Chapter 10
Risch Algorithm

We were taught at calculus classes that integration is an art, not a science (in contrast to differentiation—even a monkey can be trained to take derivatives). And we were taught wrong. The Risch algorithm (which is known for decades) allows one to find, in a finite number of steps, if a given indefinite integral can be taken in elementary functions, and if so, to calculate it. This algorithm has been constructed in works by an American mathematician Risch near 1970; many cases were not analyzed completely in these works and were later considered by other mathematicians. The algorithm is very complicated, and no computer algebra system implements it fully. Its implementation in *Mathematica* is rather complete, even with extensions to some classes of special functions, but details are not publicly known. Strictly speaking, it is not quite an algorithm, because it contains algorithmically unsolvable subproblems, such as finding out if a given combination of elementary functions vanishes. But in practice computer algebra systems are quite good in solving such problems. Here we shall consider, at a very elementary level, the main ideas of the Risch algorithm; see [16] for more details.

10.1 Rational Functions

We begin with a very simple case—integration of rational functions. Better methods than the partial fraction decomposition exist for this problem. And these methods can be generalized to much wider classes of integrands. Let's consider an integral

$$\int \frac{N(x)}{D(x)}\,\mathrm{d}x\,,$$

where $N(x)$ and $D(x)$ are polynomials. If $\deg N \geq \deg D$, we can divide with remainder; integration of a polynomial is trivial. Therefore we'll assume $\deg N < \deg D$. The integration result consists of a rational part and a logarithmic one:

A. Grozin, *Introduction to Mathematica® for Physicists*, Graduate Texts in Physics,
DOI 10.1007/978-3-319-00894-3_10,

$$\int \frac{N(x)}{D(x)}\,\mathrm{d}x = \frac{P(x)}{\hat{D}(x)} + \sum c_i \log\,(x-a_i)\;,$$

where a_i are the roots of the denominator $D(x)$, and c_i are constants. If

$$D(x) = \prod (x-a_i)^{d_i}\;,$$

then

$$\hat{D}(x) = \prod (x-a_i)^{d_i-1}\;.$$

Indeed, at $x \to a_i$ the rational part has a pole of the order $d_i - 1$; when differentiated, it becomes a pole of the order d_i, as needed. The numerator $P(x)$ is a polynomial of degree $\deg P < \deg \hat{D}$:

$$P(x) = \sum p_n x^n\,.$$

Substituting all these parts and differentiating, we can find the unknown coefficients c_i and p_n by solving a linear system.

For example, let's calculate

$$\int \frac{\mathrm{d}x}{x^2(x-1)} = \frac{p_0}{x} + c_1 \log(x) + c_2 \log(x-1)\,.$$

In[1] := Res = *p*[0]/*x* + *c*[1] ∗ Log[*x*] + *c*[2] ∗ Log[*x* − 1]
Out[1] = $c[2]\mathrm{Log}[-1+x] + c[1]\mathrm{Log}[x] + \frac{p[0]}{x}$
In[2] := Eq = Together[*x*^2 ∗ (*x* − 1) ∗ *D*[Res, *x*] − 1]
Out[2] = $-1 - xc[1] + x^2c[1] + x^2c[2] + p[0] - xp[0]$
In[3] := Eqs = Table[Coefficient[Eq, *x*, *n*] == 0, {*n*, 0, 2}]
Out[3] = $\{-1 + p[0] == 0, -c[1] - p[0] == 0, c[1] + c[2] == 0\}$
In[4] := Sol = Solve[Eqs, {*p*[0], *c*[1], *c*[2]}][[1]]
Out[4] = $\{p[0] \to 1, c[1] \to -1, c[2] \to 1\}$
In[5] := Res/.Sol
Out[5] = $\frac{1}{x} + \mathrm{Log}[-1+x] - \mathrm{Log}[x]$
In[6] := Clear[Res, Eq, Eqs, Sol]

10.2 Logarithmic Extension

Now we begin to extend the class of integrands and consider

$$\int \frac{N(x,y)}{D(x,y)}\,\mathrm{d}x\,,$$

where y depends on x (in a nonrational way). The extension is called algebraic if y is a root of a polynomial equation $p(x,y) = 0$. For example, $y = \sqrt[n]{p(x)/q(x)}$ is a root of the equation $q(x)y^n - p(x) = 0$. An algorithm for integration of expressions belonging to algebraic extensions has been constructed, but it requires an advanced mathematical apparatus [17], and we shall not discuss it here.

Extensions which are not algebraic are called transcendental. There are two important classes of such extensions. A logarithmic extension $y = \log r(x)$ (where $r(x)$ is a rational function) is characterized by the property $y' = r'/r$. An exponential extension $y = \exp r(x)$—by $y' = r'y$.

If an integral of an expression from a logarithmic extension with some $y = \log r(x)$ can be taken in elementary functions, it has the form

$$\int \frac{N(x,y)}{D(x,y)}\,\mathrm{d}x = \frac{P(x,y)}{\hat{D}(x,y)} + \sum c_i \log q_i\,,$$

where

$$D = \prod D_i^{d_i} \quad \Rightarrow \quad \hat{D} = \prod D_i^{d_i-1}\,,$$

q_i are the irreducible factors of all D_i, c_i are constants,

$$P(x,y) = \sum p_n(x)y^n = \sum p_{mn}x^m y^n$$

is a polynomial with unknown (so far) coefficients. Differentiating this general form of the result, putting everything over a common denominator, and equating coefficients of $x^m y^n$, we obtain a linear system for finding all unknown coefficients. If this system is incompatible, this means that the integral cannot be taken in elementary functions.

Example 1

We shall consider several examples. Let $y = \log x$ so that $y' = 1/x$. Let's calculate the integral

$$\int y\,\mathrm{d}x = p_2(x)y^2 + p_1(x)y + p_0(x)\,.$$

When differentiated, the degree in y reduces by 1, so that the result is quadratic in y (there is no denominator, and hence no q_i).

In[7] := y'[x_] := 1/x
In[8] := Res = Sum[p[n][x] * y[x]^n, {n, 0, 2}]
Out[8] = $p[0][x] + y[x]p[1][x] + y[x]^2 p[2][x]$
In[9] := Eq = D[Res, x] − y[x]
Out[9] = $-y[x] + \frac{p[1][x]}{x} + \frac{2y[x]p[2][x]}{x} + p[0]'[x] + y[x]p[1]'[x] + y[x]^2 p[2]'[x]$

In[10] := Eqs = Table[Coefficient[Eq, y[x], n] == 0, {n, 0, 2}]

$\text{Out}[10] = \left\{ \frac{p[1][x]}{x} + p[0]'[x] == 0, -1 + \frac{2p[2][x]}{x} + p[1]'[x] == 0, p[2]'[x] == 0 \right\}$

We see that $p_2(x)$ is a constant:

In[11] := p[2][x_] := p[2,0]; Eqs

$\text{Out}[11] = \left\{ \frac{p[1][x]}{x} + p[0]'[x] == 0, -1 + \frac{2p[2,0]}{x} + p[1]'[x] == 0, \text{True} \right\}$

Since $p_1(x)$ is a polynomial, the second equation can be satisfied only if $p_{20} = 0$:

In[12] := p[2,0] = 0; Eqs

$\text{Out}[12] = \left\{ \frac{p[1][x]}{x} + p[0]'[x] == 0, -1 + p[1]'[x] == 0, \text{True} \right\}$

The second equation gives

In[13] := p[1][x_] := x + p[1,0]

Now the first equation

In[14] := Eq1 = ExpandAll[Eqs[[1]]]

$\text{Out}[14] = 1 + \frac{p[1,0]}{x} + p[0]'[x] == 0$

gives $p_{10} = 0$:

In[15] := p[1,0] = 0; Eq1

$\text{Out}[15] = 1 + p[0]'[x] == 0$

Therefore $p_0(x) = -x$ (omitting the integration constant):

In[16] := p[0][x_] := −x; Eq1

$\text{Out}[16] = \text{True}$

In[17] := Res

$\text{Out}[17] = -x + xy[x]$

In[18] := Clear[Res, Eq, Eqs, Eq1, p]

We have derived the well-known result

$$\int \log(x)\mathrm{d}x = x\log(x) - x\,.$$

Consider the integrals

$$\int x\log(x)\mathrm{d}x\,, \qquad \int \log^2(x)\mathrm{d}x$$

in a similar way.

Example 2

Let's calculate the integral

$$\int \frac{y}{x}\,\mathrm{d}x = p_2(x)y^2 + p_1(x)y + p_0(x)\,.$$

Here $\hat{D} = 1$; it seems that there is a single q, namely x, but $\log(x) = y$, so that the logarithmic part contributes nothing new.

In[19] := Res = Sum[p[n][x] * y[x]^n, {n,0,2}];

In[20] := Eq = D[Res,x] − y[x]/x

$\mathrm{Out}[20] = -\frac{y[x]}{x} + \frac{p[1][x]}{x} + \frac{2y[x]p[2][x]}{x} + p[0]'[x] + y[x]p[1]'[x] + y[x]^2 p[2]'[x]$

In[21] := Eqs = Table[Coefficient[Eq,y[x],n] == 0, {n,0,2}]

$\mathrm{Out}[21] = \left\{\frac{p[1][x]}{x} + p[0]'[x] == 0, -\frac{1}{x} + \frac{2p[2][x]}{x} + p[1]'[x] == 0, p[2]'[x] == 0\right\}$

In[22] := p[2][x_] := p[2,0]; Eqs

$\mathrm{Out}[22] = \left\{\frac{p[1][x]}{x} + p[0]'[x] == 0, -\frac{1}{x} + \frac{2p[2,0]}{x} + p[1]'[x] == 0, \mathrm{True}\right\}$

From the second equation, $p_{20} = 1/2$; then p_1 is a constant:

In[23] := p[2,0] = 1/2; Eqs[[2]]

$\mathrm{Out}[23] = p[1]'[x] == 0$

In[24] := p[1][x_] = p[1,0]; Eqs

$\mathrm{Out}[24] = \left\{\frac{p[1,0]}{x} + p[0]'[x] == 0, \mathrm{True}, \mathrm{True}\right\}$

From the first equation, $p_{10} = 0$; then p_0 is a constant (which may be omitted):

In[25] := p[1,0] = 0; Eqs[[1]]

$\mathrm{Out}[25] = p[0]'[x] == 0$

In[26] := p[0][x_] := 0; Res

$\mathrm{Out}[26] = \frac{y[x]^2}{2}$

In[27] := Clear[Res, Eq, Eqs, p]

We have derived the well-known result

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

Example 3

Let's change the previous integral a little:

$$\int \frac{y}{x+1}\,\mathrm{d}x = p_2(x)y^2 + p_1(x)y + p_0(x) + c\log(x+1)\,.$$

In[28] := Res = Sum[p[n][x] * y[x]^n, {n,0,2}] + c * Log[x + 1]

$\mathrm{Out}[28] = c, \mathrm{Log}[1+x] + p[0][x] + y[x]p[1][x] + y[x]^2 p[2][x]$

In[29] := Eq = D[Res,x] − y[x]/(x + 1)

$\mathrm{Out}[29] = \frac{c}{1+x} - \frac{y[x]}{1+x} + \frac{p[1][x]}{x} + \frac{2y[x]p[2][x]}{x} + p[0]'[x] + y[x]p[1]'[x] + y[x]^2 p[2]'[x]$

In[30] := Eqs = Table[Coefficient[Eq, y[x], n] == 0, {n, 0, 2}]

$$\text{Out[30]} = \left\{\frac{c}{1+x} + \frac{p[1][x]}{x} + p[0]'[x] == 0, -\frac{1}{1+x} + \frac{2p[2][x]}{x} + p[1]'[x] == 0,\right.$$
$$\left. p[2]'[x] == 0\right\}$$

As in the previous examples, $p_2(x)$ is a constant:

In[31] := p[2][x_] := p[2, 0]; Eqs

$$\text{Out[31]} = \left\{\frac{c}{1+x} + \frac{p[1][x]}{x} + p[0]'[x] == 0, -\frac{1}{1+x} + \frac{2p[2,0]}{x} + p[1]'[x] == 0,\right.$$
$$\left. \text{True}\right\}$$

A polynomial $p_1(x)$ satisfying the second equation does not exist. Therefore, this integral cannot be taken in elementary functions.

In[32] := Clear[Res, Eq, Eqs, p]

Example 4

Let's consider

$$\int \frac{\mathrm{d}x}{y} = p_1(x)y + p_0(x) + c\log(y)$$

(it is not quite clear what the degree of the right-hand side in y should be; we shall see in a moment that this is irrelevant).

In[33] := Res = Sum[p[n][x] * y[x]^n, {n, 0, 1}] + c * Log[y[x]]

$\text{Out[33]} = c, \text{Log}[y[x]] + p[0][x] + y[x]p[1][x]$

In[34] := Eq = Expand[y[x] * D[Res, x] − 1]

$$\text{Out[34]} = -1 + \frac{c}{x} + \frac{y[x]p[1][x]}{x} + y[x]p[0]'[x] + y[x]^2 p[1]'[x]$$

In[35] := Eqs = Table[Coefficient[Eq, y[x], n] == 0, {n, 0, 2}]

$$\text{Out[35]} = \left\{-1 + \frac{c}{x} == 0, \frac{p[1][x]}{x} + p[0]'[x] == 0, p[1]'[x] == 0\right\}$$

From the last equation, $p_1(x)$ is a constant. The previous equation shows that it is 0, and $p_0(x)$ is a constant (it may be set to 0). It is clear that if we started from some other degree in y, we would all the same find that all $p_n(x) = 0$. And the first equation cannot be solved for c.

In[36] := Clear[Res, Eq, Eqs]

Consider the integrals

$$\int \frac{\mathrm{d}x}{xy}, \qquad \int \frac{\mathrm{d}x}{y+1}$$

in a similar way.

10.3 Exponential Extension

Now we shall consider an exponential extension with some $y = \exp r(x)$. If an integral of an expression from this extension can be taken in elementary functions, it has the form

$$\int \frac{N(x,y)}{D(x,y)} \mathrm{d}x = \frac{P(x,y)}{\hat{D}(x,y)} + \sum c_i \log q_i \,,$$
$$D = \prod D_i^{d_i} \quad \Rightarrow \quad \hat{D} = \prod D_i^{\hat{d}_i} \,,$$

where $\hat{d}_i = d_i - 1$ always except the case $D_i = y$ in which $\hat{d}_i = d_i$. This is because the derivative of $1/y$ is proportional to $1/y$, the degree of y in the denominator does not increase. We exclude y from the list of the factors q_i (if it was present, of course) because $\log y = r(x)$ is a rational function, and such a contribution is already accounted for. As usual, we differentiate the result and equate to the integrand to obtain a linear system. If it cannot be solved, then the integral does not exist in elementary functions.

Example 1

Let $y = \mathrm{e}^x$:
In[37] := y′[x_] := y[x]
Let's calculate

$$\int y \,\mathrm{d}x = p_1(x)y + p_0(x)$$

(the degree in y does not change when differentiating; therefore the polynomial in y in the right-hand side should have the same degree as the integrand).
In[38] := Res = Sum[p[n][x] ∗ y[x]^n, {n, 0, 1}]
Out[38] = p[0][x] + y[x]p[1][x]
In[39] := Eq = D[Res, x] − y[x]
Out[39] = −y[x] + y[x]p[1][x] + p[0]′[x] + y[x]p[1]′[x]
In[40] := Eqs = Table[Coefficient[Eq, y[x], n] == 0, {n, 0, 1}]
Out[40] = {p[0]′[x] == 0, −1 + p[1][x] + p[1]′[x] == 0}
In[41] := p[1][x_] := Sum[p[1, m] ∗ x^m, {m, 0, 1}]; Eqs[[2]]
Out[41] = −1 + p[1, 0] + p[1, 1] + xp[1, 1] == 0
From this equation:
In[42] := p[1, 1] = 0; p[1, 0] = 1; Eqs
Out[42] = {p[0]′[x] == 0, True}
Therefore $p_0(x)$ is a constant (which may be omitted):
In[43] := p[0][x_] := 0; Res
Out[43] = y[x]

We've got the expected result.
In[44] := Clear[Res, Eq, Eqs, *p*]

Example 2

Now let's calculate

$$\int xy\,\mathrm{d}x = p_1(x)y + p_0(x)\,.$$

In[45] := Res = Sum[*p*[*n*][*x*] ∗ *y*[*x*]^*n*, {*n*, 0, 1}]
Out[45] = $p[0][x] + y[x]p[1][x]$
In[46] := Eq = *D*[Res, *x*] − *x* ∗ *y*[*x*]
Out[46] = $-xy[x] + y[x]p[1][x] + p[0]'[x] + y[x]p[1]'[x]$
In[47] := Eqs = Table[Coefficient[Eq, *y*[*x*], *n*] == 0, {*n*, 0, 1}]
Out[47] = $\{p[0]'[x] == 0, -x + p[1][x] + p[1]'[x] == 0\}$
In[48] := *p*[1][*x*_] := Sum[*p*[1, *m*] ∗ *x*^*m*, {*m*, 0, 1}]; Eqs[[2]]
Out[48] = $-x + p[1,0] + p[1,1] + xp[1,1] == 0$
From this equation:
In[49] := *p*[1, 1] = 1; *p*[1, 0] = −1; Eqs
Out[49] = $\{p[0]'[x] == 0, \mathrm{True}\}$
In[50] := *p*[0][*x*_] := 0; Res
Out[50] = $(-1 + x)y[x]$
And without integration by parts!
In[51] := Clear[Res, Eq, Eqs, *p*]
Consider

$$\int x^2 y\,\mathrm{d}x$$

in a similar way.

Example 3

Now let's try to calculate

$$\int \frac{y}{x}\,\mathrm{d}x = p_1(x)y + p_0(x) + c\log x\,.$$

In[52] := Res = Sum[*p*[*n*][*x*] ∗ *y*[*x*]^*n*, {*n*, 0, 1}] + *c* ∗ Log[*x*]
Out[52] = $c, \mathrm{Log}[x] + p[0][x] + y[x]p[1][x]$
In[53] := Eq = *D*[Res, *x*] − *y*[*x*]/*x*
Out[53] = $\frac{c}{x} - \frac{y[x]}{x} + y[x]p[1][x] + p[0]'[x] + y[x]p[1]'[x]$

In[54] := Eqs = Table[Coefficient[Eq, *y*[*x*], *n*] == 0, {*n*, 0, 1}]

$$\text{Out}[54] = \left\{\frac{c}{x} + p[0]'[x] == 0, -\frac{1}{x} + p[1][x] + p[1]'[x] == 0\right\}$$

A polynomial $p_1(x)$ satisfying the second equation does not exist. Therefore this integral cannot be taken in elementary functions.

In[55] := Clear[Res, Eq, Eqs]

Example 4

Let's calculate

$$\int \frac{\mathrm{d}x}{y} = \frac{p_1(x)y + p_0(x)}{y}$$

(here $\hat{D} = y$, and there are no q_i). Of course, we could denote e^{-x} as y, and the problem would reduce to the Example 1 with trivial modifications; but we want to observe how the algorithm works in this new case.

In[56] := Res = Sum[*p*[*n*][*x*] ∗ *y*[*x*]^*n*, {*n*, 0, 1}]/*y*[*x*]

$$\text{Out}[56] = \frac{p[0][x] + y[x]p[1][x]}{y[x]}$$

In[57] := Eq = Expand[*y*[*x*] ∗ *D*[Res, *x*] − 1]

$\text{Out}[57] = -1 - p[0][x] + p[0]'[x] + y[x]p[1]'[x]$

In[58] := Eqs = Table[Coefficient[Eq, *y*[*x*], *n*] == 0, {*n*, 0, 1}]

$\text{Out}[58] = \{-1 - p[0][x] + p[0]'[x] == 0, p[1]'[x] == 0\}$

From the second equation, $p_1(x)$ is a constant, which may be omitted—this is the integration constant.

In[59] := *p*[1][*x*_] := 0

In[60] := *p*[0][*x*_] := Sum[*p*[0, *m*] ∗ *x*^*m*, {*m*, 0, 1}]; Eqs[[1]]

$\text{Out}[60] = -1 - p[0,0] + p[0,1] - xp[0,1] == 0$

Therefore

In[61] := *p*[0, 1] = 0; *p*[0, 0] = −1; Res

$$\text{Out}[61] = -\frac{1}{y[x]}$$

In[62] := Clear[Res, Eq, Eqs, *p*]

Example 5

Let's consider

$$\int \frac{\mathrm{d}x}{y-1} = p_1(x)y + p_0(x) + c\log(y-1)$$

(now $\hat{D} = 1$, and there is a single q, namely $y - 1$).

In[63] := Res = Sum[p[n][x] * y[x]^n, {n,0,1}] + c * Log[y[x] − 1]
Out[63] = c, Log[−1 + y[x]] + p[0][x] + y[x]p[1][x]
In[64] := Eq = Cancel[(y[x] − 1) * D[Res,x]] − 1
Out[64] = −1 + cy[x] − y[x]p[1][x] + y[x]^2p[1][x] − p[0]′[x] + y[x]p[0]′[x] −
y[x]p[1]′[x] + y[x]^2p[1]′[x]
In[65] := Eqs = Table[Coefficient[Eq,y[x],n] == 0, {n,0,2}]
Out[65] = {−1 − p[0]′[x] == 0, c − p[1][x] + p[0]′[x] − p[1]′[x] == 0,
p[1][x] + p[1]′[x] == 0}
In[66] := p[1][x_] := Sum[p[1,m] * x^m, {m,0,1}]; Eqs[[3]]
Out[66] = p[1,0] + p[1,1] + xp[1,1] == 0
Therefore
In[67] := p[1,1] = 0; p[1,0] = 0; Eqs
Out[67] = {−1 − p[0]′[x] == 0, c + p[0]′[x] == 0, True}
From the first equation, $p_0(x) = -x$ (omitting the integration constant).
In[68] := p[0][x_] := −x; Eqs
Out[68] = {True, −1 + c == 0, True}
In[69] := c = 1; Res
Out[69] = −x + Log[−1 + y[x]]
In[70] := Clear[Res, Eq, Eqs, p, c]

Consider

$$\int \frac{x}{y-1}\,\mathrm{d}x$$

in a similar way and demonstrate that this integral does not exist in elementary functions.

Example 6

Of course, the method can be also used for other exponential extensions. For example, let $y = \exp x^2$:
In[71] := y′[x_] := 2 * x * y[x]
Let's consider

$$\int y\,\mathrm{d}x = p_1(x)y + p_0(x)\,.$$

In[72] := Res = Sum[p[n][x] * y[x]^n, {n,0,1}];
In[73] := Eq = D[Res,x] − y[x]
Out[73] = −y[x] + 2xy[x]p[1][x] + p[0]′[x] + y[x]p[1]′[x]
In[74] := Eqs = Table[Coefficient[Eq,y[x],n] == 0, {n,0,1}]
Out[74] = {p[0]′[x] == 0, −1 + 2xp[1][x] + p[1]′[x] == 0}
In[75] := p[1][x_] := Sum[p[1,m] * x^m, {m,0,1}]; ExpandAll[Eqs[[2]]]
Out[75] = −1 + 2xp[1,0] + p[1,1] + 2$x^2$$p$[1,1] == 0
This equation cannot be solved. A larger degree of $p_1(x)$ does not help. Therefore this integral cannot be taken in elementary functions.
In[76] := Clear[Res, Eq, Eqs, p]

Example 7

And what about

$$\int xy\,\mathrm{d}x = p_1(x)y + p_0(x)\ ?$$

In[77] := Res = Sum[*p*[*n*][*x*] ∗ *y*[*x*]^*n*, {*n*, 0, 1}];
In[78] := Eq = *D*[Res, *x*] − *x* ∗ *y*[*x*]
Out[78] = $-xy[x] + 2xy[x]p[1][x] + p[0]'[x] + y[x]p[1]'[x]$
In[79] := Eqs = Table[Coefficient[Eq, *y*[*x*], *n*] == 0, {*n*, 0, 1}]
Out[79] = $\{p[0]'[x] == 0, -x + 2xp[1][x] + p[1]'[x] == 0\}$
In[80] := *p*[1][x_] := Sum[*p*[1, *m*] ∗ *x*^*m*, {*m*, 0, 1}]; ExpandAll[Eqs[[2]]]
Out[80] = $-x + 2xp[1,0] + p[1,1] + 2x^2 p[1,1] == 0$
From this equation, $p_{11} = 0$, $p_{10} = 1/2$. The first equation says that $p_0(x)$ is a constant (it may be omitted).
In[81] := *p*[1, 1] = 0; *p*[1, 0] = 1/2; *p*[0][*x*_] := 0; Res
Out[81] = $\dfrac{y[x]}{2}$
In[82] := Clear[Res, Eq, Eqs, *p*]

Consider

$$\int x^2 y\,\mathrm{d}x\,, \qquad \int \frac{y}{x}\,\mathrm{d}x$$

in a similar way.

10.4 Elementary Functions

A tower of extensions can be constructed. We start from the set of rational functions $N(x)/D(x)$. Then we introduce y_1, which is either a root of a polynomial equation $p(x, y_1) = 0$, or logarithm or exponent of some rational function of x, and we obtain a first extension—the set of functions $N(x, y_1)/D(x, y_1)$. Then we introduce y_2, which is either a root of a polynomial equation $p(x, y_1, y_2) = 0$, or logarithm or exponent of some function from the previous extension, and we get a next extension—the set of rational functions of x, y_1, y_2. And so on. We should take care that an extension which seems transcendental is not in fact algebraic; if we neglect this, the methods designed for transcendental extensions may break down, e.g., produce divisions by 0. For example, if $y_1 = \mathrm{e}^x$, then it would not be a good idea to introduce the exponential extension with $y_2 = e^{2x}$, because y_2 is not algebraically independent: $y_2 = y_1^2$. Similarly, if $y_1 = \log x$, then it's not reasonable to introduce the logarithmic extension with $y_2 = \log 2x$, because $y_2 = y_1 + \log 2$ (in this case the field of constants needs to be extended by the transcendental number $\log 2$). In simple cases such restrictions are obvious, but in complicated ones it is necessary to decide if some function from some extension is identical 0, and this problem is algorithmically unsolvable in general.

Such towers of algebraic, logarithmic, and exponential extensions include all functions called elementary. And even some extra ones: an algebraic extension can be defined, e.g., by a root of a fifth degree polynomial unsolvable in radicals. Indeed, trigonometric functions reduce to exponentials, and inverse trigonometric ones—to logarithms. For each elementary function there exists a tower of extensions to which it belongs (it is not unique).

Suppose we want to integrate an elementary function. We construct a tower of extensions to which it belongs. If the indefinite integral exists in elementary functions, it belongs to some further extension of our tower by some extra logarithms (their number may be zero). The Risch algorithm allows one to decide in a finite number of steps if the result exists in this further extension, and if so, to find it; if it does not exist, the algorithm proves this fact. In its classical form, the algorithm is recursive in extensions—it calls itself for solving integration subproblems in previous (smaller) extensions, until rational functions. There is a simpler and more efficient version of the Risch algorithm—to write down the general form of the result with unknown coefficients, differentiate it and equate to the integrand. Then the problem reduces to solving a linear system. This approach is guaranteed to be correct if we know upper bounds on the degrees of the polynomial $P(x, y_1, y_2, \dots)$ in its variables. But such upper bounds are not always known (as we have seen in the examples, they are known if there is no denominator). Therefore some heuristic rules to bound the degrees of P are used. This can give a situation when no result is found, though it really exists (but has a larger degree in some variables).

The Risch algorithm is an outstanding achievement of mathematics in the twentieth century. But it does not solve all problems with indefinite integration. The answer that no result exists in elementary functions is not very useful. It would be much better to get the result with some special functions. There were attempts to generalize the Risch algorithm to some special functions (the error function, polylogarithms). Some of them are implemented in *Mathematica*.