Polyhedral realizations of crystal bases and convex-geometric Demazure operators

Naoki Fujita

Published online: 13 November 2019
© Springer Nature Switzerland AG 2019

Abstract
The main object in this paper is a family of rational convex polytopes whose lattice points give a polyhedral realization of a highest weight crystal basis. Every polytope in this family is identical to a Newton–Okounkov body of a flag variety, and it gives a toric degeneration. In this paper, we prove that a specific class of polytopes in this family is given by Kiritchenko’s Demazure operators on polytopes. This implies that polytopes in this class are all lattice polytopes. As an application, we give a sufficient condition for the corresponding toric variety to be Gorenstein Fano.

Keywords Nakashima–Zelevinsky’s polyhedral realization · Crystal basis · Demazure operator · Toric degeneration

Mathematics Subject Classification Primary 05E10; Secondary 14M15 · 14M25 · 52B20

Contents

1 Introduction ............................................. 2
2 Convex-geometric Demazure operators ............................................. 4
3 Polyhedral realizations of crystal bases ............................................. 7
4 Main result .............................................. 16
   4.1 Statement of the main result ................................... 16
   4.2 Proof of Theorem 4.1 ...................................... 18
   4.3 Immediate consequences .................................... 27
5 Crystal structures ........................................... 27
6 Geometric applications ........................................ 30
References ................................................ 34

The work was partially supported by Grant-in-Aid for JSPS Fellows (No. 16J00420).

Naoki Fujita
nfujita@ms.u-tokyo.ac.jp

1 Graduate School of Mathematical Sciences, The University of Tokyo, 3-8-1 Komaba, Meguro-ku, Tokyo 153-8914, Japan
1 Introduction

The theory of crystal bases \cite{18,19} gives a combinatorial skeleton of a representation of a semisimple Lie algebra. In the theory of crystal bases, it is important to give their concrete realizations. Until now, many useful realizations have been discovered; Nakashima–Zelevinsky’s polyhedral realization \cite{34,37} is one of them. It realizes a highest weight crystal basis as the set of lattice points in some rational convex polytope. This polytope is called a Nakashima–Zelevinsky polytope. The author and Naito \cite{10} proved that the Nakashima–Zelevinsky polytope is identical to a Newton–Okounkov body of a flag variety. The theory of Newton–Okounkov bodies was introduced by Okounkov \cite{38–40}, and afterward developed independently by Kaveh–Khovanskii \cite{24,25} and by Lazarsfeld–Mustata \cite{30}. A remarkable fact is that the theory of Newton–Okounkov bodies gives a systematic method of constructing toric degenerations [2, Theorem 1]; in particular, there exists a flat degeneration of the flag variety to the normal toric variety associated with the Nakashima–Zelevinsky polytope. In this paper, we relate Nakashima–Zelevinsky polytopes with Demazure operators on polytopes.

To be more precise, let $\mathfrak{g}$ be a semisimple Lie algebra, $P_+$ the set of dominant integral weights, $I = \{1, \ldots, n\}$ an index set for the vertices of the Dynkin diagram, and $\{\alpha_i \mid i \in I\}$ the set of simple roots. For $\lambda \in P_+$, we denote by $V(\lambda)$ the irreducible highest weight $\mathfrak{g}$-module with highest weight $\lambda$, and by $B(\lambda)$ the crystal basis for $V(\lambda)$. Fix a reduced word $i = (i_1, \ldots, i_N) \in I^N$ for the longest element $w_0$ in the Weyl group. We associate to $i$ a specific parametrization $\Psi_i: B(\lambda) \hookrightarrow \mathbb{Z}^N$ of $B(\lambda)$, which gives an explicit description of the crystal structure; see Sect. 3 for the precise definition. Nakashima–Zelevinsky \cite{37} and Nakashima \cite{34} described explicitly the image $\Psi_i(B(\lambda))$ under some technical assumptions on $\lambda$. The author and Naito \cite{10} proved that the image $\Psi_i(B(\lambda))$ is identical to the set of lattice points in some rational convex polytope $\Delta_i(\lambda)$ without any assumptions on $\lambda$. We call $\Delta_i(\lambda)$ the Nakashima–Zelevinsky polytope associated with $i$ and $\lambda$.

The theory of Demazure operators on polytopes was introduced by Kiritchenko \cite{26} to construct a (possibly virtual) convex polytope, whose lattice points yield the character of $V(\lambda)$. For instance, Gelfand–Zetlin polytopes \cite{12} and Grossberg–Karshon’s twisted cubes \cite{14} are obtained in a uniform way (see \cite{26}). For $i \in I$ and $1 \leq k \leq N$ with $i_k = i$, let $D_i^{(k)}$ denote the corresponding Demazure operator on polytopes; see Sect. 2 for the precise definition. This operator is defined for a specific class of polytopes, called parapolytopes. Our purpose is to compute $D_{i_N}^{(N)} \cdots D_{i_1}^{(1)}(a)$ for specific $a \in \mathbb{R}^N$. Note that $D_{i_N}^{(N)} \cdots D_{i_1}^{(1)}(a)$ is not necessarily well-defined as we will see in Example 2.4. For $i \in I$, we denote by $d_i$ the number of $1 \leq k \leq N$ such that $i_k = i$. For $\lambda \in P_+$, we write $\lambda = \sum_{i \in I} \lambda_i d_i \alpha_i$; regard this identity as the definition of coefficients $\lambda_i$. We set

$$a_\lambda := -\Psi_i(b_{w_0, \lambda}) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}),$$

where $b_{w_0, \lambda} \in B(\lambda)$ is the lowest weight element. For subsets $X, Y \subset \mathbb{R}^N$, we define $X + Y$ to be the Minkowski sum:
X + Y := \{x + y \mid x \in X, \ y \in Y\}.

The following are the main results of this paper.

**Theorem 1** (Theorem 4.1) Let $i = (i_1, \ldots, i_N) \in I^N$ be a reduced word for $w_0$, and $\lambda \in P_+$. Assume that the Nakashima–Zelevinsky polytope $\Delta_i(\lambda)$ is a parapolytope.

1. The polytope $\Delta_i(\lambda)$ is a lattice polytope.
2. The polytope $D_{i_1}^{(N)} \cdots D_{i_1}^{(1)}(a_{i\lambda})$ is well-defined.
3. The following equality holds:

$$D_{i_1}^{(N)} \cdots D_{i_1}^{(1)}(a_{i\lambda}) = -\Delta_i(\lambda) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}).$$

**Theorem 2** (Theorem 4.10) Let $i \in I^N$ be a reduced word for $w_0$, and $\lambda, \mu \in P_+$. Assume that the polytopes $\Delta_i(\lambda)$, $\Delta_i(\mu)$, and $\Delta_i(\lambda + \mu)$ are all parapolytopes. Then, the following equalities hold:

$$\Psi_i(B(\lambda + \mu)) = \Psi_i(B(\lambda)) + \Psi_i(B(\mu)), \text{ and } \Delta_i(\lambda + \mu) = \Delta_i(\lambda) + \Delta_i(\mu).$$

We give some examples of $\Delta_i(\lambda)$ which are parapolytopes.

**Example 3** (Examples 4.2, 4.3, 4.4) The Nakashima–Zelevinsky polytope $\Delta_i(\lambda)$ is a parapolytope for all $\lambda \in P_+$ if

(i) $\varphi$ is of type $A_n$, and $i = (1; 2, 1; 3, 2, 1; \ldots; n, n - 1, \ldots, 1)$;

(ii) $\varphi$ is of type $B_n$ or $C_n$, and $i = (n, n - 1, \ldots, 1; n, n - 1, \ldots, 1; \ldots; n, n - 1, \ldots, 1) \in I^{\binom{n}{2}}$;

(iii) $\varphi$ is of type $D_n$, and $i = (n, n - 1, \ldots, 1; n, n - 1, \ldots, 1; \ldots; n, n - 1, \ldots, 1) \in I^{\binom{n(n-1)}{2}}$;

(iv) $\varphi$ is of type $G_2$, and $i = (1, 2, 1, 2, 1, 2)$ or $i = (2, 1, 2, 1, 2, 1)$.

Let $G/B$ be the full flag variety associated with $\varphi$, and $X(\Delta_i(\lambda))$ the normal toric variety associated with the rational convex polytope $\Delta_i(\lambda)$. Then, we obtain a flat degeneration of $G/B$ to $X(\Delta_i(\lambda))$ by the theory of Newton–Okounkov bodies [2]; such a degeneration to a toric variety is called a toric degeneration. Toric degenerations of $G/B$ have been studied from various points of view such as standard monomial theory [5,13], string parametrizations of dual canonical bases [1,4], Newton–Okounkov bodies [7,9,23,27], and so on; see [8] for a survey on this topic. Let $P_{++} \subset P_+$ denote the set of regular dominant integral weights. In this paper, we apply Alexeev–Brion’s argument [1] to $\Delta_i(\lambda)$, which implies that the toric varieties $X(\Delta_i(\lambda)), \lambda \in P_{++}$, are all identical and Gorenstein Fano if

(i) $\Delta_i(\lambda + \mu) = \Delta_i(\lambda) + \Delta_i(\mu)$ for all $\lambda, \mu \in P_+$;

(ii) the polytope $\Delta_i(2\rho)$ is a lattice polytope,

where $\rho$ is the half sum of the positive roots. Hence we obtain the following by Theorems 1, 2.
Corollary 4  Take \( g \) and \( i \) as in Example 3. Then, the toric varieties \( X(\Delta_1(\lambda)) \), \( \lambda \in P_{++} \), are all identical and Gorenstein Fano.

If \( g \) is of type \( A_n \), and \( i = (1; 2, 1; 3, 2, 1; \ldots; n, n - 1, \ldots, 1) \), then the Nakashima–Zelevinsky polytope \( \Delta_1(\lambda) \) is identical to the corresponding Gelfand–Zetlin polytope (see Example 3.12). Hence in this case, Theorems 1, 2 and Corollary 4 are not new (see [1, 26]).

In addition, we mention that a relation between convex-geometric Demazure operators and the additivity with respect to the Minkowski sum is discussed in [28].

This paper is organized as follows. In Sect. 2, we recall the definition of Kiritchenko’s Demazure operators on polytopes. In Sect. 3, we review some basic facts about crystal bases and their polyhedral realizations. In Sect. 4, we prove Theorems 1, 2 above. In Sect. 5, we study the crystal structure on the set of lattice points in \( \Delta_1(\lambda) \). Section 6 is devoted to some applications to toric varieties associated with Nakashima–Zelevinsky polytopes; in particular, we show Corollary 4 above.

2 Convex-geometric Demazure operators

Let \( G \) be a connected, simply-connected semisimple algebraic group over \( \mathbb{C} \), \( g \) its Lie algebra, \( W \) the Weyl group, \( I = \{1, \ldots, n\} \) an index set for the vertices of the Dynkin diagram, and \((c_{i,j})_{i,j \in I}\) the Cartan matrix. We fix a reduced word \( i = (i_1, \ldots, i_N) \in I^N \) for the longest element \( w_0 \in W \). For \( i \in I \), let \( d_i \) denote the number of \( 1 \leq k \leq N \) such that \( i_k = i \). We identify \( \mathbb{R}^N \) with the direct sum \( \mathbb{R}^{d_1} \oplus \cdots \oplus \mathbb{R}^{d_n} \) as follows:

\[
\mathbb{R}^N \sim \mathbb{R}^{d_1} \oplus \cdots \oplus \mathbb{R}^{d_n},
\]

\[
(a_1, \ldots, a_N) \mapsto (a_1^{(1)}, \ldots, a_1^{(d_1)}, \ldots, a_n^{(n)}),
\]

where we set \((a_1^{(i)}, \ldots, a_{d_i}^{(i)}) := (a_k)_{1 \leq k \leq N; i_k = i}\). If we define an \( \mathbb{R} \)-linear subspace \((\mathbb{R}^{d_i})^\perp \subset \mathbb{R}^N\) to be

\[
(\mathbb{R}^{d_i})^\perp := \bigoplus_{1 \leq j \leq n; j \neq i} \mathbb{R}^{d_j},
\]

then we have \( \mathbb{R}^N = (\mathbb{R}^{d_i})^\perp \oplus \mathbb{R}^{d_i} \). A subset \( P \subset \mathbb{R}^N \) is called a convex polytope if it is the convex hull of a finite number of points. Let \( \mathcal{P}_N \) denote the set of convex polytopes in \( \mathbb{R}^N \). This set is endowed with a commutative semigroup structure by the Minkowski sum of convex polytopes:

\[
P_1 + P_2 := \{p_1 + p_2 \mid p_1 \in P_1, \; p_2 \in P_2\}.
\]

For \( c \in \mathbb{R}_{\geq 0} \) and a convex polytope \( P \subset \mathbb{R}^N \), define a convex polytope \( cP \subset \mathbb{R}^N \) by \( cP := \{cp \mid p \in P\} \). We denote by \( F(\mathbb{R}^N) \) the set of \( \mathbb{R} \)-valued functions on \( \mathbb{R}^N \). For a convex polytope \( P \subset \mathbb{R}^N \), let \( \mathbb{I}_P \in F(\mathbb{R}^N) \) be the characteristic function of \( P \), that is,
\[ \mathbb{P}(x) = \begin{cases} 1 & \text{if } x \in P, \\ 0 & \text{otherwise.} \end{cases} \]

**Definition 2.1** [26, Definition 2] A convex polytope \( P \subset \mathbb{R}^N \) is called a *parapolytope* if for all \( i \in I \) and \( c \in \mathbb{R}^N \), there exist \( \mu = (\mu_1, \ldots, \mu_{d_i}), \nu = (\nu_1, \ldots, \nu_{d_i}) \in \mathbb{R}^{d_i} \) such that

\[ P \cap (c + \mathbb{R}^{d_i}) = c + \Pi(\mu, \nu) \]

or \( P \cap (c + \mathbb{R}^{d_i}) = \emptyset \), where \([\mu_k, \nu_k] := \{x \in \mathbb{R} | \mu_k \leq x \leq \nu_k\} \subset \mathbb{R}\) for \( 1 \leq k \leq d_i \), and

\[ \Pi(\mu, \nu) := [\mu_1, \nu_1] \times \cdots \times [\mu_{d_i}, \nu_{d_i}] \subset \mathbb{R}^{d_i}. \]

Let \( \mathcal{P} \subset \mathcal{P}_N \) denote the set of parapolytopes in \( \mathbb{R}^N \). For \( 1 \leq k \leq N \), we set

\[ \mathcal{P}(k) := \{ P \in \mathcal{P} | \text{the coordinate function } a_k \text{ is constant on } P \}. \]

For \( i \in I \), define an \( \mathbb{R} \)-linear function \( l_i : \mathbb{R}^N \to \mathbb{R} \) by

\[ l_i(a) := -\sum_{j \in I : j \neq i} c_{i,j}(a_1^{(j)} + \cdots + a_{d_j}^{(j)}). \]

Following [26, Sect. 2.3], we define a *convex-geometric Demazure operator* \( D_i^{(k)} : \mathcal{P}(k) \to F(\mathbb{R}^N) \) for \( i \in I \) and \( 1 \leq k \leq N \) such that \( i_k = i \) as follows. We take \( P \in \mathcal{P}(k) \), and denote by \( m_k \) the number of \( 1 \leq l \leq k \) such that \( i_l = i_k \). By definition, we have \( 1 \leq m_k \leq d_i \).

First, we consider the case \( P \subset c + \mathbb{R}^{d_i} \) for some \( c \in (\mathbb{R}^{d_i})^\perp \). Write

\[ P = c + \Pi(\mu, \nu) = c + [\mu_1, \nu_1] \times \cdots \times [\mu_{d_i}, \nu_{d_i}], \]

and set

\[ v_{m_k}' := v_{m_k} + l_i(c) - \sum_{1 \leq l \leq d_i} (\mu_l + \nu_l). \]

We define \( v' \in \mathbb{R}^{d_i} \) (resp., \( \mu' \in \mathbb{R}^{d_i} \)) by replacing \( v_{m_k} \) in \( v \) (resp., \( \mu_{m_k} \) in \( \mu \)) by \( v_{m_k}' \). If \( v_{m_k}' \geq v_{m_k} \), then we set

\[ D_i^{(k)}(P) := \mathbb{P}(c + \Pi(\mu, v')). \]

If \( v_{m_k}' < v_{m_k} \), then we set

\[ D_i^{(k)}(P) := -\mathbb{P}(c + \Pi(\mu, v')) + \mathbb{P} + \mathbb{P}', \]
where $P'$ is the facet of $c + \Pi(\mu', \nu)$ parallel to $P$.

In general, we define $D_i^{(k)}(P) \in F(\mathbb{R}^N)$ by

$$D_i^{(k)}(P)|_{c+\mathbb{R}^d_i} := D_i^{(k)}(\overline{P \cap (c + \mathbb{R}^d_i)})$$

for $c \in (\mathbb{R}^d_i)\perp$.

**Definition 2.2** Let $1 \leq k \leq N$, $i := i_k$, and $P \in \mathcal{P}_\square(k)$. If the function $D_i^{(k)}(P)$ is identical to the characteristic function $\mathbb{I}_Q$ of a convex polytope $Q$, then by abuse of notation, we write $Q = D_i^{(k)}(P)$.

**Remark 2.3** In the paper [26], she defined convex-geometric Demazure operators for convex parachains. Even for parapolytopes, our definition of convex-geometric Demazure operators is slightly different from hers since we specify which direction we expand in.

See [26, Sect. 2.4] for examples of functions constructed by convex-geometric Demazure operators. Our purpose is to compute $D_{i_N}^{(N)} \cdots D_{i_1}^{(1)}(a)$ for specific $a \in \mathbb{R}^N$. Note that $D_{i_N}^{(N)} \cdots D_{i_1}^{(1)}(a)$ is not necessarily well-defined as the following example shows.

**Example 2.4** Let $G = SL_4(\mathbb{C})$, and $i = (2, 1, 2, 3, 2, 1) \in I^6$, which is a reduced word for $w_0$. Then, the functions $l_i$, $i \in I$, are given by

$$l_1(a) = l_3(a) = a_1^{(2)} + a_2^{(2)} + a_3^{(2)} \text{ and } l_2(a) = a_1^{(1)} + a_2^{(1)} + a_1^{(3)}$$

for $a = (a_1^{(1)}, a_2^{(1)}, a_1^{(2)}, a_2^{(2)}, a_3^{(2)}, a_1^{(3)}) \in \mathbb{R}^6 = \mathbb{R}^2 \oplus \mathbb{R}^3 \oplus \mathbb{R}$. If we set

$$a_{\text{low}} := -\left(\frac{5}{4}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{3}{2}\right) \in \mathbb{R}^2 \oplus \mathbb{R}^3 \oplus \mathbb{R},$$

then we have $D_2^{(1)}(a_{\text{low}}), D_2^{(1)}D_2^{(1)}(a_{\text{low}}), D_3^{(1)}D_2^{(1)}(a_{\text{low}}) \in \mathcal{P}_\square$ and $D_3^{(4)}D_2^{(3)}D_1^{(2)}D_2^{(1)}(a_{\text{low}}) \in \mathcal{P}_6$. In addition, the polytope $D_3^{(4)}D_2^{(3)}D_1^{(2)}D_2^{(1)}(a_{\text{low}})$ is given by the following conditions:

$$(a_2^{(1)}, a_3^{(2)}) = \left(-\frac{1}{4}, -\frac{4}{3}\right), \quad -\frac{1}{3} \leq a_2^{(2)} \leq \frac{2}{3}, \quad -\frac{5}{4} \leq a_1^{(1)} \leq a_1^{(2)} + \frac{1}{12},$$

$$-\frac{1}{3} \leq a_2^{(2)} \leq \min\left\{a_1^{(1)} + \frac{11}{12}, \frac{2}{3}\right\}, \quad -\frac{3}{2} \leq a_3^{(3)} \leq a_1^{(2)} + a_2^{(2)} + \frac{1}{6}.$$}

Hence for $c := (-\frac{1}{4}, -\frac{1}{3}, 0, 0, 0,\frac{1}{3}) \in (\mathbb{R}^d_4)\perp$, the intersection $D_3^{(4)}D_2^{(3)}D_1^{(2)}D_2^{(1)}(a_{\text{low}}) \cap (c + \mathbb{R}^d_2)$ is identified with the set of $(a_1^{(2)}, a_2^{(2)}, a_3^{(2)}) \in \mathbb{R}^3$ satisfying the following conditions:

$$-\frac{1}{3} \leq a_1^{(2)} \leq \frac{2}{3}, \quad -a_2^{(2)} + \frac{1}{3} \leq a_2^{(2)} \leq \frac{2}{3}, \quad a_3^{(2)} = -\frac{4}{3}.$$
Since this is not of the form \( \Pi(\mu, \nu) \), we deduce that \( D_3^{(4)} D_2^{(3)} D_1^{(2)} D_2^{(1)}(a_{\text{low}}) \) is not a parapolytope, and hence that \( D_2^{(5)} D_3^{(4)} D_2^{(3)} D_1^{(2)} D_2^{(1)}(a_{\text{low}}) \) is not well-defined.

## 3 Polyhedral realizations of crystal bases

In this section, we review some fundamental properties of polyhedral realizations of crystal bases, following \([10, 34, 37]\). We start with recalling the definition of abstract crystals, introduced in \([21]\). Choose a Borel subgroup \( B \subset G \) and a maximal torus \( T \subset B \). Denote by \( t \) the Lie algebra of \( T \), by \( t^* := \text{Hom}_\mathbb{C}(t, \mathbb{C}) \) its dual space, and by \( \langle \cdot, \cdot \rangle : t^* \times t \to \mathbb{C} \) the canonical pairing. Let \( \{\alpha_i \mid i \in I\} \subset t^* \) be the set of simple roots, \( \{h_i \mid i \in I\} \subset t \) the set of simple coroots, and \( P \subset t^* \) the weight lattice.

### Definition 3.1 \([21, \text{Definition 1.2.1}]\)

A crystal \( B \) is a set equipped with maps

\[
\begin{align*}
\varepsilon_i : B &\to \mathbb{Z} \cup \{-\infty\}, \quad \varphi_i : B \to \mathbb{Z} \cup \{-\infty, \infty\} \text{ for } i \in I, \\
\tilde{\varepsilon}_i : B &\to B \cup \{0\}, \quad \tilde{f}_i : B \to B \cup \{0\} \text{ for } i \in I,
\end{align*}
\]

satisfying the following conditions:

(i) \( \varphi_i(b) = \varepsilon_i(b) + \langle \text{wt}(b), h_i \rangle \) for \( i \in I \),

(ii) \( \text{wt}(\tilde{\varepsilon}_i b) = \text{wt}(b) + \alpha_i, \quad \varepsilon_i(\tilde{\varepsilon}_i b) = \varepsilon_i(b) - 1, \) and \( \varphi_i(\tilde{\varepsilon}_i b) = \varphi_i(b) + 1 \) for \( i \in I \) and \( b \in B \) such that \( \tilde{\varepsilon}_i b \in B \),

(iii) \( \text{wt}(\tilde{f}_i b) = \text{wt}(b) - \alpha_i, \quad \varepsilon_i(\tilde{f}_i b) = \varepsilon_i(b) + 1, \) and \( \varphi_i(\tilde{f}_i b) = \varphi_i(b) - 1 \) for \( i \in I \) and \( b \in B \) such that \( \tilde{f}_i b \in B \),

(iv) \( b' = \tilde{\varepsilon}_i b \) if and only if \( b = \tilde{f}_i b' \) for \( i \in I \) and \( b, b' \in B \),

(v) \( \tilde{\varepsilon}_i b = \tilde{f}_i b = 0 \) for \( i \in I \) and \( b \in B \) such that \( \varphi_i(b) = -\infty \).

Here, \(-\infty\) and 0 are additional elements that are not contained in \( \mathbb{Z} \) and \( B \), respectively.

The maps \( \tilde{\varepsilon}_i \) and \( \tilde{f}_i \) are called the Kashiwara operators.

### Example 3.2

For \( \lambda \in P \), let \( R_\lambda = \{r_\lambda\} \) be a crystal consisting of only one element, given by: \( \text{wt}(r_\lambda) = \lambda, \varepsilon_i(r_\lambda) = -\langle \lambda, h_i \rangle, \varphi_i(r_\lambda) = 0, \) and \( \tilde{\varepsilon}_i r_\lambda = \tilde{f}_i r_\lambda = 0 \).

### Example 3.3

For \( i \in I \), we define a crystal \( B_i := \{x_i \mid x \in \mathbb{Z}\} \) as follows:

\[
\begin{align*}
\text{wt}(x_i) &:= -x_i, \quad \varepsilon_i(x_i) := x, \quad \varphi_i(x_i) := -x, \quad \tilde{\varepsilon}_i(x_i) := (x - 1)_i, \\
\tilde{f}_i(x_i) &:= (x + 1)_i, \quad \text{and} \\
\varepsilon_j(x_i) &:= \varphi_j(x_i) := -\infty, \quad \tilde{\varepsilon}_j(x_i) = \tilde{f}_j(x_i) := 0 \text{ for } j \neq i.
\end{align*}
\]

### Definition 3.4 \([21, \text{Sect. 1.2}]\)

Let \( B_1, B_2 \) be two crystals. A map

\[
\psi : B_1 \cup \{0\} \to B_2 \cup \{0\}
\]

is called a strict morphism of crystals from \( B_1 \) to \( B_2 \) if it satisfies the following conditions:
(i) \( \psi(0) = 0 \),
(ii) \( \wt(\psi(b)) = \wt(b), \varepsilon_i(\psi(b)) = \varepsilon_i(b), \) and \( \varphi_i(\psi(b)) = \varphi_i(b) \) for \( i \in I \) and \( b \in B_1 \) such that \( \psi(b) \in B_2 \),
(iii) \( \tilde{e}_i\psi(b) = \psi(\tilde{e}_ib) \) and \( \tilde{f}_i\psi(b) = \psi(\tilde{f}_ib) \) for \( i \in I \) and \( b \in B_1 \);

here, if \( \psi(b) = 0 \), then we set \( \tilde{e}_i\psi(b) = \tilde{f}_i\psi(b) = 0 \). An injective strict morphism is called a strict embedding of crystals.

Consider the total order \( < \) on \( \mathbb{Z} \cup \{-\infty\} \) given by the usual order on \( \mathbb{Z} \), and by \( -\infty < s \) for all \( s \in \mathbb{Z} \). For two crystals \( B_1, B_2 \), we can define another crystal \( B_1 \otimes B_2 \), called the tensor product of \( B_1 \) and \( B_2 \), as follows (see [21, Sect. 1.3]):

\[
B_1 \otimes B_2 := \{ b_1 \otimes b_2 \ | \ b_1 \in B_1, \ b_2 \in B_2 \},
\]
\[
\wt(b_1 \otimes b_2) := \wt(b_1) + \wt(b_2),
\]
\[
\varepsilon_i(b_1 \otimes b_2) := \max\{\varepsilon_i(b_1), \varepsilon_i(b_2) - (\wt(b_1), h_i)\},
\]
\[
\varphi_i(b_1 \otimes b_2) := \max\{\varphi_i(b_2), \varphi_i(b_1) + (\wt(b_2), h_i)\},
\]
\[
\tilde{e}_i(b_1 \otimes b_2) := \begin{cases} 
\tilde{e}_ib_1 \otimes b_2 & \text{if } \varphi_i(b_1) \geq \varepsilon_i(b_2), \\
\tilde{e}_ib_1 \otimes \tilde{e}_ib_2 & \text{if } \varphi_i(b_1) < \varepsilon_i(b_2),
\end{cases}
\]
\[
\tilde{f}_i(b_1 \otimes b_2) := \begin{cases} 
\tilde{f}_ib_1 \otimes b_2 & \text{if } \varphi_i(b_1) > \varepsilon_i(b_2), \\
\tilde{f}_ib_1 \otimes \tilde{f}_ib_2 & \text{if } \varphi_i(b_1) \leq \varepsilon_i(b_2);
\end{cases}
\]

here, \( b_1 \otimes b_2 \) stands for an ordered pair \( (b_1, b_2) \), and we set \( b_1 \otimes 0 = 0 \otimes b_2 = 0 \).

Let \( P_+ \subset P \) be the set of dominant integral weights, \( B^- \subset G \) the Borel subgroup opposite to \( B \), and \( e_i, f_i, h_i \in \mathfrak{g} \), \( i \in I \), the Chevalley generators such that \( \{e_i, h_i \ | \ i \in I\} \subset \mathrm{Lie}(B) \) and \( \{f_i, h_i \ | \ i \in I\} \subset \mathrm{Lie}(B^-) \). For \( \lambda \in P_+ \), we denote by \( V(\lambda) \) the irreducible highest weight \( G \)-module over \( \mathbb{C} \) with highest weight \( \lambda \) and with highest weight vector \( v_\lambda \). Lusztig [31–33] and Kashiwara [18–20] constructed a specific \( \mathbb{C} \)-basis of \( V(\lambda) \) via the quantized enveloping algebra associated with \( \mathfrak{g} \). This is called (the specialization at \( q = 1 \) of) the lower global basis (= the canonical basis), and denoted by \( \{G^\text{low}_\lambda(b) \ | \ b \in B(\lambda)\} \subset V(\lambda) \). The index set \( B(\lambda) \) has a crystal structure, which satisfies the following conditions:

\[
\wt(b_\lambda) = \lambda,
\]
\[
\varepsilon_i(b) = \max\{k \in \mathbb{Z}_{\geq 0} \ | \ \tilde{e}_i^k b \neq 0\},
\]
\[
\varphi_i(b) = \max\{k \in \mathbb{Z}_{\geq 0} \ | \ \tilde{f}_i^k b \neq 0\},
\]
\[
e_i \cdot G^\text{low}_\lambda(b) \in \mathbb{C}^\times \mathbb{G}^\text{low}_\lambda(\tilde{e}_ib) + \sum_{b' \in B(\lambda); \ wt(b') = \wt(b) + \alpha_i; \ \varphi_i(b') > \varphi_i(b) + 1} \mathbb{C}G^\text{low}_\lambda(b'),
\]
\[
f_i \cdot G^\text{low}_\lambda(b) \in \mathbb{C}^\times \mathbb{G}^\text{low}_\lambda(\tilde{f}_ib) + \sum_{b' \in B(\lambda); \ wt(b') = \wt(b) - \alpha_i; \ \varepsilon_i(b') > \varepsilon_i(b) + 1} \mathbb{C}G^\text{low}_\lambda(b')
\]
for $i \in I$ and $b \in B(\lambda)$, where $\mathbb{C}^\times := \mathbb{C} \setminus \{0\}$, $G^\text{low}_\lambda(0) := 0$ if $\tilde{e}_i b = 0$ or $\tilde{f}_i b = 0$, and $b_\lambda \in B(\lambda)$ is given by $G^\text{low}_\lambda(b_\lambda) \in \mathbb{C}^\times v_\lambda$. We call $B(\lambda)$ the crystal basis for $V(\lambda)$; see [22] for a survey on lower global bases and crystal bases.

**Example 3.5** Let $G = SL_3(\mathbb{C})$, and $\lambda = \varpi_1 + \varpi_2 \in P_+$, where $\varpi_1, \varpi_2 \in P_+$ denote the fundamental weights. Then, the crystal basis $B(\lambda)$ is given as follows, where $b \rightarrow b'$ if and only if $b' = \tilde{f}_i b$:

In addition, the definition of the tensor product of crystals implies that $B(\varpi_1) \otimes B(\varpi_2)$ is obtained from the crystal bases $B(\varpi_1)$ and $B(\varpi_2)$ as follows:

By this graph, we know the irreducible decomposition of the tensor product $G$-module $V(\varpi_1) \otimes V(\varpi_2)$. Indeed, since the graph of $B(\varpi_1) \otimes B(\varpi_2)$ is the disjoint union of those of $B(\lambda)$ and $B(0)$, we see that

$$V(\varpi_1) \otimes V(\varpi_2) \simeq V(\lambda) \oplus V(0)$$

as $G$-modules.

Fix a reduced word $i = (i_1, \ldots, i_N) \in I^N$ for the longest element $w_0 \in W$, and consider a sequence $j = (\ldots, j_k, \ldots, j_{N+1}, j_N, \ldots, j_1)$ of elements in $I$ such that $j_k = i_{N-k+1}$ for $1 \leq k \leq N$, $j_k \neq j_{k+1}$ for all $k \geq 1$, and the cardinality of $\{k \geq 1 \mid j_k = i\}$ is $\infty$ for every $i \in I$. Following [21] and [37], we associate to $j$ a crystal structure on

$$\mathbb{Z}^\infty := \{(\ldots, a_k, \ldots, a_2, a_1) \mid a_k \in \mathbb{Z} \text{ for } k \geq 1 \text{ and } a_k = 0 \text{ for } k \gg 0\}$$
as follows. For \( k \geq 1, i \in I, \) and \( a = (\ldots, a_1, \ldots, a_2, a_1) \in \mathbb{Z}^\infty, \) we set
\[
\sigma_k(a) := a_k + \sum_{l > k} c_{j_k, j_l} a_l \in \mathbb{Z},
\]
\[
\sigma^{(i)}(a) := \max \{\sigma_k(a) \mid k \geq 1, \ j_k = i\} \in \mathbb{Z}, \text{ and}
\]
\[
M^{(i)}(a) := \{k \geq 1 \mid j_k = i, \ \sigma_k(a) = \sigma^{(i)}(a)\}.
\]
Since \( a_l = 0 \) for \( l \gg 0, \) the integers \( \sigma_k(a), \sigma^{(i)}(a) \) are well-defined; also, we have \( \sigma^{(i)}(a) \geq 0. \) Moreover, \( M^{(i)}(a) \) is a finite set if and only if \( \sigma^{(i)}(a) > 0. \) Define a crystal structure on \( \mathbb{Z}^\infty \) by
\[
\text{wt}(a) := - \sum_{k=1}^\infty a_k \alpha_j, \quad \epsilon_i(a) := \sigma^{(i)}(a), \quad \varphi_i(a) := \epsilon_i(a) + \langle \text{wt}(a), h_i \rangle,
\]
\[
\tilde{e}_i a := \begin{cases} (a_k - \delta_{k, \max M^{(i)}(a)})_{k \geq 1} & \text{if } \sigma^{(i)}(a) > 0, \\ 0 & \text{otherwise}, \end{cases}
\]
\[
\tilde{f}_i a := (a_k + \delta_{k, \min M^{(i)}(a)})_{k \geq 1}
\]
for \( i \in I \) and \( a = (\ldots, a_1, \ldots, a_2, a_1) \in \mathbb{Z}^\infty, \) where \( \delta_{k, l} \) is the Kronecker delta; we denote this crystal by \( \mathbb{Z}_j^\infty. \) For \( k \geq 1, \) we set \( j_{\geq k} := (\ldots, j_1, \ldots, j_{k+1}, j_k). \) Then, we see that the crystal \( \mathbb{Z}_j^\infty \) is naturally isomorphic to the tensor product \( \mathbb{Z}_{j_{\geq k}}^\infty \otimes \tilde{\mathbb{B}}_{j_{k-1}} \otimes \cdots \otimes \tilde{\mathbb{B}}_{j_1} \) for all \( k \geq 2. \)

**Proposition 3.6** (see [34, Theorem 3.2] and [35, Proposition 3.1]) For \( \lambda \in P_+, \) the following hold.

1. There exists a unique strict embedding of crystals
\[
\tilde{\Psi}_j : \mathcal{B}(\lambda) \hookrightarrow \mathbb{Z}_j^\infty \otimes R_\lambda
\]
such that \( \tilde{\Psi}_j(b_\lambda) = (\ldots, 0, \ldots, 0, 0) \otimes r_\lambda. \)

2. If \( (\ldots, a_k, \ldots, a_2, a_1) \otimes r_\lambda \in \tilde{\Psi}_j(\mathcal{B}(\lambda)), \) then \( a_k = 0 \) for all \( k > N. \)

The embedding \( \tilde{\Psi}_j \) (resp., the image \( \tilde{\Psi}_j(\mathcal{B}(\lambda)) \)) is called the Kashiwara embedding (resp., the polyhedral realization) of \( \mathcal{B}(\lambda) \) with respect to \( j. \)

**Remark 3.7** We may regard Proposition 3.6 (1) as a definition of the crystal \( \mathcal{B}(\lambda), \) that is, \( \mathcal{B}(\lambda) \) is identified with
\[
\{\tilde{f}_{k_1} \cdots \tilde{f}_{k_l} (((\ldots, 0, 0) \otimes r_\lambda) \mid l \geq 0, \ k_1, \ldots, k_l \in I \}\setminus \{0\} \subset \mathbb{Z}_j^\infty \otimes R_\lambda
\]
as a set, and its crystal structure is given by that on \( \mathbb{Z}_j^\infty \otimes R_\lambda. \)

**Definition 3.8** We define \( \Psi_j : \mathcal{B}(\lambda) \hookrightarrow \mathbb{Z}^N, \ b \mapsto (a_1, a_2, \ldots, a_N), \) by
\[
\tilde{\Psi}_j(b) = (\ldots, 0, 0, a_1, a_2, \ldots, a_N) \otimes r_\lambda; 
\]
this is also called the *Kashiwara embedding* of $B(\lambda)$ with respect to $i$.

Note that the embedding $\Psi_i$ is independent of the choice of an extension $j$ by [37, Sect. 2.4].

**Definition 3.9** (See [10, Definition 2.15]) Let $i \in I^N$ be a reduced word for $w_0$, and $\lambda \in P_+$. Define a subset $S_i(\lambda) \subset \mathbb{Z}_{>0} \times \mathbb{Z}^N$ by

$$S_i(\lambda) := \bigcup_{k>0} \{(k, \Psi_i(b)) \mid b \in B(k\lambda)\},$$

and denote by $C_i(\lambda) \subset \mathbb{R}_{\geq 0} \times \mathbb{R}^N$ the smallest real closed cone containing $S_i(\lambda)$. Now let us define a subset $\Delta_i(\lambda) \subset \mathbb{R}^N$ by

$$\Delta_i(\lambda) := \{a \in \mathbb{R}^N \mid (1, a) \in C_i(\lambda)\}.$$

The set $\Delta_i(\lambda)$ is called the *Nakashima–Zelevinsky polytope* associated with $i$ and $\lambda$.

**Proposition 3.10** [10, Corollaries 2.18 (2), 2.20, and 4.3] Let $i \in I^N$ be a reduced word for $w_0$, and $\lambda \in P_+$.

1. The real closed cone $C_i(\lambda)$ is a rational convex polyhedral cone, and the equality $S_i(\lambda) = C_i(\lambda) \cap (\mathbb{Z}_{>0} \times \mathbb{Z}^N)$ holds.
2. The Nakashima–Zelevinsky polytope $\Delta_i(\lambda)$ is a rational convex polytope, and the equality $\Delta_i(\lambda) \cap \mathbb{Z}^N = \Psi_i(B(\lambda))$ holds.

**Remark 3.11** In the case that $(j, \lambda)$ is ample (see [34, Sect. 4.2] for the definition), a system of explicit linear inequalities defining $\Delta_i(\lambda)$ is given by [34, Theorem 4.1] (see also [10, Corollary 5.3]). Note that in order to prove Proposition 3.10, the ampleness of $(j, \lambda)$ is not necessary.

**Example 3.12** [34] Let $G = SL_{n+1}(\mathbb{C})$, and $\lambda \in P_+$. We consider a specific reduced word $i = (1; 2, 1; 3, 2, 1; \ldots; n, n-1, 1, \ldots, 1)$ for $w_0$. Then, by Nakashima [34, Theorem 6.1] (see also [36, Corollary 2.7]), the Nakashima–Zelevinsky polytope $\Delta_i(\lambda)$ is identical to the set of $(a_1^{(1)}, a_{n-1}^{(2)}, a_{n-1}^{(1)}, \ldots, a_1^{(n)}, \ldots, a_1^{(1)}) \in \mathbb{R}^N$ satisfying the following conditions:

$$
\begin{array}{ccccccc}
\lambda_{\geq 1} & \lambda_{\geq 2} & \cdots & \lambda_{\geq n} & 0 \\
 a_1^{(1)} & a_2^{(1)} & \cdots & a_n^{(1)} \\
 & a_1^{(2)} & \cdots & a_{n-1}^{(2)} \\
 & & \cdots & a_{n-1}^{(n-1)} \\
 & & & a_1^{(n)}
\end{array}
$$

where $N := \frac{n(n+1)}{2}$, $\lambda_{\geq k} := \sum_{k \leq l \leq n} \langle \lambda, h_l \rangle$ for $1 \leq k \leq n$, and the notation

$$
\begin{array}{ccccccc}
a & c & \\
b
\end{array}$$
means that $a \geq b \geq c$. This implies that the translation

$$\Delta_1(\lambda) + (0, 0, \lambda_{\geq n}, 0, \lambda_{\geq n}, \lambda_{\geq n-1}, \ldots, 0, \lambda_{\geq n}, \lambda_{\geq n-1}, \ldots, \lambda_{\geq 2})$$

of the Nakashima–Zelevinsky polytope is identical to the Gelfand–Zetlin polytope $GZ(\bar{\lambda})$ associated with the non-increasing sequence $\bar{\lambda} := (\lambda_{\geq 1}, \lambda_{\geq 2}, \ldots, \lambda_{\geq n}, 0)$.

**Example 3.13** Let $G = SL_3(\mathbb{C})$, $i = (1, 2, 1)$, and $\lambda = \varpi_1 + \varpi_2 \in P_+$. Then, the Nakashima–Zelevinsky polytope $\Delta_1(\lambda)$ is given by

$$\Delta_1(\lambda) = \{(a_1, a_2, a_3) \in \mathbb{R}^3 | 0 \leq a_1 \leq 1, 0 \leq a_3 \leq 1, a_1 \leq a_2 \leq a_3 + 1\};$$

see Fig. 1.

We take an extension $j = (\ldots, j_k, \ldots, j_{N+1}, j_N, \ldots, j_1)$ of $i$, which satisfies the conditions above. Then, by using the crystal structure on $\mathbb{Z}_\infty^N \otimes \mathbb{R}_\lambda$, we can compute the graph of $B(\lambda)$ in Example 3.5 from the set $\Delta_1(\lambda) \cap \mathbb{Z}^3$ of lattice points as follows:

For $w \in W$ and $\lambda \in P_+$, let $v_{w\lambda} \in V(\lambda)$ be a weight vector of weight $w\lambda$, which is called an extremal weight vector. We define a $B$-submodule $V_w(\lambda) \subset V(\lambda)$ (resp., a $B^-$-submodule $V^w(\lambda) \subset V(\lambda)$) by
\[ V_w(\lambda) := \sum_{b \in B} C b v_{w\lambda}, \]

(resp., \( V^w(\lambda) := \sum_{b \in B^-} C b v_{w\lambda} \));

this is called the Demazure module (resp., the opposite Demazure module) associated with \( w \in W \). By [21, Proposition 3.2.3 (i) and equation (4.1)], there uniquely exists a subset \( B_w(\lambda) \) (resp., \( B^w(\lambda) \)) of \( B(\lambda) \) such that

\[ V_w(\lambda) = \sum_{b \in B_w(\lambda)} C G_{\lambda}^{\text{low}}(b) \]

(resp., \( V^w(\lambda) = \sum_{b \in B^w(\lambda)} C G_{\lambda}^{\text{low}}(b) \));

this subset \( B_w(\lambda) \) (resp., \( B^w(\lambda) \)) is called a Demazure crystal (resp., an opposite Demazure crystal). Let \( b_{w\lambda} \in B(\lambda) \) denote the extremal weight element of weight \( w\lambda \), that is, \( b_{w\lambda} \) is a unique element in \( B(\lambda) \) such that \( G_{\lambda}^{\text{low}}(b_{w\lambda}) \in \mathbb{C} \times v_{w\lambda} \). Then, we have

\[ B_w(\lambda) \cap B^w(\lambda) = \{ b_{w\lambda} \}. \]

Let \( \{ s_i \mid i \in I \} \subset W \) be the set of simple reflections. The following is a collection of fundamental properties of Demazure crystals and opposite Demazure crystals.

**Proposition 3.14** [21, Propositions 3.2.3 (ii), (iii) and 4.2] Let \( w \in W \), and \( \lambda \in P_+ \).

1. \( \tilde{e}_i B_w(\lambda) \subset B_w(\lambda) \cup \{ 0 \} \) and \( \tilde{f}_i B^w(\lambda) \subset B^w(\lambda) \cup \{ 0 \} \) for all \( i \in I \).
2. If \( s_i w < w \), then

\[ B_w(\lambda) = \bigcup_{k \geq 0} \tilde{f}_i B_{s_i w}(\lambda) \setminus \{ 0 \}. \]

\[ B^{s_i w}(\lambda) = \bigcup_{k \geq 0} \tilde{e}_i B^w(\lambda) \setminus \{ 0 \}. \]

3. Let \( i = (i_1, \ldots, i_r) \in I^r \) be a reduced word for \( w \in W \). Then,

\[ B_w(\lambda) = \{ \tilde{f}_{i_1}^{a_1} \cdots \tilde{f}_{i_r}^{a_r} b_{\lambda} \mid a_1, \ldots, a_r \in \mathbb{Z}_{\geq 0} \} \setminus \{ 0 \}. \]

4. Let \( i = (i_1, \ldots, i_r) \in I^r \) be a reduced word for \( w w_0 \in W \). Then,

\[ B^w(\lambda) = \{ \tilde{e}_{i_1}^{a_1} \cdots \tilde{e}_{i_r}^{a_r} b_{w_0 \lambda} \mid a_1, \ldots, a_r \in \mathbb{Z}_{\geq 0} \} \setminus \{ 0 \}. \]

For \( \lambda \in P_+ \), the crystal \( B(-w_0 \lambda) \) is identified with the dual crystal of \( B(\lambda) \) (see [21, Sect. 1.2] for more details). Under this identification, the opposite Demazure crystals of \( B(\lambda) \) correspond to the Demazure crystals of \( B(-w_0 \lambda) \). For \( i \in I \), a subset
Lemma 3.17

The following equalities hold for there exists \( b_{S}^{\text{high}} \in S \) such that \( \tilde{e}_{i} b_{S}^{\text{high}} = 0 \), and such that

\[
S = \{ f_{i}^{k} b_{S}^{\text{high}} \mid k \in \mathbb{Z}_{\geq 0}\} \setminus \{0\}.
\]

This element \( b_{S}^{\text{high}} \) is called the highest weight element of \( S \); similarly, the lowest weight element \( b_{S}^{\text{low}} \in S \) is defined by \( \tilde{f}_{i} b_{S}^{\text{low}} = 0 \). The following is called the string property of Demazure crystals and opposite Demazure crystals.

**Proposition 3.15** (See [21, Proposition 3.3.5]) Let \( w \in W, \lambda \in P_{+}, \) and \( i \in I. \)

1. For each i-string \( S \) of \( B(\lambda) \) with highest weight element \( b_{S}^{\text{high}} \), the intersection \( B_{w}(\lambda) \cap S \) is either \( \emptyset \), \( S \), or \( \{ b_{S}^{\text{high}} \} \).
2. For each i-string \( S \) of \( B(\lambda) \) with lowest weight element \( b_{S}^{\text{low}} \), the intersection \( B_{w}(\lambda) \cap S \) is either \( \emptyset \), \( S \), or \( \{ b_{S}^{\text{low}} \} \).

Let \( i = (i_{1}, \ldots, i_{N}) \in I_{N}^{N} \) be a reduced word for \( w_{0} \), and \( \lambda \in P_{+} \). We write \( w_{\geq k} := s_{i_{k}} \cdots s_{i_{N}} \in W \) and \( x_{k} := -\langle w_{k} \lambda, h_{i_{k}} \rangle \) for \( 1 \leq k \leq N \).

**Theorem 3.16** [35, Theorem 4.1] Let \( i = (i_{1}, \ldots, i_{N}) \in I_{N}^{N} \) be a reduced word for \( w_{0} \), \( \lambda \in P_{+}, \) and \( 1 \leq k \leq N \). Then, the image \( \Psi_{i}(b_{w_{\geq k}}) \) is given by

\[
\Psi_{i}(b_{w_{\geq k}}) = (0, \ldots, 0, x_{k}, \ldots, x_{N}).
\]

For \( 1 \leq k \leq N \), we define

\[
\pi_{\geq k} : B(\lambda) \rightarrow \widetilde{B}_{i_{k}} \otimes \cdots \otimes \widetilde{B}_{i_{N}} \otimes R_{\lambda} \text{ and}
\]

\[
\pi_{\leq k} : B(\lambda) \rightarrow Z_{\geq N+1}^{\infty} \otimes \widetilde{B}_{i_{1}} \otimes \cdots \otimes \widetilde{B}_{i_{k}}
\]

by \( \pi_{\geq k}(b) := b_{2} \) and \( \pi_{\leq k}(b') := b'_{1} \) for \( b, b' \in B(\lambda) \) such that

\[
\widetilde{\Psi}_{j}(b) = b_{1} \otimes b_{2} \in (Z_{\geq N+1}^{\infty} \otimes \widetilde{B}_{i_{1}} \otimes \cdots \otimes \widetilde{B}_{i_{k-1}}) \otimes (\widetilde{B}_{i_{k}} \otimes \cdots \otimes \widetilde{B}_{i_{N}} \otimes R_{\lambda}) \text{ and}
\]

\[
\widetilde{\Psi}_{j}(b') = b'_{1} \otimes b'_{2} \in (Z_{\geq N+1}^{\infty} \otimes \widetilde{B}_{i_{1}} \otimes \cdots \otimes \widetilde{B}_{i_{k}}) \otimes (\widetilde{B}_{i_{k+1}} \otimes \cdots \otimes \widetilde{B}_{i_{N}} \otimes R_{\lambda}),
\]

respectively. In addition, we set \( \pi_{\geq 0} = \pi_{\leq N+1} = \widetilde{\Psi}_{j} \), and

\[
\pi_{\geq N+1} : B(\lambda) \rightarrow R_{\lambda}, \ b \mapsto r_{\lambda},
\]

\[
\pi_{\leq 0} : B(\lambda) \rightarrow Z_{\geq N+1}^{\infty}, \ b \mapsto (\ldots, 0, \ldots, 0, 0)
\]

We write \( x_{\geq k} := \pi_{\geq k}(b_{w_{\geq k}}) \) for \( 1 \leq k \leq N \).

**Lemma 3.17** The following equalities hold for \( 2 \leq k \leq N \) :

\[
e_{i_{k-1}}(x_{\geq k-1}) = x_{k-1}, \ \tilde{e}_{i_{k-1}}^{x_{k-1}} x_{\geq k-1} = (0)_{i_{k-1}} \otimes x_{\geq k},
\]

\[
\varphi_{i_{k-1}}((0)_{i_{k-1}} \otimes x_{\geq k}) = x_{k-1}, \ f_{i_{k-1}}^{x_{k-1}}((0)_{i_{k-1}} \otimes x_{\geq k}) = x_{\geq k-1}.
\]
Proof Since $e_{i_{k-1}}(b_{w_{\geq k-1}\lambda}) = x_{k-1} = e_{i_{k-1}}((x_{k-1})_{i_{k-1}})$, and
\[ \tilde{\Psi}_j(b_{w_{\geq k}\lambda}) = e_{i_{k-1}}(b_{w_{\geq k-1}\lambda}) \tilde{\Psi}_j(b_{w_{\geq k-1}\lambda}) 
= e_{i_{k-1}}^{-1} \tilde{\Psi}_j(b_{w_{\geq k-1}\lambda}), \]

Theorem 3.16 and the tensor product rule for crystals imply that $e_{i_{k-1}}(x_{\geq k-1}) = x_{k-1}$, $e_{i_{k-1}}^{-1} x_{\geq k-1} = (0)_{i_{k-1}} \otimes x_{\geq k}$. The other assertions of the lemma follow from these and $\varphi_{i_{k-1}}(b_{w_{\geq k}\lambda}) = x_{k-1}$. \hfill $\Box$

Proposition 3.18 The following equality holds for $1 \leq k \leq N$:
\[ \Psi_1(B^{w_{\geq k}}(\lambda)) = \{ a = (a_1, \ldots, a_N) \in \Psi_1(B(\lambda)) \mid a_l = x_l \text{ for all } k \leq l \leq N \}. \]

Proof We will prove that
\[ B^{w_{\geq k}}(\lambda) = \{ b \in B(\lambda) \mid \pi_{\geq k}(b) = x_{\geq k} \} \]
for $1 \leq k \leq N$. We proceed by induction on $k$. If $k = 1$, then the assertion is obvious since $B^{w_{\geq 1}}(\lambda) = B^{w_0}(\lambda) = \{ b_{w_0}\lambda \}$ and $\Psi_1(b_{w_0}\lambda) = (x_1, \ldots, x_N)$ by Theorem 3.16. We assume that $k > 1$, and that
\[ B^{w_{\geq k-1}}(\lambda) = \{ b \in B(\lambda) \mid \pi_{\geq k-1}(b) = x_{\geq k-1} \}. \]

Take $b \in B^{w_{\geq k}}(\lambda)$. Then, we see by Proposition 3.14 that $f_{i_{k-1}}^\varphi(b) \in B^{w_{\geq k-1}}(\lambda)$; hence the equality $\pi_{\geq k-1}(f_{i_{k-1}}^\varphi(b)) = x_{\geq k-1}$ holds. From this and Lemma 3.17, we deduce that
\[ \pi_{\geq k}(b) = \pi_{\geq k}(e_{i_{k-1}}^\varphi(b) f_{i_{k-1}}^\varphi(b)) = x_{\geq k}. \]

Conversely, take $b \in B(\lambda)$ such that $\pi_{\geq k}(b) = x_{\geq k}$. Then, we have $\tilde{\Psi}_j(b) = \pi_{\leq k-2}(b) \otimes (a)_{i_{k-1}} \otimes x_{\geq k}$ for some $0 \leq a \leq x_{k-1}$. By Lemma 3.17, it follows that $\varphi_{i_{k-1}}((a)_{i_{k-1}} \otimes x_{\geq k}) = x_{k-1} - a$, and that $f_{i_{k-1}}^\varphi((a)_{i_{k-1}} \otimes x_{\geq k}) = x_{\geq k-1}$. Hence by the tensor product rule for crystals, we deduce that
\[ f_{i_{k-1}}^\varphi(b) \tilde{\Psi}_j(b) = f_{i_{k-1}}^\varphi(b) \pi_{\leq k-2}(b) \otimes x_{\geq k-1}. \]

From this, it follows that $\pi_{\geq k-1}(f_{i_{k-1}}^\varphi(b)) = x_{\geq k-1}$, and hence that $f_{i_{k-1}}^\varphi(b) \in B^{w_{\geq k-1}}(\lambda)$. By Proposition 3.14, this implies that $b \in B^{w_{\geq k}}(\lambda)$. These prove the proposition. \hfill $\Box$

Corollary 3.19 For all $\lambda, \mu \in P_+$ and $1 \leq k \leq N$, the following holds:
\[ \Psi_1(B^{w_{\geq k}}(\lambda)) + \Psi_1(B^{w_{\geq k}}(\mu)) \subseteq \Psi_1(B^{w_{\geq k}}(\lambda + \mu)). \]
Proof Since $-\langle w_{\geq l}(\lambda + \mu), h_{i_l} \rangle = -\langle w_{\geq l} \lambda, h_{i_l} \rangle - \langle w_{\geq l} \mu, h_{i_l} \rangle$ for $k \leq l \leq N$, Proposition 3.18 implies that it suffices to show that $\Psi_i(B^{w_{\geq k}}(\lambda)) + \Psi_i(B^{w_{\geq k}}(\mu)) \subset \Psi_i(B(\lambda + \mu))$. However, this follows immediately by the additivity of $\Psi_i$ (see [10, Theorem 4.1]).

4 Main result

4.1 Statement of the main result

Let $i = (i_1, \ldots, i_N) \in I^N$ be a reduced word for $w_0$. For $i \in I$, recall that $d_i$ is the number of $1 \leq k \leq N$ such that $i_k = i$. For $\lambda \in P_+$, we write $\lambda = \sum_{i \in I} \hat{\lambda}_i d_i \alpha_i$, and set

$x_\lambda := \Psi_i(b_{w_0 \lambda}) = (x_1, \ldots, x_N),
\quad a_\lambda := -x_\lambda + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}).$

The following is the main result of this paper.

**Theorem 4.1** Let $i = (i_1, \ldots, i_N) \in I^N$ be a reduced word for $w_0$, and $\lambda \in P_+$. Assume that the Nakashima–Zelevinsky polytope $\Delta_i(\lambda)$ is a parapolytope.

1. The polytope $\Delta_i(\lambda)$ is a lattice polytope.
2. The polytope $D_{i_N}^{(N)} \cdots D_{i_1}^{(1)}(a_\lambda)$ is well-defined.
3. The following equality holds:

$$D_{i_N}^{(N)} \cdots D_{i_1}^{(1)}(a_\lambda) = -\Delta_i(\lambda) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}).$$

We prove Theorem 4.1 in the next subsection. In the rest of this subsection, we give some examples of $\Delta_i(\lambda)$ which are parapolytopes.

**Example 4.2** Let $G = SL_{n+1}(\mathbb{C})$, and $i = (1; 2, 1; 3, 2, 1; \ldots; n, n-1, \ldots, 1)$. Then, the Nakashima–Zelevinsky polytope $\Delta_i(\lambda)$ is a parapolytope for all $\lambda \in P_+$ by Example 3.12.

**Example 4.3** [16] Let $G$ be of type $B_n$, $C_n$, or $D_n$. We identify the set of vertices of the Dynkin diagram with $\{1, \ldots, n\}$ as follows:

\[
\begin{align*}
B_n & \quad 1 \quad 2 \quad \ldots \quad n-1 \quad n, \\
C_n & \quad 1 \quad 2 \quad \ldots \quad n-1 \quad n, \\
D_n & \quad 1 \quad 2 \quad \ldots \quad n-2 \quad n-1 \quad n.
\end{align*}
\]
We take a reduced word $i$ for $w_0$ to be

$$i = (n, n - 1, \ldots, 1; n, n - 1, \ldots, 1; \ldots; n, n - 1, \ldots, 1),$$

where $i \in I^2$ if $G$ is of type $B_n$ or $C_n$, and $i \in I^{n(n-1)}$ if $G$ is of type $D_n$.

If $G$ is of type $B_n$, then we see by Hoshino [16, Sect. III.A] that $\Delta_i(\lambda)$ is identical to the set of

$$(a_{n}^{(n)}, a_{n}^{(n-1)}, \ldots, a_{1}^{(1)}, a_{n}^{(n-1)}, \ldots, a_{1}^{(1)}) \in \mathbb{R}_{\geq 0}$$

satisfying the following inequalities:

$$a_{i}^{(i)} \geq a_{j}^{(j-1)} \geq \cdots \geq a_{i}^{(1)} \text{ for } 2 \leq i \leq n - 1,$$

$$a_{j}^{(n)} \geq a_{j+1}^{(n-1)} \geq \cdots \geq a_{j}^{(j)} \text{ for } 1 \leq j \leq n - 1,$$

$$a_{j}^{(n-j+1)} \geq a_{j}^{(n-j+2)} \geq \cdots \geq a_{j}^{(n)} \text{ for } 2 \leq j \leq n,$$

$$\lambda_i \geq a_{j}^{(i-j+1)} - a_{j}^{(i-j)} \text{ for } 1 \leq j \leq i \leq n - 1,$$

$$\lambda_n \geq a_{l}^{(n)} - 2a_{l}^{(n-1)} + 2 \sum_{1 \leq k \leq l-1} (a_{n-\mu_k+1}^{(n)} - a_{\mu_k+k-1}^{(n-\mu_k)}) \text{ for } l \geq 1,$$

$$n \geq \mu_1 > \cdots > \mu_l = 1,$$

$$\lambda_n \geq -a_{l}^{(n)} + 2 \sum_{1 \leq k \leq l} (a_{n-\mu_k+1}^{(n)} - a_{k+k-1}^{(n-\mu_k)}) \text{ for } l \geq 1, \ n \geq \mu_1 > \cdots > \mu_l > 1.$$

If $G$ is of type $C_n$, then it follows by Hoshino [16, Sect. III.B] that $\Delta_i(\lambda)$ is identical to the set of

$$(a_{n}^{(n)}, a_{n}^{(n-1)}, \ldots, a_{1}^{(1)}, a_{n}^{(n-1)}, \ldots, a_{1}^{(1)}) \in \mathbb{R}_{\geq 0}$$

satisfying the following inequalities:

$$a_{i}^{(i)} \geq a_{j}^{(i-j+1)} - a_{j}^{(i-j)} \text{ for } 2 \leq i \leq n - 1,$$

$$2a_{j}^{(n)} \geq a_{j+1}^{(n-1)} \geq \cdots \geq a_{j}^{(j)} \text{ for } 1 \leq j \leq n - 1,$$

$$a_{j}^{(n-j+1)} \geq a_{j}^{(n-j+2)} \geq \cdots \geq a_{j}^{(n)} \geq 2a_{j}^{(n)} \text{ for } 2 \leq j \leq n,$$

$$\lambda_i \geq a_{j}^{(i-j+1)} - a_{j}^{(i-j)} \text{ for } 1 \leq j \leq i \leq n - 1,$$

$$\lambda_n \geq a_{l}^{(n)} - a_{l}^{(n-1)} + \sum_{1 \leq k \leq l-1} (a_{n-\mu_k+1}^{(n)} - a_{\mu_k+k-1}^{(n-\mu_k)}) \text{ for } l \geq 1, \ n \geq \mu_1 > \cdots > \mu_l = 1,$$

$$\lambda_n \geq -a_{l}^{(n)} + \sum_{1 \leq k \leq l} (a_{n-\mu_k+1}^{(n)} - a_{\mu_k+k-1}^{(n-\mu_k)}) \text{ for } l \geq 1, \ n \geq \mu_1 > \cdots > \mu_l > 1.$$
If $G$ is of type $D_n$, then it follows by Hoshino [16, Sect. III.C] that $\Delta_1(\lambda)$ is identical to the set of

$$(a_{n-1}^{(n)}, a_{n-1}^{(n-1)}, \ldots, a_{n-1}^{(1)}, a_1^{(n)}, a_1^{(n-1)}, \ldots, a_1^{(1)}) \in \mathbb{R}_{\geq 0}^{n(n-1)}$$

satisfying the following inequalities:

- $a_1^{(i)} \geq a_2^{(i-1)} \geq \cdots \geq a_j^{(1)}$ for $2 \leq i \leq n - 2$,
- $a_j^{(n-1)} + a_j^{(n)} \geq a_{j+1}^{(n-2)} \geq a_{j+2}^{(n-3)} \geq \cdots \geq a_{n-1}^{(1)}$ for $1 \leq j \leq n - 2$,
- $a_j^{(n-j)} \geq a_j^{(n-j+1)} \geq a_j^{(n-j+2)} \geq \cdots \geq a_j^{(n-2)} \geq a_j^{(n-1)} + a_j^{(n)}$ for $2 \leq j \leq n - 1$,
- $a_1^{(n-1)} \geq a_2^{(n-1)} \geq a_3^{(n-1)} \geq a_4^{(n)} \geq \cdots$,
- $a_1^{(n)} \geq a_2^{(n-1)} \geq a_3^{(n-1)} \geq a_4^{(n-1)} \geq \cdots$,
- $\lambda_i \geq a_j^{(i-j+1)} - a_j^{(i-j)}$ for $1 \leq j \leq i \leq n - 2$,
- $\lambda_{n-1} \geq a_1^{(n-1)} - a_1^{(n-2)}$,
- $\lambda_n \geq a_1^{(n)} - a_1^{(n-2)}$,
- $\lambda_{n-1} \geq \max\{-a_2^{(n-1)}, a_2^{(n)} - a_2^{(n-2)}\} + \sum_{1 \leq k \leq 2l-1} (a_{\mu_k}^{(n-\mu_k)} - a_{\mu_k+k-1}^{(n-\mu_k-1)})$,
- $\lambda_n \geq \max\{-a_2^{(n-1)}, a_2^{(n-1)} - a_2^{(n-2)}\} + \sum_{1 \leq k \leq 2l-1} (a_{\mu_k}^{(n-\mu_k)} - a_{\mu_k+k-1}^{(n-\mu_k-1)})$

for $l \geq 1$, $n - 1 \geq \mu_1 > \cdots > \mu_{2l-1} > 1$,

- $\lambda_{n-1} \geq \max\{-a_2^{(n-1)}, a_2^{(n-1)} - a_2^{(n-2)}\} + \sum_{1 \leq k \leq 2l} (a_{\mu_k}^{(n-\mu_k)} - a_{\mu_k+k-1}^{(n-\mu_k-1)})$,
- $\lambda_n \geq \max\{-a_2^{(n-1)}, a_2^{(n)} - a_2^{(n-2)}\} + \sum_{1 \leq k \leq 2l} (a_{\mu_k}^{(n-\mu_k)} - a_{\mu_k+k-1}^{(n-\mu_k-1)})$

for $l \geq 1$, $n - 1 \geq \mu_1 > \cdots > \mu_{2l} > 1$.

In all cases, the Nakashima–Zelevinsky polytopes $\Delta_1(\lambda), \lambda \in P_+$, are parapolytopes.

**Example 4.4** Let $G$ be of type $G_2$. We set $\mathfrak{i} := (1, 2, 1, 2, 1, 2)$ and $\mathfrak{i}^{\text{op}} := (2, 1, 2, 1, 2, 1)$. By [34, Theorem 5.1], the Nakashima–Zelevinsky polytopes $\Delta_1(\lambda)$ and $\Delta_{\mathfrak{i}^{\text{op}}}(\lambda)$ are parapolytopes for all $\lambda \in P_+$.

### 4.2 Proof of Theorem 4.1

We set $(\mathbb{Z}^{d_i})^\perp := (\mathbb{R}^{d_i})^\perp \cap \mathbb{Z}^N$ for $i \in I$. Since $\Psi_1(B(\lambda)) = \Delta_1(\lambda) \cap \mathbb{Z}^N$ by Proposition 3.10 (2), for $i \in I$ and $\mathbf{c} \in (\mathbb{Z}^{d_i})^\perp$ such that $\Psi_1(B(\lambda)) \cap (\mathbf{c} + \mathbb{Z}^{d_i}) \neq \emptyset$, there uniquely exist $\mu^{(i)}(\mathbf{c}) = (\mu_1^{(i)}(\mathbf{c}), \ldots, \mu_{d_i}^{(i)}(\mathbf{c}))$, $\nu^{(i)}(\mathbf{c}) = (\nu_1^{(i)}(\mathbf{c}), \ldots, \nu_{d_i}^{(i)}(\mathbf{c})) \in \mathbb{Z}^{d_i}$
such that

\[ \Psi_1(B(\lambda)) \cap (c + \mathbb{Z}^{d_i}) = c + \Pi_\mathbb{Z}(\mu^{(i)}(c), v^{(i)}(c)), \]

where we write

\[ \Pi_\mathbb{Z}(\mu^{(i)}(c), v^{(i)}(c)) := \{(a^{(i)}_1, \ldots, a^{(i)}_{d_i}) \in \mathbb{Z}^{d_i} | \mu^{(i)}_l(c) \leq a^{(i)}_l \leq v^{(i)}_l(c), 1 \leq l \leq d_i\}. \]

Note that the subset \((B(\lambda) \cap \Psi^{-1}_1(c + \mathbb{Z}^{d_i})) \cup \{0\}\) of \(B(\lambda) \cup \{0\}\) is stable under \(\tilde{e}_i\) and \(\tilde{f}_i\) by the crystal structure on \(\mathbb{Z}_j^\infty \otimes R_\lambda\). For \(0 \leq k \leq N, i \in I,\) and \(c \in (\mathbb{Z}^{d_i})^\perp\) such that \(\Psi_1(B^{w \geq k+1}(\lambda)) \cap (c + \mathbb{Z}^{d_i}) \neq \emptyset\), Proposition 3.18 implies that there uniquely exist

\[ \mu^{(i,k)}(c) = (\mu^{(i,k)}_1(c), \ldots, \mu^{(i,k)}_{d_i}(c)), \quad v^{(i,k)}(c) = (v^{(i,k)}_1(c), \ldots, v^{(i,k)}_{d_i}(c)) \in \mathbb{Z}^{d_i} \]

such that

\[ \Psi_1(B^{w \geq k+1}(\lambda)) \cap (c + \mathbb{Z}^{d_i}) = c + \Pi_\mathbb{Z}(\mu^{(i,k)}(c), v^{(i,k)}(c)), \]

where we define \(w_{\geq N+1} \in W\) to be the identity element. For \(1 \leq k \leq N\) and \(c = (c_s)_{1 \leq s \leq N; i_s \neq i_k} \in (\mathbb{Z}^{d_i})^\perp\) such that \(\Psi_1(B^{w \geq k}(\lambda)) \cap (c + \mathbb{Z}^{d_i}) \neq \emptyset\), we define \(L_k(c) \in \mathbb{Z}\) by

\[ L_k(c) := -\langle \lambda, h_{i_k} \rangle + \sum_{1 \leq l \leq d_{i_k}} (\mu^{(i_k,k-1)}_l(c) + v^{(i_k,k-1)}_l(c)) + \sum_{1 \leq s \leq N; i_s \neq i_k} c_{i_k,i_s}c_s. \]

**Lemma 4.5** The integer \(L_k(c)\) is nonnegative.

**Proof** Let \(b_{\text{high}} \in B^{w \geq k}(\lambda) \cap \Psi^{-1}_1(c + \mathbb{Z}^{d_i})\) be the unique element such that \(\Psi_1(b_{\text{high}}) = c + \mu^{(i_k,k-1)}(c)\). Then, we see that

\[ \text{wt}(b_{\text{high}}) = \lambda - \sum_{1 \leq l \leq d_{i_k}} \mu^{(i_k,k-1)}_l(c)\alpha_{i_k} - \sum_{1 \leq s \leq N; i_s \neq i_k} c_s\alpha_{i_s}, \]

and hence that

\[ \langle \text{wt}(b_{\text{high}}), h_{i_k} \rangle = \langle \lambda, h_{i_k} \rangle - 2 \sum_{1 \leq l \leq d_{i_k}} \mu^{(i_k,k-1)}_l(c) - \sum_{1 \leq s \leq N; i_s \neq i_k} c_{i_k,i_s}c_s. \]

From this, it follows that

\[ L_k(c) = -\langle \text{wt}(b_{\text{high}}), h_{i_k} \rangle + \sum_{1 \leq l \leq d_{i_k}} (v^{(i_k,k-1)}_l(c) - \mu^{(i_k,k-1)}_l(c)). \]
Since we have \( f_{ik}^{(b_{\text{high}})} \) \( b_{\text{high}} \in \mathcal{B}^{w \geq k}(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}d_{ik}) \) by Proposition 3.14 (1), the equality
\[
\Psi_i(\mathcal{B}^{w \geq k}(\lambda)) \cap (c + \mathbb{Z}d_{ik}) = c + \prod_{\mathbb{Z}}(\mu^{(i_k,k-1)}(c), \nu^{(i_k,k-1)}(c))
\]
implies that \( \varphi_{ik}(b_{\text{high}}) \leq \sum_{1 \leq l \leq d_{ik}} (v_{l}^{(i_k,k-1)}(c) - \mu_{l}^{(i_k,k-1)}(c)) \), and hence that
\[
\langle \text{wt}(b_{\text{high}}), h_{ik} \rangle = \varphi_{ik}(b_{\text{high}}) - \epsilon_{ik}(b_{\text{high}}) \leq \sum_{1 \leq l \leq d_{ik}} (v_{l}^{(i_k,k-1)}(c) - \mu_{l}^{(i_k,k-1)}(c)).
\]
This proves the lemma. \( \square \)

We set
\[
\{ s_{i}^{(k)} \} = \{ 1 \leq s \leq N \mid i_s = i_k \}
\]
for \( 1 \leq k \leq N \), and define \( 1 \leq m_k \leq d_{ik} \) by \( s_{m_k}^{(k)} = k \).

**Lemma 4.6** For \( c \in (\mathbb{Z}d_{ik})^\perp \), it follows that \( \Psi_i(\mathcal{B}^{w \geq k}(\lambda)) \cap (c + \mathbb{Z}d_{ik}) \neq \emptyset \) if and only if \( \Psi_i(\mathcal{B}^{w \geq k+1}(\lambda)) \cap (c + \mathbb{Z}d_{ik}) \neq \emptyset \). In this case, the following equalities hold:
\[
\begin{align*}
\mu_{l}^{(i_k,k)}(c) &= \mu_{l}^{(i_k,k-1)}(c), \quad v_{l}^{(i_k,k)}(c) = v_{l}^{(i_k,k-1)}(c) \quad \text{for} \quad 1 \leq l < m_k, \\
\mu_{m_k}^{(i_k,k)}(c) &= x_k - L_k(c), \quad v_{m_k}^{(i_k,k)}(c) = x_k, \quad \text{and} \\
\mu_{l}^{(i_k,k)}(c) &= v_{l}^{(i_k,k)}(c) = x_{s_{l}^{(k)}} \quad \text{for} \quad m_k < l \leq d_{ik}.
\end{align*}
\]

**Proof** Since
\[
\mathcal{B}^{w \geq k+1}(\lambda) = \bigcup_{a \geq 0} \mathbb{Z}a \mathcal{B}^{w \geq k}(\lambda) \setminus \{0\} \quad \text{(4.1)}
\]
by Proposition 3.14 (2), the first assertion follows immediately by the crystal structure on \( \mathbb{Z}^\infty \otimes R_k \). By Proposition 3.18, we have \( \mu_{l}^{(i_k,k)}(c) = v_{l}^{(i_k,k)}(c) = x_{s_{l}^{(k)}} \) for \( m_k < l \leq d_{ik} \), and
\[
\Psi_i(\mathcal{B}^{w \geq k}(\lambda)) = \{ a \in \Psi_i(\mathcal{B}^{w \geq k+1}(\lambda)) \mid a_k = x_k \}. \quad \text{(4.2)}
\]
By (4.1) and (4.2), there exists \( \tilde{L}_k(c) \in \mathbb{Z} \) such that
\[
\begin{align*}
\mu_{l}^{(i_k,k)}(c) &= \mu_{l}^{(i_k,k-1)}(c), \quad v_{l}^{(i_k,k)}(c) = v_{l}^{(i_k,k-1)}(c) \quad \text{for} \quad 1 \leq l < m_k, \quad \text{and} \\
\mu_{m_k}^{(i_k,k)}(c) &= x_k - \tilde{L}_k(c), \quad v_{m_k}^{(i_k,k)}(c) = x_k.
\end{align*}
\]
Hence for the second assertion of the lemma, it suffices to show that \( \tilde{L}_k(c) = L_k(c) \). For \( i \in I \), let us consider the Demazure operator \( D_i : \mathbb{Z}[P] \rightarrow \mathbb{Z}[P] \) given by
\[
D_i(e^\lambda) := \frac{e^\lambda - e^{s_i(\lambda) + \alpha_i}}{1 - e^{\alpha_i}}
\]
for $\lambda \in P$. For $\lambda \in P$ with $\langle \lambda, h_i \rangle \leq 0$, we have
\[
D_i(e^\lambda) = e^\lambda + e^{\lambda + \alpha_i} + \cdots + e^{\mu_i(\lambda)}.
\]
By the string property of $B^{w \geq k}(\lambda)$ (Proposition 3.15 (2)) and the equality
\[
B^{w \geq k+1}(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}d_{ik}) = \bigcup_{a \geq 0} \varepsilon_k^a (B^{w \geq k}(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}d_{ik})) \setminus \{0\},
\]
we deduce that
\[
\text{ch}(B^{w \geq k+1}(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}d_{ik})) = D_i(\text{ch}(B^{w \geq k}(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}d_{ik}))).
\]  
(4.3)
Set $\Pi_1 := \prod_{\mathbb{Z}}(\mu(i_k, k)(c), v(i_k, k)(c))$ and $\Pi_2 := \prod_{\mathbb{Z}}(\mu(i_{k-1}, k)(c), v(i_{k-1}, k)(c))$, where we define $\mu(i, k)(c)$ by replacing $\mu(i, k-1)(c) = x_k$ in $\mu(i, k-1)(c)$ by $x_k - L_k(c)$. Then, it follows that
\[
D_i \left( \text{ch}(B^{w \geq k}(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}d_k)) \right)
= e^{\lambda - \sum_{1 \leq i \leq s N: i \neq k} c_i \alpha_i} \times \sum_{(a_1(i), \ldots, a_d_k(i)) \in \Pi_2} e^{(a_1(i) + \cdots + a_d_k(i)) \alpha_{ik}},
\]  
(4.4)
see [29, Proposition 6.3]. From the equalities (4.3) and (4.4), we see that
\[
\sum_{(a_1(i), \ldots, a_d_k(i)) \in \Pi_1} e^{(a_1(i) + \cdots + a_d_k(i)) \alpha_{ik}} = \sum_{(a_1(i), \ldots, a_d_k(i)) \in \Pi_2} e^{(a_1(i) + \cdots + a_d_k(i)) \alpha_{ik}}.
\]
By comparing the number of terms, we deduce that $\tilde{L}_k(c) = L_k(c)$. This proves the lemma. \hfill \Box

**Lemma 4.7** Let $i = (i_1, \ldots, i_N) \in I^N$ be a reduced word for $w_0$, and $\lambda_1, \lambda_2 \in P_+$. Assume that the polytopes $\Delta_i(\lambda_1), \Delta_i(\lambda_2)$, and $\Delta_i(\lambda_1 + \lambda_2)$ are all parapolytopes. Then, the following equality holds for all $1 \leq k \leq N + 1$:
\[
\Psi_i(B^{w \geq k}(\lambda_1 + \lambda_2)) = \Psi_i(B^{w \geq k}(\lambda_1)) + \Psi_i(B^{w \geq k}(\lambda_2)).
\]

**Proof** We proceed by induction on $k$. If $k = 1$, then the assertion is obvious since
\[
\Psi_i(b_{w_0(\lambda_1 + \lambda_2)}) = \Psi_i(b_{w_0 \lambda_1}) + \Psi_i(b_{w_0 \lambda_2})
\]
by Theorem 3.16. Let $1 \leq k \leq N$, and assume that
\[
\Psi_i(B^{w \geq k}(\lambda_1 + \lambda_2)) = \Psi_i(B^{w \geq k}(\lambda_1)) + \Psi_i(B^{w \geq k}(\lambda_2)).
\]  
(4.5)
By Corollary 3.19, for the inductive step, it suffices to prove that

\[ \Psi_1(\mathcal{B}^{w\geq k+1}(\lambda_1 + \lambda_2)) \subset \Psi_1(\mathcal{B}^{w\geq k+1}(\lambda_1)) + \Psi_1(\mathcal{B}^{w\geq k+1}(\lambda_2)). \]

Fix \( c \in \left( \mathbb{Z}^{d_k} \right) \perp \) such that \( \Psi_1(\mathcal{B}^{w\geq k}(\lambda_1 + \lambda_2)) \cap (c + \mathbb{Z}^{d_k}) \neq \emptyset \). We denote \( \mu^{(i_k, k-1)}(c), v^{(i_k, k-1)}(c), L_k(c) \) for \( \mathcal{B}^{w\geq k}(\lambda) \) by \( \mu^{(i_k, k-1)}(\lambda, c), v^{(i_k, k-1)}(\lambda, c), L_k(\lambda, c) \), respectively, where \( \lambda = \lambda_1, \lambda_2, \lambda_1 + \lambda_2 \). The equality (4.5) implies that

\[ \Psi_1(\mathcal{B}^{w\geq k}(\lambda_1 + \lambda_2)) \cap (c + \mathbb{Z}^{d_k}) = \bigcup_{c_1, c_2 \in (\mathbb{Z}^{d_k})^\perp; c_1 + c_2 = c} \left( \Psi_1(\mathcal{B}^{w\geq k}(\lambda_1)) \cap (c_1 + \mathbb{Z}^{d_k}) + \Psi_1(\mathcal{B}^{w\geq k}(\lambda_2)) \cap (c_2 + \mathbb{Z}^{d_k}) \right). \]

and hence that \( \Pi_{\mathbb{Z}}(\mu^{(i_k, k-1)}(\lambda_1 + \lambda_2, c), v^{(i_k, k-1)}(\lambda_1 + \lambda_2, c)) \) equals

\[ \bigcup_{c_1, c_2 \in (\mathbb{Z}^{d_k})^\perp; c_1 + c_2 = c} \Pi_{\mathbb{Z}}(\mu^{(i_k, k-1)}(\lambda_1, c_1) + \mu^{(i_k, k-1)}(\lambda_2, c_2), v^{(i_k, k-1)}(\lambda_1, c_1) + v^{(i_k, k-1)}(\lambda_2, c_2)). \]

From this, there exist \( c_1, c_2 \in (\mathbb{Z}^{d_k})^\perp \) such that \( c_1 + c_2 = c \), and such that

\[ v^{(i_k, k-1)}(\lambda_1 + \lambda_2, c) = v^{(i_k, k-1)}(\lambda_1, c_1) + v^{(i_k, k-1)}(\lambda_2, c_2), \]

\[ \mu^{(i_k, k-1)}(\lambda_1 + \lambda_2, c) \leq \mu^{(i_k, k-1)}(\lambda_1, c_1) + \mu^{(i_k, k-1)}(\lambda_2, c_2) \]

for all \( 1 \leq l \leq d_k \).

Since

\[ \Psi_1(\mathcal{B}^{w\geq k+1}(\lambda_1)) + \Psi_1(\mathcal{B}^{w\geq k+1}(\lambda_2)) \subset \Psi_1(\mathcal{B}^{w\geq k+1}(\lambda_1 + \lambda_2)) \]

by Corollary 3.19, we have

\[ \Pi_{\mathbb{Z}}(\mu^{(i_k, k)}(\lambda_1, c_1) + \mu^{(i_k, k)}(\lambda_2, c_2), v^{(i_k, k)}(\lambda_1, c_1) + v^{(i_k, k)}(\lambda_2, c_2)) \subset \Pi_{\mathbb{Z}}(\mu^{(i_k, k)}(\lambda_1 + \lambda_2, c), v^{(i_k, k)}(\lambda_1 + \lambda_2, c)). \]

(4.7)

Also, Lemma 4.6 implies that

\[ v^{(i_k, k)}(\lambda_1 + \lambda_2, c) = v^{(i_k, k-1)}(\lambda_1 + \lambda_2, c) = v^{(i_k, k-1)}(\lambda_1, c_1) + v^{(i_k, k-1)}(\lambda_2, c_2) \]

(4.8)

and that
\[ \mu_{m_k}^{(i_k, k)}(\lambda_1 + \lambda_2, c) = -\langle w_{\geq k}(\lambda_1 + \lambda_2), h_{i_k} \rangle - L_k(\lambda_1 + \lambda_2, c) \]
\[ \geq -\langle w_{\geq k}(\lambda_1), h_{i_k} \rangle - (w_{\geq k}\lambda_2, h_{i_k}) - (L_k(\lambda_1, c_1) + L_k(\lambda_2, c_2)) \]
(by (4.6) and the definition of \( L_k(c) \))
\[ = \mu_{m_k}^{(i_k, k)}(\lambda_1, c_1) + \mu_{m_k}^{(i_k, k)}(\lambda_2, c_2). \]

In addition, this inequality becomes the equality if and only if \( \mu_l^{(i_k, k-1)}(\lambda_1 + \lambda_2, c) = \mu_l^{(i_k, k-1)}(\lambda_1, c_1) + \mu_l^{(i_k, k-1)}(\lambda_2, c_2) \) for all \( l \leq d_{i_k} \). However, the inclusion relation (4.7) implies that \( \mu_{m_k}^{(i_k, k)}(\lambda_1, c_1) + \mu_{m_k}^{(i_k, k)}(\lambda_2, c_2) \geq \mu_{m_k}^{(i_k, k)}(\lambda_1 + \lambda_2, c) \), and hence that \( \mu_{m_k}^{(i_k, k)}(\lambda_1, c_1) + \mu_{m_k}^{(i_k, k)}(\lambda_2, c_2) = \mu_{m_k}^{(i_k, k)}(\lambda_1 + \lambda_2, c) \). This proves \( \mu^{(i_k, k)}(\lambda_1 + \lambda_2, c) = \mu^{(i_k, k)}(\lambda_1, c_1) + \mu^{(i_k, k)}(\lambda_2, c_2) \) by Lemma 4.6, which implies by (4.8) that
\[ c + \Pi \mu^{(i_k, k)}(\lambda_1 + \lambda_2, c), v^{(i_k, k)}(\lambda_1 + \lambda_2, c) \]
\[ = \left( c_1 + \Pi \mu^{(i_k, k)}(\lambda_1, c_1), v^{(i_k, k)}(\lambda_1, c_1) \right) \]
\[ + \left( c_2 + \Pi \mu^{(i_k, k)}(\lambda_2, c_2), v^{(i_k, k)}(\lambda_2, c_2) \right) \]
\[ \subset \Psi_1(B^{w_{\geq k+1}}(\lambda_1)) + \Psi_1(B^{w_{\geq k+1}}(\lambda_2)). \]

Hence we conclude that \( \Psi_1(B^{w_{\geq k+1}}(\lambda_1 + \lambda_2)) \subset \Psi_1(B^{w_{\geq k+1}}(\lambda_1)) + \Psi_1(B^{w_{\geq k+1}}(\lambda_2)) \). This proves the lemma. \( \Box \)

Note that \( \Delta_1(m \lambda) = m \Delta_1(\lambda) \) for \( m \in \mathbb{Z}_{>0} \) by the additivity of \( \Psi_1 \) (see [10, Theorem 4.1]). Hence if \( \Delta_1(\lambda) \) is a parapolytope, then the polytopes \( \Delta_1(m \lambda) \), \( m \in \mathbb{Z}_{>0} \), are all parapolytopes. By Lemma 4.7, this implies that
\[ \Psi_1(B(m \lambda)) = \frac{\Psi_1(B(\lambda)) + \cdots + \Psi_1(B(\lambda))}{m} \]
for \( m \in \mathbb{Z}_{>0} \), and hence that the equality \( \Delta_1(\lambda) = \text{Conv}(\Psi_1(B(\lambda))) \) holds. This proves part (1) of Theorem 4.1.

For \( c = (c_s)_{1 \leq s \leq N} : i_s \neq k \in (\mathbb{R}^{d_{i_k}})^- \) and \( l = k-1, k \) such that \( \text{Conv}(\Psi_1(B^{w_{\geq l+1}}(\lambda))) \cap (c + \mathbb{R}^{d_{i_k}}) \neq \emptyset \), there uniquely exist
\[ \mu_+^{(i_k, l)}(c) = (\mu_+^{(i_k, l)}(c), \ldots, \mu_+^{(i_k, l)}(c)), \]
\[ \nu_+^{(i_k, l)}(c) = (\nu_+^{(i_k, l)}(c), \ldots, \nu_+^{(i_k, l)}(c)) \in \mathbb{R}^{d_{i_k}} \]
such that
\[ \text{Conv}(\Psi_1(B^{w_{\geq l+1}}(\lambda))) \cap (c + \mathbb{R}^{d_{i_k}}) = c + \Pi (\mu_+^{(i_k, l)}(c), \nu_+^{(i_k, l)}(c)). \]

If we set
If we denote by
\[
\tilde{\mu}_+(i_k, l) (c) = (\mu_{i_k, l} (c), \ldots, \mu_{i_k, l} (c)) := (\hat{\lambda}_{i_k}, \ldots, \hat{\lambda}_{i_k}) - \mu_{i_k, l}(c), \quad \text{and}
\]
\[
\tilde{v}_+ (i_k, l) (c) = (v_{i_k, l} (c), \ldots, v_{i_k, l} (c)) := (\hat{\lambda}_{i_k}, \ldots, \hat{\lambda}_{i_k}) - \mu_{i_k, l}(c)
\]
for \( l = k - 1, k \), then we have
\[
\left( -\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda))) + (\hat{\lambda}_{i_k}, \ldots, \hat{\lambda}_{i_k}) \right) \cap (c + \mathbb{R}^{d_{i_k}}) = \tilde{c} + \Pi (\tilde{\mu}_+(i_k, l) (c), \tilde{v}_+(i_k, l) (c)).
\]

**Lemma 4.8** For \( c \in (\mathbb{R}^{d_{i_k}})^\perp \), it follows that \( \text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda))) \cap (c + \mathbb{R}^{d_{i_k}}) \neq \emptyset \) if and only if \( \text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda))) \cap (c + \mathbb{R}^{d_{i_k}}) \neq \emptyset \).

**Proof** If we denote by \( P_k : \mathbb{R}^N \rightarrow (\mathbb{R}^{d_{i_k}})^\perp \) the canonical projection, then we have
\[
P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) = \{ c \in (\mathbb{R}^{d_{i_k}})^\perp \mid \text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda))) \cap (c + \mathbb{R}^{d_{i_k}}) \neq \emptyset \}
\]
for \( l = k, k + 1 \). Hence it suffices to prove that \( P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) = P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) \).

Since \( B^{w_{i_k}+1}(\lambda) \subseteq B^{w_{i_k}+1}(\lambda) \), we have \( P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) \subseteq P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) \). Let \( c \in (\mathbb{Z}^{d_{i_k}})^\perp \) be a vertex of the lattice polytope \( P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) \). Then, it follows that \( \text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda))) \cap (c + \mathbb{R}^{d_{i_k}}) \neq \emptyset \), and that \( \mu_{i_k}(c), \nu_{i_k}(c) \in \mathbb{Q}^{d_{i_k}} \). We take \( l \in \mathbb{Z}_{>0} \) such that \( l\mu_{i_k}(c), l\nu_{i_k}(c) \in \mathbb{Z}^{d_{i_k}} \).

Since we have
\[
\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda))) \cap (\ell c + \mathbb{Z}^{d_{i_k}}) = \ell \left( \text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda))) \cap (c + \mathbb{R}^{d_{i_k}}) \right)
\]
\[
= \ell c + \Pi (l\mu_{i_k}(c), l\nu_{i_k}(c)),
\]
it follows that
\[
\Psi_1(B^{w_{i_k}+1}(\lambda)) \cap (\ell c + \mathbb{Z}^{d_{i_k}}) = \left( \text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda))) \cap (\ell c + \mathbb{R}^{d_{i_k}}) \right) \cap \mathbb{Z}^N \neq \emptyset.
\]

Hence Lemma 4.6 implies that
\[
\Psi_1(B^{w_{i_k}+1}(\lambda)) \cap (\ell c + \mathbb{Z}^{d_{i_k}}) \neq \emptyset,
\]
and hence that
\[
\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda))) \cap (c + \mathbb{R}^{d_{i_k}}) = \frac{1}{\ell} \left( \text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda))) \cap (\ell c + \mathbb{R}^{d_{i_k}}) \right) \neq \emptyset.
\]

Thus, the vertices of \( P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) \) are contained in \( P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) \). From this and the convexity of \( P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) \), we obtain \( P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) \subseteq P_k (\text{Conv}(\Psi_1(B^{w_{i_k}+1}(\lambda)))) \), which proves the lemma. \( \square \)
Lemma 4.9 The polytope \( \text{Conv}(\Psi_1(\mathcal{B}^{w \geq k}(\lambda))) \) is a parapolytope for all \( 1 \leq k \leq N + 1 \), and the following equality holds for all \( 1 \leq k \leq N \):

\[
-\text{Conv}(\Psi_1(\mathcal{B}^{w \geq k+1}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}) = D_{i_k}^{(k)} \left( -\text{Conv}(\Psi_1(\mathcal{B}^{w \geq k}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}) \right).
\]

Proof Since \( \Delta_1(\lambda) = \text{Conv}(\Psi_1(\mathcal{B}(\lambda))) \), Proposition 3.18 implies that

\[
\text{Conv}(\Psi_1(\mathcal{B}^{w \geq k}(\lambda))) = \{ (a_1, \ldots, a_N) \in \Delta_1(\lambda) \mid a_k = x_k, \ldots, a_N = x_N \},
\]

and hence that this is a parapolytope. In particular, a function

\[
D_{i_k}^{(k)} \left( -\text{Conv}(\Psi_1(\mathcal{B}^{w \geq k}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}) \right)
\]

is well-defined. We will show that

\[
\left( -\text{Conv}(\Psi_1(\mathcal{B}^{w \geq k+1}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}) \right) \cap (\mathbf{c} + \mathbb{R}^{d_{i_k}}) = D_{i_k}^{(k)} \left( \left( -\text{Conv}(\Psi_1(\mathcal{B}^{w \geq k}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}) \right) \cap (\mathbf{c} + \mathbb{R}^{d_{i_k}}) \right)
\]

for all \( \mathbf{c} \in (\mathbb{R}^{d_{i_k}})^\perp \) such that \( \text{Conv}(\Psi_1(\mathcal{B}^{w \geq k}(\lambda))) \cap (\mathbf{c} + \mathbb{R}^{d_{i_k}}) \neq \emptyset \).

First, we consider the case \( \mathbf{c} \in (\mathbb{Q}^{d_{i_k}})^\perp \), where we set \((\mathbb{Q}^{d_{i_k}})^\perp := (\mathbb{R}^{d_{i_k}})^\perp \cap \mathbb{Q}^N\). In this case, we have

\[
\mu_{+}^{(i_k,k-1)}(\mathbf{c}), \nu_{+}^{(i_k,k-1)}(\mathbf{c}), \mu_{+}^{(i_k,k)}(\mathbf{c}), \nu_{+}^{(i_k,k)}(\mathbf{c}) \in \mathbb{Q}^{d_{i_k}}.
\]

By the definition of \( D_{i_k}^{(k)} \), it suffices to prove that there exists \( l \in \mathbb{Z}_{\geq 0} \) such that

\[
\left( l \left( -\text{Conv}(\Psi_1(\mathcal{B}^{w \geq k+1}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}) \right) \right) \cap (l\mathbf{c} + \mathbb{R}^{d_{i_k}}) = D_{i_k}^{(k)} \left( \left( l \left( -\text{Conv}(\Psi_1(\mathcal{B}^{w \geq k}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}) \right) \right) \cap (l\mathbf{c} + \mathbb{R}^{d_{i_k}}) \right).
\]

From this, we may assume that \( \mathbf{c} \in (\mathbb{Z}^{d_{i_k}})^\perp \), \( \mu_{+}^{(i_k,k-1)}(\mathbf{c}), \nu_{+}^{(i_k,k-1)}(\mathbf{c}), \mu_{+}^{(i_k,k)}(\mathbf{c}), \nu_{+}^{(i_k,k)}(\mathbf{c}) \in \mathbb{Z}^{d_{i_k}} \). Then, the following equalities hold:

\[
\mu_{+}^{(i_k,k-1)}(\mathbf{c}) = \mu^{(i_k,k-1)}(\mathbf{c}), \quad \nu_{+}^{(i_k,k-1)}(\mathbf{c}) = \nu^{(i_k,k-1)}(\mathbf{c}),
\]

\[
\mu_{+}^{(i_k,k)}(\mathbf{c}) = \mu^{(i_k,k)}(\mathbf{c}), \quad \nu_{+}^{(i_k,k)}(\mathbf{c}) = \nu^{(i_k,k)}(\mathbf{c}).
\]

We set

\[
\tilde{\mu}^{(i_k,l)}(\mathbf{c}) = (\tilde{\mu}^{(i_k,l)}_{+}(\mathbf{c}), \ldots, \tilde{\mu}^{(i_k,l)}_{d_{i_k}}(\mathbf{c})) := (\hat{\lambda}_{i_k}, \ldots, \hat{\lambda}_{i_k}) - \nu^{(i_k,l)}(\mathbf{c}), \text{ and}
\]

\[
\tilde{\nu}^{(i_k,l)}(\mathbf{c}) = (\tilde{\nu}^{(i_k,l)}_{+}(\mathbf{c}), \ldots, \tilde{\nu}^{(i_k,l)}_{d_{i_k}}(\mathbf{c})) := (\hat{\lambda}_{i_k}, \ldots, \hat{\lambda}_{i_k}) - \mu^{(i_k,l)}(\mathbf{c})
\]
for \(l = k - 1, k\). By the definition of \(D_{l_k}^{(k)}\), the polytope

\[
D_{l_k}^{(k)} \left( \left( - \text{Conv}(\Psi_1(B^{w \geq k}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}) \right) \cap (\tilde{c} + \mathbb{R}^d_{l_k}) \right) = D_{l_k}^{(k)} \left( \tilde{c} + \Pi((\tilde{\mu}^{(i_k,k-1)}(c), \tilde{\nu}^{(i_k,k-1)}(c))) \right)
\]

is given by replacing \(\tilde{\nu}^{(i_k,k-1)}(c)\) in \(\tilde{\nu}^{(i_k,k-1)}(c)\) with

\[
\tilde{\nu}^{(i_k,k-1)}(c) := \tilde{\nu}^{(i_k,k-1)}(c) - \sum_{1 \leq s \leq N; i_s \neq i_k} c_{i_k,i_s} (\hat{\lambda}_{i_s} - c_s)
\]

\[
- \sum_{1 \leq l \leq d_{l_k}} (\mu^{(i_k,k-1)}(c) + \nu^{(i_k,k-1)}(c))
\]

if \(\tilde{\nu}^{(i_k,k-1)}(c) \geq \tilde{\nu}^{(i_k,k-1)}(c)\). Note that

\[
\tilde{\nu}^{(i_k,k-1)}(c) - \tilde{\nu}^{(i_k,k-1)}(c) = - \sum_{1 \leq s \leq N} c_{i_k,i_s} \hat{\lambda}_{i_s} + \sum_{1 \leq s \leq N; i_s \neq i_k} c_{i_k,i_s} c_s
\]

\[
+ \sum_{1 \leq l \leq d_{l_k}} (\mu^{(i_k,k-1)}(c) + \nu^{(i_k,k-1)}(c))
\]

\[
= L_k(c)
\]

since \(\sum_{1 \leq s \leq N} c_{i_k,i_s} \hat{\lambda}_{i_s} = (\lambda, h_{e_k})\) by \(\lambda = \sum_{i \in I} \hat{\lambda}_i d_i a_i\). Since \(L_k(c) \geq 0\) by Lemma 4.5, it follows that \(D_{l_k}^{(k)}(\tilde{c} + \Pi((\tilde{\mu}^{(i_k,k-1)}(c), \tilde{\nu}^{(i_k,k-1)}(c)))\) is the polytope given by replacing \(\tilde{\nu}^{(i_k,k-1)}(c)\) in \(\tilde{c} + \Pi((\tilde{\mu}^{(i_k,k-1)}(c), \tilde{\nu}^{(i_k,k-1)}(c)))\) with \(\tilde{\nu}^{(i_k,k-1)}(c) = \tilde{\nu}^{(i_k,k-1)}(c) + L_k(c)\), which implies by Lemma 4.6 that

\[
D_{l_k}^{(k)}(\tilde{c} + \Pi((\tilde{\mu}^{(i_k,k-1)}(c), \tilde{\nu}^{(i_k,k-1)}(c))) = \tilde{c} + \Pi((\tilde{\mu}^{(i_k,k)}(c), \tilde{\nu}^{(i_k,k)}(c))).
\]

Second, we consider the case \(c \in (\mathbb{R}^d_{l_k})^\perp\). We regard \(\tilde{\mu}^{(i_k,k-1)}(c), \tilde{\nu}^{(i_k,k-1)}(c), \tilde{\mu}^{(i_k,k)}(c), \tilde{\nu}^{(i_k,k)}(c)\) for \(1 \leq l \leq d_{l_k}\) as \(\mathbb{R}\)-valued functions on the lattice polytope

\[
P_{l_k}(\text{Conv}(\Psi_1(B^{w \geq k+1}(\lambda)))) = P_{l_k}(\text{Conv}(\Psi_1(B^{w \geq k}(\lambda))))
\]

see the proof of Lemma 4.8. Since \(-\text{Conv}(\Psi_1(B^{w \geq k+1}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N})\) and \(-\text{Conv}(\Psi_1(B^{w \geq k}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N})\) are convex, the functions \(\tilde{\mu}^{(i_k,k-1)}(c), \tilde{\nu}^{(i_k,k-1)}(c), \tilde{\mu}^{(i_k,k)}(c), \tilde{\nu}^{(i_k,k)}(c)\) are (upper or lower) convex on each line segment \(S \subset P_{l_k}(\text{Conv}(\Psi_1(B^{w \geq k+1}(\lambda))))\); hence they are continuous on the relative interior of \(S\). From this and the assertion in the case \(c \in (\mathbb{Q}^d_{l_k})^\perp\), we deduce that
for all $c \in P_k(\text{Conv}(\Psi_i(B^{w_{k+1}}(\lambda))))$. This proves the lemma. 

Since we have

$$-\text{Conv}(\Psi_i(B^{w_{k+1}}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}) = -x_\lambda + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}) = a_\lambda,$$

Lemma 4.9 implies that $D^{(k)}_{i_k} \cdots D^{(1)}_{i_1}(a_\lambda)$ is a well-defined parapolytope for $1 \leq k \leq N$, and that the following equality holds for $1 \leq k \leq N$:

$$D^{(k)}_{i_k} \cdots D^{(1)}_{i_1}(a_\lambda) = -\text{Conv}(\Psi_i(B^{w_{k+1}}(\lambda))) + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}).$$

From these, we obtain parts (2), (3) of Theorem 4.1.

### 4.3 Immediate consequences

By Theorem 4.1 (1) and Lemma 4.7, we obtain the following.

**Theorem 4.10** Let $i \in I^N$ be a reduced word for $w_0$, and $\lambda, \mu \in P_+$. Assume that the polytopes $\Delta_i(\lambda)$, $\Delta_i(\mu)$, and $\Delta_i(\lambda + \mu)$ are all parapolytopes. Then, the following equalities hold:

$$\Psi_i(B(\lambda + \mu)) = \Psi_i(B(\lambda)) + \Psi_i(B(\mu)), \text{ and}$$

$$\Delta_i(\lambda + \mu) = \Delta_i(\lambda) + \Delta_i(\mu).$$

The proof of Theorem 4.1 implies the following.

**Proposition 4.11** Let $i = (i_1, \ldots, i_N) \in I^N$ be a reduced word for $w_0$, $\lambda \in P_+$, and $2 \leq k \leq N$. Assume that the face $\{a \in \Delta_i(\lambda) \mid a_k = x_k, \ldots, a_N = x_N\}$ of $\Delta_i(\lambda)$ is a parapolytope.

1. The face $\{a \in \Delta_i(\lambda) \mid a_k = x_k, \ldots, a_N = x_N\}$ is a lattice polytope.
2. The polytope $D^{(l)}_{i_l} \cdots D^{(1)}_{i_1}(a_\lambda)$ is well-defined for $1 \leq l \leq k - 1$.
3. The following equality holds for all $1 \leq l \leq k - 1$:

$$D^{(l)}_{i_l} \cdots D^{(1)}_{i_1}(a_\lambda) = -\{a \in \Delta_i(\lambda) \mid a_{l+1} = x_{l+1}, \ldots, a_N = x_N\} + (\hat{\lambda}_{i_1}, \ldots, \hat{\lambda}_{i_N}).$$

### 5 Crystal structures

In this section, we study the crystal structure on the set of lattice points in $\Delta_i(\lambda)$. Recall that $e_i, f_i, h_i \in g$, $i \in I$, are the Chevalley generators such that
$i \in I \subset \text{Lie}(B)$ and $\{f_i, h_i \mid i \in I\} \subset \text{Lie}(B^-)$. For $i \in I$, let $g_i$ be the Lie subalgebra of $g$ generated by $e_i, f_i, h_i$, which is isomorphic to $\mathfrak{sl}_2(\mathbb{C})$ as a Lie algebra. For $m \in \mathbb{Z}_{\geq 0}$, we denote by $B^{(i)}(m)$ the crystal basis for the $(m+1)$-dimensional irreducible $g_i$-module with highest weight element $b_m$. We fix $i \in I$ and $c \in (\mathbb{Z}^{d_i})^-$ such that $\Psi_i(B(\lambda)) \cap (c + \mathbb{Z}^{d_i}) \neq \emptyset$. Recall that $\mu_i(c) = (\mu_{i1}(c), \ldots, \mu_{id_i}(c))$, $\nu_i(c) = (\nu_{i1}(c), \ldots, \nu_{id_i}(c)) \in \mathbb{Z}^{d_i}$ are uniquely determined by

$$\Psi_i(B(\lambda)) \cap (c + \mathbb{Z}^{d_i}) = c + \Pi_{\mathbb{Z}}(\mu_i(c), \nu_i(c)).$$

We define a bijective map

$$\eta_i : B(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}^{d_i}) \sim B^{(i)}(\nu_{i1}(c) - \mu_{i1}(c)) \otimes \cdots \otimes B^{(i)}(\nu_{id_i}(c) - \mu_{id_i}(c))$$

by

$$\eta_i(b) := f_i^{a_i^{(i)}(c)} b_{\nu_{i1}(c)-\mu_{i1}(c)} \otimes \cdots \otimes f_i^{a_{d_i}^{(i)}(c)} b_{\nu_{id_i}(c)-\mu_{id_i}(c)}$$

when $\Psi_i(b) = c + (\{a^{(i)}_1, \ldots, a^{(i)}_{d_i}\})$ in $c + \mathbb{Z}^{d_i}$.

**Proposition 5.1** The map $\eta_i$ is an isomorphism of $g_i$-crystals.

**Proof** It suffices to prove that $\eta_i$ is compatible with the actions of $\tilde{e}_i$ and $\tilde{f}_i$. We show that $\eta_i(\tilde{e}_ib) = \tilde{e}_i\eta_i(b)$ for all $b \in B(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}^{d_i})$, where we set $\eta_i(0) := 0$ if $\tilde{e}_ib = 0$; a proof of the compatibility with $\tilde{f}_i$ is similar. Let $b_{\text{high}}$ (resp., $b_{\text{low}}$) be the unique element in $B(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}^{d_i})$ such that $\Psi_i(b_{\text{high}}) = c + \mu^{(i)}(c)$ (resp., $\Psi_i(b_{\text{low}}) = c + \nu^{(i)}(c)$). Considering the weights of elements in the $g_i$-crystal $B(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}^{d_i})$, the standard representation theory of $\mathfrak{sl}_2(\mathbb{C})$ implies that $b_{\text{high}}$ is the highest weight element in the $i$-string through $b_{\text{low}}$. By the crystal structure on $\mathbb{Z}_j^{\infty} \otimes R_\lambda$, this implies that

$$\eta_i(\hat{e}_i^k b_{\text{low}}) = \hat{e}_i^k \eta_i(b_{\text{low}})$$

(5.1)

for all $k \in \mathbb{Z}_{\geq 0}$. We set

$$\{s_1 < \cdots < s_{d_i}\} := \{1 \leq s \leq N \mid i_s = i\}.$$

For $b \in B(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}^{d_i})$, define $\Upsilon_1(b), \Upsilon_2(b), \ldots, \Upsilon_{d_i+1}(b)$ by

$$\tilde{\Upsilon}_j(b) = \Upsilon_1(b) \otimes \Upsilon_2(b) \otimes \cdots \otimes \Upsilon_{d_i+1}(b)$$

$$\in (\mathbb{Z}_{j \geq N+1} \otimes \tilde{B}_{i_1} \otimes \cdots \otimes \tilde{B}_{i_{s_1}}) \otimes (\tilde{B}_{i_{s_1}+1} \otimes \cdots \otimes \tilde{B}_{i_{s_2}}) \otimes \cdots \otimes (\tilde{B}_{i_{s_{d_i}+1}} \otimes \cdots \otimes \tilde{B}_{i_N} \otimes R_\lambda),$$
and set $\Upsilon_{\leq k}(b) := \Upsilon_1(b) \otimes \Upsilon_2(b) \otimes \cdots \otimes \Upsilon_k(b)$ for $1 \leq k \leq d_i$. In addition, for $b \in B^{(i)}(v^{(i)}_1(c) - \mu^{(i)}_1(c)) \otimes \cdots \otimes B^{(i)}(v^{(i)}_{d_i}(c) - \mu^{(i)}_{d_i}(c))$, we define $\Upsilon_1(b), \Upsilon_2(b), \ldots, \Upsilon_{d_i}(b)$ by

$$b = \Upsilon_1(b) \otimes \Upsilon_2(b) \otimes \cdots \otimes \Upsilon_{d_i}(b)$$

$$\in B^{(i)}(v^{(i)}_1(c) - \mu^{(i)}_1(c)) \otimes B^{(i)}(v^{(i)}_2(c) - \mu^{(i)}_2(c)) \otimes \cdots \otimes B^{(i)}(v^{(i)}_{d_i}(c) - \mu^{(i)}_{d_i}(c)),$$

and set $\Upsilon_{\leq k}(b) := \Upsilon_1(b) \otimes \Upsilon_2(b) \otimes \cdots \otimes \Upsilon_k(b)$ for $1 \leq k \leq d_i$. By the tensor product rule for crystals, it suffices to prove that

$$\varepsilon_i(\Upsilon_{\leq k}(b)) = \varepsilon_i(\Upsilon_{\leq k}(\eta_i(b))), \quad 1 \leq k \leq d_i,$$

$$\varepsilon_i(\Upsilon_k(b)) - \varphi_i(\Upsilon_{\leq k-1}(b)) = \varepsilon_i(\Upsilon_k(\eta_i(b))) - \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b))), \quad 2 \leq k \leq d_i,$$

for $b \in B(\lambda) \cap \Psi_i^{-1}(c + \mathbb{Z}^{d_i})$. We proceed by induction on $k$.

If $k = 1$, then we take $b'$ in the $i$-string through $b_{\text{low}}$ such that $\Upsilon_1(b') = \Upsilon_1(b)$; the existence of $b'$ follows by (5.1). Then, we deduce that

$$\varepsilon_i(\Upsilon_{\leq 1}(b)) = \varepsilon_i(\Upsilon_1(b')) = \varepsilon_i(\Upsilon_1(\eta_i(b'))) \quad \text{(by (5.1))}$$

$$= \varepsilon_i(\Upsilon_1(\eta_i(b))) \quad \text{(by the definition of } \eta_i)$$

$$= \varepsilon_i(\Upsilon_{\leq 1}(\eta_i(b))).$$

If $k \geq 2$, then we take $b''$ in the $i$-string through $b_{\text{low}}$ such that

$$\Upsilon_{\leq k-1}(b'') = \Upsilon_{\leq k-1}(b_{\text{low}}), \quad \Upsilon_k(b'') = \Upsilon_k(b);$$

the existence of $b''$ follows by (5.1). Then, it follows that

$$\varepsilon_i(\Upsilon_k(b)) - \varphi_i(\Upsilon_{\leq k-1}(b)) = \varepsilon_i(\Upsilon_k(b')) - \varphi_i(\Upsilon_{\leq k-1}(b'')) + \varphi_i(\Upsilon_{\leq k-1}(b'')) - \varphi_i(\Upsilon_{\leq k-1}(b))$$

$$= \varepsilon_i(\Upsilon_k(b'')) - \varphi_i(\Upsilon_{\leq k-1}(b'')) + \varphi_i(\Upsilon_{\leq k-1}(b_{\text{low}})) - \varphi_i(\Upsilon_{\leq k-1}(b)) \quad \text{(5.2)}$$

$$= \varepsilon_i(\Upsilon_k(b'')) - \varphi_i(\Upsilon_{\leq k-1}(b'')) + \varepsilon_i(\Upsilon_{\leq k-1}(b_{\text{low}})) - \varepsilon_i(\Upsilon_{\leq k-1}(b))$$

$$+ \langle \text{wt}(\Upsilon_{\leq k-1}(b_{\text{low}})), h_i \rangle - \langle \text{wt}(\Upsilon_{\leq k-1}(b)), h_i \rangle.$$ 

Note that the following equality holds by (5.1):

$$\varepsilon_i(\Upsilon_k(b'')) - \varphi_i(\Upsilon_{\leq k-1}(b'')) = \varepsilon_i(\Upsilon_k(\eta_i(b''))) - \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b'))) \quad \text{(5.3)}$$

In addition, we deduce by (5.1) and by the induction hypothesis that

$$\varepsilon_i(\Upsilon_{\leq k-1}(b_{\text{low}})) - \varepsilon_i(\Upsilon_{\leq k-1}(b)) = \varepsilon_i(\Upsilon_{\leq k-1}(\eta_i(b_{\text{low}}))) - \varepsilon_i(\Upsilon_{\leq k-1}(\eta_i(b))) \quad \text{(5.4)}$$
If we write $\Psi_i(b) = c + (d_{i}^{(t)}, \ldots, d_{i}^{(t)})$ in $c + \mathbb{Z}^d$, then we have

$$\langle \text{wt}(\Upsilon_{\leq k-1}(b_{\text{low}})), h_i \rangle - \langle \text{wt}(\Upsilon_{\leq k-1}(b)), h_i \rangle$$

$$= 2 \sum_{1 \leq i \leq k-1} (d_{i}^{(t)} - \nu_{i}^{(t)}(c))$$

$$= \langle \text{wt}(\Upsilon_{\leq k-1}(\eta_i(b_{\text{low}}))), h_i \rangle - \langle \text{wt}(\Upsilon_{\leq k-1}(\eta_i(b))), h_i \rangle$$

by the definition of $\eta_i$. By (5.2)–(5.5), it follows that

$$\varepsilon_i(\Upsilon_k(b)) - \varphi_i(\Upsilon_{\leq k-1}(b))$$

$$= \varepsilon_i(\Upsilon_k(\eta_i(b''))) - \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b'')))$$

$$+ \varepsilon_i(\Upsilon_{\leq k-1}(\eta_i(b_{\text{low}}))) - \varepsilon_i(\Upsilon_{\leq k-1}(\eta_i(b)))$$

$$+ \langle \text{wt}(\Upsilon_{\leq k-1}(\eta_i(b_{\text{low}}))), h_i \rangle - \langle \text{wt}(\Upsilon_{\leq k-1}(\eta_i(b))), h_i \rangle$$

$$= \varepsilon_i(\Upsilon_k(\eta_i(b''))) - \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b''))) + \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b_{\text{low}})))$$

$$- \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b)))$$

$$= \varepsilon_i(\Upsilon_k(\eta_i(b'''))) - \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b'''))) + \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b_{\text{low}}))) - \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b)))$$

(by the definition of $\eta_i$)

$$= \varepsilon_i(\Upsilon_k(\eta_i(b))) - \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b))),$$

and hence that

$$\varepsilon_i(\Upsilon_{\leq k}(b)) = \max\{\varepsilon_i(\Upsilon_{\leq k-1}(b)), \varepsilon_i(\Upsilon_{\leq k-1}(b)) + \varepsilon_i(\Upsilon_k(b)) - \varphi_i(\Upsilon_{\leq k-1}(b))\}$$

(by the tensor product rule for crystals)

$$= \max\{\varepsilon_i(\Upsilon_{\leq k-1}(\eta_i(b))), \varepsilon_i(\Upsilon_{\leq k-1}(\eta_i(b)))$$

$$+ \varepsilon_i(\Upsilon_k(\eta_i(b))) - \varphi_i(\Upsilon_{\leq k-1}(\eta_i(b)))\}$$

$$= \varepsilon_i(\Upsilon_{\leq k}(\eta_i(b))).$$

This proves the proposition. \qed

### 6 Geometric applications

In this section, we discuss toric degenerations arising from Nakashima–Zelevinsky polytopes by the theory of Newton–Okounkov bodies [2]. We start with recalling the main result of [10], which states that $\Delta_i(\lambda)$ is identical to the Newton–Okounkov body of the full flag variety $G/B$ associated with a specific valuation. For $\lambda \in P_+$, we define a line bundle $\mathcal{L}_{\lambda}$ on $G/B$ by

$$\mathcal{L}_{\lambda} := (G \times \mathbb{C})/B,$$
where $B$ acts on $G \times \mathbb{C}$ on the right as follows:

$$(g, c) \cdot b = (gb, \lambda(b)c)$$

for $g \in G$, $c \in \mathbb{C}$, and $b \in B$. Take a reduced word $i = (i_1, \ldots, i_N) \in I^N$ for the longest element $w_0 \in W$. We see by [17, Ch. II.13] that the morphism

$$\mathbb{C}^N \to G/B, \quad (t_1, \ldots, t_N) \mapsto \exp(t_1 f_{i_1}) \cdots \exp(t_N f_{i_N}) \mod B,$$

is birational. Hence the function field $\mathbb{C}(G/B)$ is identified with the rational function field $\mathbb{C}(t_1, \ldots, t_N)$.

**Definition 6.1** We define a lexicographic order $\prec$ on $\mathbb{Z}^N$ as follows: $(a_1, \ldots, a_N) \prec (a_1', \ldots, a_N')$ if and only if there exists $1 \leq k \leq N$ such that $a_N = a_N'$, and $a_{k+1} = a_{k+1}'$. The lexicographic order $\prec$ on $\mathbb{Z}^N$ induces a total order (denoted by the same symbol $\prec$) on the set of monomials in the polynomial ring $\mathbb{C}[t_1, \ldots, t_N]$ as follows: $t_1^{a_1} \cdots t_N^{a_N} \prec t_1^{a_1'} \cdots t_N^{a_N'}$ if and only if $(a_1, \ldots, a_N) \prec (a_1', \ldots, a_N')$. Let us define a valuation $v_{i, \prec}^{\text{high}} : \mathbb{C}(G/B) \setminus \{0\} \to \mathbb{Z}^N$ by $v_{i, \prec}^{\text{high}}(f/g) := v_{i, \prec}^{\text{high}}(f) - v_{i, \prec}^{\text{high}}(g)$ for $f, g \in \mathbb{C}(t_1, \ldots, t_N) \setminus \{0\}$, and by

$$v_{i, \prec}^{\text{high}}(f) := -(a_1, \ldots, a_N)$$

for $f = ct_1^{a_1} \cdots t_N^{a_N} + (\text{lower terms}) \in \mathbb{C}[t_1, \ldots, t_N] \setminus \{0\}$, where $c \in \mathbb{C} \setminus \{0\}$, and we mean by “lower terms” a linear combination of monomials smaller than $t_1^{a_1} \cdots t_N^{a_N}$ with respect to the total order $\prec$.

**Definition 6.2** (See [23, Sect. 1.2] and [25, Definition 1.10]) Let $i \in I^N$ be a reduced word for $w_0$, and $\lambda \in P_+$. Take a nonzero section $\tau \in H^0(G/B, L_\lambda)$. We define a subset $S(G/B, L_\lambda, v_{i, \prec}^{\text{high}}, \tau) \subset \mathbb{Z}_{>0} \times \mathbb{Z}^N$ by

$$S(G/B, L_\lambda, v_{i, \prec}^{\text{high}}, \tau) := \bigcup_{k>0} \{(k, v_{i, \prec}^{\text{high}}(\sigma/k)) \mid \sigma \in H^0(G/B, L_\lambda^{\boxtimes k}) \setminus \{0\} \},$$

and denote by $C(G/B, L_\lambda, v_{i, \prec}^{\text{high}}, \tau) \subset \mathbb{R}_{\geq 0} \times \mathbb{R}^N$ the smallest real closed cone containing $S(G/B, L_\lambda, v_{i, \prec}^{\text{high}}, \tau)$. Let us define a subset $\Delta(G/B, L_\lambda, v_{i, \prec}^{\text{high}}, \tau) \subset \mathbb{R}^N$ by

$$\Delta(G/B, L_\lambda, v_{i, \prec}^{\text{high}}, \tau) := \{a \in \mathbb{R}^N \mid (1, a) \in C(G/B, L_\lambda, v_{i, \prec}^{\text{high}}, \tau) \};$$

this is called the Newton–Okounkov body of $G/B$ associated with $L_\lambda, v_{i, \prec}^{\text{high}}$, and $\tau$.

We define an $\mathbb{R}$-linear automorphism $\omega : \mathbb{R} \times \mathbb{R}^N \to \mathbb{R} \times \mathbb{R}^N$ by $\omega(k, a) := (k, -a)$. 
Theorem 6.3 (See [10, Sect. 4]) Let \( \mathbf{i} \in I^N \) be a reduced word for \( w_0 \), and \( \lambda \in P_+ \). Then, there exists a nonzero section \( \tau_\lambda \in H^0(G/B, L_\lambda) \) such that the following equalities hold:

\[
S_i(\lambda) = \omega(S(G/B, L_\lambda, v_{i,\leq}^\text{high}, \tau_\lambda)), \quad C_i(\lambda) = \omega(C(G/B, L_\lambda, v_{i,\leq}^\text{high}, \tau_\lambda)), \quad \text{and} \quad \Delta_i(\lambda) = -\Delta(S(G/B, L_\lambda, v_{i,\leq}^\text{high}, \tau_\lambda)).
\]

Remark 6.4 The author and Oya [11] proved that \( \Delta_i(\lambda) \) is also identical to the Newton–Okounkov body of \( G/B \) associated with a geometrically natural valuation, which is given by counting the orders of zeros along a specific sequence of Schubert subvarieties.

We say that \( G/B \) admits a flat degeneration to a variety \( X \) if there exists a flat morphism

\[
\pi : \mathcal{X} \to \text{Spec}(\mathbb{C}[t])
\]

of schemes such that the scheme-theoretic fiber \( \pi^{-1}(t) \) (resp., \( \pi^{-1}(0) \)) over a closed point \( t \in \mathbb{C} \setminus \{0\} \) (resp., the origin \( 0 \in \mathbb{C} \)) is isomorphic to \( G/B \) (resp., \( X \)). By Theorem 6.3 and [2, Theorem 1] (see also [15, Corollary 3.14]), there exists a flat degeneration of \( G/B \) to \( \text{Proj}(\mathbb{C}[S_i(\lambda)]) \), where the \( \mathbb{Z}_{\geq 0} \)-grading of \( S_i(\lambda) \) induces a \( \mathbb{Z}_{\geq 0} \)-grading of \( \mathbb{C}[S_i(\lambda)] \). By Proposition 3.10 (1) and [6, Theorem 1.3.5], we see that \( \text{Proj}(\mathbb{C}[S_i(\lambda)]) \) is normal; hence it is identical to the normal toric variety \( X(\Delta_i(\lambda)) \) associated with the rational convex polytope \( \Delta_i(\lambda) \). Thus, we obtain the following.

Theorem 6.5 There exists a flat degeneration of \( G/B \) to the normal toric variety \( X(\Delta_i(\lambda)) \) associated with the Nakashima–Zelevinsky polytope \( \Delta_i(\lambda) \).

We apply Alexeev–Brion’s argument [1] to this flat degeneration.

Definition 6.6 Let \( \mathbf{i} \in I^N \) be a reduced word for \( w_0 \), and write \( P_\mathbb{R} := P \otimes_{\mathbb{Z}} \mathbb{R} \). Define a subset \( S_i \subset P_+ \times \mathbb{Z}^N \) by

\[
S_i := \bigcup_{\lambda \in P_+} \{(\lambda, \Psi_i(b)) \mid b \in B(\lambda)\},
\]

and denote by \( C_i \subset P_\mathbb{R} \times \mathbb{R}^N \) the smallest real closed cone containing \( S_i \).

In a way similar to the proof of [10, Corollaries 2.18 and 4.3], we deduce the following.

Proposition 6.7 Let \( \mathbf{i} \in I^N \) be a reduced word for \( w_0 \). Then, the real closed cone \( C_i \) is a rational convex polyhedral cone, and the equality \( S_i = C_i \cap (P_+ \times \mathbb{Z}^N) \) holds.

Let \( \{\varpi_i \mid i \in I\} \subset P_+ \) be the set of fundamental weights, and \( P_{\mathbb{R},+} \subset P_\mathbb{R} \) the closure of the fundamental Weyl chamber with respect to the Euclidean topology, that
is,
\[ P_{\mathbb{R},+} := \sum_{i \in I} \mathbb{R}_{\geq 0} \omega_i. \]

Denote by \( \pi_1 : P_{\mathbb{R}} \times \mathbb{R}^N \to P_{\mathbb{R}} \) the first projection, which maps the rational convex polyhedral cone \( C_i \) onto \( P_{\mathbb{R},+} \). Then, for \( \lambda \in P_+ \), the Nakashima–Zelevinsky polytope \( \Delta_i(\lambda) \) is identical to the fiber \( C_i \cap \pi_1^{-1}(\lambda) \). Imitating \cite[Definition 4.1]{1}, we define a fan \( \Sigma_i \) from \( C_i \). For \( \lambda \in P_{\mathbb{R},+} \), we set
\[ F_\lambda := \{ \text{faces } \tau \mid \lambda \in \pi_1(\tau^0) \}, \]
\[ \sigma_\lambda^0 := \bigcap_{\tau \in F_\lambda} \pi_1(\tau^0), \]
where \( \tau^0 \) is the relative interior of \( \tau \). Denote by \( \sigma_\lambda \) the closure of \( \sigma_\lambda^0 \) in \( P_{\mathbb{R},+} \) with respect to the Euclidean topology. Then, a fan \( \Sigma_i \) with support \( P_{\mathbb{R},+} \) is defined to be
\[ \Sigma_i := \{ \sigma_\lambda \mid \lambda \in P_{\mathbb{R},+} \}; \]
the fan \( \Sigma_i \) is said to be trivial if it consists only of the faces of \( P_{\mathbb{R},+} \). Let \( P_{++} \subset P_+ \) denote the set of regular dominant integral weights. For \( \lambda \in P_{++} \), the line bundle \( L_\lambda \) on \( G/B \) is very ample (see, for instance, \cite[Sect. II.8.5]{17}); hence we see by \cite[Corollary 3.2]{25} that the real dimension of \( \Delta_i(\lambda) \) equals \( N \). In a way similar to the argument in \cite[1]{1}, we obtain the following.

**Proposition 6.8** (cf. \cite[Lemma 4.2 and Corollary 4.3]{1}) Let \( i \in I^N \) be a reduced word for \( w_0 \).

1. Two weights \( \lambda, \mu \in P_+ \) lie in the same cone of \( \Sigma_i \) if and only if \( \Delta_i(\lambda + \mu) \) is the Minkowski sum of \( \Delta_i(\lambda) \) and \( \Delta_i(\mu) \).
2. If the fan \( \Sigma_i \) is trivial, then the polytopes \( \Delta_i(\lambda), \lambda \in P_{++} \), have the same normal fan.

The following is an immediate consequence of Theorem 4.10 and Proposition 6.8.

**Corollary 6.9** If \( \Delta_i(\lambda) \) is a parapolytope for all \( \lambda \in P_+ \), then the toric varieties \( X(\Delta_i(\lambda)), \lambda \in P_{++} \), are all identical.

We say that \( X(\Delta_i(\lambda)) \) is Gorenstein Fano if the anti-canonical class \( -K_X(\Delta_i(\lambda)) \) is Cartier and ample (see \cite[Sect. 8.3]{6}). Let \( \mathcal{O}(K_{G/B}) \) denote the canonical bundle of \( G/B \). By \cite[Proposition 2.2.7 (ii)]{3}, we have \( \mathcal{O}(K_{G/B}) \simeq L_{-2\rho} \), where \( \rho \in P_{++} \) is the half sum of the positive roots. By the argument in the proof of \cite[Proposition 2.4]{1} (see also \cite[Theorem 3.8]{1}), the anti-canonical sheaf \( \mathcal{O}(-K_X(\Delta_i(2\rho))) \) is the limit of \( L_{2\rho} \simeq \mathcal{O}(K_{G/B}) \) under the flat degeneration of \( G/B \) to \( X(\Delta_i(2\rho)) \) in Theorem 6.5. Hence we obtain the following by Theorem 4.1 (1).

**Corollary 6.10** If \( \Delta_i(2\rho) \) is a parapolytope, then the toric variety \( X(\Delta_i(2\rho)) \) is Gorenstein Fano, that is, \( \Delta_i(2\rho) \) is reflexive.

By Corollaries 6.9 and 6.10, we obtain Corollary 4 in Introduction.
Acknowledgements  The author is greatly indebted to Satoshi Naito for numerous helpful suggestions and fruitful discussions. The author would also like to express his gratitude to Dave Anderson and Valentina Kiritchenko for useful comments and suggestions. At the conference “Algebraic Analysis and Representation Theory” in June 2017, the author gave a poster presentation on the result of this paper. But there was a gap in the proof at that time, and the condition of the main result has been corrected from the one at the conference.

References

1. Alexeev, V., Brion, M.: Toric degenerations of spherical varieties. Sel. Math. (N.S.) 10, 453–478 (2004)
2. Anderson, D.: Okounkov bodies and toric degenerations. Math. Ann. 356, 1183–1202 (2013)
3. Brion, M.: Lectures on the geometry of flag varieties. In: Pragacz, P. (ed.) Topics in Cohomological Studies of Algebraic Varieties. Trends in Mathematics, pp. 33–85. Birkhäuser, Basel (2005)
4. Caldero, P.: Toric degenerations of Schubert varieties. Transform. Groups 7, 51–60 (2002)
5. Chirivi, R.: LS algebras and application to Schubert varieties. Transform. Groups 5, 245–264 (2000)
6. Cox, D., Little, J., Schenck, H.: Toric Varieties. Graduate Studies in Mathematics, vol. 124. American Mathematical Society, Providence, RI (2011)
7. Fang, X., Fourier, G., Littelmann, P.: Essential bases and toric degenerations arising from birational sequences. Adv. Math. 312, 107–149 (2017)
8. Fang, X., Fourier, G., Littelmann, P.: On toric degenerations of flag varieties. In: Representation Theory—Current Trends and Perspectives, EMS Series of Congress Reports, pp. 187–232. European Mathematical Society, Zürich (2017)
9. Feigin, E., Fourier, G., Littelmann, P.: Favourable modules: filtrations, polytopes, Newton–Okounkov bodies and flat degenerations. Transform. Groups 22, 321–352 (2017)
10. Fujita, N., Naito, S.: Newton–Okounkov convex bodies of Schubert varieties and polyhedral realizations of crystal bases. Math. Z. 285, 325–352 (2017)
11. Fujita, N., Oya, H.: A comparison of Newton–Okounkov polytopes of Schubert varieties. J. Lond. Math. Soc. (2) 96, 201–227 (2017)
12. Gelfand, I.M., Zetlin, M.L.: Finite-dimensional representations of the group of unimodular matrices. Doklady Akad. Nauk SSSR (N.S.) 71, 825–828 (1950)
13. Gonciulea, N., Lakshmibai, V.: Degenerations of flag and Schubert varieties to toric varieties. Transform. Groups 1, 215–248 (1996)
14. Grossberg, M., Karshon, Y.: Bott towers, complete integrability, and the extended character of representations. Duke Math. J. 76, 23–58 (1994)
15. Harada, M., Kaveh, K.: Integrable systems, toric degenerations, and Okounkov bodies. Invent. Math. 202, 927–985 (2015)
16. Hoshino, A.: Polyhedral realizations of crystal bases for quantum algebras of finite types. J. Math. Phys. 46, 113514 (2005)
17. Jantzen, J.C.: Representations of Algebraic Groups. Mathematical Surveys and Monographs, vol. 107, 2nd edn. American Mathematical Society, Providence, RI (2003)
18. Kashiwara, M.: Crystallizing the $q$-analogue of universal enveloping algebras. Commun. Math. Phys. 133, 249–260 (1990)
19. Kashiwara, M.: On crystal bases of the $q$-analogue of universal enveloping algebras. Duke Math. J. 63, 465–516 (1991)
20. Kashiwara, M.: Global crystal bases of quantum groups. Duke Math. J. 69, 455–485 (1993)
21. Kashiwara, M.: The crystal base and Littelmann’s refined Demazure character formula. Duke Math. J. 71, 839–858 (1993)
22. Kashiwara, M.: On crystal bases. In: Representations of Groups (Banff, AB, 1994), CMS Conference Proceedings, vol. 16, pp. 155–197. American Mathematical Society, Providence, RI (1995)
23. Kaveh, K.: Crystal bases and Newton–Okounkov bodies. Duke Math. J. 164, 2461–2506 (2015)
24. Kaveh, K., Khovanskii, A.G.: Convex bodies and algebraic equations on affine varieties. Preprint, arXiv:0804.4095v1; a short version with title Algebraic equations and convex bodies appeared in Perspectives in Analysis, Geometry, and Topology, Progress in Mathematics, vol. 296. Birkhäuser, New York 2012, 263–282 (2008)
25. Kaveh, K., Khovanskii, A.G.: Newton–Okounkov bodies, semigroups of integral points, graded algebras and intersection theory. Ann. Math. 176, 925–978 (2012)
26. Kiritchenko, V.: Divided difference operators on convex polytopes. In: Naruse, H., Ikeda, T., Masuda, M., Tanisaki, T. (eds.) Schubert Calculus—Osaka 2012. Advanced Studies in Pure Mathematics, vol. 66, pp. 161–184. Mathematical Society of Japan, Tokyo (2016)
27. Kiritchenko, V.: Newton–Okounkov polytopes of flag varieties. Transform. Groups 22, 387–402 (2017)
28. Kiritchenko, V.: Newton–Okounkov polytopes of Bott–Samelson varieties as Minkowski sums. Preprint arXiv:1801.00334v1 (2018)
29. Kiritchenko, V., Smirnov, E., Timorin, V.: Schubert calculus and Gelfand–Tsetlin polytopes. Russ. Math. Surv. 67, 685–719 (2012)
30. Lazarsfeld, R., Mustata, M.: Convex bodies associated to linear series. Ann. Sci. l’ENS 42, 783–835 (2009)
31. Lusztig, G.: Canonical bases arising from quantized enveloping algebras. J. Am. Math. Soc. 3, 447–498 (1990)
32. Lusztig, G.: Quivers, perverse sheaves, and quantized enveloping algebras. J. Am. Math. Soc. 4, 365–421 (1991)
33. Lusztig, G.: Introduction to Quantum Groups. Modern Birkhäuser Classics. Birkhäuser, New York (2010). (reprint of the 1994 edition)
34. Nakashima, T.: Polyhedral realizations of crystal bases for integrable highest weight modules. J. Algebra 219, 571–597 (1999)
35. Nakashima, T.: Polytopes for crystallized Demazure modules and extremal vectors. Commun. Algebra 30, 1349–1367 (2002)
36. Nakashima, T.: Decorated geometric crystals, polyhedral and monomial realizations of crystal bases. In: Chari, V., Greenstein, J., Misra, K.C., Raghavan, K.N., Viswanath, S. (eds.) Recent Developments in Algebraic and Combinatorial Aspects of Representation Theory. Contemporary Mathematics, vol. 602, pp. 143–163. American Mathematical Society, Providence, RI (2013)
37. Nakashima, T., Zelevinsky, A.: Polyhedral realizations of crystal bases for quantized Kac–Moody algebras. Adv. Math. 131, 253–278 (1997)
38. Okounkov, A.: Brunn–Minkowski inequality for multiplicities. Invent. Math. 125, 405–411 (1996)
39. Okounkov, A.: Multiplicities and Newton polytopes. In: Kirillov’s Seminar on Representation Theory, American Mathematical Society Translations: Series 2, vol. 181, Advances in the Mathematical Sciences, vol. 35, pp. 231–244. American Mathematical Society, Providence, RI (1998)
40. Okounkov, A.: Why would multiplicities be log-concave? In: Duval, C., Guieu, L., Ovsienko, V. (eds.) The Orbit Method in Geometry and Physics. Progress in Mathematics, vol. 213, pp. 329–347. Birkhäuser, Boston (2003)

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.