogarithmicPart[y,x(x^8+1),{y^2-(x^8+1)},{x,y}]

         Log[-1 + y]   Log[1 + y]
Out[1]= ----------- - ----------
              8            8
```