The Newton interpolation formula, with more variables

Alain Lascoux ⋆

Newton addressed the question of transforming discrete sets of data, say the positions of planets at different times, into algebraic functions, before submitting them to the differential calculus that he had just constructed.

Of course, anybody faced with the sequence $1, 2, 4, 8, 16, \ldots$ will exclaim 32, except for the unwise Khalif who pledged to fill the squares of a chess-board with wheat, doubling the number of grains from one square to the other. To recognize the function $f(n) = an^2 + bn + c$, $n = 1, 2, \ldots$, which is not much more complicated, one already needs to have recourse to *finite differences*, that is, to evaluate the functions $g(n) := f(n) - f(n-1)$ and iterate. Indeed, polynomials can be characterized by the fact that iterating finite differences ultimately produces the constant sequence $0, 0, 0, \ldots$.

However, comets are not likely to appear at regularly spaced lapses of time, and to handle their seemingly erratic apparitions, Newton found the solution of normalizing differences of positions by the interval of time to which they correspond.

In other words, he writes, starting from the table $f(t_1)$, $f(t_2)$, $f(t_3)$, $\ldots$ the normalized differences

$$[1,2] = \frac{f(t_1) - f(t_2)}{t_1 - t_2} \ , \ [2,3] = \frac{f(t_2) - f(t_3)}{t_2 - t_3} \ , \ [3,4] = \frac{f(t_3) - f(t_4)}{t_3 - t_4} \ldots$$

then the differences

$$[1,2,3] = \frac{[1,2] - [2,3]}{t_1 - t_3} \ , \ [2,3,4] = \frac{[2,3] - [3,4]}{t_2 - t_4}, \ldots$$

Since already at the first step one obtains functions of two variables, the proper way to interpret Newton's operations is to define *divided differences* ∂_i as operators on functions of, say $b_1, b_2, b_3, \ldots$:

$$\partial_i \ : \ f \longrightarrow \frac{f - f^{s_i}}{b_i - b_{i+1}} \ , \ i = 1, 2, \ldots \ ,$$

denoting by s_i the transposition of b_i and b_{i+1}.

Divided differences are "local" (they act only on two variables at a time), rational (the coefficients $1/(b_i - b_{i+1})$ are rational), and decrease degree by 1. More general operators of the same kind can be found in [LS2].

⋆ text written during the Conference *Applications of the Macdonald Polynomials*, at the Newton Institute in April 2001.

They provide a discrete version of differential calculus, and one would have avoided two centuries of controversy, initiated by Bishop Berkeley, by sticking to them instead of introducing such puzzling entities as epsilons vanishing at different orders.

Given a function of one variable, let

$$f^{\partial}\,,\, f^{\partial\partial}\,,\, f^{\partial\partial\partial}\,,\, \ldots$$

denote

$$\partial_1(f(b_1))\,,\, \partial_2\big(\partial_1(f(b_1))\big)\,,\, \partial_3\Big(\partial_2\big(\partial_1(f(b_1))\big)\Big)\,,\, \ldots$$

respectively.

Then the formula of Newton to interpolate $f(t)$ from its values at time $b_1,\, b_2,\, \ldots$ is

$$f(t) = \\ f(b_1)+f^{\partial}\,(t-b_1)+f^{\partial\partial}\,(t-b_1)(t-b_2)+f^{\partial\partial\partial}\,(t-b_1)(t-b_2)(t-b_3)+\cdots \qquad (\star)$$

the expansion being exact if $n+1$ points b_i are used and f is a polynomial of degree $\leq n$.

Modernists will object that the set of polynomials

$$\{t^0,\, (t-b_1),\, (t-b_1)(t-b_2),\, \ldots\}$$

is triangular in the usual basis $\{t^0,\, t^1,\, \ldots\}$, and therefore, there is no mystery in the fact that it is a linear basis of polynomials. However, what is remarkable in Newton's formula is the process by which he obtained the coefficients of his polynomials $(t-b_1)\cdots(t-b_n)$.

When $b_1,\, b_2,\, \ldots$ all collapse to 0, then $(\star)$ becomes *Taylor's formula*

$$f(t) = f(0) + f'(0)\, t + f''(0)/2!\, t^2 + \cdots\ ,$$

the factorials in the denominator being clear by putting $b_i = i\epsilon$ and letting ϵ tend to 0.

Since divided differences are the discrete analogues of derivatives, the question is :

What is the discrete analogue of Taylor's formula in several variables ?

In other words, what are the coefficients of the images of $f(a_1, a_2, \ldots)$ under the possible different divided differences (evaluated in $b_1, b_2, b_3, \ldots$) in the expansion of $f(a_1, a_2, \ldots)$?

The answer, that I obtained with Marcel-Paul Schützenberger some years ago (cf. for example [LS1]), is as simple as in the case of one variable. The universal coefficients, called *Schubert polynomials* due to their relevance to geometry, can be defined recursively as follows:

1) They are polynomials $Y_v(\mathbb{A}, \mathbb{B})$ in two infinite totally ordered sets of variables $\mathbb{A} = \{a_1, a_2, \ldots\}$, $\mathbb{B} = \{b_1, b_2, \ldots\}$, indexed by vectors $v \in \mathbb{N}^\infty$ (from now on, we shall suppose that they have only a finite number of non-zero components).

2) They are globally stable under divided differences in the a_i's, that is, the image of a Schubert polynomial is either 0 or a Schubert polynomial.

3) When v is weakly decreasing (one says v is *dominant*, or is a *partition*, deleting the terminal zeros), then

$$Y_v = \prod_{(i,j) \in Diagr(v)} (a_i - b_j) \ , \tag{$\diamond$}$$

product on all boxes in the diagram of v: one stacks $v_1, v_2, v_3, \ldots$ boxes in successive rows, packing them to the left. A box in position (i, j) gives a factor $(a_i - b_j)$.

4) Given $v \in \mathbb{N}^\infty$ and $i \in \mathbb{N}$ such that $v_i > v_{i+1}$, writing vs_i for $[v_1, \ldots, v_{i-1}, v_{i+1}, v_i - 1, v_{i+2}, \ldots]$, one has

$$\partial_i(Y_v) = Y_{vs_i} \ . \tag{$\diamond\diamond$}$$

One does not write what happens when $v_i \leq v_{i+1}$, because in that case Y_v is the image of some $Y_{v'}$ under ∂_i, (precisely $v' = [v_1, \ldots, v_{i-1}, v_{i+1} + 1, v_i, v_{i+2}, \ldots]$). The vanishing $\partial_i^2 = 0$ implies that $\partial_i(Y_v) = 0$ in that case.

On the indices of Schubert polynomials, the action of divided differences amounts to sorting (and decreasing), in other words, amounts to use the symmetric group. Thus to check whether the equations ($\diamond\diamond$) are consistent essentially reduces to check that

$$\partial_i \partial_{i+1} \partial_i (Y_v) = \partial_{i+1} \partial_i \partial_{i+1} (Y_v) \tag{$\clubsuit$}$$

for any i, any v, because $s_i s_{i+1} s_i = s_{i+1} s_i s_{i+1}$, and $s_i s_j = s_j s_i$, $|j - i| \neq 1$ are the defining equations of the symmetric group, as generated by simple transpositions s_i. There is no need to check $\partial_i \partial_j (Y_v) = \partial_j \partial_i (Y_v)$ because operations on two disjoint pairs of variables clearly commute.

Relation ($\clubsuit$) is in fact a property of divided differences, substituting in it Schubert polynomials is irrelevant . One indeed has

$$\partial_i \partial_{i+1} \partial_i = \partial_{i+1} \partial_i \partial_{i+1} \ , \ i \geq 1 \tag{$\clubsuit\clubsuit$}$$

To check this, one can without loss of generality take i to be 1. Functions of $a_1, a_2, a_3, \ldots$ can be expressed as linear combinations of the six monomials

$$1, \ a_1, \ a_2, \ a_1^2, \ a_1 a_2, \ a_1^2 a_2 \ ,$$

with coefficients which are functions symmetrical in a_1, a_2, a_3 and arbitrary in $a_4, a_5, \ldots$.

Now, the divided differences ∂_1, ∂_2 commute with multiplication by symmetric functions in a_1, a_2, a_3, and thus the identity $\partial_1\partial_2\partial_1 = \partial_2\partial_1\partial_2$ needs only to be tested on the above set. However, both $\partial_1\partial_2\partial_1$ and $\partial_2\partial_1\partial_2$ decrease the degree by 3. Thus it is only necessary to check, and we leave this pleasure to the reader, that $\partial_1\partial_2\partial_1(a_1^2 a_2)$ and $\partial_2\partial_1\partial_2(a_1^2 a_2)$ are the same constant (which is 1), the other images being null for reason of degree.

To get a general v, one can start from different possible dominant vectors. Because the set of partitions is a lattice (with respect to intersection and union of diagrams), compatibility of the different choices boils down to the following lemma.

Lemma. *Given a dominant* $v \in \mathbb{N}^\infty$, *and* $n \in \mathbb{N}$ *such that* $v \le \rho := [n-1, n-2, \ldots, 1, 0, 0, \ldots]$ *(componentwise), then there exists at least one chain of divided differences such at each step a* ∂_i *is applied to a dominant Schubert polynomial of the type* $(a_i - b_j)\, g$, *with* g *symmetrical in* a_i, a_{i+1} *(and therefore the image is a dominant Schubert polynomial corresponding to a diagram with the box of coordinate* (i, j) *erased, the factor* $(a_i - b_j)$ *becoming* 1 *under* ∂_i*).*

The proof of the lemma consists in peeling off boxes of a staircase diagram, in such a way as to have, at a step where ∂_i will be used, columns i and $i+1$ of lengths differing by 1. This can be realized by erasing boxes from top to bottom, in successive diagonals, as indicated by the following example for $v = [6, 6, 6, 2, 2, 0, 0, \ldots]$, $n = 9$ (boxes are numbered in the order they are peeled off).

1							
8	2						
13	9	3					
		10	4				
		14	11	5			
						6	
						12	7

Now, if Y_ρ was the product $\prod_{i,j,i+j\le n}(a_i - b_j)$, then Y_v, as obtained by the above process, will be the product of factors $(a_i - b_j)$ corresponding to the boxes which are left. Therefore, the definition of Schubert polynomials is consistent.

Newton's polynomials could be characterized by their vanishing properties, because they are written in terms of their roots. Similarly, when v is dominant, and different from $[0, 0, \ldots]$, one can also easily write the vanishing properties of Y_v.

Indeed we shall only need the property that the specialization $a_1 = b_1$, $a_2 = b_2$, $a_3 = b_3, \ldots$ in $Y_v(\mathbb{A}, \mathbb{B})$, which we shall write as $Y_v(\mathbb{B}, \mathbb{B})$, vanishes (all the factors in the main diagonal of the diagram vanish).

More generally, one can prove (this will be the only property in this text that we admit. Otherwise, we would need to appeal to properties of the Ehresmann-Bruhat order on the symmetric group, for which we refer to Macdonald [M]), that

$$v \in \mathbb{N}^\infty,\ v \neq [0, 0, \ldots] \Rightarrow Y_v(\mathbb{B}, \mathbb{B}) = 0\ . \qquad (\star\star)$$

The relation $\partial_1\partial_2\partial_1 = \partial_2\partial_1\partial_2$ shows that products of divided differences are not independent. We shall only use the following products (which are, in fact, canonical representatives of all possible products of divided differences). Given $K = [k_1, k_2, \ldots] \in \mathbb{N}^n$, let

$$\partial^k := (\partial_1 \cdots \partial_{k_1})\,(\partial_2 \cdots \partial_{k_2+1})\,(\partial_3 \cdots \partial_{k_3+2})\ \ldots\ ,$$

that is, ∂^k consists in products of blocks of consecutive divided differences, of respective lengths $k_1, k_2, \ldots$, starting with ∂_1, ∂_2, $\ldots$ respectively, *divided differences operating on their left.*

We can now state the multivariate Newton interpolation formula. For any polynomial in $a_1, a_2, \ldots$, one has :

$$f(\mathbb{A}) = \sum_{K \in \mathbb{N}^\infty} f^{\partial^K}(\mathbb{B})\, Y_K(\mathbb{A}, \mathbb{B})\ . \qquad (\star\star\star)$$

Proof. One has to test $(\star\star\star)$ on a linear basis of polynomials in $\mathbb{A}$ (with coefficients in anything, for example functions of $\mathbb{B}$). One takes the Schubert polynomials $\{Y_J\}$ as such a basis. The question is : Is is true that

$$Y_J(\mathbb{A}, \mathbb{B}) = \sum_{K \in \mathbb{N}^\infty} (Y_J)^{\partial^K}(\mathbb{B})\, Y_K(\mathbb{A}, \mathbb{B})\ . \qquad (??)$$

But the $Y_J^{\partial^K}$ are either zero, or Schubert polynomials, and they all vanish under the specialization $\mathbb{A} = \mathbb{B}$, except for $Y_{[0,0,\ldots]} = 1$. Now, $Y_J^{\partial^K}$ can be $Y_{[0,0,\ldots]}$ only when $k_1 + k_2 + \cdots = j_1 + j_2 + \cdots$, for degree reason. By recursion on the righmost non-zero component of J, say j_ℓ, one sees that

$k_i = 0$, $i > \ell$ and $k_\ell = j_\ell$. Finally $K = J$ and equation (??) becomes the irrefutable identity

$$Y_J(\mathbb{A}, \mathbb{B}) = Y_{[0,0,\ldots]}(\mathbb{B}, \mathbb{B})\, Y_J(\mathbb{A}, \mathbb{B}) \ .$$

This proves $(\star\star\star)$ in full generality.

The original Newton formula has not been lost. When f depends only on a_1, then f^{∂^K} is 0 if K is not of the type $K = [n, 0, 0, \ldots]$ for some integer n. In that case, Y_K is dominant, corresponding to a row diagram with n boxes, and indeed the Newton polynomial is

$$Y_{[n,0,0,\ldots]} = (a_1 - b_1)\,(a_1 - b_2) \cdots (a_1 - b_n) \ .$$

References

[LS1] A. Lascoux, M.P. Schützenberger, *Interpolation de Newton à plusieurs variables*, Séminaire d'Algèbre M.P.Malliavin 1983-84, Springer L.N. **1146** (1985) 161-175.

[LS2] A. Lascoux, M.P. Schützenberger, *Symmetrization operators on polynomial rings*, Funct.Anal. **21**(1987) 77-78.

[M] I.G. Macdonald, *Notes on Schubert polynomials*, Publications du LACIM, Montréal, (1991).

[N] I. Newton, *Philosophiæ Naturalis Principia Mathematica*, Liber III, Lemma V, London (1687).

C.N.R.S., Institut Gaspard Monge
Université de Marne-la-Vallée,
5 Bd Descartes, Champs sur Marne,
77454 Marne La Vallée Cedex 2 FRANCE
`Alain.Lascoux@univ-mlv.fr`
`http://phalanstere.univ-mlv.fr/~al`