q-deformed binomial coefficients of words

Antoine Renard, Michel Rigo, and Markus A. Whiteland

Université de Liège, Belgique

In mathematics, a q-deformation of a concept (which can be a theorem, a function, an equality, etc.) is a generalisation of the latter involving a new parameter q, and that gives back the original object when letting q tend towards 1. In the more specific case of a counting function f, a q-analogue f_q of this function corresponds to a polynomial in q with non-negative integer coefficients which, when evaluated at $q = 1$, gives the value of the original function f. Gaussian binomial coefficients are a well-known example of such a deformation: given two integers $n \geq k \geq 0$, we define

$$\binom{n}{k}_q = \frac{(1-q^n)(1-q^{n-1})\cdots(1-q^{n-k+1})}{(1-q^k)(1-q^{k-1})\cdots(1-q)}.$$

In particular, $\lim_{q\to 1}\binom{n}{k}_q = \binom{n}{k}$.

On the other side, binomial coefficients of finite words are well-studied objects in combinatorics on words. Recall that, given two words $u, v \in A^*$, where A is an alphabet (*i.e.*, a finite set of letters), the binomial coefficient $\binom{u}{v}$ counts the number of occurrences of v as a scattered subword of u. In other words,

$$\binom{u}{v} = \#\left\{i_{|v|} > \cdots > i_1 \mid u_{i_{|v|}} \cdots u_{i_2} u_{i_1} = v\right\},$$

where $|v|$ denotes the length of v. For a reference on binomial coefficients of words, see [4].

With these two notions in mind, we defined a q-analogue of binomial coefficients of words [8]. We were, among others, inspired by the work of Morier-Genoud and Ovsienko on q-deformed rational and real numbers [6,7].

1 Definitions and first results

Definition 1. *We recursively define the* q-deformation $\binom{\cdot}{\cdot}_q$ *— an element of* $\mathbb{N}[q]$ *— of the binomial coefficients on* $A^* \times A^*$ *as follows. For all words* $u, v \in A^*$ *and letters* $a, b \in A$*:*

$$\binom{u}{\varepsilon}_q = 1, \quad \binom{\varepsilon}{v}_q = 0 \text{ if } v \neq \varepsilon, \quad \text{and} \quad \binom{ua}{vb}_q = \binom{u}{vb}_q \cdot q^{|vb|} + \delta_{a,b}\binom{u}{v}_q,$$

where $\delta_{a,b}$ *is the Kronecker delta on the letters of* A*.*

This generalises the Pascal-like formula satisfied by classical binomial coefficients of words:

$$\binom{ua}{vb} = \binom{u}{vb} + \delta_{a,b}\binom{u}{v}, \quad \forall u, v \in A^*, a, b \in A. \tag{1}$$

We note that the polynomial $\binom{u}{v}_q$ evaluated at 1 gives back the usual binomial coefficient $\binom{u}{v}$; this follows immediately from (1) and the definition above. Also, considering those q-deformed coefficients on a unary alphabet, we recover the Gaussian binomial coefficients.

The following theorem gives a combinatorial interpretation of this q-analogue.

Theorem 1 ([8]). *Let u be a word over A, $k \geq 0$, and $a_1, \ldots, a_k \in A$. Then*

$$\binom{u}{a_1 \cdots a_k}_q = \sum_{\substack{u_0, u_1, \ldots, u_k \in A^* \\ u = u_0 a_1 \cdots u_{k-1} a_k u_k}} q^{\sum_{i=1}^k i|u_i|}.$$

In other words, each occurrence of v as a subword of u contributes to $\binom{u}{v}_q$ with a term q^α where α is the sum, over all letters of v, of the number of letters at the right of them and not being part of that specific occurrence of the subword v.

Example 1. Taking $u = abbab$ and $v = ab$, we get $\binom{abbab}{ab}_q = q^6 + q^5 + q^3 + 1$. Indeed, we have to consider all factorisations of $abbab$ of the form $u_0\, a\, u_1\, b\, u_2$:

(u_1, u_2)	(ε, bab)	(b, ab)	(bba, ε)	$(\varepsilon, \varepsilon)$
$\|u_1\| + 2\|u_2\|$	6	5	3	0

This result implies, for instance, that the constant term $\binom{u}{v}_q\big|_{q=0}$ equals 1 if and only if v is a suffix of u; otherwise, it equals 0. Similarly, the degree of $\binom{u}{v}_q$ is less than or equal to $|v|(|u| - |v|)$ and the coefficient of the monomial $q^{|v|(|u|-|v|)}$ is 1 if and only if v is a prefix of u; otherwise, it equals 0. We thus have more information with $\binom{u}{v}_q$ than with $\binom{u}{v}$, as shown by the following result.

Proposition 1 ([8]). *Let $u \in A^*$ and $1 \leq k \leq |u|$. The sequence $(\binom{u}{x}_q)_{x \in A^k}$ uniquely determines the word u.*

Of course, many properties of the classical binomial coefficient of words can be generalised to our q-deformation. For instance, we have the following q-analogues of the Chu-Vandermonde identity and the sums of binomial coefficients over words of a given length.

Proposition 2 ([8]). *For all words $x, y, u \in A^*$, we have*

$$\binom{xy}{u}_q = \sum_{\substack{u = u_1 u_2 \\ u_1, u_2 \in A^*}} q^{|u_1|(|y| - |u_2|)} \binom{x}{u_1}_q \binom{y}{u_2}_q.$$

Proposition 3 ([8]). *Let $u, v \in A^*$ be words and $n \geq 1$ be an integer. We have*

$$\sum_{v \in A^n} \binom{u}{v}_q = \binom{|u|}{n}_q \quad \text{and} \quad \sum_{u \in A^n} \binom{u}{v}_q = (\#A)^{n-|v|} \binom{n}{|v|}_q,$$

where on the right-hand sides we have Gaussian binomial coefficients of integers.

An interesting application of these q-deformations occurs in a generalisation of a celebrated theorem of Eilenberg, also credited to Schützenberger [3]. It provides a characterisation of p-group languages using binomial coefficients of words. More precisely, a language is a p-group language if and only if it is a Boolean combination of languages of the form

$$L_{v,r,p} := \left\{ u \in A^* \mid \binom{u}{v} \equiv r \pmod{p} \right\},$$

where p is a prime, $0 \leq r \leq p-1$ is an integer and $v \in A^*$. Recall that a p-group language is a language recognised by a p-group, *i.e.* there exist a finite p-group M, a subset $S \subseteq M$ and a morphism μ such that $L = \mu^{-1}(S)$.

Theorem 2 ([8]). *Let p be a prime and $\mathfrak{M} = a(q-1)^d$ with $d \geq 1$ an integer and a non-zero $a \in \mathbb{F}_p = \mathbb{Z}/p\mathbb{Z}$. A language is a p-group language if and only if it is a Boolean combination of languages of the form*

$$L_{v,\mathfrak{R},\mathfrak{M}} = \left\{ u \in A^* \mid \binom{u}{v}_q \equiv \mathfrak{R} \pmod{\mathfrak{M}} \right\},$$

where $v \in A^*$ and $\mathfrak{R} \in \mathbb{F}_p[q]$ is a polynomial of degree less than $\deg(\mathfrak{M})$.

2 q-Parikh matrices

When talking about binomial coefficients of words, one often thinks about Parikh matrices [5]. Given a finite word $u \in \{1, \ldots, k\}^*$, it is an upper triangular matrix $M(u)$ of size $(k+1) \times (k+1)$ whose elements are binomial coefficients of the form $\binom{u}{i(i+1)\cdots j}$, $1 \leq i \leq j \leq k$. In particular, the second diagonal corresponds to the *Parikh vector* of u. Şerbănuţă has generalised Parikh matrices to matrices induced by a word $z = z_1 \cdots z_n$ [11]. These matrices have size $(|z|+1) \times (|z|+1)$ and they contain elements of the form $\binom{u}{z_i z_{i+1} \cdots z_j}$, $1 \leq i \leq j \leq |z|$. Taking $z = 12 \cdots k$ leads back to the definition of classical Parikh matrices.

It is therefore natural to suggest a q-deformation of those Parikh matrices. Already introduced by Eğecioğlu and Ibarra [1,2], we propose here a definition that matches our q-analogue of binomial coefficients of words [9].

Definition 2. *Let $z = z_1 \cdots z_\ell$ be a word and A be the alphabet of z, i.e., the set of letters occurring in z. For $d \in A$ and $j \geq 0$, we let $\mathcal{M}_{d,j}$ denote the upper triangular matrix having 1's on the diagonal and the only non-zero elements above the diagonal are $(\mathcal{M}_{d,j})_{i,i+1} = q^j$ for all i such that $z_i = d$. We now define the map*

$$\mathcal{P}_z : A^* \to (\mathbb{N}[q])^{(|z|+1)\times(|z|+1)},\ u_k u_{k-1} \cdots u_1 u_0 \mapsto \mathcal{M}_{u_k,k} \cdots \mathcal{M}_{u_0,0}.$$

For a word u, we say that $\mathcal{P}_z(u)$ is the q-Parikh matrix *of u induced by z.*

In particular, given a word z and $u \in A^*$ (with A the alphabet of z), the matrix $\mathcal{P}_z(u)$ gathers q-deformed binomial coefficients of the form $\binom{u}{v}_q$, where v is a factor of z, up to multiplication by a power of q. Furthermore, this also provides another way of computing q-binomial coefficients of words.

Theorem 3 ([9]). *Let z be a word of length $\ell \geq 1$ whose alphabet is A. Let $u \in A^*$. The corresponding $(\ell+1)\times(\ell+1)$ q-Parikh matrix is such that*

- ⋆ $(\mathcal{P}_z(u))_{i,j} = 0$, *for all* $1 \leq j < i \leq \ell+1$,
- ⋆ $(\mathcal{P}_z(u))_{i,i} = 1$, *for all* $1 \leq i \leq \ell+1$.
- ⋆ *for* $r \in \{1,\ldots,\ell\}$, *and all* $1 \leq i \leq \ell-r+1$, $(\mathcal{P}_z(u))_{i,i+r} = q^{\mathsf{s}(r-1)}\binom{u}{z_i\, z_{i+1}\cdots z_{i+r-1}}_q$, *where* $\mathsf{s}(n) = \sum_{i=1}^{n} i$.

Example 2. Take back the words of Example 1, that is $u = abbab$ and $z = ab$, we have

$$\mathcal{P}_z(u) = \begin{pmatrix} 1 & q^4+q & q^7+q^6+q^4+q \\ 0 & 1 & q^3+q^2+1 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & \binom{u}{a}_q & q\binom{u}{ab}_q \\ 0 & 1 & \binom{u}{b}_q \\ 0 & 0 & 1 \end{pmatrix}.$$

In particular, $\binom{u}{z}_q = \frac{q^7+q^6+q^4+q}{q} = q^6+q^5+q^3+1$.

As a consequence of this result, we can also mention the following:

Theorem 4 ([9]). *The q-binomial $\binom{u^n}{z}_q$, with $n \in \mathbb{N}$, can be expressed as*

$$\frac{1}{q^{\mathsf{s}(|z|-1)}} \sum_{k=1}^{m} R_k(q)\, \frac{1-q^{c_k n|u|}}{1-q^{c_k|u|}},$$

where m and c_k are positive integers and R_k are rational functions whose denominators only have factors of the form $(1-q^{t|u|})$ for some integer t. Moreover, these quantities c_k and R_k can be effectively computed. In particular, the sequence $(\binom{u^n}{z}_q)_{n\geq 0}$ converges in $\mathbb{N}[[q]]$ to the formal power series $\mathfrak{s}_{\mathbf{u},z}(q)$ expressed by the rational function

$$\frac{1}{q^{\mathsf{s}(|z|-1)}} \sum_{k=1}^{m} R_k(q)\, \frac{1}{1-q^{c_k|u|}}.$$

Corollary 1 ([9]). *The sequence of q-binomials $(\binom{u^n}{z}_q)_{n\geq 0}$ satisfies a linear recurrence relation with polynomial coefficients. In particular, the sequence of binomials $(\binom{u^n}{z})_{n\geq 0}$ satisfies a linear recurrence relation with constant coefficients.*

This generalises Salomaa's result [10], which we can recover by taking $q = 1$.

Example 3. The sequence $\left(\binom{(abba)^n}{ab}_q\right)_{n\geq 0}$ converges to the series

$$q^3+2q^4+q^5+q^7+2q^8+q^9+2q^{11}+4q^{12}+2q^{13}+2q^{15}+4q^{16}+2q^{17}+3q^{19}+6q^{20}+\cdots,$$

which corresponds to the rational function $R(q) = \frac{q^3}{(q-1)^2(q^2+1)^2(q^4+1)}$. One can then notice that the sequence also satisfies the relation

$$p_{n+3} = (1+q^4+q^8)p_{n+2} - (q^4+q^8+q^{12})p_{n+1} + q^{12}p_n,$$

so that the integer sequence $\left(\binom{(abba)^n}{ab}\right)_{n\geq 0}$ satisfies

$$p_{n+3} = 3p_{n+2} - 3p_{n+1} + p_n.$$

References

1. Eğecioğlu, Ö.: A q-matrix encoding extending the Parikh matrix mapping. In: Proc. Int. Conf. on Computers and Communications (Oradea, Romania) (2004)
2. Eğecioğlu, Ö., Ibarra, O.H.: A matrix q-analogue of the Parikh map. In: Exploring new frontiers of theoretical informatics. IFIP 18th world computer congress, TC1 3rd international conference on theoretical computer science (TCS2004), 22–27 August 2004, Toulouse, France., pp. 125–138. Boston, MA: Kluwer Academic Publishers (2004)
3. Eilenberg, S.: Automata, languages, and machines. Vol. B. Pure and Applied Mathematics, Vol. 59, Academic Press [Harcourt Brace Jovanovich, Publishers], New York-London (1976)
4. Lothaire, M.: Combinatorics on Words. Cambridge Mathematical Library. Cambridge University Press (1997). `https://doi.org/10.1017/CBO9780511566097`
5. Mateescu, A., Salomaa, A., Salomaa, K., Yu, S.: A sharpening of the Parikh mapping. RAIRO Theor. Informatics Appl. **35**(6), 551–564 (2001). `https://doi.org/10.1051/ITA:2001131`
6. Morier-Genoud, S., Ovsienko, V.: q-deformed rationals and q-continued fractions. Forum Math. Sigma **8**, Paper No. e13, 55 (2020). `https://doi.org/10.1017/fms.2020.9`
7. Morier-Genoud, S., Ovsienko, V.: On q-deformed real numbers. Exp. Math. **31**(2), 652–660 (2022). `https://doi.org/10.1080/10586458.2019.1671922`
8. Renard, A., Rigo, M., Whiteland, M.: Introducing q-deformed binomial coefficients of words (2024). `https://doi.org/https://doi.org/10.48550/arXiv.2402.05838`, submitted
9. Renard, A., Rigo, M., Whiteland, M.: q-parikh matrices and q-deformed binomial coefficients of words (2024). `https://doi.org/https://doi.org/10.48550/arXiv.2402.05657`, submitted
10. Salomaa, A.: Subword histories and associated matrices. Theoret. Comput. Sci. **407**(1-3), 250–257 (2008). `https://doi.org/10.1016/j.tcs.2008.05.023`
11. Şerbănuţă, T.F.: Extending Parikh matrices. Theoret. Comput. Sci. **310**(1-3), 233–246 (2004). `https://doi.org/10.1016/S0304-3975(03)00396-7`