INTEGERS: ELECTRONIC JOURNAL OF COMBINATORIAL NUMBER THEORY 7 (2007), #A34

AN INVERSE OF THE FAÀ DI BRUNO FORMULA

Gottlieb Pirsic[1]
Johann Radon Institute of Computational and Applied Mathematics (RICAM), Austrian Academy of Sciences, Altenbergerstr. 69, 4040 Linz, Austria
gottlieb.pirsic@oeaw.ac.at

Received: 2/7/07, Revised: 5/25/07, Accepted: 6/25/07, Published: 7/20/07

Abstract

For sufficiently differentiable univariate functions f, g and their composite $h = g(f)$ we prove that

$$\frac{d^n g(f(x))}{df(x)^n} = \sum_{j=0}^{n} \frac{d^j h(x)}{dx^j} \sum_{k=0}^{n-j} (-1)^k \frac{df(x)}{x}^{-(n+k)} R_{n,j,k}(x),$$

$$R_{n,j,k}(x) = \sum_{\substack{b_1+\cdots+b_k= \\ =n+k-j, \\ b_i \geq 2}} \frac{1}{k!} \binom{n+k-1}{j-1, b_1, \ldots, b_k} \frac{d^{b_1} f(x)}{dx^{b_1}} \cdots \frac{d^{b_k} f(x)}{dx^{b_k}}$$

using combinatorial considerations (set partitions in particular). We also discuss the case of multivariate f.

1. Introduction

Faà di Bruno's formula, the chain rule of higher derivatives, can be given in a convenient form as

$$\frac{d^k g(f(x))}{dx^k} = \sum_{i=0}^{k} B_{k,i}(f(x), f'(x), \ldots, f^{(k-i+1)}(x)) \frac{d^i g(f(x))}{df(x)^i},$$

where the $B_{k,i}$ are the Bell polynomials (see e.g., [3, Section 2]).

Our aim is to find rational functions $C_{k,i}$ such that

$$\frac{d^k g(f(x))}{df(x)^k} = \sum_{i=0}^{k} C_{k,i}(f(x), f'(x), \ldots, f^{(k-i+1)}(x)) \frac{d^i g(f(x))}{dx^i}.$$

[1]Research partly supported by the FWF project P19004-N18 and the Austrian grant BRST-29.

Setting $h(x) = g(f(x))$ (and taking care of the proper derivation variables), we can say the Faà di Bruno formula expresses $h^{(k)}(x)$ as a linear combination of $g^{(i)}(f(x))$, whilst we try the reverse, to express $g^{(k)}(f(x))$ as a linear combination of the $h^{(i)}(x)$.

The problem has been treated previously by Hess [1], of which there is a summary review in [2]. The motivation of this paper is to give an alternative proof and additionally to give some more insight into the combinatorial aspects of the problem.

Faà di Bruno's formula, especially its multivariate form, admits a nice combinatorial interpretation in terms of set partitions that are in a one-to-one relation with monomials in the derivatives [5]. If, as a kind of inverse relation, we try to express the derivatives $d^k g(f(x))/df(x)^k$ in terms of $d^k g(f(x))/dx^k$, similar combinatorial relations hold, but the partition structure is a bit more involved.

2. Set Partitions

We define $\pi_{n,j,k}$ as the set of partitions of the set $\{1, \dots, n+k-1\}$ into one distinguished block of length $j-1 \geq 0$ and k blocks of length at least 2. Here the blocks are not necessarily adjacent.

Let us examine the extreme cases.

- $n = j = k = 0$: The empty set admits exactly one partition, namely the empty set itself: $\pi_{0,0,0} = \{\{\emptyset\}\} = \{\{\{\}\}\}$.
- $j \leq 0$: Since we would need the distinguished set to be of negative length, we get $\pi_{n,0,k} = \emptyset$ unless we are partitioning the empty set, i.e., $n = k = 0$ which we already considered.
- $k \leq 0$: Setting $k = 0$, for $\pi_{n,j,0}$ we have to consider a partition of $\{1, \dots, n-1\}$ into one set of length $j-1$ and no additional sets. So for $n = j$ we get $\pi_{n,n,0} = \{\{\{1, \dots, n-1\}\}\}$ and $\pi_{n,j,0} = \emptyset$ else.
- $j + k \geq n + 1$: We would have to partition $\{1, \dots, n+k-1\}$ into one set of $j-1$ elements and k blocks that total to $n - j + k \leq 2k - 1$ which is not possible since these blocks are required to have length at least 2.

As more generic examples, consider

$$\pi_{2,1,1} = \{\{\emptyset, \{1,2\}\}\} \text{ and}$$
$$\pi_{3,2,1} = \{\{\{1\}, \{2,3\}\}, \{\{2\}, \{1,3\}\}, \{\{3\}, \{1,2\}\}\}.$$

One of the key points of the proof is that the $\pi_{n,j,k}$ can be defined recursively in n. We make a case distinction according to the set in which the new element $n+k-1$ is contained. (This section and the next are modeled closely after [4]).

1. This set may be the distinguished set of length $j-1$. Removing it from this set, we are left with a set partition from $\pi_{n-1,j-1,k}$.
2. The new element may be in a two-element set. In this case we have the choice of $n+k-2$ elements for the second member of that set. Removing this two-element set from the partition we are left with a partition that is in one-to-one correspondence with an element of $\pi_{n-1,j,k-1}$ (one set less, and the special set still has the same length).
3. Finally, the new element may appear in a non-distinguished set of size ≥ 3. The old partition will then have been an element of $\pi_{n-1,j,k}$.

3. Derivative Monomials

As next step we introduce the map between set partitions and monomials. To a partition $\pi = \{\pi_0, \dots, \pi_k\} \in \pi_{n,j,k}$, where without loss of generality π_0 is the special set of cardinality $j-1$, we associate the monomial $f^{(\pi)} := f^{(|\pi_1|)} f^{(|\pi_2|)} \cdots f^{(|\pi_k|)}$. Setting $R_{n,j,k} := \sum_{\pi \in \pi_{n,j,k}} f^{(\pi)}$, we use the recursive relations of $\pi_{n,j,k}$ to derive a recursion for $R_{n,j,k}$.

1. The partitions of $\pi_{n-1,j-1,k}$ correspond directly to the terms of $R_{n-1,j-1,k}$, so these are terms of $R_{n,j,k}$ as well (since there is change only in the special set that does not affect the monomials).
2. A two-element set of the partition corresponds to a term $f''(x)$, hence the terms corresponding to this case are $(n+k-2)f''(x)R_{n-1,j,k-1}$.
3. The term in this case is $(d/dx)R_{n-1,j,k}$, since derivation adds one summand for each factor of a monomial with the differentiation increased by one for that factor. This increase mirrors the addition of the new element to a set of the partition.

All in all, from these considerations we get

$$R_{n,j,k} = R_{n-1,j-1,k} + (n+k-2)f'' R_{n-1,j,k-1} + \frac{d}{dx} R_{n-1,j,k}. \tag{1}$$

4. Result

Theorem 1 *For sufficiently differentiable univariate functions f, g and $h = g \circ f$*

$$g^{(n)}(f(x)) = \sum_{j=0}^{n} h^{(j)}(x) \sum_{k=0}^{n-j} (-1)^k f'(x)^{-(n+k)} R_{n,j,k}(x), \tag{2}$$

$$R_{n,j,k}(x) = \sum_{\substack{b_1+\cdots+b_k=\\=n+k-j,\\ b_i \geq 2}} \frac{1}{k!} \binom{n+k-1}{j-1, b_1, \ldots, b_k} f^{(b_1)}(x) \cdots f^{(b_k)}(x),$$

where $R_{n,0,k}(x) := \mathbb{1}\big(n = k = 0\big)$, $R_{n,j,0}(x) := \mathbb{1}\big(n = j\big)$ *and* $R_{n,j,k}(x) := 0$ *else.*

Proof. The proof is done by induction (again, compare [4]). To begin with, observe that the $R_{n,j,k}$ given in the statement is just the explicit version of the combinatorial version given previously, and that the definitions of $R_{n,j,0}, R_{n,0,k}$ are compatible with the combinatorial definition and the usual conventions about empty sets, sums and products.

Applying $d/df(x) = 1/(df(x)/dx) \cdot d/dx$ to the equation gives (we use $n^* = n+1, j^* = j+1, k^* = k+1$)

$$\begin{aligned}
g^{(n+1)}(f(x)) =& \sum_{j=0}^{n} h^{(j+1)}(x) f'(x)^{-1} \sum_{k=0}^{n-j} (-1)^k f'(x)^{-(n+k)} R_{n,j,k} \\
&+ \sum_{j=0}^{n} h^{(j)}(x) \sum_{k=1}^{n-j} (-1)^k (-n-k) f'(x)^{-(n+k+2)} f''(x) R_{n,j,k} \\
&+ \sum_{j=0}^{n} h^{(j)}(x) \sum_{k=1}^{n-j} (-1)^k f'(x)^{-(n+k+2)} f'(x) \frac{d}{dx} R_{n,j,k} \\
=& \sum_{j^*=0}^{n^*} h^{(j^*)}(x) \sum_{k=0}^{n^*-j^*} (-1)^k f'(x)^{-n^*-k} R_{n^*-1,j^*-1,k} \\
&+ \sum_{j=0}^{n^*} h^{(j)}(x) \sum_{k^*=0}^{n^*-j} (-1)^{k^*-1} (-n^* - k^* + 2) \cdot \\
&\qquad \cdot f'(x)^{-n^*-k^*} f''(x) R_{n^*-1,j,k^*-1} \\
&+ \sum_{j=0}^{n^*} h^{(j)}(x) \sum_{k=0}^{n^*-j} (-1)^k f'(x)^{-n^*-k} \frac{d}{dx} R_{n^*-1,j,k} \\
=& \sum_{j^*=0}^{n^*} h^{(j)}(x) \sum_{k^*=0}^{n^*-j^*} (-1)^{k^*} f'(x)^{-n^*-k^*} R_{n^*,j^*,k^*}.
\end{aligned}$$

In the last equation we made use of the recursion for $R_{n,j,k}$.

Note that we did indeed not introduce any additional terms when extending the summation ranges (this follows from the discussion of the extreme cases of $\pi_{n,j,k}$). E.g., in the second sum (where the factor $(-n^*-k^*+2)$ appears) we have additional summands with $j = n+1$ or $k = -1$ or $k = n-j+1$. However, in all these cases the corresponding partition is empty, hence $R_{n,j,k} = 0$.

To conclude the proof we need to check the induction start. For $n = 0$ we get

$$g(f(x)) = h^{(0)}(x)f'(x)^0(-1)^0R_{0,0,0} = h(x),$$

which is a trivially true assertion (by usual conventions, e.g., $f'(x)^0 \equiv 1$). For $n > 0$, the $j = 0$ term is always zero by definition, so for $n = 1$ we get

$$g'(f(x)) = h^{(1)}(x)(-1)^0f'(x)^{-1-0}R_{1,1,0} = h'(x)/f'(x),$$

which is also true. Finally let us also check $n = 2$, where we have

$$\begin{aligned}
g''(f(x)) &= \sum_{j=1}^{2} h^{(j)}(x) \sum_{k=0}^{2-j} (-1)^k f'(x)^{-2-k} R_{2,j,k} \\
&= \frac{h'(x)}{f'(x)^2}R_{2,1,0} - \frac{h'(x)}{f'(x)^3}R_{2,1,1} + \frac{h''(x)}{f'(x)^2}R_{2,2,0} \\
&= 0 - \frac{h'(x)f''(x)}{f'(x)^3} + \frac{h''(x)}{f'(x)^2}.
\end{aligned}$$

This ends the proof. □

Observe that if we take $g = f^{-1}$, i.e., the inverse function of f, we get the formula for the higher derivatives of the inverse function (see e.g., [4]). Perhaps other choices of g may lead to interesting special cases as well.

5. The Multivariate Case

Now we consider a function f in n variables, a function g in one variable and let $h = g(f)$ be their composite. We introduce the abbreviations

$$g_i := \frac{d^i}{df(x)^i}g(f(x)), \quad h_i = h_i(x_1, \ldots, x_n) := \frac{\partial^i}{\partial x_1 \cdots \partial x_i}h(x_1, \ldots, x_n),$$

and

$$f_{i_1 \cdots i_m} = f_{\{i_1,\ldots,i_m\}} := \frac{\partial^{i_1+\cdots+i_m}}{\partial x_{i_1} \cdots \partial x_{i_m}}f(x_1, \ldots, x_m).$$

Note the special cases $g_0 = g(f(x)) = h(x) = h_0$, and $f_\emptyset = f$.

For a set of partitions T we define

$$f_T := \sum_{\tau\in T}\prod_{\sigma\in\tau} f_\sigma, \qquad \text{e.g., } f_{\{\{\{1,2\},\{3\}\},\{\{1,2,3\}\}\}} = f_{12}f_3 + f_{123}.$$

Let $\tau_{n,j}$ denote the set of all partitions of $\{1,\dots,n\}$ into j nonempty blocks. We point out the special cases $f_{\tau_{0,0}} = 1, f_{\tau_{n,0}} = 0$ for $n > 0$ which follow from $\tau_{0,0} = \{\emptyset\}$ and $\tau_{n,0} = \emptyset$ for $n > 0$ and the usual conventions on empty sums and products.

With these notations the multivariate forward Faà di Bruno formula (see e.g., [5]) can be written very concisely.

$$h_n = \sum_{j=0}^{n} g_j f_{\tau_{n,j}} \quad \left(= \sum_{j=0^n} \frac{d^j g(f(x))}{df(x)^j} \sum_{\tau\in\tau_{n,j}} \prod_{\sigma\in\tau} \frac{\partial^{|\sigma|} f(x)}{\prod_{i\in\sigma}\partial x_i}\right).$$

To state the inverse case in a similar manner we need some further notation. We set $\theta_{n,k} := \prod_{i=k}^{n} f_{\tau_{i,i}}$, so e.g.,

$$\theta_{3,2} = f_1^2 f_2^2 f_3, \quad \theta_{5,3} = f_1^3 f_2^3 f_3^3 f_4^2 f_5.$$

Further set

$$\rho_{n,k} := \nVdash(n=k) + \sum_{\substack{\{u_{k+1},\dots,u_n\}= \\ =\{k,\dots,n-1\}, \\ u_j\le j}} (-1)^{|\{i:k<i\le n, i\ne u_i\}|} \prod_{i=k+1}^{n} f_{\tau_{i,u_i}}.$$

Theorem 2 *The inverse multivariate Faà di Bruno formula is given by*

$$g_n = \sum_{k=0}^{n} h_k \frac{\rho_{n,k}}{\theta_{n,k}}. \tag{3}$$

This is actually a simple corollary of the following formula for the inverse of a lower triangular matrix.

Theorem 3 *Let $a_{i,j}, (i,j = 1,\dots,n)$ be the entries of a lower triangular matrix, denote the entries of its inverse matrix by $b_{i,j}$ and define $\alpha_{i,j} := 1/(a_{j,j}\cdots a_{i,i})$. Then*

$$b_{i,j} = \alpha_{i,j}\Big(\nVdash(i=j) + \sum_{\{u_k\}\in H_{i,j}} \prod_{k=j+1}^{i} (-1)^{\nVdash(u_k\ne k)} a_{k,u_k}\Big), \tag{4}$$

where $H_{i,j}$ is the set of all $(u_{j+1},\dots,u_i)$ that are permutations of $(j,\dots,i-1)$ and fulfill $u_k \le k$.

Proof. We use induction on $i-j$ to prove this where the base cases $b_{i,i} = 1/a_{i,i}$ and $b_{i,i-1} = -a_{i,i-1}/(a_{i-1,i-1}a_{i,i})$ correctly coincide with the formulas.

Now since $(b_{i,j})$ is the inverse of the lower triangular matrix $(a_{i,j})$ we get the recursion

$$\Vdash(i=j) = \sum_{k=1}^{n} a_{i,k}b_{k,j} = \sum_{k=j}^{i} a_{i,k}b_{k,j} \iff$$

$$a_{i,i}b_{i,j} = \Vdash(i=j) - \sum_{k=j}^{i-1} a_{i,k}b_{k,j}.$$

Using the induction hypothesis this equals

$$a_{i,i}b_{i,j} = -\sum_{k=j}^{i-1} a_{i,k}\alpha_{k,j} \sum_{\{u_l\}\in H_{k,j}} \prod_{l=j+1}^{k} (-1)^{\Vdash(j_k\neq k)} a_{l,u_l}.$$

We will employ the following recursion for $H_{i,j}$, the set of all admissible $(u_{j+1},\dots,u_i)$. The recursion is also in $i-j$, so the initial values are $H_{i,i}=\emptyset$. Now observe that if $u_i = k$ it is necessary that $u_{k+1} = k+1,\dots,u_{i-1} = i-1$ because of the condition $u_t \leq t$. The remaining $(u_{j+1},\dots,u_k)$ fulfill the conditions for $H_{k,j}$, so

$$H_{i,j} = \bigcup_{k=j}^{i-1} \big\{(\mathbf{h}, k+1, k+2, \dots, i-2, i-1, k),\ \mathbf{h}\in H_{k,j}\big\}.$$

(The elements of the union set are made up of these three components: $\mathbf{h}$, a range of integers and the final k, where the first two components might be empty.) It follows that

$$\begin{aligned} a_{i,i}b_{i,j} &= \alpha_{i-1,j}\sum_{k=j}^{i-1}\sum_{\{u_l\}\in H_{k,j}} -a_{i,k}a_{i-1,i-1}\cdots a_{k+1,k+1}\prod_{l=j+1}^{k}(-1)^{\Vdash(j_k\neq k)}a_{l,u_l} \\ &= \alpha_{i-1,j}\sum_{\{u_l\}\in H_{i,j}}^{*}\prod_{l=j+1}^{i}(-1)^{\Vdash(j_k\neq k)}a_{l,u_l}, \end{aligned}$$

which gives the stated formula after division by $a_{i,i}$. □

Returning to the Faà di Bruno formula, the forward case may be written as $\boldsymbol{h} = M\boldsymbol{g}$, where $\boldsymbol{h} = (h_0,\dots,h_n)^\top, \boldsymbol{g} = (g_0,\dots,g_n)^\top$ are the vectors of derivatives of h, g respectively and $M = (f_{\tau_{i,j}})_{i,j\geq 0}^{n}$ is a lower triangular matrix. So (4) applies and by $\boldsymbol{g} = M^{-1}\boldsymbol{h}$, equation (3) follows directly from (4).

To better illustrate the structure of the formulas we give examples for $n=3$

$$\begin{aligned} g_3 =& \frac{h_3}{f_1f_2f_3} - \frac{h_2}{f_1^2f_2^2f_3}\big(f_1f_{23}+f_2f_{13}+f_3f_{12}\big) \\ &+ \frac{h_1}{f_1^3f_2^2f_3}\big(f_{12}(f_1f_{23}+f_2f_{13}+f_3f_{12}) - f_1f_2f_{123}\big) \\ =& \frac{h_3}{\theta_{3,3}} - \frac{h_2}{\theta_{3,2}}f_{\tau_{3,2}} + \frac{h_1}{\theta_{3,1}}\big(f_{\tau_{2,1}}f_{\tau_{3,2}} - f_{\tau_{2,2}}f_{\tau_{3,1}}\big), \end{aligned}$$

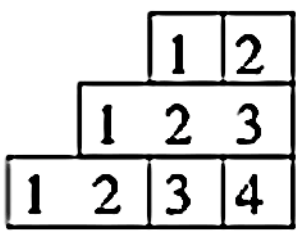
and for $n = 4$

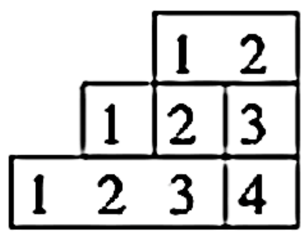
$$\begin{aligned} g_4 =& \frac{h_4}{\theta_{4,4}} - \frac{h_3}{\theta_{4,3}} f_{\tau_{4,3}} + \frac{h_2}{\theta_{4,2}} \big(f_{\tau_{3,2}} f_{\tau_{4,3}} - f_{\tau_{3,3}} f_{\tau_{4,2}} \big) \\ &- \frac{h_1}{\theta_{4,1}} \big(f_{\tau_{2,1}} f_{\tau_{3,2}} f_{\tau_{4,3}} - f_{\tau_{2,2}} f_{\tau_{3,1}} f_{\tau_{4,3}} - f_{\tau_{2,1}} f_{\tau_{3,3}} f_{\tau_{4,2}} + f_{\tau_{2,2}} f_{\tau_{3,3}} f_{\tau_{4,1}} \big) \end{aligned}$$

(this last formula has 45 terms in full notation, $n = 5$ has 363).

5.1. Remarks

There is a difference to the forward case, since the coefficients of the derivative monomials are not only ± 1 but can assume other integer values. For instance, in g_4 the term $f_M, M = \{1, 2, 3, 4, 12, 123\}$ appears with coefficient $+2$ since f_M is a term of $f_{\tau_{2,i}} f_{\tau_{3,j}} f_{\tau_{4,k}}$ for $\{i, j, k\}$ equal to $\{2, 1, 3\}$ and $\{1, 3, 2\}$. This can be expressed by writing M in two different tableaus built from the same blocks.





It would be interesting to find a general formula for the coefficients.

There is an alternative form for $H_{i,j}$ that may be more convenient for generating all possible $\{u_k\}$. Any admissible $\{u_k\}$ is in one-to-one correspondence with strictly monotone sequences: as discussed, $u_i < i$ and $u_k = k$ for $k = u_i + 1, \ldots, i - 1$. Set $k_0 = u_i$. Then also $u_{k_0} < k_0$ and again $u_k = k$ for $k = u_{k_0} + 1, \ldots, k_0 - 1$. We obtain a strictly descending sequence $k_0 = i, k_1, k_2, \ldots$ until $u_{k_t} = j$. Its length can range from 1 to $i - j$. On the other hand any such sequence gives rise to an admissible $\{u_k\}$ by setting $u_i = k_0, u_{k_t} = k_{t+1}$ and $u_k = k$ for the remaining k. So

$$\begin{aligned} H_{i,j} :=& \Big\{ (u_{j+1}, \ldots, u_i) \Big| \{u_{j+1}, \ldots, u_i\} = \{j, \ldots, i-1\}, u_k \le k \Big\} \\ =& \bigcup_{m=1}^{i-j} \bigcup_{\substack{k_0, \ldots, k_m \\ i = k_0 > \cdots > k_m = j}} \Big\{ \{u_k\} \Big| u_{k_t} = k_{t+1}, 0 \le t < m, \, u_k = k \text{ else} \Big\} \\ =& \bigcup_{m=1}^{i-j} \bigcup_{\substack{k_0, \ldots, k_m \\ j = k_0 < \cdots < k_m = i}} \Big\{ \{u_k\} \Big| u_{k_t} = k_{t-1}, 1 \le t \le m, \, u_k = k \text{ else} \Big\}. \end{aligned}$$

As one of the referees pointed out, there is yet another, even simpler alternative view. Observe that there is also a one-to one correspondence between $H_{i,j}$ and the subsets of $\{j+1,\ldots,i-1\}$. Given a set of $H_{i,j}$ consider the subset of indices such that $u_s = s$. Conversely, given any subset, set $u_s = s$ for all indices in this subset, the remaining values of u_s are uniquely determined by considerations as in the previous paragraph. As a consequence of this we also immediately see that the size of $H_{i,j}$ is 2^{i-j-1} if $i > j$ and 0 for $i \leq j, (i,j) \neq (0,0)$.

Finally note that if some x_i are identified, for example, $f(x_1, x_2, x_3) = f(x_1, x_2, x_2) = f(x_1, x_3, x_3) = f(x_1, x_2)$, the formula will stay valid. In particular, formula (2) is the special case $f(x_1, \ldots, x_n) = f(x, \ldots, x)$ of (3). To obtain the proper multiplicity (i.e., the coefficients) of the occurring terms after identifications, possibly the enumeration considerations of [5] can be applied here as well. However, as stated above, we do not even know the coefficients of the general case; perhaps then this is a way to obtain these unknown quantities.

Acknowledgments

I would like to thank the anonymous referees for their comments, which helped to improve the content and presentation of the paper.

References

[1] E. Hess, Zur Theorie der Vertauschung der unabhängigen Variablen, Zeitschrift für Mathematik und Physik, **17**, 1872, 1–12.

[2] T.Muir, The Theory of Determinants in the Historical Order of Development, Vol. **3**, Macmillan, London, 1920 [Reprint Edition: Dover, New York, 1960], 226–227.

[3] W.P. Johnson, The Curious History of Faà di Bruno's Formula. Am. Math. Monthly **109**(3), 2002, 217–234.

[4] W.P. Johnson, Combinatorics of Higher Derivatives of Inverses. Am. Math. Monthly **109**(3), 2002, 273–277.

[5] M. Hardy, Combinatorics of Partial Derivatives, Elect. J. of Combinat. **13**, 2006.