CDMTCS
Research
Report
Series

Generalized Number Derivatives

Michael Stay
Department of Computer Science
University of Auckland
New Zealand

CDMTCS-249
August 2004

Centre for Discrete Mathematics and
Theoretical Computer Science

GENERALIZED NUMBER DERIVATIVES

MICHAEL STAY

Abstract. We generalize the concept of a number derivative, and examine one particular instance of a deformed number derivative for finite field elements. We find that the derivative is linear when the deformation is a Frobenius map and go on to examine some of its basic properties.

1. Introduction

The so-called "number derivative" seems to have been invented independently at least three times [3, 1, 4]. Here we present a generalization of the number derivative that applies to nearly anything one might reasonably call a number. Afterwards, we examine the case of a specific number derivative on finite fields and some of its basic properties.

We generalize the concept of a number derivative to the following algorithm; in order to illustrate each step, we will present the corresponding step from the standard number derivative, denoted N, and our number derivative, denoted S.

(1) Choose a parameterized canonical form. In the case of N, this consists of representing each integer as a product of prime powers; the parameters are the primes. In the case of S, we choose a generator θ of the finite field $\mathrm{GF}(p^k)$ and express each finite field element as θ^n. Here, the parameter is θ.
(2) Convert this canonical form into a function. The algorithm N takes each prime power $p_i^{k_i}$ to a function $x_i^{k_i}(y_i) = y_i^{k_i}$. The algorithm S replaces θ^n with the function $x^n(y) = y^n$.
(3) Differentiate the function with respect to the parameters. The algorithm N computes $D(f) = (\sum_i \frac{\partial}{\partial x_i})(f)$. The algorithm S computes the s-derivative $D_s(f)$.
(4) Evaluate the derivative at some function of the parameters, typically the identity function. The algorithm N computes $D(f)|_{y_i=p_i}$. The algorithm S computes $D_s(f)|_{y=\theta}$.

The notation we use below requires some care. Multiplication is denoted by a dot $[\cdot]$ or by concatenation of symbols: $x \cdot x^2 = xx^2 = x^3$. x^n is a *function*:

$$x^n(y) = y^n,$$

so $x(y) = y$ is the identity function and $x^0(y) = 1$ for all y. Parentheses, when preceded by a function or operator, denote composition or application, respectively:

$$x^2(x^n) = (x^n)^2 = x^{2n}.$$

MSC2000: 05A30

Application is left associative:

$$f(g)(h) = (f(g))(h),$$

and takes precedence over multiplication:

$$gh(f) \neq g(f) \cdot h(f).$$

2. Exponential quantum calculus

We begin with the operator $M_s(f) = f(x^s)$. The s-differential is then $d_s = M_s - x$ and the s-derivative is

$$D_s(f) = \frac{d_s(f)}{d_s(x)}.$$

The s-derivative of an element $x^n(\theta)$ is

$$D_s(x^n)(\theta) = \frac{M_s(x^n) - x^n}{M_s(x) - x}(\theta) = \frac{x^{ns} - x^n}{x^s - x}(\theta) = ([n]x^{n-1})(\theta),$$

where

$$[n] \equiv \frac{x^{(s-1)n} - x^0}{x^{s-1} - x^0}.$$

The s-deformation has many similarities to the q-deformation that results in the quantum calculus [2]. To get the s-deformation from the q-deformation, one replaces the constant q by the function x^{s-1}. Since this is the same transformation we chose to use in the second step of the algorithm S, both derivatives give rise to the same number derivative.

Since the notation is somewhat simpler for the q-derivative, we will adopt it through most of the paper. The operator $M_q = M_s$:

$$M_s(f) = f(x^s) = f(x^{s-1}x) = f(qx) = M_q(f).$$

The q-differential is $d_q = M_q - x$ and the q-derivative is

$$D_q(f) = \frac{d_q(f)}{d_q(x)}.$$

The q-derivative of an element $x^n(\theta)$ is

$$D_q(x^n)(\theta) = \frac{M_q(x^n) - x^n}{M_q(x) - x}(\theta) = \frac{q^n x^n - x^n}{qx - x}(\theta) = ([n]x^{n-1})(\theta),$$

where

$$[n] \equiv \frac{q^n - q^0}{q^1 - q^0}.$$

Also, in the portions of the paper directly concerning the algorithm S, we will usually omit the final application of the functions to θ.

3. Identities

For what functions $q = x^{s-1}$, if any, is this number derivative linear? Let $\theta^a + \theta^b = \theta^c$. Then

$$D_q(x^c)(\theta) = \frac{x^{sc} - x^c}{x^s - x}(\theta) = \frac{(\theta^a + \theta^b)^s - \theta^c}{\theta^s - \theta}.$$

On the other hand,

$$(D_q(x^a) + D_q(x^b))(\theta) = \frac{x^{as} + x^{bs} - (x^a + x^b)}{x^s - x}(\theta) = \frac{\theta^{as} + \theta^{bs} - \theta^c}{\theta^s - \theta},$$

so we want the cross terms in the binomial $(\theta^a + \theta^b)^s$ to be zero modulo p. This only occurs when s is a power of p, so the derivative is linear iff M_s is a Frobenius map. In the rest of the paper, we will only consider $q = x^{s-1}$ of this form.

The derivation of the product rule is the same as that for the q-derivative:

$$\begin{aligned}
D_q(fg) &= \frac{M_q(fg) - fg}{M_q x - x} \\
&= \frac{M_q(f)M_q(g) - M_q(g)f + M_q(g)f - fg}{M_q(x) - x} \\
&= M_q(g)\frac{M_q(f) - f}{M_q(x) - x} + f\frac{M_q(g) - g}{M_q(x) - x} \\
&= M_q(g)D_q(f) + fD_q(g) \qquad (1) \\
&= gD_q(f) + M_q(f)D_q(g), \qquad (2)
\end{aligned}$$

where (2) follows by symmetry.

The same is true for the quotient rule. Since by (1),

$$\begin{aligned}
D_q(f) &= D_q(g\frac{f}{g}) \\
&= M_q(g)D_q(\frac{f}{g}) + \frac{f}{g}D_q(g),
\end{aligned}$$

we have

$$\begin{aligned}
D_q(\frac{f}{g}) &= \frac{gD_q(f) - fD_q(g)}{gM_q(g)} \qquad (3) \\
&= \frac{M_q(g)D_q(f) - M_q(f)D_q(g)}{gM_q(g)} \qquad (4)
\end{aligned}$$

where (4) follows from (2) instead.

Note that while there is not a general chain rule for the standard q-derivative, we can use the fact that every element is of the form $x^n(\theta)$ to find one for this derivative:

$$\begin{aligned}
D_q(g(x^n)) &= \frac{M_q(g(x^n)) - g(x^n)}{M_q(x) - x} \cdot \frac{M_q(x^n) - x^n}{M_q(x^n) - x^n} \\
&= \frac{M_q(g(x^n)) - g(x^n)}{M_q(x^n) - x^n} \cdot \frac{M_q(x^n) - x^n}{M_q(x) - x} \\
&= \frac{M_q(g(x^n)) - g(x^n)}{M_q(x(x^n)) - x(x^n)} \cdot D_q(x^n) \\
&= D_q(g)(x^n) \cdot D_q(x^n)
\end{aligned}$$

While the product and quotient rules (1)-(4) are the same as those typically given [2], this rule differs: since q is the function x^{p^j-1} instead of a constant, we evaluate it at x^n rather than take the q^n-derivative of g in the first term.

Finally, the q-numbers $[n]$ satisfy

$$[n+1] = q^0 + q[n] \text{ and } [n+1] - [n] = q^n.$$

4. Constants

Under what conditions does

$$d_q(x^n) = 0? \tag{5}$$

We have

$$M_q(x^n) - x^n = 0$$

which implies

$$q^n x^n = x^n$$

and

$$q^n = x^{n(p^j-1)} = x^0 = 1$$

if $n \neq -\infty$. Therefore, $D_q x^n = 0$ if $(p^k - 1)|n(p^j - 1)$. We call elements satisfying (5) "constants."

Constants behave as one might expect. Adding a constant obviously does not change the derivative; multiplying by a constant m scales the derivative by the same amount:

$$\begin{aligned} D_q(mf) &= fD_q(m) + M_q(m)D_q(f) \\ &= f \cdot 0 + mD_q(f) \\ &= mD_q(f) \end{aligned}$$

5. The exponential function exp

Consider the equation $D_q x^e = x^e$. Then

$$D_q x^e = [e]x^{e-1} = x^e$$

$$[e]x^{-1} = x^0$$

$$\frac{x^{es} - x^0}{x^s - x} = x^0$$

$$x^{es} = x^s - x + 1 \tag{6}$$

so if $\theta^s - \theta + 1$ is generated by θ^s then the equation will hold for at least one e. We may then define the function $\exp = x^e$; there is no reason to prefer one solution over another.

We use exp to illustrate a subtlety of the chain rule. One might conclude that $D_q(\exp^m) = [m]x^{me+m-1}$:

$$\begin{aligned} D_q x^{me} &= D_q(x^e(x^m)) \\ &= D_q(x^e)(x^m) \cdot D_q(x^m) \\ &= x^e(x^m) \cdot [m]x^{m-1} \\ &= x^m e \cdot [m]x^{m-1} \\ &= [m]x^{me+m-1} \end{aligned} \tag{7}$$

but (7) does not follow. It is only when applied directly to θ that $D_q x^e = x^e$. Here, $D_q x^e$ is applied to the function x^m and then to θ.

The true equation may be found by examining the derivatives of the first few powers of exp:

$$\begin{aligned} D_q(\exp^2) &= D_q(x^{2e}) \\ &= D_q(x^e \cdot x^e) \\ &= x^e D_q(x^e) + M_q(x^e) D_q(x^e) \\ &= x^{2e} + q^e x^{2e} \\ &= (q^0(x^e) + q^1(x^e)) \cdot (x^e)^2 \\ &= ([2]x^2)(x^e) \end{aligned}$$

$$\begin{aligned} D_q(\exp^3) &= D_q(x^{3e}) \\ &= D_q(x^e \cdot x^{2e}) \\ &= x^{2e} D_q(x^e) + M_q(x^e) D_q(x^{2e}) \\ &= x^{3e} + q^e \cdot (q^0 + q^e) x^{3e} \\ &= (q^0(x^e) + q^1(x^e) + q^2(x^e))(x^e)^3 \\ &= ([3]x^3)(x^e) \end{aligned}$$

The pattern is immedately clear: $D_q(\exp^m) = ([m]x^m)(\exp)$, as one would hope.

We can now prove the result by induction. Assume that $D_q(\exp^{(m-1)})$ is of the form $([m-1]x^{m-1})(\exp)$. Then

$$\begin{aligned} D_q(\exp^m) &= D_q(x^{me}) \\ &= D_q(x^e x^{(m-1)e}) \\ &= x^{(m-1)e} D_q(x^e) + M_q(x^e) D_q(x^{(m-1)e}) \\ &= x^{me} + q^e x^e \cdot ([m-1]x^{m-1})(x^e) \\ &= ((q^0 + q[m-1])x^m)(x^e) \\ &= ([m]x^m)(x^e) \\ &= ([m]x^m)(\exp). \end{aligned}$$

6. Commutation

As with the standard q-derivative, $[D_q, x\cdot] = M_q$:

$$\begin{aligned} [D_q, x\cdot](f) &= D_q(xf) - xD_q(f) \\ &= fD_q(x) + M_q(x)D_q(f) - xD_q(f) \\ &= f + qxD_q(f) - xD_q(f) \\ &= f + d_q(x)D_q(f) \\ &= (x + d_q(x)D_q)(f) \\ &= (x + d_q(x)\frac{d_q}{d_q(x)})(f) \\ &= (x + d_q)(f) \\ &= M_q(f) \end{aligned}$$

If we define the q-commutator $[f, g]_q \equiv fg - M_q(g)f$, then we find that

$$\begin{aligned}[D_q, x\cdot]_q(f) &= D_q(xf) - M_q(x)D_q(f) \\ &= fD_q(x) + M_q(x)D_q(f) - M_q(x)D_q(f) \\ &= f.\end{aligned}$$

We can define a Hamiltonian operator via the anticommutator $H = \{D_q, x\cdot\}$ to get

$$\begin{aligned}Hf &= D_q(xf) + xD_q(f) \\ &= fD_q(x) + qxD_q(f) + xD_q(f) \\ &= f + [2]xD_q(f),\end{aligned}$$

so the "energy" of a finite field element $x^n(\theta)$ is

$$\begin{aligned}Hx^n &= x^n + [2]xD_q(x^n) \\ &= (1 + [2][n])x^n \\ &= (1 + (1 + q)[n])x^n \\ &= (1 + [n] + q[n])x^n \\ &= ([n + 1] + [n])x^n\end{aligned}$$

7. q-Antiderivative

The q-derivative of a finite field element is an element itself. If we add the constant 1, the derivative does not change, so at most half of the elements have antiderivatives. If an element has an antiderivative, then it is unique up to an additive constant: suppose f has two antiderivatives F_1 and F_2. Then let $\phi = F_1 - F_2$. Now $D_q(\phi) = 0$; but any function for which that holds true is a constant by definition.

The integral operator $\int_q(d_q\cdot)$ is the Moore-Penrose inverse of D_q. Thus the equation $D_q(F) = f$ has a solution iff $f = D_q(\int_q(fd_q)))$.

8. Higher derivatives

Because $[n]$ is a function of x, there are correction terms on the higher derivatives. For instance,

$$\begin{aligned}D_q^2(x^n) &= D_q(D_q(x^n)) \\ &= D_q([n]x^{n-1}) \\ &= [n]D_q(x^{n-1}) + M_q(x^{n-1})D_q([n]) \\ &= [n][n-1]x^{n-2} + (qx)^{n-1}D_q([n])\end{aligned}$$

It is these extra terms that give rise to trigonometric-like functions. We've already seen exp; there are others like sinh and cosh with larger periods.

There will be a subspace, however, for which iterated derivatives eventually yield zero. This subspace always includes the vectors $\{x, 1\}$, and may include more.

We can define an inner product in this subspace. Let $J_q = \int_q(d_q\cdot)$ and without loss of generality, let $n \geq m$. Then

$$< J_q^n, J_q^m >=< 1, D_q^n J_q^m >= \delta_{n,m}.$$

The function exp is an eigenvector of D_q, so it is orthogonal to the subspace:

$$< J_q^n, \exp >=< 1, D_q^n \exp >=< 1, \exp >= \exp .$$

Other trigonometric functions are defined by the period with which they repeat. sinh, for example, is an eigenvector of D_q^2, and a similar identity holds.

9. Examples

We consider the field $\mathrm{GF}(2^4)$ with the field polynomial x^4-x-1. There are three possible values q may take: x^1, x^3, and x^7. Each gives rise to different structures.

n	θ^n	$D_x(\theta^n)$	$D_{x^3}(\theta^n)$	$D_{x^7}(\theta^n)$
$-\infty$	0000	0000	0000	0000
0	0001	0000	0000	0000
1	0010	0001	0001	0001
4	0011	0001	0001	0001
2	0100	0110	0001	0111
9	0101	0110	0001	0111
5	0110	0111	0000	0110
11	0111	0111	0000	0110
3	1000	1111	0111	0101
15	1001	1111	0111	0101
10	1010	1110	0110	0100
7	1011	1110	0110	0100
6	1100	1001	0110	0010
13	1101	1001	0110	0010
12	1110	1000	0111	0011
14	1111	1000	0111	0011

9.1. **$q = x$.** We have constants 0, 1. "Trig" functions include $\theta^{11} = 0111 = \exp$, $\theta^{14} = 1111 = \sinh$, and $\theta^3 = 1000 = \cosh$. The names we've chosen are fairly arbitrary; they are only meant to reflect the period with which the derivative returns to itself. θ has no antiderivative, so we have an inner product acting on the subspace $\{1, x\}$ of the space $\{1, x, \exp, \sinh\}$.

9.2. **$q = x^3$.** Nonzero constants are $\theta^0 = 1, \theta^5 = 0110$, and $\theta^{10} = 0111$, the cube roots of 1. There are no trig functions. A basis for the space is $\{1, x\}$.

9.3. **$q = x^7$.** We have the constants 0, 1 and the trig function exp. In this case, $J_q\theta = \theta^6$, so we have the inner product on a three-dimensional subspace $\{1, x, x^6\}$, while the complete basis is $\{1, x, x^6, \exp\}$.

References

[1] Cohen, G.L, and D. E. Iannucci, Derived sequences. *Journal of Integer Sequences* **6**. 2003.
[2] Kac, Victor and Pokman Cheung, *Quantum Calculus*. Springer-Verlag. 2002.
[3] Kurokawa, Nobushige, Hiroyuki Ochiai and Masato Wakayama, Absolute derivations and zeta functions. *Documenta Mathematica* **Extra Volume Kato**. 2003. pp 565-584.
[4] Ufnarovski, Victor and Bo Åhlander, How to differentiate a number. *Journal of Integer Sequences* **6**. 2003.