AAECC (2008) 19:1–25
DOI 10.1007/s00200-008-0059-z

Decomposition of ordinary differential polynomials

Xiao-Shan Gao · Mingbo Zhang

Received: 26 May 2005 / Revised: 13 July 2007 / Published online: 30 January 2008
© Springer-Verlag 2008

Abstract In this paper, we present a complete algorithm to decompose nonlinear differential polynomials in one variable and with coefficients in a computable differential field $\mathcal{K}$ of characteristic zero. The algorithm provides an efficient reduction of the problem to the factorization of LODOs over the same coefficient field. Besides arithmetic operations, the algorithm needs decomposition of algebraic polynomials, factorization of multi-variable polynomials, and solution of algebraic linear equation systems. The algorithm is implemented in Maple for the *constant field* case. The program can be used to decompose differential polynomials with thousands of terms effectively.

Keywords Decomposition · Differential polynomial · Pseudo linear differential polynomial · Differential degree

1 Introduction

The study on functional decomposition algorithms started with the decomposition of univariate polynomials. The first algorithm to find decompositions of polynomials was presented by Barton–Zippel [2] and Alagar–Tanh [1]. This was soon followed by the work of Kozen and Landau [17] who proposed a polynomial time decomposition algorithm. Gutierrez et al. gave an almost quadratic complexity decomposition

This article was partially supported by a National Key Basic Research Project of China (NO. G1998030600) and by a USA NSF grant CCR-0201253.

X.-S. Gao (✉) · M. Zhang
Key Laboratory of Mathematics Mechanization, Institute of Systems Science, AMSS, Academia Sinica, Beijing 100080, China
e-mail: xgao@mmrc.iss.ac.cn

algorithm in [11]. Gathen proposed algorithms to find univariate decompositions in both of the tame and the wild cases [31,32]. Gathen et al. proposed an algorithm to find univariate decomposition factors of multivariate polynomials [12,33]. Decomposition algorithms and unirational fields were discussed in [13,14]. In [34], Zippel presented polynomial time algorithms to decompose a given univariate rational function over an arbitrary field.

The decomposition of differential polynomials is much more difficult, even in the linear case, which is equivalent to the factorization of linear ordinary differential operators (LODOs). The first algorithms to factor LODOs proposed by Bronstein, Petkovšek, and Schwarz [3,21] were based on the classical work of Beke. These algorithms reduce the construction of the right factor into the problem of finding hyper-exponential solutions of the so-called associated equations. Another approach based on the eigenring associated with the LODO was proposed by van der Put and Singer [28]. In [29,30], van Hoeij proposed algorithms to factor an LODO with rational functions as well as power series coefficients. This algorithm was extended to factor LODOs over an exponential extension of a base field by Fredet [5]. In [26], Tsarev presented an algorithm for complete enumeration of all factorizations of an LODO. In [9], Giesbrecht and Zhang presented methods to find factorizations for the more general Ore polynomials over the rational function field of positive characteristic. In [10], Grigor'ev showed that the worst case complexity of factoring LODOs is exponential.

The problem of decomposing nonlinear differential polynomials was discussed in the classic work by Königsberger [16]. In [27], Tsarev considered differential polynomials of the form $y_n - R(x, y, \ldots, y_{n-1})$ and gave a decision procedure for the existence of a decomposition. The decomposition of differential polynomials is reduced to the solution of a system of nonlinear differential equations, which is generally considered a very difficult task. Sosnin proposed an algorithm to find a non-parametric decomposition of differential polynomials [25]. In [7], we gave an algorithm to find decompositions of differential polynomials over a constant field. In this paper, a complete decomposition algorithm for differential polynomials over any field $\mathcal{K}$ of characteristic zero will be presented.

Decomposition algorithms could be used to simplify the solution of differential equations. Applications of the factorization of LODOs to compute closed form solutions and to determine the Galois group were discussed by Singer and Ulmer [22,23]. In [18], Hu solved a problem of equilibrium of elastic body with the decomposition of differential equations.

The basic idea to decompose a given differential polynomial $f = g \circ h$ is as follows. We first try to find a polynomial decomposition for f, where g is a univariate algebraic polynomial. If f does not have a polynomial decomposition, we may find a set of possible separants of h. For each possible separant S of h, we try to find a possible right factor h. This procedure is divided into two phases: in the first phase presented in Sect. 3, we can find a differential polynomial p such that if h is a right decomposition factor of f with separant S, then h is also a right decomposition factor of p and the left decomposition factor of p w.r.t. h is pseudo linear (for definition, see Sect. 3). In the second phase presented in Sect. 4, the pseudo linear case is reduced to factorization of LODOs and the solution of algebraic parametric linear equation systems. The second

phase is the key contribution of the paper. After the right decomposition factor h is obtained, it is easy to find g.

The algorithm is implemented in Maple for the constant coefficient field case. Our experiments show that the program can be used to decompose differential polynomials with thousands of terms efficiently. To implement our algorithm for the case of rational function field, we need to have an implementation which could give all the possible factorizations for an LODO. This problem itself is also a difficult task and we do not find such an implementation. Based on our experiments on the constant field case and the analysis of our algorithm, our algorithm provides an efficient reduction of the decomposition of nonlinear differential polynomials to linear ones.

The paper is organized as follows. In Sect. 2, we present the necessary notations and preliminary results. In Sect. 3, we show how to reduce the general case to pseudo linear case. In Sect. 4, the pseudo linear case is solved. In Sect. 5, we give the main algorithm and experimental results. In Sect. 6, we give the conclusions.

2 Decomposition of differential polynomials

Let $\mathcal{K}$ be a computable differential field of characteristic zero, $C_{\mathcal{K}}$ the constant field of $\mathcal{K}$, y a differential indeterminate, $\mathcal{K}\{y\}$ the ordinary differential polynomial ring over $\mathcal{K}$. An element in $\mathcal{K}\{y\}$ is called a differential polynomial. We denote by y_i the ith derivative of y. For a differential polynomial f not in $\mathcal{K}$, let y_o be the highest derivative appearing in f. Then o is called the *order* of f and is denoted by o_f. Let o_f be the order of f, d_f the degree of f in y_{o_f}, i_f the coefficient of $y_{o_f}^{d_f}$ in f, $s_f = \frac{\partial f}{\partial y_{o_f}}$. Then d_f, i_f, and s_f are called the *degree, initial, and separant* of f respectively.

A *monomial* in $\mathcal{K}\{y\}$ is always arranged in the form $\prod_{0\le i\le r} y_i^{\alpha_i}$ where $\alpha_i \in \mathbb{N}$. The number $\sum_{0\le i\le r} \alpha_i$ is called the *total degree* and $\sum_{0\le i\le r} i\cdot\alpha_i$ is called the *differential degree* of the monomial. The maximal total degree and differential degree for all monomials in a differential polynomial f are called the *total degree* and *differential degree* of f, denoted by $\mathrm{tdeg}(f)$ and $\mathrm{ddeg}(f)$ respectively. If the total (differential) degrees of the monomials in f are all equal, then f is called *total* (*differential*) *degree homogeneous*. If $\mathrm{tdeg}(f) = k$ and f is total degree homogeneous, f is called k-total degree homogeneous. Furthermore, if $k = 1$, we say that f is *linear*.

We define a *rank* between two monomials according to the pure lexicographical order induced by the variable order $y < y_1 < y_2 < \ldots$. The monomial with the highest rank is called the *leading term* of f and is denoted as $\mathrm{lt}(f)$.

For a differential polynomial f and a nonnegative integer k, let $f_{(k)}$ be the kth derivative of f. Then for $k > 0$, we have

$$f_{(k)} = s_f y_{o_f+k} + R_f \tag{1}$$

where R_f is a differential polynomial of order lower than $o_f + k$.

Let $g, h \in \mathcal{K}\{y\}$. We use $g \circ h$ to denote the *composition* of two differential polynomials g and h, which is obtained by substituting y_i in g by $h_{(i)} (0 \le i \le o_g)$. If $f = g \circ h$, g, h are called the *left* and *right decomposition factor* of f respectively. A *decomposition* $f = g \circ h$ is called *nontrivial*, if both g and h are not in the form

$ay + b$ where a and b are in $\mathcal{K}$. Two decompositions $f = g_1 \circ h_1$ and $f = g_2 \circ h_2$ are called *equivalent* if there exist $a, b \in \mathcal{K}$ such that $h_1 = (ay + b) \circ h_2$. Here is an example.

Example 2.1 Let $\mathcal{K} = \mathbb{Q}(t)$ be the differential field of rational functions with the derivative $\frac{d}{dt}$.

$$\begin{aligned} f &= y_1^4 + 2tyy_1^2 + t^2y^2 + 2ty_1^2 + 2y_1y_2 + 2t^2y + ty_1 + y + t^2 + 1 \\ &= (y_1 + y^2) \circ (ty + y_1^2 + t) \\ &= (t^2y^2 + ty_1 + 2t^2y + y + t^2 + 1) \circ \left(y + \frac{1}{t}y_1^2\right). \end{aligned}$$

are two equivalent decompositions of f. In fact, we have $ty + y_1^2 + t = (ty + t) \circ (y + \frac{1}{t}y_1^2)$.

In this paper, we will consider the following problem.

Decomposition problem *Let f be a differential polynomial in $\mathcal{K}\{y\}$ with positive order, find $g, h \in \mathcal{K}\{y\}$ such that $f = g \circ h$.*

The composition operation has the following basic properties. The proofs for these properties are quite simple and are omitted. The readers may consult [8] for the proofs.

- The composition operation is associate:

$$f \circ (g \circ h) = (f \circ g) \circ h. \tag{2}$$

- If $f = g \circ h$, we have

$$o_f = o_g + o_h, \quad s_f = (s_g \circ h) \cdot s_h. \tag{3}$$

- If $f = g \circ h$ and $o_g = 0$, we have

$$d_f = d_g d_h, \quad i_f = (i_g \circ h) i_h^{d_g}. \tag{4}$$

- If $f = g \circ h$ and $o_g > 0$, we have

$$d_f = d_g, \quad i_f = (i_g \circ h) \cdot s_h^{d_f}. \tag{5}$$

- If $f = g \circ h$, $d = \mathrm{tdeg}(f)$, $d_1 = \mathrm{tdeg}(g)$, $d_2 = \mathrm{tdeg}(h)$ and F_d, G_{d_1}, H_{d_2} are the sums of monomials in f, g, h with total degrees d, d_1, d_2 respectively, then

$$d = d_1 \cdot d_2, \quad F_d = G_{d_1} \circ H_{d_2}. \tag{6}$$

- Let $f = f_{d_f} y_{o_f}^{d_f} + f_{d_f-1} y_{o_f}^{d_f-1} + \cdots + f_1 y_{o_f} + f_0$. If $f = g \circ h$ and $o_g > 0$, then

$$s_h^i \text{ divides } f_i, 1 \le i \le d_f. \tag{7}$$

For any $c \in \mathcal{K}$, we have $f = g \circ h = (g \circ (y + c)) \circ ((y - c) \circ h)$ by (2). So we can always assume that h has no term in $\mathcal{K}$. In this case, the term of f in $\mathcal{K}$ is equal to that of g. *In this paper, we always assume that in any decomposition $f = g \circ h$, f, g, h have no term in $\mathcal{K}$.*

In our main algorithm for decomposing differential polynomials, two basic operations are often needed: (1) to check whether a given differential polynomial is a right decomposition factor of another differential polynomial and (2) to find a polynomial decomposition of a differential polynomial, in which the left decomposition factor is in $\mathcal{K}[y]$.

Algorithm 2.2 *Input: differential polynomials f, h.*
Output: a differential polynomial g such that $f = g \circ h$ if such a g exists.

S1 If $f \in \mathcal{K}$, return $g = f$.
S2 If $o_f < o_h$, g does not exist. The algorithm terminates.
S3 If $o_f = o_h$, then we have $o_g = 0$ by (3) and $d_g = d_f/d_h$ by (4). If d_g is not an integer, g does not exist and the algorithm terminates. Otherwise, let $q = \frac{i_f}{i_h^{d_g}}$. By (4), q should be $i_g \circ h$. Go to S5.
S4 If $o_f > o_h$, by (3) and (5) the order o_g and leading degree d_g for g can be computed as follows: $o_g = o_f - o_h$, $d_g = d_f$. Let $q = \frac{i_f}{s_h^{d_g}}$. By (5), q should be $i_g \circ h$.
S5 If q is not a differential polynomial in $\mathcal{K}\{y\}$, g does not exist. The algorithm terminates. Otherwise, we will find $g = i_g y_{o_g}^{d_g} + g_1$ as follows. Let $f_1 = f - q \cdot h_{(o_g)}^{d_g} = g \circ h - q \cdot h_{(o_g)}^{d_g} = (i_g \circ h) h_{(o_g)}^{d_g} + g_1 \circ h - q \cdot h_{(o_g)}^{d_g} = g_1 \circ h$.
Call Algorithm 2.2 with f_1, h and q, h as inputs. Let the outputs be g_1 and i_g respectively. If i_g and g_1 exist, return $g = i_g y_{o_g}^{d_g} + g_1$. Otherwise, g does not exist.

The correctness of Algorithm 2.2 is clear.

Algorithm 2.3 *Input: a differential polynomial f.*
Output: a univariate polynomial g and a differential polynomial h such that $f = g \circ h$ if such g and h exist.

Note that a polynomial decomposition $f = g \circ h$ of a differential polynomial f is the same as a univariate decomposition of f as a multiple variable polynomial in $\mathcal{K}[y, y_1, \ldots, y_{o_f}]$. We use the algorithm in [33] directly to solve this problem. At the same time, if we limit the separant of the right decomposition factor to be a given differential polynomial S, then the polynomial decomposition is unique and it can be obtained directly.

3 Reduction to pseudo linear case

A differential polynomial p in $\mathcal{K}\{y\}$ is called *pseudo linear* if p is of the form $p = y_{o_p} + p_1$ where p_1 is of order lower than o_p.

Consider the decomposition problem on page 4 and write f, g, h as the following form.

$$\begin{aligned} f &= f_{d_f} y_{o_f}^{d_f} + f_{d_f-1} y_{o_f}^{d_f-1} + \cdots + f_1 y_{o_f} + f_0 \\ g &= g_{d_g} y_{o_g}^{d_g} + g_{d_g-1} y_{o_g}^{d_g-1} + \cdots + g_1 y_{o_g} + g_0 \\ h &= i_h y_{o_h}^{d_h} + h_1 \end{aligned} \tag{8}$$

where $f_{d_f} = i_f$, $g_{d_g} = i_g$, and h_1 is of degree lower than d_h in y_{o_h}.

Lemma 3.1 *Let $f = g \circ h$ and $o_g > 0$. Denote $d_f (= d_g)$ by d, $\frac{\partial^k f}{\partial^k y_{o_f}}$ by $s^{(k)}(f)$ and $\frac{\partial^k g}{\partial^k y_{o_g}}$ by $s^{(k)}(g) (1 \le k \le d)$ respectively. Then we have $s^{(i)}(f)/s_h^i = s^{(i)}(g) \circ h$ $(1 \le i \le d)$. Namely, h is a right decomposition factor of $s^{(i)}(f)/s_h^i$.*

Proof By (3), we have $s^{(1)}(f)/s_h = s_f/s_h = s_g \circ h$. This proves the case for $i = 1$. Suppose that the lemma is valid for $i = k$, that is, we have $s^{(k)}(f)/s_h^k = s^{(k)}(g) \circ h$. Using (3) to this equality again, we have $s^{(k+1)}(f)/s_h^{k+1} = \frac{\partial (s^{(k)}(f)/s_h^k)}{\partial y_{o_f}} s_h^{-1} = (\frac{\partial (s^{(k)}(g))}{\partial y_{o_g}} \circ h) s_h s_h^{-1} = s^{(k+1)}(g) \circ h$. We may move s_h out from the scope of the partial differentiation, because $o_{s_h} < o_f = o_{s_f}$ when $d_f > 1$. This proves the case for $i = k + 1$. □

When $o_g > 0$, by (7) we have $s_h^i | f_i (1 \le i \le d_f)$. From this, we may find a finite number of candidates S for s_h. Our algorithm will start with such a possible separant S.

Proposition 3.2 *For differential polynomials f and S, we can either find a decomposition $f = g \circ h$ or a $p \in \mathcal{K}\{y\}$ such that if h is a right decomposition factor of f and $s_h = S$, then h is a right decomposition factor of p and the corresponding left decomposition factor of p w.r.t. h is pseudo linear, or we can confirm that f has no decomposition such that the separant of the right decomposition factor is S.*

Proof We first use Algorithm 2.3 to find a polynomial decomposition $f = g \circ h$ such that $s_h = S$. If such a decomposition exists, we already have a decomposition for f; otherwise, we need to find decompositions $f = g \circ h$ such that $o_g > 0$. When $o_g > 0$, we have $s_h^i | f_i (1 \le i \le d_f)$. Then we check whether $S^i | f_i (1 \le i \le d_f)$ is true. If it is not true, a decomposition satisfying $f = g \circ h$ and $s_h = S$ does not exist; else, suppose that h is a possible right decomposition factor of f such that $s_h = S$, there are two cases:
Case 1 $\frac{f_d}{S^d} = \frac{i_f}{S^d} \notin \mathcal{K}$. By (5), h is a right decomposition factor of $\frac{f_d}{S^d}$. Then we can consider the decomposition of $\frac{f_d}{S^d}$ recursively. Note that after an iteration, the left decomposition factor of $\frac{f_d}{S^d}$ could be a polynomial and the condition $o_g > 0$ is not valid anymore. We need to use Algorithm 2.3 to test whether $\frac{f_d}{S^d}$ has a polynomial decomposition. If $\frac{f_d}{S^d}$ has such a polynomial decomposition, we will check whether the right decomposition factor of it is a right decomposition factor of f. If $\frac{f_d}{S^d}$ does

not have such a polynomial decomposition or the right decomposition factor of it is not a right decomposition factor of f, we can repeat the above procedure recursively. *Case 2* $\frac{f_d}{S^d} = c \in \mathcal{K}$. By (5), we have $i_g = g_d \in \mathcal{K}$. Now if $f = g \circ h$, let $d = d_f (= d_g)$ and $i = d-1$ in Lemma 3.1, we have $s^{(d-1)}(f)/S^{d-1} = s^{(d-1)}(g) \circ h$. Using notations in (8), by direct computation, we have

$$s^{(d-1)}(f) = \frac{\partial^{(d-1)} f}{\partial^{d-1} y_{o_f}} = d! f_d y_{o_f} + (d-1)! f_{d-1}$$
$$s^{(d-1)}(g) = d! g_d y_{o_g} + (d-1)! g_{d-1}.$$

So we have

$$\left(f_d y_{o_f} + \frac{1}{d} f_{d-1}\right) \Big/ S^{d-1} = \left(g_d y_{o_g} + \frac{1}{d} g_{d-1}\right) \circ h \tag{9}$$

By (9) and (1) we have

$$S y_{o_f} + \frac{1}{cd} \cdot \frac{f_{d-1}}{S^{d-1}} = \left(\frac{g_d}{c} y_{o_g} + \frac{1}{cd} g_{d-1}\right) \circ h = \frac{g_d}{c}(S y_{o_f} + R_h) + \frac{1}{cd} g_{d-1} \circ h.$$

Comparing the coefficients of y_{o_f}, we have $S = \frac{g_d}{c} \cdot S$ and then $\frac{g_d}{c} = 1$. Let $w = \frac{1}{cd} \cdot \frac{f_{d-1}}{S^{d-1}}$, $g' = \frac{1}{cd} g_{d-1}$. Then

$$p = S y_{o_f} + w = (y_{o_g} + g') \circ h. \tag{10}$$

So, the left decomposition factor of p w.r.t. h is pseudo linear. Notice that we have the set relation $\{h : s_h = S \text{ and } h \text{ is a right decomposition factor of } f\} \subseteq \{h : s_h = S \text{ and } h \text{ is a right decomposition factor of } p\}$. □

The above analysis leads to the following algorithm.

Algorithm 3.3 *Input: differential polynomials f, S.*

Output: (1) differential polynomials g and h such that $f = g \circ h$; (2) a differential polynomial p such that if h is a right decomposition factor of f with separant S, then there exists $r \in \mathcal{K}\{y\}$ such that $p = r \circ h$ and r is pseudo linear; or (3) an empty set which means that f has no right decomposition factor with separant S.

S1 $q := f$.

S2 Let $d = d_q$. Write q as the form $q = q_d y_{o_q}^d + \cdots + q_1 y_{o_q} + q_0$. If $o_q < o_S$, then output the empty set and terminate the algorithm.

S3 If $o_q = o_S$, execute Algorithm 2.3 with input q to find a decomposition $q = r \circ h$ such that $o_r = 0$, $d_r = d_q/(d_S + 1)$. If such a decomposition exists, go to step S5; otherwise go to step S6.

S4 If $o_q > o_S$, execute Algorithm 2.3 with input q to find a decomposition $q = r \circ h$ such that $o_r = 0$, $d_r = d_f$. If such a decomposition exists, go to step S5; otherwise go to step S6.

S5 If $s_h/S \in \mathcal{K}$, then use Algorithm 2.2 with input f, h to find g such that $f = g \circ h$. If such a g exists, output g, h and terminate the algorithm. Otherwise, go to the next step.

S6 For $i = 1, \ldots, d$, if there exists some i such that $S^i \nmid q_i$, then output the empty set and terminate the algorithm.

S7 Let $c = q_d/S^d$. If $c \in \mathcal{K}$, let $w = \frac{1}{c \cdot d} \frac{q_{d-1}}{S^{d-1}}$ and $p = Sy_{o_q} + w$. Using Algorithm 2.2 with input f, p to find g such that $f = g \circ p$. If such a g exists, output g and p; otherwise output p. This step is based on Case 2 in the proof of Proposition 3.2.

S8 Let $q = c - c_0$, where c_0 denotes the term of c in $\mathcal{K}$. Go to S2. This step is based on Case 1 in the proof of Proposition 3.2.

Proof of correctness of Algorithm 3.3 The algorithm basically follows the proof of Proposition 3.2. Since h is a right decomposition factor of q and $s_h = S$, we have $o_q \geq o_h \geq o_{s_h} = o_S$. Thus, if $o_q < o_S$, h does not exist in step S2.

Steps S3, S4, and S5 treat the polynomial decomposition case. In step S3, since $o_q \geq o_h \geq o_S$ and $o_q = o_S$, we have $o_h = o_S = o_{s_h}$. In this case, we have $d_h = d_{s_h} + 1 = d_S + 1$. By (4), if $q = r \circ h$ is a polynomial decomposition, $d_r = d_q/d_h = d_q/(d_S + 1)$. In step S4, if $q = r \circ h$ is a polynomial decomposition and $o_q > o_S = o_{s_h}$, then h must be of the form $h = Sx_{o_h} + u$ and $o_u < o_h$. In this case, we have $d_r = d_q$. Furthermore, such a decomposition is unique. Therefore, if a polynomial decomposition exists, it can be found in this way.

The correctness for steps S6, S7, and S8 follows directly from the proof of Proposition 3.2. In step S7, since $s_p = S$, we need to check that whether p itself is a right decomposition factor of f. □

Example 3.4 Let $\mathcal{K} = \mathbb{Q}(t)$, $S = y_1$, and

$$\begin{aligned} f = {} & 2ty_1(t^3y_1^4 + 2t^2yy_1^2 + ty^2 + 2ty_1y_2 + y_1^2 + y_1)y_3 + t^3y_1^4y_2 \\ & + 2t^4y_1^4y_2^2 + 4t^3y_1^5y_2 + 2t^2yy_1^2y_2 + 4t^3yy_1^2y_2^2 \\ & + 8t^2yy_1^3y_2 + ty^2y_2 + 2t^2y^2y_2^2 + 4ty^2y_1y_2 + 4ty_1y_2^2 \\ & + 4t^2y_1y_2^3 + 10ty_1^2y_2^2 + 5y_1^2y_2 + 4y_1^3y_2 + y_1y_2 + ty_1^2 + y. \end{aligned}$$

In S1, $q := f$, we have $d_q = 1$, $q_1 = 2ty_1(t^3y_1^4 + 2t^2yy_1^2 + ty^2 + 2ty_1y_2 + y_1^2 + y_1)$.
From S2 to S6, the algorithm does nothing on q.
In S7, $c = \frac{q_1}{S} = 4t^2y_1y_2 + 2t^4y_1^4 + 4t^3yy_1^2 + 2t^2y^2 + 2ty_1^2 + 2ty_1$.
In S8, we return to S2 with $q = 4t^2y_1y_2 + 2t^4y_1^4 + 4t^3yy_1^2 + 2t^2y^2 + 2ty_1^2 + 2ty_1$.
In S2, $d = d_q = 1$ and $o_q = 2 > o_S$. From S2 to S6, the algorithm does nothing on q.
In S7, $c = \frac{q_1}{S} = 4t^2 \in \mathcal{K}$ and by Algorithm 2.2 we find that q is not a right decomposition factor of f, so we output

$$p = q/c = y_1y_2 + \frac{1}{2}t^2y_1^4 + tyy_1^2 + \frac{1}{2}y^2 + \frac{1}{2t}y_1^2 + \frac{1}{2t}y_1.$$

By Proposition 3.2, we have $\{h : s_h = y_1 \text{ and } h \text{ is a right decomposition factor of } f\} \subseteq \{h : s_h = y_1 \text{ and } h \text{ is a right decomposition factor of } p\}$.

4 Solving the pseudo linear case

In the previous section, with a possible separant S, we either have found a decomposition of f, or we can confirm that f has no right decomposition factor with separant S, or we obtain a differential polynomial $p = Sy_{o_p} + w$ satisfying $\{h : s_h = S$ and h is a right decomposition factor of $f\} \subseteq \{h : s_h = S$ and h is a right decomposition factor of $p\}$. In this section, we will solve the following key problem of this paper: *find the right decomposition factor h of $p = Sy_{o_p} + w(o_w < o_p)$ such that $s_h = S$.*

By (3), if such a decomposition $p = r \circ h$ exists, then $s_r \circ h = \frac{s_p}{s_h} = 1$, which implies that r is pseudo linear. We write differential polynomials p, r, h as the sum of total degree homogeneous parts:

$$\begin{aligned} p &= P_d + P_{d-1} + \cdots + P_2 + P_1 \\ r &= R_{d_1} + R_{d_1-1} + \cdots + R_2 + R_1 \\ h &= H_{d_2} + H_{d_2-1} + \cdots + H_2 + H_1 \end{aligned} \tag{11}$$

where d, d_1, d_2 denote the total degrees of p, r, h respectively. Notice that p, r, h have no terms in $\mathcal{K}$. Since r is pseudo linear, $R_1 \neq 0$ and $o_{R_i} < o_{R_1}$ for $2 \leq i \leq d_1$.

For any differential polynomial f, denote $\boxed{f}_k$ to be the sum of the monomials included in f with total degree k. We consider two cases: $P_1 \neq 0$ and $P_1 = 0$.

4.1 Case 1: $P_1 \neq 0$

Use the notations introduced in (11). We assume $d_1 < d_2$ (the case $d_1 \geq d_2$ is similar). Comparing the sum of the monomials with total degree l $(1 \leq l \leq d)$ in $p = r \circ h$, we have

$$\begin{cases} P_1 = R_1 \circ H_1 \\ P_2 = R_1 \circ H_2 + R_2 \circ H_1 \\ \vdots \\ P_k = R_1 \circ H_k + R_k \circ H_1 + \sum\limits_{1<i<k} \boxed{R_i \circ \left(\sum\limits_{1\leq j\leq k-1} H_j \right)}_k & (2 \leq k \leq d_1) \\ \vdots \\ P_s = R_1 \circ H_s + \sum\limits_{1<i\leq d_1} \boxed{R_i \circ \left(\sum\limits_{1\leq j\leq s-1} H_j \right)}_{d_2} & (d_1 < s \leq d_2) \\ \vdots \\ P_v = \boxed{\sum\limits_{1<i\leq d_1} R_i \circ \left(\sum\limits_{1\leq j\leq v} H_j \right)}_v & (v > d_2) \\ \vdots \\ P_d = R_{d_1} \circ H_{d_2} \end{cases} \tag{12}$$

The basic idea of our algorithm is as follows: find R_1, H_1 from the first equation in (12) and substitute R_1, H_1 into the second equation to obtain $P_2' = P_2 - (R_1 \circ H_2 + R_2 \circ H_1)$. From $P_2' = 0$, we obtain a system of linear equations about the coefficients of R_2 and H_2. Solving this linear equation system, we may obtain R_2 and H_2. Substituting R_1, H_1, R_2, H_2 into the third equation, we may find R_3, H_3, and so on. To make this idea into an algorithm, we need to solve the following problems.

Q1. How to determine d_1 and d_2?
Q2. How to determine R_1, H_1 from the first equation of (12), that is, how to factor LODOs over $\mathcal{K}$?
Q3. Can $R_k, H_k (2 \le k \le \min\{d_1, d_2\})$ be determined uniquely by R_1 and H_1? As mentioned before, we need to determine R_k and H_k from $R_1 \circ H_k + R_k \circ H_1 = P_k'$ where P_k' is given. This will lead to a differential equation system in the coefficients of H_k and R_k. If the answer is no, then some new parameters over $\mathcal{K}$ will be introduced and we may face the difficult problem of solving algebraic differential equations about the coefficients in the next step.
Q4. If $R_k, H_k (1 \le k \le \min\{d_1, d_2\})$ can be determined uniquely by R_1 and H_1, how to find them efficiently?
Q5. If $d_1 < d_2$, we will obtain $R_i, H_i (1 \le i \le d_1)$ firstly and then compute $H_j (d_1 < j \le d_2)$ from (12). We need to know whether $H_j (d_1 < j \le d_2)$ can be determined uniquely. Similarly, how about the case of $d_1 \ge d_2$?

For problem Q1, by (6) we have $d = d_1 d_2$. It is obvious that

$$d_2 = \operatorname{tdeg}(h) \ge \operatorname{tdeg}(S) + 1 \tag{13}$$

where S is the possible separant of h introduced in Sect. 3. So we will search all possible pairs (d_1, d_2) satisfying these two conditions.

For problem Q2, there exist practical algorithms to factor an LODO over certain fields, such as the field of rational functions [29], the field of formal Laurent series [30], etc. Here we need an algorithm to find all decompositions of P_1. We can use Beke's algorithm directly or the method given in [26]. When we compute R_1 and H_1 from $P_1 = R_1 \circ H_1$, (R_1, H_1) could contain parameters and have many forms as shown by Example 4.1, but the number of the parameters and the enumerations of decompositions are both finite [26]. We will find all possible enumerations of (R_1, H_1) and use each pair to find possible solutions of h and r in (11).

Example 4.1 Let $\mathcal{K} = \mathbb{Q}(t)$. $y_2 \in \mathcal{K}\{y\}$ has only the following two decompositions

$$y_2 = y_1 \circ y_1 = \left(y_1 + \frac{1}{t+c} y\right) \circ \left(y_1 - \frac{1}{t+c} y\right)$$

where c is an arbitrary constant taking values in $\mathbb{Q}$.

It is well known that when $\mathcal{K} = C_{\mathcal{K}}$ is a constant field, factoring an LODO is equivalent to factoring a univariate polynomial over $\mathcal{K}$. This is stated as the following lemma.

Lemma 4.2 *Let* $q = \sum_{k=0}^{n} a_k y_k$, $u = \sum_{i=0}^{m} b_i y_i$, $v = \sum_{j=0}^{n-m} c_j y_j$, *where* $m \leq n$ *and* $a_k, b_i, c_j \in C_{\mathcal{K}}$. *Then* $q = u \circ v$ *if and only if* $\hat{q} = \hat{u}\hat{v}$, *where* $\hat{q} = \sum_{k=0}^{n} a_k y^k$, $\hat{u} = \sum_{i=0}^{m} b_i y^i$, $\hat{v} = \sum_{j=0}^{n-m} c_j y^j$.

Theorem 4.5 below answers problem Q3 affirmatively. To prove it, we first give a lemma.

Lemma 4.3 *If* $y_1 \circ r = s \circ u$, *where* r, s, u *are total degree homogeneous and* $\mathrm{tdeg}(s) > 1$, *then there exists a differential polynomial* s' *such that* $s = y_1 \circ s'$.

Proof Suppose that the result is not valid and let s be a counterexample whose leading term has the lowest rank. Since $y_1 \circ r = s \circ u$, we have $d_s = 1$. Let $s = s_1 y_{o_s} + s_0$ with $o_{s_0} < o_s$. We consider two cases:
Case 1 $o_{s_1} < o_s - 1$. Then

$$y_1 \circ \left(r - (s_1 \cdot y_{o_s-1}) \circ u\right) = \left(s - y_1 \circ (s_1 \cdot y_{o_s-1})\right) \circ u = \left(s_0 - (y_1 \circ s_1) \cdot y_{o_s-1}\right) \circ u = f_1 \circ u$$

where $f_1 = s_0 - (y_1 \circ s_1) \cdot y_{o_s-1}$. If f_1=0, then we have $r = (s_1 \cdot y_{o_s-1}) \circ u$ and so $s = y_1 \circ (s_1 \cdot y_{o_s-1})$, which contradicts the hypothesis. If $f_1 \neq 0$, then f_1 is total degree homogeneous ($\mathrm{tdeg}(f_1) = \mathrm{tdeg}(s) > 1$) and $o_{f_1} < o_s$. By the hypothesis, there is a differential polynomial f_1' such that $f_1 = y_1 \circ f_1'$. So $s = y_1 \circ (s_1 y_{o_s-1} + f_1')$ which contradicts the hypothesis, too.
Case 2 $o_{s_1} = o_s - 1$. Let $s_1 = i_{s_1} \cdot y_{o_s-1}^{d_{s_1}} + s_1'$. Then we have

$$y_1 \circ \left(r - \left(\frac{1}{d_{s_1}+1} i_{s_1} y_{o_s-1}^{d_{s_1}+1}\right) \circ u\right) = \left(s_1' y_{o_s} + s_0 - \frac{1}{d_{s_1}+1}(y_1 \circ i_{s_1}) y_{o_s-1}^{d_{s_1}+1}\right) \circ u = f_2 \circ u$$

where $f_2 = s_1' y_{o_s} + s_0 - \frac{1}{d_{s_1}+1}(y_1 \circ i_{s_1}) y_{o_s-1}^{d_{s_1}+1}$. If $f_2 = 0$, then $s = y_1 \circ \frac{1}{d_{s_1}+1} i_{s_1} y_{o_s-1}^{d_{s_1}+1}$, which contradicts the hypothesis. If $f_2 \neq 0$, then f_2 is still total degree homogeneous ($\mathrm{tdeg}(f_2) = \mathrm{tdeg}(s)$) and the leading term of f_2 is of lower rank than s. By the hypothesis, there is a differential polynomial f_2' such that $f_2 = y_1 \circ f_2'$. Then $s = y_1 \circ (\frac{1}{d_{s_1}+1} i_{s_1} y_{o_s-1}^{d_{s_1}+1} + f_2')$, which contradicts the hypothesis. □

Corollary 4.4 *If* $y_m \circ r = s \circ u$, *where* r, s, u *are total degree homogeneous* ($\mathrm{tdeg}(s) > 1$) *and* m *is a positive integer, then there exists a differential polynomial* s' *such that* $s = y_m \circ s'$.

Proof Lemma 4.3 proves the case for $m = 1$. Assume that it is correct for $m = k - 1$. When $m = k$, if we have $y_k \circ r = y_1 \circ (y_{k-1} \circ r) = s \circ u$, then by Lemma 4.3, there exists an s_1 such that $s = y_1 \circ s_1$. Since $k \geq 2$, $y_{k-1} \circ r$ contains no terms in $\mathcal{K}$. Then $y_{k-1} \circ r = s_1 \circ u$. By the hypothesis for $m = k - 1$, there exists an s' such that $s_1 = y_{k-1} \circ s'$. So we have $s = y_k \circ s'$. This proves the case for $m = k$. □

Theorem 4.5 *Let R_1, H_1, P_k be total degree homogeneous differential polynomials in $\mathcal{K}\{y\}$ with total degree $1, 1, k$ respectively. If there exist k-total degree homogeneous differential polynomials R_k, H_k over $\mathcal{K}$ such that $o_{R_k} < o_{R_1}$ and $P_k' = R_1 \circ H_k + R_k \circ H_1$, then R_k and H_k are unique.*

Proof If (R_{k1}, H_{k1}), (R_{k2}, H_{k2}) both satisfy the conditions, then

$$P_k' = R_1 \circ H_{k1} + R_{k1} \circ H_1 = R_1 \circ H_{k2} + R_{k2} \circ H_1. \tag{14}$$

Since R_1 is linear, from (14) we have

$$R_1 \circ (H_{k1} - H_{k2}) + (R_{k1} - R_{k2}) \circ H_1 = 0.$$

Since $R_{k1} - R_{k2}$ and $H_{k1} - H_{k2}$ are still k-total degree homogeneous, we need only to prove that there exist no nonzero R_k and H_k such that $R_1 \circ H_k = R_k \circ H_1$.

Suppose such H_k, R_k exist. Let the leading terms of R_1, H_1 be $r_m y_m, h_n y_n$, the sums of monomials included in R_k, H_k with differential degree $\mathrm{ddeg}(R_k), \mathrm{ddeg}(H_k)$ be $\tilde{R}_k, \tilde{H}_k$ respectively, and

$$\tilde{R}_k = \sum_I a_I y^{i_0} y_1^{i_1} \cdots y_{o_1}^{i_{o_1}}, \quad \tilde{H}_k = \sum_J b_J y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}} \tag{15}$$

where $o_1 = o_{\tilde{R}_k}$, $o_2 = o_{\tilde{H}_k}$, I and J are in two index sets. Notice that differential degrees have the following properties: $\mathrm{ddeg}(p) < \mathrm{ddeg}(q) \Rightarrow \mathrm{ddeg}(y_1 \circ p) < \mathrm{ddeg}(y_1 \circ q)$ and $\mathrm{ddeg}(p) < \mathrm{ddeg}(y_1 \circ p)$. Comparing the sum of monomials with maximal differential degree in both sides of $R_1 \circ H_k = R_k \circ H_1$, since R_1, H_1 are linear, we have

$$r_m \sum_J b_J \left(y_m \circ y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}} \right) = h_n^k \sum_I \left(a_I y^{i_0} y_1^{i_1} \cdots y_{o_1}^{i_{o_1}} \right) \circ y_n. \tag{16}$$

Let W be the set of all the coefficients r_m, h_n, a_I, b_J, $W_1 = \{v : v \in W \text{ and } v \notin C_{\mathcal{K}}\}$ the set of the coefficients not in the constant field of $\mathcal{K}$. There are two cases:

(1) $W_1 = \emptyset$. Then from (16) we have:

$$y_m \circ \left(r_m \sum_J b_J y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}} \right) = \left(\sum_I a_I y^{i_0} y_1^{i_1} \cdots y_{o_1}^{i_{o_1}} \right) \circ (h_n y_n).$$

That is, $y_m \circ (r_m \tilde{H}_k) = \tilde{R}_k \circ (h_n y_n)$. By Corollary 4.4, there is a differential polynomial Y_1 such that $\tilde{R}_k = y_m \circ Y_1$. So we have $o_{R_k} \geq o_{\tilde{R}_k} \geq m = o_{R_1}$, which contradicts the hypothesis.

(2) $W_1 \neq \emptyset$. Then each $v \in W_1$ is transcendental over $C_{\mathcal{K}}$. Otherwise, let $\alpha(y) = \sum_{0 \leq i \leq l} h_i y^i$ be the minimal polynomial of v. Then $(\sum_{1 \leq i \leq l} i h_i v^{i-1}) v_{(1)} = 0$, where $v_{(1)}$ denotes the derivative of v. But neither of the product factor equals 0, a contradiction. Let $\{c_1, c_2, \ldots, c_e\}$ be a set of transcendental basis of W over $C_{\mathcal{K}}$. Then each

$v \in W$ is algebraic over the field $C_{\mathcal{K}}(c_1, \ldots, c_e)$. Let $\beta_v(z) = \sum_{0 \le i \le l_v} \beta_{vi}(c_1, \ldots, c_e) z^i$ be the "minimal" polynomial of v (without the request of monic), where $\beta_{vi}(c_1, \ldots, c_e)$ is a polynomial in $c_1, \ldots, c_e$ with coefficients in $C_{\mathcal{K}}$. Since c_j is transcendentally independently over $C_{\mathcal{K}}$ and (16) is an identity in $C_{\mathcal{K}}\{y\}$ for c_j, replacing c_j with some nonzero elements $k_j (1 \le j \le e)$ in $C_{\mathcal{K}}$, we have

$$\tilde{r}_m \sum_J \tilde{b}_J \left(y_m \circ y^{j_0} y_1^{j_1} \cdots y_{o2}^{j_{o2}} \right) = \tilde{h}_n^k \sum_I \left(\tilde{a}_I y^{i_0} y_1^{i_1} \cdots y_{o1}^{i_{o1}} \right) \circ y_n$$

where each of the coefficient $\tilde{v} (v \in W)$ is an algebraic element over $C_{\mathcal{K}}$ satisfying the equation $\sum_{0 \le i \le l_v} \beta_{vi}(k_1, \ldots, k_e) \tilde{v}^i = 0$. Obviously, we can choose k_j such that $\tilde{r}_m \ne 0$. So we have

$$y_m \circ \left(\tilde{r}_m \sum_J \tilde{b}_J y^{j_0} y_1^{j_1} \cdots y_{o2}^{j_{o2}} \right) = \left(\sum_I \tilde{a}_I y^{i_0} y_1^{i_1} \cdots y_{o1}^{i_{o1}} \right) \circ (\tilde{h}_n y_n). \tag{17}$$

We should explain that the derivative appears in (17) is the differential operator of $\tilde{\mathcal{K}}$, but not that of $\mathcal{K}$, where $\tilde{\mathcal{K}}$ is the differential overfield of $\mathcal{K}$ generated by adding all the algebraic elements over $C_{\mathcal{K}}$ to $\mathcal{K}$. The derivative on $\tilde{\mathcal{K}}$ is uniquely determined by that on $\mathcal{K}$ and the constant field of $\tilde{\mathcal{K}}$ is the algebraic closure of $C_{\mathcal{K}}$. Again by Corollary 4.4, there is a differential polynomial $Y_2 \in \tilde{\mathcal{K}}\{y\}$ such that $0 \ne R' = \sum_I \tilde{a}_I y^{i_0} y_1^{i_1} \cdots y_{o1}^{i_{o1}} = y_m \circ Y_2$. So we have $o_{R_k} \ge o_{\tilde{R}_k} \ge o_{R'} \ge m = o_{R_1}$, which contradicts the hypothesis.

□

Remark Theorem 4.5 asserts that no new parameters will be introduced in both of R_k and H_k $(1 < k \le d_1)$. In particular, if there are no parameters in the decomposition $P_1 = R_1 \circ H_1$, then there are no parameters in H_k, R_k. If R_1, H_1 contain parameters, H_k, R_k may contain these parameters too (and they can be determined uniquely by the parameters). Furthermore, the identical equation $R_1 \circ H_k + R_k \circ H_1 = P_k'$ may add new algebraic constraints on the parameters in R_1 and H_1. □

As mentioned before, to determine R_k and H_k from $R_1 \circ H_k + R_k \circ H_1 = P_k'$, it seems that we need to solve a linear differential system. In fact, we can reduce the problem to solving an algebraic parametric linear equation system by comparing the differential degree, and hence answers problem Q4.

Proposition 4.6 *Use notations introduced in* (11)*. For each* $k > 1$*, we may determine* R_k *and* H_k *from* R_1*,* H_1 *and* p *by solving an algebraic parametric linear equation system in the coefficients of* R_k *and* H_k*.*

Proof Suppose that R_i, H_i $(1 \le i \le k-1)$ have been obtained. Based on (12), we hope to find R_k, H_k from

$$R_1 \circ H_k + R_k \circ H_1 = P_k' = P_k - \sum_{1 < i < k} \boxed{R_i \circ \left(\sum_{1 \le j \le k-1} H_j \right)}_k . \tag{18}$$

Denote $\tilde{R}_k$, $\tilde{H}_k$, $\tilde{P}'_k$ to be the sums of the terms in R_k, H_k, $\tilde{P}'_k$ with maximal differential degree respectively. Using the notions in (15) and comparing the sum of the terms with maximal differential degree on both sides, we have:

$$r_m \sum_J b_J \left(y_m \circ y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}}\right) + h_n^k \sum_I \left(a_I y^{i_0} y_1^{i_1} \cdots y_{o_1}^{i_{o_1}}\right) \circ y_n - \tilde{P}'_k = 0. \quad (19)$$

Notice that in the identity (19), all the coefficients a_I, b_J are algebraically linear and the parameters may appear in the coefficients of P'_k. We can obtain $\tilde{R}_k$, $\tilde{H}_k$ by solving the algebraic parametric linear system (by Theorem 4.5, $\tilde{R}_k$ and $\tilde{H}_k$ can be expressed by the parameters and further more, we get more constraints on the parameters).

After $\tilde{R}_k$ and $\tilde{H}_k$ are found, we may let $R_k = \tilde{R}_k + R'_k$ and $H_k = \tilde{H}_k + H'_k$ and substitute them into (18). Since R_1 is linear, we have

$$R_1 \circ H'_k + R'_k \circ H_1 = P'_k - R_1 \circ \tilde{H}_k + \tilde{R}_k \circ H_1.$$

Repeat the above procedure for the above equation, we can finally obtain R_k, H_k. □

Problem Q5 is similar to problems Q3 and Q4. When $d_1 \geq d_2$, we may obtain all $H'_i s$ first with Proposition 4.6. Then we have obtained the right decomposition factor h of p and the corresponding left decomposition factor is certainly uniquely determined by the coefficients of h. When $d_1 < d_2$, we obtain all $R'_i s$ and $H_i (1 \leq i \leq d_1)$ firstly. From (12) we have

$$R_1 \circ H_s = P'_s = P_s - \sum_{1<i\leq d_1} \boxed{R_i \circ \left(\sum_{1\leq j\leq s-1} H_j\right)}_{d_2} \quad (d_1 < s \leq d_2). \quad (20)$$

If there is a differential polynomial H'_s such that $R_1 \circ H'_s = P'_s$, then $R_1 \circ (H_s - H'_s) = 0$ and $H_s = H'_s$. So h is uniquely determined. Denote $\tilde{H}_s$, $\tilde{P}'_s$ to be the sums of the terms in H_s, $\tilde{P}'_s$ with maximal differential degree respectively. Using the notions in (15) and comparing the sum of the terms with maximal differential degree on both sides of (20), we have

$$r_m \sum_J b_J (y_m \circ y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}}) - \tilde{P}'_k = 0 \quad (21)$$

where $m = o_{R_1}$, $h_m = i_{R_1}$. Similar to Proposition 4.6, in order to find $\tilde{H}_s$, we need only to solve a linear equation system and this system has a unique solution. It is now clear that we can treat problem Q4 and problem Q5 uniformly by setting $R_i = 0, i = d_2 + 1, \ldots, d_1$ when $d_1 > d_2$ and $H_i = 0, i = d_1 + 1, \ldots, d_2$ when $d_2 > d_1$ in problem Q5.

The above analysis leads to the following algorithm.

Algorithm 4.7 *Input: Differential polynomials R_1, H_1, p as shown in (11), integers d_1, d_2 satisfying $d_1 d_2 = d$ and $\mathbb{P}_1 = \{p_1 = 0, \ldots, p_u = 0\}$ which denotes the constraints on the parameters appearing in R_1. R_1, H_1 are linear, $i_R = 1$.*

Output: A differential polynomial h and a set $\mathbb{P}$, if they exist. If p has a right decomposition factor of form (11), then it must be h. h may be expressed by parameters and the parameters satisfy the algebraic equations in $\mathbb{P}$.

S1 Use notations introduced in (11). Let $m = o_{R_1}, n = o_{H_1}, h_n = i_{H_1}, o_2 = o_p - m$, $\mathbb{P} = \mathbb{P}_1, h = 0$. For $k = 2, \ldots, d_2$, set $R_k = H_k = 0$.

S2 Assume that the differential polynomials $R_i, H_i (1 \le i \le k-1)$ has been obtained. Let $P'_k = P_k - \sum_{1<i<k} \boxed{R_i \circ (\sum_{1\le j\le k-1} H_j)}_k$, compute R_k, H_k by steps S3 to S6.

S3 If $P'_k = 0$, return R_k, H_k. If $\mathrm{ddeg}(P'_k) < \max\{m, n\cdot k\}$, then R_k, H_k do not exist, terminate the algorithm. Let $\tilde{P}'_k$ be the sum of the terms included in P'_k with maximal differential degree.

S4 Let

$$T_1 = \{(i_0, i_1, \ldots, i_{m-1}) : \sum_{0\le s\le m-1} i_s = k, \sum_{0\le s\le m-1} s\cdot i_s = \mathrm{ddeg}(P'_k) - m\}$$

$$T_2 = \{(j_0, j_1, \ldots, j_{o_2}) : \sum_{0\le l\le o_2} j_l = k, \sum_{0\le l\le o_2} l\cdot j_l = \mathrm{ddeg}(P'_k) - n\cdot k\}$$

$$\tilde{R}_k = \sum_{I\in T_1} a_I y^{i_0} y_1^{i_1}\cdots y_{m-1}^{i_{m-1}}, \quad \tilde{H}_k = \sum_{J\in T_2} b_J y^{j_0} y_1^{j_1}\cdots y_{o_2}^{j_{o_2}}.$$

Then we have

$$\sum_{J\in T_2} b_J \left(y_m \circ y^{j_0} y_1^{j_1}\cdots y_{o_2}^{j_{o_2}}\right) - h_n^k \sum_{I\in T_1} \left(a_I y^{i_0} y_1^{i_1}\cdots y_{m-1}^{i_{m-1}}\right) \circ y_n - \tilde{P}'_k = 0. \tag{22}$$

If T_1 (T_2) is empty, then $R_k := 0$ ($H_k := 0$) and the corresponding sum is defined to be zero.

S5 Comparing the coefficients of the monomials on both sides of (22), we obtain a parametric linear equation system LS about a_I and b_J. Solve this system with methods from [24,6].
If LS does not have a solution, then R_k, H_k do not exist and we terminate the algorithm; otherwise, we obtain $\tilde{R}_k, \tilde{H}_k$ and $\mathbb{T}$, where $\mathbb{T}$ is the new constraints on the parameters. $\mathbb{P} := \mathbb{P} \bigcup \mathbb{T}$.

S6 $R_k := R_k + \tilde{R}_k$, $H_k := H_k + \tilde{H}_k$, and $P'_k := P'_k - R_1 \circ \tilde{H}_k - \tilde{R}_k \circ H_1$, go to S3.

S7 For $k = 2, \ldots, d_2$, compute R_k, H_k one by one by step S2 (If $d_1 < d_2$, let $R_k = 0, d_1 + 1 \le k \le d_2$). Output $h = H_1 + H_2 + \cdots + H_{d_2}$ and $\mathbb{P}$.

Proof of correctness of Algorithm 4.7 Since $R_1 \circ H_k + R_k \circ H_1 = P'_k$ and r is pseudo linear, $\mathrm{ddeg}(P'_k) = \max\{\mathrm{ddeg}(R_1 \circ H_k), \mathrm{ddeg}(R_k \circ H_1)\} = \max\{o_{R_1} + \mathrm{ddeg}(H_k), k\cdot o_{H_1} + \mathrm{ddeg}(R_k)\}$. Hence if $\mathrm{ddeg}(P'_k) < \max\{m, n\cdot k\}$, then R_k, H_k do not exist in step S2.

The correctness of other steps follow directly from the proof of Proposition 4.6. □

4.2 Case 2: $P_1 = 0$

Let $k = \min\{i : P_i \neq 0\}$. Then $k = \min\{i : H_i \neq 0\}$. Comparing the sum of the monomials with total degree l $(k \leq l \leq d)$ in $p = r \circ h$, we have

$$\begin{cases} P_k = R_1 \circ H_k \\ P_{k+1} = R_1 \circ H_{k+1} \\ \vdots \\ P_{2k-1} = R_1 \circ H_{2k-1} \\ P_{2k} = R_1 \circ H_{2k} + R_2 \circ H_k \\ P_{2k+1} = R_1 \circ H_{2k+1} + \boxed{R_2 \circ (H_k + H_{k+1})}_{2k+1} \\ \vdots \\ P_d = R_{d_1} \circ H_{d_2} \end{cases} \tag{23}$$

Generally, for $m, n \leq \min\{k \cdot d_1, d_2\}$, we have:

$$P_m = R_1 \circ H_m + \sum_{2 \leq l \leq [\frac{m}{k}]} \boxed{R_l \circ \left(\sum_{k \leq j \leq [\frac{m}{l}]} H_j \right)}_m \quad (k \nmid m) \tag{24}$$

$$P_n = R_1 \circ H_n + R_{\frac{n}{k}} \circ H_k + \sum_{2 \leq l < \frac{n}{k}} \boxed{R_l \circ \left(\sum_{k \leq j \leq [\frac{n}{l}]} H_j \right)}_n \quad (k \mid n) \tag{25}$$

where $[\frac{m}{l}]$ denotes the maximal integer not larger than $\frac{m}{l}$.

We will obtain R_1 from the first k equations $R_1 \circ H_j = P_j$, $(k \leq j \leq 2k-1)$ by decomposing linear differential polynomials with coefficients in $\mathcal{K}$ and obtain other R_i, H_j by solving linear systems similar to Algorithm 4.7. The selection of (d_1, d_2) and the uniqueness of R_k and H_k are quite similar to that in the case $P_1 \neq 0$. When R_1 is given, the uniqueness for H_j $(k \nmid j)$ is obvious by (24) and the uniqueness for $R_{\frac{j}{k}}$, $H_j (k \mid j)$ in (25) is guaranteed by Theorem 4.8.

Theorem 4.8 *R_1, P_{ik}, H_k are total degree homogeneous differential polynomials with total degree $1, i \cdot k, k$ respectively. If there exist an $i-$total degree homogeneous differential polynomial R_i and an $i \cdot k$-total degree homogeneous differential polynomial H_{ik} such that $o_{R_i} < o_{R_1}$ and $P_{ik} = R_1 \circ H_{ik} + R_i \circ H_k$, then H_{ik} and R_i are unique.*

Proof Since R_1 is linear, we need only to prove that there exist no nonzero H_{ik} and R_i such that $R_1 \circ H_{ik} = R_i \circ H_k$, similar to Theorem 4.5. Suppose that such H_{ik}, R_i exist. Let $m = o_{R_1}$, the sum of the monomials included in R_i, H_k, H_{ik} with maximal differential degrees be $\tilde{R}_i, \tilde{H}_k, \tilde{H}_{ik}$ respectively and

$$\tilde{R}_i = \sum_I a_I y^{i_0} y_1^{i_1} \cdots y_{o_1}^{i_{o_1}}, \quad \tilde{H}_k = \sum_J b_J y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}}, \quad \tilde{H}_{ik} = \sum_L c_L y^{l_0} y_1^{l_1} \cdots y_{o_3}^{l_{o_3}}$$

where $o_1 = o_{\tilde{R}_i}$, $o_2 = o_{\tilde{H}_k}$, $o_3 = o_{\tilde{H}_{ik}}$ and I, J, L are in three index sets. Comparing the sum of monomials with maximal differential degree of both sides, we have

$$r_m \sum_L c_L \left(y_m \circ y^{l_0} y_1^{l_1} \cdots y_{o_3}^{j_{o_3}}\right) = \sum_I a_I \left(\sum_J b_J y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}}\right)^{i_0} \times \left(\sum_J b_J y_1 \circ y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}}\right)^{i_1} \cdots \left(\sum_J b_J y_{o_1} \circ y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}}\right)^{i_{o_1}} \tag{26}$$

where r_m is the initial of R_1. Similar to Theorem 4.5, we can get an identity

$$\bar{r}_m \sum_L \bar{c}_L \left(y_m \circ y^{l_0} y_1^{l_1} \cdots y_{o_3}^{j_{o_3}}\right) = \sum_I \bar{a}_I \left(\sum_J \bar{b}_J y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}}\right)^{i_0} \times \left(\sum_J \bar{b}_J y_1 \circ y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}}\right)^{i_1} \cdots \left(\sum_J \bar{b}_J y_{o_1} \circ y^{j_0} y_1^{j_1} \cdots y_{o_2}^{j_{o_2}}\right)^{i_{o_1}} \tag{27}$$

in $\tilde{\mathcal{K}}\{y\}$ by replacing some of the coefficients with values in $C_{\mathcal{K}}$ such that the left side of (26) will not become zero, where $\tilde{\mathcal{K}}$ is the differential extension field of $\mathcal{K}$ generated by adding all algebraic elements over $C_{\mathcal{K}}$ to $\mathcal{K}$ and the coefficients $\bar{r}_m, \bar{a}_I, \bar{b}_J, \bar{c}_L$ are in the algebraic closure of $C_{\mathcal{K}}$. The derivative appearing in (27) is the differential operator on $\tilde{\mathcal{K}}$. So we have

$$y_m \circ \left(\bar{r}_m \sum_L \bar{c}_L y^{l_0} y_1^{l_1} \cdots y_{o_3}^{j_{o_3}}\right) = \left(\sum_I \bar{a}_I y^{i_0} y_1^{i_1} \cdots y_{o_1}^{i_{o_1}}\right) \circ \left(\sum_J \bar{b}_J y^{j_0} y_1^{j_1}\right).$$

Let $R' = \sum_I \bar{a}_I y^{i_0} y_1^{i_1} \cdots y_{o_1}^{i_{o_1}}$. Then by Corollary 4.4, there exists a differential polynomial $Y_3 \in \tilde{\mathcal{K}}\{y\}$ such that $R' = y_m \circ Y_3$. So $o_{R_i} \geq o_{\tilde{R}_i} \geq o_{R'} \geq m = o_{R_1}$, which contradicts the hypothesis. □

Now we show how to obtain R_1 from $P_j = R_1 \circ H_j (k \leq j \leq 2k-1)$, which is a key step of our algorithm. We consider the following general problem: *given a total degree homogeneous differential polynomial q (*$\mathrm{tdeg}(q) > 1$*), how to find the decomposition $q = s \circ r$ such that r is linear?* The solution to this problem is given in Proposition 4.10. We first prove a lemma.

Lemma 4.9 *If $l \circ q = u \circ y_m$, where l, q, u are total degree homogeneous, l is linear,* $\mathrm{tdeg}(q) > 1$*, and m is a positive integer, then there exists a differential polynomial q' such that $q = q' \circ y_m$.*

Proof Let k be the minimal subscript such that y_k appears in q and $q = q_d y_k^d + q_{d-1} y_k^{d-1} + q^{(1)}$, where d is the degree of q in y_k and $q^{(1)}$ is the sum of the monomials in q with degrees in y_k less than $d-1$. We need to prove that $k \geq m$. There are two cases:
Case 1 $l \circ q_d \neq 0$. Then $l \circ q = (l \circ q_d) y_k^d + q^{(2)}$, where the degree of $q^{(2)}$ in y_k is less than d. So y_k appears in $l \circ q$.
Case 2 $l \circ q_d = 0$. Then $q_d \in \mathcal{K}$, and $d \geq 2$ since $\mathrm{tdeg}(q) > 1$ and q is homogenous. Let $n = o_l$ be the order of l and $l = \sum_{0 \leq i \leq n} a_i y_i$ $(a_i \in \mathcal{K})$. The coefficient of y_k^{d-1} in $l \circ q$ is $l \circ q_{d-1} + (\sum_{1 \leq i \leq n} a_i y_{i-1}) \circ (d q_d y_{k+1})$.

Note that the order of q_{d-1} is greater than k, so the order of $l \circ q_{d-1}$ is greater than that of $(\sum_{1 \leq i \leq n} a_i y_{i-1}) \circ (d q_d y_{k+1})$ and $l \circ q_{d-1} + (\sum_{1 \leq i \leq n} a_i y_{i-1}) \circ (d q_d y_{k+1}) \neq 0$. y_k appears in $l \circ q$, too.

Since y_k appears in $l \circ q$ and the minimal subscript j with y_j appearing in $l \circ q = u \circ y_m$ is not less than m, we have $k \geq m$. □

Proposition 4.10 *Let q be a total degree homogeneous differential polynomial. Then we can construct a linear differential polynomial l such that if s is a linear left decomposition factor of q then s is a left decomposition factor of l.*

Proof Assume that q can be decomposed as $q = s \circ r$, where s is linear and r is total degree homogeneous. Let $v = \min\{i : y_i \text{ appears in } q\}$, that is, there exists a differential polynomial q' such that $q = q' \circ y_v$ and y appears in q'. By Lemma 4.9, there exists a differential polynomial r' such that $r = r' \circ y_v$ and we have $q' = s \circ r'$. So we need only to consider the case of $v = 0$.

Let $a = \deg_y q$, $e = \deg_y r$, $n = o_s$. We can write q, s, r as follows:

$$\begin{aligned} q &= q_a y^a + \cdots + q_1 y + q_0 \\ s &= s_n y_n + \cdots + s_1 y_1 + s_0 y \\ r &= r_e y^e + \cdots + r_1 y + r_0 \end{aligned}$$

where $o_{r_e} \geq 1$ or $r_e = 1$ $\left(\text{if } r_e \in \mathcal{K}, \text{then} s \circ r = (s \circ r_e y) \circ (\frac{r}{r_e})\right)$. We have

$s \circ r = (s \circ r_e) y^e + (s \circ r_{e-1} + s_n y_{n-1} \circ (e r_e y_1) + s_{n-1} y_{n-2} \circ (e r_e y_1) + \cdots + s_1 (e r_e y_1)) y^{e-1} + u_1$

$$= \begin{cases} (s \circ r_e) y^e + u_2 & o_{r_e} \geq 1 \\ s_0 y^e + (s \circ (r_{e-1} + ey) - e s_0 y)\, y^{e-1} + u_3 & r_e = 1 \text{ and } s_0 \neq 0 \\ (s \circ (r_{e-1} + ey))\, y^{e-1} + u_4 & r_e = 1 \text{ and } s_0 = 0 \end{cases}$$

where u_1, u_3, u_4 are of degree lower than $e-1$ in y and u_2 is of degree lower than e in y.

From $q = s \circ r$, we have

1. If $q_a \notin \mathcal{K}$, then s satisfies $s \circ r_e = q_a$ or $s \circ (r_{e-1} + ey) = q_a$. In this case, s is a linear left decomposition factor of q_a. So, we will find the linear left decomposition factors of q_a.

2. If $q_a \in \mathcal{K}$, then $s_0 = q_a$, $e = a$ and $q_{a-1} = s \circ (r_{e-1} + ey) - es_0 y$. So $q_{a-1} + aq_a y = s \circ (r_{e-1} + ey)$ and we will consider the decomposition of the linear differential polynomial $q_{a-1} + aq_a y$.

That is, if s is a linear left decomposition factor of q, then s is a linear left decomposition factor of q_a or $q_{a-1} + aq_a y$. If $q_a \notin \mathcal{K}$, repeat the above analysis for q_a, we will finally obtain a linear differential polynomial l satisfying the condition of the proposition. □

The above analysis leads to the following algorithm.

Algorithm 4.11 *Input: k total degree homogeneous differential polynomials* $P_k, \ldots, P_{2k-1}$.

Output: a linear differential polynomial l such that if s is a common linear left decomposition factor of $P_i (k \leq i \leq 2k-1)$*, then s is a left decomposition factor of l.*

S1 For i from k to $2k-1$, do steps S2 and S3 to construct the linear differential polynomial L_i.
S2 $v = \min\{i : y_i \text{ appears in } P_i\}$. Find P_i' such that $P_i = P_i' \circ y_v$, that is, replace y_i with y_{i-v} in P_i. Let $a = \deg_y P_i'$, $P_{i,a}$ and $P_{i,a-1}$ be the coefficient of y^a and y^{a-1} in P_i' respectively.
S3 If $P_{i,a} \in \mathcal{K}$, we have found $L_i := P_{i,a-1} + aP_{i,a}y$; otherwise, $P_i := P_{i,a}$, go to S2.
S4 Compute the greatest common left decomposition factor s of all L_i by left Euclid algorithm [15] and output s.

4.3 Solve the pseudo linear case

Now we could solve the problem proposed at the beginning of this section. But remember that our aim is to compute the decomposition of f and the right decomposition factor of p is just a possible right decomposition factor of f.

Algorithm 4.12 *Input: Differential polynomials* f, p, S *with* $d_p = 1, i_p = S$.

Output: The set of differential polynomials g, h *such that* $f = g \circ h$*, and* $s_h = S$*. If such g and h do not exist, return the empty set.*

S1 Let $d = \text{tdeg}(p)$. Write p as $p = \sum_{1 \leq i \leq d} P_i$ where $P_i = \boxed{P}_i$.
S2 Let $k = \min\{i : P_i \neq 0\}$. If $k = 1$, let $s = P_1$. If $k > 1$, execute Algorithm 4.11 with input $P_k, \ldots, P_{2k-1}$ and with output s.
S3 Let Δ be the set of all $(R_1, \mathbb{P}_1)$ where R_1 is a left decomposition factor of s with initial 1 and $\mathbb{P}_1$ the corresponding constraints for the parameters in R_1. If $\Delta = \emptyset$, the algorithm terminates. If there is no parameter in R_1 or there is no constraint on the parameters appearing in R_1, we set $\mathbb{P} = \emptyset$.
S4 Let $M = \{(R_1, \mathbb{P}_1, d_1, d_2) : (R_1, \mathbb{P}_1) \in \Delta, d_1 d_2 = d \text{ and } d_2 \geq \text{tdeg}(S) + 1\}$.
S5 If $M \neq \emptyset$, select a $(R_1, \mathbb{P}_1, d_1, d_2)$ from M and let $M := M - \{(R_1, \mathbb{P}_1, d_1, d_2)\}$. Otherwise, the algorithm terminates.

S6 If $k = 1$, we can find H_1 by $R_1 \circ H_1 = P_1$. Execute Algorithm 4.7 with input $R_1, H_1, p, d_1, d_2, \mathbb{P}_1$ to find an h and a set $\mathbb{P}$. If $k > 1$, we can also find an h and a set $\mathbb{P}$ similar to Algorithm 4.7 but using equations (23). If h does not exist, go to S5.

S7 Compute a g such that $f = g \circ h$: let $g = \sum_I g_I y^{i_0} y_1^{i_1} \cdots y_c^{i_c}$, where $c = o_f$ and $i_0 + i_1 + \cdots + i_c \le \text{tdeg}(f)$. Substituting f, g, h in $f = g \circ h$ and comparing the coefficients of the monomials on both sides, we obtain a parametric linear system LS about g_I. Solve $LS = 0$ with methods from [24,6]. If $LS = 0$ has a solution, output g, h; else go to S5.

Remark In step S7, if h constains no parameter, then we can compute g by Algorithm 2.2. If the coefficients of h contain parameters, then the order and total degree of h may be changed by the parameters and we cannot know the order or the total degree of g exactly. But $o_g \le o_f$, $\text{tdeg}(g) \le \text{tdeg}(f)$ are always true. The size of the linear system might be large. We can reduce the size by comparing the total degree homogeneous parts and the differential homogeneous parts. In most examples, there exist no or very few parameters and some of the coefficients of h are independent of the parameters. This can also be used to reduce the size of the linear system.

Example 4.13 Continue from Example 3.4. We have obtained a p from f and S. Use f, p, S as the input to Algorithm 4.12.

S1. $d = \text{tdeg}(p) = 4$, write p as $p = P_4 + P_3 + P_2 + P_1$, where $P_4 = \frac{1}{2t^2} y_1^4$, $P_3 = t y y_1^2$, $P_2 = \frac{1}{2} y^2 + y_1 y_2 + \frac{1}{2t} y_1^2$, $P_1 = \frac{1}{2t} y_1$.

S2. By $k = 1$, we have $s = P_1 = \frac{1}{2t} y_1$.

S3. By $P_1 = R_1 \circ H_1$ and $i_{R_1} = 1$, we have $R_1 = y_1 + \frac{1}{t} y$, $H_1 = \frac{1}{2t} y$, $\Delta = \{(y_1 + \frac{1}{t} y, \emptyset)\}$.

S4. By $\text{tdeg}(S) = 1$, $M = \{(y_1 + \frac{1}{t} y, \emptyset, 2, 2), (y_1 + \frac{1}{t} y, \emptyset, 1, 4)\}$.

S5. Choose $(y_1 + \frac{1}{t} y, \emptyset, 2, 2) \in M$ and let $M = \{(y_1 + \frac{1}{t} y, \emptyset, 1, 4)\}$.

S6. Let $R_1 = y_1 + \frac{1}{t} y$, $d_2 = 2$. Substituting these into (12), we have:

$$P_2 = \left(y_1 + \frac{1}{t} y\right) \circ H_2 + R_2 \circ \left(\frac{1}{2t} y\right) \tag{28}$$

$$P_3 = \boxed{R_2 \circ (H_1 + H_2)}_3 \tag{29}$$

$$P_4 = R_2 \circ H_2. \tag{30}$$

Since $o_{R_2} < 1, o_{H_2} \le o_f - o_{R_1} = 2 - 1 = 1$ and R_2, H_2 are 2−total degree homogeneous, we can assume

$$R_2 = a y^2, H_2 = b_1 y^2 + b_2 y y_1 + b_3 y_1^2.$$

Substitute them in (28), by Algorithm 4.7 we have:

$$b_3 = \frac{1}{2}, \quad b_2 = 0, \quad b_1 = 0, \quad a = 2t^2.$$

So we obtain $R_2 = 2t^2y^2$, $H_2 = \frac{1}{2}y_1^2$. Now substituting R_1, R_2, H_1, H_2 in (29) and (30), we find that they really hold. So we have the following decomposition for p:

$$p = (R_1 + R_2) \circ (H_1 + H_2) = \left(y_1 + \frac{1}{t}y + 2t^2y^2\right) \circ \left(\frac{1}{2t}y + \frac{1}{2}y_1^2\right)$$

and $h = \frac{1}{2t}y + \frac{1}{2}y_1^2$.

S7. Since there is no parameter in h, by Algorithm 2.2, we find $g = 16t^3y^2y_1 + 8t^4y^2y_2 + 8yy_1 + 4tyy_2 + 8ty_1^2 + 4t^2y_1y_2 + 2ty$ such that $f = g \circ h$.

If we choose $\{(y_1 + \frac{1}{t}y, \emptyset, 1, 4) \in M$ to begin the computation firstly, we will find that there is no solution for g, h.

5 The algorithm and experimental results

As a summary, we give the algorithm to decompose any differential polynomial.

Algorithm 5.1 *Input: a differential polynomial f in $\mathcal{K}\{y\}$.*
Output: a nontrivial decomposition $f = g \circ h$, if such g, h exist.

S1 Let $f := f - f_0$, where f_0 is the part of f in $\mathcal{K}$.

S2 Find a polynomial $g \in \mathcal{K}[y]$ and an $h \in \mathcal{K}\{y\}$ such that $f = g \circ h$ with Algorithm 2.3. If such g and h exist, output $g + f_0, h$. Otherwise, go to next step.

S3 Let $d = d_f$ and write f as the form $f_d y_{o_f}^d + f_{d-1}y_{o_f}^{d-1} + \cdots + f_1 y_{o_f} + f_0$, where f_i denotes the coefficient of $y_{o_f}^i$ in f $(0 \le i \le d)$. Let $T = \{S : S = 1$ or S^i is a factor of f_i for all i satisfying $1 \le i \le d\}$. Here two factors $u, v \in T$ are defined to be equivalent if $\frac{u}{v} \in \mathcal{K}$. T contains non-equivalent factors.

S4 If $T \neq \emptyset$, choose an $S \in T$, go to next step; otherwise, terminate the algorithm and return "no nontrivial decomposition exists".

S5 S could serve as a candidate for the separant of h.
Execute Algorithm 3.3 with input f, S. There are three cases:
(1) We obtain a decomposition of $f = g \circ h$. Output $g + f_0, h$ and terminate the algorithm.
(2) We obtain a differential polynomial p with $d_p = 1$ and $i_p = S$ such that if h is a right decomposition factor of f with separant S, then h is a right decomposition factor of p. Go to next step.
(3) The output is the empty set. Let $T := T - \{S\}$, go to S4;

S6 Execute Algorithm 4.12 with input f, p, S. If we obtain a decomposition $f = g \circ h$, then output $g + f_0$ and h; otherwise $T := T - \{S\}$, go to S4.

Example 5.2 Let f be the differential polynomial given in Example 4.13

S2. Execute Algorithm 2.3, we find that f does not have a polynomial decomposition.

S3. $d = d_f = 1$, $f_1 = 2ty_1(t^3y_1^4 + 2t^2yy_1^2 + ty^2 + 2ty_1y_2 + y_1^2 + y_1)$, $T = \{1, y_1, t^3y_1^4 + 2t^2yy_1^2 + ty^2 + 2ty_1y_2 + y_1^2 + y_1, y_1(t^3y_1^4 + 2t^2yy_1^2 + ty^2 + 2ty_1y_2 + y_1^2 + y_1)\}$.

Table 1 Decomposing randomly generated differential polynomials

(o_f, t_f, l_f)	Time (s)	(o_f, t_f, l_f)	Time (s)
(2, 10, 33.6)	0.2	(2, 20, 84.5)	0.88
(2, 30, 224.2)	2.41	(2, 40, 457.9)	5.3
(2, 50, 387.6)	4.9	(3, 10, 181.1)	1.6
(3, 12, 258.7)	1.8	(3, 15, 348.2)	3.4
(4, 8, 263.3)	1.7	(4, 10, 437.1)	3.1
(5, 8, 602.5)	4.7	(6, 6, 403.4)	2.1
(7, 6, 1021.9)	6.6	(8, 6, 1349.2)	9.8
(9, 4, 301.5)	2.1	(9, 6, 2512.4)	23.7
(10, 5, 1905.4)	15.4	(10, 6, 3373.4)	38.4

S4. Select $S = 1$.

S5. Execute Algorithm 3.3, we obtain $p = y_1$.

S6. By Algorithm 4.12, we get no nontrivial decomposition for f.

Now, return to S4. Select $S = y_1$.

S5. Execute Algorithm 3.3, and we obtain $p = y_1 y_2 + \frac{1}{2}t^2 y_1^4 + t y y_1^2 + \frac{1}{2}y^2 + \frac{1}{2t}y_1^2 + \frac{1}{2t}y_1$(Example 3.4).

S6. By Algorithm 4.12, we have $g = 16t^3 y^2 y_1 + 8t^4 y^2 y_2 + 8 y y_1 + 4t y y_2 + 8t y_1^2 + 4t^2 y_1 y_2 + 2t y,\ h = \frac{1}{2t}y + \frac{1}{2}y_1^2$. Thus we obtain a decomposition $f = g \circ h$.

We may use Algorithm 5.1 recursively to find an irreducible decomposition of any differential polynomial, in which each decomposition factor has no nontrivial decomposition.

We implemented Algorithm 5.1 in the *constant field case* $\mathcal{K} = \mathbb{Q}$ in Maple. In Table 1, for each pair of (o_f, t_f), we generate ten differential polynomials of order o_f and total degree t_f randomly and decompose them. l_f is the average length of the ten differential polynomials generated randomly. The second and the fourth columns denote the average time on decomposing the ten differential polynomials. All the randomly generated differential polynomials in Table 1 are indecomposable. In Tables 2 and 3, we generate two differential polynomials g and h randomly and decompose $f = g \circ h$. o_f, t_f, l_f are the order, the total degree, and the number of terms in f respectively. The running times are collected on a PC with a 3.2 GHz CPU and 512 M memory and are given in seconds. From these results, we may conclude that our algorithm is efficient in handling differential polynomials with hundreds, even thousands, of terms.

The worst case complexity of our algorithm is exponential due to combinatorial selections of all the possible separants of h in step S3 of Algorithm 5.1, all factors of LODOs in step S2 of Algorithm 4.12, solving parametric linear system in Algorithm 4.7. It is worth noting that the complexity of factoring LODOs, which is equivalent to decomposing linear differential polynomials, is already exponential [10]. The practical efficiency of our algorithm is mainly due to two reasons. First, for each selection of the possible separant of h and a pair of factors of an LODO, our algorithm is very fast, and for most of the practical problems, the number of selections is not very large. Second,

Table 2 Randomly generated g and h

g_1	$-27-10y^8-2y^7-23y^5+21y^4-47y^2-43y$
h_1	$-39y^2-28y^3+27y^4-33y_1y^5+17y^7-37y^6y_1+17y_1^2y^5+12y^5y_1^3$
g_2	$-7y+13y_1-15y_1^3-46y^4-48y^3y_1-35y^2y_1^2-34yy_1^5$
h_2	$-13y_1^2+36y_1y^2-10y^2y_1^2-10yy_1^3-26y^5y_1^3$
g_3	$41yy_1-39y_1y^2-47yy_1^3+50y_1^4$
h_3	$-34y_1^2y+32yy_2^2+35y^2y_1^2+13y^2y_1y_2$
g_4	$-36y^2-44y_1^3+28y^3y_1+35y^2y_1^2+7y_1^4$
h_4	$-49-24y_1-26y_2+6yy_1-12y_2y+46yy_3+26y_1y_3-13y_2y_3-21y_3^2$
g_5	$16y^2-4y_1^4$
h_5	$-3y^2+26yy_3^2-27y_1y_3y_2+34yy_1^2y_2+y_1^2y_2^2+14y_3^2y_2^2$
g_6	$-25yy_2^2+50y_1^2y_2+49y_2^2y_1+15y^2y_1^2+49y^2y_1y_2+44yy_2^3+3y_1^4+24y_2y_1^3-38y_2^4$
h_6	$-18y^2+46y_1^2y-35y^4-28y^3y_1+14yy_1^3$
g_7	$-y^3+46y_1^2y-19yy_1y_2+37yy_2^2$
h_7	$-36y_1+5y_1^2+14y_2^2+23y^2y_2-35y_1^2y+43yy_1y_2-48y_2^3-13y^2y_2^2+22y_2^3y_1+17y_2^4$
g_8	$13y_2^2-9y_1^2y-43y_1^3+44y_2^2y_1-33y_2^3$
h_8	$12-10y-48y_3+45yy_1+6yy_3-38y_1y_3+37y_2^2-40y_2y_3+11y_3^2$
g_9	$38y-46y_1+39y_3+9yy_1-16y_2y+47y_1^2+31y_2^2+19y_2y_3$
h_9	$-33y_1^2+12y^3+y^3y_1-9y^2y_1^2+44yy_1^3$
g_{10}	$35yy_1+42y_2y-3yy_3+5y_1y_3$
h_{10}	$-49y-37y_1^2+11y^2y_2+11y_2^2y_1-5y^2y_2^2+$ $44yy_1y_2^2-40yy_2^3+17y_1^4-28y_2y_1^3-47y_1^2y_2^2-5y_2^3y_1$

Table 3 Decompose $f = g \circ h$

g, h	(o_g, t_g)	(o_h, t_h)	(o_f, t_f, l_f)	Time (s)
g_1, h_1	(0,8)	(1,8)	(1,64,639)	2.48
g_2, h_2	(1,6)	(1,8)	(2,48,1174)	14.85
g_3, h_3	(1,4)	(2,4)	(3,16,458)	9.32
g_4, h_4	(1,4)	(3,2)	(4,8,994)	9.67
g_5, h_5	(1,4)	(3,4)	(4,16,970)	17.9
g_6, h_6	(2,4)	(1,4)	(3,16,1229)	24.8
g_7, h_7	(2,3)	(2,4)	(4,12,1360)	11.3
g_8, h_8	(2,3)	(3,2)	(5,6,709)	6.5
g_9, h_9	(3,2)	(1,4)	(4,8,231)	1.78
g_{10}, h_{10}	(3,2)	(2,4)	(5,8,535)	2.17

since we consider the constant coefficient case, our implementation uses Lemma 4.2 to factor LODOs.

To implement Algorithm 5.1 in the case of rational function field, we need an implementation which gives all the possible factorizations for an LODO. This problem

itself is also a difficult task and we do not find such an implementation. Based on our experiments on the constant field case and the analysis of our algorithm, we expect that our algorithm provides an efficient reduction of the decomposition of nonlinear differential polynomials to linear ones.

6 Conclusions

In this paper, we give a complete and practical algorithm for decomposing nonlinear differential polynomials in one variable and with coefficients in a differential field $\mathcal{K}$ of characteristic zero. Besides arithmetic operations, the algorithm needs complete enumeration of all factorizations of LODOs, polynomial decomposition, factorization of multi-variable polynomials and solution of algebraic parametric linear equation systems over $\mathcal{K}$. The algorithm can give all the enumerations of nonequivalent decompositions. If there are parameters in the decomposition factors, then the parameters come from the factorization of an LODO with order not greater than that of the given differential polynomial to be decomposed and they are constrained by some algebraic equations. Experiments show that the algorithm is quite efficient.

Many problems on the decomposition of differential polynomials are still open. The irreducible decomposition of a polynomial is unique in certain sense [20,4]. Similar results were proved for Ore polynomials and hence for LODOs [19]. It is interesting to see whether this property is correct for differential polynomial decomposition.

References

1. Alagar, V.S., Thanh, M.: Fast decomposition algorithms. In: Proc. EUROCAL, vol. 85(2), pp. 150–153. Springer, Heidelberg (1985)
2. Barton, D.R., Zippel, R.E.: Polynomial decomposition. In: Proc. SYMSAC, vol. 76, pp. 356–358 (1976)
3. Bronstein, M., Petkovšek, M.: On Ore rings, linear operators and factorization. Program. Comput. Softw. **20**(1), 27–44 (1994)
4. Dorey, F., Whaples, G.: Prime and composite polynomials. J. Algebra **28**, 88–101 (1974)
5. Fredet, A.: Factorization of linear differential operators in exponential extensions. In: Proc. of ISSAC'03, pp. 103–110. ACM Press, New York (2003)
6. Gao, X.S., Chou, S.C.: Solving parametric algebraic systems. In: Proc. of ISSAC'92, pp. 335–341. ACM Press, New York (1992)
7. Gao, X.S., Zhang, M.: Decomposition of differential polynomials with constant coefficients. Proc. ISSAC'04, 175-182, ACM Press, New York (2004)
8. Gao, X.S., Zhang, M.: Decomposition of differential polynomials with rational function coefficients. MM-Preprints, vol. 23, pp. 92–112. KLMM, Beijing (2004)
9. Giesbrecht, M., Zhang, Y.: Factoring and decomposing Ore polynomials over $Fq(t)$. In: Proc. of ISSAC'03, pp. 127–135. ACM Press, New York (2003)
10. Grigor'ev, D.Y.: Complexity of factoring and calculating the GCD of linear ordinary differential operators. J. Symb. Comput. **10**, 7–37 (1990)
11. Gutierrez, J., Recio T., de Velasco C.R.: Polynomial decomposition of almost quadratic complexity. In: Proc. AAECC, vol. 6, pp. 471–475. Springer, Heidelberg (1989)
12. Gutierrez, J., Rubio, R., Sevilla, D.: On multivariate rational function decomposition. J. Symb. Comput. **33**(5), 545–562 (2002)
13. Gutierrez, J., Sevilla, D.: Computation of unirational fields. J. Symb. Comput. **41**(11), 1222–1244 (2006)

14. Gutierrez, J., Rubio, R., Sevilla, D.: Unirational fields of transcendence degree one and functional decomposition. In: Proc. of ISSAC'01, pp. 167–174. ACM Press, New York (2001)
15. Hu, H.C.: On the equilibrium of a transversely isotropic elastic body under body forces. Acta Phys. Sin. **11**, 219–230 (1955)
16. Königsberger, L.: Allgemeine Untersuchungen aus der Theorie der Differentialgleichungen. Teubner, Leipzig (1882)
17. Kozen, D., Landau, S.: Polynomial decomposition algorithms. J. Symb. Comput. **7**, 445–456 (1989)
18. Li, Z.: A subresultant theory for Ore polynomials with applications. In: Proc. ISSAC'98, pp. 132–139. ACM Press, New York (1998)
19. Ore, O.: Theory of noncommutative polynomials. Ann. Math. **34**(3), 480–508 (1933)
20. Ritt, J.F.: Prime and composite polynomials. Trans. AMS **23**, 51–66 (1922)
21. Schwarz, F.: Efficient factorization of linear ODE's. SIGSAM Bull. **28**(1), 9–17 (1994)
22. Singer, M.F.: Liouillian solutions of nth order homogeneous linear differential equations. Am. J. Math. **103**(4), 661–682 (1981)
23. Singer, M.F., Ulmer, F.: Galois groups of second and third order linear differential equations. J. Symb. Comput. **16**, 9–36 (1993)
24. Sit, W.Y.: An algorithm for solving parametric linear systems. J. Symb. Comput. **13**, 353–394 (1992)
25. Sosnin, M.V.: An algorithm for nonparametric decomposition of differential polynomials. Program. Comput. Softw. **27**(1), 43–49 (2001)
26. Tsarev, S.P.: An algorithm for complete enumeration of all factorizations of a linear ordinary differential operator. In: Proc. of ISSAC'96, pp. 226–231. ACM Press, New York (1996)
27. Tsarev, S.P.: On factorization of nonlinear ordinary differential equations. In: Proc. ISSAC'99, pp. 159–164. ACM Press, New York (1999)
28. Van der Put, M., Singer, M.F.: Galois Theory of Linear Differential Equations. Springer, Berlin (2003)
29. van Hoeij, M.: Factorization of differential operators with rational functions coefficients. J. Symb. Comput. **24**(5), 537–561 (1997)
30. van Hoeij, M.: Formal solutions and factorization of differential operators with power series coefficients. J. Symb. Comput. **24**, 1–30 (1997)
31. von zue Gathen, J.: Functional decomposition of polynomials: the wild case. J. Symb. Comput. **9**, 437–452 (1990)
32. von zue Gathen, J.: Functional decomposition of polynomials: the tame case. J. Symb. Comput. **9**, 281–299 (1990)
33. von zur Gathen, J., Gutierrez, J., Rubio, R.: Multivariate polynomial decomposition. Appl. Algebra Eng. Commun. Comput. **14**(1) (2003)
34. Zippel, R.E.: Rational function decomposition. In: Proc. of ISSAC'91, pp. 1–6. ACM Press, New York (1991)