

A Tutorial on Closed Difference Forms

Burkhard Zimmermann

RISC, Linz (Austria)

January 15, 2001

Summary by Frédéric Chyzak

Abstract

Zeilberger’s theory of closed difference forms provides with a deeper understanding of the creative telescoping method used to prove many (q -)hypergeometric (multi-)sum identities, and of “companion” or “dual” identities. By introducing new types of summation domains, the closed form approach allows to discover new identities of the form “sum equals sum,” including new summatory representations of $\zeta(3)$. A transform similar to a pullback (change of variables) of differential forms is introduced, and permits to find more new identities. This summary is freely inspired by [1, 2, 4, 5] and the talk.

1. Comparison Between Differential and Difference Calculi

By mimicking differential calculus [2], Zeilberger has developed a complete *difference calculus* [4]. This theory, which we recall here, culminates with a discrete analogue to Stokes’s theorem.

Given a \mathbb{C} -vector space V , which will take the role of a tangent space momentarily, an *alternate multilinear p -form* on V is just a multilinear map $\phi : V^p \rightarrow \mathbb{C}$ that satisfies the rule

$$\phi(v_1, \dots, v_{i+1}, v_i, \dots, v_p) = -\phi(v_1, \dots, v_p).$$

This represents a p -volume measure, in the sense that it assigns an (oriented) volume to the parallelepipedic polyhedron determined by the vectors v_i . By a natural convention, 0-forms are just constants. To a p -form ϕ and a q -form ψ , one associates a $(p+q)$ -form, i.e., a $(p+q)$ -volume measure, by means of the *exterior product* $\phi \wedge \psi$:

$$(\phi \wedge \psi)(v_1, \dots, v_{p+q}) = \sum_{\sigma \in S_{p,q}} \epsilon(\sigma) \phi(v_{\sigma(1)}, \dots, v_{\sigma(p)}) \psi(v_{\sigma(p+1)}, \dots, v_{\sigma(p+q)})$$

where $S_{p,q}$ denotes the set of permutations of $\{1, \dots, p+q\}$ with $\sigma(1) < \dots < \sigma(p)$ and $\sigma(p+1) < \dots < \sigma(p+q)$, and where $\epsilon(\sigma)$ denotes the signature of the permutation σ . Consider the direct sum $\mathcal{A}(V) = \bigoplus_{p \geq 0} \mathcal{A}_p(V)$ of the vector spaces $\mathcal{A}_p(V)$ of alternate p -forms. By extending the exterior product by linearity, we obtain an associative multiplication on $\mathcal{A}(V)$, which becomes a graded algebra with the product rule $\psi \wedge \phi = (-1)^{pq} \phi \wedge \psi$ for a p -form ϕ and a q -form ψ .

Next, an *alternate difference p -form*, or for short a *difference p -form*, is a map ω which to each element ξ of a real manifold M associates a multilinear p -form $\omega(\xi)$ on the tangent space $V = T_\xi M$. Exterior products of difference forms are defined pointwise. At this point, difference forms and differential forms share the same definition. In the following however, we focus to the case when M is a submanifold of \mathbb{R}^d : each $\omega(\xi)$ is then an alternate form on $V = \mathbb{R}^d$. By imposing the additional property $\omega(\xi_1, \dots, \xi_d) = \omega([\xi_1], \dots, [\xi_d])$, we obtain forms that are piecewise constant, as well as their coefficients. (Compare this situation with the theory in the differential setting,

where one insists in having C^∞ forms and C^∞ coefficients.) The possible variations of forms with ξ is at the origin of the notions of *exterior differential* and *exterior difference* introduced below.

In the differential setting, a kind of a derivation is defined on differential forms in the following way. One starts with the usual derivative ω' , which satisfies the asymptotic relation $\omega(\xi + v) = \omega(\xi) + \omega'(\xi)(v) + o(v)$ as $v \rightarrow 0$. Each $\omega'(\xi)$ is a linear map from $V = \mathbb{R}^d$ to the vector space $\mathcal{A}_p(V)$, and can be viewed as a multilinear map from V^{p+1} to \mathbb{C} that is not alternate, but alternate in its last p variables only. Making it alternate by an averaging technique, we obtain the *exterior differential* $d\omega$ given by

$$(d\omega)(\xi)(v_0, \dots, v_p) = \sum_{i=0}^p (-1)^i (\omega'(\xi)(v_i))(v_0, \dots, \hat{v}_i, \dots, v_p).$$

In the difference case, we start with another linearization instead of the derivative ω' to define the exterior difference of ω , namely by secants instead of tangents. Let $\omega^\Delta(\xi)$ be the linear map on V defined by $\omega(\xi + v) = \omega(\xi) + \omega^\Delta(\xi)(v) + R(v)$ and $R(v)$ is zero for each element $v = e_i$ of the canonical basis of $V = \mathbb{R}^d$. Again, $(v_0, \dots, v_p) \mapsto \omega^\Delta(\xi)(v_0)(v_1, \dots, v_p)$ is alternate in its last p variables only, but the full alternate nature is recovered by the *exterior difference* $d\omega$ defined by

$$(d\omega)(\xi)(v_0, \dots, v_p) = \sum_{i=0}^p (-1)^i (\omega^\Delta(\xi)(v_i))(v_0, \dots, \hat{v}_i, \dots, v_p).$$

As opposed to the classical exterior differential, exterior difference heavily depends on the choice of a basis on V ; but like it, it satisfies $d \circ d = 0$.

Denote (n_1, \dots, n_d) the dual basis of the canonical basis of the manifold \mathbb{R}^d that contains M . As in the differential setting, the exterior difference dn_i of the restriction of n_i to M (i.e., or the i th coordinate function on M) plays a special role: the dn_i form a basis for the ring of difference form, and the $dn_{i_1} \wedge \dots \wedge dn_{i_p}$ for $i_1 < \dots < i_p$ span the vector space (respectively, free module) of p -forms. Exterior differential and exterior difference share a formally simple, easy-to-memorize formulation on the canonical basis (dn_1, \dots, dn_d) : for $\omega = f dn_{i_1} \wedge \dots \wedge dn_{i_r}$, we get

$$d\omega = df \wedge dn_{i_1} \wedge \dots \wedge dn_{i_r}$$

where the exterior differential is $df = \sum_{i=1}^d \frac{\partial f}{\partial \xi_i} dn_i$, and the exterior difference $df = \sum_{i=1}^d (\Delta_i f) dn_i$, where Δ_i is the finite difference operator defined by $(\Delta_i f)(\xi_1, \dots, \xi_d) = f(\xi_1, \dots, \xi_i + 1, \dots, \xi_d) - f(\xi_1, \dots, \xi_d)$.

In order to make the link between difference forms and summation, we restrict to hypercubic manifolds given by setting some of the coordinates ξ_i to 0 and letting all others vary freely in $[0, 1)$, and to the manifolds obtained after translating the latter by vectors with integer entries. Note that all those elementary manifolds (in various dimensions) have volume 1, and that we have restricted difference forms to be constant on such sets. As a consequence, the integral of a form $f dn_1 \wedge \dots \wedge dn_d$ on $[0, 1)^d$ is just $f(0, \dots, 0)$, as is for $i_1 < \dots < i_r$ the integral of $f dn_{i_1} \wedge \dots \wedge dn_{i_r}$ on the hypercube defined by $0 \leq \xi_j < 1$ for each $j = i_k$ and $\xi_j = 0$ for all other j . By integration over a union of elementary manifolds, we are naturally led to integral representing sums; for example:

$$\int_{\mathbb{R}^d} f dn_1 \wedge \dots \wedge dn_d = \sum_{(n_1, \dots, n_d) \in \mathbb{Z}^d} f(n_1, \dots, n_d).$$

We are now ready to derive a difference variant of Stokes's theorem: consider the oriented hypercube $\Omega = [0, 1)^d$ and its boundary $\partial\Omega$ defined as usual as a formal linear combination of $2d$ faces,

$$\partial\Omega = F(\xi_1 = 0) - F(\xi_2 = 0) + \dots + (-1)^{d+1} F(\xi_d = 0) - F(\xi_1 = 1) + F(\xi_2 = 1) + \dots + (-1)^d F(\xi_d = 1),$$

where $F(\xi_i = a)$ is the (oriented) face $\Omega \cap \{\xi \mid \xi_i = a\}$. Boundaries of other elementary manifolds are obtained by translating $\partial\Omega$, keeping the same coefficients. In this way, we can define the integral of a form over a linear combination of manifolds to be the very same linear combination of integrals of the same form over the manifolds. For

$$(1) \quad \omega = \sum_{i=1}^d f_i dn_1 \wedge \cdots \wedge \hat{dn}_i \wedge \cdots \wedge dn_d$$

we get

$$\begin{aligned} \int_{\partial\Omega} \omega &= \sum_{i=1}^d (-1)^i \int_{F(\xi_i=1)-F(\xi_i=0)} f_i dn_1 \wedge \cdots \wedge \hat{dn}_i \wedge \cdots \wedge dn_d \\ &= \left(\sum_{i=1}^d (-1)^i f_i(0, \dots, 1, \dots, 0) - \sum_{i=1}^d (-1)^i f_i(0, \dots, 0) \right) dn_1 \wedge \cdots \wedge dn_d \\ &= \sum_{i=1}^d (-1)^i (\Delta_i f_i)(0, \dots, 0) dn_1 \wedge \cdots \wedge dn_d = \int_{\Omega} d\omega. \end{aligned}$$

We could have as well considered forms ω defined on the integer lattice \mathbb{Z}^d , and defined their sums $\sum_{\Omega} \omega$ on a manifold Ω by the integrals $\int_{\Omega} \omega$ of the form ω extended to \mathbb{R}^d by $\omega(\xi_1, \dots, \xi_d) = \omega([\xi_1], \dots, [\xi_d])$. We shall adopt this equivalent viewpoint from the next section on. By linearity with respect to manifolds, we obtain the following discrete variant of Stokes's formula [4].

Theorem 1 (Zeilberger–Stokes formula). *For any difference p -form ω such that $\omega(\xi_1, \dots, \xi_d) = \omega([\xi_1], \dots, [\xi_d])$ on any manifold Ω that is a linear combination of elementary hypercubic manifolds, we have $\sum_{\partial\Omega} \omega = \sum_{\Omega} d\omega$.*

2. Closed Form Identities (Pun Intended!)

An interesting situation is that of a *closed (difference) form*, which by definition is a difference form ω such that $d\omega = 0$. In this case, the sum $\sum_{\partial\Omega} \omega = 0$ for any manifold Ω on all of which Ω is defined, owing to Theorem 1 above. If more specifically ω is given by (1), we obtain

$$\sum_{i=1}^d \sum_{\partial\Omega} f_i dn_1 \wedge \cdots \wedge \hat{dn}_i \wedge \cdots \wedge dn_d = 0,$$

in other words a relation between a priori infinite sums! Using the leeway available in the choice of Ω yields several kinds of identities: sum equals constant, sum equals sum, etc. In the following, we detail this situation in the special case $r = 2$. Let us denote dn and dk for dn_1 and dn_2 , respectively, and consider a closed 1-form $\omega = g dn + f dk$, so that $\Delta_n f = \Delta_k g$.

2.1. Stripe-shaped manifolds. Consider $\Omega = \mathbb{R}^+ \times [0, n] = \{(x, y) \mid x \geq 0 \text{ and } 0 \leq y \leq n\}$ and the closed form ω obtained for

$$f(n, k) = \binom{m}{k} \binom{n}{k} \binom{p+n+m-k}{n+m} \quad \text{and} \quad g(n, k) = \frac{mk - p(n+1)}{(n+m+1)(n+1-k)} f(n, k).$$

Stokes's theorem on Ω then yields (after elementary manipulations of binomial sums)

$$\sum_{k=0}^n \binom{m}{k} \binom{n}{k} \binom{p+n+m-k}{n+m} = \sum_{k \in \mathbb{N}} f(n, k) = \sum_{k \in \mathbb{N}} f(0, k) + \sum_{l=0}^n g(l, 0) = \binom{m+p}{m} \binom{n+p}{n}.$$

More generally, many *closed-form* identities like the one above, where “closed form” now means that both the summand and the sum are hypergeometric sequences, correspond to a “closed form” that involves the summand as one of its coefficients. Hence Zeilberger’s “pun intended.”

But some magic takes place here: changing Ω to $[0, k] \times \mathbb{R}^+$ and summing with respect to n instead of k , the same method sometimes yields a *companion identity*. Moreover, the more variables there are, the more amplified this phenomenon is: for r variables and in lucky cases where all summations make sense, a single closed difference $(r - 1)$ -form with hypergeometric coefficients can be viewed as a simultaneous encoding of r closed form summation identities [4].

2.2. Triangular-shaped manifolds. Zeilberger observed that for a closed form $\omega_1 = g_1 dn + f_1 dk$, the functions $f_s(n, k) = f_1(sn, k)$ and $g_s(n, k) = g_1(sn, k) + g_1(sn + 1, k) + \cdots + g_1(sn + s - 1, k)$ provide for each $s > 1$ with another closed form $\omega_s = g_s dn + f_s dk$. Basing on this, Amdeberhan and Zeilberger [1] derived the following representations for $\zeta(3)$:

$$\begin{aligned} \zeta(3) &= \frac{5}{2} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{\binom{2n}{n} n^2} = \frac{1}{4} \sum_{n=1}^{\infty} \frac{(-1)^{n-1} (56n^2 - 32n + 5)}{(2n-1)^2 \binom{3n}{n} \binom{2n}{n} n^3} \\ &= \frac{1}{72} \sum_{n=0}^{\infty} \frac{(-1)^n (5265n^4 + 13878n^3 + 13761n^2 + 6120n + 1040)}{(4n+3)(4n+1)(3n+2)^2(3n+1)^2(n+1) \binom{4n}{n} \binom{3n}{n}}. \end{aligned}$$

Specifically, they considered $\Omega = \{(x, y) \mid y \geq \lfloor x + 1 \rfloor\}$ and the functions

$$f_1(n, k) = (-1)^k \frac{k!^2(n-k-1)!}{(n+k+1)!(k+1)} \quad \text{and} \quad g_1(n, k) = 2(-1)^k \frac{k!^2(n-k)!}{(n+k+1)!(n+1)^2}.$$

The representations above have respectively been obtained for $s = 1, 2$, and 3 ; their general terms decrease like $O(n^{-3/2}4^{-n})$, $O(n^{-2}27^{-n})$, $O(n^{-2}64^{-n})$, respectively—at the cost of more and more operations for each term, though! Changing Ω to $\Omega_s = \{(x, y) \mid y \geq s\lfloor x + 1 \rfloor\}$ leads to other representations [1], like, for $s = 2$,

$$\zeta(3) = \sum_{n=0}^{\infty} \frac{(-1)^n P(n)}{80(5n+4)(5n+3)(5n+2)(5n+1)(4n+3)^2(4n+1)^2(2n+1)^2(n+1) \binom{5n}{n} \binom{4n}{n}}$$

where $P = 1613824n^8 + 7638016n^7 + 15700096n^6 + 18317312n^5 + 13278552n^4 + 6131676n^3 + 1763967n^2 + 289515n + 20782$. The general term is now $O(n^{-2}(27/3125)^{-n})$, with $27/3125 \approx 115.74$.

To sketch the proof, we apply Stokes’s theorem to ω_s on Ω_s , and obtain:

$$\sum_{n=0}^{\infty} g_s(n, 0) + \sum_{k=0}^{\infty} f_s(sk + s, k) + \sum_{k=0}^{\infty} (g_s(sk, k) + \cdots + g_s(sk + s - 1, k)) = 0.$$

Next, noting that $g_1(n, 0) = 2/(n+1)^3$ and grouping the sums over k yields the announced identity.

2.3. Finite triangular-shaped and rectangular-shaped manifolds. Other identities like

$$\frac{\Gamma(x+n)\Gamma(y+n)}{\Gamma(n)\Gamma(x+y+n)} {}_3F_2\left(\begin{matrix} x, y, v+n-1 \\ v, x+y+n \end{matrix} \middle| 1\right) = \frac{\Gamma(x+k)\Gamma(y+k)}{\Gamma(k)\Gamma(x+y+k)} {}_3F_2\left(\begin{matrix} x, y, v+k-1 \\ v, x+y+k \end{matrix} \middle| 1\right)$$

and $\sum_{n+m=s} \binom{2n}{n} \binom{2m}{m} = 4^s$ are based on other choices for Ω , like a rectangle $[0, k] \times [0, n]$ or a “triangle” $\{(x, y) \mid \lfloor x \rfloor + \lfloor y \rfloor \leq s\}$ for Ω [5].

3. Closed Forms with Holonomic Coefficients

Consider a closed form $\omega = g \, dn + f \, dk$ with hypergeometric coefficients. Since f is hypergeometric in n , one can find some rational function R of (n, k) such that $\Delta_k g = \Delta_n f = Rf$. It is also well-known that if a hypergeometric sequences h has a hypergeometric anti-difference H , there has to be some rational function S such that $H = Sh$. Here we get $g = S\Delta_n f = SRf$. This situation extends to more variables, which legitimates Zeilberger's focus to closed forms whose coefficients are all multiples of the same hypergeometric sequence f by polynomials in the variables; he called such forms *WZ forms* [4]. Here we extend this situation to forms whose coefficients are rational multiples of the same holonomic sequence, and make the link between closed forms and creative telescoping explicit.

Let a summation identity $\sum_{k=a}^b f_{n,k} = F_n$ be given, where both f and F are holonomic ∂ -finite sequences. In view of verifying it, knowing F allows to compute a non-zero operator $P_0(n, S_n)$ such that $P_0 \cdot F = 0$. Proving the identity thus reduces to proving $\sum_{k=a}^b (P_0 \cdot f)(n, k) = 0$. By restricting to holonomic hypergeometric summands and right-hand sides, Zeilberger's presentation essentially only dealt with the case $P_0 = S_n - 1$: F can always be assumed to be 1, otherwise we replace $f(n, k)$ with $f(n, k)/F(n)$. In this spirit, we now require that P_0 be a right multiple of $S_n - 1$ and write $P_0 = (S_n - 1)R$ this factorization.

The holonomy of f ensures that there exists a pair (P, Q) with non-zero P such that

$$(2) \quad (P + (S_k - 1)Q) \cdot f = 0.$$

Provided that there exists such a pair for $P = P_0$, the operator Q can be computed by Chyzak's ∂ -finite extension of Gosper's algorithm [3]. Let \mathcal{A} be the algebra of difference operators with respect to n and k with coefficients that are rational functions in n and k , and introduce the module $\mathfrak{M} = \mathcal{A} \cdot f$. The form

$$(3) \quad \omega = (R \cdot f) \, dk - (Q \cdot f) \, dn,$$

whose coefficients all lie in \mathfrak{M} is closed:

$$d\omega = ((S_n - 1)R \cdot f) \, dn \wedge dk - ((S_k - 1)Q \cdot f) \, dk \wedge dn = \left((P + (S_k - 1)Q) \cdot f \right) \, dn \wedge dk = 0.$$

Conversely, assume that there exists a closed form ω (with coefficients in \mathfrak{M}) given by (3). By closedness, we have $((S_n - 1)R + (S_k - 1)Q) \cdot f = 0$, whence after summation over k , and provided that R involves neither k nor S_k ,

$$(S_n - 1)R \cdot \sum_{k=a}^b f(n, k) = 0.$$

More generally, if the r -form $f \, dk_1 \wedge \dots \wedge dk_r + \sum_{i=1}^r (P_i \cdot f) \, dn \wedge dk_1 \wedge \dots \wedge \hat{dk}_i \wedge \dots \wedge dk_r$ is closed,

$$\text{i.e., } (S_n - 1) \cdot f + (S_{k_1} - 1)P_1 \cdot f + \dots + (S_{k_r} - 1)P_r \cdot f = 0,$$

the r -fold summation $\sum_{k_1, \dots, k_r} f$ yields a constant with respect to n .

4. Extended WZ Cohomology

Is it easily shown that any 1-form with coefficients defined on \mathbb{Z}^r is exact. Even more is true: any 1-form with holonomic coefficients derives from a holonomic sequence. More specifically, a 1-form ω given by (3) is exact if and only if there exists a function $\phi(n, k)$ such that $\omega = d\phi$, or more explicitly

$$-(Q \cdot f) = (S_n - 1) \cdot \phi \quad \text{and} \quad R \cdot f = (S_k - 1) \cdot \phi.$$

This always holds if we look for unconstrained ϕ : simply define ϕ by

$$\phi(n, k) = \sum_{i=0}^{k-1} (R \cdot f)(0, i) - \sum_{j=0}^{n-1} (Q \cdot f)(j, k).$$

The non-trivial problem is to impose $\phi \in \mathfrak{M}$. (For example, when f is hypergeometric, all coefficients of ω as well as ϕ have to be rational multiples of f .) Then, not all 1-forms ω remain exact. Viewing closed forms modulo exact forms we are led to a cohomology that Zeilberger named *WZ cohomology* in [4] in the case of hypergeometric f , and that we call *extended WZ cohomology* in the more general case of holonomic ∂ -finite f . Following Zeilberger [4], we suggest the following extended research problem: characterize those holonomic ∂ -finite sequences f for which there exists a non-exact closed form with coefficients in $\mathfrak{M} = \mathcal{A} \cdot f$ and compute the corresponding cohomology.

5. Pullbacks

In the differential case, the notion of *pullback* propagates a change of variables in functions to the level of differential forms, thus permitting change of variables in integrals: for a differentiable map ϕ from a manifold N to another manifold M , one gets a mapping ϕ^* that transforms a p -form ω on M to a p -form on N while preserving closedness of forms by simply requiring

$$(4) \quad (\phi^* \omega)(\xi)(v_1, \dots, v_p) = \omega(\phi(\xi))(\phi'(\xi)(v_1), \dots, \phi'(\xi)(v_p)).$$

In the difference case, a simple example of a pullback has already been given in Section 2.2: the closed form ω_s is the pullback of the closed form ω_1 under the map given by $\phi(n, k) = (sn, k)$. However, no simple definition of a pullback seems possible: the obvious guess that mimicks (4), substituting ϕ^Δ for ϕ' , unfortunately does not preserve closedness (taking finite differences is not a local operation). Zimmermann [5] and Gessel independently gave a definition for the case of a linear mapping ϕ that maps integer points to integer points.

The key observation is that for a linear transform $l = \phi(n)$, defined by $l_i = \sum_j a_{i,j} n_j$, shifting by 1 with respect to n_j after performing the substitution induced by ϕ is equivalent to doing shifts with respect to each l_i before substituting, as detailed by the formula $S_{l_j} \phi^* = \phi^* S_{n_1}^{a_{1,j}} \dots S_{n_r}^{a_{r,j}}$. It then follows from a technical but easy calculation that $\Delta_{l_j} \phi^* = \phi^* \sum_i P_{i,j} \Delta_{n_i}$ for some operators $P_{i,j}$. Imposing the natural relations $\phi^*(f) = f \circ \phi$ and $\phi^*(df) = d(\phi^* f)$ for 0-forms f leads to

$$\sum_i \phi^*((\Delta_{n_i} f) dn_i) = \sum_j (\Delta_{l_j} (\phi^* f)) dl_j = \sum_{i,j} \phi^*(P_{i,j} \Delta_{n_i} f) dl_j.$$

Choosing f such that $df = (\Delta_{n_i} f) dn_i$, we get $\phi^*(g dn_i) = \sum_j \phi^*(P_{i,j} g) dl_j$, a definition that proves to preserve closedness.

Bibliography

- [1] Amdeberhan (Tewodros) and Zeilberger (Doron). – Hypergeometric series acceleration via the WZ method. *Electronic Journal of Combinatorics*, vol. 4, n° 2, 1997, p. Research Paper 3. 4 pages. – The Wilf Festschrift (Philadelphia, PA, 1996).
- [2] Cartan (Henri). – *Cours de calcul différentiel*. – Hermann, Paris, 1977, *Collection Méthodes*.
- [3] Chyzak (Frédéric). – An extension of Zeilberger's fast algorithm to general holonomic functions. *Discrete Mathematics*, vol. 217, n° 1-3, 2000, pp. 115–134. – Formal power series and algebraic combinatorics (Vienna, 1997).
- [4] Zeilberger (Doron). – Closed form (pun intended!). In *A tribute to Emil Grosswald: number theory and related analysis*, pp. 579–607. – American Mathematical Society, Providence, RI, 1993.
- [5] Zimmermann (Burkhard). – *Difference forms and hypergeometric summation*. – Master's thesis, Universität Wien, Vienna, Austria, February 2000.