

# Chapter 2

## Integration of rational functions

Franz Winkler

Computer Analysis, SS 2021

The problem we consider in this section is the integration of rational functions with rational coefficients, i.e. to compute

$$\int \frac{p(x)}{q(x)} dx,$$

where  $p(x), q(x) \in \mathbb{Q}[x]$ ,  $\gcd(p, q) = 1$ , and  $q(x)$  is monic. We exclude the trivial case  $q = 1$ .

From classical calculus we know that this integral can be expressed as

$$\int \frac{p(x)}{q(x)} dx = \frac{g(x)}{q(x)} + c_1 \cdot \log(x - \alpha_1) + \cdots + c_n \cdot \log(x - \alpha_n), \quad (2.1)$$

where  $g(x) \in \mathbb{Q}[x]$ ,  $\alpha_1, \dots, \alpha_n$  are the different roots of  $q$  in  $\mathbb{C}$ , and  $c_1, \dots, c_n \in \mathbb{Q}(\alpha_1, \dots, \alpha_n)$ . This requires factorization of  $q$  over  $\mathbb{C}$  into its linear factors, decomposing  $p/q$  into its complete partial fraction decomposition, and computation in the potentially extremely high degree algebraic extension  $\mathbb{Q}(\alpha_1, \dots, \alpha_n)$ . Then the solution (2.1) is achieved by integration by parts and C. Hermite's reduction method.

However, as we will see in the sequel, complete factorization of the denominator can be avoided, resulting in a considerable decrease in computational complexity. Instead of factoring  $q$  we will only use its squarefree factors.

The material for this chapter is taken from Chapter 4 of [Win96].

## 2.1 Squarefree factorization and partial fraction decomposition

A Euclidean domain in which quotient and remainder are computable by algorithms `quot` and `rem` admits an algorithm for computing the greatest common divisor  $g$  of any two elements  $a, b$ . This algorithm has originally been stated by Euclid for the domain of the integers. In fact, it can be easily extended to compute not only the gcd but also the coefficients  $s, t$  in the linear combination  $g = s \cdot a + t \cdot b$ , i.e. the Bezout cofactors.

$\mathbb{Q}[x]$  is such a Euclidean domain. So the algorithm `E_EUCLID` computes the gcd and the Bezout cofactors of two polynomials.

## GCD\_EUCLID

for given non-zero polynomials  $a, b \in K[x]$ ,  
compute the GCD  $g$  and the Bézout cofactors  $s, t$

- (1)  $(r_0, r_1, s_0, s_1, t_0, t_1) := (a, b, 1, 0, 0, 1);$   
 $i := 1;$
- (2) **while**  $r_i \neq 0$  **do**  
     $q_i :=$  quotient of  $r_{i-1}$  on division by  $r_i$ ;  
     $(r_{i+1}, s_{i+1}, t_{i+1}) := (r_{i-1}, s_{i-1}, t_{i-1}) - q_i \cdot (r_i, s_i, t_i);$   
     $i := i + 1$   
**endwhile** ;
- (3)  $(g, s, t) := (r_{i-1}, s_{i-1}, t_{i-1})$  ; **return**

**Theorem 2.1.1.** *Let  $a, b$  be non-zero polynomials in  $\mathbb{Q}[x]$  with  $\deg(a) \geq \deg(b) > 0$ . Let  $g, s, t$  be the result of applying  $E\_EUCLID$  to  $a$  and  $b$ . Then  $\deg(s) < \deg(b) - \deg(g)$  and  $\deg(t) < \deg(a) - \deg(g)$ .*

**Corollary.** *Let  $a, b$  be non-zero, relatively prime polynomials in  $\mathbb{Q}[x]$ . Let  $c \in \mathbb{Q}[x]$  such that  $\deg(c) < \deg(a \cdot b)$ . Then  $c$  can be represented uniquely as  $c = u \cdot a + v \cdot b$ , where  $\deg(u) < \deg(b)$  and  $\deg(v) < \deg(a)$ .*

By a few GCD computations we can determine the squarefree factorization of a polynomial in  $\mathbb{Q}[x]$ . For a proof of the following theorem we refer to [Win96], Theorem 4.4.1.

**Theorem 2.1.2.** *Let  $K$  be a field of characteristic 0, and  $a(x)$  a non-constant polynomial in  $K[x]$ . Then  $a$  is squarefree if and only if  $\gcd(a, a') = 1$  (where  $a'$  denotes the derivative of  $a$  w.r.t  $x$ ).*

From this theorem we easily derive an algorithm for determining the squarefree factorization of a polynomial  $a(x)$  in  $\mathbb{Q}[x]$ :

$$a(x) = \prod_{i=1}^s a_i(x)^i,$$

for squarefree, pairwise relatively prime factors  $a_i(x)$ . Details are given in [Win 96].

**Example 2.1.3.** Let us determine the squarefree factorization of

$$a(x) = x^5 + 6x^4 + 11x^3 + 2x^2 - 12x - 8 = \underbrace{(x+1)(x-1)}_{a_1} \cdot \underbrace{(1)}_{a_2}^2 \cdot \underbrace{(x+2)}_{a_3}^3$$

> **a := expand((x+1)\*(x-1)\*(x+2)^3);**

$$a := x^5 + 6x^4 + 11x^3 + 2x^2 - 12x - 8$$

> **b0 := a;**

$$b0 := x^5 + 6x^4 + 11x^3 + 2x^2 - 12x - 8$$

> **b1 := gcd(b0, diff(b0, x));**

> **factor(b1);**

$$(x+2)^2$$

> **c1 := simplify(b0/b1);**

> **factor(c1);**

$$(x-1)(x+1)(x+2)$$

> **b2 := gcd(b1, diff(b1, x));**

$$b2 := x + 2$$

> **c2 := simplify(b1/b2);**

$$c2 := x + 2$$



**> a1 := simplify(c1/c2);**

$$a1 := x^2 - 1$$

**> b3 := gcd(b2,diff(b2,x));**

$$b3 := 1$$

**> c3 := simplify(b2/b3);**

$$c3 := x + 2$$

**> a2 := simplify(c2/c3);**

$$a2 := 1$$

**> b4 := gcd(b3,diff(b3,x));**

$$b4 := 1$$

**> c4 := simplify(b3/b4);**

$$c3 := 1$$

**> a3 := simplify(c3/c4);**

$$a3 := x + 2$$

So we have determined the squarefree factors  $a_1, a_2, a_3$  of  $a$ .

**Definition 2.1.4.** Let  $p(x)/q(x)$  be a *proper rational function* over in  $\mathbb{Q}(x)$ ; i.e.,  $p, q \in \mathbb{Q}[x]$ ,  $\gcd(p, q) = 1$ , and  $\deg(p) < \deg(q)$ . Let  $q = q_1 \cdot q_2^2 \cdots q_k^k$  be the squarefree factorization of  $q$ . Let  $a_1(x), \dots, a_k(x) \in K[x]$  be such that

$$\frac{p(x)}{q(x)} = \sum_{i=1}^k \frac{a_i(x)}{q_i(x)^i} \quad \text{with} \quad \deg(a_i) < \deg(q_i^i) \text{ for } 1 \leq i \leq k. \quad (2.1.1)$$

Then the right hand side of (2.1.1) is called the *incomplete squarefree partial fraction decomposition (ispfd)* of  $p/q$ .

Let  $b_{ij}(x) \in \mathbb{Q}[x]$ ,  $1 \leq j \leq i \leq k$ , be such that

$$\frac{p(x)}{q(x)} = \sum_{i=1}^k \sum_{j=1}^i \frac{b_{ij}(x)}{q_i(x)^j} \quad \text{with} \quad \deg(b_{ij}) < \deg(q_i) \text{ for } 1 \leq j \leq i \leq n. \quad (2.1.2)$$

Then the right hand side of (2.1.2) is called the *(complete) squarefree partial fraction decomposition (spfd)* of  $p/q$ . □

Both the incomplete and the complete squarefree partial fraction decomposition of a proper rational function are uniquely determined. For any proper rational function  $p/q$  the ispdf can be computed by the following algorithm.

**algorithm ISPDF(in:  $p, q$ ; out:  $D$ );**

$[p/q]$  is a proper rational function in  $K(x)$ ,

$D = [[a_1, q_1], \dots, [a_k, q_k]]$  is the ispdf of  $p/q$ , i.e.  $p/q = \sum_{i=1}^k (a_i/q_i^i)$  with  $\deg(a_i) < \deg(q_i^i)$  for  $1 \leq i \leq k$ .]

(1)  $[q_1, \dots, q_k] := \text{SQFR\_FACTOR}(q);$

(2)  $c_0 := p; d_0 := q; i := 1;$

(3) **while**  $i < k$  **do**

$\{ d_i := d_{i-1}/q_i^i;$

    determine  $c_i, a_i$  such that

$\deg(c_i) < \deg(d_i), \deg(a_i) < \deg(q_i^i),$

        and  $c_i \cdot q_i^i + a_i \cdot d_i = c_{i-1} \}$ ;

$a_k := c_{k-1};$

**return**



**Theorem 2.1.5.** *The algorithm ISPFD is correct.*

*Proof:* Immediately before execution of the body of the **while** statement for  $i$ , the relation

$$\frac{p}{q} = \frac{a_1}{q_1} + \cdots + \frac{a_{i-1}}{q_{i-1}^{j-1}} + \frac{c_{i-1}}{d_{i-1}}, \quad \text{where} \quad d_{i-1} = q_i^j \cdots q_k^k,$$

holds, as can easily be seen by induction on  $i$ .

The polynomials  $c_i$  and  $a_i$  in step (3) can be computed by application of the Corollary to Theorem 2.1.1. □

Once we have the incomplete spfd we can rather easily get the complete spfd by successive division. Namely if  $a_i = s \cdot q_i + t$ , then

$$\frac{a_i}{q_i^j} = \frac{s}{q_i^{j-1}} + \frac{t}{q_i^j}.$$

**Example 2.1.6.** Consider the proper rational function

$$\frac{p(x)}{q(x)} = \frac{4x^8 - 3x^7 + 25x^6 - 11x^5 + 18x^4 - 9x^3 + 8x^2 - 3x + 1}{3x^9 - 2x^8 + 7x^7 - 4x^6 + 5x^5 - 2x^4 + x^3}.$$

The squarefree factorization of  $q(x)$  is

$$q(x) = (3x^2 - 2x + 1)(x^2 + 1)^2 x^3,$$

Application of ISPDF yields the incomplete spfd

$$\frac{p(x)}{q(x)} = \frac{4x}{3x^2 - 2x + 1} + \frac{-x^3 + 2x + 2}{(x^2 + 1)^2} + \frac{x^2 - x + 1}{x^3}.$$

By successive division of the numerators by the corresponding  $q_i$ 's we finally get the complete spfd

$$\frac{p(x)}{q(x)} = \frac{4x}{3x^2 - 2x + 1} + \frac{-x}{x^2 + 1} + \frac{3x + 2}{(x^2 + 1)^2} + \frac{1}{x} + \frac{-1}{x^2} + \frac{1}{x^3}. \quad \square$$

## 2.2 The integration algorithm

The problem we consider in this section is the integration of rational functions with rational coefficients, i.e. to compute

$$\int \frac{p(x)}{q(x)} dx = \frac{g(x)}{q(x)} + c_1 \cdot \log(x - \alpha_1) + \cdots + c_n \cdot \log(x - \alpha_n), \quad (2.2.1)$$

where  $p(x), q(x) \in \mathbb{Q}[x]$ ,  $\gcd(p, q) = 1$ , and  $q(x)$  is monic. We exclude the trivial case  $q = 1$ .

First we compute the squarefree factorization of the denominator  $q$ , i.e.

$$q = f_1 \cdot f_2^2 \cdot \dots \cdot f_r^r,$$

where the  $f_i \in \mathbb{Q}[x]$  are squarefree,  $f_r \neq 1$ ,  $\gcd(f_i, f_j) = 1$  for  $i \neq j$ . Based on this squarefree factorization we compute the squarefree partial fraction decomposition of  $p/q$ , i.e.

$$\frac{p}{q} = g_0 + \sum_{i=1}^r \sum_{j=1}^i \frac{g_{ij}}{f_i^j} = g_0 + \frac{g_{11}}{f_1} + \frac{g_{21}}{f_2} + \frac{g_{22}}{f_2^2} + \dots + \frac{g_{r1}}{f_r} + \dots + \frac{g_{rr}}{f_r^r}, \quad (2.2.2)$$

where  $g_0, g_{ij} \in \mathbb{Q}[x]$ ,  $\deg(g_{ij}) < \deg(f_i)$ , for all  $1 \leq j \leq i \leq r$ .

Integrating  $g_0$  is no problem, so let us consider the individual terms in (2.2.2).

Now let  $\frac{g}{f^n}$  be one of the non-trivial terms in (1.2.1) with  $n \geq 2$ , i.e.  $f$  is squarefree and  $\deg(g) < \deg(f)$ . We reduce the computation of

$$\int \frac{g(x)}{f(x)^n} dx$$

to the computation of an integral of the form

$$\int \frac{h(x)}{f(x)^{n-1}} dx, \quad \text{where} \quad \deg(h) < \deg(f).$$

This is achieved by a reduction process due to C. Hermite.



Since  $f$  is squarefree, we have  $\gcd(f, f') = 1$ . By the extended Euclidean algorithm E-EUCLID and the corollary to Theorem 2.1.1 compute  $c, d \in \mathbb{Q}[x]$  such that

$$g = c \cdot f + d \cdot f', \quad \text{where} \quad \deg(c), \deg(d) < \deg(f).$$

By integration by parts we can now reduce <sup>1</sup>

$$\begin{aligned} \int \frac{g}{f^n} &= \int \frac{c \cdot f + d \cdot f'}{f^n} = \int \frac{c}{f^{n-1}} + \int \frac{d \cdot f'}{f^n} \\ &= \int \frac{c}{f^{n-1}} - \frac{d}{(n-1) \cdot f^{n-1}} + \int \frac{d'}{(n-1) \cdot f^{n-1}} \\ &= -\frac{d}{(n-1) \cdot f^{n-1}} + \int \overbrace{\frac{c + d'/(n-1)}{f^{n-1}}}^h, \end{aligned}$$

where  $\deg(h) < \deg(f)$ .

---

<sup>1</sup> integration by parts:  $\int u \cdot v' = u \cdot v - \int u' \cdot v$ .

Choose:  $u = -d$ ,  $v = 1/((n-1)f^{n-1})$ ; so  $u' = -d'$ ,  $v' = -f'/f^n$ .

Now we collect all the rational partial results and the remaining integrals and put everything over a common denominator, so that we get polynomials  $g(x), h(x) \in \mathbb{Q}[x]$  such that

$$\int \frac{p}{q} = \int g_0 + \frac{g}{\underbrace{f_2 \cdot f_3^2 \cdots f_r^{r-1}}_{\bar{q}}} + \int \frac{h}{\underbrace{f_1 \cdot \cdots \cdot f_r}_{q^*}}, \quad (2.2.3)$$

where  $\deg(g) < \deg(\bar{q})$  and  $\deg(h) < \deg(q^*)$ .

We could also determine  $g$  and  $h$  in (2.2.3) by first choosing undetermined coefficients for these polynomials, differentiating (4.6.3), and then solving the resulting linear system for the undetermined coefficients. However, the Hermite reduction process is usually faster. Let us prove that the decomposition in (2.2.3) is unique.

**Lemma 2.2.1.** *Let  $p, q, u, v \in \mathbb{Q}[x]$ ,  $\gcd(p, q) = 1$ ,  $\gcd(u, v) = 1$ , and  $p/q = (u/v)'$  (so  $u/v$  is the integral of  $p/q$ ). Let  $w \in \mathbb{Q}[x]$  be a squarefree factor of  $q$ . Then  $w$  divides  $v$ , and the multiplicity of  $w$  in  $q$  is strictly greater than the multiplicity of  $w$  in  $v$ .*

*Proof:* Clearly we can restrict ourselves to  $w$  being irreducible (otherwise apply the lemma for all irreducible factors of  $w$ ). Now, since

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2} = \frac{p}{q},$$

$w$  must divide  $v$ . Assume now that  $v = w^r \hat{w}$  with  $\gcd(w, \hat{w}) = 1$ . We show that  $w^r$  does not divide  $u'v - uv'$ . Suppose it does. Since  $w^r$  divides  $u'v$  and  $\gcd(w, u) = 1$ ,  $w^r$  would have to divide  $v' = rw^{r-1}w'\hat{w} + w^r\hat{w}'$ . Hence,  $w$  would have to divide  $w'\hat{w}$ . But this is impossible since  $w$  is irreducible. Therefore  $w^{r+1}$  must divide the reduced denominator of  $(u/v)'$ . □

**Theorem 2.2.2.** *The solution  $g, h$  to equation (2.2.3) is unique.*

*Proof:* Suppose there were two solutions. By subtraction we would get a solution for  $p = 0$ ,

$$\int 0 dx = \frac{g}{q} + \int \frac{h}{q^*} dx.$$

So  $(g/\bar{q})' = -h/q^*$ . By Lemma 2.2.1, every factor in the denominator of  $h/q^*$  must have multiplicity at least 2. This is impossible, since  $q^*$  is squarefree. □

The integral  $\int h/q^*$  can be computed in the following well-known way: Let  $q^*(x) = (x - \alpha_1) \cdots (x - \alpha_n)$ , where  $\alpha_1, \dots, \alpha_n$  are the distinct roots of  $q^*$ . Then

$$\int \frac{h(x)}{q^*(x)} dx = \sum_{i=1}^n \int \frac{c_i}{x - \alpha_i} dx = \sum_{i=1}^n c_i \log(x - \alpha_i) \quad (2.2.4)$$

with  $c_i = \frac{h(\alpha_i)}{q^{*'}(\alpha_i)}, \quad 1 \leq i \leq n.$

No part of the sum of logarithms in (2.2.4) can be a rational function, as we can see from the following theorem in [Hardy 1916], p. 14.

**Theorem 2.2.3.** Let  $\alpha_1, \dots, \alpha_n$  be distinct elements of  $\mathbb{C}$  and  $c_1, \dots, c_n \in \mathbb{C}$ . If  $\sum_{i=1}^n c_i \log(x - \alpha_i)$  is a rational function, then  $c_i = 0$  for all  $1 \leq i \leq n$ . □

**Example 2.2.4.** Let us integrate  $x/(x^2 - 2)$  according to (2.2.4).

$$\begin{aligned} \int \frac{x}{x^2-2} dx &= \int \frac{1/2}{x-\sqrt{2}} dx + \int \frac{1/2}{x+\sqrt{2}} dx \\ &= \frac{1}{2} (\log(x - \sqrt{2}) + \log(x + \sqrt{2})) \\ &= \frac{1}{2} \log(x^2 - 2). \end{aligned}$$

So obviously we do not always need the full splitting field of  $q^*$  in order to express the integral of  $h/q^*$ . In fact, whenever we have two logarithms with the same constant coefficient, we can combine these logarithms. □

The following theorem, which has been independently discovered by M. Rothstein [Rothstein 1976] and B. Trager [Trager 1976], answers the question of what is the smallest field in which we can express the integral of  $h/q^*$ .

**Theorem 2.2.5.** *Let  $p, q \in \mathbb{Q}[x]$  be relatively prime,  $q$  monic and squarefree, and  $\deg(p) < \deg(q)$ . Let*

$$\int \frac{p}{q} = \sum_{i=1}^n c_i \log v_i, \quad (2.2.5)$$

*where the  $c_i$  are distinct non-zero constants and the  $v_i$  are monic squarefree pairwise relatively prime elements of  $\overline{\mathbb{Q}}[x]$ . Then the  $c_i$  are the distinct roots of the polynomial*

$$r(c) = \operatorname{res}_x(p - c \cdot q', q) \in \mathbb{Q}[c],$$

*and*

$$v_i = \gcd(p - c_i \cdot q', q), \quad \text{for } 1 \leq i \leq n.$$



*Proof:* Let  $u_i = (\prod_{j=1}^n v_j)/v_i$ , for  $1 \leq i \leq n$ . Then by differentiation of (2.2.5) we get

$$p \cdot \prod_{i=1}^n v_i = q \cdot \sum_{i=1}^n c_i v_i' u_i.$$

So  $q \mid \prod_{i=1}^n v_i$  and on the other hand each  $v_i \mid q v_i' u_i$ , which implies that each  $v_i \mid q$ . Hence,

$$q = \prod_{i=1}^n v_i, \quad \text{and} \quad p = \sum_{i=1}^n c_i v_i' u_i.$$

Consequently, for each  $j$ ,  $1 \leq j \leq n$ , we have

$$\begin{aligned} v_j &= \gcd(0, v_j) = \gcd(p - \sum_{i=1}^n c_i v_i' u_i, v_j) \\ &= \gcd(p - c_j v_j' u_j, v_j) = \gcd(p - c_j \sum_{i=1}^n v_i' u_i, v_j) \\ &= \gcd(p - c_j q', v_j), \end{aligned}$$

and for  $l \neq j$  we have

$$\gcd(p - c_j q', v_l) = \gcd(p - c_j v_l' u_l, v_l) = \gcd((c_l - c_j) v_l' u_l, v_l) = 1.$$

Thus we conclude that

$$v_i = \gcd(p - c_i q', q), \quad \text{for } 1 \leq i \leq n. \quad (2.2.6)$$

(2.2.6) implies that  $\text{res}_x(p - c_i q', q) = 0$  for all  $1 \leq i \leq n$ .

Conversely, if  $c \in \overline{\mathbb{Q}}$  and  $\text{res}_x(p - c q', q) = 0$ , then

$\gcd(p - c q', q) = s(x) \in \overline{\mathbb{Q}}[x]$  with  $\deg(s) > 0$ . Thus, any irreducible factor  $t(x)$  of  $s(x)$  divides

$p - c q' = \sum_{i=1}^n c_i v'_i u_i - c \sum_{i=1}^n v'_i u_i$ . Since  $t$  divides one and only one  $v_j$ , we get  $t \mid (c_j - c) v'_j u_j$ , which implies that  $c_j - c = 0$ . Thus, the  $c_j$  are exactly the distinct roots of  $r(c)$ . □

**Example 2.2.4. (continued)** We apply Theorem 2.2.5.

$$r(c) = \operatorname{res}_x(p - cq', q) = \operatorname{res}_x(x - c(2x), x^2 - 2) = -2(2c - 1)^2.$$

There is only one root of  $r(c)$ , namely  $c_1 = 1/2$ . We get the argument of the corresponding logarithm as

$$v_1 = \gcd(x - \tfrac{1}{2}(2x), x^2 - 2) = x^2 - 2. \text{ So}$$

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



**Example 2.2.6.** Let us consider integrating the rational function

$$\frac{p(x)}{q(x)} = \frac{4x^8 - 3x^7 + 25x^6 - 11x^5 + 18x^4 - 9x^3 + 8x^2 - 3x + 1}{3x^9 - 2x^8 + 7x^7 - 4x^6 + 5x^5 - 2x^4 + x^3}.$$

The squarefree factorization of  $q(x)$  is

$$q(x) = (3x^2 - 2x + 1)(x^2 + 1)^2x^3,$$

so the squarefree partial fraction decomposition of  $p/q$  is

$$\frac{p(x)}{q(x)} = \frac{4x}{3x^2 - 2x + 1} + \frac{-x}{x^2 + 1} + \frac{3x + 2}{(x^2 + 1)^2} + \frac{1}{x} + \frac{-1}{x^2} + \frac{1}{x^3}.$$

Now let us consider the third term of this decomposition, i.e. we determine

$$\int \frac{3x+2}{(x^2+1)^2} dx.$$

By the extended Euclidean algorithm we can write

$$3x+2 = 2 \cdot (x^2+1) + \left(-x + \frac{3}{2}\right) \cdot (2x).$$

Integration by parts yields

$$\begin{aligned} \int \frac{3x+2}{(x^2+1)^2} dx &= \int \frac{2}{x^2+1} dx + \int \frac{(-x+\frac{3}{2}) \cdot (2x)}{(x^2+1)^2} dx \\ &= \int \frac{2}{x^2+1} dx + \frac{(-x+\frac{3}{2}) \cdot (-1)}{x^2+1} - \int \frac{1}{x^2+1} dx \\ &\quad + \frac{x-\frac{3}{2}}{x^2+1} + \int \frac{1}{x^2+1} dx \end{aligned}$$

The remaining integral is purely logarithmic, namely

$$\int \frac{1}{x^2+1} dx = \frac{i}{2} \cdot \log(1-ix) - \frac{i}{2} \cdot \log(1+ix) = \arctan(x). \quad \square$$

Thank you for your attention!

