PROCEEDINGS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 137, Number 2, February 2009, Pages 597–606
S 0002-9939(08)09493-8
Article electronically published on August 22, 2008

ON THE ANALYTIC SOLUTION OF THE CAUCHY PROBLEM

XIANG-DONG HOU

(Communicated by Jim Haglund)

Abstract. Derivatives of a solution of an ODE Cauchy problem can be computed inductively using the Faà di Bruno formula. In this paper, we exhibit a noninductive formula for these derivatives. At the heart of this formula is a combinatorial problem, which is solved in this paper. We also give a more tractable form of the Magnus expansion for the solution of a homogeneous linear ODE.

1. Introduction

Consider the Cauchy problem

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} f_1(y_1,\dots,y_n) \\ \vdots \\ f_n(y_1,\dots,y_n) \end{bmatrix}, \qquad \begin{bmatrix} y_1(0) \\ \vdots \\ y_n(0) \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{1.1}$$

where $f_1, \dots, f_n$ have continuous partial derivatives of total order up to v in a neighborhood of $(0,\dots,0)$. The $(v+1)$st order derivatives of the unique solution of (1.1) in a neighborhood of $(0,\dots,0)$ are given by

$$\frac{d^{v+1}y_i}{dt^{v+1}} = \frac{d^v}{dt^v} f_i(y_1,\dots,y_n), \qquad v \geq 0. \tag{1.2}$$

The right side of (1.2) is given by the multivariate Faà di Bruno formula [2, 12]. (For more about the Faà di Bruno formula, see [1, 3, 6, 10].) Altogether, we have for $v \geq 0$ that

$$\frac{d^{v+1}y_i}{dt^{v+1}} = v! \sum_{\substack{(\lambda_1,\dots,\lambda_n)\in\mathbb{N}^n \\ 1\le\lambda_1+\cdots+\lambda_n\le v}} \frac{\partial^{\lambda_1+\cdots+\lambda_n} f_i}{\partial y_1^{\lambda_1}\cdots\partial y_n^{\lambda_n}} \sum_{\substack{(\mu_{jk})\in\mathbb{N}^{v\times n} \\ \sum_j \mu_{jk}=\lambda_k,\ \sum_{j,k} j\mu_{jk}=v}} \prod_{j,k} \frac{1}{\mu_{jk}!}\Big(\frac{1}{j!}\frac{d^j y_k}{dt^j}\Big)^{\mu_{jk}}, \tag{1.3}$$

where $\mathbb{N} = \{0,1,2,\dots\}$. Equation (1.3) allows us to compute the derivatives $\frac{d^{v+1}y_i}{dt^{v+1}}$ inductively on v. If $f_1,\dots,f_n$ are analytic in a neighborhood of $(0,\dots,0)$, then (1.3) also allows us to compute the coefficients of the power series solution inductively.

Received by the editors April 13, 2007, and, in revised form, January 24, 2008.
2000 *Mathematics Subject Classification.* Primary 34A25, 05A15.
Key words and phrases. Cauchy problem, ODE, Magnus expansion, partial order, combinatorics, Bruno's formula.

The derivatives of the solution of (1.1) can also be expressed in terms of the linear differential operator

$$D = \sum_{j=1}^{n} f_j(y_1, \dots, y_n) \frac{\partial}{\partial y_j}.$$

We have

$$\frac{d^v y_i}{dt^v} = D^v y_i, \qquad v \geq 0; \tag{1.4}$$

see [5, 4, 7]. However, the right side of (1.4) is not explicit in terms of the partial derivatives of $f_1, \dots, f_n$.

The purpose of the present paper is to point out that there is a formula for the derivatives of the solution of (1.1) which is noninductive and is explicit in terms of the partial derivatives of $f_1, \dots, f_n$. The formula is stated in Section 2. At the heart of the formula is a combinatorial problem, which is solved in Section 3. Section 4 discusses a connection between a byproduct of Section 3 and the Magnus expansion for the solution of a homogeneous linear ODE.

2. The main result

For $(j_1, \dots, j_{v-1}) \in \mathbb{N}^{v-1}$ and $(i_1, \dots, i_v) \in \mathbb{N}^v$, we say that $(j_1, \dots, j_{v-1}) \prec (i_1, \dots, i_v)$ if $(i_1, \dots, i_v) = (j_1, \dots, j_{u-1}, j_u + 1, 0, j_{u+1}, \dots, j_{v-1})$ for some $1 \leq u \leq v-1$ or $(i_1, \dots, i_v) = (j_1, \dots, j_{v-1}, 1)$. Define $\mathbb{N}^0 = \{\emptyset\}$ and define $\emptyset \prec (1) \in \mathbb{N}^1$. Using transitivity, we extend $\prec$ to a partial order in $\bigcup_{v \geq 0} \mathbb{N}^v$.

Let

$$\mathcal{I}_v = \big\{(i_1, \dots, i_v) \in \mathbb{N}^v : i_1 + \dots + i_u \geq u \text{ for } 1 \leq u \leq v \text{ and } i_1 + \dots + i_v = v\big\}.$$

Also define $\mathcal{I}_0 = \{\emptyset\}$. If $(j_1, \dots, j_{v-1}) \prec (i_1, \dots, i_v)$, then $(j_1, \dots, j_{v-1}) \in \mathcal{I}_{v-1}$ if and only if $(i_1, \dots, i_v) \in \mathcal{I}_v$.

Let $f_1(y_1, \dots, y_n), \dots, f_n(y_1, \dots, y_n)$ be functions of $y_1, \dots, y_n$ which are differentiable enough times. Write

$$\boldsymbol{f} = \begin{bmatrix} f_1(y_1, \dots, y_n) \\ \vdots \\ f_n(y_1, \dots, y_n) \end{bmatrix}.$$

Moreover, for $i > 0$, let $\frac{\partial^i \boldsymbol{f}}{\partial \boldsymbol{y}^i}$ be an $n \times n^i$ matrix whose columns are indexed by $(\beta_1, \dots, \beta_i) \in \{1, \dots, n\}^i$ lexicographically and whose $\big(\alpha, (\beta_1, \dots, \beta_i)\big)$-entry is

$$\frac{\partial^i f_\alpha}{\partial y_{\beta_1} \cdots \partial y_{\beta_i}}.$$

For example, if $\boldsymbol{f} = [f_1(y_1, y_2), f_2(y_1, y_2)]^T$, then

$$\frac{\partial^2 \boldsymbol{f}}{\partial \boldsymbol{y}^2} = \begin{bmatrix} \dfrac{\partial^2 f_1}{\partial y_1 \partial y_1} & \dfrac{\partial^2 f_1}{\partial y_1 \partial y_2} & \dfrac{\partial^2 f_1}{\partial y_2 \partial y_1} & \dfrac{\partial^2 f_1}{\partial y_2 \partial y_2} \\ \dfrac{\partial^2 f_2}{\partial y_1 \partial y_1} & \dfrac{\partial^2 f_2}{\partial y_1 \partial y_2} & \dfrac{\partial^2 f_2}{\partial y_2 \partial y_1} & \dfrac{\partial^2 f_2}{\partial y_2 \partial y_2} \end{bmatrix}.$$

Theorem 2.1. *The derivatives of the solution of* (1.1) *are given by*

(2.1)
$$\frac{d^{v+1}}{dt^{v+1}}\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{(i_1,\ldots,i_v)\in\mathcal{I}_v} a(i_1,\ldots,i_v)\frac{\partial^{i_1}\boldsymbol{f}}{\partial\boldsymbol{y}^{i_1}}\Big(I_{n^{i_1-1}}\otimes\frac{\partial^{i_2}\boldsymbol{f}}{\partial\boldsymbol{y}^{i_2}}\Big)\cdots\Big(I_{n^{i_1+\cdots+i_{v-1}-(v-1)}}\otimes\frac{\partial^{i_v}\boldsymbol{f}}{\partial\boldsymbol{y}^{i_v}}\Big)\boldsymbol{f},$$

where $a(i_1,\ldots,i_v)\in\mathbb{Z}^+$, $(i_1,\ldots,i_v)\in\mathcal{I}_v$, *is defined inductively by*

$$\begin{cases} a(i_1,\ldots,i_v) = \displaystyle\sum_{(j_1,\ldots,j_{v-1})\prec(i_1,\ldots,i_v)} a(j_1,\ldots,j_{v-1}), \\ a(\emptyset)=1. \end{cases} \tag{2.2}$$

Proof. For $(i_1,\ldots,i_v)\in\mathcal{I}_v$, let $i_{v+1}=0$ and let

$$F_{(i_1,\ldots,i_v,0)} = \prod_{u=1}^{v+1}\Big(I_{n^{i_1+\cdots+i_{u-1}-(u-1)}}\otimes\frac{\partial^{i_u}\boldsymbol{f}}{\partial\boldsymbol{y}^{i_u}}\Big), \tag{2.3}$$

where the factors in the product appear from left to right in the order of $u = 1,2,\ldots,v+1$. Then (2.1) can be written as

$$\frac{d^{v+1}}{dt^{v+1}}\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \sum_{(i_1,\ldots,i_v)\in\mathcal{I}_v} a(i_1,\ldots,i_v)F_{(i_1,\ldots,i_v,0)}. \tag{2.4}$$

To prove (2.4) and (2.2), we use induction on v. The initial case $v=0$ needs no proof. Since

$$\frac{d}{dt}\Big(\frac{\partial^i f_\alpha}{\partial y_{\beta_1}\cdots\partial y_{\beta_i}}\Big) = \sum_{\beta_{i+1}}\frac{\partial^{i+1}f_\alpha}{\partial y_{\beta_1}\cdots\partial y_{\beta_i}\partial y_{\beta_{i+1}}}\frac{dy_{\beta_{i+1}}}{dt} = \sum_{\beta_{i+1}}\frac{\partial^{i+1}f_\alpha}{\partial y_{\beta_1}\cdots\partial y_{\beta_i}\partial y_{\beta_{i+1}}}f_{\beta_{i+1}},$$

we have

$$\frac{d}{dt}\Big(\frac{\partial^i\boldsymbol{f}}{\partial\boldsymbol{y}^i}\Big) = \frac{\partial^{i+1}\boldsymbol{f}}{\partial\boldsymbol{y}^{i+1}}(I_{n^i}\otimes\boldsymbol{f}).$$

Thus

$$\begin{aligned} &\frac{d}{dt}\Big[I_{n^{i_1+\cdots+i_{u-1}-(u-1)}}\otimes\frac{\partial^{i_u}\boldsymbol{f}}{\partial\boldsymbol{y}^{i_u}}\Big] \\ &= I_{n^{i_1+\cdots+i_{u-1}-(u-1)}}\otimes\Big[\frac{\partial^{i_u+1}\boldsymbol{f}}{\partial\boldsymbol{y}^{i_u+1}}(I_{n^{i_u}}\otimes\boldsymbol{f})\Big] \\ &= \Big[I_{n^{i_1+\cdots+i_{u-1}-(u-1)}}\otimes\frac{\partial^{i_u+1}\boldsymbol{f}}{\partial\boldsymbol{y}^{i_u+1}}\Big]\big[I_{n^{i_1+\cdots+i_{u-1}+(i_u+1)-u}}\otimes\boldsymbol{f}\big]. \end{aligned} \tag{2.5}$$

By (2.3) and (2.5), we have

$$
\begin{aligned}
&\frac{d}{dt}F_{(i_1,\dots,i_v,0)}\\
&=\sum_{u=1}^{v+1}\Big[\prod_{s=1}^{u-1}\Big(I_{n^{i_1+\cdots+i_{s-1}-(s-1)}}\otimes\frac{\partial^{i_s}\boldsymbol{f}}{\partial\boldsymbol{y}^{i_s}}\Big)\Big]\Big[\frac{d}{dt}\Big(I_{n^{i_1+\cdots+i_{u-1}-(u-1)}}\otimes\frac{\partial^{i_u}\boldsymbol{f}}{\partial\boldsymbol{y}^{i_u}}\Big)\Big]\\
&\quad\times\Big[\prod_{s=u+1}^{v+1}\Big(I_{n^{i_1+\cdots+i_{s-1}-(s-1)}}\otimes\frac{\partial^{i_s}\boldsymbol{f}}{\partial\boldsymbol{y}^{i_s}}\Big)\Big]\\
&=\sum_{u=1}^{v+1}F_{(i_1,\dots,i_{u-1},i_u+1,0,i_{u+1},\dots,i_v,0)}.
\end{aligned}
$$

Therefore, assuming (2.4), we have

$$
\begin{aligned}
\frac{d^{v+2}}{dt^{v+2}}\begin{bmatrix}y_1\\ \vdots\\ y_n\end{bmatrix}&=\sum_{(i_1,\dots,i_v)\in\mathcal{I}_v}a(i_1,\dots,i_v)\sum_{u=1}^{v+1}F_{(i_1,\dots,i_{u-1},i_u+1,0,i_{u+1},\dots,i_v,0)}\\
&=\sum_{(j_1,\dots,j_{v+1})\in\mathcal{I}_{v+1}}a(j_1,\dots,j_{v+1})F_{(j_1,\dots,j_{v+1},0)},
\end{aligned}
$$

where

$$
a(j_1,\dots,j_{v+1})=\sum_{(i_1,\dots,i_v)\prec(j_1,\dots,j_{v+1})}a(i_1,\dots,i_v).
$$

So the induction is complete. □

Remark. Let R be a commutative ring of characteristic 0 and $R[[Y_1,\dots,Y_n]]$ be the ring of formal power series in $Y_1,\dots,Y_n$ over R. Let $f_1,\dots,f_n\in R[[Y_1,\dots,Y_n]]$ and $y_1,\dots,y_n\in R[[X]]$ be the unique solution of (1.1). Then (1.3) and (2.2) also hold. Both formulas give the coefficients of $y_1,\dots,y_n$ in terms of the coefficients of $f_1,\dots,f_n$, (1.3) inductively and (2.2) explicitly.

3. Determination of $a(i_1,\dots,i_v)$

For each $(i_1,\dots,i_v)\in\mathcal{I}_v$, we define a *walk* $w(i_1,\dots,i_v)$ to be a sequence of points

$$(0,0),\ (0,i_1),\ (1,i_1),\ (1,i_1+i_2),\ \dots,\ (v-1,i_1+\cdots+i_v),\ (v,v)$$

in $\mathbb{N}^2$. It helps to think of the walk $w(i_1,\dots,i_v)$ as line segments connecting the points in the above sequence. Therefore, the walk $w(i_1,\dots,i_v)$ starts from $(0,0)$ and moves i_1 units up, 1 unit right, i_2 units up, 1 unit right and so on until it reaches (v,v). For example, $w(2,0,3,2,0,2,0,0,1,0)$ is shown in Figure 1. Note that since $(i_1,\dots,i_v)\in\mathcal{I}_v$, the walk $w(i_1,\dots,i_v)$ is above the line $y=x$. Let R denote the closed region of $[0,v]\times[0,v]$ below the walk $w(i_1,\dots,i_v)$. For each $0\le i\le v$ let L_i denote the line $y=x+i$. Then $L_i\cap R$ consists of line segments; each line segment is further divided into subsegments by certain vertices of the walk. In Figure 1, $L_1\cap R$ consists of two segments (from lower left to upper right): $\overline{(0,1)\,(1,2)}$ and $\overline{(2,3)\,(9,10)}$. The segment $\overline{(2,3)\,(9,10)}$ consists of two subsegments: $\overline{(2,3)\,(8,9)}$ and $\overline{(8,9)\,(9,10)}$. If a segment consists of subsegments of length $l_1\sqrt{2},l_2\sqrt{2},\dots$

(from lower left to upper right), we say that the segment is of type $(l_1, l_2, \dots)$. In Figure 1,

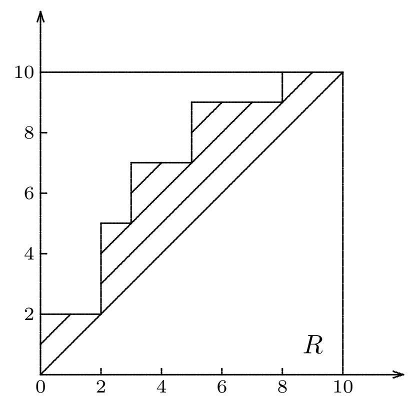

FIGURE 1. The walk $w(2,0,3,2,0,2,0,0,1,0)$

- L_0 has 1 segment of type $(2,8)$;
- L_1 has 2 segments of type $(1),(6,1)$;
- L_2 has 1 segment of type $(1,2,2)$;
- L_3 has 2 segments of type $(1),(1)$.

Theorem 3.1. *Let $(i_1,\dots,i_v) \in \mathcal{I}_v$ and let R be the closed region of $[0,v]\times[0,v]$ below the walk $w(i_1,\dots,i_v)$. Assume that $L_s \cap R$ consists of segments of type $(l_1^{s,1}, l_2^{s,1},\dots),(l_1^{s,2},l_2^{s,2},\dots),\dots$. Then*

(3.1)
$$a(i_1,\dots,i_v) = \prod_s \prod_u \Big[\binom{-1+l_1^{s,u}+l_2^{s,u}+\cdots}{-1+l_1^{s,u}} \binom{-1+l_2^{s,u}+\cdots}{-1+l_2^{s,u}} \cdots \Big]$$
$$= \prod_s \prod_u \binom{l_1^{s,u}+l_2^{s,u}+\cdots}{l_1^{s,u},l_2^{s,u},\cdots} \frac{l_1^{s,u} l_2^{s,u}\cdots}{(l_1^{s,u}+l_2^{s,u}+\cdots)(l_2^{s,u}+l_3^{s,u}+\cdots)\cdots}.$$

For example, Theorem 3.1 applied to the walk in Figure 1 gives

$$a(2,0,3,2,0,2,0,0,1,0) = \binom{9}{1}\binom{6}{5}\binom{4}{0}\binom{3}{1} = 162.$$

Proof of Theorem 3.1. From (2.2), it is clear that $a(i_1,\dots,i_r)$ is the number of sequences $\emptyset = \alpha_0 \prec \alpha_1 \prec \cdots \prec \alpha_v = (i_1,\dots,i_v)$, where $\alpha_u \in \mathcal{I}_u$, $0 \le u \le v$. The sequence $(i_1,\dots,i_v,0)$ can be reduced to 0 through v reduction steps; each step is a replacement of a string $(i,0)$ in $(i_1,\dots,i_v,0)$ by $i-1$, where $i>0$. Therefore, $a(i_1,\dots,i_v)$ is the number of sequences of reduction steps which reduce $(i_1,\dots,i_v,0)$ to 0.

The sequence $(i_1,\dots,i_v,0)$ consists of strings of maximal length of the form $(j_1,\dots,j_s,\underbrace{0,\dots,0}_{t})$, where $j_1\cdots j_s \neq 0$ and $t>0$. Each such string can be reduced in a unique way through $\min\{t, j_1 + \cdots + j_s\}$ reduction steps to

$$\begin{cases} \underbrace{(0,\dots,0)}_{t-(j_1+\cdots+j_s)+1} & \text{if } t \ge j_1 + \cdots + j_s, \\ (*,\dots,*) & \text{if } t < j_1 + \cdots + j_s, \end{cases}$$

where the $*$'s are positive and sum to $j_1 + \cdots + j_s - t$. (See Figure 2.)

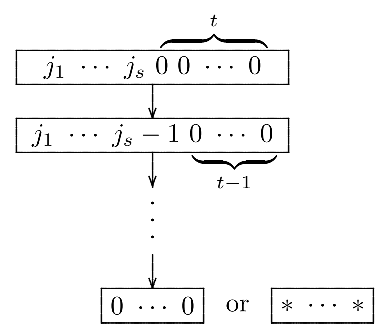

FIGURE 2. Reduction of $(j_1, \dots, j_s, 0, \dots, 0)$

Before the reduction of $(j_1, \dots, j_s, 0, \dots, 0)$ reaches $(0, \dots, 0)$ or $(*, \dots, *)$, the result of each intermediate step is a string of the form $(*, \dots, 0)$, $* \neq 0$, which cannot be combined with neighboring strings for further reduction. After $(j_1, \dots, j_s, 0, \dots, 0)$ is reduced to $(0, \dots, 0)$ or $(*, \dots, *)$, for further reduction, $(0, \dots, 0)$ must be combined with a left neighboring string and $(*, \dots, *)$ must be combined with a right neighboring string. Therefore, the reduction steps that reduce $(i_1, \dots, i_v, 0)$ to 0 are unique although these steps may be performed in different orders. (See Figure 3.)

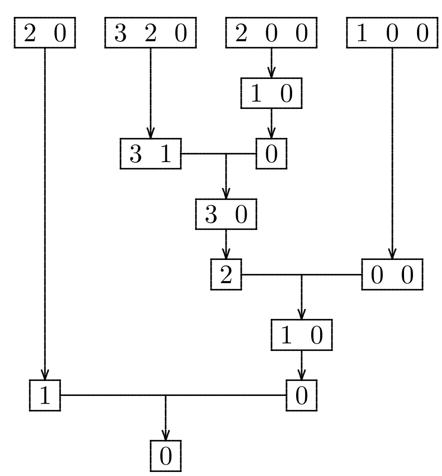

FIGURE 3. Reduction of $(2, 0, 3, 2, 0, 2, 0, 0, 1, 0, 0)$

Let R_1 and R_2 be two reduction steps of $(i_1, \ldots, i_v, 0)$. We say $R_1 \le R_2$ if R_1 needs to be performed before R_2 can be performed. Therefore, $a(i_1, \ldots, i_v)$ is the number of ways to order the v reduction steps of $(i_1, \ldots, i_v, 0)$ so that the partial order $\le$ is preserved.

We prove (3.1) by induction on v. The case $v = 0$ is obvious. Assume $v > 0$.

Case 1. Assume that $i_1 = 1$. In this case, $(i_2, \ldots, i_v) \in \mathcal{I}_{v-1}$. In the reduction of $(i_1, i_2, \ldots, i_v, 0)$, the string $(i_2, \ldots, i_v, 0)$ must be reduced to 0 before it can be combined with i_1 for the final reduction step. So, $a(i_1, i_2, \ldots, i_v) = a(i_2, \ldots, i_v)$ and (3.1) holds by the induction hypothesis.

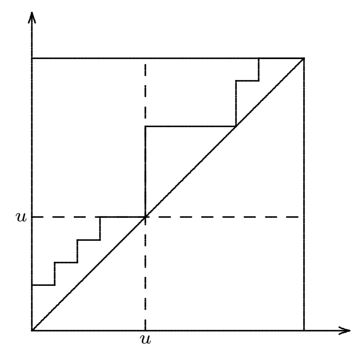

FIGURE 4. A walk $w(i_1, \ldots, i_{u-1}, 0, i_{u+1}, \ldots, i_v)$

Case 2. Assume that $i_1 > 1$ but the walk $w(i_1, \ldots, i_v)$ contains a point (u, u) with $0 < u < v$. We assume u is the smallest value with this property; see Figure 4. Then $i_u = 0$, so

$$(i_1, \ldots, i_v, 0) = (i_1, \ldots, i_{u-1}, 0, i_{u+1}, \ldots, i_v, 0),$$

where $(i_1, \ldots, i_{u-1}, 0) \in \mathcal{I}_u$ and $(i_{u+1}, \ldots, i_v) \in \mathcal{I}_{v-u}$. The string $(i_1, \ldots, i_{u-1}, 0)$ is reduced to 1 in $u-1$ steps before it can be combined with a right neighboring string for further reduction; the string $(i_{u+1}, \ldots, i_v, 0)$ is reduced to 0 in $v-u$ steps before it can be combined with a left neighboring string for further reduction. (See Figure 5.) Therefore,

$$a(i_1, \ldots, i_v) = \binom{v-1}{u-1} a(i_1, \ldots, i_{u-1}, 0) a(i_{u+1}, \ldots, i_v). \tag{3.2}$$

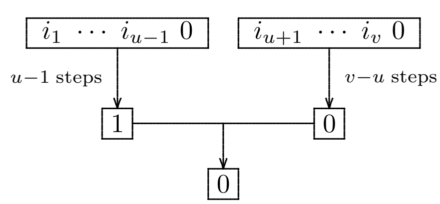

FIGURE 5. Reduction of $(i_1, \ldots, i_{u-1}, 0, i_{u+1}, \ldots, i_v, 0)$

The walk $w(i_1,\dots,i_{u-1},0)$ is the part of $w(i_1,\dots,i_v)$ from $(0,0)$ to (u,u); the walk $w(i_{u+1},\dots,i_v)$ is a translate of the part of $w(i_1,\dots,i_v)$ from (u,u) to (v,v). Since (3.1) holds for $a(i_1,\dots,i_{u-1},0)$ and $a(i_{u+1},\dots,i_v)$ (induction hypothesis), it follows from (3.2) that (3.1) also holds for $a(i_1,\dots,i_v)$.

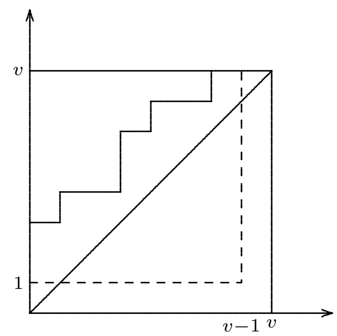

FIGURE 6. A walk $w(i_1,\dots,i_v)$ not touching L_0

Case 3. Assume that $w(i_1,\dots,i_v)$ does not contain any point (u,u) with $0<u<v$; see Figure 6. Then $i_1>1$, $i_v=0$, and $(i_1-1,i_2,\dots,i_{v-1})\in\mathcal{I}_{v-1}$. When reducing $(i_1,\dots,i_{v-1},0,0)$ to 0, the last step is always $(1,0)\to 0$. However, it is easy to see that the number of ways to reduce $(i_1,\dots,i_{v-1},0,0)$ to $(1,0)$ equals the number of ways to reduce $(i_1-1,i_2,\dots,i_{v-1},0)$ to 0. So

$$a(i_1,\dots,i_{v-1},0)=a(i_1-1,i_2,\dots,i_{v-1}). \tag{3.3}$$

The walk $w(i_1-1,i_2,\dots,i_{v-1})$ is a translate of the part of $w(i_1,\dots,i_v)$ from $(0,1)$ to $(v-1,v)$; see Figure 6. So by (3.3) and the induction hypothesis, (3.1) holds for $a(i_1,\dots,i_v)$. □

Since $a(i_1,\dots,i_v)$ is the number of chains $\emptyset=\alpha_0\prec\alpha_1\prec\cdots\prec\alpha_v=(i_1,\dots,i_v)$, where $\alpha_u\in\mathcal{I}_u$, $0\le u\le v$, and since for each $\alpha_u\in\mathcal{I}_u$, there are exactly u $\alpha_{u+1}\in\mathcal{I}_{u+1}$ such that $\alpha_u\prec\alpha_{u+1}$, we have

$$\sum_{(i_1,\dots,i_v)\in\mathcal{I}_v} a(i_1,\dots,i_v)=v!.$$

4. The set $\mathcal{I}_v$ and the Magnus expansion

The set $\mathcal{I}_v$ has an interesting combinatorial interpretation. Let $\mathcal{I}'_v=\{(j_1,\dots,j_v):(j_v,\dots,j_1)\in\mathcal{I}_v\}$. Let $\cdot$ be a nonassociative multiplication defined on a set $\mathcal{A}$. For $(a_1,\dots,a_{v+1})\in\mathcal{A}^{v+1}$, $(i_1,\dots,i_v)$, $(j_1,\dots,j_v)\in\mathbb{N}^v$, let $(a_1,\dots,a_{v+1})^{i_1,\dots,i_v}_{j_1,\dots,j_v}$ be the expression $a_1\cdots a_{v+1}$ with i_u "(" before a_u and j_u ")" after a_{u+1}. For example,

$$(a_1,a_2,a_3,a_4,a_5)^{2,0,1,1}_{0,2,0,2}=((a_1\ a_2\ (a_3))(a_4\ a_5)).$$

Note that $(a_1,\dots,a_{v+1})^{i_1,\dots,i_v}_{j_1,\dots,j_v}$ may not be a well-defined product as in the above example. However, if $(a_1,\dots,a_{v+1})^{i_1,\dots,i_v}_{j_1,\dots,j_v}$ is a well-defined product, then $(j_1,\dots,j_v)$ is determined by $(i_1,\dots,i_v)$ and vice versa. Assume that the rightmost "(" occurs before a_u. Then its matching ")" must occur after a_{u+1}, turning $(a_u\,a_{u+1})$ into a group. The matching ")" of the remaining "(" are determined the same way. By this argument and induction on v, one can easily see that $(a_1, \ldots, a_{v+1})^{i_1,\ldots,i_v}$ (an expression with "(" at indicated places) can be completed to a well-defined product $(a_1, \ldots, a_{v+1})^{i_1,\ldots,i_v}_{j_1,\ldots,j_v}$ if and only if $(i_1, \ldots, i_v) \in \mathcal{I}_v$. In the same way, $(a_1, \ldots, a_{v+1})_{j_1,\ldots,j_v}$ (an expression with ")" at indicated places) can be completed to a well-defined product if and only if $(j_1, \ldots, j_v) \in \mathcal{I}'_v$. If $(a_1, \ldots, a_{v+1})^{i_1,\ldots,i_v}_{j_1,\ldots,j_v}$ is a well-defined product, it will also be denoted by $(a_1, \ldots, a_{v+1})^{i_1,\ldots,i_v}$ or $(a_1, \ldots, a_{v+1})_{j_1,\ldots,j_v}$. Therefore, $(i_1, \ldots, i_v) \mapsto (a_1, \ldots, a_{v+1})^{i_1,\ldots,i_v}$ ($(j_1, \ldots, j_v) \mapsto (a_1, \ldots, a_{v+1})_{j_1,\ldots,j_v}$, respectively) is a bijection from $\mathcal{I}_v$ ($\mathcal{I}'_v$, respectively) to the set of all well-defined products $a_1 \cdots a_{v+1}$ with suitable associations. By [9, 13],

$$|\mathcal{I}_v| = |\mathcal{I}'_v| = \frac{(2v)!}{(v+1)!v!}.$$

The set $\mathcal{I}'_v$ has an application in the Magnus expansion for the solution of a homogeneous linear ODE. Let $M_n(\mathbb{R})$ denote the set of all $n \times n$ matrices over $\mathbb{R}$. Consider the Cauchy problem of the homogeneous linear ODE

$$\frac{dY}{dt} = A(t)Y, \qquad Y(0) = I_n, \tag{4.1}$$

where $A(t)$ and $Y(t)$ are $M_n(\mathbb{R})$-valued functions of t. Assume that $A(t)$ is continuous in a neighborhood of 0. Then in a (possibly smaller) neighborhood of 0, the unique solution of (4.1) can be expressed as

$$Y(t) = e^{\Omega(t)}$$

for some $M_n(\mathbb{R})$-valued function $\Omega(t)$. The Magnus expansion expresses $\Omega(t)$ as an infinite series involving $A(t)$; see Magnus [11] and Iserles and Nørsett [8]. For any two continuous $M_n(\mathbb{R})$-valued functions $B(t)$ and $C(t)$, define

$$B(t) \cdot C(t) = \Big[\int_0^t B(\tau)d\tau,\ C(t)\Big],$$

where [ ] is the Lie bracket. Then using the notation of the last paragraph, the Magnus expansion can be written as

$$\Omega(t) = \sum_{v=0}^{\infty} \sum_{(j_1,\ldots,j_v)\in\mathcal{I}'_v} f_{j_1} \cdots f_{j_v} \int_0^t \big(\underbrace{A(\tau) \cdots A(\tau)}_{v+1}\big)_{j_1,\ldots,j_v} d\tau, \tag{4.2}$$

where f_j is given by

$$\sum_{j=0}^{\infty} f_j z^j = \frac{z}{e^z - 1}.$$

Equation (4.2) can be easily derived from the Magnus expansion in [8, Eq. (2.6)]. One only has to note that using the notation of the present paper, the inductive formula for the coefficients of the Magnus expansion in [8, Eq. (2.14)] can be made explicit. In the Magnus expansion in [8], the terms are indexed by certain binary trees and the coefficients are given inductively. In (4.2), the terms are indexed by $\mathcal{I}'_v$, which is more tractable, and the coefficients are explicit.

References

[1] F. Faà di Bruno, *Note sur un nouvelle formule de calcul différentiel*, Quarterly Journal of Pure and Applied Mathematics **1** (1857), 359–360.

[2] G. M. Constantine and T. H. Savits, *A multivariate Faà di Bruno formula with applications*, Trans. Amer. Math. Soc. **348** (1996), 503–520. MR1325915 (96g:05008)

[3] A. D. D. Craik, *Prehistory of Faá di Bruno's formula*, Amer. Math. Monthly **112** (2005), 119–130. MR2121322 (2006a:40001)

[4] W. Gröbner, *Die Lie-Reihen und ihre Anwendungen*, Deutscher Verlag der Wissenschaften, Berlin, 1967. MR0217392 (36:482)

[5] O. M. Gvozdetskiĭ and V. P. Igumnov, *Representation of solutions of ordinary differential equations in the form of Lie series*, Ukrainian Math. J. **38** (1986), 192–194 (English translation of Ukrain. Mat. Zh.). MR841056 (87e:34011)

[6] M. Hardy, *Combinatorics of partial derivatives*, Electron. J. Combin. **13** (2006), #R1. MR2200529 (2007d:26007)

[7] V. P. Igumnov, *Representation of solutions of differential equations by modified Lie series*, Differential Equations **20** (1984), 688–694 (English translation of Differentsial′nye Uravneniya). MR751841 (86a:34016)

[8] A. Iserles and S. P. Nørsett, *On the solution of linear differential equations in Lie groups*, R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci. **357** (1999), 983–1019. MR1694700 (2000d:34022)

[9] N. Jacobson, *Lectures in Abstract Algebra. Vol. I*, Springer-Verlag, New York-Heidelberg, 1975, p. 19. MR0392227 (52:13044)

[10] W. P. Johnson, *The curious history of Faà di Bruno's formula*, Amer. Math. Monthly **109** (2002), 217–234. MR1903577 (2003d:01019)

[11] W. Magnus, *On the exponential solution of differential equations for a linear operator*, Comm. Pure Appl. Math. **7** (1954), 649–673. MR0067873 (16:790a)

[12] R. Most, *Ueber die höheren differentialquotienten*, Mathematische Annalen **4** (1871), 499–504.

[13] I. Niven, *Formal power series*, Amer. Math. Monthly **76** (1969), 871–889. MR0252386 (40:5606)

Department of Mathematics, University of South Florida, Tampa, Florida 33620
E-mail address: xhou@math.usf.edu