GENERALIZED DESCENT PATTERNS IN PERMUTATIONS 
AND ASSOCIATED HOPF ALGEBRAS

JEAN-CHRISTOPHE NOVELLI, CHRISTOPHE REUTENAUER, JEAN-YVES THIBON

Abstract. Descents in permutations or words are defined from the relative position of two consecutive letters. We investigate a statistic involving patterns of \( k \) consecutive letters, and show that it leads to Hopf algebras generalizing noncommutative symmetric functions and quasi-symmetric functions.

1. Introduction and Background

Recall that the standardized word \( \text{std}(w) \) of a word \( w \in A^* \) over an ordered alphabet \( A \) is the permutation obtained by iteratively scanning \( w \) from left to right, and labelling 1, 2, \ldots the occurrences of its smallest letter, then numbering the occurrences of the next one, and so on. For example, \( \text{std}(bbacab) = 341625 \).

A permutation \( \sigma \in \mathfrak{S}_n \) is said to have a descent at \( i \) if \( \sigma_i > \sigma_{i+1} \). One can alternatively say that the standardization of the two-letter word \( \sigma_i \sigma_{i+1} \) is 21, and the descent set of \( \sigma \) can be encoded by the descent pattern \( (\text{std}(\sigma_i \sigma_{i+1})_{i=1}^{n-1}) \), produced by scanning \( \sigma \) with a sliding window of width two.

An obvious generalization of this notion would be to use a window of arbitrary width \( k \). For example, the permutation \( \sigma = (85736124) \) would have as 3-descent pattern the sequence

\[
(1) \quad p = (312, 231, 312, 231, 312, 123) .
\]

This idea immediately raises a couple of questions. It is known that the sums of permutations of \( \mathfrak{S}_n \) having the same descent pattern span a subalgebra \( \Sigma_n \) of the group algebra (Solomon’s descent algebra \([12]\)) and that the direct sum of all the \( \Sigma_n \) has a natural Hopf algebra structure (Noncommutative symmetric functions), inherited from that of the Hopf algebra of permutations \([8, 2]\). Are there analogs of these facts for the generalized descent classes?

We will show that, although the \( k \)-descent classes do not span a subalgebra of the group algebra, the Hopf algebra construction still works, and leads to some interesting combinatorics.

2. Generalized descent patterns and codes

For ordinary descents in \( \mathfrak{S}_n \), a classical encoding of the patterns is by compositions of \( n \). Recall that if the descents of \( \sigma \) form the set \( D = \{d_1, \ldots, d_{r-1}\} \), we encode it by the composition \( I = (i_1, \ldots, i_r) \) of \( n \) such that \( d_j = i_1 + i_2 + \cdots + i_j \). Then, in
the algebra of noncommutative symmetric functions, the concatenation of compositions corresponds to the (outer) multiplication in various bases, appropriately called \textit{multiplicative bases}.

2.1. \textbf{The $k$-descent code.} Another encoding having the same property would be to represent the descent set $D$ by a binary word $b = b_1 b_2 \cdots b_n$, with $b_i = 1$ if $i \in D$ and $b_i = 0$ otherwise (note that $b_n$ is always 0). It is this kind of encoding which is most easily generalized to $k$-descent patterns.

We regard $\sigma \in S_n$ as a permutation of $\mathbb{Z}$ whose support is contained in $[1,n]$ and associate with it the sequence $(d_i \in [1,k])_{i \in \mathbb{Z}}$, where $p_i$ is the relative position of $\sigma(i)$ w.r.t. its $k-1$ predecessors $\sigma(i-1), \ldots, \sigma(i-k+1)$, that is, $d_i = j$ if $\sigma(i)$ is the $j$th element of the sequence $\sigma(i), \sigma(i-1), \ldots, \sigma(i-k)$ sorted in increasing order. Hence $d_i = k$ for $i \leq 1$ and for $i > n$.

We can therefore identify the $k$-descent pattern with the word of length $n$

\begin{equation}
DC_k(\sigma) = d_1 d_2 \cdots d_n \in [1,k]^n,
\end{equation}

which we will call the $k$-descent code. Indeed, starting from the $k$-descent pattern $p = (p_1, \ldots, p_r)$ of $\sigma$, one recovers the $k$-descent code by first computing the $k$-descent code of $p_1$ and appending to it the final letters of all words $p_2$ up to $p_r$. Conversely, $d_1 \ldots d_k$ gives back $p_1$ since, in that case the sequence $d_1 \ldots d_k$ represents the complement to $k$ of the code of $p_1$ (obtained by changing $i$ into $k-i$), and then gives back $p_k$ from both $p_{k-1}$ and $d_k$ since $p_k$ is the word obtained by removing the first letter of $p_{k-1}$, then adding 1 to all values greater than or equal to $d_k$, and concatenating $d_k$ to the resulting word.

For example, one can check on $DC_3(85736124) = 32212123$ that the above algorithm gives back the sequence $\square$.

Here are two more examples of the $k$-descent code.

\begin{align*}
(3) \quad & DC_3(426135) = 323123 \quad \text{and} \quad DC_4(426135) = 434133.
\end{align*}

2.2. \textbf{The $k$-recoil code and equivalence classes.} The $k$-descent code of the inverse permutation will be called the $k$-recoil code: $RC_k(\sigma) := DC_k(\sigma^{-1})$. It can be computed without inverting permutations in the following way: first see $\sigma$ as a permutation of $\mathbb{Z}$ as before. Then, for all $i$, restrict $\sigma$ to the values between $i$ and $i-k+1$. And then associate with $i$ the position of $i$ in this new word. For example, $RC_3(425163) = 323123$ which is coherent with the previous example since $425163^{-1} = 426135$.

Since we only need to compare letter $i$ with the $k$ preceding letters in the lexicographic order, we can rephrase this construction with the help of the standardization process. Indeed, two permutations $\sigma$ and $\tau$ have same $k$-recoil code iff

\begin{equation}
\forall i \leq n-k+1, \quad std(\sigma_{[i,i+k-1]}) = std(\tau_{[i,i+k-1]}),
\end{equation}

where $\sigma_{[a,b]}$ means the restriction of $\sigma$ to its values in the interval $[a,b]$. We shall then write $\sigma \equiv_k \tau$. 

This can be extended to words over an ordered alphabet: we set \( u \equiv_k v \) iff \( u \) is a rearrangement of \( v \) and \( \text{std}(u) \equiv_k \text{std}(v) \). For \( k = 2 \), this is the hypoplactic congruence (see [6, 9]).

For example, with \( k = 3 \), each equivalence class in \( \mathfrak{S}_n \) with \( n \leq 3 \) has one element, and there are 18 classes in \( \mathfrak{S}_4 \), among which 6 non-singleton classes:

\[(5) \ [1423, 4123], \ [1432, 4132], \ [2143, 2413], \ [2314, 2341], \ [3142, 3412], \ [3214, 3241].\]

For example, the 3-recoil code of both 2314 and 2341 is 3223.

Note that the first letter of the \( k \)-descent (or recoil) code is always \( k \) and the next one is either \( k \) or \( k - 1 \). More generally, a \( k \)-recoil code is a word \( I = (i_1, \ldots, i_r) \) satisfying \( i_\ell \in [\max(k-\ell+1, 1), k] \) for all \( \ell \). Conversely, given a word \( I \) satisfying these conditions, one easily builds a permutation with \( I \) as \( k \)-recoil code. By induction, there exists a permutation \( \sigma \) with \( k \)-recoil code \( I' = (i_1, \ldots, i_{r-1}) \). Now, place \( r \) anywhere between the \( i_r \)-th and the \( 1+i_r \)-th element of \( \sigma \) in the interval \( [r-k+1, r-1] \). This permutation has \( I \) as \( k \)-recoil code. Note that in particular this allows one to build easily the smallest (resp. the largest) elements for the lexicographic order of each equivalence class: put at each step letter \( r \) at the rightmost (resp. leftmost) possible spot.

**Proposition 2.1.** The number \( N(k,n) \) of \( k \)-descent (or recoil) classes of \( \mathfrak{S}_n \) is

\[
N(k,n) = \begin{cases} 
n! & \text{if } n \leq k, \\
k!k^{n-k} & \text{if } n \geq k.
\end{cases}
\]

More precisely, the \( k \)-descent (or recoil) codes \( I = (i_1, \ldots, i_r) \) satisfy

\[
(7) \begin{cases} 
\ell \leq k & \text{if } \ell \leq k, \\
1 \leq \ell \leq k & \text{if } \ell \geq k.
\end{cases} \iff i_\ell \in [\max(k-\ell+1, 1), k].
\]

For example, the 3-recoil codes of all permutations of \( \mathfrak{S}_3 \) and \( \mathfrak{S}_4 \) are, taking the permutations in lexicographic order:

\[(8)\]
\[
333, 332, 323, 322, 331, 321.
\]

\[(9)\]
\[
3333, 3332, 3323, 3322, 3331, 3321, 3323, 3223, 3232, 3222, 3313, 3312, 3213, 3312, 3212, 3321, 3321, 3221, 3311, 3211.
\]

In particular, one can check that the codes 3331, 3321, 3232, 3223, 3312, and 3213 occur twice, and correspond to the six non-singleton 3-classes in \( \mathfrak{S}_4 \) (see Equation (5)).

**2.3. Classes of permutations having the same \( k \)-recoil code.** The following proposition generalizes the fact that two permutations have same recoil code \( (k = 2) \) iff one can go from one to the other by successively exchanging non-consecutive adjacent values.

**Proposition 2.2.** Two permutations \( \sigma \) and \( \mu \) have same \( k \)-recoil code iff one can go from \( \sigma \) to \( \tau \) by successively exchanging adjacent values whose difference is at least \( k \).
Proof – Let us write $\sigma \equiv_k' \tau$ if one can go from $\sigma$ to $\tau$ by exchanging adjacent values whose difference is at least $k$. Then it is obvious that

$$\sigma \equiv_k' \tau \Rightarrow \sigma \equiv_k \tau.$$  

(10) So each $\equiv_k'$ class is contained in an $\equiv_k$ class. In particular, the number of $\equiv_k'$ classes is at least equal to the number of $\equiv_k$ classes, and if those numbers are equal, then so are the equivalence classes.

Now, each $\equiv_k'$ class has a minimal element for the lexicographic order and this element has no consecutive letters $i$ and $j$ so that $i - j \geq k$. Let us denote by $W(k, n)$ the set of words such that there are no consecutive letters $i$ and $j$ so that $i - j \geq k$.

There are at most as many $\equiv_k'$ classes as elements in $W(k, n)$. Observe now that if one removes $n$ from any word of $W(k, n)$, one obtains a word in $W(k, n - 1)$. Moreover, given a word of $W(k, n - 1)$, in order to get a word in $W(k, n)$, one can put $n$ at $n$ different spots if $n \leq k$ or at $k$ different spots (after $n - k + 1, \ldots, n - 1$ or at the end) if $n \geq k$. So $|W(k, n)|$ is equal to $N(k, n)$, the number of $\equiv_k$ classes thanks to the formula relating $N(k, n)$ and $N(k, n - 1)$.

This argument unravels a simple characterization of the minimal element of an equivalence class:

**Corollary 2.3.** The minimal elements of the classes of $\equiv_k$ are the permutations $\sigma$ such that no two adjacent letters satisfy $\sigma_i - \sigma_{i+1} \geq k$.

By symmetry, the maximal elements of $\equiv_k$ are the permutations $\sigma$ such that no two adjacent letters satisfy $\sigma_{i+1} - \sigma_i \geq k$.

Moreover, the set of maximal elements is obtained from the set of minimal elements of $S_n$ by the transformation $\tau' := (n + 1 - \tau_1, \ldots, n + 1 - \tau_n)$.

The proposition implies that the $\equiv_k$ classes split the right permutohedron into connected components. We can be more precise:

**Proposition 2.4.** The set of permutations having a given k-recoil (respectively descent) code is an interval of the right (resp. left) weak order.

Proof – Let $C$ be a $k$-recoil class and let $\alpha_C$ (respectively $\omega_C$) be its minimal (resp. maximal) element. If $\sigma$ has same $k$-RC as $\alpha_C$ (and $\omega_C$), then, thanks to Proposition 2.2, if $\sigma$ is not minimal, $\sigma$ has a pair of adjacent letters $\sigma_i$ and $\sigma_{i+1}$ such that $\sigma_i - \sigma_{i+1} \geq k$. Then one can exchange those two letters and iterate the process until one reaches the minimal element, so that $\sigma \geq \alpha_C$ for the right permutohedron. The same argument proves that $\sigma \leq \omega_C$.

Conversely, all permutations of the interval $[\alpha_C, \omega_C]$ of the right permutohedron have same $k$-RC: their inversion sets are contained in the inversion set of $\omega_C$ and contain the inversion set of $\alpha_C$, so that letters $i, i + 1, \ldots, i + k - 1$ have same relative positions.
As in the case of ordinary descents, the order ideals defined by maximal elements are unions of classes. This property is essential for defining multiplicative bases in the associated Hopf algebras.

**Proposition 2.5.** Let $\omega_C$ be the maximal element of an $\equiv_k$ class. Then the interval $[\text{id}, \omega_C]$ of the right permutohedron is an union of $\equiv_k$ classes, where $\text{id}$ is the identity permutation.

Moreover, the interval $[\alpha_C, \omega]$ of the right permutohedron is also an union of $\equiv_k$ classes, where $\alpha_C$ is the minimal element of an $\equiv_k$ class and $\omega$ is the maximal permutation.

**Proof** – By Corollary 2.3 the second statement is equivalent to the first one.

Thus, we must prove that if $x \leq \omega_C$ then the maximal element $\omega_C'$ of its $\equiv_k$ class satisfies $\omega_C' \leq \omega_C$. If $x$ is maximal, we are done. Otherwise, thanks to the characterization of the maximal elements, we have

$$x = \ldots i j \ldots,$$

where $j - i \geq k$. We shall prove that the permutation $x'$ obtained from $x$ by exchanging $i$ and $j$ also satisfies $x' \leq \omega_C$. Then, since all classes are intervals, we see by induction on the distance from $x$ to $\omega_C'$ that $\omega_C'$ also satisfies $\omega_C' \leq \omega_C$. Consider the subset of the elements of the permutohedron greater than $x$ such that $i$ and $j$ are not exchanged, that is the set of elements greater than $x$ and not greater than $x'$. This set does not contain any maximal element: the values between $i$ and $j$ in such permutations can only be either smaller than $i$ or greater than $j$ so that there are always two consecutive values with difference at least $k$.

---

2.4. **$k$-Eulerian numbers and polynomials.** The classical Eulerian polynomials count permutations by their number of descents, or equivalently, by the number of 2s in their 2-descent code.

In this form, the definition can be easily generalized. We define the $k$-Eulerian polynomial $E_{n,k}$ as the sum over $S_n$ of the product of $t_i$ where $i$ runs through all entries except the first one of their $k$-recoil code.

For example, with $k = 3$, we obtain from (8) and (9)

$$E_{1,3} = 1; \quad E_{2,3} = t_2 + t_3; \quad E_{3,3} = t_1 t_2 + t_1 t_3 + 2 t_2 t_3 + t_2^2 + t_3^2;$$

(12) $$E_{4,3} = t_1^2 t_2 + t_1^2 t_3 + 2 t_1 t_2^2 + 7 t_1 t_2 t_3 + 3 t_1 t_3^2 + t_2^3 + 5 t_2^2 t_3 + 3 t_2 t_3^2 + t_3^3.$$

2.5. **$k$-Major index.** The classical major index is the sum of the positions of the descents. We can replace the monomial $q^{\text{maj}(\sigma)}$ by the product $\prod_i q^{i-1}_{C_i}$ where $C = \text{RC}(\sigma)$.

For example, with $k = 3$, the $k$-major index polynomials of the first symmetric groups are

$$M_{1,3} = 1; \quad M_{2,3} = q_2 + q_3; \quad M_{3,3} = q_1^2 q_2 + q_1^2 q_3 + q_2^2 q_3 + q_2^3 + q_2 q_2^2 + q_3^3.$$
3. Associated combinatorial Hopf algebras

The equivalence relation $\equiv_k$ can be used to define subalgebras and quotients of the Hopf algebra of permutations, generalizing respectively noncommutative symmetric functions and quasi-symmetric functions (cf. [3]).

Recall that the Hopf algebra of permutations, introduced in [8], can be realized as the algebra $\text{FQSym}$ (Free Quasi-Symmetric functions, cf. [2]), spanned by the polynomials

\[ G_\sigma(A) = \sum_{\text{std}(w) = \sigma} w. \]

3.1. Subalgebras. Imitating the case $k = 2$, we define generalized ribbons by

\[ R_C = \sum_{\text{DC}_k(\sigma) = C} G_\sigma \]

for a $k$-descent code $C$.

This basis generalizes the classical (strict) descent classes. We can also generalize the large descent classes (permutations whose descent set is contained in a prescribed one), corresponding to the multiplicative basis $S_I$ of $\text{Sym}$. We set

\[ S^C := S^{\omega C} = \sum_{\sigma \leq \omega C} G_\sigma \quad \text{and} \quad E^C := E^{\alpha C} = \sum_{\sigma \geq \alpha C} G_\sigma, \]

where $\leq$ and $\geq$ correspond to the left weak order.

The products of the $S^\sigma$ and the $E^\sigma$ in $\text{FQSym}$ are well-known [1]:

\[ S^\sigma S^\tau = S^{\mu[I]} \cdot \tau, \quad E^\sigma E^\tau = E^{\mu[\sigma]}, \]

where $\mu[i] = (\mu_1 + i, \ldots, \mu_n + i)$.

We can then state:

**Theorem 3.1.** The $S^C$ and the $E^C$ are multiplicative bases of a subalgebra $\text{DSym}^{(k)}$ of $\text{FQSym}$. The $R_C$ are also a basis of $\text{DSym}^{(k)}$.

Moreover, $\text{DSym}^{(k)}$ is free as an algebra over the $S^\sigma$ (respectively $E^\sigma$) indexed by permutations that are both mirror images of connected permutations and maximal (resp. minimal) elements of $\equiv_k$ classes.

**Proof** – Since the shifted concatenation $\sigma \cdot \tau[|\sigma|]$ of two minimal elements of an $\equiv_k$ class is a minimal element, the product of two $E^C$ is internal, so $\text{DSym}^{(k)}$ is a subalgebra of $\text{FQSym}$ and the $E^C$ are a multiplicative basis. By inclusion-exclusion, we get that the $R_C$ (and the $S^C$ too) are bases of $\text{DSym}^{(k)}$. The fact that the $S$ are multiplicative imply that the $S$ are also multiplicative.

Moreover, since the $E^\sigma$ indexed by permutations that are mirror images of connected permutations generated $\text{FQSym}$, the free subalgebra $\text{DSym}^{(k)}$ is free over a set of generators indexed by permutations that are both mirror images of connected permutations and minimal elements of $\equiv_k$ classes. The same holds for the $S^\sigma$. $\blacksquare$

For example,

\[ R_{321}R_{3321} = R_{3211221} + R_{3211321} + R_{3212321} + R_{3213321}. \]
The Hilbert series of $\text{DSym}^{(k)}$ is

$$(20) \quad H_k(t) = \sum_{j=0}^{k-1} j!t^j + \frac{k!t^k}{1-kt}$$

and the generating series $G_k(t)$ for the number of generators by degree is given by

$$(21) \quad \frac{1}{1 - G_k(t)} = H_k(t)$$

With $k = 3$ and 4, one finds

$$(22) \quad G_3(t) = t + t^2 + 3t^3 + 7t^4 + 17t^5 + 99t^7 + 239t^8 + 577t^9 + 1393t^{10} + 3363t^{11} + 8119t^{12} + 19601t^{13} + \ldots$$

which is Sequence A001333 of [11].

$$(23) \quad G_4(t) = t + t^2 + 3t^3 + 13t^4 + 47t^5 + 173t^6 + 639t^7 + 2357t^8 + 8695t^9 + 32077t^{10} + 118335t^{11} + 436549t^{12} + 1610471t^{13} + \ldots$$

which is Sequence A084519 of [11]. With $k = 5$, the sequence is not (yet) in [11].

**Theorem 3.2.** $\text{DSym}^{(k)}$ is a Hopf subalgebra of $\text{FQSym}$. 

**Proof** – We already know that $\text{DSym}^{(k)}$ is a subalgebra of $\text{FQSym}$. So there only remains to prove that $\text{DSym}^{(k)}$ is a subcoalgebra of $\text{FQSym}$. 

Thanks to the definition of the coproduct of $G_{\sigma}$ in $\text{FQSym}$, this amounts to a trivial combinatorial property: if two permutations of size $n$ have same standardized word on each factor of a given size, then, adding given letters all greater than $n$ (or all smaller than 1) at the same position in both permutations does not change the property: the resulting permutations also have same standardized word on each factor of the previous size. 

**Corollary 3.3.** $\text{DSym}^{(k)}$ is a Hopf subalgebra of $\text{DSym}^{(l)}$ for $k < l$. In particular, the $\text{DSym}^{(k)}$ interpolate between $\text{Sym} = \text{DSym}^{(2)}$ and $\text{FQSym} = \text{DSym}^{\infty}$. 

### 3.2. Duality.

Let us denote by $\text{DQSym}^{(k)} = \text{DSym}^{(k)^*}$, the dual bialgebra of $\text{DSym}^{(k)}$. Dualizing Theorem 3.2 we obtain:

**Theorem 3.4.** $\text{DQSym}^{(k)} = \text{DSym}^{(k)^*} := \text{FQSym}/ \equiv_k$ is a noncommutative (for $k > 2$) and non-cocommutative Hopf algebra. It is also a (non-free) dendriform quotient of $\text{FQSym}$. 

We shall write $F_C = \underline{F}_{\sigma}$ where $C$ is the $k$-descent code of $\sigma$. For example, $52413 \equiv_3 21543$, so that 42531 and 21543 have same 3-descent composition. We have

$$(24) \quad \Delta F_{42531} = F_{42531} \otimes 1 + F_{3142} \otimes F_1 + F_{213} \otimes F_{21} + F_{21} \otimes F_{321} + F_1 \otimes F_{2431} + 1 \otimes F_{42531},$$
and
\[
\Delta F_{21543} = F_{21543} \otimes 1 + F_{2143} \otimes F_1 + F_{213} \otimes F_{21} + F_{21} \otimes F_{321} + F_1 \otimes F_{1432} + 1 \otimes F_{21543}.
\]
(25)

Moreover,
\[
F_{42531} F_1 = F_{425316} + F_{425631} + F_{426531} + F_{462531} + F_{642531},
\]
(26)
and
\[
F_{21543} F_1 = F_{215436} + F_{215463} + F_{215643} + F_{216543} + F_{261543} + F_{621543},
\]
(27)
and one easily checks that the indices of both expressions match in a one-to-one correspondence by the \(\equiv_3\) relation on the inverse permutations.

References

[1] G. Duchamp, F. Hivert, J.-C. Novelli, and J.-Y. Thibon, Noncommutative Symmetric Functions VII: Free Quasi-Symmetric Functions Revisited, preprint, math.CO/0809.4479.
[2] G. Duchamp, F. Hivert, and J.-Y. Thibon, Noncommutative symmetric functions VI: free quasi-symmetric functions and related algebras, Internat. J. Alg. Comput. 12 (2002), 671–717.
[3] I.M. Gelfand, D. Krob, A. Lascoux, B. Leclerc, V. S. Retakh, and J.-Y. Thibon, Noncommutative symmetric functions, Adv. in Math. 112 (1995), 218–348.
[4] F. Hivert and N. Thiéry, MuPAD-Combinat, an open-source package for research in algebraic combinatorics, Sém. Lothar. Combin. 51 (2004), 70p. (electronic).
[5] D. Krob, B. Leclerc, and J.-Y. Thibon, Noncommutative symmetric functions II: Transformations of alphabets, Internal J. Alg. Comput. 7 (1997), 181–264.
[6] D. Krob and J.-Y. Thibon, Noncommutative symmetric functions IV: Quantum linear groups and Hecke algebras at \(q = 0\), J. Alg. Comb. 6 (1997), 339–376.
[7] I.G. Macdonald, Symmetric functions and Hall polynomials, 2nd ed., Oxford University Press, 1995.
[8] C. Malvenuto and C. Reutenauer, Duality between quasi-symmetric functions and Solomon descent algebra, J. Algebra, 177 (1995), 967–982.
[9] J.-C. Novelli, On the hypoplactic monoid, Disc. Math. 217 (2000), 315–336.
[10] C. Reutenauer, Free Lie algebras, Oxford University Press, 1993.
[11] N.J.A. Sloane, The On-Line Encyclopedia of Integer Sequences, http://www.research.att.com/~njas/sequences/
[12] L. Solomon, A Mackey formula in the group ring of a Coxeter group, J. Algebra, 41, (1976), 255–268.