

NEWTON–OKOUNKOV BODIES AND PICARD NUMBERS ON SURFACES

JULIO-JOSÉ MOYANO-FERNÁNDEZ, MATTHIAS NICKEL, AND JOAQUIM ROÉ

Abstract: We study the shapes of all Newton–Okounkov bodies $\Delta_v(D)$ of a given big divisor D on a surface S with respect to all rank 2 valuations v of $K(S)$.

We obtain upper bounds for, and in many cases we determine exactly, the possible numbers of vertices of the bodies $\Delta_v(D)$. The upper bounds are expressed in terms of Picard numbers and they are *birationally invariant*, as they do not depend on the model \tilde{S} where the valuation v becomes a flag valuation.

We also conjecture that the set of all Newton–Okounkov bodies of a single ample divisor D determines the Picard number of S , and prove that this is the case for Picard number 1, by an explicit characterization of surfaces of Picard number 1 in terms of Newton–Okounkov bodies.

2020 Mathematics Subject Classification: Primary: 14C20; Secondary: 14E15, 14C22, 13A18.

Key words: valuation, blowup, Newton–Okounkov body, algebraic surface, Picard number.

1. Introduction

Newton–Okounkov bodies were introduced by A. Okounkov ([24]) as a tool in representation theory; later, K. Kaveh and A. G. Khovanskii ([17, 18]) on the one hand, and R. Lazarsfeld and M. Mustață ([23]) on the other, developed a general theory with applications to both convex and algebraic geometry. Very quickly they gained a central role in the asymptotic theory of linear series on projective varieties, and they have been used to address questions of arithmetic geometry, combinatorics, Diophantine approximation, mirror symmetry, and representation theory; see S. Boucksom’s excellent exposition [3] for the general theory. Blum and Jonsson ([2]), Boucksom and Chen ([4]), Fang, Fourier, and Littelmann ([10]), Harada and Kaveh ([15]), or Rietsch and Williams ([25]) are a sample of recent applications.

To every big divisor D on a normal projective variety X of dimension d , and to every valuation v of the field $K(X)$ of rank d , one associates a convex body in \mathbb{R}^d , called the Newton–Okounkov body $\Delta_v(D)$ of D with respect to v ; it contains all normalized values of divisors in the complete linear systems $|kD|$, $k \gg 0$. Its volume agrees (up to a factor of $d!$) with the volume of D for every v , but its shape depends on the valuation v . A lot of effort has been put into trying to understand this dependence, and more generally the information encoded in the shapes of Newton–Okounkov bodies (see [20], [6], [9], [14], and [26]). By works of Lazarsfeld and Mustață ([23])

During the elaboration and revision of this work, the first author was partially funded by MCIN/AEI/10.13039/501100011033, by “ERDF A way of making Europe”, and by “European Union NextGeneration EU/PRTR”, grants PID2022-138906NB-C22 and TED2021-130358B-I00, as well as by Universitat Jaume I, grants UJI-B2021-02 and GACUJIMA-2023-06. The second author was partially supported by the LOEWE grant “Uniformized Structures in Algebra and Geometry”. The third author was partially supported by MCIN/AEI/10.13039/501100011033, grants MTM2016-75980-P and PID2020-116542GB-I00, and by the Direcció General de Recerca, grant 2021 SGR 01468.

and Jow ([16]), we know that the collection of all Newton–Okounkov bodies of D is a complete numerical invariant of D , and by work of Küronya, Lozovanu, and Maclean ([20]), we know that on a surface S all Newton–Okounkov bodies are polygons.

Every valuation v of maximal rank of the field $K(S)$ can be built from a smooth flag of subvarieties on some (smooth) birational model $\tilde{S} \rightarrow S$ of S , and it is often simpler to focus on flags living on S itself: the proof of polygonality of two-dimensional Newton–Okounkov bodies, for instance, relies on the flag construction (and Zariski decompositions). It follows from the proof that the number of vertices of a Newton–Okounkov polygon with respect to a smooth flag on S is at most $2\rho(S) + 2$, where $\rho(S)$ is the Picard number of S . However, this bound is often not sharp; in [27], the third author and T. Szemberg determined the number of vertices a Newton–Okounkov polygon may have, where D is required to be ample, S is smooth, and the valuation is induced by a smooth flag on S .

The main thrust of the present work is to showcase the benefits of considering valuations determined by flags on higher models \tilde{S} and not just on S . Although this does often lead to slightly more involved arguments, one in fact gains clarity as statements become simpler; this is the case for instance in our first main result. It could be expected that, since higher smooth models $\tilde{S} \rightarrow S$ have larger Picard number, the number of vertices of Newton–Okounkov polygons with respect to flags lying on \tilde{S} could grow; somewhat surprisingly, this is not the case:

Theorem A. *Let S be a smooth projective algebraic surface, and let D be a big Cartier divisor on S . For every rank 2 valuation v of $K(X)$, the Newton–Okounkov body $\Delta_v(D)$ has at most $2\rho(S) + 2$ vertices.*

There are geometric restrictions which usually make it impossible to attain the maximal number of vertices $2\rho(S)+2$ with flags lying on S ; see [27]. These restrictions loosen when we consider arbitrary valuations of maximal rank, and we in fact hope they disappear:

Theorem B. *Let S be a smooth projective algebraic surface with Picard number $\rho(S) = 1$, and let D be an ample Cartier divisor on S . There exist rank 2 valuations v_3, v_4 of $K(S)$, such that the Newton–Okounkov body $\Delta_{v_k}(D)$ has exactly k vertices.*

Conversely, if S is a smooth projective algebraic surface with Picard number $\rho(S) > 1$, then there exist rank 2 valuations v such that $\Delta_v(D)$ has at least five vertices.

In fact, we propose the following:

Conjecture. *Let S be a smooth projective algebraic surface, and let D be an ample Cartier divisor on S . For every natural number k , $3 \leq k \leq 2\rho(S) + 2$, there exists a rank 2 valuation v of $K(S)$, such that the Newton–Okounkov body $\Delta_v(D)$ has exactly k vertices.*

Theorem B on one hand shows that the upper bound can always be attained if the Picard number is 1 (generalizing results of [6, 14] for \mathbb{P}^2) and on the other hand characterizes surfaces with Picard number 1 in terms of Newton–Okounkov bodies. Our conjecture means that we expect both facts to generalize for arbitrary Picard number, and so that Newton–Okounkov bodies determine Picard numbers.

One sees in Theorem A that if D is the pullback of a divisor by a birational morphism $S \rightarrow S'$ of smooth projective surfaces, then the Picard number of S' (smaller

than that of S) may be used to bound the number of vertices of Newton–Okounkov bodies of D . This suggests that curves contracted by the maps $S \dashrightarrow \mathbb{P}^N$ (which are birational to the image for large k and big D) “should not count” in the Picard number upper bound; the appropriate notion is the following.

Definition 1.1. Let S be a smooth projective surface, and D a big and nef Cartier divisor on S . We define the Picard number of S relative to D , or simply the Picard number of D , written $\rho_D(S)$, as the Picard number of S minus the number of irreducible curves C on S with $C \cdot D = 0$. (If D is big but not nef, we define $\rho_D(S)$ as $\rho_{P_D}(S)$, where P_D is the positive part in the Zariski decomposition of D ; see Definition 2.5.)

In Section 2 we show that this Picard number is invariant under pullback by birational morphisms; this allows us to extend the notion to singular surfaces using resolution of singularities. Then we can prove a stronger version of Theorem A.

Theorem A#. *Let S be a normal projective algebraic surface, and let D be a big Cartier divisor on S . Then for every rank 2 valuation v on $K(S)$, the Newton–Okounkov body $\Delta_v(D)$ has at most $2\rho_D(S) + 2$ vertices.*

It is plausible that the conjecture above holds for big (not necessarily ample) divisors replacing ρ with ρ_D , but we have little evidence to support this assertion.

In the proofs of Theorems A# and B we use the description of Newton–Okounkov polygons in terms of Zariski decompositions introduced by Lazarsfeld and Mustață in [23] and we also employ techniques from [20] and [27]. However, these techniques do not immediately carry over to valuations whose flags live in higher models $\tilde{S} \rightarrow S$. One important fact that comes to our help is that *curves orthogonal to D never contribute vertices* to a polygon $\Delta_v(D)$; cf. Lemma 3.1. We also investigate Newton–Okounkov bodies with respect to smooth flags on a fixed model, but since our objective is to apply this study to higher models $\tilde{S} \rightarrow S$, it becomes essential to pay attention to “contracted curves”, i.e., irreducible curves whose intersection number with the given big divisor D is not positive. Flags whose divisorial part does intersect D positively turn out to behave in a way similar to smooth flags, which allows us to extend some results of [27] for ample divisors to the general setting of big divisors, with suitable modifications to take care of contracted curves. We attach to every big divisor D on a smooth surface S a number $mv(D)$, computed in terms of configurations of negative curves (Definition 4.1), and invariant under pullbacks, which allows us to describe all Newton–Okounkov bodies of D whose flag lives on a given model and intersects D positively:

Theorem C. *On every smooth projective surface S , for every big divisor D and every integer δ with $3 \leq \delta \leq mv(D)$, there exists a smooth flag Y_\bullet in S such that the Newton–Okounkov polygon $\Delta_{Y_\bullet}(D)$ has exactly δ vertices.*

On every normal projective surface S , for every big and nef divisor D , there is a model $\tilde{S} \rightarrow S$ and a flag $Y_\bullet = \{\tilde{S} \supset C \supset \{p\}\}$ such that $D \cdot \pi(C) > 0$ and Δ_{Y_\bullet} is a k -gon if and only if $3 \leq k \leq mv(D)$.

If D is big but not nef, a version of the second statement still holds, only replacing $D \cdot \pi(C)$ with $P_D \cdot \pi(C)$, where P_D stands for the positive part in the Zariski decomposition $D = P_D + N_D$.

The number $mv(D)$ is usually smaller than $2\rho_D(S) + 2$, so there are still geometric restrictions making it impossible to reach the maximum number of vertices; this can only be achieved by using flags (on some model $\tilde{S} \rightarrow S$) whose divisorial part does not intersect (the pullback of) D . These *contracted* or *infinitesimal flags* are extremely interesting, and are the ones allowing us to prove Theorem B, but it is also challenging to work with them; in particular, it seems that the path to proving the conjecture above should go through the construction of adequate infinitesimal flags, which we were unable to do at this point.

It follows from our analysis that there are important qualitative differences between the Newton–Okounkov bodies of a fixed divisor D with respect to flags whose divisorial part has zero or non-zero intersection with D . We would like to point out that such a distinction appeared already in the work of Choi, Park, and Won [5, Theorem 1.1], although we are not aware of any direct connection between both phenomena, since they work with Newton–Okounkov bodies of non-big pseudoeffective divisors, which are just segments and thus the question of the number of points is moot for them.

The paper is organized as follows. In Section 2 we recall necessary notions and relevant facts from existing literature. Section 3 is devoted to the proof of Theorem A#. In Section 4 we study Newton–Okounkov bodies with respect to flags whose divisorial part intersects D positively and prove Theorem C. In Sections 5 and 6 we study Newton–Okounkov bodies with respect to flags whose divisorial part does not intersect D ; first we exhibit examples that show the differences in behavior with respect to positively intersecting flags, and finally we prove Theorem B.

We work over the complex field.

2. Newton–Okounkov bodies on surfaces

In this section we recall the terminology used throughout the paper, according to the conventions in [26], [6], or [14]. For a general reference on positivity the reader is referred to [21, 22].

Let X be a normal, projective, algebraic variety of dimension n defined over the complex numbers with function field $K(X)$. The group of Cartier divisors on X will be denoted as $\text{Div}(X)$, whereas $\text{Num}(X)$ will be the subgroup of divisors numerically equivalent to 0, and $N^1(X) = \text{Div}(X)/\text{Num}(X)$ will be the Néron–Severi group of X . Let us recall [21, 1.1.16]: the Néron–Severi group is a free abelian group of finite rank $\rho(X)$, the Picard number of X .

A Cartier divisor D on X is called *big* if $h^0(X, \mathcal{O}_X(kD))$ grows like k^n . This is equivalent to saying that $\mathcal{O}_X(D)$ has maximal Iitaka dimension. Bigness makes sense also for \mathbb{Q} - and \mathbb{R} -divisors, and it only depends on the numerical equivalence class of D ; see [23, (0.3)]. The convex cone in $N^1(X)_{\mathbb{R}}$ of all big \mathbb{R} -divisor classes on X is called the *big cone*, and it will be denoted by $\text{Big}(X)$.

Let $v: K(X) \setminus \{0\} \rightarrow \mathbb{Z}_{\text{lex}}^n$ be a discrete valuation of rank n on X whose value group is \mathbb{Z}^n endowed with the lexicographic ordering. For any such valuation and any big divisor D on X one defines the Newton–Okounkov body of D with respect to v as

$$\Delta_v(D) = \overline{\left\{ \frac{v(s)}{k} : s \in H^0(kD) \subset K(X), k \in \mathbb{N} \right\}},$$

where $\overline{\{\cdot\}}$ stands for the closure with respect to the usual topology of \mathbb{R}^n .

The simplest valuations of maximal rank, and the most used in the context of Newton–Okounkov bodies, are defined via flags. Let us consider a flag of irreducible subvarieties

$$Y_{\bullet} = \{X = Y_0 \supset Y_1 \supset \cdots \supset Y_n = \{p\}\}.$$

The flag Y_\bullet is said to be *full* and *admissible* if Y_i has codimension i in X and is smooth at the point p . The point p is called the *center* of the flag. It is possible to construct a discrete valuation of rank n from Y_\bullet as follows. For every $\phi \in K(X)$ we set $\phi_0 = \phi$, and

$$v_i(\phi) = \text{ord}_{Y_i}(\phi_{i-1}), \quad \phi_i = \frac{\phi_{i-1}}{g_i^{v_i(\phi)}} \Big|_{Y_i} \quad \text{for } i = 1, \dots, n,$$

where g_i is a local equation of Y_i in Y_{i-1} around p . It is clear that $v_{Y_\bullet} = (v_1, \dots, v_n)$ defines a rank n discrete valuation $K(X) \setminus \{0\} \rightarrow \mathbb{Z}_{\text{lex}}^n$. Now we can define the Newton–Okounkov body of D with respect to the flag Y_\bullet as $\Delta_{Y_\bullet}(D) := \Delta_{v_{Y_\bullet}}(D)$.

This definition can be extended to the case of an admissible flag on a birational model \tilde{X} of X , where $\pi: \tilde{X} \rightarrow X$ is a proper birational morphism. Such a flag

$$Y_\bullet = \{\tilde{X} = Y_0 \supset Y_1 \supset \dots \supset Y_n = \{p\}\}$$

is called *infinitesimal* if $\pi(Y_1) = \pi(p)$ and *proper* if $\text{codim } \pi(Y_i) = i$. If $\tilde{X} = X$ and $\pi = \text{id}_X$, then the flag is said to be *smooth* at p . The Newton–Okounkov bodies given by infinitesimal resp. proper resp. smooth flags are called infinitesimal resp. proper resp. smooth. By [6, Theorem 2.9], every discrete valuation of rank n , and hence every Newton–Okounkov body, is of one of these types.

We are interested in the case $n = 2$. Let S be therefore a complex projective normal surface.

Remark 2.1. For the computations of Newton–Okounkov bodies, we shall use the Zariski decomposition of effective divisors, which needs S to be smooth; this is not really restrictive, because passing to a resolution of singularities $\pi: \tilde{S} \rightarrow S$, one has $\Delta_v(D) = \Delta_v(\pi^*(D))$. On the other hand, by [6, Theorem 2.9] for every rank 2 valuation v there is a model $\pi: \tilde{S} \rightarrow S$ and a smooth flag $Y_\bullet = \{\tilde{S} \supset C \supset \{p\}\}$ on \tilde{S} such that v is the valuation associated to Y_\bullet . Hence we may also assume that v is a flag valuation. Note that the choice of flag is uniquely determined on every model where it exists: if $\pi': \tilde{S}' \rightarrow \tilde{S}$ is a higher model, then the restriction of π' to the strict transform C' of C in \tilde{S}' is an isomorphism, so we have a uniquely defined flag $Y'_\bullet = \{\tilde{S}' \supset C' \supset \{p'\}\}$ on \tilde{S}' inducing the same valuation, namely the one given by the unique point p' on C' above p .

When S is smooth, every big \mathbb{Q} -divisor D admits a Zariski decomposition $D = N_D + P_D$, where P_D (the positive part of D) is a nef \mathbb{Q} -divisor and N_D (the negative part of D) is an effective \mathbb{Q} -divisor with the property that, whenever kD and kN_D are integral divisors, multiplication by the section defining kN_D induces an isomorphism

$$H^0(S, \mathcal{O}(kP_D)) \longrightarrow H^0(S, \mathcal{O}(kD)).$$

Zariski decompositions can also be defined for big \mathbb{Q} - and \mathbb{R} -divisors. They are determined by the numerical equivalence class of D , so it makes sense to study them as functions on the Néron–Severi space [1].

The first explicit examples of Newton–Okounkov bodies appeared in [23, Section 6]. In fact, [23, Lemma 6.3 and Theorem 6.4] gives a description of those bodies associated to any big divisor on S , in terms of Zariski decompositions:

Proposition 2.2 (Lazarsfeld–Mustață). *Let D be a big divisor on a smooth projective surface S and let $Y_\bullet = \{S \supset C \supset \{p\}\}$ be an admissible flag on S . Let $D_t = P_t + N_t$ be the Zariski decomposition of $D_t = D - tC$, $t \in \mathbb{R}$. Let ν be the coefficient of C in N_0 , and let $\mu = \sup\{t > 0 : D - tC \text{ is big}\} > \nu \geq 0$. Then the Newton–Okounkov body of D with respect to Y_\bullet is given by*

$$\Delta_{Y_\bullet} = \{(t, y) \in \mathbb{R}^2 : \nu \leq t \leq \mu, \alpha(t) \leq y \leq \beta(t)\},$$

where

$$\alpha(t) = \text{ord}_p(N_t|_C), \quad \beta(t) = \alpha(t) + (C \cdot P_t).$$

Remark 2.3. Since by [6, Theorem 2.9] every rank 2 valuation on a surface S is a flag valuation on some model $\tilde{S} \rightarrow S$, the previous proposition holds for every Newton–Okounkov body $\Delta_\nu(D)$, provided that the Zariski decomposition is done on \tilde{S} .

In the study of Zariski decompositions, Bauer, Küronya, and Szemberg introduced in [1] the concept of the Zariski chamber as a subcone building a partition of the big cone

$$\text{Big}(X) = \bigcup_{P \text{ big} \wedge \text{ nef}} \Sigma_P,$$

where

$$\begin{aligned} \Sigma_P &= \{D \in \text{Big}(X) : \text{Neg}(D) = \text{Null}(P)\}, \\ \text{Neg}(D) &= \{C : C \text{ is an irreducible component of } N_D\}, \\ \text{Null}(D) &= \{C : C \text{ is an irreducible curve with } C \cdot D = 0\}. \end{aligned}$$

Note that the definitions give $\text{Neg}(D) \cup \text{Null}(D) \subset \text{Null}(P_D)$ (the union may be disjoint or not). By exploiting the Zariski chamber decomposition it is not hard to prove that Zariski decompositions depend continuously on the big divisor D :

Proposition 2.4 ([1, Proposition 1.16]). *Let (D_n) be a sequence of big divisors converging in $N^1(X)_\mathbb{R}$ to a big divisor D . If $D_n = P_n + N_n$ is the Zariski decomposition of D_n , and if $D = P + N$ is the Zariski decomposition of D , then the sequences (P_n) and (N_n) converge to P and N respectively.*

This fact guarantees, for instance, that the functions α and β above are continuous, and it will be essential for our arguments later on.

Definition 2.5. Let D be a big divisor on a projective algebraic surface S . We define the *Picard number of S relative to D* , if S is smooth, as $\rho_D(S) = \rho(S) - \#\text{Null}(P_D)$.

The definition can be extended to the case of a possibly singular normal surface S , as follows. Observe first that the definition for smooth surfaces is invariant under pullback by proper birational morphisms: if S is smooth and $\pi: \tilde{S} \rightarrow S$ is the blowup centered at a point p , then $\pi^*(P_D) = P_{\pi^*(D)}$, and by the projection formula [11, 8.1.7],

$$\text{Null}(\pi^*(P_D)) = \{\tilde{C} : C \in \text{Null}(D)\} \cup \{E_p\},$$

therefore

$$(1) \quad \#\text{Null}(\pi^*(D)) - \#\text{Null}(D) = 1 = \rho(\tilde{S}) - \rho(S).$$

It follows that $\rho_{\pi^*D}(\tilde{S}) = \rho_D(S)$. Now, if D is a big and nef divisor on a normal surface S , we may define $\rho_D(S)$ as $\rho_{\pi^*(D)}(\tilde{S})$ for any resolution of singularities $\pi: \tilde{S} \rightarrow S$; equation (1) guarantees that this definition does not depend on the choice of a resolution.

Remark 2.6. When S is smooth, as D is big, P_D is big as well, and in particular $P_D^2 > 0$. The index theorem implies that the intersection form restricted to $P_D^\perp \supset \langle \text{Null}(P_D) \rangle$ is negative definite, and the number of curves $\#\text{Null}(P_D)$ is bounded above by $\rho(S) - 1$. Therefore $1 \leq \rho_D(S) \leq \rho(S)$, with $\rho_D(S) = \rho(S)$ if and only if D is ample.

The description of Newton–Okounkov bodies of Proposition 2.2 was further investigated by Küronya, Lozovanu, and Maclean in [20] and by Roé and Szemberg in [27]; we summarize their results as the following theorem:

Theorem 2.7 (Küronya–Lozovanu–Maclean, Roé–Szemberg). *Let D be a big divisor on a smooth surface S , and consider a smooth flag $Y_\bullet = \{S \supset C \supset \{p\}\}$. Let $D_t = D - tC$ be with Zariski decomposition $D_t = P_t + N_t$ as above, and let $\alpha(t)$, $\beta(t)$ be the functions determining the lower and upper boundaries of $\Delta_{Y_\bullet}(D)$ given by Proposition 2.2. Then,*

- (i) $\alpha(t)$ is a non-decreasing function [20, Theorem B].
- (ii) $\alpha(t)$ and $\beta(t)$ are piecewise linear functions with finitely many pieces [20, Theorem B].
- (iii) The function $\alpha(t)$ or $\beta(t)$ has a point of non-differentiability at $t = t_0$ if and only if the negative part of $D_t = D - tC$ acquires one or more irreducible components at t_0 , i.e., for every $\varepsilon > 0$ the supports of N_{t_0} and $N_{t_0+\varepsilon}$ differ.
- (iv) The function $\alpha(t)$ has a point of non-differentiability at $t = t_0$ if and only if one or more irreducible components acquired at t_0 are contained in the connected component of $N_{t_0+\varepsilon}$ that passes through p [27, Proposition 3.3].
- (v) The function $\beta(t)$ has a point of non-differentiability at $t = t_0$ if and only if one or more irreducible components acquired at t_0 are contained in a connected component of $N_{t_0+\varepsilon}$ that intersects C away from p [27, Proposition 3.3].

Proof: The only statement that is not immediately clear from the references is (iii). By (iv) and (v) we only need to show that if all irreducible components acquired at t_0 do not pass through p , then one or more of them is contained in a connected component of $N_{t_0+\varepsilon}$ that intersects C away from p . Finally, this is true since otherwise the support of N_{t_0} and $N_{t_0+\varepsilon}$ would not change. \square

The theorem immediately extends to a normal surface as long as Y_\bullet is a smooth flag, by computing the Zariski decomposition of the pullback of D to any resolution of singularities of S .

The fact that N_t acquires a new irreducible component at t_0 can be equivalently expressed by saying that D_t crosses a Zariski wall between chambers, i.e., D_{t_0} and $D_{t_0+\varepsilon}$ belong to different cones Σ_P for every $\varepsilon > 0$.

These results were then used to connect the number of vertices of the Newton–Okounkov body with the Picard number of the surface.

Corollary 2.8. *Let S be a smooth surface, D a big divisor on S , and $Y_\bullet = \{S \supset C \supset \{p\}\}$ a smooth flag. Let $D_t = D - tC$ be with Zariski decomposition $D_t = P_t + N_t$ as above, and let $\mu = \sup\{t > 0 : D_t \text{ is big}\}$.*

Then $\Delta_{Y_\bullet}(D)$ is a polygon with at most $2\rho(S) + 1$ vertices. If N_μ has $\rho(S) - 1$ irreducible components, then $\Delta_{Y_\bullet}(D)$ has a single rightmost vertex.

In the statement of Corollary 2.8, one can “almost” replace $\rho(S)$ with $\rho_D(S)$, as we shall see in Theorem 3.2. This has two advantages; first, and most obvious, the bound is sharper for non-ample D , and second, $\rho_D(S)$ is invariant under proper birational morphisms π (replacing D with $\pi^*(D)$). We include a short sketch of the proof of Corollary 2.8, since we will use the ideas therein extensively.

Proof: The description of the Newton–Okounkov body in Proposition 2.2 together with the decomposition of the big cone into Zariski chambers [1] show that the number of vertices of $\Delta_{Y_\bullet}(D)$ is bounded by two times the number of Zariski walls that $D_t = D - tC$ crosses for $0 \leq \nu < t < \mu$ and the leftmost and rightmost vertices (at most two of each).

Let us prove the second claim first. If N_ν consists of $\rho - 1$ irreducible components, then by the index theorem the divisor P_μ , which is orthogonal to all components of N_μ , must be a multiple of a big class. Since P_μ is not big, it has to be zero, and so $\Delta_{Y_\bullet}(D)$ has only one rightmost vertex.

Next, the minimality of the negative part in the Zariski decomposition ensures that $\text{supp}(N_t) \subset \text{supp}(N_\mu)$ for $0 \leq \nu \leq t \leq \mu$. As N_μ has negative definite intersection matrix, by the index theorem the support of N_μ can only contain $\rho - 1$ curves at most, and therefore at most $\rho - 1$ Zariski walls are crossed. Therefore $\Delta_{Y_\bullet}(D)$ is a polygon with at most $2(\rho - 1) + 4 = 2\rho(S) + 2$ vertices. However, if $\Delta_{Y_\bullet}(D)$ had exactly $2\rho(S) + 2$ vertices, N_ν would consist of $\rho - 1$ irreducible components, and therefore $\Delta_{Y_\bullet}(D)$ would have only one rightmost vertex, so $\Delta_{Y_\bullet}(D)$ has at most $2\rho(S) + 1$ vertices, as claimed. \square

Moreover, [27] shows that the bound is attained, i.e., for every positive integer ρ there exist a smooth projective surface S , a big divisor D , and a smooth flag Y_\bullet such that $\Delta_{Y_\bullet}(D)$ has $2\rho(S) + 1$ vertices. To do so, they determine the numbers of vertices that can be obtained as Newton–Okounkov bodies of every *ample* divisor D . We can partially extend these results of [27] to the case of non-ample divisors, applying techniques from the theory of stable base loci on surfaces (see [1, 28]).

Given an effective divisor D , let $\text{Fix}(D)$ denote the fixed divisor of $|D|$ (i.e., every D' linearly equivalent to D satisfies $D' \geq \text{Fix}(D)$ and the base locus of $|D - \text{Fix}(D)|$ is zero-dimensional). Let us denote by $SB(D)$ the stable base locus of D , i.e., the intersection of the base loci of $|kD|$ for all $k \geq 1$. We also use the notation $\text{supp}(D)$ for the support of an effective divisor D . The following lemma collects a few known facts:

Lemma 2.9. *Let D be a big divisor on a smooth surface S and let $D = P_D + N_D$ be its Zariski decomposition. Then*

- (i) *The stable base locus $SB(D)$ decomposes as $SB(D) = \text{supp}(N_D) \cup SB(P_D)$.*
- (ii) *The stable base locus $SB(P_D)$ is contained in $\text{supp}(\text{Fix}(P_D))$, and it has no isolated points.*
- (iii) *The stable base locus $SB(P_D)$ is the union of some of the components of $\text{Null}(P_D)$.*
- (iv) *For every component E of $SB(P_D)$, the coefficient of E in $\text{Fix}(kP_D)$ is bounded, i.e., there exists $\ell \in \mathbb{N}$ such that for every component E of $SB(P_D)$, and every $k \geq 0$, the coefficient of E in $\text{Fix}(kP_D)$ is less than ℓ .*

All these facts are due to Zariski [28]. See [1, Section 2], especially Proposition 2.5 and its proof, for a modern account with notations consistent with ours.

Notation 2.10. Let S be a projective normal surface, and assume a big divisor D and an admissible flag $Y_\bullet = \{S \supset C \supset \{p\}\}$ are given. We fix the following notations for the rest of the paper.

- \diamond $D_t = D - tC$ for every $t \in \mathbb{R}$.
- \diamond P_t resp. N_t is the positive resp. negative part of the Zariski decomposition $D_t = P_t + N_t$ of D_t .
- \diamond $\nu = \nu_C(D) \in \mathbb{Q}$ is the coefficient of C in N_0 .

- ◇ $\mu = \mu_C(D) = \sup\{t > 0 : D - tC \text{ is big}\}$.
- ◇ For every (irreducible, reduced) curve C_i in the support of N_μ ,

$$t_{C_i} = t_{C_i}(D, C) = \inf\{t \in \mathbb{R} : \text{ord}_E(N_t) > 0\}.$$

Note that the support of every N_t with $0 \leq t \leq \mu$ is contained in the support of N_μ .

- ◇ A vertex $p = (t, s)$ of the polygon $\Delta_{Y_\bullet}(D)$ is called *leftmost*, *rightmost*, or *interior* when $t = \nu$, $t = \mu$, or $\nu < t < \mu$, respectively.

3. Upper bound for the number of vertices

One might expect that Newton–Okounkov bodies with respect to flags living in birational models of S with larger Picard number could have more vertices. However, extensive computations on blowups of \mathbb{P}^2 (see [6, 14]) have shown that this is not the case for $S = \mathbb{P}^2$ and suggest that a bound similar to the ones above might hold not only for smooth flags, but for arbitrary Newton–Okounkov polygons. In this section we prove Theorem A#, to this effect. Simultaneously, we shall obtain a sharp bound that extends the bound of [27] for ample divisor classes to arbitrary big divisor classes.

The main step towards the proof is the observation that, in the association of negative curves to vertices of the polygon, described above, those curves in $\text{Null}(D)$ do not contribute vertices:

Lemma 3.1. *Let D be a big and nef divisor on a smooth surface S , and consider the admissible flag $Y_\bullet := \{S \supset C \supset \{p\}\}$ on S . Let us write as above $D_t = D - tC$ with Zariski decomposition $D_t = P_t + N_t$. Let (t, s) be an interior vertex of $\Delta_{Y_\bullet}(D)$ (i.e., with $0 < t < \mu_C(D)$). Then there exists a curve F_i , belonging neither to $\text{Null}(D)$ nor to N_t , with $t_{F_i} = t$.*

Proof: By Theorem 2.7 we know that the supports of $N_{t+\varepsilon}$ and N_t differ. Let

$$N_{t+\varepsilon} = a_1 F_1 + \cdots + a_k F_k$$

be the negative part of $D_{t+\varepsilon} = D - (t + \varepsilon)C$ for $|\varepsilon|$ small enough, where $a_i = a_i(\varepsilon)$ is a continuous function, piecewise linear, with a point of non-differentiability at $\varepsilon = 0$ (at least for *some* a_i). Assume that F_i, \dots, F_k are the irreducible components of $N_{t+\varepsilon}$ not contained in N_t , i.e., $a_j(0) = 0$ if and only if $j \geq i$. We need to prove that at least one of these does not belong to $\text{Null}(D)$.

By continuity of the Zariski decomposition (see Proposition 2.4), no curve H satisfying $P_t \cdot H > 0$ belongs to the negative part $N_{t+\varepsilon}$ for small ε . Therefore, the components F_i, \dots, F_k all belong to $\text{Null}(P_t)$. Next we claim that at least one curve in F_i, \dots, F_k intersects C or a component of N_t . Indeed, if we assume by way of contradiction that none of them does, then $N'_{t+\varepsilon} := a_1 F_1 + \cdots + a_{i-1} F_{i-1}$ satisfies

$$(D_{t+\varepsilon} - N'_{t+\varepsilon}) \cdot F_j = \begin{cases} (D_{t+\varepsilon} - N_{t+\varepsilon}) \cdot F_j = 0 & \text{for } j < i, \\ D \cdot F_j \geq 0 & \text{for } j \geq i; \end{cases}$$

and for any curve E not among the F_j , we have that $(D_{t+\varepsilon} - N'_{t+\varepsilon}) \cdot E \geq P_{t+\varepsilon} \cdot E \geq 0$, hence $D_{t+\varepsilon} - N'_{t+\varepsilon}$ is nef, contradicting the minimality of $N_{t+\varepsilon}$ in the Zariski decomposition.

So we can assume without loss of generality that F_i intersects either C or a component of N_t . Then F_i does not belong to $\text{Null}(D)$, because if it did, we would have

$$0 = P_t \cdot F_i = (D - tC - N_t) \cdot F_i = -(tC + N_t) \cdot F_i < 0,$$

a contradiction. □

Theorem 3.2 (Theorem A#). *Let S be a normal projective algebraic surface, and D a big Cartier divisor on S . Then for every rank 2 valuation v on $K(S)$, the poly-gon $\Delta_v(D)$ has at most $2\rho_D(S) + 2$ vertices.*

Proof: Since S is normal, for every resolution of singularities $\pi: \tilde{S} \rightarrow S$ one has isomorphisms

$$H^0(\mathcal{O}_{\tilde{S}}(k\pi^*(D))) \cong H^0(\mathcal{O}_S(kD))$$

for all k . The pullback divisor $\pi^*(D)$ on \tilde{S} is therefore big, and the valuations of $K(\tilde{S}) = K(S)$ applied to sections of $\mathcal{O}_S(kD)$ and $\mathcal{O}_{\tilde{S}}(k\pi^*(D))$ are the same, so $\Delta_v(\pi^*(D)) = \Delta_v(D)$. On the other hand, $\rho_{\pi^*(D)}(\tilde{S}) = \rho_D(S)$ by definition. So we can replace S with \tilde{S} and assume that S is smooth.

By the same token, since v is given by a smooth flag $Y_\bullet = \{\tilde{S} \supset C \supset \{p\}\}$, on some blowup $\pi: \hat{S} \rightarrow S$, we may replace S with \hat{S} and assume that v is the flag valuation of an admissible flag.

Now let $D = P_D + N_D$ be the Zariski decomposition of D . The body $\Delta_v(D)$ is a translate of the body $\Delta_v(P_D)$ (a well-known fact which holds even in higher dimension; see [19, Theorem C, 3] together with Remark 2.1). Moreover, $\rho_D(S) = \rho_{P_D}(S)$; so we can replace D with P_D and assume that D is nef.

Write as above $D_t = D - tC$ and let $D_t = P_t + N_t$ be its Zariski decomposition for $t \in [0, \mu]$. By continuity of the Zariski decomposition, N_μ has the same support as $N_{\mu-\varepsilon}$ for $\varepsilon > 0$ small enough. Since $P_{\mu-\varepsilon}$ is big and nef, by the index theorem the number of components of $\text{Null}(P_{\mu-\varepsilon})$ (which includes all components of N_μ) is at most $\rho(S) - 1$. For every component E of $\text{Null}(D)$,

$$0 \leq P_{\mu-\varepsilon} \cdot E = D \cdot E - \mu - \varepsilon C \cdot E - N_{\mu-\varepsilon} \cdot E = (-\mu - \varepsilon) C \cdot E - N_{\mu-\varepsilon} \cdot E,$$

so E is either equal to C , to a component of $N_{\mu-\varepsilon}$, or to a component of $\text{Null}(P_{\mu-\varepsilon})$. Since all components of $N_{\mu-\varepsilon}$ are components of $\text{Null}(P_{\mu-\varepsilon})$, either $E = C$ or E is a component of $\text{Null}(P_{\mu-\varepsilon})$.

We now distinguish two cases. If C is not a component of $\text{Null}(D)$, the number of components of N_μ which are not components of $\text{Null}(D)$ is bounded above by

$$\rho(S) - 1 - \#\text{Null}(D) = \rho_D(S) - 1,$$

so the number of interior vertices in $\Delta_v(D)$ is bounded above by $2\rho_D(S) - 2$, and the same argument as in the proof of Corollary 2.8 implies that the total number of vertices is at most $2\rho_D(S) + 1$. On the other hand, if $C = E$ belongs to $\text{Null}(D)$, the number of components of N_μ which are not components of $\text{Null}(D)$ is bounded above by

$$\rho(S) - 1 - (\#\text{Null}(D) - 1) = \rho_D(S),$$

and the number of interior vertices in $\Delta_v(D)$ is bounded above by $2\rho_D(S)$. Moreover, again the same argument as in the proof of Corollary 2.8 shows that, if the number of interior vertices equals $2\rho_D(S)$, then there is just one rightmost vertex, and since C is a component of $\text{Null}(D)$ we have $\beta(0) - \alpha(0) = D \cdot C = 0$; therefore there is just one leftmost vertex, for a total number of vertices equal to $2\rho_D(S) + 2$. This completes the proof. \square

4. Newton–Okounkov bodies with respect to D -positive flags

For the case of an *ample* divisor D , it was determined in [27] the possible numbers of vertices of Newton–Okounkov bodies of D with respect to *smooth* flags. In this section we extend this result to bodies of *big and nef* divisors on smooth surfaces, with respect to smooth flags whose divisorial part intersects D positively. Note that

the latter hypothesis is automatically satisfied if D is ample. As a consequence, we can also deal with singular proper flags for ample divisors.

We begin by recalling the result of [27]. Let S be a smooth projective surface. For a configuration of curves $\mathcal{N} = (C_1, \dots, C_k)$ with negative definite intersection matrix, let $mc(\mathcal{N})$ denote the largest number of irreducible components of a connected subconfiguration of \mathcal{N} , and define

$$mv(\mathcal{N}) = \begin{cases} k + mc(\mathcal{N}) + 4 & \text{if } k < \rho(S) - 1, \\ k + mc(\mathcal{N}) + 3 & \text{if } k = \rho(S) - 1. \end{cases}$$

Then the maximum number of vertices of a Newton–Okounkov body of an ample divisor D on S is the maximum of all $mv(\mathcal{N})$ for \mathcal{N} with negative definite intersection matrix [27, Theorem 1.1].

Now assume D is merely big and nef rather than ample, and let $\text{Null}(D) = \{E_1, \dots, E_\ell\}$. After the results of the previous sections, it should not be a surprise that $\text{Null}(D)$ needs to be taken into account to define a version of mv appropriate for the present setting.

Definition 4.1. For a configuration of curves $\mathcal{N} = (C_1, \dots, C_k, E_1, \dots, E_\ell)$ with negative definite intersection matrix that contains all curves of $\text{Null}(D)$, let $mc_D(\mathcal{N})$ denote the largest number of irreducible components *not in* $\text{Null}(D)$ which belong to a single connected component of \mathcal{N} . Note that the connected component may include irreducible components belonging to $\text{Null}(D)$, but these do not contribute to mc_D . Then we define

$$mv_D(\mathcal{N}) = \begin{cases} k + mc_D(\mathcal{N}) + 4 & \text{if } k < \rho_D(S) - 1, \\ k + mc_D(\mathcal{N}) + 3 & \text{if } k = \rho_D(S) - 1. \end{cases}$$

We will prove below that the maximum number of vertices of a Newton–Okounkov body of D with respect to a proper flag whose divisorial part intersects D positively is

$$(2) \quad mv(D) := \max\{mv_D(\mathcal{N}) : \mathcal{N} \supseteq \text{Null}(D), \text{ negative definite}\}.$$

It may be possible to obtain a Newton–Okounkov body $\Delta_{Y_\bullet}(D)$ with more than $mv(D)$ vertices using flags $Y_\bullet = \{S \supset C \supset \{p\}\}$ whose divisorial part C belongs to $\text{Null}(D)$; see Example 5.2.

If D is a big but not necessarily nef divisor, we let $mv(D) = mv(P_D)$, where P_D is the positive part in the Zariski decomposition $D = P_D + N_D$.

Lemma 4.2. *Let S be a smooth projective surface, D a big and nef divisor on S , and $\pi_p: \tilde{S} \rightarrow S$ be the blowup centered at a point $p \in S$. Then $mv(D) = mv(\pi_p^*(D))$.*

Proof: First we observe that if $\text{Null}(D) = \{E_1, \dots, E_\ell\}$, then we have $\text{Null}(\pi_p^*(D)) = \{\tilde{E}_1, \dots, \tilde{E}_\ell, E_p\}$, where for any curve E on S we will denote by \tilde{E} its strict transform in \tilde{S} .

Second, a configuration of curves $\mathcal{N} = (C_1, \dots, C_k, E_1, \dots, E_\ell)$ has negative definite intersection matrix on S if and only if $\tilde{\mathcal{N}} = (\tilde{C}_1, \dots, \tilde{C}_k, \tilde{E}_1, \dots, \tilde{E}_\ell, E_p)$ has negative definite intersection matrix on \tilde{S} . Indeed, there is an obvious equality of spans

$$\langle \tilde{C}_1, \dots, \tilde{C}_k, \tilde{E}_1, \dots, \tilde{E}_\ell, E_p \rangle = \langle \pi_p^*(C_1), \dots, \pi_p^*(C_k), \pi_p^*(E_1), \dots, \pi_p^*(E_\ell), E_p \rangle$$

and, since E_p is orthogonal to all pullbacks and has negative self-intersection, the intersection form on the span $\langle \pi_p^*(C_1), \dots, \pi_p^*(C_k), \pi_p^*(E_1), \dots, \pi_p^*(E_\ell), E_p \rangle$ is negative definite if and only if it is so on $\langle C_1, \dots, C_k, E_1, \dots, E_\ell \rangle$.

Moreover, the connected components of \mathcal{N} which contain curves not in $\text{Null}(D)$ are in bijection with the connected components of $\tilde{\mathcal{N}}$ which contain curves not in $\text{Null}(\pi_p^*(D))$.

Then, for every configuration $\mathcal{N} = (C_1, \dots, C_k, E_1, \dots, E_\ell)$ with $mv_D(\mathcal{N}) = mv(D)$, the configuration of strict transforms $\tilde{\mathcal{N}} := (\tilde{C}_1, \dots, \tilde{C}_k, \tilde{E}_1, \dots, \tilde{E}_\ell, E_p)$ has $mv_{\pi_p^*(D)}(\tilde{\mathcal{N}}) = mv_D(\mathcal{N})$; this shows that $mv(\pi_p^*(D)) \geq mv(D)$. Finally, for every configuration $\tilde{\mathcal{N}} = (\tilde{C}_1, \dots, \tilde{C}_k, \tilde{E}_1, \dots, \tilde{E}_\ell, E_p)$ with $mv_{\pi_p^*(D)}(\tilde{\mathcal{N}}) = mv(\pi_p^*(D))$, the configuration of push-forwards $\mathcal{N} = (C_1, \dots, C_k, E_1, \dots, E_\ell)$ has $mv_D(\mathcal{N}) = mv_{\pi_p^*(D)}(\tilde{\mathcal{N}})$; this proves the inequality $mv(D) \geq mv(\pi_p^*(D))$, and the proof is complete. \square

By Lemma 4.2, mv is a birational invariant, and this allows us to extend the definition to possibly singular surfaces. If D is a big and nef divisor on a normal projective surface S , then we define $mv(D)$ to be equal to $mv(\pi^*(D))$, where $\pi: \tilde{S} \rightarrow S$ is a resolution of singularities of S . By the lemma and the fact that two arbitrary resolutions are dominated by some smooth model, $mv(\pi^*(D))$ does not depend on the chosen resolution π .

Lemma 4.3. *Let B be a big and nef \mathbb{Q} -divisor on a smooth surface S . For every integer m large and divisible enough, $C_m = mB - \text{Fix}(mB)$ is a big and nef effective divisor with $\text{Null}(C_m) \subset \text{Null}(B)$.*

Proof: Replacing B with a suitable multiple, we may assume that B is an effective divisor on S . Then $\text{Fix}(B)$ is an effective divisor $a_1F_1 + \dots + a_kF_k$ such that for every positive integer m , the support of $\text{Fix}(mB)$ is contained in $\text{Fix}(B)$. Moreover, by Lemma 2.9(iv), the coefficients of the components F_i of $\text{Fix}(B)$ in $\text{Fix}(mB)$ are bounded by a constant a independent of m , i.e., $\text{Fix}(mB) \leq a(F_1 + \dots + F_k)$ for every m . As the big cone is open, for m large enough the divisor

$$Z_m = mB - a(F_1 + \dots + F_k)$$

is big and effective, so $C_m \geq Z_m$ is big and effective as well. On the other hand, $\text{Fix}(C_m) = 0$ by definition of C_m , so C_m is nef.

Now fix m_0 such that Z_{m_0} is big and effective, and let $Z_{m_0} = P + N$ be its Zariski decomposition. For every m , $E_m = C_m - Z_m$ is effective and supported on $\text{Fix}(B)$, so if $m \geq m_0$, we can write

$$C_m = E_m + Z_m = E_m + P + N + (m - m_0)B.$$

For every curve E which is not a component of $\text{Fix}(B)$ or $\text{Null}(P)$, one has $E_m \cdot E \geq 0$, $N \cdot E \geq 0$ (because every component of N belongs to $\text{Null}(P)$), $B \cdot E \geq 0$, and $P \cdot E > 0$, so $C_m \cdot E > 0$. It follows that every curve in $\text{Null}(C_m)$ for m large enough is either a component of $\text{Fix}(B)$ or $\text{Null}(P)$, which is a finite collection of curves.

Now to show that $\text{Null}(C_m) \subset \text{Null}(B)$ for m large enough, observe that, for every curve E not in $\text{Null}(B)$,

$$C_m \cdot E = mB \cdot E - \text{Fix}(mB) \cdot E \geq m(B \cdot E) - a \left(\sum_{i=1}^k F_i \cdot E \right)$$

is positive for m larger than some integer m_E . It follows that the claim is satisfied by every integer m larger than m_0 and the finitely many numbers m_E , where E is a component of $\text{Fix}(B)$ or $\text{Null}(P)$ not in $\text{Null}(B)$. \square

Lemma 4.4. *Let S be a smooth projective surface, and D a big and nef divisor on S with $\text{Null}(D) = \{E_1, \dots, E_\ell\}$. Let C_1, \dots, C_k be irreducible curves on S such that the intersection matrix of $\mathcal{N} = (C_1, \dots, C_k, E_1, \dots, E_\ell)$ is negative definite. Then there is an irreducible curve C with $\text{Null}(C) \subseteq \text{Null}(D)$ with the following properties (using Notation 2.10).*

- (i) C intersects each C_i , $i = 1, \dots, k$, in at least two points.
- (ii) For every t with $D_t = D - tC$ pseudoeffective, the negative part N_t of its Zariski decomposition is supported on \mathcal{N} .
- (iii) Every C_i , $i = 1, \dots, k$, is a component of N_μ .
- (iv) $t_{C_1} < \dots < t_{C_k}$.

Moreover, C can be chosen in the linear span of D , the C_i , and the E_j , and satisfying $C \cdot D > 0$.

Proof: This proof is adapted to the present setting from [27, Lemma 5.3].

We will prove by induction on k that there are positive rational numbers a_1, \dots, a_k such that, denoting by

$$D - a_1C_1 - \dots - a_kC_k = P + N$$

the Zariski decomposition of the divisor on the left hand side, then P is big and nef with $\text{Null}(P) \subseteq \text{Null}(D)$, and for every sufficiently large m and every irreducible curve $C \in |mP - \text{Fix}(mP)|$ properties (iii) and (iv) are satisfied. By Lemma 2.9 and Bertini's theorem, for m large and divisible enough there are irreducible curves in $|mP - \text{Fix}(mP)|$ that intersect D positively and each C_i in at least two points, so by Lemma 4.3 we shall be done.

If $k = 1$, choose a positive integer a such that the divisor class $B = D - (1/a)C_1$ is big and its Zariski decomposition $B = N + P$ satisfies $\text{Null}(P) = \text{Null}(D)$. This is certainly possible: on one hand, because the big cone is open, B is big for $1/a$ small enough; on the other hand, by the continuity of the Zariski decomposition and [1, Proposition 1.3], for $1/a$ small enough we have $N \cup \text{Null}(P) \subseteq \text{Null}(D)$, and every component E_j of $\text{Null}(D)$ which does not satisfy $C_1 \cdot E_j = 0$ belongs to N , so $N \cup \text{Null}(P) \supseteq \text{Null}(D)$.

Then for every m and every $C \in |mP - \text{Fix}(mP)|$,

$$D_{1/m} = D - (1/m)C = (1/a)C_1 + N + (1/m)\text{Fix}(mP)$$

is exactly equal to $N_{1/m}$, as it is an effective divisor with negative definite intersection matrix because $\text{supp}(N + (1/m)\text{Fix}(mP)) \subseteq \text{Null}(mP) = \text{Null}(D)$ since P is nef (this follows from [28, Theorem 8.1]). Therefore $0 < t_{C_1} < 1/m$, and we are done.

Now assume the claim is true for $\mathcal{N}' = (C_1, \dots, C_{k-1}, E_1, \dots, E_\ell)$, and let a, a_1, \dots, a_{k-1} be positive rational numbers such that the Zariski decomposition

$$B' = D - a_1C_1 - \dots - a_{k-1}C_{k-1} = P' + N'$$

has P' big with $\text{Null}(P') = \text{Null}(D)$ and moreover

- if we denote by N'_t the negative part of the Zariski decomposition of $D'_t = D - tP'$, then $\sup\{t \in \mathbb{Q} : C_i \text{ is not contained in } N'_t\}$ is a finite positive real number t'_{C_i} , and
- $t'_{C_1} < \dots < t'_{C_{k-1}}$.

In this case, since P' is big and nef, Wilson's theorem ([21, Theorem 2.3.9]) guarantees that, for $m \gg 0$ and every $C' \in |mP' - \text{Fix}(mP')|$, the numbers t_{C_i} such that C_i is a component of N_t for $t > t_{C_i}$ satisfy $t_{C_1} < \dots < t_{C_{k-1}} < 1$.

Moreover, for every $t \in [0, 1]$, we have $D_t = (1-t)D + t(ma_1C_1 + \cdots + ma_{k-1}C_{k-1} + mN' + \text{Fix}(mP'))$, with $(1-t)D$ nef and $ma_1C_1 + \cdots + ma_{k-1}C_{k-1} + \text{Fix}(mP')$ effective, so by the extremality properties of the Zariski decomposition it follows that

$$N_t \leq t(ma_1C_1 + \cdots + ma_{k-1}C_{k-1} + N' + \text{Fix}(mP'))$$

(with equality if and only if $t = 1$). In particular, all components of N_t are among the C_i or the E_j .

Choose rational numbers s_i with $0 = s_0 < t'_1 < s_1 < t'_2 < \cdots < s_{k-2} < t'_{k-1} < s_{k-1} < 1$. The choices made guarantee that the irreducible components of N'_{s_i} not in $\text{Null}(D)$ are exactly C_1, \dots, C_i , and $P'_{s_i} \cdot C_j > 0$ for all $i < j \leq k$. Therefore by continuity of the Zariski decomposition (see Proposition 2.4) there exist $\varepsilon_1, \dots, \varepsilon_k > 0$ such that, choosing $a_k \in \mathbb{Q}$ with $a_k \leq \varepsilon_i$, the irreducible components not in $\text{Null}(D)$ of the negative part in the Zariski decomposition of $D - s_i(P' - a_kC_k)$ are also exactly C_1, \dots, C_i . Thus by choosing a rational number a_k smaller than $\varepsilon_0, \dots, \varepsilon_{k-1}$ and setting

$$\begin{aligned} D - a_1C_1 - \cdots - a_kC_k &= P + N, \\ C &\in |mP - \text{Fix}(mP)|, \end{aligned}$$

again Wilson's theorem guarantees that for $m \gg 0$ the irreducible components not in $\text{Null}(D)$ of the negative part in the Zariski decomposition $P_{s_i} + N_{s_i}$ of $D - s_iC$ are C_1, \dots, C_i . On the other hand, in the Zariski decomposition $D - C = P_1 + N_1$ one has $N_1 = ma_1C_1 + \cdots + ma_kC_k + N + \text{Fix}(mP)$ and therefore

$$t_{C_{k-1}} < s_{k-1} < t_{C_k} < 1,$$

which completes the induction step.

For the last claim, we first observe that the class P thus constructed is a combination of D , the C_i , and the E_j , and hence (since $\text{Null}(P) = \text{Null}(D)$) the class of C' is also a combination of D , the C_i , and the E_j . Moreover, $C' \cdot D \geq 0$, and we can slightly modify P to obtain a P'' to guarantee that $C'' \cdot D > 0$ (and still satisfy the properties) by taking $P'' = P + \varepsilon D$ for ε small enough. \square

Lemma 4.5. *Let D be a big and nef divisor on S with $\text{Null}(D) = \{E_1, \dots, E_\ell\}$. Assume $k < \rho_D(S) - 1$ and C_1, \dots, C_k are irreducible curves on S such $\mathcal{N} = (C_1, \dots, C_k, E_1, \dots, E_\ell)$ is a maximal effective divisor with negative definite intersection matrix, i.e., such that there exists no curve C' distinct from $C_1, \dots, C_k, E_1, \dots, E_\ell$ with $\mathcal{N} + C'$ having negative definite intersection matrix. Then the irreducible curve C from Lemma 4.4 can be assumed to have a numerical class linearly independent of $\langle D, E_1, \dots, E_\ell, C_1, \dots, C_k \rangle$ in $N^1(S)_{\mathbb{R}}$.*

Proof: The argument from [27, Lemma 5.4] applies verbatim, with $B = P$. \square

Theorem 4.6. *On every smooth projective surface S , for every big and nef divisor D and every integer δ , $3 \leq \delta \leq mv(D)$, there exists a smooth flag $Y_\bullet = \{S \supset C \supset \{p\}\}$ with $D \cdot C > 0$ such that the Newton–Okounkov polygon $\Delta_{Y_\bullet}(D)$ has exactly δ vertices.*

Remark 4.7. By its definition in (2), the number $mv(D)$ only depends on the set $\text{Null}(D)$, i.e., all big and nef divisors D with the same set $\text{Null}(D)$ have Newton–Okounkov bodies with $3, \dots, mv(D)$ vertices.

Proof: Write as above $\text{Null}(D) = \{E_1, \dots, E_\ell\}$. Choose a configuration $\mathcal{N}_0 = (C_1, \dots, C_k, E_1, \dots, E_\ell)$ with negative definite intersection matrix such that $mv(D) = mv_D(\mathcal{N}_0)$.

Assume moreover that its components have been ordered in such a way that one connected component of \mathcal{N}_0 contains all C_i for $1 \leq i \leq mc(\mathcal{N}_0)$. By the definition of $mv(\mathcal{N})$, it is not restrictive to assume that \mathcal{N}_0 is maximal, i.e., there exists no curve C' with $\mathcal{N}_0 + C'$ having negative definite intersection matrix.

Then, if $k < \rho_D(S) - 1$, for every $mc_D(\mathcal{N}_0) \leq i \leq k$, $mv_D(C_1, \dots, C_i, E_1, \dots, E_\ell) = mv_D(\mathcal{N}_0) - k + i$, and for $0 \leq i \leq mc_D(\mathcal{N}_0)$, $mv_D(C_1, \dots, C_i, E_1, \dots, E_\ell) = mv_D(\mathcal{N}_0) - k - mc_D(\mathcal{N}_0) + 2i$. On the other hand, if $k = \rho_D(S) - 1$, then

$$mv_D(C_1 + \dots + C_i) = \begin{cases} mv_D(\mathcal{N}_0) - k + i + 1 & \text{for every } mc_D(\mathcal{N}_0) \leq i < k, \\ mv_D(\mathcal{N}_0) - k - mc_D(\mathcal{N}_0) + 2i + 1 & \text{for } 0 \leq i \leq mc_D(\mathcal{N}_0). \end{cases}$$

In any event,

$$\{3, \dots, mv(D)\} = \{mv_D(\mathcal{N}) : \mathcal{N} \leq \mathcal{N}_0\} \cup \{mv_D(\mathcal{N}) - 1 : \mathcal{N} \leq \mathcal{N}_0\}.$$

Therefore, it will be enough to prove that, for every choice of irreducible curves C_1, \dots, C_k on S such that the intersection matrix of $\mathcal{N} = (C_1, \dots, C_k, E_1, \dots, E_\ell)$ is negative definite,

- If \mathcal{N} is maximal, there is a flag Y_\bullet such that $\Delta_{Y_\bullet}(D)$ has $mv_D(\mathcal{N})$ vertices.
- If \mathcal{N} is non-zero or has fewer than $\rho_D(S) - 1$ components, there is a flag Y_\bullet such that $\Delta_{Y_\bullet}(D)$ has $mv_D(\mathcal{N}) - 1$ vertices.
- If \mathcal{N} is non-zero and has fewer than $\rho_D(S) - 1$ components, there is a flag Y_\bullet such that $\Delta_{Y_\bullet}(D)$ has $mv_D(\mathcal{N}) - 2$ vertices.

Moreover, in each case the divisorial of Y_\bullet can be assumed to intersect D positively.

In the case of a maximal \mathcal{N} with fewer than $\rho_D(S) - 1$ components, choose an irreducible curve C satisfying the conditions of Lemma 4.5, and let p be one of the intersection points of C and C_1 (unless $\mathcal{N} = 0$, in which case we choose an arbitrary $p \in C$). We claim that $D, Y_\bullet : S \supset C \supset \{p\}$ give a body with $mv_D(\mathcal{N})$ vertices. On the one hand, since $C \cdot D > 0$, it follows that $P_0 \cdot C > 0$, so $\nu = 0$ and $\Delta_{Y_\bullet}(D)$ has two leftmost vertices. Moreover, Theorem 2.7 ensures that $\Delta_{Y_\bullet}(D)$ has two interior vertices with first coordinate equal to the number t_i given by Lemmas 4.4 and 4.5 for $i = 1, \dots, mc_D(\mathcal{N})$, whereas it only has an upper interior vertex for $mc_D(\mathcal{N}) < i \leq k$. Finally, as the numerical class of C is independent of $(D, E_1, \dots, E_\ell, C_1, \dots, C_k)$ in $N^1(S)_\mathbb{R}$, by Corollary 2.8 $\Delta_{Y_\bullet}(D)$ has two rightmost vertices. So, the total number of vertices is $mv_D(\mathcal{N})$.

Now choose C verifying the conditions of Lemma 4.4, so that the class of C belongs to the span of D , the C_i , and the E_j . The shape of $\Delta_{Y_\bullet}(D)$ is as before, but with a single rightmost vertex; if \mathcal{N} has $\rho_D(S) - 1$ components (in particular \mathcal{N} is maximal), the total number of vertices is $mv_D(\mathcal{N})$, otherwise it is $mv_D(\mathcal{N}) - 1$.

Finally, if \mathcal{N} is non-zero, we can pick p differently, while keeping the same curve C that satisfies the conditions of Lemma 4.4. If $mc_D(\mathcal{N}) = 1$, we let p be a point of C not on \mathcal{N} , and if $mc_D(\mathcal{N}) > 1$, then we take p to be one of the intersection points of C with C_2 . In this way we obtain one lower point less, so if \mathcal{N} has $\rho_D(S) - 1$ components, the total number of vertices is $mv_D(\mathcal{N}) - 1$, otherwise it is $mv_D(\mathcal{N}) - 2$. \square

Corollary 4.8. *Let D be a big and nef divisor on the normal projective surface S . Then there is a proper flag $Y_\bullet = \{\tilde{S} \supset C \supset \{p\}\}$ over S such that $D \cdot \pi(C) > 0$ and Δ_{Y_\bullet} is a k -gon if and only if $3 \leq k \leq mv(D)$.*

Recall that Y_\bullet is a proper flag over S if $\pi: \tilde{S} \rightarrow S$ is a proper birational morphism, p is a smooth point of C and \tilde{S} , and $\pi(C)$ is a curve on S . Note that the corollary describes all Newton–Okounkov bodies of ample divisors with respect to proper flags.

Proof: By the definition of $mv(D)$, we may replace S with a resolution of singularities and assume S is smooth.

The existence part is then given by Theorem 4.6, so we only need to show that for every flag $Y_\bullet = \{\tilde{S} \supset C \supset \{p\}\}$ where $\pi: \tilde{S} \rightarrow S$ is a proper birational map between smooth surfaces, such that $D \cdot \pi(C) > 0$, the number of vertices of $\Delta_{Y_\bullet}(D)$ is bounded above by $mv(D)$.

By Lemma 4.2, it holds that $mv(D) = mv(\pi^*(D))$, and by the projection formula [11, 8.1.7], $\pi^*(D) \cdot C = D \cdot \pi(C) > 0$, whereas $\Delta_{Y_\bullet}(D) = \Delta_{Y_\bullet}(\pi^*(D))$ as already observed in the proof of Theorem 3.2. Therefore we may replace S with \tilde{S} and it is enough to prove the claim for smooth flags $Y_\bullet = \{S \supset C \supset \{p\}\}$.

Keeping Notation 2.10, let as above $\text{Null}(D) = \{E_1, \dots, E_\ell\}$, and let $\{C_1, \dots, C_k\}$ be the components of N_μ not in $\text{Null}(D)$. Choose a number $\mu' < \mu$ such that the support of $N_{\mu'}$ is the same as the support of N_μ . Since $P_{\mu'} = D - \mu'C - N_{\mu'} \leq D$ and by hypothesis $D \cdot C > 0$, for every component E_i of $\text{Null}(D)$ which is not a component of $N_{\mu'}$ we have $P_{\mu'} \cdot E_i = D \cdot E_i - \mu'C \cdot E_i - N_{\mu'} \cdot E_i \leq 0$, which, $P_{\mu'}$ being nef, implies $P_{\mu'} \cdot E_i = 0$. Therefore all curves in the configuration $\mathcal{N} = (C_1, \dots, C_k, E_1, \dots, E_\ell)$ have intersection number 0 with $P_{\mu'}$, which is big (because $\mu' < \mu$) so by the index theorem \mathcal{N} has negative definite intersection matrix.

Then by [27, Proposition 3.3] and Lemma 3.1, the number of vertices of $\Delta_{Y_\bullet}(D)$ is at most $mv_D(\mathcal{N})$. \square

5. Newton–Okounkov bodies with respect to D -orthogonal flags

It is a subtler problem to determine the number of vertices of the Newton–Okounkov bodies $\Delta_{Y_\bullet}(D)$, where the divisorial part of the flag Y_\bullet belongs to $\text{Null}(D)$. To begin with, there are only finitely many curves in $\text{Null}(D)$, which entails that there are only finitely many such bodies, and the facts explained in Theorem 4.6 and Remark 4.7 do not hold anymore.

Example 5.1. There exist surfaces S with divisors D, D' satisfying $\text{Null}(D) = \text{Null}(D')$ for which the numbers k such that some k -gon arises as a Newton–Okounkov body of D or D' with respect to flags with divisorial part in $\text{Null}(D)$ differ.

Let S be the blowup of $\mathbb{P}^1 \times \mathbb{P}^1$ at a point p . There are three curves of negative self-intersection on S , namely the exceptional divisor E_p and the strict transforms F_1, F_2 of the two rulings passing through p . Consider the divisors $D = 2E_p + F_1 + F_2$ and $D' = 3E_p + 2F_1 + F_2$. It is immediate that $\text{Null}(D) = \text{Null}(D') = \{E_p\}$. In addition, consider flags of the form $Y_\bullet = \{S \supset E_p \supset \{q\}\}$. According to Galindo, Monserrat, and Moreno-Ávila [12, Theorem 3.6], the first component v_1 of the valuation $v_{Y_\bullet} = (v_1, v_2)$ is a non-positive at infinity divisorial valuation, hence we can compute the Newton–Okounkov bodies of D resp. D' with respect to flags Y_\bullet using their explicit calculations presented in [13]. More precisely, we distinguish three cases:

- (1) If $q \in E_p \cap F_1$, then [13, Theorem 3.12] applies so that both $\Delta_{Y_\bullet}(D)$ and $\Delta_{Y_\bullet}(D')$ are quadrilaterals.
- (2) If $q \in E_p \cap F_2$, then [13, Theorem 3.13(a)] implies again that both $\Delta_{Y_\bullet}(D)$ and $\Delta_{Y_\bullet}(D')$ are quadrilaterals.
- (3) Finally, if $q \in E_p$ belongs to neither $E_p \cap F_1$ nor $E_p \cap F_2$, then the Newton–Okounkov body $\Delta_{Y_\bullet}(D)$ is a triangle, whereas $\Delta_{Y_\bullet}(D')$ is a quadrilateral by [13, Theorem 3.12].

The cases (1) and (3) are illustrated in Figure 1, whereas case (2) behaves in the same manner as (1). These results agree with Theorem 4.6.

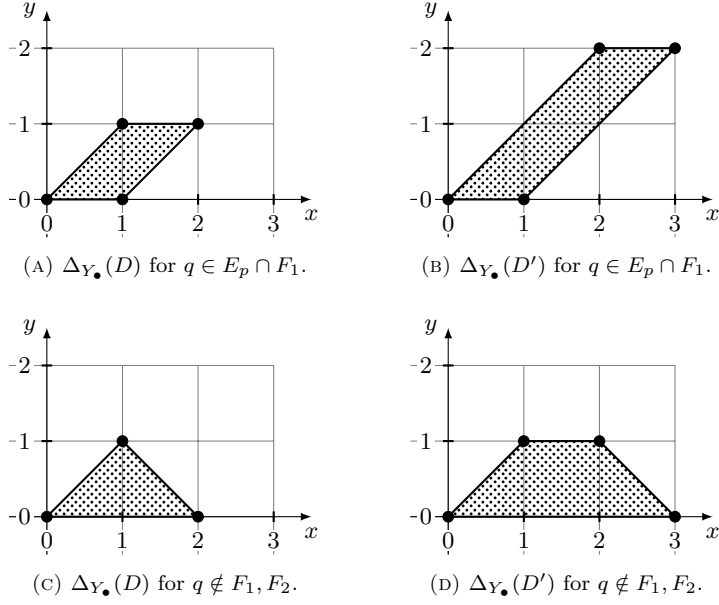


FIGURE 1. Newton–Okounkov bodies in Example 5.1.

Example 5.2. It is also possible to obtain Newton–Okounkov bodies of D with more than $mv(D)$ vertices by using flags whose divisorial part is orthogonal to D .

Let $S \rightarrow \mathbb{P}^2$ be a proper birational map (which can be factored as a composition of point blowups), and let D be the pullback to S of the class of a line. Obviously $\text{Null}(D)$ consists of all the exceptional components, so $\rho_D(S) = 1$ and only $\mathcal{N} = \text{Null}(D)$ satisfies the conditions of Definition 4.1. Therefore $mv(D) = mv_D(\mathcal{N}) = 3$. However, it was shown in [6] and [14] that there are choices of S that support flags Y_\bullet such that $\Delta_{Y_\bullet}(D)$ is a quadrilateral.

Note that in Example 5.2 the maximum number $2\rho_D(S) + 2 = 4$ of vertices allowed by Theorem 3.2 is attained. However, for many surfaces S and divisors D (e.g., the line on \mathbb{P}^2 , but also its pullback to most of the blowups considered in [6] and [14]) there is no flag Y_\bullet such that $\Delta_{Y_\bullet}(D)$ has $2\rho_D(S) + 2$ vertices. A general upper bound which takes into account the geometry on S can be given following the ideas of the previous section using the following definition:

Definition 5.3. Assume D is a big and nef divisor such that $\text{Null}(D) = \{E_1, \dots, E_\ell\}$ is non-empty. For a configuration of curves $\mathcal{N} = (C_1, \dots, C_k, E_{i_1}, \dots, E_{i_{\ell-1}})$ with negative definite intersection matrix that contains all curves of $\text{Null}(D)$ but one, let $mc_D(\mathcal{N})$ denote the largest number of irreducible components *not* in $\text{Null}(D)$ which belong to a single connected component of \mathcal{N} . Then we define

$$mv_D(\mathcal{N}) = \begin{cases} k + mc_D(\mathcal{N}) + 3 & \text{if } k < \rho_D(S), \\ k + mc_D(\mathcal{N}) + 2 & \text{if } k = \rho_D(S), \end{cases}$$

and

$$mv^{\text{Null}}(D) := \max\{mv_D(\mathcal{N}) : \mathcal{N} \text{ negative definite, } \#(\text{Null}(D) \setminus \mathcal{N}) = 1\}.$$

Proposition 5.4. *On every smooth projective surface S , for every big and nef divisor D , and every admissible flag Y_\bullet on S whose divisorial part belongs to $\text{Null}(D)$, the polygon $\Delta_{Y_\bullet}(D)$ has at most $mv^{\text{Null}}(D)$ vertices.*

The proof of the upper bound in Corollary 4.8 carries over to this setting, and we omit it for brevity. On the other hand, the analogous result to the existence part (i.e., Theorem 4.6) does not hold for D -orthogonal flags.

Invariance under pullback by birational morphisms (i.e., Lemma 4.2) also fails for mv^{Null} . However, this added complexity will turn out to be useful in the next section.

Example 5.5. Let $\pi: S \rightarrow \mathbb{P}^2$ be the blowup centered at a point p , and let D be the pullback of a line by π . Then $\text{Null}(D)$ consists of the exceptional divisor E_p only, so $mv^{\text{Null}}(D) = 3$. However, there exist further blowups $\eta: \tilde{S} \rightarrow S$ that allow flags Y_\bullet such that $\Delta_{Y_\bullet}(\eta^*(D))$ is a quadrilateral, cf. Example 5.2. Therefore, Proposition 5.4 implies

$$mv^{\text{Null}}(\eta^*(D)) \geq 4 > mv^{\text{Null}}(D).$$

6. Infinitesimal flags with maximal negative configurations

In the same way that results on Newton–Okounkov bodies with respect to D -positive admissible flags were used in Section 4 to describe the Newton–Okounkov bodies with respect to proper (not necessarily admissible) flags, we next apply our results on Newton–Okounkov bodies with respect to D -orthogonal flags to describe Newton–Okounkov bodies with respect to infinitesimal flags. However, working with infinitesimal flags provides additional flexibility (replacing S with a suitable blowup) which will turn out to be enough to characterize surfaces of Picard number 1 in terms of Newton–Okounkov bodies.

It will simplify the presentation to use the notion of *relative Zariski decomposition* inspired by [8, Section 8].

Definition 6.1. Let D be an arbitrary divisor on the smooth projective surface S , and let $\mathcal{N} = \{E_1, \dots, E_k\}$ be a configuration of (irreducible, reduced) curves with negative definite intersection matrix. There is a unique effective \mathbb{Q} -divisor $N_{\mathcal{N}} = a_1 E_1 + \dots + a_k E_k$ satisfying

- (i) $(D - N_{\mathcal{N}}) \cdot E_i \geq 0$ for all $i = 1, \dots, k$.
- (ii) $(D - N_{\mathcal{N}}) \cdot E_i = 0$ for every i such that $a_i > 0$.

We then set $P_{\mathcal{N}} = D - N_{\mathcal{N}}$ and call $D = P_{\mathcal{N}} + N_{\mathcal{N}}$ the Zariski decomposition of D with respect to \mathcal{N} .

The properties of the relative Zariski decomposition that we will need are contained in the following lemmas:

Lemma 6.2. *If D is pseudoeffective and $D = N + P$ is its Zariski decomposition, then for every negative definite configuration \mathcal{N} , the relative Zariski decomposition $D = N_{\mathcal{N}} + P_{\mathcal{N}}$ satisfies $N_{\mathcal{N}} \leq N$.*

Lemma 6.3. *Let $D = N_{\mathcal{N}} + P_{\mathcal{N}}$ be the relative Zariski decomposition of a divisor D with respect to some negative definite configuration \mathcal{N} . If $P_{\mathcal{N}}^2 > 0$, then D is big.*

Both lemmas are elementary and we leave their proofs to the reader.

Fix a big and nef divisor D on a smooth projective surface S , with relative Picard number $\rho_D(S) = \rho$. Our construction starts from an irreducible nodal curve C on S , which for simplicity we assume has positive self-intersection, and we will show that there is an infinitesimal flag Y_\bullet centered at the node p of the curve C , such that C induces two internal vertices on the Newton–Okounkov body of D with respect

to Y_\bullet . This allows us to prove that D admits Newton–Okounkov bodies with four vertices, and also with five if $\rho > 1$.

For each $k > 0$, let us denote by $S_k \rightarrow S$ the composition of the k successive blowups

$$S_k \longrightarrow S_{k-1} \longrightarrow \cdots \longrightarrow S_1 \longrightarrow S_0 = S,$$

where $S_1 \rightarrow S_0$ is the blowup with center at p , $S_2 \rightarrow S_1$ is the blowup with center at one of the two points in $C_1 \cap E_1$, and for every $i > 1$, $S_{i+1} \rightarrow S_i$ is the blowup with center at the point $p_i = C_i \cap E_i$, where C_i denotes the strict transform of C on S_i , and E_i stands for the exceptional divisor of $S_i \rightarrow S_{i-1}$.

Slightly abusing notations, for every divisor Z on a surface X we shall denote by the same symbol Z its pullback by a blowup morphism (the spirit of this abuse of notation is consistent with a birational point of view which thinks of Z as a b -divisor; see [7]).

Denoting by $E_{i,k}$ the strict transform on S_k of the exceptional divisor E_i , the pullback of C on S_k (which—as said above—we continue to write C for simplicity of notation) is $C_k + 2E_{1,k} + 3E_{2,k} + \cdots + (k+1)E_{k,k}$ and the pullback of E_i (written E_i for simplicity) is $E_{i,k} + E_{i+1,k} + \cdots + E_{k,k}$.

All divisors on S_k we shall be concerned with belong to numerical equivalence classes which are linear combinations of $[D], [C], [E_1], \dots, [E_k]$. They are moreover orthogonal to $E_{i,k}$ for $i = 2, \dots, k-1$ and thus they actually belong to the span $\langle [D], [C], [E], [E_1] \rangle$, where $E = \sum_{i=1}^k iE_{i,k}$. In particular, $C_k = C - E - E_1$. Note also that C_k intersects E at the components $E_{1,k}$ and $E_{k,k}$ of E .

Proposition 6.4. *Let D be a big and nef divisor on a smooth projective surface S , and C an irreducible curve on S with a node at the point p and not belonging to $\text{Null}(D)$. For every integer $k > 0$ let $S_k \rightarrow S$ be the composition of blowups determined by C as described above, and consider $Y_\bullet = \{S_k \supset E_k \supset \{p_k\}\}$, where $p_k = C_k \cap E_k$. If $k > (D \cdot C)/\sqrt{D^2} - 1$, then $\Delta_{Y_\bullet}(D)$ has at least two internal vertices.*

Note that choosing $p'_k = E_k \cap E_{k-1,k}$ would lead to the same conclusion.

Proof: The determination of $\Delta_{Y_\bullet}(D)$ relies on the understanding of the Zariski decomposition of $D_t = D - tE_k$, which we denote by $D_t = N_t + P_t$.

Clearly, the divisors $E_{i,k}$ for $i = 1, \dots, k-1$ belong to N_t as soon as $t > 0$. In fact, for $\mathcal{N} = \{E_1, \dots, E_{k-1}\}$, it is easy to check that the relative Zariski decomposition $D_t = P_{t,\mathcal{N}} + N_{t,\mathcal{N}}$ of D_t with respect to \mathcal{N} has negative part $N_{t,\mathcal{N}} = \frac{t}{k}E - tE_k$, and we know (by Lemma 6.2) that $N_t \geq N_{t,\mathcal{N}}$. Therefore, for every $t > 0$ such that D_t is pseudoeffective, either C_k is a component of N_t or $C_k \cdot P_{t,\mathcal{N}}$ is non-negative, which gives

$$0 \leq C_k \cdot \left(D_t - \left(\frac{t}{k}E - tE_k \right) \right) = (C - E - E_1) \cdot \left(D - \frac{t}{k}E \right) = D \cdot C - t \frac{k+1}{k}.$$

In other words, for $t > \frac{kD \cdot C}{k+1}$ such that D_t is pseudoeffective, C_k belongs to N_t .

On the other hand, by Lemma 6.3 we know that D_t is big whenever

$$0 < P_{t,\mathcal{N}}^2 = \left(D_t - \left(\frac{t}{k}E - tE_k \right) \right)^2 = \left(D - \frac{t}{k}E \right)^2 = D^2 - t^2/k.$$

The hypothesis $k > (D \cdot C)/\sqrt{D^2} - 1$ guarantees the existence of a $t > \frac{kD \cdot C}{k+1}$ such that $D^2 - t^2/k > 0$, i.e., with D_t big and therefore pseudoeffective. Hence for t large enough, C_k does belong to N_t , whereas for t close to zero C_k does not belong to N_t , because $D_0 \cdot C_k = D \cdot C_k > 0$. Thus there is a value t_{C_k} (less than or equal to $\frac{kD \cdot C}{k+1}$) such that

C_k belongs to N_t if and only if $t > t_{C_k}$, hence there are interior vertices on $\Delta_{Y_\bullet}(D)$ with abscissa t_{C_k} . Since C_k meets the configuration \mathcal{N} at the component $E_{1,k}$, N_t is connected, so the connected component of N_t to which C_k belongs is simply N_t , and it intersects E_k at two points (at least), namely $C_k \cap E_k$ and $E_{k-1} \cap E_k$. The first point is equal to p so C_k is contained in the connected component of N_t that passes through p and by statement (iv) of Theorem 2.7 the function $\alpha(t)$ has a point of non-differentiability at $t = t_0$, i.e., there is an internal lower vertex at $t = t_0$. The second intersection point $E_{k-1} \cap E_k$ of N_t with E_k is different from p , so the connected component containing C_k intersects E_k away from p and by statement (v) of Theorem 2.7 the function $\beta(t)$ has a point of non-differentiability at $t = t_0$ and there is an internal upper vertex at $t = t_0$. Thus $\Delta_{Y_\bullet}(D)$ has two internal vertices with abscissa t_{C_k} . \square

Theorem 6.5. *For every ample divisor D on a normal projective surface S there exist rank 2 valuations v such that $\Delta_v(D)$ has at least four vertices. If $\rho(S) > 1$, then there exist rank 2 valuations v such that $\Delta_v(D)$ has more than four vertices.*

Corollary 6.6. *Among all projective smooth surfaces S , those with Picard number 1 are characterized by the fact that for every big divisor D and every rank 2 valuation v of $K(S)$, the Newton–Okounkov polygon $\Delta_v(D)$ has at most four vertices.*

For the proof of Theorem 6.5 we need the following lemma.

Lemma 6.7. *Let D be an ample divisor and C a nodal curve on the surface S . Let k be an integer such that $k > (D \cdot C)/\sqrt{D^2} - 1$ and $D - \frac{D \cdot C}{k+1}C$ is ample. Consider the construction above of surfaces S_k with infinitesimal flags $Y_\bullet = \{S_k \supset E_k \supset \{p_k\}\}$, and suppose that $\Delta_{Y_\bullet}(D)$ has exactly four vertices. Then for every t such that $D_t = D - tE_k$ is pseudoeffective, the negative part N_t in the Zariski decomposition $D_t = P_t + N_t$ is supported on $\mathcal{N} = \{E_{1,k}, \dots, E_{k-1,k}, C_k\}$.*

Proof: We argue by contradiction. Assume that there is an irreducible component Z_k of N_μ not in \mathcal{N} . In particular, since D is ample, Z_k is the strict transform in S_k of a curve Z in S . Let us denote by Z_i the strict transform of Z in the sequence of blowups $S_i \rightarrow S_{i-1} \rightarrow \dots \rightarrow S_1 \rightarrow S_0 = S$.

By the results above, the four vertices of $\Delta_{Y_\bullet}(D)$ must be one leftmost vertex, one rightmost vertex, and two vertices with abscissa $t_C = \inf\{t : C_k \text{ is a component of } N_t\}$. Since $D_0 \cdot Z_k = D \cdot Z > 0$ because D is ample, Z_k does not belong to N_t for t near zero. Therefore $\inf\{t : Z_k \text{ is a component of } N_t\} = t_C$ (otherwise Z_k would create some additional vertex on $\Delta_{Y_\bullet}(D)$ with abscissa different from 0, $t_C \cdot \mu$), and in particular $P_{t_C} \cdot Z_k = P_{t_C} \cdot C_k = 0$. Moreover, for every $t < t_C$, the Zariski decomposition $D_t = P_t + N_t$ agrees with the Zariski decomposition of D_t relative to $\mathcal{N} \setminus \{C_k\}$ computed above, and in particular

$$P_t = D - \frac{t}{k}E \quad \text{for all } 0 \leq t \leq t_C.$$

The equality $P_{t_C} \cdot C_k = 0$ now allows us to determine $t_C = \frac{kD \cdot C}{k+1}$, which in turn implies a strong restriction on Z_k : let us denote by $m_{Z,i}$ the multiplicity at p_i of the strict transform Z_i of Z in S_i . Then we have

$$(3) \quad 0 = P_{t_C} \cdot Z_k = \left(D - \frac{kD \cdot C}{k+1}E \right) \cdot \left(Z - \sum_i m_{Z,i}E_i \right) = D \cdot Z - \frac{D \cdot C}{k+1} \sum_i m_{Z,i},$$

and on the other hand $0 \leq C_k \cdot C' = C_k \cdot Z_k = C \cdot Z - 2m_{Z,1} - \sum_{i=2}^k m_{Z,i}$; therefore we deduce that $\sum_{i=1}^k m_{Z,i} \leq C \cdot Z$, and equation (3) yields

$$0 \geq D \cdot Z - \frac{D \cdot C}{k+1} C \cdot Z = \left(D - \frac{D \cdot C}{k+1} C \right) \cdot Z.$$

But this is a contradiction, because $D - \frac{D \cdot C}{k+1} C$ is ample. \square

Proof of Theorem 6.5: The Newton–Okounkov bodies $\Delta_{Y_\bullet}(D)$ constructed in Proposition 6.4 have at least two internal vertices, a rightmost and a leftmost vertex, so they have at least four vertices.

Now assume $\rho(S) > 1$, and choose a nodal curve C with $C^2 > 0$ whose numerical equivalence class is not a rational multiple of $[D]$. We will show that in the construction above of an infinitesimal flag $Y_\bullet = \{S_k \supset E_k \supset \{p_k\}\}$ it is possible to choose k such that the Newton–Okounkov body $\Delta_{Y_\bullet}(D)$ has five or more vertices.

We argue by contradiction. Assume that for every $k \geq (D \cdot C)/\sqrt{D^2} - 1$ the polygon $\Delta_{Y_\bullet}(D)$ has exactly four vertices, namely one leftmost vertex, one rightmost vertex, and two vertices with abscissa $t_C = \inf\{t : C_k \text{ is a component of } N_t\}$. By Lemma 6.7 this implies that, if k is large enough, then for all t such that $D_t = D - tE_k$ is pseudoeffective, the negative part N_t in the Zariski decomposition $D_t = P_t + N_t$ is supported on $\mathcal{N} = \{E_{1,k}, \dots, E_{k-1,k}, C_k\}$, and hence P_t can be computed as $P_{t,\mathcal{N}}$. This is a straightforward (if somewhat cumbersome) linear algebra computation leading to

$$(4) \quad P_\mu = D - \frac{(k+1)\mu - k(D \cdot C)}{k^2 - k(C^2 - 1)} C - \frac{\mu C^2 - k(D \cdot C)}{k^2 - k(C^2 - 1)} E.$$

The assumption that there is a unique rightmost vertex implies that $P_\mu \cdot E_k = 0$. The facts that $D \cdot E_k = C \cdot E_k = 0$ and $E \cdot E_k = -1$, setting $P_\mu \cdot E_k = 0$, imply $\mu = k(D \cdot C)/C^2$. Substituting this value in (4) and taking into account the intersection numbers between D , C , and E , one gets

$$P_\mu^2 = \frac{1}{C^2} \det \begin{pmatrix} D^2 & D \cdot C \\ D \cdot C & C^2 \end{pmatrix},$$

which is negative by the index theorem. This is a contradiction, because P_μ is nef by definition. \square

Conjecture 6.8. *Let S be a smooth projective algebraic surface, and D a big divisor on S . There is a smooth projective surface \tilde{S} with a proper projective morphism $\pi: \tilde{S} \rightarrow S$ and a flag $Y_\bullet = \{\tilde{S} \supset E \supset \{p\}\}$ such that $\Delta_{Y_\bullet}(D)$ has exactly $2\rho + 2$ vertices.*

Note that the divisorial part E of the flag belongs necessarily to $\text{Null}(D)$. The use of a different model \tilde{S} provides “new” curves on $\text{Null}(D)$ to use in the flag, as was done above to prove Theorem 6.5, but it is a challenge to construct appropriate blowups $\pi: \tilde{S} \rightarrow S$ where one can guarantee the maximal number of internal vertices.

Acknowledgments

We are grateful to Alex Küronya and Carlos Jesús Moreno-Ávila for several discussions on the subject. The first and second authors wish to thank the UAB (Universitat Autònoma de Barcelona), where the first ideas of this paper came out, for the invitation and hospitality.

References

- [1] TH. BAUER, A. KÜRONYA, AND T. SZEMBERG, Zariski chambers, volumes, and stable base loci, *J. Reine Angew. Math.* **576** (2004), 209–233. DOI: 10.1515/crll.2004.090.
- [2] H. BLUM AND M. JONSSON, Thresholds, valuations, and K-stability, *Adv. Math.* **365** (2020), 107062, 57 pp. DOI: 10.1016/j.aim.2020.107062.
- [3] S. BOUCKSOM, Corps d’Okounkov (d’après Okounkov, Lazarsfeld–Mustață et Kaveh–Khovanskii), *Astérisque* **361** (2014), Exp. no. 1059, vii, 1–41.
- [4] S. BOUCKSOM AND H. CHEN, Okounkov bodies of filtered linear series, *Compos. Math.* **147**(4) (2011), 1205–1229. DOI: 10.1112/S0010437X11005355.
- [5] S. R. CHOI, J. PARK, AND J. WON, Okounkov bodies and Zariski decompositions on surfaces, *Bull. Korean Math. Soc.* **54**(5) (2017), 1677–1697. DOI: 10.4134/BKMS.b160686.
- [6] C. CILIBERTO, M. FARNIK, A. KÜRONYA, V. LOZOVANU, J. ROÉ, AND C. SHRAMOV, Newton–Okounkov bodies sprouting on the valuative tree, *Rend. Circ. Mat. Palermo (2)* **66**(2) (2017), 161–194. DOI: 10.1007/s12215-016-0285-3.
- [7] A. CORTI, ED., *Flips for 3-folds and 4-folds*, Oxford Lecture Ser. Math. Appl. **35**, Oxford University Press, Oxford, 2007. DOI: 10.1093/acprof:oso/9780198570615.001.0001.
- [8] S. D. CUTKOSKY AND V. SRINIVAS, On a problem of Zariski on dimensions of linear systems, *Ann. of Math. (2)* **137**(3) (1993), 531–559. DOI: 10.2307/2946531.
- [9] M. DUMNICKI, B. HARBOURNE, A. KÜRONYA, J. ROÉ, AND T. SZEMBERG, Very general monomial valuations of \mathbb{P}^2 and a Nagata type conjecture, *Comm. Anal. Geom.* **25**(1) (2017), 125–161. DOI: 10.4310/CAG.2017.v25.n1.a4.
- [10] X. FANG, G. FOURIER, AND P. LITTELMANN, On toric degenerations of flag varieties, in: *Representation Theory—Current Trends and Perspectives*, EMS Ser. Congr. Rep., European Mathematical Society (EMS), Zürich, 2017, pp. 187–232. DOI: 10.4171/171-1/8.
- [11] W. FULTON, *Intersection Theory*, Second edition, *Ergeb. Math. Grenzgeb. (3)* **2**, Springer-Verlag, Berlin, 1998. DOI: 10.1007/978-1-4612-1700-8.
- [12] C. GALINDO, F. MONSERRAT, AND C.-J. MORENO-ÁVILA, Non-positive and negative at infinity divisorial valuations of Hirzebruch surfaces, *Rev. Mat. Complut.* **33**(2) (2020), 349–372. DOI: 10.1007/s13163-019-00319-w.
- [13] C. GALINDO, F. MONSERRAT, AND C.-J. MORENO-ÁVILA, Seshadri-type constants and Newton–Okounkov bodies for non-positive at infinity valuations of Hirzebruch surfaces, *Quaest. Math.* **46**(11) (2023), 2367–2401. DOI: 10.2989/16073606.2022.2146020.
- [14] C. GALINDO, F. MONSERRAT, J. J. MOYANO-FERNÁNDEZ, AND M. NICKEL, Newton–Okounkov bodies of exceptional curve valuations, *Rev. Mat. Iberoam.* **36**(7) (2020), 2147–2182. DOI: 10.4171/rmi/1195.
- [15] M. HARADA AND K. KAVEH, Integrable systems, toric degenerations and Okounkov bodies, *Invent. Math.* **202**(3) (2015), 927–985. DOI: 10.1007/s00222-014-0574-4.
- [16] S.-Y. JOW, Okounkov bodies and restricted volumes along very general curves, *Adv. Math.* **223**(4) (2010), 1356–1371. DOI: 10.1016/j.aim.2009.09.015.
- [17] K. KAVEH AND A. G. KHOVANSKII, Mixed volume and an extension of intersection theory of divisors, *Mosc. Math. J.* **10**(2) (2010), 343–375, 479. DOI: 10.17323/1609-4514-2010-10-2-343-375.
- [18] K. KAVEH AND A. G. KHOVANSKII, Newton–Okounkov bodies, semigroups of integral points, graded algebras and intersection theory, *Ann. of Math. (2)* **176**(2) (2012), 925–978. DOI: 10.4007/annals.2012.176.2.5.
- [19] A. KÜRONYA AND V. LOZOVANU, Positivity of line bundles and Newton–Okounkov bodies, *Doc. Math.* **22** (2017), 1285–1302. DOI: 10.4171/DM/596.
- [20] A. KÜRONYA, V. LOZOVANU, AND C. MACLEAN, Convex bodies appearing as Okounkov bodies of divisors, *Adv. Math.* **229**(5) (2012), 2622–2639. DOI: 10.1016/j.aim.2012.01.013.
- [21] R. LAZARSFELD, *Positivity in Algebraic Geometry I. Classical Setting: Line Bundles and Linear Series*, *Ergeb. Math. Grenzgeb. (3)* **48**, Springer-Verlag, Berlin, 2004. DOI: 10.1007/978-3-642-18808-4.
- [22] R. LAZARSFELD, *Positivity in Algebraic Geometry II. Positivity for Vector Bundles, and Multiplier Ideals*, *Ergeb. Math. Grenzgeb. (3)* **49**, Springer-Verlag, Berlin, 2004. DOI: 10.1007/978-3-642-18810-7.
- [23] R. LAZARSFELD AND M. MUSTAȚĂ, Convex bodies associated to linear series, *Ann. Sci. Éc. Norm. Supér. (4)* **42**(5) (2009), 783–835. DOI: 10.24033/asens.2109.
- [24] A. OKOUNKOV, Brunn–Minkowski inequality for multiplicities, *Invent. Math.* **125**(3) (1996), 405–411. DOI: 10.1007/s002220050081.

- [25] K. RIETSCH AND L. WILLIAMS, Newton–Okounkov bodies, cluster duality, and mirror symmetry for Grassmannians, *Duke Math. J.* **168**(18) (2019), 3437–3527. DOI: 10.1215/00127094-2019-0028.
- [26] J. ROÉ, Local positivity in terms of Newton–Okounkov bodies, *Adv. Math.* **301** (2016), 486–498. DOI: 10.1016/j.aim.2016.05.028.
- [27] J. ROÉ AND T. SZEMBERG, On the number of vertices of Newton–Okounkov polygons, *Rev. Mat. Iberoam.* **38**(5) (2022), 1383–1397. DOI: 10.4171/rmi/1369.
- [28] O. ZARISKI, The theorem of Riemann–Roch for high multiples of an effective divisor on an algebraic surface, *Ann. of Math. (2)* **76**(3) (1962), 560–615. DOI: 10.2307/1970376.

Julio-José Moyano-Fernández

Universitat Jaume I, Campus de Riu Sec, Departament de Matemàtiques & Institut Universitari de Matemàtiques i Aplicacions de Castelló, 12071 Castellón de la Plana, Spain

E-mail address: moyano@uji.es

Matthias Nickel

Johann Wolfgang Goethe-Universität, Institut für Mathematik, Robert-Mayer-Straße 6–8, D-60325 Frankfurt am Main, Germany

E-mail address: nickel@math.uni-frankfurt.de

Joaquim Roé

Universitat Autònoma de Barcelona, Departament de Matemàtiques, 08193 Bellaterra (Barcelona), Spain

E-mail address: joaquim.roe@uab.cat

Received on July 29, 2021.

Accepted on April 17, 2024.