Combinatoire des droites et segments pour la visibilité 3D

Marc Glisse

Nancy-Université
Université Nancy 2

École doctorale IAEM Lorraine

UFR mathématiques et informatique

Combinatoire des droites et segments pour la visibilité 3D

THÈSE

présentée et soutenue publiquement le 29 octobre 2007

pour l'obtention du

Doctorat de l'université Nancy 2

(spécialité informatique)

par

Marc Glisse

Composition du jury

Président :	Paul Zimmermann
Rapporteurs :	Helmut Alt Michel Pocchiola
Examinateurs :	Hazel Everett Sylvain Lazard Raimund Seidel

Laboratoire Lorrain de Recherche en Informatique et ses Applications — UMR 7503

Loria

Résumé

Cette thèse présente principalement des résultats sur la combinatoire des droites et segments qui apparaissent naturellement dans l'étude des problèmes de visibilité en trois dimensions. Nous exposons en premier lieu des résultats sur la taille de la silhouette d'un objet vu d'un point, c'est à dire sur la complexite de l'ensemble des droites ou segments tangents à l'objet et passant par le point. Nous présentons en particulier les premières bornes théoriques non triviales pour des polyèdres non-convexes, à savoir que, sous des hypothèses raisonnables, la complexité moyenne de la silhouette est au plus la racine carrée de la complexité du polyèdre, phénomène largement observé en infographie. Nous présentons aussi des bornes, en moyenne et dans le cas le pire, sur le nombre de droites et segments tangents à quatre objets dans une scène composée d'objets polyédriques ou sphériques. Ces bornes donnent en particulier l'espoir que la complexité des structures de données globales comme le complexe de visibilité ne soit pas nécessairement prohibitive. Les bornes sur les polytopes sont également les premières à tirer parti des propriétés structurelles des scènes composées de triangles organisés en polytopes de facon réaliste, c'est à dire non nécessairement disjoints. Ces bornes induisent enfin les premières bornes non triviales sur la complexité des ombres induites par des sources lumineuses non ponctuelles. Les résultats presentés dans cette thèse améliorent significativement l'état de l'art sur les propriétés combinatoires des structures de visibilité en trois dimensions et devraient favoriser les développements algorithmiques futurs pour ces problèmes.

Abstract

This thesis mainly presents results on combinatorial problems on lines and segments that appear naturally in the study of visibility in three dimensions. We first show results on the size of the silhouette of an object seen from a viewpoint, that is the complexity of the set of lines tangent to the object that go through the point. In particular, we give the first non-trivial theoretical bounds for non-convex polyhedra, that is that under some reasonable hypotheses, the average complexity of the silhouette is at most the square root of the complexity of the polyhedron, a phenomenon widely observed in graphics. We also provide bounds, on average and in the worst case, on the number of lines tangent to four objects where the objects are either polytopes or balls. These bounds give us hope that the size of global structures like the visibility complex might not be prohibitive. The bounds on the polytopes are also the first that take advantage of the structure of scenes where triangles are organized in realistic polytopes, that is polytopes that need not be disjoint. These bounds eventually induce the first non-trivial bounds on the complexity of the umbras created by area light sources. The results presented in this thesis significantly enhance the state of the art on the combinatorial properties of visibility structures in three dimensions and should help the future algorithmic developments for these problems.

Remerciements

De nombreuses personnes doivent être remerciées pour cette thèse. Il y a pour commencer mes parents, sans qui rien n'aurait été possible. Il y a ensuite les professeurs qui, avec mes parents, m'ont fait m'intéresser aux sciences et aux mathématiques en particulier. Il faut aussi remercier Anne-Laure et Bubu qui, très indirectement, sont responsables de mon orientation vers l'informatique. Il y a ensuite Luc, qui m'a fait découvrir l'existence de la géométrie algorithmique, et Michel qui m'a donné mon premier cours sur le sujet. Il y a aussi Hervé, avec qui j'ai fait mes débuts dans la recherche, et qui m'a présenté à l'équipe avec qui j'ai finalement décidé de faire cette thèse. Il y a encore les chercheurs de renommée qui n'ont pas hésité à collaborer avec un étudiant, comme Helmut, Jeff, Olivier, Otfried. Je tiens à remercier particulièrement tous les membres de mon jury.

L'équipe dans laquelle j'ai préparé cette thèse a été formidable. Ils ont su à la fois me proposer des sujets de recherche motivants, être des collègues très sympathiques, sans oublier de faire preuve d'un minimum de fermeté quand c'était nécessaire. Je considère aujourd'hui les membres de l'équipe comme des amis plus que comme des collègues. Je dois particulièrement remercier Sylvain qui est sans doute celui qui a dû supporter la tâche la plus difficile : m'apprendre à rédiger...

Je voudrais enfin remercier les amis qui, souvent à distance, m'ont accompagné à diverses périodes de cette thèse, dont Arnaud, Bubu, Liguori, Mélanie, Roupoil, Sandra, Shaggy...

Table des matières

Première partie

Introduction

Chapitre 1

Introduction

La visibilité est un sujet très important en informatique. Elle joue un rôle dans des thématiques très diverses. F. Durand a écrit un bon panorama des différentes applications et techniques liées à la visibilité dans [54]. Je ne citerai ici que quelques-uns de ces problèmes.

En informatique graphique, le premier problème qui se pose est de déterminer les objets qui sont visibles depuis un point donné, ou de façon duale de supprimer les objets (ou parties d'objets) qui ne sont pas visibles. C'est un domaine extrêmement vaste qui a donné lieu à de très nombreux travaux. Il est à noter que c'est loin d'être le seul problème de visibilité rencontré en infographie. Une fois déterminé quel objet est vu, on a besoin de savoir de quelle couleur il est vu, c'est-à-dire en particulier s'il est éclairé ou à l'ombre. Calculer les zones d'ombre est un problème de visibilité : il s'agit de déterminer les régions qui ne voient pas les sources lumineuses. Pour obtenir un rendu réaliste, une approche est de simuler l'illumination globale de la scène, c'est-à-dire de calculer les échanges de lumière entre tous les éléments de surfaces, qui sont considérés à la fois comme des capteurs et des émetteurs de lumière. L'étape la plus complexe de ce type de calcul consiste à déterminer le facteur de forme de deux éléments de surface, une mesure de leur visibilité relative définie comme la proportion de lumière émise par l'un qui est reçue par l'autre ; à ce jour, la seule méthode effective pour approcher le facteur de forme consiste à échantillonner les éléments de surface par des nuages de points et à étudier les interactions lumineuses entre ces points.

La reconnaissance de forme est un autre problème lié à la visibilité. Il s'agit par exemple de reconnaître dans une image 2D la présence d'un objet dont on connaît un modèle 3D. Pour cela on calcule généralement d'abord l'ensemble des vues 2D combinatoires possibles de l'objet avant d'essayer de reconnaître certaines de ces vues dans l'image.

La visibilité apparaît aussi en planification de mouvement. Le plus court chemin d'un point à un autre dans le plan entre des obstacles est constitué de segments de visibilité tangentielle entre les objets, donc déterminer ce plus court chemin peut se ramener à un problème de graphe discret après avoir calculé la visibilité entre les objets.

On distingue deux approches fondamentales aux problèmes de visibilité : le lancer de rayon et l'approche globale. Les deux sont présentées ci-dessous.

Lancer de rayon

Une des techniques les plus utilisées en visibilité est le lancer de rayon. Il s'agit du problème de visibilité le plus naturel : quel objet voit-on dans cette direction ? Ou plus géométriquement : étant donnée une demi-droite, quel est l'objet coupé par cette demi-droite le plus proche de son extrémité ?

Cette requête basique permet en fait de résoudre de façon approchée la plupart des problèmes plus complexes de visibilité grâce à des échantillonnages. Pour calculer la vue d'un point, on échantillonne l'ensemble des directions et envoie un rayon dans chacune des directions sélectionnées. Pour calculer le facteur de forme entre deux éléments de surface, on prend un échantillon de points sur un des éléments et un échantillon de directions, on envoie des rayons depuis ces points dans ces directions, et on compte combien sont arrivés sur l'autre élément. Avec de bons échantillons, on obtient ainsi des réponses approximatives relativement satisfaisantes à la plupart des problèmes de visibilité.

Grâce à son utilité évidente, cette requête à fait l'objet de nombreux travaux visant à réduire le temps de réponse. Chang, dans [34], donne une liste des techniques les plus populaires utilisées pour accélérer le lancer de rayon. Une grande catégorie d'approches consiste à subdiviser l'espace afin que lorsque le rayon se situe dans une région déterminée de l'espace on n'ait besoin de tester son intersection qu'avec les rares objets qui appartiennent à cette région. Une autre approxime les objets par des objets beaucoup plus simples afin de simplifier les tests d'intersection et de n'avoir besoin d'étudier les objets dans toute leur complexité que lorsque c'est vraiment nécessaire.

La plupart des techniques développées sont des heuristiques dont il n'est pas prouvé que la complexité soit intéressante, mais qui en pratique sont extrêmement rapides. D'autres techniques ont été étudiées pour lesquelles les auteurs ont su prouver des bornes sur le temps maximal que peut prendre une requête, mais ces techniques n'ont généralement pas été implantées.

Approche globale

L'approche consistant à lancer un grand nombre de rayons pour approximer les résultats de visibilité a un certain nombre d'inconvénients. En particulier, dès que l'on modifie le moindre élément de la scène, il faut tout recalculer. Il est difficile d'exploiter la cohérence spatiale (des rayons proches ont des chances de voir la même chose) et temporelle (si les objets bougent peu on a des chances de continuer à voir la même chose) avec le lancer de rayon.

À l'opposé du lancer de rayon, une approche globale a été étudiée. Elle consiste à calculer toutes les informations de visibilité d'un certain type et les stocker dans une structure de données. Cette structure est ensuite utilisée pour répondre efficacement aux diverses requêtes de visibilité. Elle peut aussi être mise à jour lorsque la scène est modifiée.

Selon le problème que l'on cherche à résoudre, les informations de visibilité stockées dans la structure ne sont pas les mêmes. Une des structures les plus simples est le graphe de visibilité, qui est un graphe dont les nœuds sont les objets de la scène et les arêtes relient les objets qui se voient. Une structure beaucoup plus grosse est le graphe d'aspect qui

partitionne l'espace en composantes connexes de points qui ont la même vue topologique. Entre les deux on trouve le complexe de visibilité qui partitionne l'ensemble des rayons en composantes connexes de rayons qui tombent sur le même objet, et sa version simplifiée, le squelette de visibilité, dont le plongement dans l'espace ambiant correspond au graphe d'aspect. Une étude plus détaillée du sujet peut être trouvée dans trois thèses récentes : [53], [65] et [72].

Cette approche globale a une autre particularité que son caractère global : elle peut être exacte. Elle ne nécessite pas d'approximer la vue par une grille de pixels. Elle permet donc de produire des images indépendantes de la résolution et dont la validité est garantie. Ce caractère exact pourrait rendre cette méthode très utile en particulier dans la vérification des résultats obtenus par les heuristiques et approximations actuellement utilisées en pratique. Le squelette de visibilité a aussi été appliqué avec succès au calcul des limites d'ombres par Duguet et al. [52], un problème sur lequel les techniques classiques de raffinement de maillage ont montré leurs limitations.

L'étude de l'approche globale a cependant été freinée par l'impression que les structures mises en jeu sont généralement trop grosses. En effet, les premières bornes théoriques [56] découvertes sur sa taille concernaient des scènes artificiellement pathologiques. Une première implantation de ces techniques [57] a permis de commencer à regarder la taille de ces structures sur des scènes réelles, mais les limitations algorithmiques de cette implantation n'ont laissé étudier que des scènes de petite taille, donc peu représentatives de ce qui est utilisé en pratique.

Ainsi, afin de savoir dans quelles conditions une approche globale peut être raisonnable, il est important de mieux comprendre la taille des structures impliquées et de développer des algorithmes efficaces. Cette thèse porte principalement sur l'étude de la taille de telles structures.

Les objets géométriques centraux de cette étude sont les droites libres (celles qui ne coupent aucun obstacle) et surtout les segments libres maximaux, c'est-à-dire des segments (éventuellement infinis) qui ne coupent aucun obstacle mais s'appuient sur des obstacles (ou l'infini) à leurs deux extrémités. Les droites qui jouent le rôle le plus important sont les droites tangentes à des objets de la scène. En effet, elles marquent la limite entre les droites qui coupent un objet et celles qui ne le coupent pas. En 2D, on étudie ainsi les droites tangentes à un objet et surtout les droites tangentes à deux objets, qui en général sont en nombre fini. De la même façon, en 3D, on étudie aussi particulièrement les droites que leurs conditions de tangence définissent comme isolées.

L'étude des structures globales de visibilité se ramène pour l'essentiel à une étude des propriétés combinatoires des droites libres et segments libres maximaux, ce qui est le thème principal de cette thèse.

Organisation du document

Le chapitre 2 présente un résumé des différents chapitres de cette thèse. Je précise de plus à la fin de chaque section les publications dérivées de ces travaux ainsi que ma participation et les personnes avec qui j'ai eu la chance de collaborer. Les chapitres reprennent dans leur majorité les articles essentiellement tels qu'ils ont été publiés, avec parfois à la

fin une section « Bonus » qui présente un résultat supplémentaire non publié.

Dans des scènes constituées de polyèdres, un grand nombre de complexités ne dépendent pas directement de la taille des objets considérés, mais plus de la complexité de leur silhouette. Il est communément admis que la silhouette d'un polyèdre à n arêtes a généralement de l'ordre de $\sqrt{n}$ arêtes. Après avoir présenté des exemples de polytopes n'ayant pas une silhouette de faible complexité, je donne dans le chapitre 3 des conditions sous lesquelles on peut garantir que toutes les silhouettes d'un polyèdre sont de petite taille. Je prouve ensuite dans le chapitre 4 que, sous des hypothèses raisonnables, la complexité moyenne de la silhouette d'un polyèdre qui approxime une surface est $O(\sqrt{n})$, pour une direction d'observation aléatoire. À l'exception d'un théorème de Kettner et Welzl [77] qui ne s'applique qu'à des approximations convexes d'une sphère, il s'agit des premières bornes théoriques sur la complexité des silhouettes, alors que la propriété de faible taille des silhouettes est utilisée depuis très longtemps.

L'essentiel de mes travaux porte sur l'approche globale. Dans la partie III (chapitres 5 à 8), j'étudie cette approche dans des scènes constituées de polytopes.

Je regarde d'abord différentes définitions possibles du complexe de visibilité dans le chapitre 5. En effet, les précédentes études sur le sujet n'ont pas donné de définition unique de cette structure, et un certain nombre des définitions utilisées posent de sérieux problèmes.

J'ai particulièrement travaillé sur la taille des structures de visibilité, et pour cela sur le dénombrement des quadritangentes, c'est-à-dire des droites tangentes à quatre objets. En effet, des structures comme le complexe de visibilité s'appuient sur une classification des droites en fonction de leurs relations de tangence aux objets de la scène. L'espace des droites est de dimension quatre et la propriété d'être tangent à un objet est généralement de codimension 1. Donc de façon générale les droites définies de façon unique (ou au moins isolées) par leurs relations de tangence sont les quadritangentes, ce qui fait de ces droites une des clefs de l'étude des structures globales de visibilité.

Je me suis ainsi d'abord intéressé au nombre maximal de tangentes à quatre objets dans des scènes constituées de k polytopes avec n arêtes au total. Le chapitre 6 montre que ce nombre est $\Theta(n^2k^2)$ dans le cas le pire (et explique comment les construire), ce qui constitue essentiellement la première borne tirant parti de la structure de scènes réalistes (une borne similaire de $O(n^2k^2)$ avait déjà été prouvée précédemment [59], mais dans un cadre plus restrictif). J'étudie ensuite dans le chapitre 7 les applications de ces résultats à la complexité des ombres dans des scènes constituées de polytopes disjoints.

Pour implanter les algorithmes de construction des structures de visibilité, on a besoin de prédicats fiables et efficaces permettant de manipuler les droites tangentes à quatre polytopes, c'est-à-dire en général les droites définies comme transversales à quatre segments. Le chapitre 8 présente en particulier un prédicat permettant de tester l'existence d'une droite transversale à quatre segments donnés, et le degré maximal des polynômes impliqués dans cette implantation est minimal.

J'ai aussi étudié dans la partie IV les problèmes de visibilité dans des scènes constituées de sphères, un cas particulier d'objet lisse.

Je dénombre en particulier les tangentes à quatre objets dans des scènes constituées de n boules, aussi bien dans le cas le pire que dans le cadre d'objets placés aléatoirement. Le chapitre 9 présente un exemple avec $\Omega(n^3)$ droites libres tangentes à des quadruplets

de boules, et un autre avec $\Omega(n^4)$ segments libres maximaux tangents à des quadruplets de boules unité. Le chapitre 10 montre que dans une scène avec des sphères réparties uniformément, chaque objet n'interagit qu'avec un nombre borné d'autres objets.

Le paragraphe précédent concerne le calcul d'informations de visibilité pour une scène complète. Dans certains cas, on veut simplement savoir si deux boules se voient. Il est très coûteux de déterminer la réponse exacte à cette question, et on présente dans le chapitre 11 un algorithme permettant de dire plus rapidement si deux boules se voient « à ε près ». Cet algorithme est dérivé du résultat suivant sur les couvertures de disques : si des disques unité couvrent un disque unité dans le plan, en n'en gardant qu'environ $\varepsilon^{-1/2}$ on peut continuer à couvrir tout le disque sauf une surface d'aire inférieure à ε.

Je présente pour terminer deux travaux indépendants. Avant de commencer ma thèse et de me concentrer sur l'approche globale, je me suis intéressé au lancer de rayon et en particulier à une technique très classique pour l'accélérer : les octrees. En m'appuyant sur une mesure de la complexité moyenne de lancer de rayon dans une scène pour un octree donné, présentée dans [13], j'ai développé un algorithme qui construit pour une scène donnée un octree quasiment optimal (chapitre 12). J'inclus ce résultat ici car il forme un contraste intéressant avec l'approche globale.

Enfin, que serait une thèse de géométrie algorithmique sans diagramme de Voronoï ? J'ai donc étudié le diagramme de Voronoï du polygone le plus éloigné dans le plan, et le chapitre 13 montre qu'il a une taille linéaire et explique comment le calculer en temps presque linéaire. Surprenamment, ce diagramme n'avait pas été étudié jusqu'à maintenant.

Chapitre 2

Résumé des contributions

2.1 Silhouettes

La silhouette d'un polyèdre associée à un point d'observation donné est, grossièrement, l'ensemble des arêtes situées entre une face qui fait face à l'observateur et une qui lui tourne le dos. Les silhouettes apparaissent dans divers problèmes en infographie comme la suppression des surfaces cachées et les calculs d'ombres (voir [51, 52, 59]) et les algorithmes permettant de les calculer ont été largement étudiés ([74] en présente un survol). Elles sont importantes en reconnaissance de formes et constituent une des indications visuelles les plus significatives de la forme d'un objet.

Il est communément admis que la silhouette d'un polyèdre a généralement une complexité bien moindre que celle du polyèdre entier. Sander et al. [94] par exemple énoncent la propriété souvent répétée que la silhouette d'un maillage a souvent une taille $\Theta(\sqrt{n})$ où n est le nombre de faces du maillage. Une étude expérimentale de Kettner et Welzl [77] confirme ceci pour un ensemble d'objets réalistes. Ces expériences ont été étendues par McGuire [81] à un ensemble plus vaste d'objets plus gros pour lesquels il a observé en moyenne des silhouettes de taille $n^{0.8}$.

Il y a peu de résultats théoriques permettant de justifier ces observations. Kettner et Welzl [77] prouvent qu'un polyèdre convexe qui approxime une sphère avec une distance de Hausdorff ε a $\Theta(1/\varepsilon)$ arêtes et qu'une projection orthographique aléatoire d'un tel polytope est un polygone à $\Theta(1/\sqrt{\varepsilon})$ côtés, les arêtes sur le bord du polygone correspondant à celles sur la silhouette du polytope.

Les techniques utilisées ici s'appuient toutes sur la longueur de la silhouette (dans une projection orthographique, ce qui correspond à un observateur à l'infini) pour déduire des bornes sur le nombre d'arêtes qui composent cette silhouette.

On appelle épais un polytope qui est inclus dans une boule et en contient une autre telles que le rapport des rayons est borné. On dit qu'un polytope a des arêtes comparables si le rapport de longueur entre la plus grande et la plus courte est borné. Nous montrons dans le chapitre 3 la propriété suivante :

Théorème 1. *Un polytope épais, à faces triangulaires, à arêtes comparables, et de diamètre* 1 *a des arêtes de longueur* $\Omega(1/\sqrt{n})$.

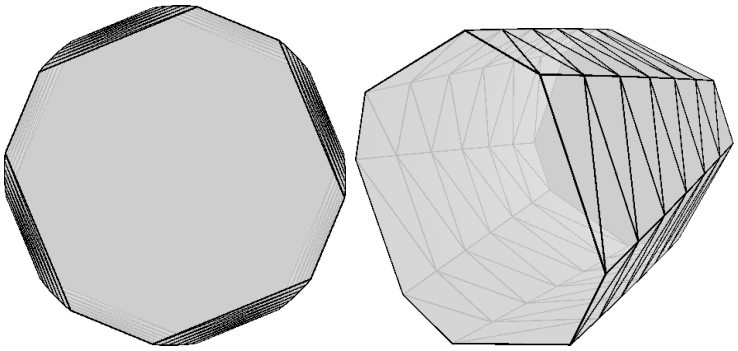


FIG. 2.1 – Une silhouette de complexité linéaire.

Un polytope épais de diamètre 1 a forcément une silhouette de longueur $\Theta(1)$. Si chaque arête contribuait à la longueur de la projection une fraction constante de sa longueur, on aurait une borne de $O(\sqrt{n})$ sur le nombre d'arêtes sur la silhouette. Ce n'est pas toujours le cas comme le montre la figure 2.1 où un nombre linéaire d'arêtes sont sur la silhouette mais presque alignées avec la direction d'observation et donc ne contribuent presque pas à la longueur de la silhouette.

Si on suppose de plus que deux arêtes adjacentes font entre elles un angle minimal, la borne $O(\sqrt{n})$ devient vraie. Alternativement, si on suppose que les directions des droites supportant les arêtes sont uniformément réparties, on peut prouver une borne de $O(n^{2/3})$ sur le nombre d'arêtes sur la silhouette.

On voit ici qu'il est assez difficile de prouver des bornes intéressantes sur la complexité de la silhouette pour un observateur particulièrement mal placé. En revanche, il est beaucoup plus simple d'étudier la complexité moyenne de la silhouette d'un polyèdre donné pour un observateur situé à un endroit aléatoire à l'infini, ou de façon équivalente une projection orthographique dans une direction aléatoire. En effet, on remarque rapidement qu'une arête e est sur la silhouette d'un polyèdre pour un ensemble de directions de mesure proportionnelle à α_e l'angle diédral externe entre les faces adjacentes à e (qui vaut 0 quand les faces se prolongent et π quand e correspond à un pli maximal). Il suffit alors d'étudier la somme de ces angles pour connaître la taille moyenne de la silhouette d'un polyèdre. On remarque de plus que la longueur moyenne que contribue une arête à la silhouette est proportionnelle à $l_e\alpha_e$ où l_e est la longueur de l'arête. En d'autres termes, lorsqu'elle apparaît sur la silhouette, une arête contribue en moyenne une fraction constante de sa longueur à la longueur de la silhouette. Connaître la longueur des arêtes permet donc de relier très simplement le nombre d'arêtes sur la silhouette à la longueur de la silhouette, et les hypothèses du théorème 1 suffisent à obtenir une borne de $O(\sqrt{n})$ sur la complexité de la silhouette.

Les résultats précédents, présentés dans le chapitre 3, se concentrent sur des polyèdres convexes. La convexité est cruciale et sans elle il est facile d'obtenir des silhouettes de grande taille. Afin de pouvoir prouver des bornes sur la silhouette de polyèdres non convexes, on a besoin d'une hypothèse forte remplaçant la convexité.

On présente dans le chapitre 4 des résultats sur des polyèdres qui approximent une surface donnée. Intuitivement, si un polyèdre approxime bien une surface lisse, deux faces

adjacentes font un angle $O(1/\sqrt{n})$ entre elles. D'autre part, si un polyèdre approxime une surface non lisse, deux faces n'ont un angle important entre elles que si la surface n'est pas lisse à cet endroit, et un élément de la silhouette du polyèdre peut être associé à un élément de silhouette de la surface approximée.

Le résultat précis est le suivant. Soit P un polyèdre avec n faces triangulaires épaisses qui est homéomorphe à une variété S de dimension 2 compacte sans bord et dont la silhouette a une longueur moyenne bornée $silh(S)$. On appelle $f : P \to S$ l'homéomorphisme. On suppose que toutes les arêtes de P mesurent au moins $\frac{\alpha}{\sqrt{n}}$ et que pour tout point x de P la distance entre x et $f(x)$ est au plus $\frac{\gamma}{n}$ où α et γ sont deux constantes positives. Alors pour tout point x de P la distance entre x et $f(x)$ est au plus $\frac{\beta\, h(x)}{\sqrt{n}}$ où $h(x)$ est la plus petite des hauteurs du triangle de P qui contient x et β est une constante positive, et la complexité moyenne de la silhouette de P est majorée par $\left(15\,\beta + \frac{24}{\alpha}\,\mathrm{silh}(S)\right)\sqrt{n}$. Je généralise ensuite ce résultat à des surfaces à bord et à d'autres distributions du point de vue.

Ce résultat est important car il s'agit, pour des polyèdres non convexes de taille n, de la première confirmation théorique de la propriété communément admise que les silhouettes sont souvent de taille $\sqrt{n}$.

Le chapitre 3 présente mon travail de DEA, publié à la Canadian Conference on Computational Geometry, rédigé en collaboration avec H. Alt et X. Goaoc et inclus ici à cause de son lien fort avec le chapitre 4. Le chapitre 4 a été publié au Symposium on Computational Geometry [63] et soumis à Discrete and Computational Geometry. J'en suis l'auteur principal et l'ai rédigé en collaboration avec S. Lazard.

2.2 Tangentes à des polytopes

Une structure intéressante pour la visibilité est le complexe de visibilité [56, 90] qui encode les relations de visibilité en partitionnant l'ensemble des segments libres maximaux en fonction des objets qu'ils touchent. Dans le chapitre 5, on s'intéresse à la définition précise de cette structure. Il s'avère en effet que les définitions proposées précédemment sont incompatibles entre elles et posent individuellement certains problèmes.

La taille de cette partition est fortement liée au nombre de segments libres maximaux tangents à quatre objets de la scène. Pour une scène de n triangles, il peut y avoir $\Theta(n^4)$ telles droites dans le cas le pire [56], même si ces triangles forment un terrain (voir [39] ou la figure 2.2). La table 2.1 résume les bornes connues sur le sujet.

Le complexe est donc potentiellement énorme, ce qui a beaucoup freiné son application en pratique. Cependant, certains indices, à la fois théoriques et pratiques, semblent indiquer que cette estimation de taille est trop pessimiste. Les bornes inférieures sont spécialement conçues pour l'occasion et peu réalistes en pratique. Pour des scènes réalistes, Durand et al. [57] observent une croissance quadratique de la taille du complexe, mais pour des scènes de petite taille. Pour des scènes aléatoires, Devillers et al. [46] montrent que la taille du complexe est bien plus petite : pour des boules unité aléatoirement distribuées l'espérance de la taille est linéaire, et pour des polygones ou polytopes épais de taille comparable elle est au plus quadratique. De plus, en 2D, même si la taille du complexe peut être quadratique dans le cas le pire, des résultats expérimentaux suggèrent

	Cas le pire	En moyenne
droites libres pour un polyèdre	$\Theta(n^4)$ (trivial)	
droites libres au dessus d'un terrain polyédrique	$O(n^3 2^{c\sqrt{\log n}})$ [67, 89]	
droites libres parmi des polytopes homothètes disjoints	$\Omega(n^3)$ [41]	
droites libres parmi des boules unité	$\Omega(n^2)$ [46], $O(n^{3+\epsilon})$ [5]	$\Theta(n)$ [46]
segments libres max. au dessus d'un terrain polyédrique	$\Theta(n^4)$ [39]	
segments libres max. isolés parmi k polyèdres convexes disjoints génériques	$\Theta(n^2k^2)$ [59, 26]	
segments libres max. parmi des boules unité	$\Omega(n^2)$ [46], $O(n^4)$	$\Theta(n)$ [46]

TAB. 2.1 – Bornes publiées sur la complexité de l'ensemble des droites libres ou des segments libres maximaux parmi des objets de complexité totale n. Les complexités en moyenne concernent une distribution uniforme des centres des sphères.

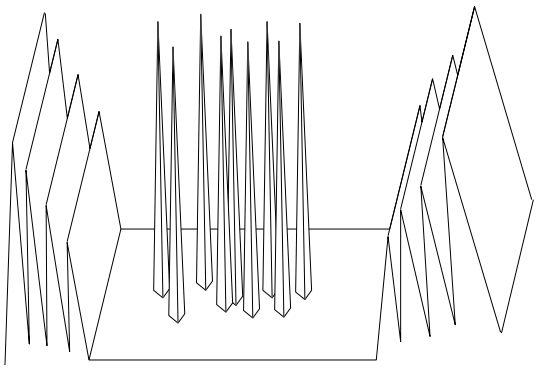


FIG. 2.2 – Un terrain de taille n avec $\Omega(n^4)$ segments libres maximaux tangents en quatre points.

fortement que la taille du complexe d'une scène consistant en des triangles éparpillés est linéaire [36].

Si ces résultats sont encourageants, il ne faut pas oublier que la plupart des scènes ne sont pas aléatoires. En fait, les scènes sont en général très structurées, et on peut en tirer parti. Une scène est souvent représentée par de nombreux triangles qui forment un nombre bien moins important de morceaux convexes. En particulier, si la scène est constituée de k polytopes disjoints avec un total de n arêtes, sous une hypothèse forte de position générique, le nombre maximal de droites tangentes à quatre polytopes est $O(n^2k^2)$. Cela se déduit directement de la borne prouvée dans [59] sur le nombre de changements combinatoires de la carte des silhouettes vue depuis un point mobile sur une droite.

Dans le chapitre 6, on présente une généralisation de ce résultat à des polytopes qui peuvent se couper et être dans des positions dégénérées, pour lesquels la borne de $O(n^2k^2)$ reste valide ; nous montrons aussi une borne de $\Omega(n^2k^2)$ dans le cas le pire. Cette généralisation est très différente du cas 2D où deux polygones disjoints n'ont que quatre bitangentes alors que deux polygones qui se coupent peuvent en avoir un nombre linéaire. On généralise aussi ces résultats en montrant que le nombre de composantes connexes

de segments libres maximaux tangents à au moins quatre polytopes est $O(n^2k^2)$, et que cela marche aussi pour des droites (non nécessairement libres). On majore encore par $O(n^2k^2)$ le nombre d'ensembles minimaux de sommets et arêtes ouvertes qui admettent une transversale commune isolée tangente aux polytopes.

L'idée principale de la preuve est de considérer une arête particulière et de balayer l'espace avec les plans qui contiennent cette arête. Dans ces plans, les polytopes sont des polygones, et on cherche les positions du plan telles que trois polygones admettent une tangente commune.

On donne enfin un algorithme qui utilise la même approche de balayage que la technique d'énumération pour construire ces tangentes en temps $O(n^2k^2 \log n)$. Obtenir un algorithme de cette complexité n'est pas trivial, car le balayage expliqué ci-dessus fait apparaître toutes les droites tangentes à quatre polytopes sans prendre en compte les occlusions possibles et il faut donc être capable d'éliminer les tangentes obstruées sans coût supplémentaire.

Ces résultats, obtenus en collaboration avec H. Brönnimann, O. Devillers, V. Dujmović, H. Everett, X. Goaoc, S. Lazard, H.-S. Na et S. Whitesides, ont été présentés à la Canadian Conference on Computational Geometry [26] et au Symposium on Computational Geometry [27]. La version complète est publiée dans le SIAM Journal on Computing [28]. J'ai contribué aux versions préliminaires, en particulier pour la borne inférieure.

Le travail sur la définition du complexe, présenté dans le chapitre 5, est plus personnel.

2.3 Ombres de polytopes

Les ombres jouent un rôle central dans la perception humaine [80, 106]. Malheureusement, calculer des ombres réalistes est un problème très difficile, en particulier quand les sources de lumière ne sont pas ponctuelles. Une variété d'approches a été considérée pour rendre les ombres (voir par exemple [54, 110] pour un panorama) et plusieurs techniques utilisent fortement les possibilités des cartes graphiques modernes (voir [68] pour une vue d'ensemble).

Un point est à l'ombre s'il ne voit aucune partie d'aucune source lumineuse. Un point est à la lumière s'il voit entièrement toutes les sources lumineuses. Les autres points sont dans la pénombre. La limite entre la pénombre et la lumière est raisonnablement bien comprise, comme on le verra dans la section 7.3. En revanche, on sait moins de choses sur la limite de l'ombre. Il existe néanmoins une littérature conséquente sur la construction de ces contours d'ombres (voir par exemple [49, 57, 55, 56, 70, 87, 101, 104]).

Dans le chapitre 7, on prouve diverses bornes résumées dans les tableaux 2.2 et 2.3 sur la complexité de l'ombre et de la pénombre projetées sur un plan par un segment ou des polytopes lumineux en présence d'obstacles en forme de polytopes. On montre en particulier qu'un segment lumineux, en présence de deux obstacles triangulaires, peut de façon surprenante générer quatre composantes connexes d'ombre. On montre aussi des bornes de $\Omega(nk^2 + k^4)$ sur le nombre maximal de composantes connexes de l'ombre et $O(nk^3)$ sur la complexité de cette ombre dans une scène avec un segment lumineux et k polytopes disjoints comme obstacles. On prouve enfin que l'ombre portée sur un plan par un polytope lumineux avec k polytopes obstacles peut avoir $\Omega(n^2k^3 + nk^5)$ composantes

Type de scène	Bornes inférieures	Bornes supérieures
Source lumineuse : segment		
2 triangles	4	$O(1)$
2 polytopes épais	$\Omega(n)$	$O(n)$
k polytopes	$\Omega(nk^2 + k^4)$	$O(nk^3)$
Source lumineuse : polytope(s)		
k polytopes	$\Omega(n^2k^3 + nk^5)$	$O(n^3k^3)$

TAB. 2.2 – Bornes inférieures sur le nombre de composantes connexes et bornes supérieures sur la complexité de l'ombre portée sur un plan par k polytopes de complexité totale $O(n)$. Dans le cas des sources lumineuses polygonales, la borne inférieure utilise un unique polytope lumineux, de complexité $O(n)$, et la borne supérieure s'applique quelle que soit la proportion des k polytopes qui sont lumineux.

Sources lumineuses	Borne inférieure	Borne supérieure
$O(1)$ polytopes de taille m	$\Omega(n\alpha(k) + km + k^2)$	$O(n\alpha(k) + km\alpha(k) + k^2)$
$O(k)$ polytopes de taille $O(n)$	$\Omega(nk + k^4)$	$O(nk\alpha(k^2) + k^4)$

TAB. 2.3 – Bornes sur la complexité de l'union de l'ombre et de la pénombre portées sur un plan par un ensemble de k polytopes disjoints de complexité totale n dont certains sont lumineux.

connexes et a une complexité $O(n^3k^3)$. Il s'agit des premiers résultats non triviaux sur la complexité des ombres dans des scènes de polytopes.

On voit alors que l'ombre, qui est délimitée par des arcs de conique, est intrinsèquement beaucoup plus complexe que la limite entre la lumière et la pénombre qui est bornée par des segments et dont on prouve que la complexité est $\Omega(nk + k^4)$ (dans le cas le pire) et $O(nk\alpha(k^2) + k^4)$ où $\alpha(k)$ désigne l'inverse de la fonction d'Ackermann. De plus, s'il n'y a qu'un nombre borné de sources lumineuses de complexité m, la complexité est $\Omega(n\alpha(k) + km + k^2)$ (dans le cas le pire) et $O(n\alpha(k) + km\alpha(k) + k^2)$.

Ce travail a été réalisé en collaboration avec J. Demouth, O. Devillers, H. Everett, S. Lazard et R. Seidel, présenté au Symposium on Computational Geometry [45] et soumis à Computational Geometry: Theory and Applications. Ma contribution porte sur les résultats de bornes inférieures et en particulier les théorèmes 60, 61 et 62.

2.4 Prédicats sur les tangentes à des polytopes

Comme tous les algorithmes géométriques, les calculs de visibilité s'appuient sur des tests géométriques de base impliquant un petit nombre d'objets. La réponse d'un prédicat est discrète, il s'agit typiquement d'un signe $+/-/0$ qui peut être interprété par quelque chose comme dedans / dehors / sur le bord, ou plus haut / moins haut / à la même hauteur... Pour passer de la formulation géométrique d'un prédicat (un point est-il à

l'intérieur d'un objet ?) à une implantation, il s'agit en général de trouver un ou plusieurs polynômes en les paramètres qui décrivent les objets, et dont le signe (ou une combinaison des signes) donne la réponse à la question géométrique.

Pour un problème géométrique donné, déterminer ce qu'est une bonne implantation est un problème difficile. Entre plusieurs formulations polynomiales, celle dont le degré maximal des polynômes est le plus petit est souvent préférée car le degré d'un polynôme est étroitement lié à la précision nécessaire pour calculer exactement le signe.

Pour les problèmes de visibilité étudiés ici, on a besoin de prédicats déterminant si quatre arêtes admettent une transversale commune, ainsi que de prédicats permettant de traiter de telles droites en vérifiant si elles coupent un triangle donné où en les triant selon leur ordre d'apparition dans le balayage d'un plan autour d'une arête.

De tels prédicats ont déjà été étudiés dans [60]. On présente dans le chapitre 8 des polynômes de degré beaucoup moins élevé. En particulier, si les arêtes sont données par les coordonnées cartésiennes de leurs deux extrémités, on détermine leur nombre de tangentes communes avec des polynômes de degré au plus 12 (contre 36 dans [60]) et on compare deux telles droites dans l'ordre de balayage autour d'une arête avec des polynômes de degré au plus 36 (contre 144).

Quand le polynôme de plus haut degré impliqué dans un prédicat est irréductible, surtout si on peut vérifier que son annulation a un sens géométrique essentiel pour le prédicat, on peut espérer avoir atteint une sorte d'optimalité. C'est effectivement le cas pour le polynôme de degré 12 mentionné ci-dessus, mais vraisemblablement pas pour celui de degré 36 dont le plus gros facteur irréductible semble être seulement de degré 24.

Ce travail est réalisé en collaboration avec O. Devillers, S. Lazard et S. Petitjean.

2.5 Tangentes à des sphères

On s'est intéressé dans les parties précédentes aux scènes dont les objets sont des polytopes. On étudie dans la partie IV les scènes dont les objets sont des sphères.

Dans le chapitre 9, je donne deux bornes inférieures : une borne de $\Omega(n^3)$ sur la complexité de l'espace des droites libres dans une scène de n boules (dont le rapport des rayons n'est pas borné), et une borne de $\Omega(n^4)$ sur la complexité de l'espace des segments libres maximaux dans une scène de boules unité. Le seul résultat non trivial précédemment connu sur le sujet est dû à Agarwal et al. [4] qui ont montré que la complexité maximale de l'espace des droites libres dans une scène de sphères unité est $O(n^{3+\varepsilon})$ pour tout $\varepsilon > 0$. Un résumé des résultats existants est donné dans le tableau 2.4.

	Droites libres		Segments libres maximaux	
Boules unité	$\Omega(n^2)$	trivial	$\Omega(n^4)$	théorème 89
	$O(n^{3+\varepsilon})$	[4]	$O(n^4)$	trivial
Boules quelconques	$\Omega(n^3)$	théorème 84	$\Omega(n^4)$	théorème 89
	$O(n^4)$	trivial	$O(n^4)$	trivial

TAB. 2.4 – Complexités connues sur les espaces libres.

Des scènes formées de boules peuvent donc avoir un complexe de visibilité de très

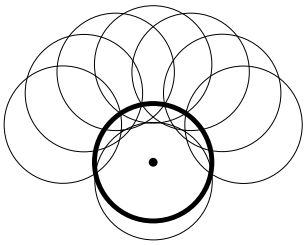


FIG. 2.3 – Une famille couvrante minimale de grande taille.

grande taille. Cependant, les scènes présentées sont très particulières et peu représentatives de ce qui se passe en pratique. Dans [46], Devillers et al. étudient une scène constituée de boules unité uniformément distribuées dans un univers en forme de boule en dimension 3 et montrent que dans ce cas le nombre moyen de droites tangentes à quatre de ces sphères est linéaire en le nombre de sphères.

Dans le chapitre 10, on généralise ce résultat pour des distributions de Poisson. On montre plus précisément que dans un univers de dimension quelconque avec des sphères uniformément distribuées, une sphère ne voit en moyenne qu'un nombre borné d'autres sphères. La borne est assez forte pour montrer aussi qu'une boule ne voit qu'un nombre borné de triplets d'autres boules, et donc ne participe qu'à un nombre borné de quadri-tangentes.

Le travail du chapitre 9 sur les bornes inférieures est personnel. Celui du chapitre 10 sur les scènes uniformes a été réalisé en collaboration avec H. Alt.

2.6 Couverture approximative de boules

On dit qu'une famille $\mathcal{F}$ d'ensembles couvre un ensemble U si U est inclus dans l'union des éléments de $\mathcal{F}$. Dans un cadre géométrique, U et les éléments de $\mathcal{F}$ sont des sous-ensembles d'un espace géométrique, par exemples des points, hyperplans, boules dans $\mathbb{R}^d$. De nombreux problèmes géométriques, en particulier en visibilité, se ramènent à décider si une famille donnée couvre un objet. Par exemple, à quelques détails près (l'ordre dans lequel une droite coupe les objets), deux objets se voient si les droites passant par ces deux objets ne sont pas couvertes par les droites intersectant les obstacles.

Dire qu'une famille $\mathcal{F}$ couvre U revient à dire que l'intersection de U avec les complémentaires des éléments de $\mathcal{F}$ est vide. Si U est convexe et $\mathcal{F}$ se compose de complémentaires d'objets convexes, on peut donc déduire du théorème de Helly [71] que la taille de la plus petite sous-famille de $\mathcal{F}$ qui couvre U est au plus $d+1$. Les cas comme celui-ci où cette taille est bornée sont rares : pour la plupart des classes d'objets il existe des familles minimales couvrantes arbitrairement grandes (la figure 2.3 montre une famille minimale de 8 disques unité qui couvrent un autre disque unité dans le plan).

Pourtant, il semble intuitivement que si une famille $\mathcal{F}$ de disques unité couvre un autre disque unité U, il devrait y avoir un petit sous-ensemble de $\mathcal{F}$ qui couvre presque tout U. On définit donc une ε-couverture de U comme une famille $\mathcal{F}$ telle que la partie de U non couverte par $\mathcal{F}$ a une mesure majorée par ε. On étudie dans le chapitre 11

la taille de la plus petite ε-couverture que l'on peut extraire d'une couverture donnée, et on montre en particulier que dans plusieurs cas la taille de cette ε-couverture peut être bornée indépendamment de la famille couvrante utilisée, soit un résultat similaire au théorème de Helly.

Plus précisément, on montre que si U et les éléments de $\mathcal{F}$ sont des boules unité dans $\mathbb{R}^d$, pour tout $\varepsilon > 0$ on peut extraire de $\mathcal{F}$ une ε-couverture de U de taille $O\left(\varepsilon^{\frac{1-d}{2}} \operatorname{polylog} \frac{1}{\varepsilon}\right)$. On donne aussi un exemple où toute ε-couverture de U à une taille $\Omega\left(\varepsilon^{\frac{1-d}{2}}\right)$, ce qui montre que la borne est presque optimale. L'idée de la preuve est de s'inspirer du théorème de Helly. Si au lieu des boules on avait des demi-espaces, le théorème de Helly dit que l'on pourrait extraire de $\mathcal{F}$ $d+1$ boules qui couvrent U. On approxime donc chaque boule par le demi-espace le plus adapté qui la contient (les demi-espaces couvrent aussi U) et on ne garde que $d+1$ boules dont les demi-espaces couvrent U. La région non couverte de U se situe entre les boules et les hyperplans, ce qui permet de borner sa taille. Si elle est supérieure à ε, on applique récursivement ce raisonnement aux régions qui ne sont pas encore couvertes.

On applique ce résultat à un problème de visibilité. On se place dans $\mathbb{R}^3$ avec des boules unité disjointes. On montre que si deux de ces boules, A et B, ne se voient pas à cause des autres boules situées entre elles, alors on peut trouver une famille de taille $O\left(\varepsilon^{-\frac{7}{2}} \operatorname{polylog} \frac{1}{\varepsilon}\right)$ de boules qui suffisent à bloquer la visibilité entre A et B sauf pour un ensemble de droites de mesure au plus ε.

On étudie aussi le problème de la couverture de cubes de $\mathbb{R}^d$. On montre que là encore, si on a un cube unité U couvert par une famille $\mathcal{F}$ de cubes unité, il existe une borne ne dépendant que de la dimension d et de ε sur la taille d'une sous-famille qui couvre ce cube sauf une région de volume au plus ε. En dimension 2 on montre même précisément une borne optimale de $O(\varepsilon)$.

Enfin, en utilisant des techniques de programmation linéaire généralisée, on obtient pour chacun des résultats précédents un algorithme permettant de trouver un point de U non couvert par la famille $\mathcal{F}$ ou de prouver que $\mathcal{F}$ est une ε-couverture de U. La complexité de cet algorithme est $O(|\mathcal{F}| f(\varepsilon))$ où $f(\varepsilon)$ est la borne sur la taille des ε-couvertures du problème concerné.

Ces travaux se distinguent des recherches précédentes qui cherchaient à extraire des sous-couvertures parfaites de taille proche de l'optimale. Trouver une sous-couverture minimale est généralement **NP**-dur, et il a été montré que dans beaucoup de cas chercher une sous-couverture dont la taille est dans un rapport borné avec la taille minimale est aussi **NP**-dur [78, 32, 37].

Ces travaux ont été réalisés en collaboration avec J. Demouth, O. Devillers et X. Goaoc et soumis au Symposium on Discrete Algorithms.

2.7 Octrees optimaux pour le lancer de rayon

Prévoir et optimiser les performances du lancer de rayon est un sujet très important en infographie vue son utilisation extensive et la quantité de calcul associée. Aronov et Fortune [14] sont les premiers à garantir une performance dans un rapport borné avec

l'optimum dans le modèle de calcul suivant : on construit une triangulation compatible avec la scène et lance des rayons en localisant le point de départ puis en traversant la triangulation jusqu'à rencontrer un objet. Cependant, les triangulations ne sont pas un modèle de calcul très populaire en infographie, contrairement à des décompositions spatiales telles que les kd-trees et les octrees. Aronov et al. ont développé [13] une mesure pour de telles décompositions et ont prouvé qu'elle estimait fiablement le coût moyen du lancer d'un rayon.

Dans le chapitre 12, nous nous intéressons au problème de construire, pour une scène donnée, un octree qui minimise la mesure de coût précédemment mentionnée. On généralise en fait à toute dimension d avec des objets qui sont des simplexes de dimension $d-1$. On donne une construction d'un arbre qui a un coût $O(M)$ où M est le minimum des coûts de tous les arbres possibles. Parfois il est important que l'arbre soit équilibré (que les cellules voisines aient une profondeur proche) : on montre aussi que rééquilibrer l'arbre que l'on a construit n'augmente pas sa taille de plus qu'un facteur multiplicatif. Ces résultats ont depuis été validés expérimentalement par Aronov et al. [12].

Ces travaux, effectués en collaboration avec H. Brönnimann, ont été présentés à la Canadian Conference on Computational Geometry [31] et au Latin American Theoretical Informatics Symposium [30] et publiés à Computational Geometry: Theory and Applications [64]. Ils datent d'avant le début de ma thèse mais constituent mon premier contact avec les problématiques de la visibilité 3D et contrastent avec l'approche globale étudiée dans le reste de la thèse.

2.8 Diagramme de Voronoï du polygone le plus éloigné

On considère une famille $\mathcal{S}$ d'objets géométriques, appelés sites, dans le plan. Le diagramme de Voronoï du site le plus éloigné de $\mathcal{S}$ partitionne le plan en régions, chaque région étant associée à un site $P \in \mathcal{S}$ et contenant les points du plan pour lesquels P est le plus éloigné des sites de $\mathcal{S}$.

Les diagrammes de Voronoï du site le plus proche ont été l'objet d'études très approfondies (voir par exemple [17] pour un survol), mais la variante sur les sites les plus éloignés a été l'objet d'une attention moindre. Le cas où les sites sont des segments qui ont le droit de se couper n'a été résolu que récemment par Aurenhammer et al. [18] ; ils donnent un algorithme de complexité $O(n \log n)$ pour calculer ce diagramme pour n segments.

Les diagrammes de Voronoï du site le plus éloigné ont un certain nombre d'applications importantes. La plus connue est probablement le problème de trouver le plus petit disque qui touche tous les sites. Ce disque peut être calculé en temps linéaire une fois connu le diagramme, car son centre se situe sur un sommet ou un arc du diagramme.

On s'intéresse ici au cas de sites complexes qui ont une description de taille non bornée. Parmi les premiers à se placer dans ce cadre, on trouve Abellanas et al. [1] : leurs sites sont des ensembles finis de points, et la distance à un site est la distance au plus proche de ces points. Cela peut se reformuler en considérant un ensemble de n points coloriés en utilisant k couleurs, et leur diagramme de Voronoï de la couleur la plus éloignée subdivise le plan en fonction de la couleur la plus éloignée. L'objectif est celui déjà mentionné de

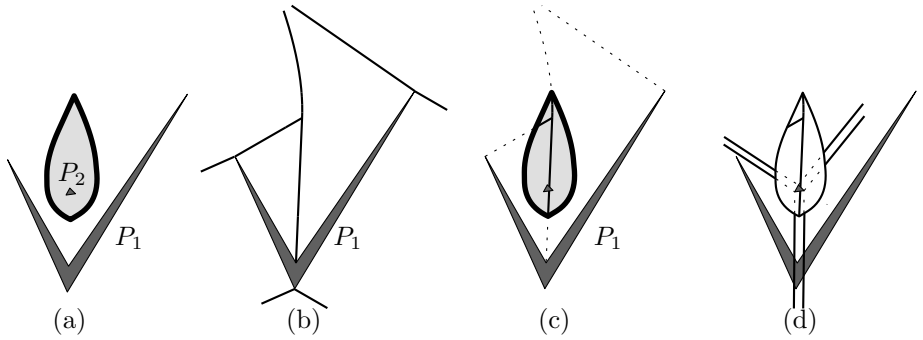


FIG. 2.4 – (a) Le bissecteur de deux polygones peut être une courbe fermée. (b) L'axe médian de P_1. (c) La région de Voronoï de P_1. (d) Le diagramme de Voronoï augmenté de P_1 et P_2.

trouver le plus petit disque qui contient un point de chaque couleur. C'est un problème de localisation de services, où l'on cherche une position aussi proche que possible de k types de services distincts (comme les écoles, les postes, les supermarchés...).

Le diagramme de Voronoï de la couleur la plus éloignée peut être vu comme la projection de l'enveloppe supérieure des k surfaces de Voronoï correspondant aux k couleurs. Huttenlocher et al. [73] montrent que cette enveloppe supérieure a une complexité $\Theta(nk)$ pour n points et peut être calculée en temps $O(nk \log n)$ (voir aussi le livre de Agarwal et Sharir [98, Section 8.7]).

Van Kreveld et Schlechter [105] considèrent le diagramme de Voronoï du site le plus éloigné pour une famille de polygones simples disjoints. À nouveau, ils s'intéressent au problème de trouver le disque de plus petit rayon qui touche tous les polygones et l'appliquent ensuite à un problème cartographique de regroupement des îles en archipels. Leur algorithme s'appuie sur l'affirmation que le diagramme de Voronoï du polygone le plus éloigné est un cas particulier de diagramme de Voronoï *abstrait* du site le plus éloigné tel que défini par Mehlhorn et al. [83]. Mais cette affirmation est fausse, car comme on le voit sur la figure 2.4, le bissecteur de deux polygones peut être une courbe fermée. En particulier, les régions peuvent être bornées, ce qui est impossible pour les diagrammes de Voronoï abstraits.

On remarque que le diagramme de Voronoï du polygone le plus éloigné peut encore être vu comme une enveloppe supérieure de surfaces, mais cela ne semble pas donner de borne intéressante sur sa taille ni d'algorithme efficace.

On montre dans le chapitre 13 que, dans une scène avec k polygones simples disjoints de complexité totale n, la taille du diagramme de Voronoï du polygone le plus éloigné est $O(n)$. Cette borne reste vraie si on subdivise le plan en régions qui ont non seulement le même polygone le plus éloigné mais aussi le même sommet ou arête le plus proche sur ce polygone.

L'idée de la preuve est d'orienter les arcs du diagramme par distance croissante au(x) polygone(s) le(s) plus éloigné(s). Une liste des différents types de sommets du diagramme est présentée dans la figure 2.5. Les sommets du diagramme ayant au moins autant d'arcs entrants que sortants se trouvent tous sur l'axe médian d'un des polygones, ce qui simplifie leur décompte. En remarquant que le nombre d'arcs entrants et d'arcs sortants du

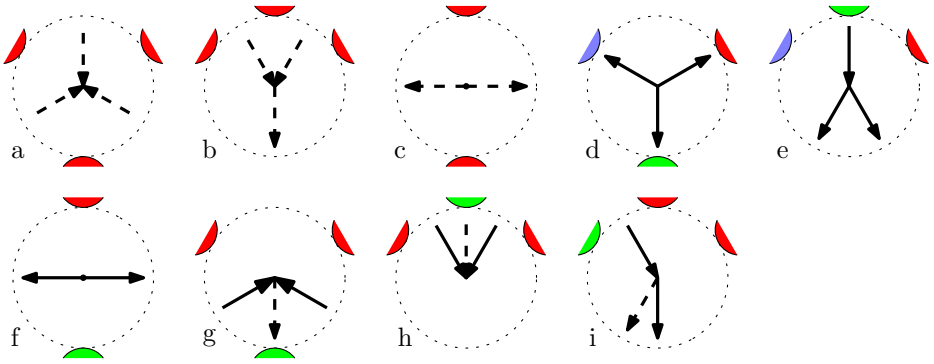


FIG. 2.5 – Les différents types de sommets du diagramme. Les bissecteurs sont en gras et les arcs des axes médians en pointillés.

diagramme entier est le même (c'est simplement le nombre d'arcs), on n'a alors pas besoin de compter les autres types de sommets.

On présente ensuite un algorithme de type « diviser pour régner » pour calculer le diagramme en temps $O(n \log^3 n)$. L'idée est de séparer les polygones en deux sous-ensembles, calculer les diagrammes de chacun de ces ensembles, construire une structure de localisation de point sur chacun de ces diagrammes, puis utiliser des techniques de recherche paramétrée pour trouver des points de départs à partir desquels seront suivies les lignes de fusion. Cet algorithme donne simplement un algorithme de complexité $O(n \log^3 n)$ pour trouver le plus petit disque qui touche tous les polygones.

Notons finalement que dans le cas d'une famille de polygones disjoints *convexes*, la situation est beaucoup plus simple et le plus petit disque qui touche tous les polygones peut être calculé en temps $O(n)$ où n est le nombre total de sommets des polygones [75].

Ces travaux, réalisés en collaboration avec O. Cheong, H. Everett, J. Gudmundsson, S. Hornus, S. Lazard, M. Lee et H.-S. Na, ont été présentés au European Symposium on Algorithms [35].

Deuxième partie

Silhouettes

Chapter 3

On the worst-case complexity of the silhouette of a polytope

Ces résultats, obtenus avec H. Alt et X. Goaoc, ont été présentés à la Canadian Conference on Computational Geometry [8].

Given a viewpoint, the apparent boundary of a polyhedron, or *silhouette*, is the set of edges incident to a visible face and an invisible one; a face whose supporting plane contains the viewpoint is considered invisible. The worst-case upper bound on the complexity of a silhouette is $O(n)$. With this definition, the silhouette of a polytope (i.e., a convex bounded polyhedron) is a simple closed curve on its surface that separates visible and invisible faces.

Silhouettes arise in various problems in computer graphics, such as hidden surface removal [59] or shadow computations [52, 50], so a better understanding of the size of the silhouette of polyhedra directly improves the theoretical complexity of algorithms in computer graphics.

Practical observations, supported by an experimental study by Kettner and Welzl [77], suggest that the number of silhouette edges of a polyhedron is often much smaller than the total number of edges. In the same paper, they proved that a polyhedral approximation of a sphere with Hausdorff distance ϵ has $\Theta(1/\epsilon)$ edges, and a random orthographic silhouette of such a polyhedron has size $\Theta(1/\sqrt{\epsilon})$.

In this paper, we investigate the worst-case size of the silhouette of a polytope observed under orthographic projection. We prove that some classes of polytopes have orthographic silhouettes with sub-linear complexity in the worst-case. We also give examples with linear-size silhouette when some of our conditions are not satisfied.

Our approach is to consider the orthogonal projection of the polytope on a plane, since the boundary of the projected polygon is the projection of the silhouette. We measure the length of the boundary of this polygon, which we call the *apparent length* of the polytope. First we show that all silhouettes of a triangulated fat object with n edges of length $\Theta(1)$ have apparent length $O(\sqrt{n})$. Secondly we derive bounds on the number of silhouette edges, using an additional condition on the repartition of the directions of the edges.

This paper is organized as follows. In Section 3.1, we review some examples of ill-shaped polytopes with silhouettes of linear complexity. Next, Section 3.2 studies the

apparent length of the silhouette, and Section 3.3 relates it to the number of silhouette edges. Finally, Section 3.4 discusses extensions and applications of our results.

3.1 Examples

The goal of this paper is to find conditions under which polytopes have sub-linear sized silhouettes in the worst-case. This section examines three examples of ill-shaped polytopes with silhouette of linear complexity, and identifies the reasons for this behavior.

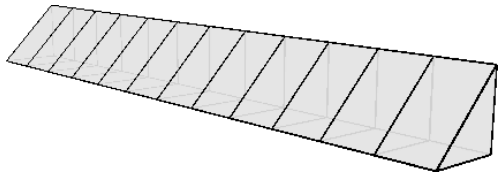


Figure 3.1: A non-fat triangulated polytope with bounded-length edges.

The example of Fig. 3.1 is characteristic of polytopes that are much longer along one dimension than along the others. This kind of behavior can be ruled out by considering *fat* polytopes, i.e., polytopes such that the ratio of the radius of the smallest enclosing to the largest enclosed sphere is $O(1)$.

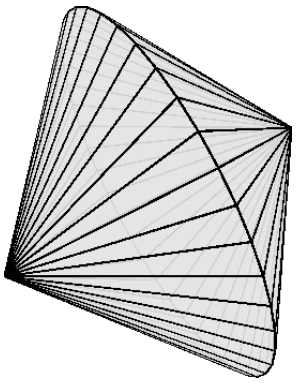


Figure 3.2: A fat triangulated polytope with uneven edges.

Our second example (see Fig. 3.2) illustrates the impact of the length of the edges on the silhouette. The ratio of the length of the longest edge to the length of the smallest is $\Omega(n)$, where n is the total number of edges. To avoid such behavior, we require that our polytopes have *bounded-length edges*, i.e., that all edges are of length $\Theta(1)$.

Our last example, in Fig. 3.3, exhibits a linear-size silhouette due to faces with order n edges. We therefore consider polytopes with faces of bounded complexity. Without loss of generality, we assume that our polytopes are triangulated.

This set of conditions is minimal in the sense that each of the previous examples satisfies all but one condition.

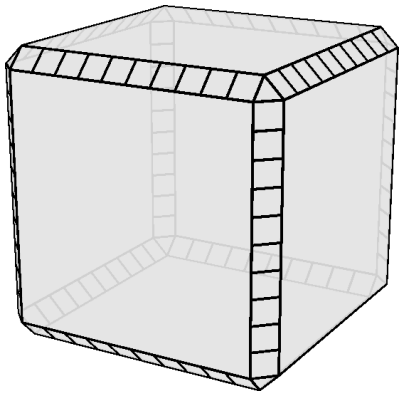


Figure 3.3: A fat polytope with bounded-length edges but with a face of large complexity.

In summary, in the rest of this paper we consider *triangulated fat* polytopes with *bounded-length edges*.

3.2 Apparent length

Recall that the apparent length of a silhouette is defined as the length of the orthogonal projection of the silhouette on a plane. In this section, we give bounds on the apparent length of the silhouette of a polytope.

We first recall a classical result on measures of convex sets. A proof can be found in [95][1].

Lemma 1. *Let $\mathcal{O}$ and $\mathcal{O}'$ be two convex objects in $\mathbb{R}^2$ (resp. $\mathbb{R}^3$) such that $\mathcal{O}$ contains $\mathcal{O}'$. Then the length (resp. area) of $\partial\mathcal{O}$ is larger than that of $\partial\mathcal{O}'$.*

For a polytope $\mathcal{P}$, let $\mathcal{A}(\mathcal{P})$ denote its surface area, and $\mathcal{L}(\mathcal{P})$ be the maximum apparent length of its silhouettes. The following lemma relates those two quantities.

Lemma 2. *If $\mathcal{P}$ is a fat polytope, then $\mathcal{L}(\mathcal{P}) = \Theta(\sqrt{\mathcal{A}(\mathcal{P})})$.*

Proof. Let r be the radius of the largest enclosed sphere of $\mathcal{P}$, and λr be the radius of the smallest enclosing sphere. Since $\mathcal{P}$ is fat, λ is $\Theta(1)$.

First, we apply Lemma 1 to $\mathcal{P}$ and its biggest enclosed sphere, and to $\mathcal{P}$ and its smallest enclosing sphere. This yields that $\mathcal{A}(\mathcal{P}) = \Theta(r^2)$. Next, consider an orthogonal projection of $\mathcal{P}$. Each of the two spheres projects onto a circle whose radius is the same as the radius of the corresponding sphere. Since the projection of $\mathcal{P}$ is convex, we can apply Lemma 1 to these circles and the boundary of that projection, and obtain that the length of that boundary is $\Theta(r)$. Taking the maximum over all possible orthogonal projections, we obtain that $\mathcal{L}(\mathcal{P}) = \Theta(r)$. It follows that $\mathcal{L} = \Theta(\sqrt{\mathcal{A}(\mathcal{P})})$.

□

The next lemma bounds the area of a polytope with bounded-length edges.

[1] In fact, the proof in [95] is much more general than our statement, and applies to any Minkowski measure, in any dimension.

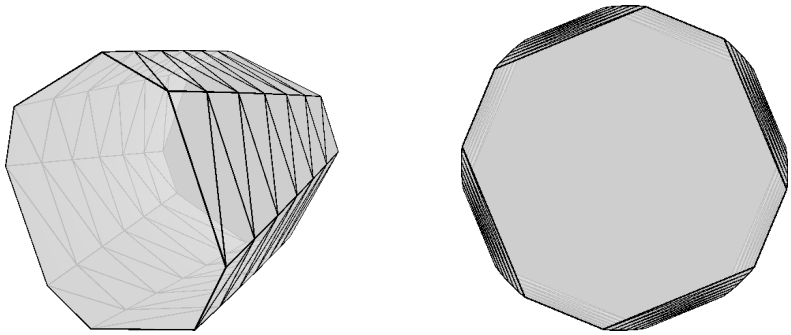


Figure 3.4: A triangulated fat polytope with bounded-length edges and a linear-size silhouette. The front and back faces, of complexity $O(\sqrt{n})$, were not triangulated for clarity.

Lemma 3. *If $\mathcal{P}$ is a triangulated polytope with bounded-length edges, then $\mathcal{A}(\mathcal{P}) = O(n)$.*

Proof. Since the polytope has bounded-length edges, the area of any of its triangles is $O(1)$. By Euler's formula, a triangulated polytope with n edges has $O(n)$ triangles, and the result follows. □

We can conclude with the following corollary, directly deduced from Lemmas 2 and 3.

Corollary 4. *If $\mathcal{P}$ is a triangulated fat polytope with n bounded-length edges, then the apparent length of any of its silhouettes is $O(\sqrt{n})$.*

3.3 Complexity of the silhouette

This section uses Corollary 4 to measure the complexity of the silhouette. To exploit the upper bound on the apparent length of the silhouette, we bound from below the contribution of silhouette edges to the apparent length. However, the contribution of an edge can be arbitrarily small, as it can be parallel to the direction of projection, and a triangulated fat polytope with bounded-length edges can have a linear number of such silhouette edges, as shown[2] in Fig. 3.4. Thus, we need to bound from above the number of silhouette edges that can be close to the direction of projection.

We give two distinct additional conditions that ensure a sub-linear size for the silhouette. The first one is local.

Lemma 5. *Let ϵ be a positive real number and $\mathcal{P}$ be a polytope with n bounded-length edges such that any two incident edges make an angle in the interval $[\epsilon, \pi - \epsilon]$. Then, any silhouette of $\mathcal{P}$ has $O(\mathcal{L}(\mathcal{P}))$ edges.*

Proof. Let $\vec{\delta}$ be a viewing direction . As any two incident edges make an angle in the interval $[\epsilon, \pi - \epsilon]$, two consecutive silhouette edges contribute $\Omega(\epsilon)$ to the apparent length of the silhouette. It follows that the number of silhouette edges is $O(\mathcal{L}(\mathcal{P}))$. Note that the constant in the O depends on ϵ. □

[2]See Appendix 3.5.1 for details.

Notice that if $O(\mathcal{L}(\mathcal{P}))$ edges do not satisfy the angle hypothesis, the same reasoning can be applied to the remaining edges on the silhouette, and the result of Lemma 5 still holds.

Combining Corollary 4 with Lemma 5 yields:

Theorem 6. *Let ϵ be a positive real number and $\mathcal{P}$ be a triangulated fat polytope with n bounded-length edges such that any two incident edges make an angle in the interval $[\epsilon, \pi - \epsilon]$. Then, any silhouette of $\mathcal{P}$ has $O(\sqrt{n})$ edges. This still holds if $O(\sqrt{n})$ edges do not satisfy the angle hypothesis.*

The second condition corresponds to a regular repartition of the directions of the edges of the polytope and is thus global. The idea is that if the directions of the edges do not accumulate along a few directions, the number of edges almost collinear with any direction is bounded, and so is the complexity of the silhouette. The meaning of this accumulation hypothesis is explained in the next Lemma.

Lemma 7. *Let $\mathcal{P}$ be a polytope with n bounded-length edges and apparent length $O(\sqrt{n})$ such that for any direction $\vec{\delta}$, the number of edges of $\mathcal{P}$ making an angle smaller than $\Theta(n^{-1/6})$ with $\vec{\delta}$ is $O(n^{2/3})$. Then any silhouette of $\mathcal{P}$ has $O(n^{2/3})$ edges.*

Proof. Let us fix a direction $\vec{\delta}$, and let α be a real number. We count separately the silhouette edges that make an angle greater than α with $\vec{\delta}$, and the others, and find the value of α yielding the best trade-off.

If we represent the set of directions by a unit sphere, the directions that make an angle smaller than α with $\vec{\delta}$ form a spherical cap of area $\Theta(\alpha^2)$. The sphere can be covered by $\Theta(1/\alpha^2)$ such spherical caps and the directions of the n edges are distributed over the sphere, so one of the caps has to contain $\Omega(\alpha^2 n)$ edge directions. This means that, for some viewing direction, there are $\Omega(\alpha^2 n)$ edges that make an angle less than α. Thus, the best we can ask is that the number of silhouette edges having a negligible contribution to the apparent length is $O(\alpha^2 n)$.

Let k denote the number of silhouette edges that make an angle greater than α with $\vec{\delta}$. The contribution of these k edges to the apparent length of the silhouette is $\Omega(k\alpha)$. Thus, $k = O(\mathcal{L}/\alpha) = O(\sqrt{n}/\alpha)$.

If we ask that at most $O(\alpha^2 n)$ edges of the polytope make an angle less than α with any given direction, then the complexity of the silhouette is bounded from above by

$$O\left(\sqrt{n}/\alpha + \alpha^2 n\right).$$

The best trade-off one can achieve is to choose $\sqrt{n}/\alpha = \Theta(\alpha^2 n)$, which means $\alpha = \Theta(n^{-1/6})$. In that case, the number of silhouette edges is $O(n^{2/3})$, and the regular distribution assumption is the one mentioned in the statement of the lemma. □

Note that the proof of Lemma 7 establishes a more general result: a weaker condition on the repartition of the directions of the edges still yields a sub-linear bound on the complexity of the silhouette, which is in between $O(n^{2/3})$ and $O(n)$. Besides, if the repartition condition is satisfied for a given direction $\vec{\delta}$, then the orthographic silhouette along this direction has a complexity $O(n^{2/3})$.

Combining Corollary 4 with Lemma 7 yields:

Theorem 8. *Let $\mathcal{P}$ be a triangulated fat polytope with n bounded-length edges such that for any direction $\vec{\delta}$, the number of edges of $\mathcal{P}$ making an angle smaller than $\Theta(n^{-1/6})$ with $\vec{\delta}$ is $O(n^{2/3})$. Then any silhouette of $\mathcal{P}$ has $O(n^{2/3})$ edges.*

The requirements on the polytopes in Lemmas 5 and 7 are strong, and may describe empty classes of polytopes. For the case of Theorem 6, Appendix 3.5.2 describes a class of polytopes that satisfy the requirements of the theorem when we allow $O(\sqrt{n})$ edges to miss the angular condition. However, to the best of our knowledge, whether there exists or not a polytope meeting the conditions of Lemma 7 is an open question.

3.4 Discussion

This section discusses our results, giving extensions as well as possible applications.

To begin with, notice that, in the results of Sections 3.2 and 3.3, the fatness assumption can be weakened: Lemma 2 holds for any polytope $\mathcal{P}$ with bounded-length edges that satisfies $d(\mathcal{P})^2 = O(\mathcal{A}(\mathcal{P}))$, where $d(\mathcal{P})$ is its diameter. This means having a fat orthogonal projection with the same diameter, i.e., to be fat along at least two dimensions.

Next, to extend our approach to the perspective case, the distance from the object to the viewpoint has to be taken into account. When the viewpoint is far from the polytope, the perspective case should behave as the orthographic case. But when the viewpoint is close to the polytope, the perspective projection introduces distortion: the length of the projection of a silhouette edge depends on its distance to the center of the view. Also, the global hypothesis on the distribution of the directions of the edges has to be adjusted accordingly.

The results of this paper are only a first step toward the understanding of the complexity of silhouettes, but they still have promising applications.

A first application is the computation of shadow boundaries. Drettakis and Duguet [52, 50] propose a solution based on a *visibility skeleton* restricted to the visual events generated by a punctual light source. In their detailed report [50], they show that their algorithm has complexity $O(ns_n)$, where n is the size of the polyhedron that casts a shadow, and s_n the size of its silhouette. Even the orthographic case is of interest, since it corresponds to a light source at infinity, a simple sun model for instance.

A second application is hidden surface removal, which has a long history as a problem difficult to address practically [48]. A solution proposed by Efrat et al. [59] is to render first the silhouettes of the objects, and then optimize the rendering in the single-object regions. They estimate the number of combinatorial changes to the rendered silhouettes of polytopes when the viewpoint moves along a line or an algebraic curve. Depending on the motion, this number depends either linearly or quadratically on the silhouette complexity, which they bound from above by the complexity of the polytope. Extension of our work to the perspective case would thus yield a direct improvement of their bounds.

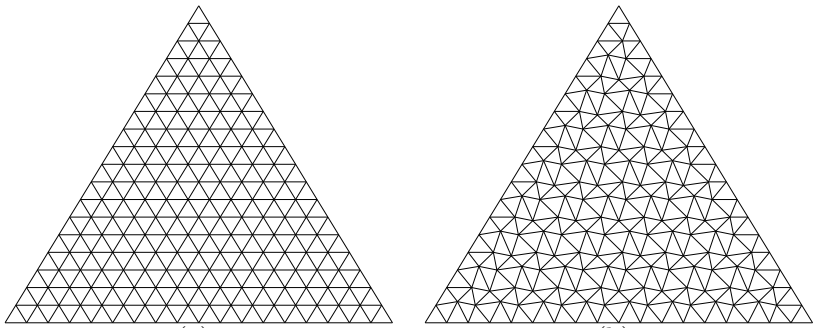


Figure 3.5: A face of a polytope for Theorem 6.

3.5 Appendix

3.5.1 The cylinder example

This appendix details the (not necessarily intuitive) construction of the polyhedron of Fig. 3.4.

First, start with two almost regular polygons of diameter $\Theta(\sqrt{n})$ facing each other at distance $\Theta(\sqrt{n})$. These polygons both have $\Theta(\sqrt{n})$ edges of alternating length $1 \pm \epsilon$. Note that corresponding edges on the two polygons have different length. Next, connect each pair of corresponding edges by a strip of length $\Theta(\sqrt{n})$ and width $\Theta(1)$. Thus, each strip is an almost rectangular trapezoïd.

We then triangulate the extremal polygons and the strips, inserting points on the edges and inside the faces, so that the triangles have edges of length $\Theta(1)$. A triangulation of the $\Theta(\sqrt{n})$-gons can be made with $\Theta(n)$ triangles. There are $\Theta(\sqrt{n})$ strips, each triangulated with $\Theta(\sqrt{n})$ triangles. So the total size of the polyhedron is $\Theta(n)$.

Now, when looking along the axis of this cylinder-like polytope, the silhouette is made of $\Theta(\sqrt{n})$ polygon edges and all the sides of the strips, that is $\Theta(\sqrt{n})$ collections of $\Theta(\sqrt{n})$ edges. This is thus an example of a polytope with a linear-size silhouette. Yet, this polytope is fat with an aspect ratio close to $\sqrt{2}$, with triangular faces, and bounded-length edges.

3.5.2 A polytope for Theorem 6

This appendix describes a class of polytopes that satisfy the requirements of Theorem 6 when we allow $O(\sqrt{n})$ edges to miss the angular condition.

Start with a regular tetrahedron, and triangulate each of its faces regularly as shown in Figure 3.5(a). Then, for each face, perturb the interior points as shown in Figure 3.5(b). The vertices on every second horizontal line are moved alternatively upward and to the left, and the remaining vertices are moved in the direction opposite to that of their top-left neighbour. The scale of the perturbation is chosen to be proportional to the size of the triangles, so that the angles between edges do not depend on the size of the triangulation.

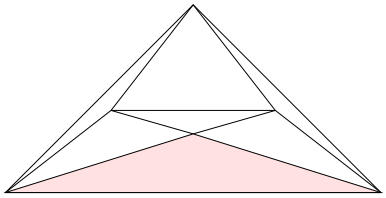


Figure 3.6: A triangulation with long edges.

Notice that this perturbation is 2-periodic along each of the 3 main directions of the triangulation

Now, we slightly inflate the faces of the perturbed polytope so that no two triangles are coplanar. The resulting polytope is fat, triangulated, has bounded-length edges and only the $O(\sqrt{n})$ edges included in the edges of the initial tetrahedron are aligned with some of their neighbours.

Notice that a similar perturbation scheme applied to the triangulation of the lateral surface of the cylinder in Figure 3.4, brings the size of its silhouette from $\Theta(n)$ down to $\Theta(\sqrt{n})$.

3.6 Bonus: a lower bound

It is also possible to give a lower bound on the complexity of the silhouette of a polytope. However, the assumptions are not exactly the same. Indeed, recursively applying the triangulation of Fig. 3.6, it is possible to triangulate an equilateral triangle with only edges of length at least one quarter of the size of the initial triangle. Applying this to a tetrahedron gives an exemple of a triangulated fat polytope with bounded-length edges and a constant size silhouette.

Theorem 9. *Let $\mathcal{P}$ be a fat triangulated polytope with n bounded-length edges and fat faces. Then all the silhouettes are of complexity $\Omega(\sqrt{n})$.*

Proof. Assume the length of the edges is $\Theta(1)$. Since all the faces are fat triangles, their area is $\Theta(1)$. The total surface of the polytope is then $\Theta(n)$; its diameter is $\Omega(\sqrt{n})$ (Lemma 1); its apparent length is $\Omega(\sqrt{n})$ and the complexity of the silhouette is $\Omega(\sqrt{n})$ since the length of the edges is $O(1)$. Notice that the fatness assumption can be reduced to the same 2-dimensional fatness discussed at the beginning of Section 3.4. □

Chapter 4

An Upper Bound on the Average Size of Silhouettes

Ces résultats ont été présentés au Symposium on Computational Geometry [63] et soumis à Discrete and Computational Geometry (en collaboration avec S. Lazard pour cette dernière version).

It is a widely observed phenomenon in computer graphics that the size of the silhouette of a polyhedron is much smaller than the size of the whole polyhedron. This paper provides, for the first time, theoretical evidence supporting this for a large class of objects, namely for polyhedra that approximate surfaces in some reasonable way; the surfaces may be non-convex and non-differentiable and they may have boundaries. We prove that such polyhedra have silhouettes of expected size $O(\sqrt{n})$ where the average is taken over all points of view and n is the complexity of the polyhedron.

4.1 Introduction

The silhouette of a polyhedron with respect to a given viewpoint is, roughly speaking, the set of edges incident to a front and a back face. Silhouettes arise in various problems in computer graphics such as hidden surface removal and shadow computations (see [51, 52, 59] for some recent references) and algorithms to compute them efficiently have been well-studied (see the survey by Isenberg et al. [74]). They are important in shape recognition; Sander et al. [94] claim that the silhouette "is one of the strongest visual cues of the shape of an object".

It is a widely accepted fact that the silhouette of a polyhedron is usually much smaller than the whole polyhedron. Sander et al. [94], for instance, state the largely repeated claim that the silhouette of a mesh is often of size $\Theta(\sqrt{n})$ where n is the number of faces of the mesh. An experimental study by Kettner and Welzl [77] confirms this for a set of realistic objects. This experimental study was extended by McGuire [81] to a larger database of larger objects for which the observed size of the silhouette is approximately $n^{0.8}$.

There are few theoretical results supporting these observations. Kettner and Welzl [77] prove that a convex polyhedron that approximates a sphere with Hausdorff distance

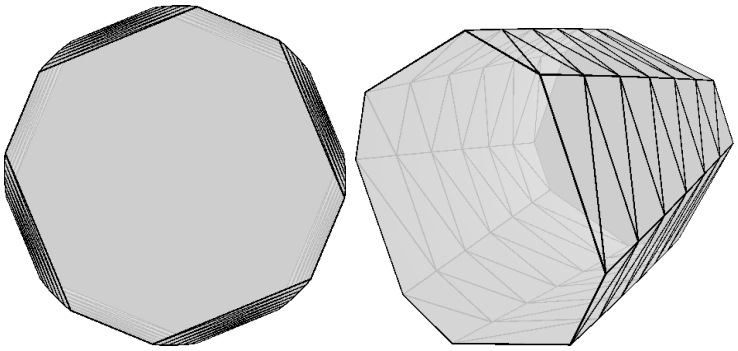


Figure 4.1: A worst-case linear silhouette (left) of a polyhedron approximating a cylinder.

ε has $\Theta(1/\varepsilon)$ edges, and a random orthographic projection of such a polytope has $\Theta(1/\sqrt{\varepsilon})$ silhouette edges. In Section 3.2, we give conditions under which it could easily be proved that the average silhouette of a *convex* polyhedron has size $O(\sqrt{n})$ and in Section 3.3 we give additional conditions under which the worst-case size is provably sub-linear.

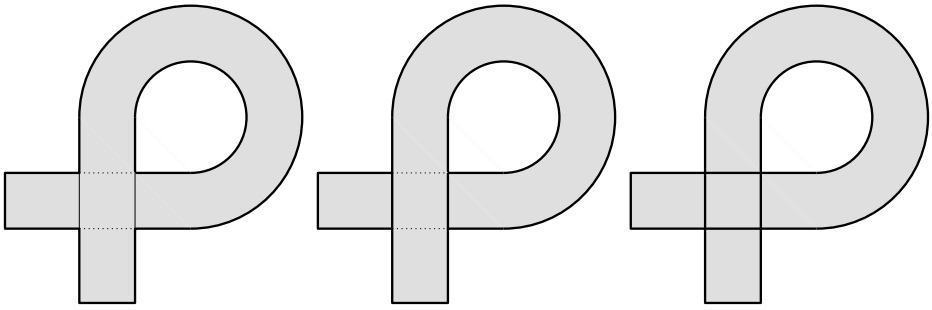
The goal of this paper is to study the average silhouette size of *non-convex* polyhedra. Convexity is a very strong assumption, which was crucial in the previous theoretical results. Here, rather, we assume that the polyhedron is a good approximation of some fixed (not necessarily convex) surface. Notice that it is very difficult to guarantee anything on the *worst-case* complexity of the silhouette of a polyhedron unless it approximates a strictly convex surface. In Section 3.3, we give an example of a polyhedral approximation of a section of a cylinder with worst-case silhouette size $\Theta(n)$ (see Figure 4.1). Moreover, this example can be modified in such a way that the surface is smooth, and its polyhedral approximation is as "nice" as one might hope (for instance, it can be required that the faces are fat triangles that all have almost the same size).

In this paper we prove an upper bound on the *expected size* of the silhouette for random viewpoints. We prove that the silhouette of a polyhedron that approximates a surface in a reasonable way has expected size $O(\sqrt{n})$. Note that the average is taken over all viewpoints for a given surface, and not on a set of surfaces.

In Section 4.2, we define precisely the notion of silhouette for polyhedra and general surfaces. We then present and prove our main result in Section 4.3 and conclude in Section 4.4.

4.2 Definitions

The term silhouette has been used in the literature to represent several different notions, depending on the application, reflecting such issues as: is the object considered opaque or transparent? Is occlusion taken into account? Is one interested by what the eye perceives, *i.e.,* a plane curve, or by the space curve which gave birth to it? In the area of photography, for instance, a silhouette (also called apparent boundary) is defined as an outline of a solid object, as cast by its shadow, that appears dark against a light background (Figure 4.2(a)). In the field of computer vision, by contrast, the silhouette (also called rim, profile or contour generator) is roughly defined as the curve on the surface



(a) (b) (c)

Figure 4.2: Three different notions of silhouettes: (a) outline of a solid, as cast by its shadow, (b) rim of an opaque object, and (c) rim of a transparent object.

that separates front face regions from the back ones, either for opaque (Figure 4.2(b)) or for transparent (Figure 4.2(c)) objects.

In this paper we prove an upper bound on the size of the transparent silhouette; since such a silhouette contains the apparent boundary and the contour, our bounds also apply to all these types of silhouettes. In the rest of the paper the term silhouette will be used to mean transparent silhouette.

In the rest of this section we give a formal definition of silhouettes of polyhedra and then provide a definition for more general surfaces.

4.2.1 Polyhedra

The (transparent) *silhouette* of a polyhedron from a viewpoint (possibly at infinity) is the set of edges that are adjacent to a front face and a back face. A face is considered a front face if the angle between its normal vector and a vector from a point of the face to the viewpoint is acute, and a back face if that angle is larger than $\pi/2$. If the point of view is in the plane containing the face, we refer to the definition of silhouettes for the case of general surfaces. The normal vectors should point outwards, but what really matters is that the orientation is consistent for the two faces that share this edge, so this definition also applies to non-orientable (necessarily self-intersecting) polyhedra.

In this paper, we call complexity of a silhouette (of a polyhedron) its number of edges.

4.2.2 General surfaces

Our objective is to bound the size of the silhouette of a polyhedron. To achieve this goal, we need to relate the silhouette of the polyhedron to the silhouette of the surface it approximates, which means we need a definition of silhouettes that applies to a larger class of objects. Although this may seem unintuitive, we first define the silhouette as a set of rays, and then relate this to the more usual concept of a set of points on the surface.

Let S be a compact 2-manifold without boundary. It separates $\mathbb{R}^3$ in two non-empty open regions; call $\mathscr{O}$ and $\mathscr{O}'$ their closures (so $\mathscr{O} \cap \mathscr{O}' = S$ and $\mathscr{O} \cup \mathscr{O}' = \mathbb{R}^3$). Let V be a viewpoint not on S but possibly at infinity. The (transparent) *silhouette* of S from V

is the set of rays R starting from V that are tangent to S in a non-crossing way (R may cross S elsewhere). More formally, we require that there exists an open segment u of R that contains a connected component of $R \cap S$ and is contained either in $\mathscr{O}$ or $\mathscr{O}'$.

This definition defines a set of rays. The silhouette can also be seen as the trace of this set of rays on the surface. More precisely, for each ray R on the silhouette, we consider the closest point to V on each connected component of $R \cap S$ that satisfies the non-crossing property. This definition is consistent with the one given for the particular case of polyhedra, and is the one we will use in this paper.

For a given viewpoint at infinity, we define the (projected) *length* of the silhouette as the length (counted with multiplicity if several points have the same projection) of the projection of the silhouette, along the direction given by the viewpoint, on an orthogonal plane.

Remark. The definition of the silhouette can be extended to cases where S is not a 2-manifold, but an immersion of a compact 2-manifold. More precisely, we have a 2-manifold S' and an application $f : S' \to \mathbb{R}^3$ such that $S = f(S')$ and for any point on S' there exists a neighborhood U of that point such that U and $f(U)$ are homeomorphic. The local orientation is sufficient to decide whether R crosses S or not (note that more complicated things can happen than crossing or being tangent, even with smooth surfaces; for instance, the surface may ripple an infinite number of times in the neighborhood of a point, making it impossible to define on which side of S R is near the intersection point). This remark extends to the whole paper and, in particular, to Theorem 10. However, we do not give either a definition or a proof of this, as it would uselessly make everything more obscure.

4.3 Main results

Let S be a compact 2-manifold without boundary whose silhouettes have finite average length, silh(S), where the average is taken over all viewpoints at infinity. Let P_n be a polyhedron with n triangular faces, that is homeomorphic to S through $f_n : P_n \to S$, such that:

1. the length of any edge of P_n is at least $\frac{\alpha}{\sqrt{n}}$ and
2. for any point x on P_n, $d(x, f_n(x)) < \frac{\beta\, h(x)}{\sqrt{n}}$ where $h(x)$ is the smallest height of the triangle(s) of P_n that contain(s) x,

where α and β are two arbitrary positive numbers and $d()$ denotes the Euclidean distance.

Theorem 10. *The expected complexity of the silhouette of P_n is $O(\sqrt{n})$, where the average is taken over all viewpoints at infinity. More precisely, for any n, the expected complexity is at most*

$$\left(15\,\beta + \frac{24}{\alpha}\, silh(S)\right)\sqrt{n}.$$

Note that the bound is valid for any n and any polyhedron P_n satisfying the above assumptions. Note also that the bound depends on S only by the average length of its silhouette.

We first clarify the meaning of the hypotheses on P_n and their implications. We then prove Theorem 10 in Section 4.3.2. We finally show in Section 4.3.3 how Theorem 10 can be generalized to surfaces with boundary and viewpoints at finite distance. In particular, we prove the following result.

Let S' be any compact two-manifold with boundary of finite length and whose silhouette has finite average length (taken over all viewpoints at infinity).

Theorem 11. *Any mesh P_n with n triangular faces that approximates S' according to Hypotheses 1 and 2 has a silhouette of expected complexity $O(\sqrt{n})$ when the viewpoint is chosen uniformly at random in a ball.*

4.3.1 Meaning of the hypotheses

Hypothesis 1 is here to avoid short edges. The main idea of the proof is to link the complexity of the silhouette to its length, and arbitrarily short edges would make this impossible. Now the $\frac{1}{\sqrt{n}}$ factor makes sense: intuitively, since the polyhedron has n faces, each face has area of order $\frac{1}{n}$, which means that the edges have length of order $\frac{1}{\sqrt{n}}$.

Hypothesis 2 is rather technical, and we discuss instead the meaning of the following two more intuitive hypotheses, which, together with Hypothesis 1, imply[3] Hypothesis 2.

3. The faces of P_n are fat.

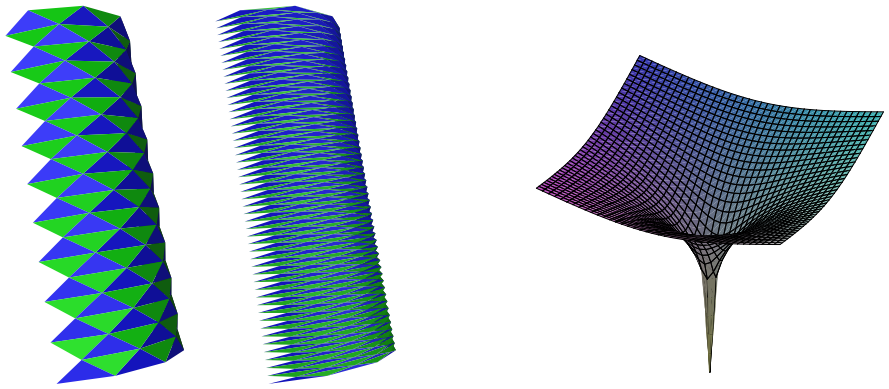
4. For any x on P_n, $d(x, f_n(x)) < \frac{\gamma}{n}$, where γ is some positive constant.

Hypothesis 3 is quite natural. Hypothesis 4 ensures that P_n approximates S. Furthermore, the $\frac{1}{n}$ factor is reasonable; indeed, in 2D, when considering a regular polygon with edge length $\Theta(\frac{1}{\sqrt{n}})$ inscribed in a circle of radius 1, the maximal distance between a point on the polygon and the circle is $\Theta(\frac{1}{n})$. The situation is the same in 3D. Basically it means that the error when approximating the surface with a plane is of the second order.

Our hypotheses (1-3-4 or 1-2) ensure that the homeomorphism f_n has good properties, that is that, roughly speaking, the polyhedron can be obtained by only a small perturbation of the surface while keeping the normal vectors in approximately the same directions. This is crucial for our proof since otherwise, for example, a cylinder can be approximated by a lantern of Schwarz [96] (see Figure 4.3(a)) whose silhouette has expected complexity $\Theta(n)$ and unbounded length.

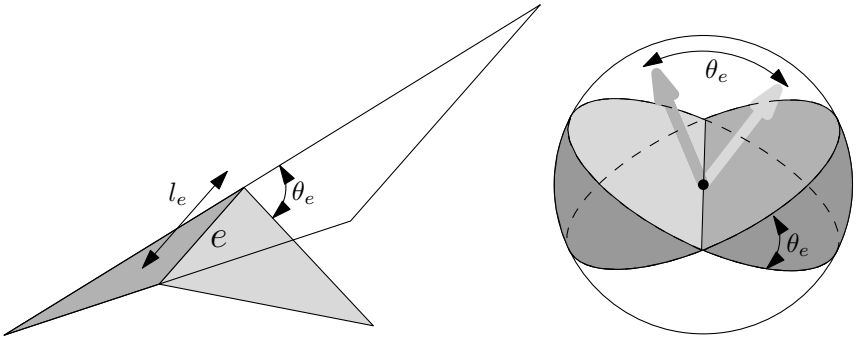
Notice that the existence of polyhedra with arbitrarily large number of edges that approximate the surface according to these hypotheses is a constraint on the surface. Not every surface admits such an approximation (think of the neighborhood of 0 in the surface defined by $z = (x^2 + y^2)^{1/8}$ as shown in Figure 4.3(b)). However, the class of surfaces for which such approximations exist is quite large. It includes, in particular, smooth surfaces and polyhedra with fat faces.

[3] Indeed, for any x in P_n, Hypotheses 1 and 3 imply that $h(x) \geqslant \delta/\sqrt{n}$ for some positive constant δ; Hypothesis 2 then follows from Hypothesis 4 since $h(x)/\sqrt{n} \geqslant \delta/n \geqslant \delta/\gamma \cdot d(x, f_n(x))$.



(a) (b)

Figure 4.3: (a) Two half lanterns of Schwarz (courtesy of Boris Thibert). (b) A surface that cannot be approximated with the right properties.



(a) (b)

Figure 4.4: (a) Length and dihedral angle of an edge; (b) set of directions for which e is on the silhouette.

4.3.2 Proof of Theorem 10

We consider a point of view chosen randomly at infinity. We call l_e the length of an edge e of polyhedron P_n and θ_e the exterior dihedral angle associated to e (see Figure 4.4).

Let T_e denote the union of the two triangles adjacent to edge e (including e but not the other edges). For any part $\mathcal{R}$ of S, let silh($\mathcal{R}$) be the average length of the part of the silhouette of S that lies in $\mathcal{R}$.

We first recall a classical formula on the expected size of silhouettes which can also be found, for instance, in [81].

An edge e is on the silhouette if the direction of view is in the dark area of the sphere of directions of Figure 4.4(b). The angular measure of this region is $4\theta_e$, which means that the probability for e to be on the silhouette is θ_e/π. The expected number of edges on the silhouette is thus

$$E = \frac{1}{\pi} \sum_{\text{edge } e} \theta_e.$$

We now state our main lemma. The general idea of this lemma is that under strong

hypotheses (S has bounded curvature, the edges have length $\Theta(\frac{1}{\sqrt{n}})$, and Hypotheses 3 and 4 are satisfied), one can prove that $\theta_e \leqslant \frac{C}{\sqrt{n}}$ for some constant C. In cases where this inequality does not hold, edge e is near some kind of edge of the surface, or at least some feature that will appear quite often on the silhouette and we are going to charge this edge to the silhouette of S.

Lemma 12. *For any edge e on P_n,*

$$\theta_e \leqslant \frac{C}{\sqrt{n}} + \frac{8\pi}{l_e}\, silh\left(f_n\left(T_e\right)\right) \quad with \quad C = 31.3\,\beta.$$

Theorem 10 follows from Lemma 12. Indeed, since P_n has $\frac{3n}{2}$ edges, each of length at least $\frac{\alpha}{\sqrt{n}}$ (by Hypothesis 1), we get that the expected complexity of the silhouette is

$$E \leqslant \frac{1}{\pi}\frac{3n}{2}\frac{C}{\sqrt{n}} + 8\,\frac{\sqrt{n}}{\alpha}\,3\ \mathrm{silh}(S),$$

because $\sum_{\text{edge } e} \mathrm{silh}\left(f_n\left(T_e\right)\right) = 3\ \mathrm{silh}(S)$ since the length of the silhouette of S that lies in the image (through f_n) of a triangle is counted three times (once per edge). Hence,

$$E \leqslant \left(15\,\beta + \frac{24}{\alpha}\,\mathrm{silh}(S)\right)\sqrt{n} = O\left(\sqrt{n}\right).$$

Proof of Lemma 12. The idea of the proof is as follows. Consider the set of directions for which e is on the silhouette. We first construct a subset Ω of these directions whose measure is a constant times $\theta_e - \frac{C}{\sqrt{n}}$ (see Figure 4.5). We then prove a lower bound on the length of the silhouette of $f_n(T_e)$ for all these directions, and deduce the result.

Let C be a positive constant, whose value will be defined later (see Equation 2). For any edge e on P_n, we can assume that $\theta_e - \frac{C}{\sqrt{n}} > 0$ since, otherwise, $\theta_e \leqslant \frac{C}{\sqrt{n}}$ and there is nothing else to prove.

The set of directions for which e is on the silhouette is the set of directions between the planes defined by the faces adjacent to e. Rotate each face about e by an angle of $\frac{C}{2\sqrt{n}}$ so that the exterior dihedral angle decreases by $\frac{C}{\sqrt{n}}$ (see Figure 4.5(a)). Ω is defined to be the set of directions between these two new planes that make an angle larger than $\pi/3$ with the line supporting e; Figure 4.5(b) shows one component of Ω, the other one consists of the symmetric set of opposite directions. The measure of the set of directions between these two planes is $4\,(\theta_e - \frac{C}{\sqrt{n}})$. Restricting this set of directions to those that make an angle larger than $\pi/3$ with the line supporting e, we get, by integrating on the sphere of directions, that the measure of Ω is $2\,(\theta_e - \frac{C}{\sqrt{n}})$.

The remaining step uses the property, which we prove in Corollary 14, that for all the directions in Ω, the silhouette of $f_n(T_e)$ has length at least $l_e/4$. Assuming this temporarily, we sum this inequality over Ω. The smaller side of the inequality is $2\frac{l_e}{4}(\theta_e - \frac{C}{\sqrt{n}})$. The larger side is the integral of the length of the silhouette of $f_n(T_e)$ over all directions in Ω, which is smaller than this same integral over all directions, that is $4\pi\,\mathrm{silh}(f_n(T_e))$. Hence $4\pi\,\mathrm{silh}(f_n(T_e)) \geqslant \frac{l_e}{2}(\theta_e - \frac{C}{\sqrt{n}})$, which concludes the proof. □

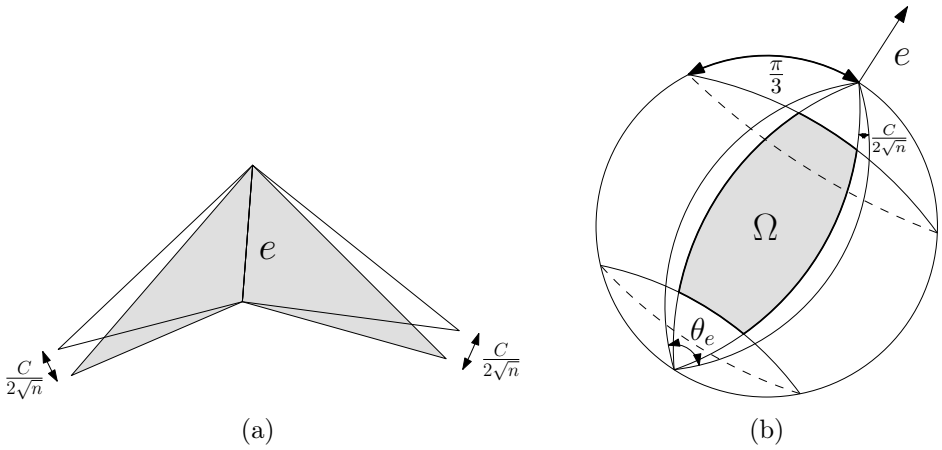


Figure 4.5: Construction of Ω.

We now state a lemma and its corollary which we used in the proof of Lemma 12 under the hypothesis that $\theta_e - \frac{C}{\sqrt{n}} > 0$. We can thus assume in the sequel that this property holds.

Let e' be the segment obtained by clipping from e all the points at distance less than $\frac{l_e}{4}$ from its extremities. Refer now to Figures 4.6(a)–(b).

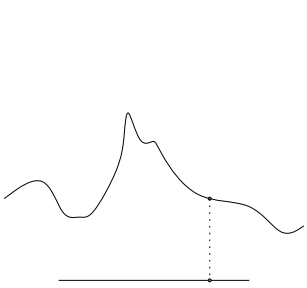
Lemma 13. *Any line with direction $d \in \Omega$ that intersects e' can be translated in a direction orthogonal to e and d until it becomes tangent to S in $f_n(T_e)$.*

Corollary 14. *For any direction d in Ω, the silhouette of $f_n(T_e)$ has length at least $\frac{l_e}{4}$.*

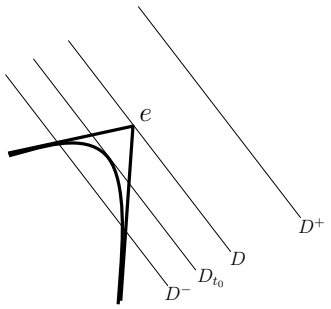
Proof. Consider the projection of e' and of the silhouette of $f_n(T_e)$ onto a plane orthogonal to d (see Figure 4.6(a)). It follows from Lemma 13 that, in that plane, each point on the projection of e' maps to a point on the projected silhouette in the direction orthogonal to e'. Hence, the projected silhouette is longer than the projection of e', which is at least $\frac{\sqrt{3}}{2}$ times the length of e' since d makes an angle of at least $\pi/3$ with e'. Thus the silhouette of $f_n(T_e)$ has length at least $\frac{\sqrt{3}}{2}\frac{l_e}{2} > \frac{l_e}{4}$. □

Proof of Lemma 13. Let D denote a line with direction $d \in \Omega$ that intersects e'. Let T_1 and T_2 denote the two triangles adjacent to e and let h_1 and h_2 denote their respective smallest heights. Let $\chi_i = \beta h_i/\sqrt{n}$, $\chi^+ = \max(\chi_1, \chi_2)$, and $\chi^- = \min(\chi_1, \chi_2)$. Refer now to Figure 4.6(b). We call D_t, $t \in [-\chi^-, \chi^+]$, the line obtained by translating D at distance $|t|$ in a direction orthogonal to the plane defined by e and d; positive values of t correspond to lines in the half-space bounded by the plane defined by e and D, and not containing T_e; negative values of t correspond to lines in the other half-space. For clarity, we denote $D_{-\chi^-}$ by D^- and D_{χ^+} by D^+.

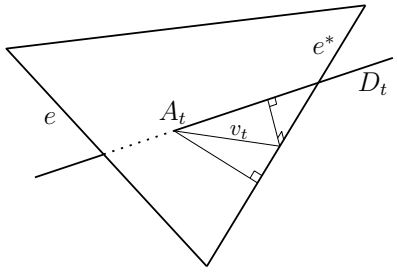
By construction, D^+ is at distance χ^+ from T_e. Thus D^+ does not intersect $f_n(T_e)$, by Hypothesis 2. We prove that D^- intersects $f_n(T_e)$ and that no line D_t intersects the boundary of $f_n(T_e)$. This will imply that, sweeping D_t from D^+ to D^-, the first line D_{t_0} that intersects $f_n(T_e)$ is tangent to $f_n(T_e)$ at one of its interior point, which will conclude the proof.



(a) Orthogonal projection along d of e' and of the silhouette of $f_n(T_e)$.



(b) For the definition of D_t.



(c)

Figure 4.6: For the proofs of Lemma 13 and Corollary 14.

We first prove that no line D_t intersects the boundary of $f_n(T_e)$. In other words, we prove that, for each edge e^* on the boundary of T_e, no line D_t intersects $f_n(e^*)$. Let T_i be the triangle (of T_e) containing e^*. By Hypothesis 2, it is sufficient to prove that the distance between D_t and e^* remains greater than or equal to χ_i for all t.

First notice that it is sufficient to prove that the distance between D_t and e^* remains greater than or equal to χ_i for all $t \in [-\chi^-, 0]$. Indeed, then, the distance between $D_0 = D$ and e^* is at least χ_i, and the distance between D_t and e^* increases for $t \geqslant 0$ (see Figure 4.6(b)).

Let Γ be the smallest angle d can make with the plane containing T_i and refer to Figure 4.6(c). Let A_t be the point of intersection between D_t and the plane containing T_i and v_t be the distance between A_t and the point on e^* that realizes the distance between D_t and e^*. The distance between D_t and e^* satisfies $d(D_t, e^*) \geqslant v_t \sin\Gamma \geqslant d(A_t, e^*) \sin\Gamma$. Hence, for proving that $d(D_t, e^*) \geqslant \chi_i$ for $t \leqslant 0$, it is sufficient to prove that $d(A_t, e^*) \geqslant \frac{\chi_i}{\sin\Gamma}$ for $t \leqslant 0$. We set $a = \frac{\chi_i}{\sin\Gamma}$ to simplify the notation.

We just proved that $d(A_t, e^*) \geqslant a$ implies $d(D_t, e^*) \geqslant \chi_i$ (for all t). Conversely, we have that $d(D_t, e^*) < \chi_i$ implies $d(A_t, e^*) < a$. Similarly, for edge e, we get that $d(D_t, e) < \chi_i$ implies $d(A_t, e) < a$. By definition of D_t, we have that $d(D_t, e) < \chi_i$ for $t \leqslant 0$, thus $d(A_t, e) < a$ for $t \leqslant 0$. Furthermore, the angle between e and segment $\{A_t \mid t \in [-\chi^-, \chi^+]\}$ is at least $\pi/3$ because this angle is at least the angle between their orthogonal projection on the plane defined by e and D that is the angle between e and D since all A_t lie in the plane spanned by D_t which projects on D; the lower bound of $\pi/3$

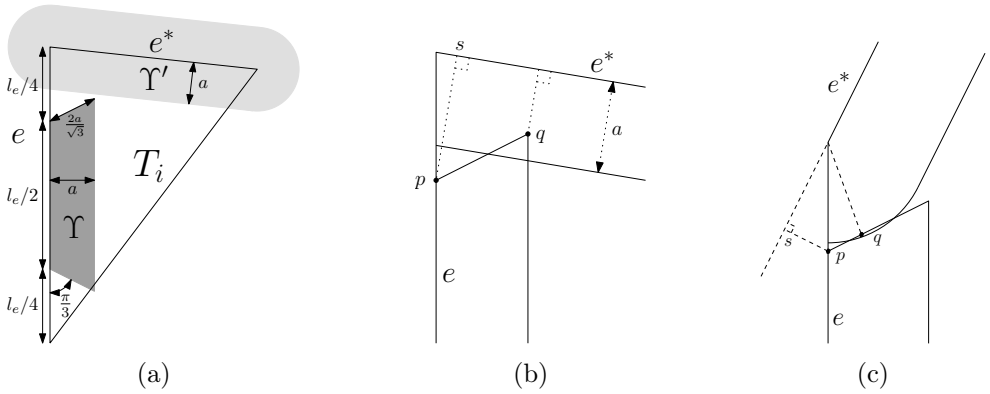


Figure 4.7: For the proof of Lemma 13.

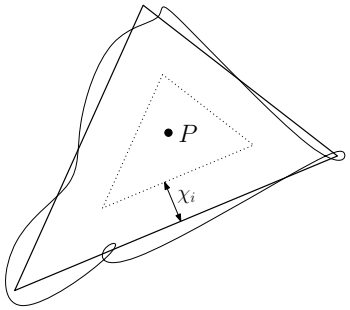


Figure 4.8: Projection of ∂T_i, $f_n(\partial T_i)$ and D^-.

follows since the angle between e and D is at least $\pi/3$ by definition of Ω. Hence, the locus of points A_t, for $t \leqslant 0$, lies in a region, denoted Υ, shown in dark gray in Figure 4.7(a). For proving that $d(A_t, e^*) \geqslant a$ for $t \leqslant 0$, it is thus sufficient to prove that this region does not intersect the set, denoted Υ', of points at distance less than a from e^* (shown in light gray in Figure 4.7(a)).

Referring to Figures 4.7(b)–(c), let p be the endpoint of e' the closest to e^* and s be its projection on the line supporting e^*. If the two regions Υ and Υ' intersect, there exists a point q in the intersection that is at distance less than or equal to $\frac{2}{\sqrt{3}}a$ from p and at distance less than or equal to a from e^*; thus $d(p,s) \leqslant d(p,e^*) \leqslant d(p,q) + d(q,e^*) \leqslant (1 + \frac{2}{\sqrt{3}})a$. On the other hand, $d(p,s)$ is one fourth of one of the heights of the triangle T_i and thus is at least $\frac{h_i}{4}$. Hence, if the two regions intersect, then $\frac{h_i}{4} \leqslant \left(1 + \frac{2}{\sqrt{3}}\right) \frac{\chi_i}{\sin\Gamma}$. We postpone to Lemma 16 the proof that, with $C = 31.3\,\beta$, we have $\frac{h_i}{4} > \left(1 + \frac{2}{\sqrt{3}}\right) \frac{\chi_i}{\sin\Gamma}$, which implies that the two regions Υ and Υ' are disjoint. This concludes the proof that no line D_t intersects the boundary of $f_n(T_e)$.

We now prove that D^- intersects $f_n(T_e)$. Consider a projection, $p()$, along the direction d onto a plane orthogonal to d. We proved that, for any of the two triangles T_i, Υ is at distance at least χ_i from each edge $e^* \neq e$ of T_i. It follows that Υ lies in triangle T_i and thus that D_t intersects T_i for all $t \leqslant 0$. Therefore, D^- intersects T_i and is at distance at least χ_i from each edge $e^* \neq e$ of T_i, for $i = 1, 2$. Furthermore, D^- is at distance $\chi^- = \min(\chi_1, \chi_2)$

from e, by definition. We now consider the triangle T_i for which $\chi_i = \chi^-$. It follows that D^- is at distance at least χ_i from all three edges of T_i. Thus D^- projects to a point $P = p(D^-)$ inside triangle $p(T_i)$, at distance at least χ_i from the three edges of $p(T_i)$ (see Figure 4.8).

Roughly speaking, by Hypothesis 2, the curve $f_n(\partial T_i)$ is at distance less than χ_i from ∂T_i (the boundary of T_i) thus its projection $p(f_n(\partial T_i))$ is at distance less than χ_i from the edges of $p(T_i)$. It is thus intuitively clear that $p(D^-)$ intersects $p(f_n(T_i))$, and thus that D^- intersects $f_n(T_i)$ (and thus $f_n(T_e)$).

More formally, consider the application g_n from the triangle $p(T_i)$ to the plane containing it such that, for any point x in T_i, the point $p(x)$ is sent to the point $g_n(p(x)) = p(f_n(x))$. We first prove that the curves $p(\partial T_i)$ and $g_n(p(\partial T_i))$ are homotopic in $\mathbb{R}^2 \setminus P$. Consider the continuous application

$$\begin{array}{rccl} F: & \partial T_i \times [0,1] & \longrightarrow & \mathbb{R}^2 \\ & (x,\lambda) & \longrightarrow & \lambda\, p(x) + (1-\lambda)\, g_n(p(x)) = \lambda\, p(x) + (1-\lambda)\, p(f_n(x)). \end{array}$$

F is an homotopy between the curves $p(\partial T_i)$ and $g_n(p(\partial T_i))$ in $\mathbb{R}^2$. We prove that the image of F does not contain P, which yields the result. The triangle inequality gives

$$d(P, F(x,\lambda)) \geqslant d(P, p(x)) - d(F(x,\lambda), p(x)).$$

We have already proved that point P is at distance at least χ_i from $p(x)$ for all points x in ∂T_i. On the other hand, the distance between $p(x)$ and $p(f_n(x))$ is larger than or equal to the distance between $p(x)$ and their barycenter $F(x,\lambda)$, for any $\lambda \in [0,1]$. Hence

$$d(P, F(x,\lambda)) \geqslant \chi_i - d(p(x), p(f_n(x))).$$

Finally, since $d(p(x), p(f_n(x))) < \chi_i$ for all $x \in T_i$, by Hypothesis 2, we have that $d(P, F(x,\lambda)) > 0$ for all (x,λ). Hence, the image of F does not contain point P and thus the curves $p(\partial T_i)$ and $g_n(p(\partial T_i))$ are homotopic in $\mathbb{R}^2 \setminus P$.

Now, we can contract $p(\partial T_i)$ to a point while remaining in $p(T_i)$. Composing this with g_n gives a contraction of $g_n(p(\partial T_i))$ in $g_n(p(T_i))$. On the other hand, there is no contraction of $p(\partial T_i)$ in $\mathbb{R}^2 \setminus P$ (since P is in $p(T_i)$), thus there is no contraction of its homotopic curve $g_n(p(\partial T_i))$ in $\mathbb{R}^2 \setminus P$. Hence, there exists a curve that is contractible in $g_n(p(T_i))$ but not in $\mathbb{R}^2 \setminus P$. It follows that $g_n(p(T_i))$ is not included in $\mathbb{R}^2 \setminus P$. Hence P is in $g_n(p(T_i)) = p(f_n(T_i))$. Therefore, D^- intersects $f_n(T_i)$ and thus $f_n(T_e)$, which concludes the proof.

□

We finally prove the two following simple technical lemmas which complete the proof of Theorem 10. Recall that Γ is the smallest angle a direction $d \in \Omega$ can make with the plane containing T_i.

Lemma 15. $\sin\Gamma = \frac{\sqrt{3}}{2} \sin\frac{C}{2\sqrt{n}}$.

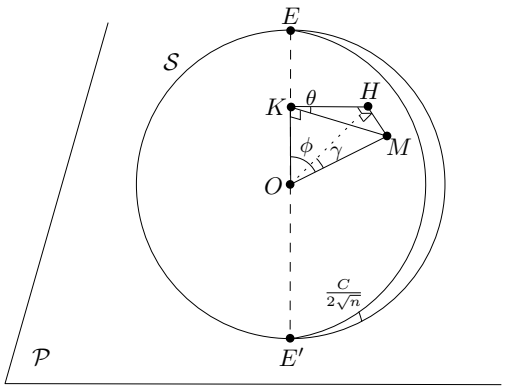


Figure 4.9: For the proof of Lemma 15

Proof. In the following, we identify the sphere of directions with a sphere $\mathcal{S}$ embedded in $\mathbb{R}^3$; let O denote its center. We assume that the embedding preserves directions (*i.e.*, for any direction d, the corresponding point M on $\mathcal{S}$ is such that d and OM have the same direction).

Refer to Figure 4.9. Let d be a direction in Ω and M be its corresponding point on $\mathcal{S}$. Consider one of the T_i and let $\mathcal{P}$ be the plane containing O and parallel to the plane containing T_i. Let H be the orthogonal projection of M onto plane $\mathcal{P}$. Let E and E' be the two points on $\mathcal{S}$ that correspond to the two (opposite) directions of segment e. Let K be the orthogonal projection of M (and H) onto the line EE'. Finally, let θ be the angle $\angle MKH$, ϕ be the angle $\angle MOK$, and γ be the angle $\angle MOH$.

It follows from these definitions that

$$\sin\gamma = \frac{HM}{OM} = \frac{HM}{KM}\,\frac{KM}{OM} = \sin\theta\,\sin\phi.$$

Now, the angle γ is also the angle between direction d and the plane that contains T_i. Thus $\Gamma = \inf_{d\in\Omega}\gamma$, by definition of Γ. The angle θ is the angle between the plane containing T_i and the plane containing e and d. It thus follows from the definition of Ω that $\inf_{d\in\Omega}\theta = \frac{C}{2\sqrt{n}}$ (see Figure 4.5(a)). The angle ϕ is the angle between d and the line containing e. It thus also follows from the definition of Ω that $\inf_{d\in\Omega}\phi = \frac{\pi}{3}$. In addition, since γ, θ and ϕ are in $[0, \frac{\pi}{2}]$, we have

$$\sin\Gamma = \inf_{d\in\Omega}\sin\gamma, \quad \inf_{d\in\Omega}\sin\theta = \sin\frac{C}{2\sqrt{n}} \quad \text{and} \quad \inf_{d\in\Omega}\sin\phi = \sin\frac{\pi}{3}.$$

Furthermore, the constraints on θ and ϕ in the definition of Ω are independent. Thus, the minima of θ and ϕ can be attained for the same direction d in Ω. It follows that

$$\inf_{d\in\Omega}(\sin\theta\,\sin\phi) = \inf_{d\in\Omega}\sin\theta\,.\,\inf_{d\in\Omega}\sin\phi.$$

We can thus conclude that

$$\sin\Gamma = \inf_{d\in\Omega}\sin\gamma = \inf_{d\in\Omega}\sin\theta\,\sin\phi = \inf_{d\in\Omega}\sin\theta\,\inf_{d\in\Omega}\sin\phi = \frac{\sqrt{3}}{2}\sin\frac{C}{2\sqrt{n}}.$$

□

Lemma 16. $\frac{h_i}{4} > \left(1 + \frac{2}{\sqrt{3}}\right) \frac{\chi_i}{\sin\Gamma}$ *with* $C = 31.3\,\beta$.

Proof. By Lemma 15, replacing χ_i and Γ by their values in the inequality $\frac{h_i}{4} > \left(1 + \frac{2}{\sqrt{3}}\right) \frac{\chi_i}{\sin\Gamma}$ gives

$$\frac{h_i}{4} > \left(1 + \frac{2}{\sqrt{3}}\right) \frac{\frac{\beta h_i}{\sqrt{n}}}{\frac{\sqrt{3}}{2} \sin\left(\frac{C}{2\sqrt{n}}\right)}$$

or equivalently

$$4\,\beta \left(1 + \frac{2}{\sqrt{3}}\right) < \sqrt{n}\,\frac{\sqrt{3}}{2} \sin\left(\frac{C}{2\sqrt{n}}\right). \tag{1}$$

Notice first that for large enough values of n, using the approximation $\sin x \approx x$ in the neighborhood of zero, we derive the sufficient condition

$$C > \frac{16\beta}{\sqrt{3}} \left(1 + \frac{2}{\sqrt{3}}\right) \sim 19.9\,\beta.$$

Now, since we want our result for all n, the computation is more complicated. Recall first that for any strictly concave function f, such that $f(0) = 0$, $f(x) > \frac{f(x_0)}{x_0}\,x$ for any $x \in (0, x_0)$. It follows that $\sin x > \frac{2}{\pi}x$ for any $x \in (0, \frac{\pi}{2})$. Since we assumed that $\theta_e - \frac{C}{\sqrt{n}} > 0$ and thus that $0 < \frac{C}{2\sqrt{n}} < \frac{\theta_e}{2} < \frac{\pi}{2}$, we get

$$\sin\left(\frac{C}{2\sqrt{n}}\right) > \frac{2}{\pi}\,\frac{C}{2\sqrt{n}}.$$

To guarantee inequality (1), it is thus sufficient to have

$$4\,\beta \left(1 + \frac{2}{\sqrt{3}}\right) \leqslant \sqrt{n}\,\frac{\sqrt{3}}{2}\,\frac{2}{\pi}\,\frac{C}{2\sqrt{n}}.$$

or equivalently

$$C \geqslant \frac{8}{3}\left(2 + \sqrt{3}\right)\pi\,\beta \sim 31.27\,\beta,$$

which concludes the proof. Note that we can set

$$C = 31.3\,\beta. \tag{2}$$

in the definition of Ω (in the proof of Lemma 12) since Inequality (1) is the only constraint on C. □

4.3.3 Generalizations

We prove here Theorem 11. We first show that Theorem 10 generalizes to the case where the viewpoint is chosen randomly at finite distance. We then show that considering surfaces with boundary does not change the asymptotic expected complexity of the silhouette.

Point of view at finite distance. We have thus far restricted ourselves to the case where the viewpoint is chosen uniformly at random at infinity. However, our result applies to any distribution of viewpoints such that the probability for an edge e to be on the transparent silhouette is $O(\theta_e)$, where θ_e is the exterior dihedral angle associated to e; indeed, the expected number of edges on the silhouette is then $\sum_{\text{edge } e} O(\theta_e)$ and we get the result by applying, as before, Lemma 12.[4] Such a distribution of viewpoints is obtained, in particular, when the point of view is chosen uniformly at random in a ball. This is also the case if S delimits a bounded region $\mathscr{O}$ and the viewpoint is chosen uniformly at random in $B \setminus \mathscr{O}$, for a ball B.

Surfaces with boundary. Let S be a 2-manifold with boundary $\mathcal{B}$. We consider that the boundary is always on the transparent silhouette and so the definition of the transparent silhouette of a 2-manifold S with boundary is exactly that of a 2-manifold without boundary plus the boundary $\mathcal{B}$.

The surface S is approximated by a triangulated mesh P_n that satisfies Hypotheses 1 and 2, as in the case without boundary, except that now the mesh may not be a polyhedron (some edges may have only one adjacent face rather than two).

To give an upper bound on the number of edges on the silhouette of the mesh, we consider the boundary edges and the other (non-boundary) edges separately. For the non-boundary edges, the same reasoning as before still holds. For the boundary edges, it is easy to see that the length (in 3D) of the boundary of P_n cannot be much larger than the length of $\mathcal{B}$. Indeed, the two are homeomorphic, and the hypotheses imply that the image of an edge e, of length l_e, is a curve whose extremities lie at distance at least $l_e - 2\beta \cdot \frac{l_e}{\sqrt{n}} = \Omega(\frac{1}{\sqrt{n}})$ apart. This means that the length of $\mathcal{B}$ is at least $\Omega(\frac{1}{\sqrt{n}})$ times the number of boundary edges of P_n. Hence, the number of boundary edges of P_n is at most $O(\sqrt{n})$ times the length of $\mathcal{B}$. So, if the length of $\mathcal{B}$ is bounded, the expected complexity of the silhouette of P_n is $O(\sqrt{n})$.

4.4 Conclusion

This paper gives an idea of why, and when, the usual claim that the silhouette of a triangulated mesh has size $O(\sqrt{n})$ is valid. In particular, we have given a set of conditions such that any triangulated mesh approximating a surface in a way that satisfies those conditions has a silhouette of expected size $O(\sqrt{n})$. Roughly speaking, the mesh should have no short edges, its faces should be fat, and the distance between it and the surface it approximates should never be too large. The surface itself is not necessarily everywhere differentiable and may have boundaries.

A natural question to ask is whether meshes satisfying those conditions exist. In fact, for smooth surfaces, the meshes produced by Boissonnat and Oudot [23] are one such example. The critical property of the meshes they compute is that the ratio between

[4] Note that, in Lemma 12, $\text{silh}(f_n(T_e))$ always refers to an expected length for a viewpoint chosen randomly at infinity.

the size of the largest and the smallest triangles remains bounded, although meshes are non-uniform with small triangles in areas of large curvature. However, in order to satisfy our conditions, non-smooth surfaces with curved sharp edges (such as a flying saucer with a sharp equatorial arc) would have to be approximated by small triangles over the whole surface. Such meshes would have silhouettes of expected size $O(\sqrt{n})$ but then n would be much larger than necessary; it would be reasonable to replace the large number of triangles used to mesh large flat portions of the surface with a smaller number of large triangles, which would give a silhouette of size closer to linear. This explains why the observed expected size of silhouettes, as shown in [81], is larger than $O(\sqrt{n})$. The fact that non-uniform meshes approximating such surfaces appear, in computer graphics, to have silhouettes of expected size much smaller than n is thus likely due to additional properties of the surfaces or the meshes.

Troisième partie

Visibilité en présence de polytopes

Chapitre 5

Complexe de visibilité de polytopes

5.1 Définition

Pour simplifier les explications, on étudiera uniquement des scènes avec des objets convexes d'intérieurs disjoints.[5] Nous travaillons avec des droites et plus souvent avec des segments de droite. On appelle *segment libre maximal* un segment qui n'intersecte l'intérieur d'aucun objet et qui est maximal au sens de l'inclusion pour cette propriété. Un tel segment a donc ses extrémités soit à l'infini soit sur le bord d'objets de la scène étudiée. Il représente une ligne de vision, et les objets vus de chaque côté. On dit qu'une droite est *tangente* à un objet si elle intersecte cet objet mais pas son intérieur[6]. De la même façon un segment est tangent à un objet si il l'intersecte et que sa droite support est tangente à l'objet.

Le *complexe de visibilité* a été introduit en 2D par Pocchiola et Vegter [90]. Il est défini comme une partition de l'espace des segments libres maximaux (muni d'une topologie adaptée) découpé le long d'arêtes (les segments tangents à un objet) et de sommets (les segments tangents à deux objets). Ceci correspond à une partition en composantes connexes de segments ayant leurs extrémités sur les mêmes objets et tangents aux mêmes objets. Comme on peut le voir dans la définition, la structure combinatoire (sommets, arêtes, faces) est essentielle.

Cette définition à été étendue au cas 3D par Durand, Drettakis et Puech [53]. Le *complexe de visibilité* est défini comme une partition de l'ensemble des segments libres maximaux en composantes connexes de segments qui touchent les mêmes objets. Une discussion plus poussée de cette généralisation peut être trouvée dans [65]. À nouveau, obtenir une description combinatoire de cette structure est essentiel. Cependant, à la différence du cas 2D, on ne va généralement pas se contenter de décrire à quel objet une droite est tangente, on va aussi préciser à quel arête de cet objet, afin d'être capable d'utiliser plus simplement les cellules. Je me contenterai ici d'étudier le cas où les objets sont des polytopes.

[5] Pour les objets plats, on utilise les mêmes définitions en considérant l'intérieur relatif de l'objet, sauf qu'on autorise plusieurs objets plats à s'intersecter s'ils sont dans le même plan.

[6] Pour un objet plat, une droite est tangente si elle l'intersecte et soit elle n'intersecte pas son intérieur relatif soit elle est dans le même plan.

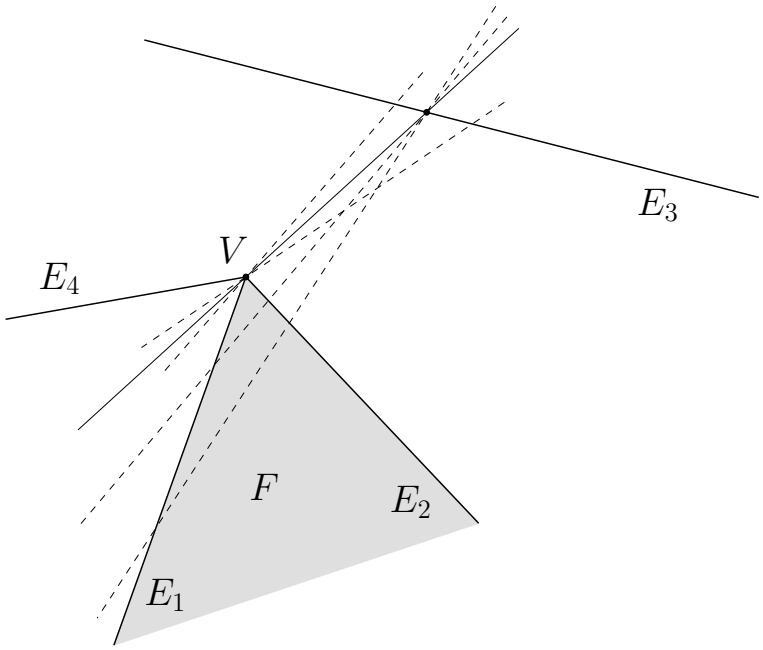


FIG. 5.1 – Exemple de EEE se terminant sur un VE.

5.2 Des supports problématiques

5.2.1 Supports naturels

On appelle *supports naturels* d'un segment l'ensemble des sommets et des arêtes ouvertes qu'il intersecte sur les polytopes auxquels il est tangent. [53] utilise aussi des faces de polytopes comme supports, mais cela n'apporte rien pour mon étude. Les cellules du complexe de visibilité sont alors définies dans [53, 65] comme des composantes connexes de droites ayant les mêmes supports. Elles sont décrites par leurs supports et leur bord, c'est-à-dire les cellules de dimension inférieure adjacentes. Les sommets du complexe sont en particulier décrits par deux sommets (VV), un sommet et deux arêtes (VEE) ou quatre arêtes ($EEEE$). Les arcs du complexe sont eux décrits par un sommet et une arête (VE) ou trois arêtes (EEE). Les cellules de dimension 2 sont décrites par un sommet (V) ou deux arêtes (EE) et celles de dimension 3 par une arête (E).

Cette description présente un gros inconvénient : elle ne représente pas une structure de complexe valide. Il existe en particulier des cellules de dimension 2 dont une partie du bord appartient simplement à une autre cellule de dimension 2 et des cellules de dimension 1 dont une extrémité n'est pas un sommet. En effet, sur la figure 5.1, on voit un arc du complexe, engendré par les générateurs E_1, E_2 et E_3 dont une extrémité a pour générateurs V et E_3. Cependant, V et E_3 engendrent eux aussi un arc du complexe et non un sommet. Notons que cette configuration est parfaitement générique. Cependant, dans certains cas il est possible que l'on ne s'intéresse qu'aux droites tangentes à trois objets ou à deux objets dans un plan bitangent (voir [65]), auquel cas cette situation ne peut se produire que dans le cas dégénéré où deux objets se touchent.

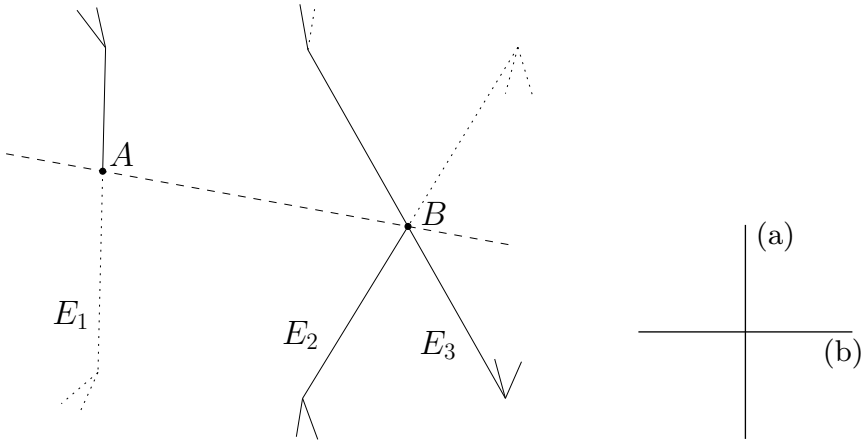


FIG. 5.2 – Trois supports qui définissent autre chose qu'un arc.

5.2.2 Supports intrinsèques

Une arête est un *support intrinsèque* d'une droite si et seulement si la droite coupe l'arête fermée et le plan défini par la droite et l'arête est tangent au polytope qui contient l'arête (si l'arête est incluse dans la droite, on la compte aussi comme support)[7]. On peut de façon équivalente demander à ce que dans tout voisinage de cette droite il existe une droite tangente au polytope en cette arête ouverte. Cette définition s'étend sans problème aux supports intrinsèques de segments. On définit à nouveau une cellule du complexe comme une composante connexe de segments libres maximaux ayant les mêmes supports intrinsèques. La définition des supports intrinsèques revient à appeler support les arêtes qui engendrent les cellules de dimension 3 adjacentes du complexe. Une arête E qui est un support naturel est automatiquement un support intrinsèque. Un sommet V correspond en général sur un polytope à deux supports intrinsèques qui dépendent de la direction du segment libre maximal. Cependant, lorsque le segment qui passe par le sommet est dans le plan d'une des faces adjacentes à ce sommet (sans intersecter l'intérieur de cette face), trois des arêtes adjacentes au sommet sont des supports intrinsèques de la droite. On peut ainsi aller jusqu'à quatre supports intrinsèques correspondant à un seul sommet quand le segment libre maximal est dans l'intersection des plans de deux faces (non adjacentes) incidentes à ce sommet, comme sur la figure 5.3.

Cette définition de support est plus fine que celle des supports naturels, et en particulier elle ne souffre pas du problème générique montré dans la figure 5.1. En effet les cellules de dimension 1 ont pour supports (E_1, E_2, E_3) et (E_2, E_3, E_4) et leur intersection a pour supports (E_1, E_2, E_3, E_4).

Malheureusement, il reste un cas dégénéré où les supports intrinsèques ne définissent pas un complexe satisfaisant. Dans la figure 5.2, on voit deux arêtes E_2 et E_3 qui s'intersectent en un point B (leurs polytopes supports s'intersectent uniquement en ce point) et une arête E_1 qui intersecte le plan défini par les arêtes E_2 et E_3 en un point A. Les droites qui passent par A et intersectent E_2 et E_3 (et donc sont dans le plan défini par ces

[7] Si les deux faces incidentes à une arête sont coplanaires, de sorte que l'on pourrait fusionner les deux faces, l'arête n'est jamais considérée comme support.

deux arêtes) forment un arc (a). Les droites qui passent par B et intersectent E_1 forment un arc (b). Les deux arcs s'intersectent en un point correspondant à la droite qui passe par A et B. On remarque alors que les deux arcs (a) et (b) et leur point d'intersection ont exactement les mêmes supports intrinsèques, à savoir E_1, E_2 et E_3. La cellule du complexe définie par ces supports correspond donc topologiquement à deux composantes de dimension 1 qui se coupent en un point qui n'est pas un sommet du complexe (voir la figure 5.2).

5.3 Tentative de réparation

5.3.1 Supports intrinsèques

Comme on vient de le voir, la notion de support intrinsèque ne permet pas de distinguer les arcs (a) et (b) de la figure 5.2.

Pour arriver à casser l'arc (a) en deux, on peut décider de subdiviser les arêtes au niveau de leurs intersections isolées avec d'autres arêtes. Ainsi la moitié de l'arc (a) aura pour supports E_1, E_2^- et E_3^- et l'autre moitié E_1, E_2^+ et E_3^+ alors que la droite (AB) aura pour supports E_1, E_2^-, E_2^+, E_3^- et E_3^+.

Pour casser l'arc (b) en deux, on peut introduire une notion de couplage entre les supports. Les arêtes E_2 et E_3 sont dites couplées pour une droite D si les trois sont coplanaires[8], ou alternativement si dans tout voisinage de la droite il existe une droite tangente aux polytopes de E_2 et E_3 en des points de E_2 et E_3 qui ne sont ni leurs extrémités ni un point d'intersection (isolé) de ces arêtes.

Avec ces deux remarques et en définissant les cellules du complexes comme des composantes connexes de droites ayant les mêmes supports et les mêmes relations de couplage entre supports, la cellule qui pose problème dans la figure 5.2 se retrouve coupée en quatre arcs et un sommet, comme il se doit.

Alors que dans la définition originale des supports intrinsèques on s'intéressait uniquement aux cellules de dimension 3 adjacentes, les relations de couplage font prendre en compte aussi des cellules de dimension 2 adjacentes.

On peut aussi noter que la notion de couplage correspond à des plans tangents à deux objets, c'est-à-dire aux événements visuels de type T++T décrits dans [65].

Cette définition modifiée a cependant perdu beaucoup de sa simplicité par rapport à celle de départ. Il n'est pas actuellement prouvé si elle définit une bonne structure de complexe cellulaire.

5.3.2 Supports naturels

Les supports naturels ne permettent pas de décrire en tant que sommet du complexe l'extrémité de l'arc EEE de la figure 5.1. La seule façon de la décrire est en tant que segment de droite ayant pour supports V et E_3 et se trouvant dans le plan de F.

[8] On pourrait aussi demander que les trois soient coplanaires et concourantes, ce qui permettrait d'isoler le point d'intersection de (a) et (b) sans découper les arêtes.

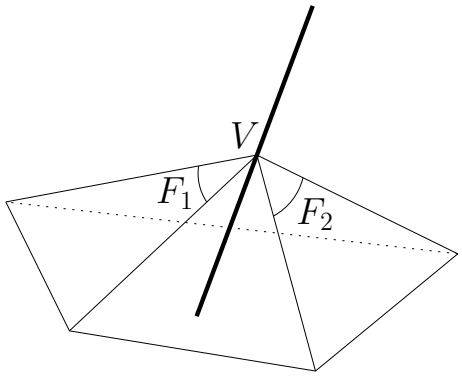


FIG. 5.3 – Exemple de V_{FF} : une droite passant par un sommet et dans l'intersection de deux des plans incidents à ce sommet sur le polytope.

On est ainsi amené à définir un nouveau type de support naturel : les V_F. Il s'agit d'une paire d'arêtes qui s'intersectent en un point unique et définissent un plan tangent à leurs deux polytopes (ou leur polytope si elles sont sur le même). Un V_F est un support naturel d'une droite si et seulement si la droite passe par le point et est dans le plan défini par les deux arêtes[9].

On obtient ainsi de nouveaux sommets du complexes engendrés par $V_F E$ et de nouveaux arcs engendrés par V_F. Il s'ajoute du coup automatiquement encore un dernier type de support, qui engendre à lui seul un sommet du complexe, le V_{FF} (voir la figure 5.3 pour un exemple), qui est une paire de V_F partageant le même point.

Cette définition modifiée résout les problèmes posés par les deux figures, mais il n'est pas non plus prouvé si elle définit correctement une structure de complexe.

5.4 Conclusion

On vient de voir que décrire les cellules du complexe de visibilité de polytopes est difficile. On propose deux définitions possibles dont on ne sait si elles sont équivalentes, ni même si elles définissent des complexes valides. Clarifier cette situation sera un point nécessaire avant de construire explicitement le complexe de visibilité, car pour stocker utilement le complexe, il faut pouvoir décrire chacune de ses cellules.

[9]Dans le cas où le point d'intersection est l'extrémité des deux arêtes, on peut ne pas compter la paire d'arêtes comme un support naturel quand la droite intersecte l'intérieur relatif de l'enveloppe convexe des deux segments, car on n'est pas sur le bord d'une cellule de droites tangentes aux deux segments ouverts.

Chapter 6

Lines and free line segments tangent to arbitrary three-dimensional convex polyhedra

Ces résultats ont été obtenus avec H. Brönnimann, O. Devillers, V. Dujmović, H. Everett, X. Goaoc, S. Lazard, H.-S. Na et S. Whitesides. Des versions préliminaires de ce travail ont été présentées à la Canadian Conference on Computational Geometry [26] et au Symposium on Computational Geometry [27]. La version complète est publiée dans Computational Geometry: Theory and Applications [28].

Motivated by visibility problems in three dimensions, we investigate the complexity and construction of the set of tangent lines in a scene of three-dimensional polyhedra. We prove that the set of lines tangent to four possibly intersecting convex polyhedra in $\mathbb{R}^3$ with a total of n edges consists of $\Theta(n^2)$ connected components in the worst case. In the generic case, each connected component is a single line, but our result still holds for arbitrarily degenerate scenes. More generally, we show that a set of k possibly intersecting convex polyhedra with a total of n edges admits, in the worst case, $\Theta(n^2k^2)$ connected components of maximal free line segments tangent to at least four polytopes. Furthermore, these bounds also hold for possibly occluded lines rather than maximal free line segments.

Finally, we present a $O(n^2k^2 \log n)$ time and $O(nk^2)$ space algorithm that, given a scene of k possibly intersecting convex polyhedra, computes all the *minimal* free line segments that are tangent to any four of the polytopes and are isolated transversals to the set of edges they intersect; in particular, we compute at least one line segment per connected component of tangent lines.

6.1 Introduction

Computing visibility relations in a 3D environment is a problem central to computer graphics and engineering tasks such as radio propagation simulation and fast prototyping. Examples of visibility computations include determining the view from a given point, and computing the umbra and penumbra cast by a light source. In many applications,

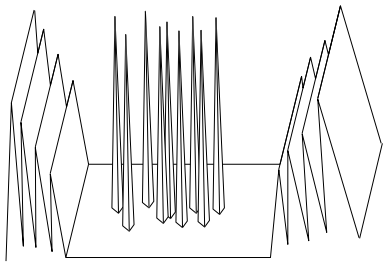


Figure 6.1: A terrain of size n with $\Omega(n^4)$ maximal free line segments tangent in four points.

visibility computations are well-known to account for a significant portion of the total computation cost. Consequently a large body of research is devoted to speeding up visibility computations through the use of data structures (see [54] for a survey).

One such structure, the visibility complex [56, 90], encodes visibility relations by partitioning the set of maximal free line segments. The size of this partition is intimately related to the number of maximal free line segments tangent to four objects in the scene; for a scene of n triangles in $\mathbb{R}^3$, the complex can have size $\Theta(n^4)$ in the worst case [56], even when the triangles form a terrain (see [39] or Figure 6.1). The complex is thus potentially enormous, which has hindered its application in practice. However, there is evidence, both theoretical and practical, that this estimation is pessimistic. The lower bound examples, which are carefully designed to exhibit the worst-case behavior, are unrealistic in practice. For realistic scenes, Durand *et al.* [57] observe a quadratic growth rate, albeit for rather small scenes. For random scenes, Devillers *et al.* [46] prove that the expected size of the visibility complex is much smaller; for uniformly distributed unit balls the expected size is linear and for polygons or polyhedra of bounded aspect ratio and similar size it is at most quadratic. Also, in 2D, while the worst-case complexity of the visibility complex is quadratic, experimental results strongly suggest that the size of the visibility complex of a scene consisting of scattered triangles is linear [36].

While these results are encouraging, most scenes are not random. In fact, most scenes have a lot of structure which we can exploit; a scene is typically represented by many triangles which form a much smaller number of convex patches. In particular, if a scene consists of k *disjoint* convex polyhedra with a total of n edges, then under a strong general position assumption, the number of maximal free line segments tangent to four of the polyhedra is at most $O(n^2k^2)$; this follows directly from the bound proved in [59] on the number of combinatorial changes of the silhouette map viewed from a point moving along a straight line, and was also later proved in [26]. We present in this paper a generalization of these results. After preliminary definitions, we give a detailed account of our results and then present related previous work.

Preliminary definitions. We consider a scene that consists of a finite number of polytopes, not necessarily disjoint, not necessarily fully dimensional, and in arbitrary position. The definitions below are standard, yet carefully phrased in a way that remains valid in those situations.

A *polytope* is the convex hull of a point set. A plane is *tangent* to a polytope if it

intersects the polytope and bounds a closed half-space that contains the polytope. A face, an edge, or a vertex of a polytope in $\mathbb{R}^3$ is the 2, 1 or 0-dimensional intersection of the polytope with a tangent plane. Note that, with this usual definition of polytopes, edges and faces are closed and they are not subdivided in any way.

A line or segment is *tangent* to a polytope (whether or not the latter is fully dimensional) if it intersects the polytope and is contained in a tangent plane. In a given plane, a line is tangent to a polygon if it intersects the polygon and bounds a closed half-plane that contains the polygon. With these definitions, given a polygon in a plane π, and a line contained in π that intersects the relative interior of this polygon, the line is tangent to the polygon when considered as a polytope in $\mathbb{R}^3$, but not tangent to the polygon in the plane π.

The set of lines in $\mathbb{R}^3$ has a natural topological structure, namely, that of Plücker space [102]. The set of lines tangent to at least four polytopes is a subspace, whose *connected components* correspond to lines that can be continuously moved one into the other while remaining tangent to at least four polytopes.[10] A line or line segment is *free* if it is tangent to each polytope that its relative interior intersects;[11] otherwise it is *occluded*. A free line segment is a *maximal free line segment* if it is not properly contained into another free line segment. The space of line segments also has a natural topological structure and the *connected components* of maximal free line segments tangent to at least four among the k polytopes are defined similarly as for lines.

A *support vertex* of a line is a polytope vertex that lies on the line. A *support edge* of a line is a polytope edge that intersects the line but has no endpoint on it (a support edge intersects the line at only one point of its relative interior). A *support* of a line is one of its support vertices or support edges. The supports of a segment are defined similarly. Notice that it follows from the definition of polytopes that any line has at most two supports in any given polytope.

A line is *isolated with respect to* a set of edges and vertices if the line cannot be moved continuously while remaining a common transversal to these edges and vertices. Furthermore, we say that a set $\mathcal{S}$ of edges and vertices *admits an isolated transversal* if these edges and vertices admit a common transversal that is isolated with respect to $\mathcal{S}$. Finally, a line is *isolated* if it is isolated with respect to a set of some, and hence all, of its supports.

Our results. In this paper we present two types of results, combinatorial bounds and algorithms.

Combinatorial bounds. We generalize the result of [26, 59] in two ways. First, we consider polytopes that may **intersect**. We show that among k polytopes of total complexity n, the number of lines tangent to any four of them is in the worst case either infinite or $\Theta(n^2k^2)$. The most surprising aspect of this result is that the bound (which is tight) is the same whether the polytopes intersect or not. This is in sharp contrast to the 2D case, where the number of tangents of two convex polygons is always 4 if disjoint, and

[10] The set of polytopes to which the line is tangent might change during the motion.

[11] When the polytopes are fully dimensional, a segment is free if it does not intersect the interior of any of them. Our definition ensures that a segment is free also when it intersects and is coplanar with a two-dimensional polytope. The endpoints of a free segment may also lie on the boundary of a polytope.

could be linear in the size of the polygons if they intersect. Second, we consider polytopes in **arbitrary position**: we drop all general position assumptions. The polytopes may intersect in any way; they may overlap or coincide. They may degenerate to polygons, segments or points. While four polytopes in general position (as defined in [26]) admit a finite number of common tangents, four polytopes in arbitrary position may admit an infinite number of common tangents which can be partitioned into connected components.

Our main results are, more precisely, the following.

Theorem 17. *Given k polytopes in $\mathbb{R}^3$ with n edges in total, there are, in the worst case, $\Theta(n^2k^2)$ connected components of maximal free line segments tangent to at least four of the polytopes. This bound also holds for connected components of possibly occluded lines tangent to at least four of the polytopes.*

These results improve the trivial bound of $O(n^4)$. Note that, when $k \neq 4$, neither of the two results stated in Theorem 17 implies the other since a line tangent to at least four among k polytopes may contain many, but does not necessarily contain any, maximal free line segments tangent to four polytopes.

When $k = 4$, Theorem 17 implies that there are $\Theta(n^2)$ connected components of lines tangent to the four polytopes, an improvement on the previously known upper bound of $O(n^3 \log n)$ which follows from the same bound on the complexity of the set of line transversals to a set of polyhedra (here four) with n edges in total [3]. Moreover, we prove a tighter bound when one of the four polytopes has few edges.

Theorem 18. *Given 3 polytopes with n edges in total and one polytope with m edges, there are, in the worst case, $\Theta(mn)$ connected components of lines tangent to the four polytopes.*

We also prove the following result which is more powerful, though more technical, than Theorem 17. Whereas Theorem 17 bounds the number of connected components of tangents, Theorem 19 bounds the number of isolated tangents with some notion of multiplicity. For example, the line in Figure 6.2 is counted $\binom{k}{2}$ times which is the number of minimal sets of vertices that admit that line as an isolated transversal. Although neither theorem implies the other, we will prove in Proposition 39 that the upper bound of Theorem 17 is easily proved using Theorem 19.

Theorem 19. *Given k polytopes in $\mathbb{R}^3$ with n edges in total, there are, in the worst case, $\Theta(n^2k^2)$ minimal sets of open edges and vertices, chosen from some of the polytopes, that admit a possibly occluded isolated transversal that is tangent to these polytopes.*

Algorithm. We now turn our attention to the computation of all free segments that are isolated transversals to their set of supports and tangent to the corresponding polytopes. Durand *et al.* [56] proposed an algorithm for this problem with worst-case time complexity $O((n^3 + p) \log n)$ where p is the output size; this algorithm, based on a double-sweep, has proven to be difficult to implement. Durand *et al.* also presented an algorithm with $\Theta(n^5)$ worst-case time complexity that incorporates interesting heuristics leading to reasonable performance in practice [57]. We present an algorithm that uses, in the worst case, $O(n^2k^2 \log n)$ time and $O(nk^2)$ space, is readily implementable, and uses only simple data

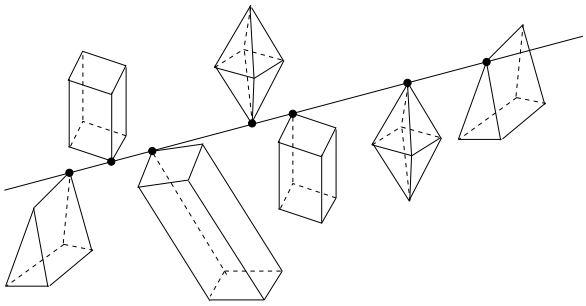


Figure 6.2: A line tangent at a vertex of each of k polytopes.

structures. The polytopes may intersect and be in arbitrary position. A preliminary version of this algorithm was described for disjoint convex polyhedra in X. Goaoc's Ph.D. thesis [65].

Theorem 20. *Given k polytopes in $\mathbb{R}^3$ with n edges in total, we can compute all the possibly occluded lines that are isolated transversals to their set of supports and tangent to the corresponding polytopes in $O(n^2k^2 \log n)$ time and $O(nk)$ space. We can also compute, in $O(n^2k^2 \log n)$ time and $O(nk^2)$ space, all the minimal free segments that are isolated transversals to their set of supports and tangent to the corresponding polytopes.*

It should be noted that our algorithm does not provide the endpoints (possibly at infinity) of the maximal free segments. Computing the endpoints of each such segment can be done by shooting rays in $O(\log^2 n)$ time per ray using $O((nk)^{2+\epsilon})$ preprocessing time and storage [6]. Such ray-shooting data structures are not, however, readily implementable. Alternatively, each ray-shooting query can be answered in $O(k \log n)$ time after $O(n \log n)$ preprocessing time and using additional $O(n)$ space by applying the Dobkin-Kirkpatrick hierarchy on each polytope [47].

To emphasize the importance of considering intersecting polytopes, observe that computer graphics scenes often contain non-convex objects. These objects, however, can be decomposed into sets of convex polyhedra. Notice that simply decomposing these objects into convex polyhedra with disjoint interiors may induce a scene of much higher complexity than a decomposition into intersecting polytopes. Moreover, the decomposition of a polyhedron into interior-disjoint polytopes may introduce new tangents which were not present in the original scene; indeed a line tangent to two polytopes along a shared face is not tangent to their union.

The importance of considering polytopes in arbitrary position comes from the fact that graphics scenes are full of degeneracies both in the sense that four polytopes may admit infinitely many tangents and that polytopes may share edges or faces. There may actually be more connected components of tangents when the objects are in degenerate position; this is, for instance, the case for line segments [29]. Also, we could not find a perturbation argument that guarantees the preservation of all (or at least a constant fraction of) the connected components of tangents and we do not believe that finding such a perturbation is a simple matter.

	Worst-case	Expected
free lines to a polyhedron	$\Theta(n^4)$ (trivial)	
free lines above a polyhedral terrain	$O(n^3 2^{c\sqrt{\log n}})$ [67, 89]	
free lines among disjoint homothetic polytopes	$\Omega(n^3)$ [41]	
free lines among unit balls	$\Omega(n^2)$ [46], $O(n^{3+\epsilon})$ [5]	$\Theta(n)$ [46]
max. free segments above a polyhedral terrain	$\Theta(n^4)$ [39]	
isolated maximal free segments among k generic disjoint convex polyhedra	$\Theta(n^2k^2)$ [59, 26]	
max. free segments among unit balls	$\Omega(n^2)$ [46], $O(n^4)$	$\Theta(n)$ [46]

Table 6.1: Published bounds on the complexity of the set of free lines or maximal free line segments among objects of total complexity n. The expected complexities are given for the uniform distribution of the balls centers.

Related results. Previous results on this topic include those that bound the complexity of sets of free lines or free line segments among different sets of objects. They are summarized in Table 6.1.

Recently, Agarwal *et al.* [5] proved that the set of free lines among n unit balls has complexity $O(n^{3+\epsilon})$. Devillers *et al.* showed a simple bound of $\Omega(n^2)$ [46] for this problem.

The complexity of the set of free line segments among n balls is trivially $O(n^4)$. Devillers and Ramos showed that the set of free line segments can have complexity $\Omega(n^3)$ (personal communication 2001, see also [46]). When the balls are unit size, the $\Omega(n^2)$ lower bound for the set of free lines holds. A lower bound of $\Omega(n^4)$ that applies to either case appears in Chapter 9.

We mention two results for polyhedral environments. Halperin and Sharir [67] and Pellegrini [89] proved that, in a polyhedral terrain with n edges, the set of free lines has near-cubic complexity. De Berg, Everett and Guibas [41] showed a $\Omega(n^3)$ lower bound on the complexity of the set of free lines (and thus free segments) among n disjoint homothetic convex polyhedra.

This paper is organized as follows. We prove the upper bounds of Theorems 17, 18, and 19 in Sections 6.2 and 6.3, and the lower bounds in Section 6.4. In section 6.5, we present our algorithm for computing free segments.

6.2 Main lemma

We prove in this section a lemma which is fundamental for the proofs of the upper bounds of Theorems 17, 18, and 19. Consider four polytopes $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$ in $\mathbb{R}^3$, with p, q, r, and $s \geqslant 1$ edges, respectively, and let e be an edge of $\mathbf{S}$.

Main Lemma. *There are $O(p+q+r)$ isolated lines intersecting e and tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$ excluding those that lie in planes that contain e and are tangent to all four polytopes.*

The proof of the Main Lemma is rather complicated because it handles polytopes which may intersect as well as all the degenerate cases. To assist the reader, we first give an overview of the proof. We then state preliminaries and definitions in Section 6.2.2. In Sections 6.2.3 and 6.2.4, we bound the number of so-called "generic tangent lines". In Sec-

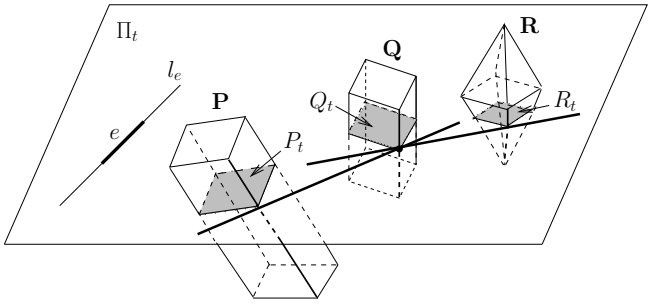


Figure 6.3: Plane Π_t contains edge e and intersects polytopes $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$ in polygons P_t, Q_t, and R_t.

tion 6.2.5, we bound the number of "non-generic tangent lines". Finally, in Section 6.2.6, we pull these results together to conclude the proof of the Main Lemma.

6.2.1 Proof overview

The proof is inspired by a method which was, to our knowledge, first used in [20] (and later in [42, 59, 26]). We present here an overview of the proof in which we do not address most of the problems arising from degeneracies. In particular, some definitions and remarks will require more elaboration in the context of the complete proof.

We sweep the space with a plane Π_t rotating about the line containing e. The sweep plane intersects the three polytopes $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$ in three, possibly degenerate or empty, convex polygons denoted P_t, Q_t, and R_t, respectively (see Figure 6.3). During the sweep, we track the *bitangents*, that is, the lines tangent to P_t and Q_t, or to Q_t and R_t, in Π_t. As the sweep plane rotates, the three polygons deform and the bitangents move accordingly. Every time two bitangents become aligned during the sweep, the common line they form is tangent to $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$.

In any given instance of the sweep plane Π_t, we consider the pairs of bitangents (one involving P_t and Q_t, and the other Q_t and R_t) that share a vertex of Q_t (see Figure 6.3). The isolated lines intersecting e and tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$ are isolated transversals with respect to a tuple of supports that consists of e and the supports of two such bitangents. We consider all *candidate* such tuples of supports as the sweep plane rotates.

Such a tuple induced by an instance of the sweep plane changes as the plane rotates only when a support of a bitangent changes. We define *critical planes* in such a way that the supports of the bitangents do not change as the sweep plane rotates between two consecutive critical planes. As the sweep plane rotates, the supports of a bitangent change if a support starts or ceases to be swept, or if, during its motion, the bitangent becomes tangent to one of the polygons along an edge of that polygon (see Figure 6.4). In the latter case, this means that the bitangent crosses a face or contains an edge of one of the polytopes. We thus define two types of critical planes: an instance of the sweep plane is critical if it contains a vertex of one of the polytopes, or if it contains a line that lies in the plane containing a face of one of the polytopes, and is tangent to another of the polytopes (see Figures 6.4 and 6.5). We will show that the number of critical planes

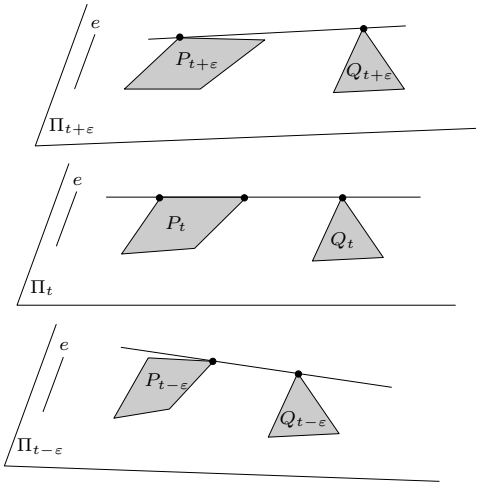


Figure 6.4: A bitangent to P_t and Q_t is tangent to P_t along an edge. The plane Π_t is F-critical.

is $O(p + q + r)$.

When the polytopes intersect there may exist a linear number of bitangents in an instance of the sweep plane (two intersecting convex polygons may admit a linear number of bitangents, as is the case for two regular n-gons where one is a rotation of the other about its center). Thus there can be a linear number of candidate tuples induced by any instance of the sweep plane, and the linear number of critical planes leads to a quadratic bound on the total number of distinct candidate tuples. In the detailed proof of the lemma, we amortize the count of candidate tuples over all the critical planes to get a linear bound on the number of distinct candidate tuples and thus on the number of isolated lines intersecting e and tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$; this bound will however not hold for those isolated lines that lie in planes that contain e and are tangent to all four polytopes. Indeed, the number of such isolated tangent lines can be quadratic, in degenerate cases; for instance, four polytopes such that a plane contains edge e and a face of linear complexity from each other polytope may admit in this plane a quadratic number of such isolated tangent lines (one through each of a quadratic number of pairs of vertices).

6.2.2 Preliminaries and definitions

We can assume without loss of generality that $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$ have *non-empty interior*. Indeed, since the set of isolated tangent lines to the four polytopes is zero-dimensional, there is always room to extend any polytope with empty interior in such a way that none of the original isolated tangent lines are lost.

We say that a line *properly* intersects a polygon if it intersects its relative interior. In the sequel, we use this definition only when the line and polygon are coplanar. Notice that a line that contains a segment is tangent to the segment as well as properly intersects it.

Let l_e be the line containing e and let Π_t denote the sweep plane parameterized by $t \in [0, \pi]$ such that Π_t contains the line l_e for all t and $\Pi_0 = \Pi_\pi$. Each plane Π_t intersects the three polytopes $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$ in three, possibly degenerate or empty, convex polygons, P_t, Q_t, and R_t, respectively (see Figure 6.3).

For any t, a *bitangent* to polygons P_t and Q_t is a line tangent to P_t and Q_t in Π_t (the line may intersect the polygon R_t in any way, possibly not at all). For any t, let a (P_t, Q_t)-*tuple* be the unordered set of all supports in $\mathbf{P}$ and $\mathbf{Q}$ of one of the bitangents to polygons P_t and Q_t. Note that a support in $\mathbf{P}$ may be identical to a support in $\mathbf{Q}$, in which case the (P_t, Q_t)-tuple does not contain duplicates. Also note that a (P_t, Q_t)-tuple consists of exactly one support in $\mathbf{P}$ and one support in $\mathbf{Q}$ (possibly identical) except when the corresponding bitangent is tangent to $\mathbf{P}$ (or $\mathbf{Q}$) along a face (either intersecting the face properly or containing one of its edges); then the (P_t, Q_t)-tuple contains two supports in $\mathbf{P}$ (or $\mathbf{Q}$) instead of one. A $\mathbf{PQ}$-*tuple* is a set of edges and vertices that is a (P_t, Q_t)-tuple for some t. We define similarly the (Q_t, R_t)-*tuples* and $\mathbf{QR}$-*tuples*.

We say that a (P_t, Q_t)-tuple is *maximal for some* t if it is not contained in any other (P_t, Q_t)-tuple, for the same t. Note that a (P_t, Q_t)-tuple is non-maximal for some t if and only if all its supports intersect Π_t in one and the same point, and P_t and Q_t are not equal to one and the same point (see Figure 6.7(b)).

For any t, let a (P_t, Q_t, R_t)-*tuple* be the union of a (P_t, Q_t)-tuple and a (Q_t, R_t)-tuple that share at least one support in $\mathbf{Q}$. A (P_t, Q_t, R_t)-tuple is maximal for some t if it is not contained in any other (P_t, Q_t, R_t)-tuple, for the same t. A $\mathbf{PQR}$-*tuple* is a set of edges and vertices that is a (P_t, Q_t, R_t)-tuple for some t. Note that a $\mathbf{PQR}$-tuple typically consists of three supports, one from each polytope, and consists, in all cases, of at most two supports in $\mathbf{P}$, at most three supports in $\mathbf{Q}$, and at most two supports in $\mathbf{R}$.

A line intersecting e and tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$ is called a *generic tangent line* if and only if it intersects $\mathbf{S}$ only on e and is tangent to P_t, Q_t, and R_t in some plane Π_t. Otherwise it is called a *non-generic tangent line*. A non-generic tangent line properly intersects a face of $\mathbf{S}$ or properly intersects P_t, Q_t, or R_t in some plane Π_t. In the latter case P_t, Q_t, or R_t is a face or an edge of $\mathbf{P}$, $\mathbf{Q}$, or $\mathbf{R}$ lying in Π_t; thus a non-generic tangent line is (in both cases) tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$ in a plane containing a face or two edges of these polytopes, a degenerate situation.

In the following three subsections, we bound the number of generic and non-generic tangent lines. It is helpful to keep in mind that, as observed earlier, two convex polygons in a plane Π_t (such as P_t and Q_t) may admit a linear number of tangents if they intersect.

6.2.3 Generic tangent lines

Lemma 21. *The set of supports in* $\mathbf{P}$, $\mathbf{Q}$, *and* $\mathbf{R}$ *of a generic tangent line is a* $\mathbf{PQR}$-*tuple.*

Proof. Any generic tangent line ℓ is tangent in Π_t to P_t, Q_t, and R_t for some value t. Thus the set of supports of ℓ in $\mathbf{P}$ and $\mathbf{Q}$ (resp. in $\mathbf{Q}$ and $\mathbf{R}$) is a (P_t, Q_t)-tuple (resp. a (Q_t, R_t)-tuple). Moreover the (P_t, Q_t)-tuple and the (Q_t, R_t)-tuple contain the same supports in $\mathbf{Q}$, and thus their union is a (P_t, Q_t, R_t)-tuple, hence a $\mathbf{PQR}$-tuple. □

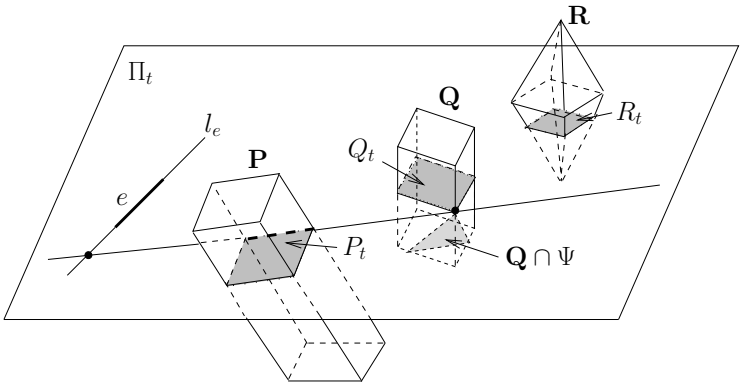


Figure 6.5: Plane Π_t is F-critical: it contains a line that lies in a plane Ψ containing a face of $\mathbf{P}$ such that the line is tangent to $\mathbf{Q} \cap \Psi$ at a point not on l_e.

We now define the *critical planes* Π_t in such a way that, as we will later prove, the set of (P_t, Q_t, R_t)-tuples is invariant for t ranging strictly between two consecutive critical values. We introduce two types of critical planes: the *V-critical* and *F-critical planes*.

A plane Π_t is *V-critical* if it contains a vertex of $\mathbf{P}$, $\mathbf{Q}$, or $\mathbf{R}$, not on l_e. (The constraint that the vertex does *not* lie on l_e ensures that the number of V-critical planes is finite even in degenerate configurations.) A plane Π_t is *F-critical* relative to an ordered pair of polytopes $(\mathbf{P}, \mathbf{Q})$ if (see Figure 6.5) it contains a line ℓ such that

(i) ℓ lies in a plane $\Psi \neq \Pi_t$ containing a face of $\mathbf{P}$, and

(ii) ℓ is tangent in Ψ to polygon $\mathbf{Q} \cap \Psi$ or $\mathbf{P} \cap \Psi$, at some point not on l_e.

For simplicity, we do not require that ℓ is tangent to $\mathbf{P}$; this leads to overestimating the number of common tangents to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$ but only by an asymptotically negligible amount. Note that not all lines in Ψ tangent to $\mathbf{Q}$ are tangent to the polygon $\mathbf{Q} \cap \Psi$ when that polygon is a face or edge of $\mathbf{Q}$ lying in Ψ. Note also that we define Π_t to be F-critical when ℓ is tangent to $\mathbf{P} \cap \Psi$ at some point not on l_e only for handling the very degenerate case where $\mathbf{Q} \cap \Psi$ is an edge of $\mathbf{Q}$ and there exists a line in Ψ that properly intersects $\mathbf{Q} \cap \Psi$ and is tangent to $\mathbf{P} \cap \Psi$ along an edge that has an endpoint on l_e (see Figure 6.6). Note finally that if $\ell \in \Pi_t$ satisfies (i) and is tangent, in Ψ, to $\mathbf{P} \cap \Psi$ at some point not on l_e then polytope $\mathbf{Q}$ plays no role and thus Π_t is F-critical relative to $(\mathbf{P}, \mathbf{Q})$ for all polytopes $\mathbf{Q}$.

F-critical planes relative to $(\mathbf{Q}, \mathbf{P})$, $(\mathbf{Q}, \mathbf{R})$, and $(\mathbf{R}, \mathbf{Q})$ are defined similarly. A plane Π_t is *F-critical* if it is F-critical relative to pairs of polytopes $(\mathbf{P}, \mathbf{Q})$, $(\mathbf{Q}, \mathbf{P})$, $(\mathbf{Q}, \mathbf{R})$, or $(\mathbf{R}, \mathbf{Q})$.

The values of t corresponding to critical planes Π_t are called *critical values*. We call *V-critical* and *F-critical events* the ordered pairs (t, o) where t is a critical value and o is a vertex or line depending on the type of critical event. In a V-critical event, o is a vertex of $\mathbf{P}$, $\mathbf{Q}$, or $\mathbf{R}$ that belongs to $\Pi_t \setminus l_e$. In an F-critical event, o is a line lying in some plane Π_t and satisfying Conditions (i-ii) above. A *critical event* is a V-critical or F-critical event.

Lemma 22. *There are at most* $\frac{2}{3}(p+q+r)$ *V-critical events and* $\frac{8}{3}(p+2\,q+r)$ *F-critical*

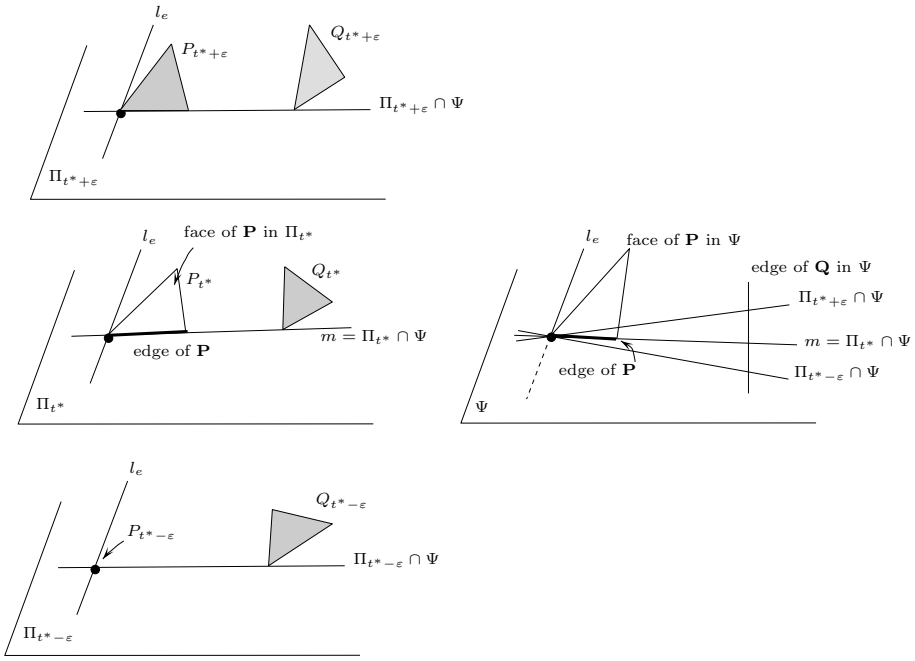


Figure 6.6: Plane Π_{t^*} contains a line m such that (i) m lies in a plane $\Psi \neq \Pi_{t^*}$ containing a face of $\mathbf{P}$, and (ii) m is tangent to polygon $\mathbf{P} \cap \Psi$ at some point not on l_e; however m is not tangent to $\mathbf{Q} \cap \Psi$. If the definition of F-critical planes was not considering such plane Π_{t^*} to be F-critical then Lemma 23 would not hold. Indeed the set u of supports of line $\Pi_{t^*-\varepsilon} \cap \Psi$ is a maximal (P_t, Q_t)-tuple for some but not all t in any open neighborhood of t^*, and, although Π_{t^*} is V-critical, there exists no V-critical event (t^*, v) such that u contains v or an edge with endpoint v.

events.

Proof. The number of V-critical events is at most the total number of vertices of $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$, and hence is less than two thirds the total number of edges of $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$. We now count the number of F-critical events relative to polytopes $(\mathbf{P}, \mathbf{Q})$. Let Ψ be a plane containing a face of $\mathbf{P}$, and suppose that for some plane Π_t, line $\ell = \Pi_t \cap \Psi$ satisfies Conditions (i-ii). Plane Ψ does not contain l_e because otherwise both l_e and ℓ lie in the two distinct planes Ψ and Π_t, so $\ell = l_e$ but then ℓ cannot satisfy Condition (ii). Furthermore ℓ and l_e intersect or are parallel since they both lie in Π_t. Thus if $\Psi \cap l_e$ is a point then ℓ contains it, and otherwise $\Psi \cap l_e = \emptyset$ and ℓ is parallel to l_e.

If $\Psi \cap l_e$ is a point, there are at most four candidates for a line ℓ in plane Ψ going through $\Psi \cap l_e$ and tangent to $\mathbf{Q} \cap \Psi$ or $\mathbf{P} \cap \Psi$ at some point not on l_e. Likewise, if $\Psi \cap l_e$ is empty, there are at most four candidates for a line ℓ in plane Ψ that is parallel to l_e and tangent to $\mathbf{Q} \cap \Psi$ or $\mathbf{P} \cap \Psi$. In either case, each candidate line is contained in a unique plane Π_t, for $t \in [0, \pi]$, since $\ell \neq l_e$ (ℓ contains a point not on l_e). Hence, a face of $\mathbf{P}$ generates at most four F-critical events relative to $(\mathbf{P}, \mathbf{Q})$. Therefore the number of critical events relative to $(\mathbf{P}, \mathbf{Q})$ is at most $\frac{8}{3}p$ since the number of faces of a polytope is at most two thirds the number of its edges. Hence the number of critical events relative

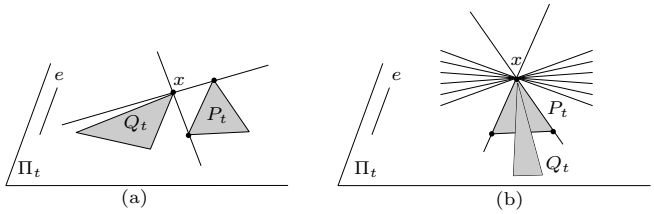


Figure 6.7: Lines through x in Π_t and tangent to P_t and Q_t.

to $(\mathbf{P},\mathbf{Q})$, $(\mathbf{Q},\mathbf{P})$, $(\mathbf{Q},\mathbf{R})$ and $(\mathbf{R},\mathbf{Q})$ is at most $\frac{8}{3}(p+2\,q+r)$. □

The following lemma states that the critical planes have the desired property. Let u_e be the set of supports of l_e in $\mathbf{P}$ and $\mathbf{Q}$ and let u denote some (P_t, Q_t)-tuple.

Lemma 23. *Let t^* be the endpoint of a maximal interval[12] throughout which $u \neq u_e$ is a maximal (P_t, Q_t)-tuple. Then t^* is a critical value. Moreover, there exists a V-critical event (t^*, v) or a F-critical event (t^*, m) such that u contains v or an edge with endpoint v, or u is contained in the set of supports of m.*

The proof of this lemma is rather long and intricate; we postpone it to Section 6.2.4. Note that, as stated, this lemma only applies under the assumptions that u is maximal and distinct from u_e. These assumptions are made in order to simplify the proof of Lemma 23; we don't suggest that the lemma is false without them.

Lemma 24. *Any edge or vertex of $\mathbf{P}$ or $\mathbf{Q}$ is in at most 2 $\mathbf{PQ}$-tuples that are maximal (P_t, Q_t)-tuples for all t in any given non-empty interval[12] of $\mathbb{R}/\pi\mathbb{Z}$.*

Proof. Let $\tilde{t}$ be an element of a non-empty interval $\mathcal{I}$ of $\mathbb{R}/\pi\mathbb{Z}$ and x be an edge or vertex of $\mathbf{P}$ or $\mathbf{Q}$. If x does not intersect $\Pi_{\tilde{t}}$ then no $(P_{\tilde{t}}, Q_{\tilde{t}})$-tuple contains x. If x intersects $\Pi_{\tilde{t}}$ in one point then there are, in general, at most two lines in $\Pi_{\tilde{t}}$ going through x and tangent to $P_{\tilde{t}}$ and $Q_{\tilde{t}}$ (see Figure 6.7(a)); in all cases there are at most 3 $(P_{\tilde{t}}, Q_{\tilde{t}})$-tuples containing x (see Figure 6.7(b)), however at most 2 of them are maximal. If x intersects $\Pi_{\tilde{t}}$ in more than one point, x is an edge lying in $\Pi_{\tilde{t}}$. Then any line in $\Pi_{\tilde{t}}$ intersecting x and tangent to $P_{\tilde{t}}$ and $Q_{\tilde{t}}$ contains an endpoint of x and thus x belongs to no $(P_{\tilde{t}}, Q_{\tilde{t}})$-tuple.

Hence at most 2 $\mathbf{PQ}$-tuples contain x and are maximal (P_t, Q_t)-tuples for $t = \tilde{t}$, and thus at most 2 $\mathbf{PQ}$-tuples contain x and are maximal (P_t, Q_t)-tuples for all t in $\mathcal{I}$. □

Lemma 25. *There are at most $O(p+q+r)$ $\mathbf{PQR}$-tuples.*

Proof. In order to count the number of distinct (P_t, Q_t, R_t)-tuples, we charge each maximal (P_t, Q_t, R_t)-tuple to a critical event. We then show that each critical event is charged at most a constant number of times. It then follows from Lemma 22 that there are $O(p+q+r)$ distinct maximal (P_t, Q_t, R_t)-tuples. A maximal (P_t, Q_t, R_t)-tuple consists of at most two supports in $\mathbf{P}$, at most three supports in $\mathbf{Q}$, and at most two supports in $\mathbf{R}$, and thus contains at most $(2^2-1)(2^3-1)(2^2-1)$ distinct subsets with at least one

[12] Such an interval could be open or closed, a single point or an interval of positive length.

support in each of **P**, **Q** and **R**. Each maximal (P_t, Q_t, R_t)-tuple thus contains at most a constant number of distinct (P_t, Q_t, R_t)-tuples, which implies the result.

Let s be a maximal (P_t, Q_t, R_t)-tuple and let $\mathcal{I}$ be any maximal connected subset of $\mathbb{R}/\pi\mathbb{Z}$ such that s is a maximal (P_t, Q_t, R_t)-tuple for all $t \in \mathcal{I}$. Let u be a maximal (P_t, Q_t)-tuple and u' a maximal (Q_t, R_t)-tuple such that the union of u and u' is s and such that u and u' share at least one support in **Q**.

First, suppose that $\mathcal{I} = \mathbb{R}/\pi\mathbb{Z}$. Then u is a maximal (P_t, Q_t)-tuple for all $t \in \mathbb{R}/\pi\mathbb{Z}$. Thus each support in u intersects Π_t for all $t \in \mathbb{R}/\pi\mathbb{Z}$ and thus intersects l_e; moreover each support in u intersects Π_t only on l_e for all $t \in \mathbb{R}/\pi\mathbb{Z}$ except possibly for one value of t. Since **P** and **Q** have non-empty interior, $P_t \cup Q_t$ is not reduced to a point for all t in some interval of positive length. For all t in such an interval, since u is maximal, the union of the supports in u intersects Π_t in at least two distinct points. These at least two distinct points lie on l_e for some values of t by the above argument. Thus, for these values of t, l_e is the only line in Π_t whose set of supports contains u. Hence u is the set of supports of l_e. The same property holds for v and thus s is also the set of supports of l_e. We can thus assume in the following that $\mathcal{I} \neq \mathbb{R}/\pi\mathbb{Z}$, and only count the maximal (P_t, Q_t, R_t)-tuples that are not the set of supports of l_e.

Interval $\mathcal{I}$ is thus a non-empty interval of $\mathbb{R}/\pi\mathbb{Z}$; it can be open or closed, a single point or an interval of positive length. Let w_0 and w_1 denote the endpoints of $\mathcal{I} \neq \mathbb{R}/\pi\mathbb{Z}$.

If s contains a vertex v, or an edge with endpoint v, such that v lies in $\Pi_{w_i} \setminus l_e$, for $i = 0$ or 1, then we charge s to the V-critical event (w_i, v). Otherwise, we charge s to an F-critical event (w_i, m) where m is a line in Π_{w_i} whose set of supports contains u or u'. Such a V-critical or F-critical event exists by Lemma 23.

We now prove that each critical event is charged by at most a constant number of distinct maximal (P_t, Q_t, R_t)-tuples. As mentioned before, that will imply the result.

Consider a V-critical event (t^*, v) that is charged by a maximal (P_t, Q_t, R_t)-tuple s. By the charging scheme, s contains a support x that is v or an edge with endpoint v, and s is a maximal (P_t, Q_t, R_t)-tuple for all t in at least one of three intervals, $\{t^*\}$ and two open intervals having t^* as endpoint; denote these intervals by $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$.

By Lemma 24, at most 2 **PQ**-tuples contain x and are maximal (P_t, Q_t)-tuples for all t in $\mathcal{I}_i$. Moreover, each of these **PQ**-tuples contains at most 2 supports in **Q**, and each of these supports belongs to at most 2 **QR**-tuples that are maximal (Q_t, R_t)-tuples for all t in $\mathcal{I}_i$. Thus at most 8 **PQR**-tuples contain x and are maximal (P_t, Q_t, R_t)-tuples for all t in $\mathcal{I}_i$, for each $i = 1, \ldots, 3$. Hence any V-critical event (t^*, v) is charged by at most 24 distinct maximal (P_t, Q_t, R_t)-tuples.

Consider now an F-critical event (t^*, m) that is charged by a maximal (P_t, Q_t, R_t)-tuple s, and define as before u and u'. By the charging scheme, the set of supports of m contains u or u' (or both); suppose without loss of generality that it contains u. The set of supports of m contains at most two supports in **P** and at most two supports in **Q**. Since u contains at least one support in **P** and at least one support in **Q**, there are at most 3^2 choices for u.

By the charging scheme, s is a maximal (P_t, Q_t, R_t)-tuple for all t in at least one of 3 intervals, $\{t^*\}$ and two open intervals having t^* as endpoint; denote by $\mathcal{I}_1, \mathcal{I}_2, \mathcal{I}_3$ these intervals. It follows from Lemma 24 that, for each support x of **Q** in u, at most 2 **QR**-tuples contain x and are maximal (Q_t, R_t)-tuples for all t in $\mathcal{I}_i$. There are at most

3^2 choices for u (as shown above), 2 for x, 3 for i and 2 for the **QR**-tuples containing x. Hence any F-critical event (t^*, m) is charged by at most $2^2 \times 3^3$ distinct maximal (P_t, Q_t, R_t)-tuples.

Therefore each critical event is charged by at most a constant number of distinct maximal (P_t, Q_t, R_t)-tuples, which concludes the proof. □

Corollary 26. *There are at most $O(p+q)$ **PQ**-tuples.*

Proof. Replace **R** by a copy of **Q** in Lemma 25. Any **PQ**-tuple is also a **PQQ**-tuple, and there are at most $O(p+q+q) = O(p+q)$ of these. □

Proposition 27. *There are $O(p+q+r)$ isolated generic tangent lines.*

Proof. A generic tangent line is transversal to e and to the edges and vertices of a **PQR**-tuple, by definition and Lemma 21. An isolated generic tangent line is thus an isolated transversal with respect to a set of edges and vertices that consists of a **PQR**-tuple and either edge e or one or both of its endpoints. The number of such sets is four times the number of **PQR**-tuples, which is in $O(p+q+r)$ by Lemma 25. The result follows since each such set consists of at most eight edges and vertices (at most two supports from each of the four polytopes) and thus admits at most eight isolated transversals [29]. □

6.2.4 Proof of Lemma 23

Recall that u_e denotes the set of supports of l_e in **P** and **Q**, and that Lemma 23 states the following.

> *Let t^* be the endpoint of a maximal interval throughout which $u \neq u_e$ is a maximal (P_t, Q_t)-tuple. Then t^* is a critical value. Moreover, there exists a V-critical event (t^*, v) or a F-critical event (t^*, m) such that u contains v or an edge with endpoint v, or u is contained in the set of supports of m.*

We can assume that *u contains no vertex v and no edge with endpoint v, such that v lies on $\Pi_{t^*} \setminus l_e$* because otherwise (t^*, v) is a V-critical event such that u contains v or an edge with endpoint v, which concludes the proof.

We prove a series of lemmas that yields Lemma 23. Indeed, we prove the existence of a line m in Π_{t^*} whose set of supports contains u (Lemma 30) such that (i) m lies in a plane $\Psi \neq \Pi_{t^*}$ containing a face of **P** (Lemma 31), and (ii) m is tangent in Ψ to polygon $\mathbf{Q} \cap \Psi$ or $\mathbf{P} \cap \Psi$, at some point not on l_e (Lemma 32). This proves that Π_{t^*} contains a line m whose set of supports contains u and such that (t^*, m) is an F-critical event, which concludes the proof.

By hypothesis, for any sufficiently small open neighborhood $\mathcal{N}$ of t^* whose endpoints are denoted by t_0 and t_1, u is not a maximal (P_t, Q_t)-tuple for some $t \in \mathcal{N}$ and u is a maximal (P_t, Q_t)-tuple for $t = t^*$ or for all $t \in (t^*, t_1)$ (or by symmetry for all $t \in (t_0, t^*)$).

We only consider in the following supports in **P** and in **Q**; polytope **R** plays no role. We start by proving two preliminary lemmas.

Lemma 28. *Each support in u intersects Π_t in exactly one point (possibly on l_e), for all t in any sufficiently small open neighborhood $\mathcal{N}$ of t^*.*

Moreover, the union of all supports in u intersects Π_t in at least two distinct points for all $t \neq t^*$ in $\mathcal{N}$. This property also holds for $t = t^*$ if u is a maximal (P_{t^*}, Q_{t^*})-tuple.

Proof. Since u is a (P_t, Q_t)-tuple for some t in every open neighborhood of t^*, each support in u intersects Π_t for some t in every open neighborhood of t^*. It thus follows from the assumption that u contains no vertex v and no edge with endpoint v, such that v lies on $\Pi_{t^*} \setminus l_e$, that each support in u intersects Π_t for all t in any sufficiently small open neighborhood $\mathcal{N}$ of t^*. It follows that each support in u either lies in l_e or intersects Π_t in exactly one point for all $t \in \mathcal{N}$. However, no edge of u lies in l_e because otherwise, if x denotes such an edge of, say, $\mathbf{P}$, then any line tangent to P_t in Π_t and intersecting x contains an endpoint of x which is a vertex of $\mathbf{P}$; thus, by definition, u does not contain x but one of its endpoints. Hence each support of u intersects Π_t in exactly one point for all $t \in \mathcal{N}$.

We now prove that the union of the supports in u intersects Π_t in at least two distinct points for any $t \in \mathcal{N}$ such that u is a maximal (P_t, Q_t)-tuple. Suppose for a contradiction that the union of the supports in u intersects Π_t in one single point v for some $t \in \mathcal{N}$ such that u is a maximal (P_t, Q_t)-tuple. Then polygons P_t and Q_t are both reduced to point v because otherwise u is not maximal (otherwise, a line in Π_t tangent to P_t and Q_t at v can be rotated about v until it becomes tangent to P_t or Q_t at some other points). Thus $v = P_t = Q_t$ is a vertex of $\mathbf{P}$ and of $\mathbf{Q}$ because the polytopes have non-empty interior. Hence $u = \{v\}$ because each support in u contains v. It follows that v lies on l_e since each support in u intersects Π_t for all $t \in \mathcal{N}$. Moreover, since P_t and Q_t are both reduced to point $v = l_e \cap \mathbf{P} = l_e \cap \mathbf{Q}$, the set u_e of supports of l_e is u, contradicting the hypotheses of Lemma 23.

Thus, if u is a maximal (P_t, Q_t)-tuple for all $t \in (t^*, t_1)$, the union of the supports in u intersects Π_t in at least two distinct points for all $t \in (t^*, t_1)$ and thus for all $t \neq t^*$ in any sufficiently small open neighborhood of t^*. Also, if u is a maximal (P_t, Q_t)-tuple for $t = t^*$, the union of the supports in u intersects Π_t in at least two distinct points for $t = t^*$ and thus for all t in any sufficiently small open neighborhood of t^*. □

Lemma 29. *If u is a maximal (P_{t^*}, Q_{t^*})-tuple then u consists of at least three supports.*

Proof. Note that it follows from Lemma 28 that u contains at least two supports. Suppose for a contradiction that u consists of only two supports. By Lemma 28, they intersect Π_t in exactly two distinct points for all t in any sufficiently small open neighborhood $\mathcal{N}$ of t^*. Thus there exists for all $t \in \mathcal{N}$ a unique line m_t in Π_t whose set of supports contains u; moreover m_t is continuous in terms of t. Since u is a maximal (P_{t^*}, Q_{t^*})-tuple, the set of supports of m_{t^*} is u. Thus, for all t in any sufficiently small $\mathcal{N}$, the set of supports of m_t is u. Thus the set of supports of m_t is invariant for $t \in \mathcal{N}$ and since m_{t^*} is tangent to P_{t^*} and Q_{t^*}, line m_t is tangent to P_t and Q_t for all $t \in \mathcal{N}$.

Hence, for all $t \in \mathcal{N}$, line m_t, whose set of supports is u, is tangent to P_t and Q_t in Π_t. Thus u is a maximal (P_t, Q_t)-tuple for all $t \in \mathcal{N}$. Moreover, m_t is the unique line in Π_t whose set of supports contains u, thus u is a maximal (P_t, Q_t)-tuple for all $t \in \mathcal{N}$, contradicting the hypotheses of the lemma. □

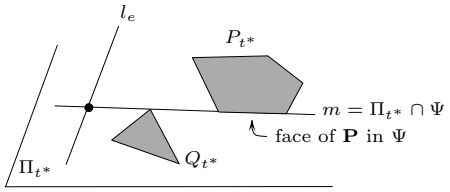


Figure 6.8: Line m is tangent to $\mathbf{P}$ along a face in plane $\Psi \neq \Pi_{t^*}$.

Lemma 30. *There exists a line m in Π_{t^*} whose set of supports contains u that is tangent to P_{t^*} and Q_{t^*} along an edge of one of them, say of P_{t^*}.*

Proof. Consider first the case where u is a maximal (P_{t^*}, Q_{t^*})-tuple. There exists in Π_{t^*} a line m tangent to P_{t^*} and Q_{t^*} whose set of supports is u. By Lemma 29, the set u of supports of m contains at least three supports, and hence at least two supports in $\mathbf{P}$ (or in $\mathbf{Q}$). Furthermore, the supports of m in one polytope intersect Π_{t^*} in distinct points (by definition of supports). Thus m intersects P_{t^*} (or Q_{t^*}) in at least two distinct points and is tangent to P_{t^*} and Q_{t^*}. The result follows since P_{t^*} (and Q_{t^*}) is convex.

Consider now the case where u is a maximal (P_t, Q_t)-tuple for all $t \in (t^*, t_1)$. Then, for all $t \in (t^*, t_1)$, there exists a line in Π_t tangent to P_t and Q_t and whose set of supports is u. Moreover, by Lemma 28, this line is unique for each $t \in (t^*, t_1)$ and varies continuously in terms of $t \in (t^*, t_1)$. When t tends to t^*, the line tends to a line m_{t^*} in Π_{t^*} which is tangent to P_{t^*} and Q_{t^*} and whose set of supports contains u. If its set of supports strictly contains u then m_{t*} is tangent to P_{t^*} and Q_{t^*} along an edge of one of them because the polygons are convex, and hence we can choose $m = m_{t^*}$ to complete the proof. Otherwise, u is a (P_{t^*}, Q_{t^*})-tuple.

We can suppose that u is a non-maximal (P_{t^*}, Q_{t^*})-tuple since we already treated the case where u is maximal. There exists in Π_{t^*} a line tangent to P_{t^*} and Q_{t^*} whose set of supports is u. Since u is non-maximal this line is tangent to P_{t^*} and Q_{t^*} at a shared vertex, and can be rotated about this vertex in Π_{t^*} until it becomes tangent to P_{t^*} and Q_{t^*} at some other points, which must occur because u is non-maximal; let m denote the resulting line. The set of supports of m contains u and m is tangent to P_{t^*} and Q_{t^*} along an edge of one of them because the polygons are convex. □

Lemma 31. *Line m lies in a plane $\Psi \neq \Pi_{t^*}$ containing a face of $\mathbf{P}$.*

Proof. By Lemma 30, m contains an edge of P_{t^*}; see Figure 6.8. This edge either intersects the relative interior of some face of $\mathbf{P}$ in which case we take Ψ to be the plane containing that face, or it is an edge of $\mathbf{P}$ in which case we take Ψ to be a plane, different from Π_{t^*}, containing one of the two faces of $\mathbf{P}$ incident to that edge. □

Let m_t be the line $\Psi \cap \Pi_t$ for all t in any sufficiently small open neighborhood $\mathcal{N}$ of t^*; line m_t is well defined since $\Psi \cap \Pi_{t^*}$ is line m by Lemmas 30 and 31.

Lemma 32. *Line m is tangent to $\mathbf{P} \cap \Psi$ or to $\mathbf{Q} \cap \Psi$, at some point not on l_e.*

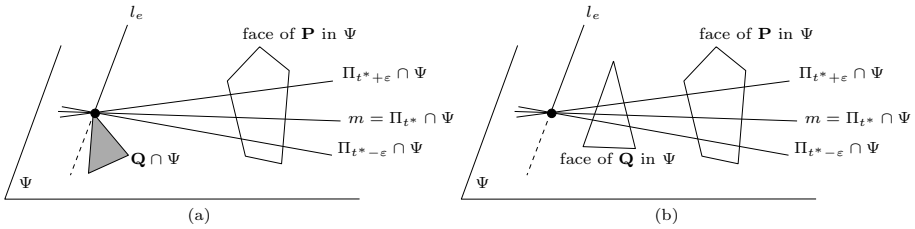


Figure 6.9: m is tangent to $\mathbf{P}$ along a face in Ψ and (a) to $\mathbf{Q} \cap \Psi$ only on l_e or (b) to $\mathbf{Q}$ along a face in Ψ.

Proof. We assume for a contradiction that line m does not satisfy the lemma, i.e., m is not tangent to $\mathbf{P} \cap \Psi$ or to $\mathbf{Q} \cap \Psi$ at any point other than on l_e. We prove that the set of supports of m is u and is a maximal (P_t, Q_t)-tuple for all t in any sufficiently small neighborhood of t^*, contradicting the hypotheses of Lemma 23 and thus proving Lemma 32.

Since m is tangent to $\mathbf{Q}$ (by Lemma 30), m is tangent to $\mathbf{Q} \cap \Psi$ only on l_e (see Figure 6.9(a)), or m properly intersects $\mathbf{Q} \cap \Psi$ which is then a face or an edge of $\mathbf{Q}$ (see Figure 6.9(b))[13]. Similarly m is tangent to $\mathbf{P} \cap \Psi$ only on l_e, or m properly intersects it; however $\mathbf{P} \cap \Psi$ is necessarily a face of $\mathbf{P}$ by Lemma 31.

The following Lemmas 33 and 34 imply that the set of supports of m_t is invariant and equal to u for all t in any sufficiently small open neighborhood $\mathcal{N}$ of t^*. Moreover, since m_t varies continuously with t and $m = m_{t^*}$ is tangent to P_{t^*} and Q_{t^*} (by Lemma 30), line m_t is tangent to P_t and Q_t for all $t \in \mathcal{N}$. Hence u is a (P_t, Q_t)-tuple for all $t \in \mathcal{N}$. We now prove that u is a maximal (P_t, Q_t)-tuple for all $t \in \mathcal{N}$.

As we have seen before, $m = m_{t^*}$ is tangent to $\mathbf{P}$ in at least two points (by Lemma 30), thus m_{t^*} intersects its supports in at least two distinct points. Moreover the set of supports of m_{t^*} is u. Thus there is a unique line in Π_{t^*} whose set of supports contains u. Hence u is a maximal (P_{t^*}, Q_{t^*})-tuple.

By Lemma 28, m_t is the unique line in Π_t whose set of supports contains u for all $t \neq t^*$ in $\mathcal{N}$. Thus u is a maximal (P_t, Q_t)-tuple for all $t \neq t^*$ in $\mathcal{N}$.

Hence u is a maximal (P_t, Q_t)-tuple for all $t \in \mathcal{N}$, contradicting the hypotheses of Lemma 23 and thus concluding the proof of Lemma 32. □

Lemma 33. *The set of supports of m_t is u for some t in any sufficiently small open neighborhood $\mathcal{N}$ of t^*.*

Proof. We first prove that the supports in u are supports of m_t for all $t \in \mathcal{N}$. A support vertex in u lies on l_e by Lemma 28 and thus lies in Π_t for all t. A support vertex in u also lies on m by Lemma 30 and thus lies in plane Ψ by Lemma 31. Hence, for all $t \in \mathcal{N}$, the support vertices in u lie on m_t, and thus are supports of m_t.

[13]Note that in these two situations, two edges of two distinct polytopes are then coplanar (in the first case an edge of $\mathbf{Q}$ and e are coplanar, and in the later case a face of $\mathbf{P}$ is coplanar with a face or an edge of $\mathbf{Q}$). Hence proving this lemma is straightforward under some general position assumption that excludes such situations.

In order to prove that the support edges in u are supports of m_t, it is sufficient (by Lemma 30) to prove that the support edges of m are supports of m_t. The support edges of m in $\mathbf{P}$ lie in plane Ψ (see Figure 6.9(b)) because Ψ contains m and a face of $\mathbf{P}$ (indeed if m intersects an edge of $\mathbf{P}$ not in Ψ then m contains one of its endpoints, and thus the edge is not a support). Thus all the support edges of m lie in Ψ and m contains none of their endpoints (by definition). Since m_t lies in Ψ for all t and $m_{t^*} = m$, line m_t intersects all the support edges of m and contains none of their endpoints for all t in any sufficiently small open neighborhood $\mathcal{N}$ of t^*. Hence the support edges of m in $\mathbf{P}$ are supports of m_t for all $t \in \mathcal{N}$.

Consider the case where $\mathbf{Q} \cap \Psi$ is a face or an edge of $\mathbf{Q}$. Similarly as for $\mathbf{P}$, the support edges of m in $\mathbf{Q}$ lie in plane Ψ, and thus are supports of m_t for all $t \in \mathcal{N}$.

Consider now the case where m is tangent to $\mathbf{Q} \cap \Psi$ only on l_e at, say, point v (see Figure 6.9(a)). Then v lies in Ψ (since $m \subset \Psi$ by Lemma 31) and also lies in Π_t for all t (since $l_e \subset \Pi_t$ for all t). Hence m_t contains v for all $t \in \mathcal{N}$. Moreover, m_t is tangent to $\mathbf{Q} \cap \Psi$ only at v for all t in any sufficiently small open neighborhood $\mathcal{N}$ of t^*. Hence the set of supports of m_t in $\mathbf{Q}$ is invariant for all $t \in \mathcal{N}$.

We have so far proved that the set of supports of m_t contains u for all $t \in \mathcal{N}$.

We now prove that the set of supports of m_t is u for some $t \in \mathcal{N}$. Consider first the case where u is a maximal (P_{t^*}, Q_{t^*})-tuple. Then, by Lemma 28, the union of the supports in u intersects Π_{t^*} in at least two distinct points, thus $m_{t^*} = m$ is the only line in Π_{t^*} whose set of supports contains u. Moreover, since u is a (P_{t^*}, Q_{t^*})-tuple, there exists a line in Π_{t^*} whose set of supports is u. Hence the set of supports of m_{t^*} is u.

Consider now the case where u is a maximal (P_t, Q_t)-tuple for all $t \in (t^*, t_1)$. By Lemma 28, for all $t \in (t^*, t_1)$, the union of the supports in u intersects Π_t in at least two distinct points, thus m_t is the only line in Π_t whose set of supports contains u. For all $t \in (t^*, t_1)$, since u is a (P_t, Q_t)-tuple there exists a line in Π_t whose set of supports is u. Hence the set of supports of m_t is u for all $t \in (t^*, t_1)$. □

Lemma 34. *The set of supports of m_t is invariant for t ranging in any sufficiently small open neighborhood $\mathcal{N}$ of t^*.*

Proof. First if $m = l_e$ then $m_t = l_e$ for all $t \in \mathcal{N}$ because Ψ contains $m = l_e$ (by Lemma 31) and Π_t contains l_e for all t (by definition). Thus the set of supports of m_t is invariant for all $t \in \mathcal{N}$. We now assume that $m \neq l_e$.

Line m is tangent to polygon P_{t^*} along an edge by Lemma 30. Thus m is tangent to $\mathbf{P}$ in at least two points. Hence, since $\mathbf{P} \cap \Psi$ is a face of $\mathbf{P}$ and m lies in Ψ, either m properly intersects $\mathbf{P} \cap \Psi$ or m is tangent to $\mathbf{P} \cap \Psi$ along one of its edges. In the later case, the edge does not lie in l_e since $m \neq l_e$, thus m is tangent to $\mathbf{P} \cap \Psi$ at some point not on l_e, contradicting our assumptions. Hence m properly intersects the face of $\mathbf{P}$ in Ψ.

It follows that, if m contains a vertex of $\mathbf{P}$, then this vertex is an endpoint of a support edge of m_t for all t in any sufficiently small open neighborhood of t^* (indeed m_t lies in Ψ and tends to m when t tends to t^*). By Lemma 33, the set of supports of m_t is u for some t in any sufficiently small open neighborhood of t^*. Hence, if m contains a vertex of $\mathbf{P}$, this vertex is an endpoint of a support edge in u. By assumption u contains no edge with endpoint on $\Pi_{t^*} \setminus l_e$, thus m contains no vertex of $\mathbf{P}$ except possibly on l_e (since m lies in Π_{t^*}). It thus follows that the set of supports of m_t in $\mathbf{P}$ is invariant for t ranging

in any sufficiently small open neighborhood of t^* (since $m_t \subset \Psi$ tends to m when t tends to t^* and all supports of m lie in Ψ).

Now consider the case where m properly intersects $\mathbf{Q} \cap \Psi$ which is a face or an edge of $\mathbf{Q}$. Similarly as for $\mathbf{P}$, m contains no vertex of $\mathbf{Q}$ except possibly on l_e and thus the set of supports of m_t in $\mathbf{Q}$ is invariant for t ranging in any sufficiently small open neighborhood of t^*.

Finally, consider the case where m is tangent to $\mathbf{Q} \cap \Psi$ only on l_e. Then, as in the proof of Lemma 33, the set of supports of m_t in $\mathbf{Q}$ is invariant for all t ranging in any sufficiently small open neighborhood of t^*, which concludes the proof. □

6.2.5 Non-generic tangent lines

We count here the number of non-generic tangent lines. Note that, as mentioned before, there are no such lines under some adequate general position assumption.

Proposition 35. *There are at most $O(p+q+r)$ isolated non-generic tangent lines except possibly for those that lie in planes that contain e and are tangent to all four polytopes.*

Proof. An isolated non-generic tangent line lies in plane Π_t for some t and contains (at least) two distinct points, each of which is a vertex of $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, or $\mathbf{S}$, or a point of tangency between the line and one of the polygons P_t, Q_t, and R_t; indeed, otherwise the line can be moved in Π_t while keeping the same supports.

We count first the isolated non-generic tangent lines that contain two distinct points of tangency with two of the polygons P_t, Q_t, and R_t in Π_t for some t. Consider such a line ℓ tangent to, say, P_t and Q_t in Π_t. Line ℓ is non-generic and thus properly intersects a face of $\mathbf{S}$ or a face or an edge of $\mathbf{R}$ lying in Π_t. If ℓ properly intersects a face of $\mathbf{S}$ or a face or an edge of $\mathbf{R}$ lying in Π_t but not entirely contained in l_e, then Π_t is one of the at most four planes tangent to $\mathbf{R}$ or $\mathbf{S}$. There are $O(p+q)$ lines tangent to P_t and Q_t in two distinct points in each of these planes and thus $O(p+q)$ such lines in total. Otherwise, Π_t intersects each of $\mathbf{R}$ and $\mathbf{S}$ in an edge contained in l_e. The supports of ℓ are thus the union of a $\mathbf{PQ}$-tuple, and of, in each of $\mathbf{R}$ and $\mathbf{S}$, the edge lying in l_e or one (or both) of its endpoint. It follows that at most a constant number of such isolated non-generic tangent lines contain a given $\mathbf{PQ}$-tuple in its set of supports. Hence the number of such lines is at most the number of $\mathbf{PQ}$-tuples, which is in $O(p+q)$ by Corollary 26. It follows that there are at most $O(p+q+r)$ isolated non-generic tangent lines that contain two distinct points of tangency with two of the polygons P_t, Q_t, and R_t in Π_t for some t. We obtain similarly that there are at most $O(p+q+r)$ isolated non-generic tangent lines that contain two distinct points of tangency with only one the polygons P_t, Q_t, and R_t.

We now count the isolated non-generic tangent lines that contain a unique vertex of $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, or $\mathbf{S}$ and a unique point of tangency with the polygons P_t, Q_t, and R_t in Π_t for some t. Each vertex v of $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, or $\mathbf{S}$ that does not lie on l_e is contained in a unique plane Π_t and there are, in that plane, at most six lines through v and tangent to P_t, Q_t, or R_t. There are thus $O(p+q+r)$ such lines in total. Consider now a line ℓ through a vertex v on l_e and tangent to P_t at $w \neq v$ in Π_t for some t. We can suppose that each of Q_t and R_t is either tangent to ℓ at w or is properly intersected by ℓ; indeed otherwise ℓ is tangent to two polygons in two distinct points. If Q_t (or R_t) is a face of $\mathbf{Q}$ (resp. $\mathbf{R}$) or

an edge not contained in l_e then Π_t is one of the at most two planes tangent to $\mathbf{Q}$ (resp. $\mathbf{R}$) and, in each of these planes, there are at most two lines through v and tangent to P_t. If Q_t (or R_t) is tangent to ℓ at w such that the support edges of ℓ in $\mathbf{P}$ and in $\mathbf{Q}$ (resp. $\mathbf{R}$) are not collinear then ℓ goes through a vertex of $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, or $\mathbf{S}$ that lies on l_e, and through a vertex of the intersection of two of these polytopes. There are at most eight vertices of $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$ on l_e and $O(p+q+r)$ vertices on the intersection of two of these polytopes. There are thus $O(p+q+r)$ such lines in total. Otherwise, Q_t (and R_t) is an edge contained in l_e or is tangent to ℓ at w such that the support edges of ℓ in $\mathbf{P}$ and in $\mathbf{Q}$ (resp. $\mathbf{R}$) are collinear; then ℓ is not isolated.

We finally bound the number of isolated non-generic tangent lines that contain no point of tangency with the polygons P_t, Q_t, and R_t in Π_t for any t (and thus contain at least two vertices of $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$). Consider such a line ℓ that lies in plane Π_t for some t. Line ℓ is tangent to $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$ and thus properly intersect P_t, Q_t, and R_t in plane Π_t which is tangent to $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$. If plane Π_t is not tangent to $\mathbf{S}$, ℓ goes through an endpoint of e (since ℓ is tangent to $\mathbf{S}$) and there are $O(p+q+r)$ such lines ℓ that go through an endpoint of e and at least another vertex of $\mathbf{P}$, $\mathbf{Q}$, or $\mathbf{R}$. If plane Π_t is tangent to $\mathbf{S}$, line ℓ lies in a plane Π_t tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$, which concludes the proof. □

Note that there can be $\Omega(n^2)$ isolated non-generic tangent lines that lie in a plane tangent to all four polytopes. Consider, for instance, four polytopes that admit a common tangent plane containing edge e, an edge e' of $\mathbf{P}$, and two faces of $\mathbf{Q}$ and $\mathbf{R}$ of linear complexity such that all the lines through a vertex of each face intersect e and e'. All these lines are isolated non-generic tangent lines.

6.2.6 Proof of the Main Lemma

Proposition 27, which handles the isolated generic tangent lines, and Proposition 35, which handles the isolated non-generic tangent lines, directly yield the Main Lemma.

6.3 Upper bounds

We prove in this section the upper bounds of Theorems 17, 18, and 19. The lower bounds are proved in Section 6.4. Consider k pairwise distinct polytopes $\mathbf{P}_1,\ldots,\mathbf{P}_k$ with $n_1,\ldots,n_k$ edges, respectively, and n edges in total.

Lemma 36. *For any edge e of $\mathbf{P}_i$, there are $O(n_j+n_l+n_m)$ sets of open edges, chosen from $\mathbf{P}_i$, $\mathbf{P}_j$, $\mathbf{P}_l$, and $\mathbf{P}_m$, that admit an isolated transversal that intersects e and is tangent to these four polytopes.*

Proof. Any isolated transversal to a set of edges is isolated with respect to the set of all its supports. It is thus sufficient to bound the number of sets of open edges, chosen from $\mathbf{P}_i$, $\mathbf{P}_j$, $\mathbf{P}_l$, and $\mathbf{P}_m$, that are intersected by an isolated line that intersects e and is tangent to these four polytopes. The Main Lemma states that there are $O(n_j+n_l+n_m)$ isolated lines intersecting e and tangent to $\mathbf{P}_i$, $\mathbf{P}_j$, $\mathbf{P}_l$, and $\mathbf{P}_m$, excluding those that lie in planes that contain e and are tangent to all four polytopes. Any of these $O(n_j+n_l+n_m)$ isolated lines

intersects at most two open edges in any polytope. Thus there are $O(n_j + n_l + n_m)$ sets of open edges (chosen from $\mathbf{P}_i$, $\mathbf{P}_j$, $\mathbf{P}_l$, and $\mathbf{P}_m$) that are intersected by one of these isolated lines. Now consider any isolated line that lies in a plane that contains e and is tangent to all four polytopes. This plane contains all the open edges that are intersected by the isolated line. Thus these edges (and any subset of them) admit no isolated transversal. □

Lemma 37. *A minimal set of open edges and vertices that admit an isolated transversal consists of (i) two vertices, (ii) one vertex and one or two edges, or (iii) two, three, or four edges.*

Proof. Consider a minimal set of open edges and vertices that admits an isolated transversal. The elements are necessarily distinct because the set is minimal. If the set contains two vertices, it contains no other element since the two vertices admit a unique transversal.

Suppose now that the set contains one vertex. None of the open edges contain the vertex because otherwise such an edge would be redundant. Thus, the vertex and any segment define either a line, and thus admit an isolated transversal, or they define a plane. If none of the other edges intersect that plane in a unique point, the vertex and all open edges admit zero or infinitely many common transversals, a contradiction. Thus there exists an edge that intersects the plane in a unique point. Hence, the vertex and two open edges admit a unique transversal, and the minimal set contains no other element.

Suppose finally that the set only contains open edges. The characterization of the transversals to a set of line segments [29] shows that either two, three or four of these line segments admit at most two transversals, or that the set of common transversals to all the open line segments can be parameterized by an open set of parameters in $\mathbb{R}^2$, $\mathbb{R}$ or $\mathbb{R}/\pi\mathbb{Z}$. In the latter case, the edges admit no isolated transversal, a contradiction. Hence, the minimal set of edges consists of two, three or four edges. (Note that two or three edges may admit an isolated transversal if that transversal contains one or two of the edges.) □

We can now prove the upper bound of Theorem 19.

Proposition 38. *There are $O(n^2k^2)$ minimal sets of open edges and vertices, chosen from some polytopes, that admit an isolated transversal that is tangent to these polytopes.*

Proof. We bound the number of minimal sets depending of their type according to Lemma 37. First, there are $O(n^2)$ pairs of vertices, pairs of edges, and sets of one vertex and one edge. Hence, at most $O(n^2)$ such pairs admit an isolated transversal.

Consider a minimal set of one vertex and two open edges, chosen from some polytopes, that admit an isolated transversal that is tangent to these polytopes. The open edges do not contain the vertex because otherwise they admit no isolated transversal. Thus the vertex and each edge define a plane. For each of the $O(n^2)$ planes defined by a vertex and an open edge not containing it, there are $O(k)$ lines in that plane that are tangent to one of the polytopes at some point other than the vertex. Hence there are $O(n^2k)$ sets of one vertex and two edges, chosen from some polytopes, that admit an isolated transversal that is tangent to these polytopes.

It is straightforward to show that three open edges admit an isolated transversal only if the line containing one of the edges intersects the two other edges. Since any line

intersects at most two open edges in any of the k polytopes, there are $O(nk^2)$ sets of three open edges that admit an isolated transversal.

Consider now the case of four edges, chosen from at most three polytopes, that admit an isolated transversal that is tangent to these polytopes. The two edges chosen from the same polytope belong to the same face, and the isolated transversal lies in the plane containing that face. Each of the two other open edges intersects that plane in one point, because otherwise the four open edges admit zero or infinitely many transversals. For each of the $O(n)$ planes containing a face of one of the polytopes, and each of the $O(n)$ edges intersecting that plane in exactly one point, there are at most $2k$ lines in that plane that contain this point and are tangent to one of the k polytopes at some other point. Hence there are $O(n^2k)$ sets of four open edges, chosen from at most three polytopes, that admit an isolated transversal that is tangent to these polytopes.

We finally bound the number of sets of four edges, no two chosen from the same polytope. By Lemma 36 and by summing over all n edges e of the polytopes, the number T of sets of four open edges, chosen from four polytopes, that admits an isolated transversal that is tangent to these four polytopes satisfies

$$T \leqslant n \sum_{j<l<m} C\,(n_j + n_l + n_m),$$

where C is some constant. Since each n_i, $1 \leqslant i \leqslant k$, appears $\binom{k-1}{2}$ times in the sum, it follows that

$$T \leqslant C\,n \sum_{1\leqslant i\leqslant k} n_i \binom{k-1}{2} = C\,n^2 \binom{k-1}{2}$$

so T is in $O(n^2k^2)$ as claimed. □

The above result implies the following upper bounds and in particular those of Theorem 17.

Proposition 39. *There are $O(n^2k^2)$ connected components of maximal free line segments tangent to at least four of the polytopes. This bound also holds for connected components of possibly occluded lines tangent to at least four of the polytopes. Furthermore, the same bound holds for isolated such segments or lines.*

Proof. We prove the proposition for possibly occluded lines tangent to at least four of the polytopes; the proof is similar for maximal free line segments. By Proposition 38, there are $O(n^2k^2)$ minimal sets of open edges and vertices, chosen from some polytopes, that admit an isolated transversal that is tangent to these polytopes. The bound on the number of connected components thus follows from the fact that any connected component of lines tangent to four polytopes contains an isolated line. Indeed, any non-isolated line can be moved while keeping the same set of supports until (at the limit) the line intersects a new edge or vertex. During the motion, the line remains tangent to all four polytopes since it keeps the same supports (except at the limit); if the line has more than one degree of freedom, this can be repeated until the line becomes isolated. □

We now prove the upper bound of Theorem 18. We start by two preliminary lemmas.

Lemma 40. *Four possibly intersecting convex polygons in $\mathbb{R}^2$ admit at most a constant number of connected components of line transversals.*

Proof. Consider the usual geometric transform where a line in $\mathbb{R}^2$ with equation $y = ax+b$ is mapped to the point $(-a, b)$ in the dual space (see e.g. [99, §8.2.1]). The transversals to a convex polygon are mapped to a region bounded from above by a convex x-monotone curve and from below by a concave x-monotone curve; such a region is called stabbing region, and the curves are referred to as the upper and lower boundaries of the stabbing region. The transversals to four polygons are mapped to the intersection of four stabbing regions. There exists no transversal of a given slope if and only if the lower boundary of a stabbing region lies above the upper boundary of another stabbing region at that slope. Two such boundaries intersect in at most two points, and thus the transversals to four polygons form at most a constant number of connected components of transversals. □

As in Section 6.2, let $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$ be four polytopes in $\mathbb{R}^3$, with p, q, r, and $s \geqslant 1$ edges, respectively, and let e be a closed edge of $\mathbf{S}$.

Lemma 41. *There are $O(p + q + r)$ connected components of lines intersecting e and tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$.*

Proof. As in the proof of Proposition 39, any connected component of lines intersecting e and tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$ contains an isolated line. The Main Lemma thus yields that there are $O(p+q+r)$ connected components of lines intersecting e and tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$ except for the components that only contain isolated lines that lie in planes that contain e and are tangent to all four polytopes.

We show that there are at most a constant number of connected components of lines intersecting e and tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$ that lie in planes that contain e and are tangent to all four polytopes. There may be infinitely many such planes that intersect $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$ only on l_e but all the lines tangent to the four polytopes in all these planes belong to the same connected component. Besides these planes there are at most two planes containing e and tangent to all four polytopes. In any such plane, the lines tangent to the four polytopes are the transversals to the four polygons that are the faces, edges, or vertices of $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$, and $\mathbf{S}$ lying in the plane. Lemma 40 thus yields the result. □

We can now prove the upper bound of Theorem 18.

Proposition 42. *Given 3 polytopes with n edges in total and one polytope with m edges, there are $O(mn)$ connected components of lines tangent to the four polytopes.*

Proof. Let $\mathbf{S}$ denote the polytope with m edges. First, if $\mathbf{S}$ consists of a single point, it is straightforward to show that there are $O(n)$ connected components of lines tangent to the four polytopes. Otherwise, by summing over all the edges of $\mathbf{S}$, Proposition 41 yields that the number of connected components of lines tangent to the four polytopes is $O(mn)$. □

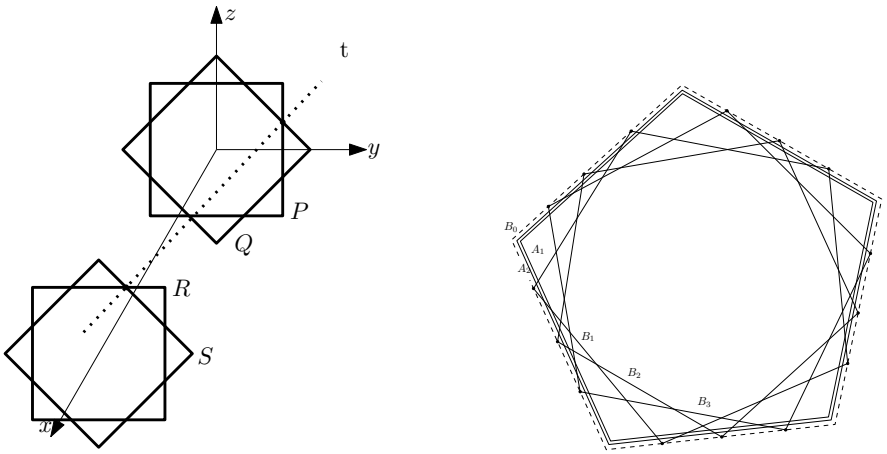


Figure 6.10: Lower bound examples for Lemmas 43 and 44.

6.4 Lower bounds

We provide in this section the lower-bound examples needed for Theorems 17, 18, and 19. The following proposition proves the lower bound of Theorem 18.

Lemma 43. *There exist four disjoint polytopes of complexity n such that the number of common tangent lines is finite and $\Omega(n^2)$. There also exist two polytopes of complexity n and two polytopes of complexity m such that the number of common tangent lines is finite and $\Omega(mn)$.*

Proof. We consider four planar regular polygons P, Q, R, and S, each with n vertices, embedded in $\mathbb{R}^3$. P is centered at the origin and parallel to the yz-plane, Q is obtained from P by a rotation of angle $\frac{\pi}{n}$ about the x-axis, and R and S are obtained from P and Q, respectively, by a translation of length 1 in the positive x-direction (see Figure 6.10). We transform the polygons P and Q into the polytopes $\mathbf{P}$ and $\mathbf{Q}$ by adding a vertex at coordinates $(\varepsilon, 0, 0)$. Similarly, we transform the polygons R and S into the polytopes $\mathbf{R}$ and $\mathbf{S}$ by adding a vertex at coordinates $(1+\varepsilon, 0, 0)$.

For ε sufficiently small, the lines tangent to $\mathbf{P}$, $\mathbf{Q}$, $\mathbf{R}$ and $\mathbf{S}$ are the lines through a vertex of $P \cap Q$ and a vertex of $R \cap S$. Since $P \cap Q$ and $R \cap S$ have $2n$ vertices each, there are $4n^2$ tangent lines. Now, moving $\mathbf{P}$ and $\mathbf{S}$ by 2ε in the x direction ensures the disjointness of the polytopes while preserving the existence of the tangents if ε is small enough.

Replacing R and S in the above construction by regular polygons each with m vertices yields the $\Omega(mn)$ lower bound in the case of two polytopes of complexity n and two polytopes of complexity m. □

We now prove the lower bounds of Theorems 17 and 19. The following proposition directly yields these bounds since the number of isolated tangents to any four of the polytopes is less or equal to the number of sets of open edges and vertices in at most four polytopes that admit an isolated transversal that is tangent to these polytopes.

Lemma 44. *There exist k disjoint polytopes of total complexity n such that the number of maximal free line segments tangent to four of them is finite and $\Omega(n^2k^2)$. Moreover these segments lie in pairwise distinct lines.*

Proof. The lower bound example is similar to the one with four polyhedra. For simplicity suppose that n and k are such that $\frac{n}{k}$ and $\frac{k}{4}$ are integers. We first take a $\frac{n}{k}$-regular polygon A_1 in the plane $x = 0$. Next we consider a copy, B_0, of A_1 scaled by a factor of $(1 + \varepsilon)$, and on each edge of B_0 we place $\frac{k}{4}$ points. Polygon B_i, $1 \leqslant i \leqslant \frac{k}{4}$, is constructed by taking the i^{th} point on each edge of B_0. If ε is small enough, the intersection points of A_1 and B_i are outside the other polygons B_j for $1 \leqslant j \leqslant \frac{k}{4}$ and $i \neq j$. Now the A_i, for $2 \leqslant i \leqslant \frac{k}{4}$, are constructed as copies of A_1 scaled by a factor $1 + \frac{i}{k}\varepsilon$ (see Figure 6.10). For the moment, all polygons lie in plane $x = 0$. We now construct 4 families of $\frac{k}{4}$ polygons each:

- P_i is a copy of A_i translated by $i\epsilon$ in the negative x direction
- Q_i is a copy of B_i translated by $i\epsilon$ in the positive x direction
- R_i is a copy of B_i translated by $1 - i\epsilon$ in the positive x direction
- S_i is a copy of A_i translated by $1 + i\epsilon$ in the positive x direction

Any choice of four polygons, one in each family P_i, Q_j, R_l and S_m, reproduces the quadratic example of Lemma 43 with polygons of size $\frac{n}{k}$ and thus with total number of tangents larger than $\left(\frac{k}{4}\right)^4 4 \left(\frac{n}{k}\right)^2 = \frac{n^2k^2}{4}$. Furthermore the lines tangent to P_i, Q_j, R_l and S_m are only occluded by $P_{i'}$ and $S_{m'}$ for $i' > i$ and $m' > m$, that is, beyond the portion of the tangents containing the contact points. The k polygons can be transformed into k convex polyhedra as in Lemma 43. □

6.5 Algorithm

Using the sweep-plane algorithm outlined in Section 2.1, we can compute in $O(n^2k^2 \log n)$ time all minimal sets of open edges and vertices, chosen from some of the polytopes, that admit a possibly occluded isolated transversal that is tangent to these polytopes. Now, for some of these lines, the segment joining the contact points with the polytopes is free. We can use standard, but complicated, ray-shooting data structures in order to determine which of these $O(n^2k^2)$ segments are free; this can be done in $O(\log^2 n)$-time per query using $O((nk)^{2+\epsilon})$ preprocessing time and storage [6].

We present in this section a solution that uses $O(n^2k^2 \log n)$ time and $O(nk^2)$ space. We adapt the algorithm outlined in Section 2.1 to directly compute the minimal sets of edges and vertices admitting an isolated line transversal that contains a free segment tangent to their respective polytopes. Our algorithm has better time and space complexities than the previously mentioned approach, and is readily implementable. Moreover, the space complexity drops to $O(nk)$ if no occlusion is taken into account. Precisely, we prove the following theorem which is more powerful, though more technical, than Theorem 20 and directly yields it.

Theorem 45. *Given k polytopes in $\mathbb{R}^3$ with n edges in total, we can compute all the minimal sets of open edges and vertices, chosen from some of the polytopes, that admit an isolated, possibly occluded, line transversal tangent to these polytopes in $O(n^2k^2 \log n)$ time and $O(nk)$ space. We can also compute, in $O(n^2k^2 \log n)$ time and $O(nk^2)$ space, all the minimal sets of open edges and vertices that admit an isolated line transversal containing a maximal free segment that is tangent to these polytopes. Furthermore, the algorithm reports which of the transversals contains such a free line segment.*

For ease of presentation, we describe a simplified version of the algorithm in which we assume that the polytopes are in generic position; see Section 6.5.2 for details. Using the same techniques as in Section 2, it is straightforward though tedious to generalize the algorithm for arbitrary situations. We also only detail the algorithm for the case of minimal sets of four edges, no two chosen from the same polytope; the other sets of at most four edges and vertices can be computed similarly.

6.5.1 Algorithm overview and data structures

The input to our algorithm is a set of possibly intersecting polytopes structured in a standard way so that classic incidence queries can be performed in constant time (see, for instance, [24, §9.1]).

We consider each polytope edge, e, in turn and sweep a plane around it between its two incident faces. During the sweep we create and maintain the following objects.

Combinatorial polygons. The sweep plane intersects each polytope in a (possibly empty) convex polygon whose vertices correspond to polytope edges. For each of these polygons, we maintain the set of vertices, each represented by its corresponding polytope edge, in a data structure that admits logarithmic-time vertex insertion, deletion and look-up operations, as well as ray-shooting queries. This can be done with a balanced binary search tree (see [88, §7.9.1]).

Combinatorial bitangents. The algorithm keeps track of the lines contained in the sweep plane and tangent to two polygons. The polytopes properly intersected by such a bitangent between its two supports are its *blockers*. A bitangent is represented by (pointers to) its two supports and a set of its blockers, ordered by polytope index, stored in a balanced binary search tree.

Polytope edges. We associate with each polytope edge a list of pointers to the combinatorial bitangents it supports in the current sweep plane.

Critical events. The sweep stops at critical events at which time combinatorial polygons and bitangents are updated. In addition to the V- and F-critical events defined in Section 6.2.3, we introduce the following two new types of events at which the set of blockers of some combinatorial bitangents may change. A *T-critical* event occurs whenever three bitangents, supported by a **PQR**-tuple, become aligned (see Figure 6.11b). An *I-critical* event occurs when the sweep plane contains a point of intersection between an edge and a face of two (distinct) polytopes (see Figure 6.12).

Each event is represented by a data structure providing pointers to the primitives that define it: a vertex for a V-event, a bitangent and a face for a F-event, three bitangents for

a T-event, and a face and an edge for a I-event. In addition, for a T-event, we store a bit of information specifying which of the line transversals to l_e and the three support edges defines the T-event. Note that the critical value of each critical event can be computed in constant time from the information associated with the event; it thus does not need to be explicitly stored.

Finally, critical events are sorted in the order in which they appear during the sweep and stored in an *event queue* supporting insertion and deletion in logarithmic time.

6.5.2 Generic position assumption

Our generic position assumption is that *the ordered set of events does not change under any arbitrarily small perturbation of the input polytopes.* This assumption corresponds to (i) the events are generic, and (ii) no two events occur in the same sweep plane, except for F- and I-critical events induced by the same pair of edge and face. The genericity of the events is ensured by (but not characterized by) the following geometric conditions:

V-critical events: no vertex lies on a line containing another edge,

F-critical events: no two edges in two distinct polytopes are coplanar,

I-critical events: if an edge intersects a face of another polytope, it does so properly and not on a line containing another edge,

T-critical events: any four lines containing polytope edges admit zero or two transversals.

6.5.3 Initialization

For each new sweep, we initialize the event queue and construct the combinatorial polygons and combinatorial bitangents as follows.

Combinatorial polygons. Computing the combinatorial polygons in the initial sweep plane can easily be done in $O(n)$ time.

Combinatorial bitangents. The bitangent lines to two polygons P and Q in the initial sweep plane through a given vertex of P can be computed by a binary search on Q in $O(\log n)$ time. The blockers of a given bitangent can be found using one ray-shooting query per combinatorial polygon, for a total time of $O(k \log n)$. Altogether, the $O(nk)$ combinatorial bitangents can thus be computed in $O(nk^2 \log n)$ time.

Event queue. There are $O(n)$ V-critical events and $O(nk)$ I-critical events, since an edge intersects a polytope in at most two faces. The $O(nk)$ edge-face intersection points are computed and stored once before the beginning of the first sweep; this computation can be done by using brute force in $O(n^2)$ time, and with $O(nk)$ space, since it is done once for all the sweeps. For each sweep, all the V- and I-critical events can then be inserted in $O(nk \log n)$ time. For each of the $O(nk)$ combinatorial bitangents, we also insert F- and T-critical events in $O(k \log n)$ time as explained in Section 6.5.4 (Lemma 46). In total, initializing the event queue takes $O(nk^2 \log n)$ time per sweep.

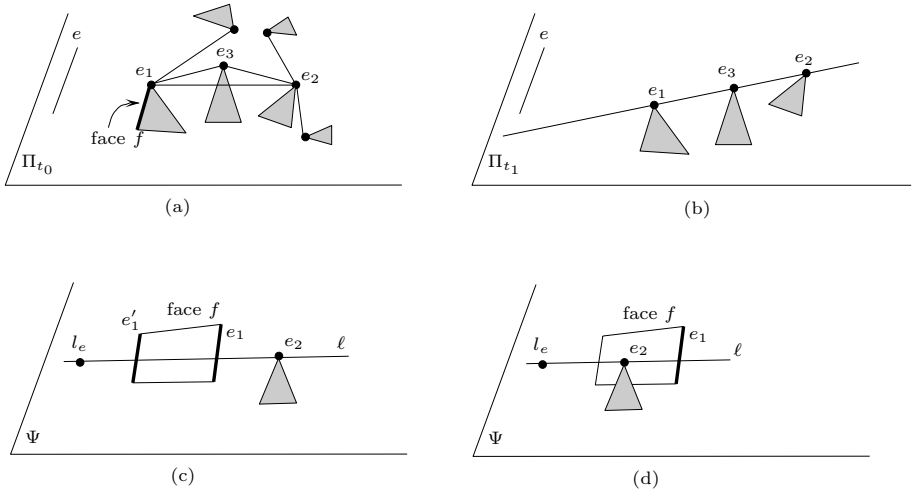


Figure 6.11: (a) The sweep plane in which the combinatorial bitangent with support edges e_1 and e_2 is created. (b) The sweep plane at a T-critical event induced by the three bitangents with support edges in e_1, e_2, and e_3. (c-d) A line ℓ that defines an F-critical event. (d) The F-event defined by ℓ occurs simultaneously with an I-critical event.

Thus, initializing all the combinatorial polygons, bitangents, and the event queue can be done in $O(nk^2 \log n)$ time per sweep plus $O(n^2)$ time overhead for a total of $O(n^2k^2 \log n)$ time as announced in Theorem 45.

6.5.4 Updating the event queue

Every time a new combinatorial bitangent is created, we compute and insert into the queue new F- and T-events as described below. Let e_1 and e_2 denote the two support edges of a new combinatorial bitangent. Let Π_{t_0} denote the critical plane at which the new combinatorial bitangent is created.

New T-critical events. See Figure 6.11a-b. Consider all the bitangents having e_1 as support edge and compute the set of support edges (distinct from e_1 and e_2) of all these bitangents. Compute the intersection of this set with the similar set for e_2; this can be done in $O(k \log k)$ time by ordering the edges by their indices. For each edge e_3 in that set, insert a T-event for each line transversal to l_e, e_1, e_2, and e_3 if the transversal is tangent to the three polytopes containing e_1, e_2, and e_3; this test can be done in constant time. Each of the at most k insertions into the event queue takes $O(\log n)$. Thus computing and inserting the new T-critical events takes $O(k \log n)$ time per new bitangent.

New F-critical events. Consider in turn each of the four faces incident to one of the two support edges. Let e_1 and f denote the considered edge and face. We compute a candidate F-event, in constant time, as follows. Compute the line ℓ (if any) that lies in the plane Ψ containing f and goes through l_e and e_2 (see Figure 6.11c). If ℓ is tangent to the polytope containing e_2, ℓ defines an F-event. We reject this F-event if ℓ does not intersect e_1 (in such a case, the edge e_1 does not intersect the sweep plane at the F-event

and thus the combinatorial bitangent to e_1 and e_2 would have been deleted at some V-event before the F-event). We also discard this F-event if it occurs at the critical value t_0 where the (considered) bitangent is created (that is Π_{t_0} contains ℓ); we discard such F-events because when a bitangent is created at an F-event, we do not re-insert the same F-event into the queue. We thus retain at most four F-events, at most one for each of the four faces incident to one of the two support edges. If no F-event is retained, the bitangent will be deleted at a V-critical event and no new F-critical event is created. If more than one F-event is retained, we need only keep the first one, since, as we shall see in Section 6.5.5.2, the combinatorial bitangent will be deleted at the first of these events.

Again, let f denote the face incident to edge e_1 that induces that F-critical event. If the other support edge, e_2, intersects face f (see Figure 6.11d), then this event will be treated as an I-critical event and again we create no new F-event. Otherwise, we insert the F-event into the queue in $O(\log n)$ time. We thus get the following lemma.

Lemma 46. *Each time a combinatorial bitangent is created, the event queue can be updated in $O(k \log n)$ time.*

6.5.5 Processing events

6.5.5.1 V-critical events

Let v denote the vertex that induces a V-critical event. As the sweep plane reaches v, all edges incident to v start or cease to be swept; we call the former *starting* edges and the latter *terminating* edges. Let $\mathbf{Q}$ denote the polytope to which v belongs and let Π_{t_0} be the sweep plane containing v. When processing a V-event, we perform the following operations.

Create and delete combinatorial bitangents. Suppose first that the critical plane through v properly intersects $\mathbf{Q}$. Consider in turn each combinatorial bitangent supported by a terminating edge, e_t, incident to v and let h denote the other support edge of this bitangent. We check all starting edges incident to v to find the edge e_s such that the line in $\Pi_{t_0+\epsilon}$ through e_s and h is tangent to $\mathbf{Q}$ for $\epsilon > 0$ arbitrarily small. We create a new combinatorial bitangent and delete the old one; in fact, we simply replace e_t by e_s in the combinatorial bitangent, create a pointer from edge e_s to the bitangent, and update the event queue. After handling the last bitangent supported by edge e_t, delete all the pointers from e_t to the bitangents.

The critical plane through v contains $O(k)$ bitangents through v, thus, by continuity, at most $O(k)$ combinatorial bitangents are deleted and created. Each deletion and creation takes linear time in the degree of v plus $O(k \log n)$ time for updating the event queue (Lemma 46). Hence, since the sum of the degrees of the vertices is $O(n)$, this step takes $O(nk^2 \log n)$ time in total for all non-extremal V-events.

Suppose now that the critical plane through v is tangent to Q and that all edges incident to v are starting. For each edge not incident to v, we can decide in constant time whether it supports a bitangent through v in the critical plane through v. If so, we check, for each edge incident to v, if the line in plane $\Pi_{t_0+\epsilon}$ that goes through these two edges is tangent to $\mathbf{Q}$ for $\epsilon > 0$ arbitrarily small. If so, we create a new combinatorial bitangent.

By continuity, $O(k)$ bitangents are created in total time $O(n + kd)$ where d is the degree of v. For each of these newly created bitangents, we compute its set of blockers in (brute force) $O(n)$ time and update the event queue in $O(k \log n)$ time (Lemma 46). This takes $O(nk \log n)$ time per event, hence $O(nk^2 \log n)$ time per sweep since there are at most two sweep planes tangent to any polytope.

Finally, if all edges incident to v are terminating, we delete all the $O(k)$ bitangents supported by these edges; for each bitangent, deleting its blockers and the pointer from the edge not incident to v can be done in $O(k)$ time. Hence, this takes $O(k^2)$ time per critical event and $O(k^3)$ time per sweep.

Update the combinatorial polygon associated with $\mathbf{Q}$. This takes $O(\log n)$ time per polytope edge incident to v, thus $O(n \log n)$ time in total for all V-events.

Hence, processing all V-events takes $O(nk^2 \log n)$ time per sweep.

6.5.5.2 F-critical events

We process an F-critical event as follows. Let b and f denote the bitangent and face associated with the event. Let e_1 and e_2 denote the two support edges of b such that e_1 is the edge that belong to f (see Figure 6.11c-d). By construction of F-events (see Section 6.5.4), e_2 does not intersect face f (see Figure 6.11c), thus the bitangent b is deleted and a new combinatorial bitangent is created.

Bitangent b is removed from the lists of bitangents supported by e_1 and e_2 in $O(k)$ time. The support edges of the new bitangent are e_2 and the edge $e_1' \neq e_1$ of f that is intersected by the line in the plane Ψ (containing f) through l_e and e_2 (see Figure 6.11c). This edge e_1' is also one of the two edges adjacent to e_1 in its combinatorial polygon. Edge e_1' can thus be computed in $O(\log n)$ time. As usual, the new bitangent is added to the lists of bitangents supported by e_1' and e_2. We then compute all the blockers of this new bitangent by performing one ray-shooting query per combinatorial polygon, for a total time of $O(k \log n)$. We finally update the event queue in $O(k \log n)$ time (Lemma 46).

There are $O(k)$ F-events associated to each polytope face, thus $O(nk)$ F-events per sweep. Hence, the total time complexity for processing all F-events is $O(nk^2 \log n)$ per sweep.

6.5.5.3 I-critical events

An I-event is associated with a face f of some polytope $\mathbf{P}$ and an edge e_1 of some other polytope $\mathbf{Q}$. Let p denote the point of intersection between f and e_1. The sweep plane, Π_{t_0}, that contains p intersects the two polytopes $\mathbf{P}$ and $\mathbf{Q}$ in two polygons P_{t_0} and Q_{t_0}. See Figure 6.12. Point p lies on an edge of P_{t_0}; the two endpoints of this edge are the intersection of two edges of $\mathbf{P}$, say e_0 and e_2. These two polytope edges can be computed in $O(\log n)$ time using the combinatorial polygon associated with $\mathbf{P}$.

Create or delete combinatorial bitangents. If the two polygons P_{t_0} and Q_{t_0} are tangent at p (see Figure 6.12a), the two combinatorial bitangents whose pairs of support edges are (e_0, e_1) and (e_1, e_2) are either created or deleted at the I-event. If these bitangents appear in the list of bitangents having edge e_1 as support, we remove them from the list

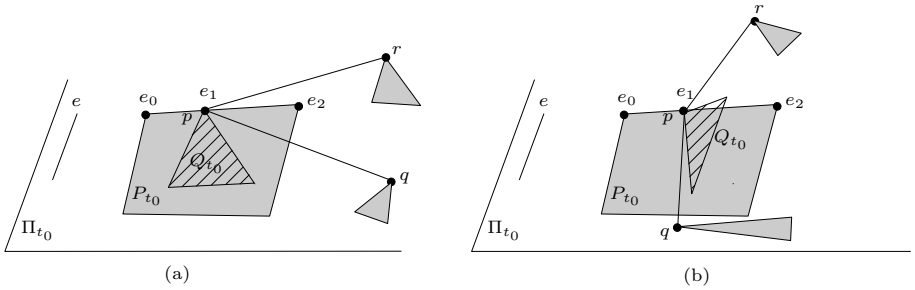


Figure 6.12: I-critical event.

and delete them; this can be done brute force in $O(k)$ time. Otherwise we create these two combinatorial bitangents. We compute their set of occluders in $O(k \log n)$ time by intersecting the bitangents with all the polytopes using their associated combinatorial polygons. Finally, we update the event queue in $O(k \log n)$ time.

Update sets of blockers. Consider now each of the $O(k)$ bitangents having e_1 as a support edge except for the two bitangents that might have just been created. We update its set of blockers as follows. First, note that only polytope $\mathbf{P}$ may have to be added to, or removed from, the set of blockers. Two situations occur: either the geometric bitangent segment joining the two support edges in Π_{t_0} properly intersects polygon P_{t_0}, or not. In the first case (e.g., segment pq in Figure 6.12), polytope $\mathbf{P}$ was and remains a blocker of the bitangent. In the second case (e.g., segment pr in Figure 6.12), $\mathbf{P}$ has to be either removed from, or added to, the set of blockers. This can be done in $O(k \log k)$ time by searching for $\mathbf{P}$ in the set (recall that polytopes are ordered by their index in a binary search tree).

Processing an I-event thus takes $O(k \log n)$ time. Since any polytope edge intersects any other polytope in at most two points, there are $O(nk)$ I-events which can be processed in $O(nk^2 \log n)$ time in total per sweep.

6.5.5.4 T-critical events

Suppose that on the line transversal to e_1, e_2, e_3 and l_e (the one associated to the T-event) edges e_1, e_2, e_3 are met in that order at points p_1, p_2, p_3. Let $\mathbf{Q}_i$ be the polytope containing e_i, $1 \leqslant i \leqslant 3$.

Update sets of blockers. Update the occluder set for the bitangent with support edges e_1 and e_3 by either removing $\mathbf{Q}_2$ (if it appears in the set) or adding $\mathbf{Q}_2$ (if it does not appear in the set); this can be done in $O(\log n)$ time.

Output. First determine if the segment p_1p_3 is unoccluded by checking if the set of blockers of the bitangent with support edges e_1 and e_3 is empty or reduced to $\mathbf{Q}_2$. If so and if the segment intersects the reference edge e, then it is a free segment transversal to the four edges e, e_1, e_2, e_3. In order to report each such transversal exactly once, we report it only if the reference edge e is smaller than e_2 for some global ordering of all edges. This can be done in constant time.

There are $O(nk^2)$ T-critical events per sweep (see the proof of Proposition 38), thus all the T-events can be processed in $O(nk^2 \log n)$ time per sweep.

6.5.6 Complexity

Note first that we assume a model of computation in which bounded-degree algebraic polynomials may be evaluated in constant time. See [60] for a detailed description of the predicates concerning line transversals that are used in this algorithm.

In this model of computation, we have described a $\Theta(n^2k^2 \log n)$-time algorithm for computing all the minimal sets of edges, no two chosen from the same polytope, that admit an isolated line transversal containing a free segment that is tangent to all these polytopes. As mentioned earlier, the sweep-plane algorithm can be easily modified to report all types of minimal support sets.

The space used by the algorithm is $\Theta(nk^2)$ in the worst case. To see this, first notice that storing the combinatorial polygons and the V-, F- and I-critical events uses $O(nk)$ space. There are also $O(nk)$ combinatorial bitangents in any sweep plane. Storing the combinatorial bitangents thus requires $\Theta(nk^2)$ space since, in the worst case, $\Theta(nk)$ of them may be intersected by $\Theta(k)$ polytopes. Furthermore, there may be $\Theta(nk^2)$ T-events in the queue since each of the $\Theta(nk)$ bitangents may share a support with $\Theta(k)$ other bitangents. This yields the bounds of Theorem 20 for computing minimal free segments.

Notice that, with a slight modification to the algorithm, and no increase in the time complexity, we can reduce the storage requirement of the T-events to $O(nk)$. To do this we maintain the bitangents sorted by polar angle around each vertex of the combinatorial polygons, which can easily be done since the cyclic ordering changes only at T-critical events or when a bitangent is created or deleted. Since two bitangents become aligned only when they are neighbors in this cyclic ordering, we only need to maintain the T-events for pairs of consecutive bitangents and there can only be $O(nk)$ of these at any one time.

Finally, the bounds of Theorem 20 that concern the computation of potentially occluded isolated lines tangent to polytopes are obtained by noticing that we need not maintain the sets of blockers of the bitangents which reduces the space requirements for the combinatorial bitangents to $O(nk)$.

6.6 Conclusion

We have presented a tight bound on the *number* of (connected components of) lines and maximal free line segments that are tangent to at least four among k possibly intersecting polytopes in arbitrary position. A problem that we leave open is to prove that the same bound holds for the *combinatorial complexity* of the set of all maximal free line segments among k polytopes.

We have also shown how to compute in near-optimal worst-case time all the *minimal* free line segments that are isolated transversals to their set of supports and tangent to the corresponding polytopes. We believe that our algorithm can also be made to report all connected sets of minimal free segments that are transversal to the same set of edges.

A problem that we have not solved, however, is to compute in the same time and space complexities, respectively, the polytopes supporting the endpoints of the corresponding *maximal* free line segments.

6.7 Bonus : complexité dépendant de la silhouette

On donne ici deux variantes des résultats de la section 6.3 prenant en compte la taille de la silhouette des polytopes de la scène.

Proposition 47. *En position générique, si on a 4 polytopes disjoints avec n arêtes au total dont un polytope avec au plus m arêtes sur sa silhouette quel que soit le point de vue, il y a $O(mn)$ droites isolées tangentes aux quatre polytopes.*

Démonstration. On appelle **S** le polytope dont on connaît la taille de la silhouette. On somme alors sur les arêtes de ce polytope. Dans la somme, on sépare le décompte des événements V de celui des événements F. Pour chaque face d'un des polytopes, les autres polytopes coupent son plan support en trois polygones convexes (éventuellement vides) et les événements F correspondent à des droites tangentes à deux de ces polygones. Le nombre d'événements F est donc $O(n)$. Pour compter les événements V, on remarque que compter le nombre de sommets intervenant pour chaque arête autour de laquelle on balaye est équivalent à compter pour chaque sommet le nombre d'arêtes de balayage pour lesquelles il intervient. Ces arêtes sont précisément celles qui sont sur la silhouette de **S** vu depuis ce sommet. Le nombre d'événements V est donc $O(mn)$. □

Proposition 48. *En position générique, si on a k polytopes disjoints avec n arêtes au total et que depuis tout sommet la somme des tailles des silhouettes des autres polytopes est bornée par m, il y a $O(mnk^2)$ droites isolées tangentes à quatre polytopes.*

Démonstration. La technique est la même que dans le lemme précédent. Le nombre total d'événements F est $O(nk^2)$. On doit maintenant compter combien de fois chaque sommet peut apparaître comme un événement V. Pour chaque arête sur la silhouette d'un des polytopes (m choix au plus) et pour chaque paire d'autres polytopes ($O(k^2)$ choix), le sommet risque d'être un événement V. Cela laisse $O(mk^2)$ possibilités par sommet, soit un total de $O(mnk^2)$ événements V. □

Pour fixer les idées, si les k polytopes du lemme précédent ont la même taille n/k et chacun une silhouette de taille majorée par $\sqrt{n/k}$ depuis tout point de vue (voir les chapitres 3 et 4 pour une raison de faire cette supposition), on a $O(n^{3/2}k^{5/2})$ droites isolées tangentes à quatre polytopes, ce qui est mieux que la borne en $O(n^2k^2)$ utilisée dans le cas général.

Chapter 7

On the Complexity of Umbra and Penumbra

Ce travail, réalisé avec J. Demouth, O. Devillers, H. Everett, S. Lazard et R. Seidel, a été présenté au Symposium on Computational Geometry [45] et soumis à Computational Geometry: Theory and Applications.

Computing shadow boundaries is a difficult problem in the case of non-point light sources. A point is in the umbra if it does not see any part of any light source; it is in full light if it sees entirely all the light sources; otherwise, it is in the penumbra. While the common boundary of the penumbra and the full light is well understood, less is known about the boundary of the umbra. In this paper we prove various bounds on the complexity of the umbra and the penumbra cast by a segment or polygonal light source on a plane in the presence of polygonal or polytopal obstacles.

In particular, we show that a single segment light source may cast on a plane, in the presence of two triangles, four connected components of umbra and that two fat convex obstacles of total complexity n can give rise to $\Omega(n)$ connected components of umbra. In a scene consisting of a segment light source and k disjoint polytopes of total complexity n, we prove an $\Omega(nk^2 + k^4)$ lower bound on the maximum number of connected components of the umbra and a $O(nk^3)$ upper bound on its complexity. We also prove that, in the presence of k disjoint polytopes of total complexity n, some of which are light sources, the umbra cast on a plane may have $\Omega(n^2k^3 + nk^5)$ connected components and has complexity $O(n^3k^3)$.

These are the first bounds on the size of the umbra in terms of both k and n. These results prove that the umbra, which is bounded by arcs of conics, is intrinsically much more intricate than the full light/penumbra boundary which is bounded by line segments and whose worst-case complexity is, as we will show, in $\Omega(nk + k^4)$ and $O(nk\alpha(k^2) + k^4)$; moreover, if there are only $O(1)$ light sources of total complexity m, then the complexity is in $\Omega(n\alpha(k) + km + k^2)$ and $O(n\alpha(k) + km\alpha(k) + k^2)$.

Scene type	Lower bounds	Upper bounds
Segment light source		
2 triangles	4	$O(1)$
2 fat polytopes	$\Omega(n)$	$O(n)$
k polytopes	$\Omega(nk^2 + k^4)$	$O(nk^3)$
Polytopal light source(s)		
k polytopes	$\Omega(n^2k^3 + nk^5)$	$O(n^3k^3)$

Table 7.1: Lower bounds on the number of connected components and upper bounds on the complexity of the umbra cast on a plane by k polytopes of total complexity $O(n)$. In the case of polygonal light sources, the lower bound holds for a single light source of complexity $O(n)$ and the upper bound when any number of the k polytopes are light sources.

Light sources	Lower bound	Upper bound
$O(1)$ polytopes of size m	$\Omega(n\alpha(k) + km + k^2)$	$O(n\alpha(k) + km\alpha(k) + k^2)$
$O(k)$ polytopes of size $O(n)$	$\Omega(nk + k^4)$	$O(nk\alpha(k^2) + k^4)$

Table 7.2: Bounds on the complexity of the union of umbra and penumbra cast on a plane by a set of k disjoint polytopes of total complexity n, some of which are light sources.

7.1 Introduction

Shadows play a central role in human perception [80, 106]. Unfortunately, computing realistic shadows efficiently is a difficult problem, particularly in the case of non-point light sources. A wide variety of approaches have been considered for rendering shadows (see, for example, the surveys [54, 110]) and many methods make extensive use of graphics hardware (see the survey [68]).

A point is in the *umbra* if it does not see any part of the light source(s); it is in *full light* if it sees entirely all the light source(s); otherwise, it is in the *penumbra*. While the boundary between the penumbra and the full light is reasonably well understood (see Section 7.3), less is known about the boundary of the umbra. Nevertheless, there is an extensive literature concerning the explicit computation of these shadow boundaries; see, for example, [49, 57, 55, 56, 70, 87, 101, 104].

In this paper we prove various bounds, summarized in Tables 7.1 and 7.2, on the complexity of the umbra and penumbra cast by segment or polygonal light source(s) on a plane in the presence of polygon or polytope (*i.e.* convex polyhedral) obstacles. In particular, we show that a single segment light source may cast, in the presence of two triangles, four connected components of umbra. We prove that the umbra defined by one segment light source and two fat convex obstacles of total complexity n can have $\Omega(n)$ connected components. We also prove an $\Omega(nk^2 + k^4)$ lower bound on the maximum number of connected components of the umbra and a $O(nk^3)$ upper bound on its complexity in a scene consisting of a segment light source and k disjoint polytopes of total complexity n.

Finally, we prove that the umbra cast on a plane by a polygonal light source and k convex obstacles can have $\Omega(n^2k^3 + nk^5)$ connected components and has complexity $O(n^3k^3)$. These are the first bounds on the size of the umbra in terms of both k and n.

These results are surprising in the sense that they show that the umbra cast by a single segment light source may have many connected components. The fact that the umbra may have four connected components in the case of two triangle obstacles comes as a total surprise. Our lower bounds of $\Omega(nk^2 + k^4)$ and $\Omega(n^2k^3 + nk^5)$ connected components, for k polytopes of total complexity n, are rather pathological in the sense that most of the obstacles are very long and thin. However, we also present a lower bound example of $\Omega(n)$ connected components in the case of two fat polygons or polytopes of complexity $O(n)$. Concerning the upper bounds of $O(nk^3)$ and $O(n^3k^3)$, even though these bounds are not *a priori* tight, they substantially improve the only previously known bounds for this problem which were the trivial $O(n^4)$ and $O(n^6)$ upper bounds. Finally, it is interesting to point out that even for the simplest case of non-point light sources, obtaining tight bounds on the complexity of the umbra and understanding its structure is a very challenging problem.

These results show that the umbra, which is bounded by arcs of conics, is intrinsically much more intricate than the full light/penumbra boundary which is bounded by line segments and for which we prove that the worst-case complexity is in $\Omega(nk + k^4)$ and $O(nk\alpha(k^2) + k^4)$, where $\alpha(k)$ denotes the pseudo-inverse of the Ackermann function; moreover, if there are only $O(1)$ light sources of total complexity m, then the complexity of the full light/penumbra boundary is in $\Omega(n\alpha(k)+km+k^2)$ and $O(n\alpha(k)+km\alpha(k)+k^2)$.

The paper is organized as follows. The next section provides notation and definitions. We give in Section 7.3 almost tight lower and upper bounds on the complexity of the boundary between full light and penumbra cast on a plane by a polygonal light source in the presence of polytopal obstacles. We present, in Section 7.4, upper bounds on the complexity of the umbra, in Section 7.5, lower bounds on the maximal number of connected components of umbra and conclude in Section 7.6.

7.2 Preliminaries

Let s be a line segment and p a point. We denote by $\langle s, p\rangle$ the set of line transversals of s through p. Similarly, for any triple of segments s_1, s_2 and s_3, we denote by $\langle s_1, s_2, s_3\rangle$ its set of line transversals. It is a well-known fact that $\langle s_1, s_2, s_3\rangle$ consists of lines belonging to the same regulus of a ruled quadric surface (see *e.g.* [92]). More precisely, the line transversals lie on a hyperboloid of one sheet when the three segments are pairwise skew and not all parallel to the same plane. If the segments are pairwise skew and all parallel to the same plane, then the line transversals lie on a hyperbolic paraboloid. Otherwise, they lie in one or two planes. Hence any set of transversals, whether $\langle s, p\rangle$ or $\langle s_1, s_2, s_3\rangle$, forms patches of a quadric (possibly degenerating to one or two planes). Moreover, the set of transversals consists of at most three patches, or more formally, at most three connected components in line space [29]. We let $\langle s, p\rangle$ and $\langle s_1, s_2, s_3\rangle$ denote not just sets of lines but also the surface patches they form.

Let $\mathcal{P}$ be a finite set of disjoint convex polygons or polytopes in $\mathbb{R}^3$ with $\mathcal{L} \subset \mathcal{P}$

identified as light sources. A surface $\sigma = \langle e, v \rangle$ is called an EV-*surface* if there exist two distinct objects $P, Q \in \mathcal{P}$ so that e is an edge of P, v a vertex of Q and σ intersects a light source. A surface $\sigma = \langle e_1, e_2, e_3 \rangle$ is called an EEE-*surface* if there exist three distinct objects $P, Q, R \in \mathcal{P}$ so that e_1, e_2 and e_3 are respectively edges of P, Q and R and σ intersects a light source.

Any plane Π intersects an EV-surface or an EEE-surface in a set of arcs of a conic (each possibly empty or possibly a line segment). Hence the intersection between Π and *all* the EV and EEE surfaces defines an arrangement of arcs of conics on Π.

Here we are interested in the arcs of conics that correspond to shadow boundaries. In particular, we are interested in arcs resulting from the intersection between Π and maximal free line segments[14] that intersect a light source and are supported by a line which is on an EV or EEE surface. The intersection of these free line segments with Π defines an arrangement of arcs of conics on Π which we call the *shadow arrangement* on the *shadow plane* Π.

A point p is in the umbra if for any point q on a light source, the segment pq intersects an object from $\mathcal{P} \setminus \mathcal{L}$. Similarly, p is in full light if for any point q on a light source, the segment pq does not intersect any object from $\mathcal{P} \setminus \mathcal{L}$. Otherwise, p is in the penumbra.

We will make extensive use of the fact that the boundary of the umbra and penumbra consists of arcs of the shadow arrangement (see, for example, [70]). Notice that not all arcs of the shadow arrangement are on the umbra or penumbra boundaries; some arcs correspond to other lighting discontinuities.

Throughout this paper, we consider the regions of umbra and penumbra on a plane cast by a segment light source or polygonal light source(s) in the presence of convex polygons or polytopes.

7.3 The penumbra boundary

We recall here some straightforward and well-known properties of the penumbra and give bounds on the complexity of the common boundary of the penumbra and the full light. In this section we refer to the union of the umbra and penumbra as the *shadow region*.

Property 1. The shadow region cast by a light source on a plane in the presence of obstacles is the union of all the shadow regions cast by each obstacle.

Property 2. The shadow region cast on a plane Π by a polygonal light source S in the presence of one polytope P is the intersection of half-planes in Π, each of which is defined as the intersection of Π with a half-space that contains P but not S, is bounded by a plane tangent to both of them, and contains an edge of one of them.

Note that these two properties imply that the boundary of the shadow region is only composed of line segments induced by EV-surfaces.

Theorem 49. *The complexity of the shadow region cast on a plane by a set of k disjoint polytopes of total complexity n, some of which are light sources, is, in the worst case, in*

[14] A *maximal free line segment* is a segment that intersects the interior of no object and whose endpoints lie on some object or at infinity.

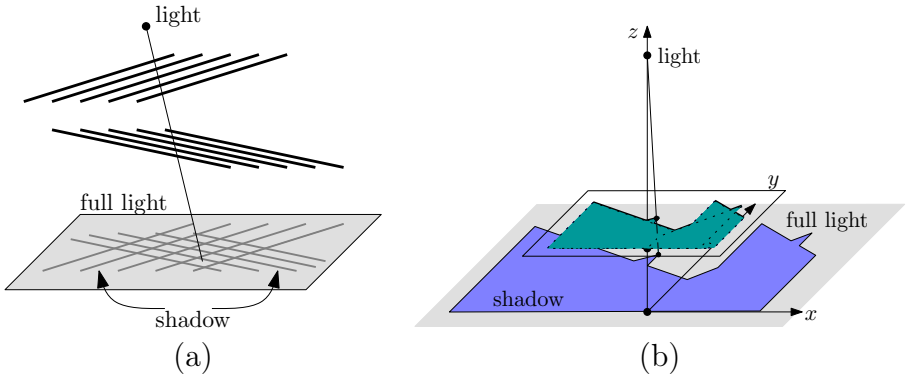


Figure 7.1: $\Omega(k^2)$ and $\Omega(k\alpha(k))$ lower bounds.

$\Omega(nk+k^4)$ and $O(nk\alpha(k^2)+k^4)$, where $\alpha(k)$ denotes the pseudo-inverse of the Ackermann function. If there are only $O(1)$ light sources of total complexity m, then the complexity of the shadow region is in $\Omega(n\alpha(k)+km+k^2)$ and $O(n\alpha(k)+km\alpha(k)+k^2)$.

Proof. By Property 2, the shadow cast on a plane Π by a polygonal light source in the presence of one polytope is a convex polygon. Furthermore, if the light source has m edges and the polyhedron has n_i edges, the shadow region in Π has $O(n_i+m)$ edges. By Property 1, the shadow region in the presence of k polytopes of total complexity n is thus the union of k convex polygons of total complexity $O(n+km)$, which has complexity $O((n+km)\alpha(k)+k^2)$ [16]. The same bounds hold if there are $O(1)$ light sources of total complexity m. Similarly, if $O(k)$ of the polytopes of the scene are light sources, of total complexity n, the shadow region is the union of $O(k^2)$ convex polygons of total complexity $O(nk)$, which has complexity $O((nk)\alpha(k^2)+k^4)$.

For the proof of the lower bounds, we consider the following collection of examples. In all constructions the shadow plane Π is the plane $z=0$.

$\Omega(k^2)$ and $\Omega(k^4)$ examples. Refer to Figure 7.1. We consider a point light source at a height z (large enough) and a grid consisting of k thin horizontal and parallel rectangles at height $z=1$ together with k other thin horizontal and parallel rectangles at height $z=2$. They form a grid of shadow on plane Π which has size $\Omega(k^2)$. Replacing the point light source by k point light sources very close to each other gives, similarly, a shadow region of complexity $\Omega(k^4)$.

$\Omega(k\alpha(k))$ example. Refer to Figure 7.1. Again, the light source is a point with large positive z-coordinate. We consider a set of k line segments in plane $z=1$ (with positive y coordinates) having, in that plane, an upper envelope of size $\Omega(k\alpha(k))$[109]. We transform each line segment into a trapezoid linking it to its projection on the $y=0$ line (in plane $z=1$). We get a set of trapezoids whose shadow, in plane $z=0$, for a point light source at large enough z is basically the upper envelope of the segments. Note that the trapezoids can easily be made disjoint by placing them in different horizontal planes very close to plane $z=1$.

$\Omega(n\alpha(k))$ example. Refer to Figure 7.2. First modify the above $\Omega(k\alpha(k))$ example such that the left "vertical" side of each trapezoid has slope γ and the right "vertical" side has

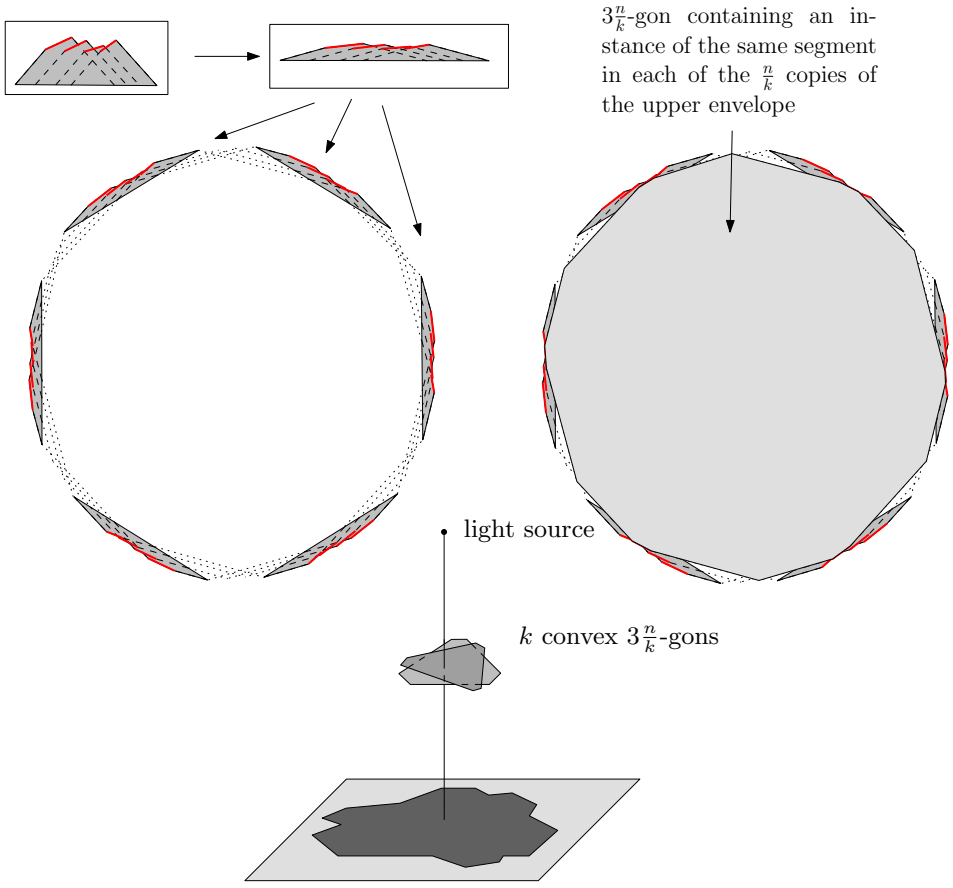


Figure 7.2: $\Omega(n\alpha(k))$ lower bound.

slope $-\gamma$, for some γ large enough. Now, by some suitable scaling, we make all slopes of the vertical walls strictly smaller than $\frac{k\pi}{n}$. Assemble $\frac{n}{k}$ copies of the previous construction into a large regular $\frac{n}{k}$-gon where each side is, in fact, a $k\alpha(k)$ upper envelope. Finally, this construction can be seen as k convex $3\frac{n}{k}$-gons by connecting all of the $\frac{n}{k}$ copies of the same trapezoid by extending their walls.

This set of k convex $3\frac{n}{k}$-gons, embedded in different horizontal planes close to the $z = 1$ plane, engender, in the presence of a point light source at large enough z, shadows of complexity $n\alpha(k)$.

Ω(mk) example. Refer to Figure 7.3. We use a horizontal m-gon as light source and a thin rectangle as obstacle. Then the shadow has $\Omega(m)$ size. Using multiple copies of the obstacle such that the different shadows are disjoint easily gives an $\Omega(mk)$ example. □

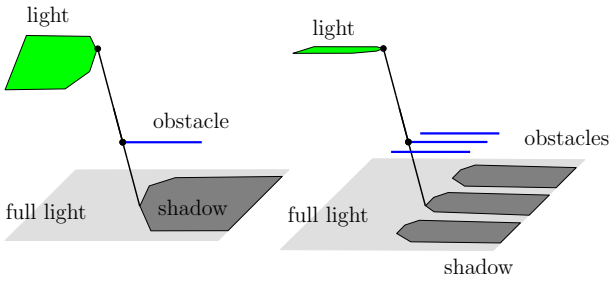


Figure 7.3: $\Omega(mk)$ lower bound.

There is still a small gap between the $\Omega(nk+k^4)$ lower bound and the $O(nk\alpha(k^2)+k^4)$ upper bound and between the $\Omega(k^2+mk+n\alpha(k))$ lower bound and the $O(k^2+mk\alpha(k)+n\alpha(k))$ upper bound. In fact we conjecture that the lower bounds are tight because the shadows of the different obstacles cast by a single light source have some similarity with homothetic projections of the light, and the union of k convex homothetic m-gons is $\Theta(mk)$ since two convex homothetic polygons intersect in at most two points [76].

7.4 Upper bounds

In this section we prove the following two upper bounds on the complexity of the umbra cast by a segment light source or polygonal light source(s) on a plane.

Theorem 50. *The complexity of the umbra cast on a plane by one segment light source in the presence of k disjoint polytopes of total complexity n is $O(nk^3)$.*

Theorem 51. *The complexity of the umbra cast on a plane by a set of k disjoint polytopes of total complexity n, some of which are light sources, is $O(n^3k^3)$.*

7.4.1 The umbra cast by a segment light source

We will actually prove an upper bound on the complexity of the shadow arrangement which yields the same bound for the complexity of the umbra. Notice that, in the case

of a single segment light source, the EEE-surfaces $\sigma = \langle e_1, e_2, e_3 \rangle$ that contribute to the shadow arrangement are such that one of e_1, e_2 or e_3 is the segment light source. Similarly, when considering EV-surfaces $\sigma = \langle e, v \rangle$ either e is the segment light source or v is one of its endpoints.

We prove Theorem 50 by considering a plane rotating about the line supporting the segment light source. First, if the segment light source, s, is not parallel to the shadow plane Π, we apply a projective transformation to the scene, sending to infinity the point of intersection between the line containing s and plane Π; this does not change the complexity of the shadow arrangement. We can thus assume in the rest of this section that the segment light source is parallel to Π. The sweep plane, denoted π, intersects the shadow plane in a line parallel to s; we will say that, throughout the sweep, this line moves from left to right.

We start with two preliminary lemmas concerning the intersections that an instance of the sweep plane can have with the shadow arrangement.

Let α be the conic that is the intersection of the shadow plane Π and the EEE-surface $\sigma = \langle s, e_1, e_2 \rangle$ where s is the segment light source, e_1 and e_2 are two other segments, and let π be a plane containing the light source s and intersecting Π.

Lemma 52. *If s, e_1, and e_2 are pairwise skew then π intersects α in at most a single point.*

Proof. First, since s is parallel to Π, any transversal to s and to a point in $\alpha \cap \pi$ lies in plane π. If the intersection between π and e_1 or e_2 is empty, there exists no line tangent to the three segments in π and thus π does not cross α. Otherwise, segments e_1 and e_2 intersect π in two points p_1 and p_2, respectively. The line defined by p_1 and p_2 is the only line that is possibly a transversal to the three segments in π (it may not intersect the segment s). Hence there exists at most one transversal in π which defines exactly one point on α. □

Notice that the preceding lemma implies that the conic arcs are monotonic in the direction of the sweep (orthogonal to s).

For the rest of the proof we consider an arrangement A of arcs of conics which contains the shadow arrangement. We will establish an upper bound of $O(nk^3)$ on the complexity of A which will yield the bound of Theorem 50. The arrangement A consists of the intersection of Π with (i) those lines that are transversal to the light source s, and the edges e_1 and e_2 of two other polytopes and that do not intersect the interior of these polytopes (the connected components of these lines form patches of EEE-surfaces) and (ii) those lines that are transversal to a vertex and an edge of two polytopes, one of which is the segment light source, and that do not intersect the interior of these polytopes (the connected components of these lines form patches of EV-surfaces).

We now count the number of crossings between an instance of the sweep plane π and the arcs in A.

Lemma 53. *The plane π properly intersects at most $O(k^2)$ arcs of A.*

Proof. The arcs of A are defined as the intersection with Π of lines ℓ which are (i) transversal to the segment light source s and tangent to two polytopes, (ii) transversal to an

endpoint of s and tangent to another polytope, or (iii) transversal to s and to a polytope vertex.

An instance π of the sweep plane never properly intersects an arc of type (iii) (since such an arc is either included in π or does not intersect it). Now, if π intersects an arc of one of the other two types, then π contains the corresponding line ℓ, which is tangent to two polygons of $\mathcal{P} \cap \pi$. Since $\mathcal{P} \cap \pi$ consists of at most k disjoint convex polygons, there are $O(k^2)$ lines in π that are tangent to two polygons of $\mathcal{P} \cap \pi$, hence the result. □

Proof of Theorem 50. We consider an orthogonal frame in plane Π whose *vertical* axis is parallel to the segment light source s; the other axis is called *horizontal*.

We first show that the number of proper intersection points between arcs of A is $O(k^2)$ times the number of arcs. It follows from Lemma 52 (along with a simple argument in the case that the three segments are not pairwise skew) that any arc of A is either horizontally monotone or is a vertical line segment. Consider an arc α_0 and its rightmost endpoint p (any one if α_0 is vertical). We charge to α_0 all points of intersection involving α_0 and all arcs whose rightmost endpoints are strictly to the right of p. Any arc properly intersects α_0 in at most $O(1)$ points so the number of intersection points charged to α_0 is bounded by the number of arcs properly intersected by the sweep plane containing p. By Lemma 53, there are at most $O(k^2)$ such arcs. Thus, each arc is charged at most $O(k^2)$ times.

We now bound the number of arcs (and thus the number of arc endpoints) generating A. Each arc corresponds either to a patch of an EV or EEE surface. Consider first the EV-surfaces. Since either the edge or the vertex is on the light source, there are at most $O(n)$ such surfaces.

Now consider the arcs generated by EEE-surfaces. Let n_i be the number of vertices of polytope P_i, $1 \leqslant i \leqslant k$. The number of EEE-surfaces involving the light source and edges from polytopes P_i and P_j is $O(n_i + n_j)$ [28, Corollary 2.6]. Then, $\sum_{1 \leqslant i < j \leqslant k} O(n_i + n_j) = O(nk)$.

There are at most $O(nk)$ arcs generating A. Since each arc is charged with at most $O(k^2)$ intersection points, there are at most $O(nk^3)$ intersection points. The total complexity of the shadow arrangement, and thus of the umbra, is then $O(nk^3)$. □

Remark. The upper bound of Theorem 50 is not known to be tight. However, we prove, in Lemma 59, that this is a tight bound on the complexity of the arrangement A.

7.4.2 The umbra cast by polygonal light sources

To prove Theorem 51 we consider an arrangement B of arcs of conics that, as in the previous section, contains the shadow arrangement. This arrangement B consists of the intersections of Π with (i) the lines that are transversal to a vertex and an edge of two polytopes and that do not intersect the interior of these polytopes (the connected components of these lines form patches of EV-surfaces) (ii) the lines that are transversal to edges of three polytopes and that do not intersect the interior of these polytopes (the connected components of these lines form patches of EEE-surfaces). Notice that B may contain arcs generated by surfaces that do not intersect the light source or possibly by surfaces that intersect the interior of other polytopes in the scene. We will establish a $O(n^3k^3)$ upper

bound on the complexity of B which yields the same bound for the complexity of the umbra.

Lemma 54. *Any line L in Π properly intersects at most $O(nk^2)$ arcs of B.*

Proof. An intersection point between L and B corresponds to a line transversal which belongs to an EV or EEE surface. Consider first EV-surfaces. The line transversal lies in a plane which contains L and a vertex, say v, of one of the polytopes. There exist $O(n)$ such planes and in each of them there are at most $O(k)$ lines through v that are tangent to a polytope (since we only consider proper intersections between L and the arcs of B). Thus there are at most $O(nk)$ points on $L \cap B$ which correspond to lines in EV-surfaces.

Now we consider EEE-surfaces. Let n_i be the number of vertices of polytope P_i, for $1 \leqslant i \leqslant k$. The number of EEE-surfaces generated by three edges of polytopes P_i, P_j and P_l, not intersecting the interior of P_i, P_j and P_l, and that intersect L is $O(n_i + n_j + n_l)$ [28, Main Lemma]. Since $\sum_{1 \leqslant i < j < l \leqslant k} O(n_i + n_j + n_l) = O(nk^2)$, there are at most $O(nk + nk^2) = O(nk^2)$ arcs of B which intersect the line L on Π. □

Proof of Theorem 51. Here, we introduce an arbitrary coordinate frame Oxy in the plane Π. We call Ox the horizontal axis and Oy the vertical axis.

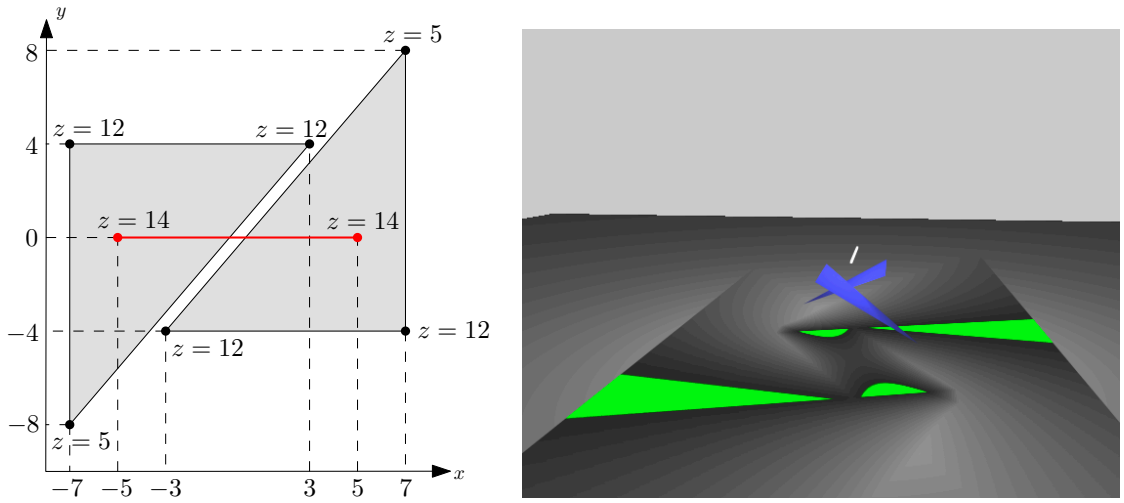
As in the proof of Theorem 50, we first show that the number of intersection points between arcs of B is $O(nk^2)$ times the number of conic arcs. We first break all conic arcs into maximal horizontally monotone pieces. This increases the number of arcs only by a constant factor. Consider a conic arc α_0 and its rightmost endpoint p along Ox (any endpoint if the arc is vertical). We charge to α_0 all points of intersection between α_0 and all conic arcs whose rightmost endpoints are strictly to the right of p. Any arc properly intersects α_0 in at most $O(1)$ points so the number of intersection points charged to α_0 is bounded by the number of arcs that are properly intersected by a vertical line in Π and containing p. By Lemma 54, there are at most $O(nk^2)$ such arcs. Thus, each arc is charged at most $O(nk^2)$ times.

We now bound the number of arcs (and thus the number of arc endpoints) generating B. Let n_i be the number of vertices of polytope P_i, $1 \leqslant i \leqslant k$ and e an edge. The number of EEE-surfaces pertinent to B and involving e and edges from polytopes P_i and P_j is $O(n_i + n_j)$ [28, Corollary 2.6]. Thus, for each edge e, there are, at most, $\sum_{1 \leqslant i < j \leqslant k} O(n_i + n_j) = O(nk)$ EEE-surfaces having e as a generating segment. Furthermore, the number of EV-surfaces involving edge e or one of its vertices is $O(n)$. Since there exist n edges, the total number of arcs in B is therefore $O(n^2k)$.

In conclusion, there are at most $O(n^2k)$ arcs generating B, each of them charged with at most $O(nk^2)$ intersection points, hence there are at most $O(n^3k^3)$ intersection points. The total complexity of B and, thus of the umbra, is $O(n^3k^3)$. □

7.5 Lower bounds

In this section we present several lower bounds on the complexity of the umbra.



(a) (b)

Figure 7.4: (a) Two triangles and a segment light source (viewed from above) that cast 4 connected components of umbra on the plane $z = 0$. (b) The scene rendered with the ray tracer `OpenRT` (the umbra is in light grey); courtesy of Andreas Dietrich.

7.5.1 The umbra cast by a segment light source

Here we concentrate on the umbra cast by a segment light source in the presence of various configurations of obstacles.

Theorem 55. *A segment light source and two triangles may cast, on a plane, four connected components of umbra.*

Proof. Consider the following scene consisting of a segment light source, s, two triangles, T_1 and T_2, and a shadow plane, Π, the horizontal plane of equation $z = 0$; see Figure 7.4(a) and (b).

Figure 7.5(a) shows a superset of the shadow arrangement generated by this configuration (the arrangement A defined in Section 7.4). Although it can be shown that the four shaded regions in the figure are exactly the umbra, we will simply argue here that there are at least four connected components. We do this by illustrating four segments in the umbra and then arguing that they are each in different connected components.

The idea is to consider a series of planes rotating about the segment light source and the intersections of those planes with the two triangles and the shadow plane; Figure 7.6 shows such a sequence. We then examine the umbra in those planes by considering the relevant bitangents.

Let P_+ be one such plane (containing s) and going through the point $(0, 7, 0)$ and L_+ the intersection of P_+ and Π. Figure 7.6(b) shows the segment s, the intersections between P_+ and the two triangles T_1 and T_2, L_+ and four bitangents that together define the umbra on L_+. Consider the two segments R_1^+ and R_2^+ as shown in Figure 7.6(b). It is easy to see, by examining the bitangents, that R_1^+ and R_2^+ are in the umbra. Hence there are two segments of umbra on the line L_+. We obtain two other segments, R_1^- and R_2^-,

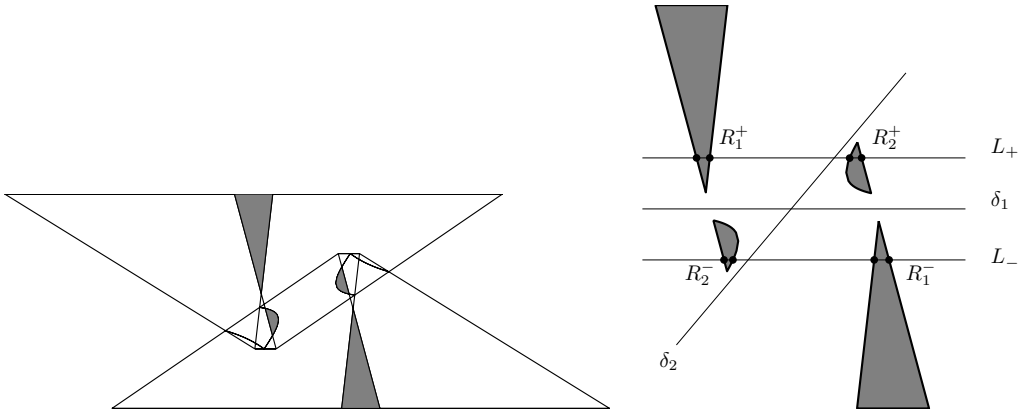


Figure 7.5: (a) Superset of the shadow arrangement on plane Π. The four shaded regions are the regions of umbra. (b) The four connected components of umbra and the four lines used in the proof of Theorem 55.

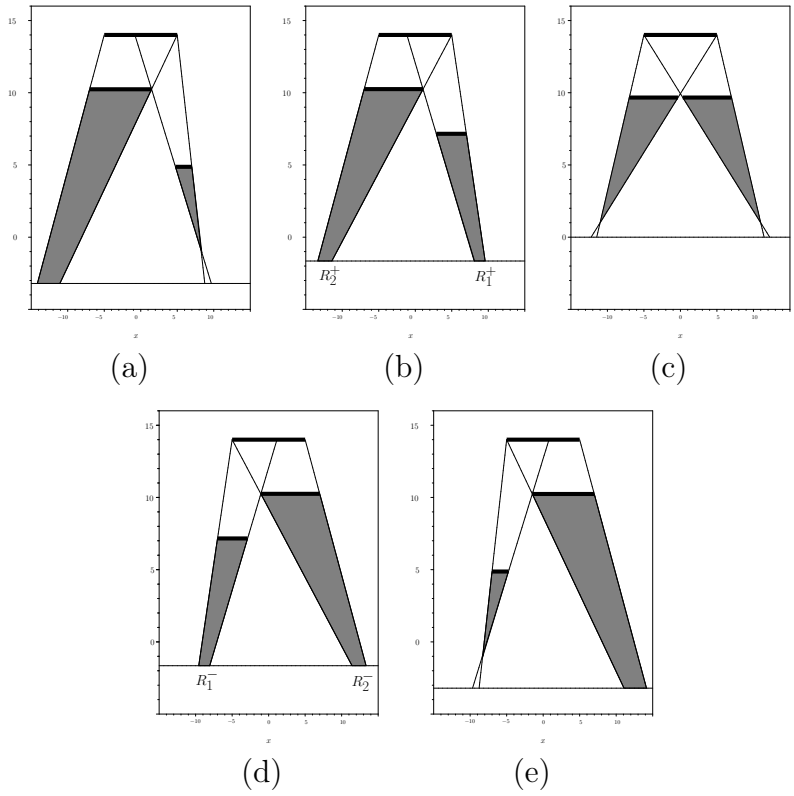


Figure 7.6: Views in the sweep plane with bitangents that define the umbra. The number of components of umbra in the intersection of the sweep plane and the plane $z = 0$ is : (a) one (sweep plane through $(0, 10, 0)$), (b) two (sweep plane P_+ through $(0, 7, 0)$), (c) zero (sweep plane $y = 0$), (d) two (sweep plane P_- through $(0, -7, 0)$), (e) one (sweep plane through $(0, -10, 0)$).

by taking the symmetric plane P_- with respect to the xz-plane (through point $(0, -7, 0)$ and whose intersections with the scene is shown on Figure 7.6(d)).

Now, we show that the four segments R_1^+, R_1^-, R_2^+ and R_2^- lie in different connected components of umbra. In order to prove this result, we exhibit two lines on Π which contain no point in the umbra and separate the four segments as shown in Figure 7.5(b).

First consider the plane $y = 0$ containing the light segment s and orthogonal to the shadow plane Π. This plane intersects Π in a line, δ_1, as shown in Figure 7.5(b), and separates R_1^+ and R_2^+ from R_1^- and R_2^- since P_+ and P_- are symmetric with the plane $y = 0$. To show that δ_1 contains no point of the umbra, consider the intersection of the $y = 0$ plane with the segment s and the two triangles T_1 and T_2; see Figure 7.6(c). A study of the bitangents reveals that no point of δ_1 lies in the umbra.

Now consider the plane orthogonal to Π, parallel to the two triangle hypotenuses and going through the midpoint of s. Let δ_2 be the intersection of this plane with Π; see Figure 7.5(b). Elementary computations show that the line δ_2 separates R_1^+ and R_2^- from R_1^- and R_2^+. There can be no point of the umbra on δ_2 since the plane intersects the light source but not the triangles (see Figure 7.4(a)). This completes the proof. □

Note that the line supporting s and the lines supporting the triangle hypotenuses are pairwise skew and not all parallel to a same plane. Thus the corresponding EEE-surface is a hyperboloid of one sheet which intersects Π in a hyperbola. We determine the equation of this hyperbolic curve to be $41y^2 - 52xy + 928 = 0$. This curve admits two asymptotes which contain no point in the umbra and which separate the connected components of umbra. One of these asymptotes is δ_1 and we could have chosen the other to be δ_2.

Note also that in our example, the light source is parallel to the shadow plane, and there are also many symmetries. None of this is critical; the example can be perturbed and the result still holds.

We now prove a lower bound for fat polytopes, polytopes whose aspect ratios are bounded from below by a positive constant when n goes to infinity.

Theorem 56. *The umbra cast on a plane by one segment light source in the presence of two fat disjoint polytopes of total complexity n can have $\Omega(n)$ connected components.*

Proof. Our lower-bound example consists of one segment light source s_1, a polytope Q_2 of size $O(n)$, and another polytope, Q_3, of constant size. Refer to Figure 7.7.

First we consider three skew lines $l_1 \supset s_1$, l_2 and l_3 and $\sigma = \langle s_1, l_2, l_3 \rangle$ the quadric patch(es) consisting of the lines stabbing s_1, l_2 and l_3. In the shadow plane Π, by adding suitable halfplanes P_2 and P_3 as obstacles limited by the lines l_2 and l_3, we obtain α, a single conic arc of $\sigma \cap \Pi$, bounding the umbra where the umbra is on the concave side of α (Figure 7.7-left).

We now consider p_1, one of the endpoints of s_1, and α_2, α_3, the intersections of planes P_2, P_3 with the cone of apex p_1 and base α (Figure 7.7-center).

Next, we reduce the obstacles P_2 and P_3 to convex polygons Q_2 and Q_3 by limiting them by a polygonal approximation of α_2 and α_3 such that Q_3 remains within (*i.e.,* on the convex side of) α_3 and Q_2 intersects α_2 n times (Figure 7.7-right). The umbra cast by s_1 on Π in the presence of Q_2 and Q_3 then consists of n connected components that

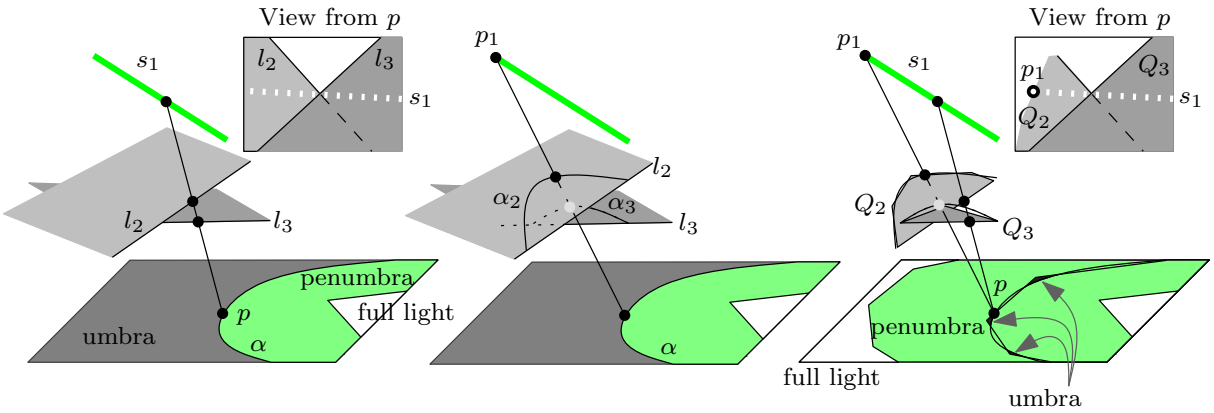


Figure 7.7: $\Omega(n)$ lower bound.

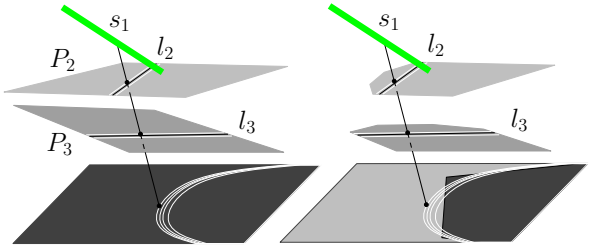


Figure 7.8: $\Omega(nk^2)$ lower bound.

are the intersection of the concave region outside α and the convex polygon that is the intersection of the cone of apex p_1 and base Q_2 with the plane Π.

Note that the polygons Q_2 and Q_3 are fat since Q_2 consists of a segment and of an approximation of a conic and Q_3 is of constant size. Finally, polygons Q_2 and Q_3 can be trivially transformed into fat polytopes without changing the umbra. □

Theorem 57. *The umbra cast on a plane by one segment light source in the presence of k disjoint polytopes of total complexity n can have $\Omega(nk^2)$ connected components.*

Proof. Consider three non-parallel segments s_1, l_2, and l_3 all parallel to the shadow plane Π and planes $P_2 \supset l_2$ and $P_3 \supset l_3$ parallel to Π, refer to Figure 7.8. The surface $\langle s_1, l_2, l_3 \rangle$ intersects Π in a conic arc α.

Now consider the following setup: s_1 is the light source; P_2 has k narrow rectangular holes (or slits) parallel and arbitrary close to l_2; similarly P_3 has k slits parallel and arbitrary close to l_3. (A plane with k such slits can be modelled by $O(k)$ rectangles.) Each pair of slits, s_2 from P_2 and s_3 from P_3, together with the light source s_1 induce a piece of penumbra in Π that is essentially a thickened copy of the conic arc α.

We thus get that the umbra covers the whole plane Π except for k^2 curves of penumbra that are all close to α (see Figure 7.8-left).

Finally, we trim the two planes P_2 and P_3, creating an n-sided convex polygon on Π such that the region outside this polygon is in light or penumbra and each edge intersects

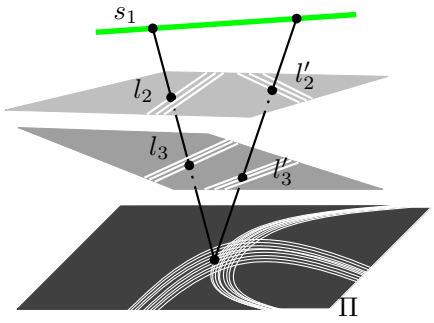


Figure 7.9: $\Omega(k^4)$ lower bound.

all the k^2 curves. The umbra then consists of $O(nk^2)$ regions inside the convex polygon and between the k^2 conics (see Figure 7.8-right). Note that the $O(k)$ convex obstacles can each be transformed into a polytope by the addition of a single vertex without changing the umbra. □

Theorem 58. *The umbra cast on a plane by a segment light source in the presence of k disjoint polytopes can have $\Omega(k^4)$ connected components.*

Proof. Refer to Figure 7.9. As in the previous lower-bound example, we create k^2 curves of penumbra using parallel thin holes. Making a second set of thin holes in each plane, we create a second family of curves of light and penumbra intersecting the first one. The umbra is now the complement of the union of these two sets of curves and it consists of $\Omega(k^4)$ connected components. □

We now present a lower bound on the complexity of the arrangement A, introduced in Section 7.4.1, which proves (together with the proof of Theorem 50) that the complexity of this arrangement is $\Theta(nk^3)$.

Lemma 59. *The arrangement A cast by one segment light source in the presence of k convex obstacles of total complexity n can have an $\Omega(nk^3)$ complexity.*

Proof. The proof is obtained by using large obstacles with small holes so that the shadow plane Π is almost completely in umbra with small curves of penumbra drawn by rays going through these holes. Such obstacles with small holes can be created using several convex polygons.

We consider a conic α drawn in Π and a line l_1 (containing the light source) intersecting Π in a point of α.

Let $x_1 = \alpha \cap l_1$ (x_1 is at infinity if l_1 is parallel to Π) and x_2, x_3, x_4, x_5 be four other points on α and x_6, x_7 be two other points on l_1, then any quadric through these seven points contains α and l_1. Since a quadric is defined by nine points, by choosing the two other points in different manners we can construct different quadrics containing l_1 and α.

First we construct a set of quadrics, σ_1, $\sigma_2 \ldots \sigma_k$, close together. On σ_i we choose two lines $l_{2,i}$ and $l_{3,i}$, such that all the lines $l_{h,i}$ for all i and a given h are close together. Then

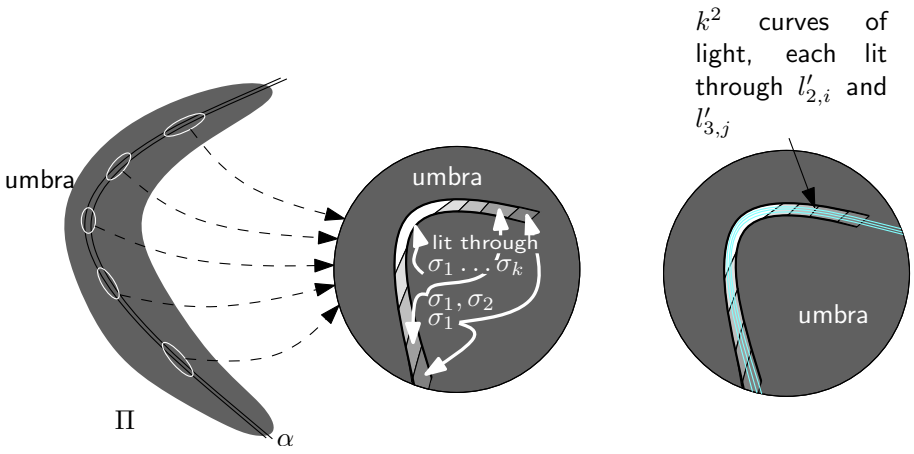


Figure 7.10: $\Omega(nk^3)$ lower bound on the complexity of arrangement A.

we can construct a (non-convex) polyhedral obstacle of size $O(k)$ with thin holes around lines $l_{h,i}$ for $h = 2, 3; 1 \leqslant i \leqslant k$. We consider a plane P above, and parallel to, Π that intersects the σ_i in a family of conics β_i that are close together. We can construct in plane P a convex n-gon Q that intersects all these conics n times each.

Before the introduction of Q, the picture in Π is a slab of penumbra around α. Adding Q cuts this penumbra in n pieces, and, in each piece, there are k different intensities; indeed, when we move a point p along α from a point such that all the k transversals through p and l_1, $l_{2,i}$, and $l_{3,i}$ are non-obstructed, these transversals through p get obstructed one after another by Q until they are all obstructed and p is in the umbra (see Figure 7.10 left and center). This defines $\Omega(nk)$ regions of penumbra of different intensities of light around α.

Now, we choose a completely different quadric σ' far from the σ_i but still containing α and l_1. Thus, we can choose two lines l'_2 and l'_3 in the same regulus as l_1 and arrange obstacles to light Π through thin holes around l'_2 and l'_3. The area lit through these holes is also a small area around α. We can arrange things such that this area is much thinner than the previous one, and thus we get a thin curve of penumbra which crosses all the $\Omega(nk)$ regions we have described above. Now, making k holes $l'_{2,i}$ around l'_2 and k holes $l'_{3,j}$ around l'_3 we transform this thin penumbra curve into k^2 "parallel" curves making the size of the arrangement $\Omega(nk^3)$. □

7.5.2 The umbra cast by a polygonal light source

Note that the lower bound of $\Omega(nk^2 + k^4)$ of Section 7.5.1 for a segment light source can easily be modified into a lower bound of $\Omega(nk^3 + k^6)$ in the case of a polygonal light source (by adding a third plane with $O(k)$ slits and a big polygonal light source). We present here a lower bound of $\Omega(n^2k^3 + nk^5)$ on the complexity of the umbra cast by a polygonal light source in the presence of k polygonal obstacles.

Theorem 60. *The umbra cast on a plane by one polygonal light source in the presence of k disjoint polytopes of total complexity n can have $\Omega(n^2k^3)$ connected components.*

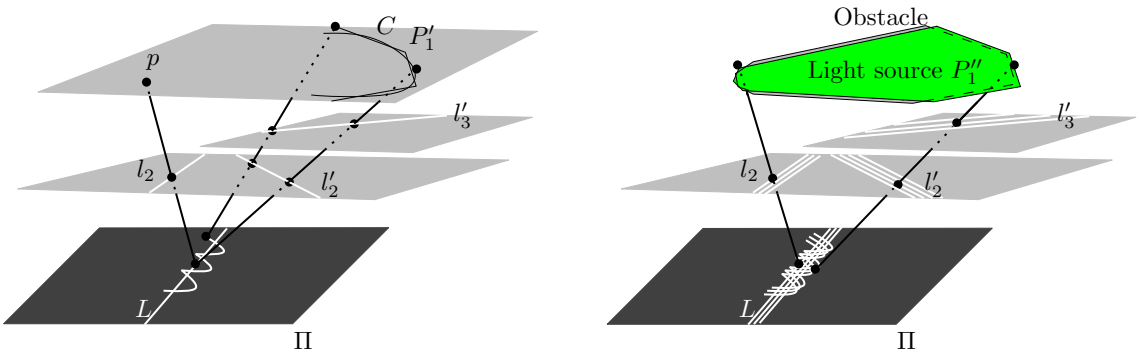


Figure 7.11: $\Omega(n^2k^3)$ lower bound.

Proof. Refer to Figure 7.11. Let p be a point and P_1 a small n-gon light source very close to p. Add an n-gon obstacle very close to the light source so that the light source behaves like n point light sources (when viewed from the correct side).

Now consider a plane obstacle with k thin holes parallel to a line l_2. This creates nk parallel thin lines of light on the shadow plane that can be made arbitrarily close to a line L (by having the k thin holes sufficiently close to each other and the n point light sources sufficiently close to each other). Note that by duplicating this construction (and thus with two polygonal light sources which behave as $2n$ point light sources) we get an arrangement of $2nk$ lines of light with n^2k^2 connected components of umbra.

Now consider two lines l_2' and l_3' that together with L admit a quadric as line transversals. Cut this quadric by a plane and approximate (a piece C of) the resulting conic by a convex n-polyline, P_1'. The set of transversals to the boundary of this polyline with l_2' and l_3' defines a curve on the shadow plane that cuts L order n times. We define a light source as the convex hull of P_1' and put an obstacle very close to it so that the light source behaves as if the polyline P_1' was the light source (when viewed from the right region). Now, replacing l_2' and l_3' by two plane obstacles with order k thin holes close and parallel to l_2' and l_3', respectively, we get k^2 curves of light, each of which intersects order n times each of the nk lines of light close to L. This give $\Omega(n^2k^3)$ connected component of the umbra.

Note that the two light sources P_1 and P_1' can be merged into one by considering P_1'' in the same plane as P_1, by noticing that there are enough degrees of freedom on l_2' and l_3' so that the convex hull of P_1 and an arc of the conic contains C on its boundary. □

Theorem 61. *The umbra cast on a plane by one polygonal light source in the presence of k disjoint polytopes of total complexity n can have $\Omega(nk^5)$ connected components.*

Proof. Refer to Figure 7.12. Consider three horizontal pairwise skew lines l_1, l_2, l_3 that lie above a horizontal plane Π and let C be the conic intersection of their common transversal with Π. Replace each of the l_i by a plane obstacle and k thin holes close to l_i and place a large (horizontal) light source S above these planes obstacles.

Consider now an n-gon P that intersects C order n times. Let s_4 and s_5 be two intersecting horizontal segments. Let P' be the symmetric of P with respect of the point

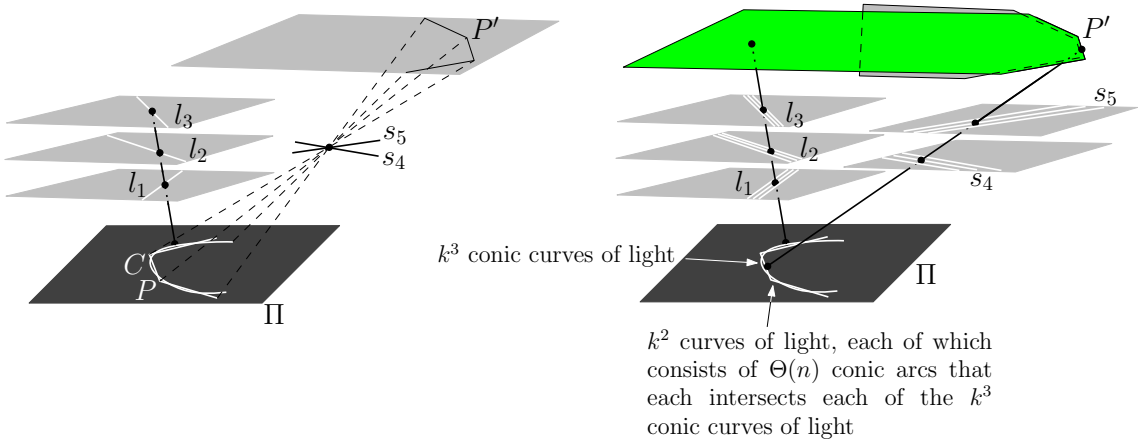


Figure 7.12: $\Omega(nk^5)$ lower bound.

of intersection between s_4 and s_5. We consider P' as a light source and put an obstacle very close to it so that it behaves as a one-dimensional polygonal light source when viewed from C. This induces on the shadow plane a polyline of light that intersects C order n times.

Now perturb segments s_4 and s_5 so that they do not intersect and replace them by (horizontal) plane obstacles with k thin holes close and parallel to s_4 and s_5, respectively. We hence get k^2 curves of light, each of which consists of order n conic arcs that each intersects C; hence each of these k^2 curves of light intersects C order n times. By chosing the holes near l_1, l_2 and l_3 sufficiently close to each other, respectively, each of the k^2 curves of light close to P intersects $O(n)$ times each of the k^3 curves of light close to C. We hence get $O(nk^5)$ connected components of umbra. □

7.6 Conclusion

The purpose of this paper is to establish the complexity of the boundaries between the umbra, penumbra and fully-lit regions on a plane in a polyhedral scene consisting of k convex objects of total complexity n.

The results presented here constitute a first step toward understanding the intrinsic structure and complexity of the umbra in this setting. We have proved that if the light is reduced to one line segment, then the umbra may have $\Omega(nk^2+k^4)$ connected components and $O(nk^3)$ complexity. We have also shown that a polygonal light source could generate an umbra with $\Omega(n^2k^3 + nk^5)$ connected components and $O(n^3k^3)$ complexity. In both cases these components of umbra are delimited by arcs of conics. These results prove that the umbra is intrinsically much more intricate than the boundary between full light and penumbra which is bounded by line segments and has complexity $O(n\alpha(k)+km\alpha(k)+k^2)$, where m is the complexity of the light source.

Our upper bounds, in fact, apply to the complexity of the arrangement of the curves where the derivative of the light intensity is discontinuous. These arrangements clearly include the boundary of the umbra, but also a lot of curves inside the penumbra that

are not relevant to the umbra. Furthermore, our upper bound on the complexity of these arrangements is tight for a segment light source. This perhaps explains why our bounds on the complexity of the umbra are not tight. Notice, however, that we do have tight bounds for small n ($n = O(k)$) and, in the case of segment light source, for small k ($k = O(1)$).

7.7 Bonus : une borne inférieure sur l'arrangement des ombres

Pour prouver la borne $O(n^3k^3)$ du théorème 51 sur la complexité de l'ombre, on a en fait borné la complexité totale de l'arrangement B des arcs de limites d'ombres. Les théorèmes 60 et 61 présentent des bornes inférieures sur la taille de l'ombre ; j'ajoute ici une borne inférieure sur la complexité de l'arrangement.

Lemme 62. *L'arrangement B projeté par une source lumineuse polygonale en présence de k polytopes disjoints dans une scène de complexité totale n peut avoir une complexité $\Omega(n^3)$.*

Démonstration. On considère d'abord une source lumineuse ponctuelle et un polygone à n côtés. Ils créent sur le plan des ombres un polygone d'ombre P_1. On choisit maintenant une seconde source ponctuelle et un second polygone à n côtés de façon que le polygone d'ombre P_2 qu'ils engendrent coupe P_1 en $2n$ points. Si on remplace les deux sources ponctuelles par de très petites sources lumineuses polygonales à n côtés, chaque sommet des sources engendre dans B un polygone très proche de celui qui était engendré par la source ponctuelle. Les polygones d'ombre de départ P_1 et P_2 avaient $2n$ points d'intersection. Chacun est maintenant remplacé par n^2 points d'intersection entre les n polygones qui remplacent P_1 et ceux qui remplacent P_2, ce qui fait $\Omega(n^3)$ points d'intersection dans l'arrangement B. Ceci montre en particulier qu'on ne peut pas espérer améliorer l'exposant de n dans la borne supérieure $O(n^3k^3)$ sur la complexité de B. □

Chapter 8

Predicates for line transversals to lines and line segments in three-dimensional space

Ce travail a été réalisé avec O. Devillers, S. Lazard et S. Petitjean.

Given 4 line segments determined by their endpoints in 3D, we wonder if there are zero, one, two or an infinite number of lines stabbing these 4 line segments. We propose an algebraic formulation of degree 12 for that predicate.

8.1 Introduction

The computation of visibility between objects in three-dimensional space is one of the main problems of computer graphics, useful for hidden parts elimination, umbra and penumbra computations, radiosity methods... [54].

The visibility complex of a set of polyhedra can be computed by sweeping the space around each edge of the scene [28] in $O(n^2k^2)$ time for a scene of k convex polyhedra of total complexity n. In this sweeping algorithm an event corresponds to a line stabbing four edges of the scene. As any geometric algorithm, visibility computation relies on basic geometric tests called predicates involving a small number of input objects; the answer of a predicate is discrete, typically it is a sign $+/-/0$ that can be interpreted by something like inside/outside/on boundary or above/below/same height... Going from the geometric formulation of a predicate (is an object above another?) to an actual implementation consists in finding one (or several) polynomial(s) in the parameters describing the objects such that the sign (or a combination of the signs) of the polynomial(s) gives the answer to the geometric question.

For a given geometric problem, what a good design is for a predicate remains a difficult question. Between several polynomial formulations, the set of polynomials with lowest maximal degree is often preferred since the degree of a polynomial is closely related to the precision of the computation needed to evaluate exactly this polynomial, and such exact evaluation may be needed to solve robustness issues using the exact computation paradigm [111].

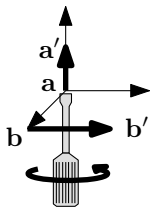


Figure 8.1: Relative orientation of two lines.

For our visibility problems, the predicates needed include: — the existence of a stabbing line to four edges (or more edges to take degeneracies into account) and — the comparison of such stabbing lines in the sweep order.

A first attempt to design these predicates using Plücker coordinates yields to a degree 36 predicate (in the Cartesian coordinates of the endpoints of the segments) for the existence of a line stabbing four line segments and of a degree 144 predicate for comparing stabbing lines [60].

In this paper, we investigate various predicates dealing with lines and line segments in three-dimensional space, we mainly use a direct formulation which gives better results than Plücker coordinates, more precisely for the two predicates mentioned above, we reduce the degrees from 36 and 144 down to 12 and 36.

If the involved polynomial of highest degree is irreducible, we can argue that we reach a kind of optimality, this is indeed the case for the degree 12 predicate while in the other case the minimal degree must be at least 24 (and is probably really 24).

8.2 Notation and preliminaries

We work in a projective setting so that two parallel lines intersect (at infinity). Let $\mathbb{P}^n$ denote the real projective space of dimension n. Lines in 3D are defined by two points in $\mathbb{P}^3$. More formally, a line $\overline{A}$ through the points $\mathbf{a} = (\mathbf{a}_x : \mathbf{a}_y : \mathbf{a}_z : \mathbf{a}_\omega)^T$ and $\mathbf{a}' = (\mathbf{a}'_x : \mathbf{a}'_y : \mathbf{a}'_z : \mathbf{a}'_\omega)^T$ in $\mathbb{P}^3$ is the set $\{\alpha\mathbf{a} + \alpha'\mathbf{a}'; (\alpha, \alpha') \in \mathbb{P}^1\}$; a point on $\overline{A}$ is thus parameterized by (α, α') in $\mathbb{P}^1$.

The line $\overline{A}$ can be split in two parts cutting it at points $\mathbf{a}$ and $\mathbf{a}'$. Since the projective space cannot be oriented, to precise these two sets we need to restrict the representative coordinates of $\mathbf{a}$ and $\mathbf{a}'$; namely we will assume in the sequel that all fourth coordinates ω are non negative. These two parts will be denoted by $A = \{\alpha\mathbf{a} + \alpha'\mathbf{a}'; \alpha\alpha' \geqslant 0\}$ and $\tilde{A} = \{\alpha\mathbf{a} + \alpha'\mathbf{a}'; \alpha\alpha' \leqslant 0\}$; if neither $\mathbf{a}$ nor $\mathbf{a}'$ is at infinity (*i.e.* $\mathbf{a}_\omega$ and $\mathbf{a}'_\omega$ are non-zero), A is the finite segment and $\tilde{A}$ the "segment" going through infinity.

If the points $\mathbf{a}$ and $\mathbf{a}'$ are finite we also like to orient the line $\overline{A}$ from $\mathbf{a}$ to $\mathbf{a}'$ that is, inside A, we consider decreasing $\frac{\alpha}{\alpha'}$. Given two oriented lines $\overline{A}$ and $\overline{B}$ we say that the orientation of $\overline{B}$ is positive with respect to $\overline{A}$ if the determinant $|\mathbf{a}\ \mathbf{b}\ \mathbf{a}'\ \mathbf{b}'| > 0$ (with the hypotheses that the fourth homogeneous coordinates are non negative). Intuitively, this means that, if the orientation of $\overline{B}$ is positive with respect to $\overline{A}$, a screwdriver oriented along $\overline{A}$ and turning according to $\overline{B}$ will turn in the positive direction along $\overline{A}$, that is it will screw (see Figure 8.1).

8.3 A direct approach to transversals to lines and line segments

8.3.1 Algebraic formulation

For any pair of parameters (α, α') and (β, β') in $\mathbb{P}^1$, there is a line through $\alpha\mathbf{a}+\alpha'\mathbf{a}'$ and $\beta\mathbf{b}+\beta'\mathbf{b}'$ that also intersects a line $\overline{C}$ if and only if the four points $\alpha\mathbf{a}+\alpha'\mathbf{a}'$, $\beta\mathbf{b}+\beta'\mathbf{b}'$, $\mathbf{c}$ and $\mathbf{c}'$ are coplanar (note that such a line transversal is not unique if $\alpha\mathbf{a}+\alpha'\mathbf{a}' = \beta\mathbf{b}+\beta'\mathbf{b}'$). This condition can be expressed as:

$$\begin{vmatrix} \alpha\mathbf{a}_x + \alpha'\mathbf{a}'_x & \beta\mathbf{b}_x + \beta'\mathbf{b}'_x & \mathbf{c}_x & \mathbf{c}'_x \\ \alpha\mathbf{a}_y + \alpha'\mathbf{a}'_y & \beta\mathbf{b}_y + \beta'\mathbf{b}'_y & \mathbf{c}_y & \mathbf{c}'_y \\ \alpha\mathbf{a}_z + \alpha'\mathbf{a}'_z & \beta\mathbf{b}_z + \beta'\mathbf{b}'_z & \mathbf{c}_z & \mathbf{c}'_z \\ \alpha\mathbf{a}_\omega + \alpha'\mathbf{a}'_\omega & \beta\mathbf{b}_\omega + \beta'\mathbf{b}'_\omega & \mathbf{c}_\omega & \mathbf{c}'_\omega \end{vmatrix} = 0.$$

This can be rewritten with a suitable notation as

$$\begin{aligned} 0 &= |\alpha\mathbf{a} + \alpha'\mathbf{a}'\ \beta\mathbf{b} + \beta'\mathbf{b}'\ \mathbf{c}\ \mathbf{c}'| \\ &= \alpha\beta\,|\mathbf{a}\ \mathbf{b}\ \mathbf{c}\ \mathbf{c}'| + \alpha\beta'\,|\mathbf{a}\ \mathbf{b}'\ \mathbf{c}\ \mathbf{c}'| + \alpha'\beta\,|\mathbf{a}'\ \mathbf{b}\ \mathbf{c}\ \mathbf{c}'| + \alpha'\beta'\,|\mathbf{a}'\ \mathbf{b}'\ \mathbf{c}\ \mathbf{c}'|\,. \end{aligned} \tag{3}$$

Similarly, we obtain that there exists a line through $\alpha\mathbf{a}+\alpha'\mathbf{a}'$ and $\beta\mathbf{b}+\beta'\mathbf{b}'$ that stabs $\overline{D}$ if and only if:

$$0 = \alpha\beta\,|\mathbf{a}\ \mathbf{b}\ \mathbf{d}\ \mathbf{d}'| + \alpha\beta'\,|\mathbf{a}\ \mathbf{b}'\ \mathbf{d}\ \mathbf{d}'| + \alpha'\beta\,|\mathbf{a}'\ \mathbf{b}\ \mathbf{d}\ \mathbf{d}'| + \alpha'\beta'\,|\mathbf{a}'\ \mathbf{b}'\ \mathbf{d}\ \mathbf{d}'|\,. \tag{4}$$

We have the following lemma.

Lemma 63. *There exists a line through $\alpha\mathbf{a}+\alpha'\mathbf{a}'$ and $\beta\mathbf{b}+\beta'\mathbf{b}'$ that intersects lines $\overline{C}$ and $\overline{D}$ if and only if $((\alpha,\alpha'),(\beta,\beta'))$ is solution of Equations (3) and (4). Furthermore, such a solution corresponds to a unique line transversal unless $\alpha\mathbf{a}+\alpha'\mathbf{a}' = \beta\mathbf{b}+\beta'\mathbf{b}'$ and that point is either coplanar with lines $\overline{C}$ and $\overline{D}$ or on $\overline{C}$ or $\overline{D}$.*

Proof. If a line through $\alpha\mathbf{a}+\alpha'\mathbf{a}'$ and $\beta\mathbf{b}+\beta'\mathbf{b}'$ intersects lines $\overline{C}$ and $\overline{D}$, the above discussion implies that $((\alpha,\alpha'),(\beta,\beta'))$ is solution of Equations (3-4). Conversely, suppose that (α,β) is solution of Equations (3-4). If the two points $\alpha\mathbf{a}+\alpha'\mathbf{a}'$ and $\beta\mathbf{b}+\beta'\mathbf{b}'$ are distinct, then the line through them is unique and Equations (3-4) imply that it intersects $\overline{C}$ and $\overline{D}$. On the other hand, if $\alpha\mathbf{a}+\alpha'\mathbf{a}' = \beta\mathbf{b}+\beta'\mathbf{b}'$, Equations (3-4) imply that there exists a line through this point and line $\overline{C}$ and a line through this point and line $\overline{D}$, but these two lines may be distinct; it is however straightforward that there exists a line through $\alpha\mathbf{a}+\alpha'\mathbf{a}' = \beta\mathbf{b}+\beta'\mathbf{b}'$ that intersects lines $\overline{C}$ and $\overline{D}$ because a point and two lines always admit a line transversal in $\mathbb{P}^3$. Indeed, if the point is on line $\overline{C}$, any line through this point and $\overline{D}$ is a transversal and there are infinitely many transversals. Otherwise, the point and $\overline{C}$ define a plane. If $\overline{D}$ is in that plane or contains the point, any transversal to the point and $\overline{C}$ also intersects $\overline{D}$ and there are infinitely many transversals. Otherwise $\overline{D}$ intersects the plane in a single point which, together with $\alpha\mathbf{a}+\alpha'\mathbf{a}'$, defines a unique transversal. □

Lemma 63 directly implies the following corollary.

Corollary 64. *There exists a line that intersects segments A and B and lines $\overline{C}$ and $\overline{D}$ if and only if Equations (3)-(4) admit a solution $((\alpha,\alpha'),(\beta,\beta'))$ such that $\alpha\alpha'$ and $\beta\beta'$ are non-negative.*

We now show how Equations (3) and (4) can be transformed into degree-2 equations in (α,α') or (β,β') and into the equation of a line.

By considering Equations (3) and (4) as linear equations in variable (β,β') with parameter (α,α') the system admits (at least) a solution only if

$$\begin{aligned}
0 &= \begin{vmatrix} \alpha\,|\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| + \alpha'\,|\mathbf{a}'\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & \alpha\,|\mathbf{a}\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| + \alpha'\,|\mathbf{a}'\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ \alpha\,|\mathbf{a}\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| + \alpha'\,|\mathbf{a}'\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & \alpha\,|\mathbf{a}\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| + \alpha'\,|\mathbf{a}'\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} \\
&= \alpha^2 \begin{vmatrix} |\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} + \alpha\alpha' \left(\begin{vmatrix} |\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} + \right. \\
&\qquad \left. \begin{vmatrix} |\mathbf{a}'\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}'\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} \right) + \alpha'^2 \begin{vmatrix} |\mathbf{a}'\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}'\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} \\
&= F_\alpha\alpha^2 + G_\alpha\alpha\alpha' + H_\alpha\alpha'^2. \qquad (5)
\end{aligned}$$

Similarly considering (α,α') as the variable, we get

$$\begin{aligned}
0 &= \begin{vmatrix} \beta\,|\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| + \beta'\,|\mathbf{a}\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & \beta\,|\mathbf{a}\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| + \beta'\,|\mathbf{a}\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \\ \beta\,|\mathbf{a}'\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| + \beta'\,|\mathbf{a}'\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & \beta\,|\mathbf{a}'\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| + \beta'\,|\mathbf{a}'\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} \\
&= \beta^2 \begin{vmatrix} |\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}'\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} + \beta\beta' \left(\begin{vmatrix} |\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}'\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} + \right. \\
&\qquad \left. \begin{vmatrix} |\mathbf{a}\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}'\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} \right) + \beta'^2 \begin{vmatrix} |\mathbf{a}\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}'\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} \\
&= F_\beta\beta^2 + G_\beta\beta\beta' + H_\beta\beta'^2. \qquad (6)
\end{aligned}$$

Finally by eliminating $\alpha\beta$ between Equations (3) and (4) we get

$$\begin{aligned}
0 &= \alpha\beta' \begin{vmatrix} |\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} + \alpha'\beta \begin{vmatrix} |\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}'\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} \\
&\qquad + \alpha'\beta' \begin{vmatrix} |\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}'\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix} \\
&= F_\alpha\alpha\beta' + F_\beta\alpha'\beta + H_{\alpha\beta}\alpha'\beta'. \qquad (7)
\end{aligned}$$

Note that the coefficients F_α and F_β in Equation (7) are the same as in Equations (5) and (6). Note also that, if one of these coefficients is non-zero, Equation (7) is the equation of a line whose equation in affine space is $F_\alpha\frac{\alpha}{\alpha'} + F_\beta\frac{\beta}{\beta'} + H_{\alpha\beta} = 0$. Finally, note that $H_{\alpha\beta} = \frac{G_\alpha+G_\beta}{2}$; indeed, by exchanging the diagonals in the two determinants defining G_α, we get that

$$G_\alpha = H_{\alpha\beta} + H' \quad \text{where} \quad H' = \begin{vmatrix} |\mathbf{a}'\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}'\,\mathbf{b}\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \end{vmatrix}$$

and similarly that $G_\beta = H_{\alpha\beta} - H'$.

Let $\Delta = G_\alpha^2 - 4F_\alpha H_\alpha$ be the discriminant of Equation (5). Note that one can easily prove that Δ is also the discriminant $G_\beta^2 - 4F_\beta H_\beta$ of Equation (6) by developing the two expressions.

Remark 65. *The discriminant Δ also appears naturally as the determinant of the matrix whose coefficients are the pairwise side operators of the lines $\overline{A}$, $\overline{B}$, $\overline{C}$ and $\overline{D}$:*

$$\begin{vmatrix} 0 & |\mathbf{a}\,\mathbf{a}'\,\mathbf{b}\,\mathbf{b}'| & |\mathbf{a}\,\mathbf{a}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{a}\,\mathbf{a}'\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}\,\mathbf{a}'\,\mathbf{b}\,\mathbf{b}'| & 0 & |\mathbf{b}\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{b}\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}\,\mathbf{a}'\,\mathbf{c}\,\mathbf{c}'| & |\mathbf{b}\,\mathbf{b}'\,\mathbf{c}\,\mathbf{c}'| & 0 & |\mathbf{c}\,\mathbf{c}'\,\mathbf{d}\,\mathbf{d}'| \\ |\mathbf{a}\,\mathbf{a}'\,\mathbf{d}\,\mathbf{d}'| & |\mathbf{b}\,\mathbf{b}'\,\mathbf{d}\,\mathbf{d}'| & |\mathbf{c}\,\mathbf{c}'\,\mathbf{d}\,\mathbf{d}'| & 0 \end{vmatrix} = \Delta.$$

With Δ written this way, it is easy to see that exchanging the role of any two lines among $\overline{A}$, $\overline{B}$, $\overline{C}$, $\overline{D}$ does not change Δ.

We now prove the following lemmas stating some equivalences between these equations.

Lemma 66. *If $F_\alpha F_\beta \neq 0$ then Systems (3)-(4), (5)-(7), and (6)-(7) are equivalent.*

Proof. We have already seen that Equations (3) and (4) imply Equations (5), (6), and (7). Assume now that $F_\alpha F_\beta \neq 0$ and Equations (5) and (7) are satisfied. Since $F_\alpha \neq 0$, any solution of Equation (5) is such that $\alpha' \neq 0$. One can easily verify that equations (3), (5) and (7) satisfy the relation

$$\alpha' F_\beta.(3) + \beta' \,|\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'|\,.(5) = (\alpha\,|\mathbf{a}\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'| + \alpha'\,|\mathbf{a}'\,\mathbf{b}\,\mathbf{c}\,\mathbf{c}'|).(7).$$

Since $\alpha' F_\beta \neq 0$, Equation (3) must be satisfied. The proof that Equation (4) is satisfied is similar, as well as the equivalence between Systems (3)-(4) and (6)-(7). □

Remark 67. *Note that exchanging the roles of $\mathbf{a}$ and $\mathbf{a}'$, which is equivalent to exchanging α and α', yields very similar equations; precisely, it exchanges F_α and H_α, leaves G_α and Δ unchanged, changes the sign of F_β, G_β and H_β (which means that Eq. (6) is unchanged), and $H_{\alpha\beta}$ takes the new value of $\frac{G_\alpha + G_\beta}{2}$.*

Remark 68. *The above remark implies that, if F_α or H_α is non-zero and F_β or H_β is non-zero, then after exchanging the role of $\mathbf{a}$ and $\mathbf{a}'$ if $F_\alpha = 0$ and exchanging the role of b and b' if $F_\beta = 0$, the new equations are such that $F_\alpha F_\beta \neq 0$ and Lemma 66 applies.*

Lemma 69. *If Equation (5) identically vanishes then System (3)-(4) has infinitely many solutions. If $F_\alpha = H_\alpha = 0$ and $G_\alpha \neq 0$, then System (3)-(4) has two solutions.*

Proof. If Eq. (5) identically vanishes then System (3)-(4) admits infinitely many solutions since for any solution of Eq. (5), that is for any (α, α'), there exists a (β, β') such that $((\alpha, \alpha'), (\beta, \beta'))$ is solution of the system or (α, α') cancel the coefficients of β in Equations (3) and (4); in the latter case, $((\alpha, \alpha'), (1, 0))$ is solution of the system (see [40, Th. 3, Ch. 3, §6] for a reference on resultants).

If $F_\alpha = H_\alpha = 0$ and $G_\alpha \neq 0$, then the solutions of Eq. (5) are $(1, 0)$ and $(0, 1)$. Substituting these solutions in Eq. (3)-(4) gives the solutions. When $(\alpha, \alpha') = (1, 0)$, if

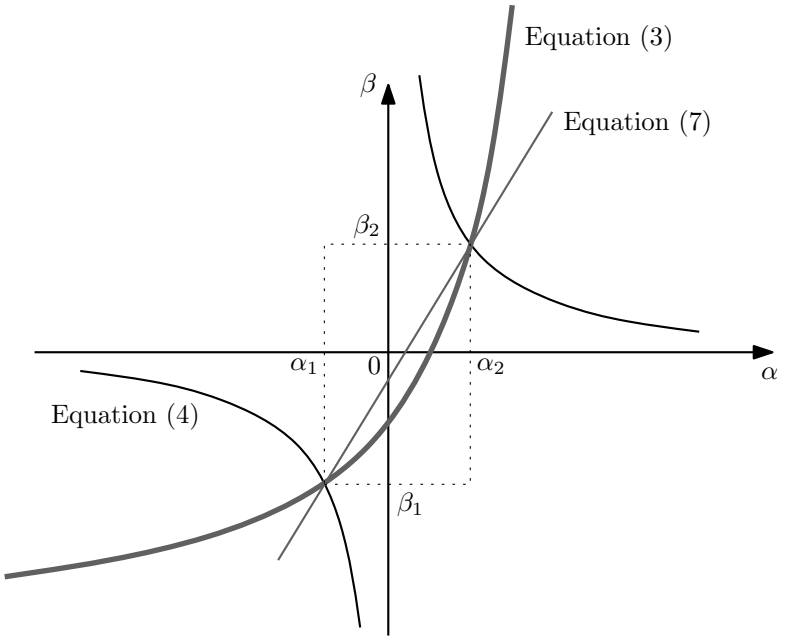


Figure 8.2: Geometric interpretation in $\alpha - \beta$-space.

Equation (3) does not identically vanish, it yields $(\beta, \beta') = (|\mathbf{a}\ \mathbf{b}'\ \mathbf{c}\ \mathbf{c}'|, -|\mathbf{a}\ \mathbf{b}\ \mathbf{c}\ \mathbf{c}'|)$ and otherwise Equation (4) yields $(\beta, \beta') = (|\mathbf{a}\ \mathbf{b}'\ \mathbf{d}\ \mathbf{d}'|, -|\mathbf{a}\ \mathbf{b}\ \mathbf{d}\ \mathbf{d}'|)$ (which is necessarily well defined in $\mathbb{P}^1$ since, if all four determinants are zero, then $G_\alpha = 0$). Similarly, for $(\alpha, \alpha') = (0, 1)$, $(\beta, \beta') = (|\mathbf{a}'\ \mathbf{b}'\ \mathbf{c}\ \mathbf{c}'|, -|\mathbf{a}'\ \mathbf{b}\ \mathbf{c}\ \mathbf{c}'|)$ or $(|\mathbf{a}\ \mathbf{b}'\ \mathbf{c}\ \mathbf{c}'|, -|\mathbf{a}\ \mathbf{b}\ \mathbf{c}\ \mathbf{c}'|)$. □

We get the following lemma by exchanging the role of lines $\overline{A}$ and $\overline{B}$ in Lemma 69.

Lemma 70. *If Equation (6) identically vanishes then System (3)-(4) has infinitely many solutions. If $F_\beta = H_\beta = 0$ and $G_\beta \neq 0$, then System (3)-(4) has two solutions.*

8.3.2 Geometric interpretation

We give here a geometric interpretation of the above algebraic formulation. We give this interpretation in the affine parameter space, that is in the $\alpha\beta$-plane with $\alpha' = \beta' = 1$. We assume that the four lines or segments are defined by Euclidean points given in homogeneous form with their fourth coordinates equal to one. Recall that a point $\alpha\mathbf{a} + \mathbf{a}'$ on line $\overline{A}$ is on segment A for $\alpha \in [0, +\infty]$ and similarly for points on $\overline{B}$. We first consider the case of four lines or segments and then the case where a fifth line or segment is added.

8.3.2.1 Four lines or line segments

Refer to Figure 8.2. A point in the $\alpha\beta$-space corresponds to the 3D line going through points $\alpha\mathbf{a} + \mathbf{a}'$ and $\beta\mathbf{b} + \mathbf{b}'$ (if these two points are distinct). Equation (3) defines in the $\alpha\beta$-space the conic corresponding to the lines passing through $\overline{A}$, $\overline{B}$, and $\overline{C}$. Similarly, Equation (4) defines the conic corresponding to the lines passing through $\overline{A}$, $\overline{B}$, and $\overline{D}$.

These two conics and the line defined in $\alpha\beta$-space by Equation (7) are in the same pencil and intersect in the same set of points. Hence, if the two conics intersect, they intersect in only two points, say, (α_1, β_1) and (α_2, β_2), which lie on the line defined by Equation (7). These two points correspond to 3D lines that intersect the four lines $\overline{A}$, $\overline{B}$, $\overline{C}$ and $\overline{D}$. Furthermore α_1, α_2 (resp. β_1, β_2) are the two solutions of Equation (5) (resp. (6)). If $\alpha_i \in [0, +\infty]$ then the line corresponding to (α_i, β_i) does not only intersect the line $\overline{A}$ but also the line segment A; similarly if $\beta_i \in [0, +\infty]$, the line intersects segment B.

If we embed the $\alpha\beta$-space in 3D by adding a third dimension ζ, the conic (3) can be viewed as the projection of the intersection of the quadric $\zeta = \alpha\beta$ with the plane of equation

$$\zeta \left|\mathbf{a}\ \mathbf{b}\ \mathbf{c}\ \mathbf{c}'\right| + \alpha \left|\mathbf{a}\ \mathbf{b}'\ \mathbf{c}\ \mathbf{c}'\right| + \beta \left|\mathbf{a}'\ \mathbf{b}\ \mathbf{c}\ \mathbf{c}'\right| + \left|\mathbf{a}'\ \mathbf{b}'\ \mathbf{c}\ \mathbf{c}'\right| = 0 \tag{8}$$

obtained by substituting ζ to $\alpha\beta$ in Equation (3).

The line through $\alpha\mathbf{a} + \mathbf{a}'$ and $\beta\mathbf{b} + \mathbf{b}'$ stabs $\overline{C}$ if the point $(\alpha, \beta, \alpha\beta)$ in $\alpha - \beta - \zeta$-space belongs to plane (8). Locating this point with respect to the plane gives the orientation of the line; assuming $|\mathbf{a}\ \mathbf{b}\ \mathbf{c}\ \mathbf{c}'| > 0$, the line turns positively around $\overline{C}$ if the point is below and it turns negatively around $\overline{C}$ if the point is above.

8.3.2.2 Adding a fifth line

If we add a fifth line $\overline{E}$, then we get another conic in $\alpha\beta$-space:

$$\alpha\beta \left|\mathbf{a}\ \mathbf{b}\ \mathbf{e}\ \mathbf{e}'\right| + \alpha \left|\mathbf{a}\ \mathbf{b}'\ \mathbf{e}\ \mathbf{e}'\right| + \beta \left|\mathbf{a}'\ \mathbf{b}\ \mathbf{e}\ \mathbf{e}'\right| + \left|\mathbf{a}'\ \mathbf{b}'\ \mathbf{e}\ \mathbf{e}'\right| = 0. \tag{9}$$

This conic can belong to the same pencil as Equations (3), (4) and (7). Then the two stabbing lines of $\overline{A}$, $\overline{B}$, $\overline{C}$ and $\overline{D}$ stab also $\overline{E}$. Otherwise locating points (α_i, β_i) with respect to this conic allows to determine if $\overline{E}$ stabs or turn positively or negatively around the stabbing lines.

The three Equations (3), (4) and (9) define three planes in $\alpha\beta\zeta$-space whose intersection is a point M if the planes are affinely independent.

The five lines admit a common stabbing line if the point M lies on the quadric $\zeta = \alpha\beta$ which can be verified by solving the system of Equations (3), (4) and (9) and substituting in the quadric.

8.4 Number of transversals to lines and segments

8.4.1 Number of transversals to four lines

Lemma 71. *Given four lines, there is a predicate of degree* 12 *in the Cartesian coordinates of their defining points to determine whether they admit 0, 1, 2, or infinitely many transversals.*

Proof. We first consider the number of solutions to the system of Equations (3) and (4); let S denote that system. By Lemmas 69 and 70, if Equation (5) or (6) identically vanishes then S has infinitely many solutions; otherwise, if $F_\alpha = H_\alpha = 0$ or $F_\beta = H_\beta = 0$, then S has two solutions. Otherwise, at least one of F_α or H_α is non-zero and at least one of F_β or

H_β is non-zero. By Remark 68, we can then apply Lemma 66 after exchanging, if needed, the roles of α and α' and of β and β'. The number, zero, one, or two, of solutions to S is then given by the sign (negative, null, or positive) of the discriminant Δ of Equation (5). Moreover, by Remark 67, this discriminant is not modified by the exchange of α and α' and of β and β', thus the number of transversals is given by the sign of the discriminant of Equation (5) without any exchange.

Now, by Lemma 63, the number of solutions of S is the number of transversals to the four lines except if lines $\overline{A}$ and $\overline{B}$ intersect and either lines $\overline{C}$ and $\overline{D}$ are coplanar with the point of intersection of $\overline{A}$ and $\overline{B}$ or the point of intersection of $\overline{A}$ and $\overline{B}$ belongs to line $\overline{C}$ or $\overline{D}$ in which cases there are infinitely many line transversals. We notice that in both of these cases, $\Delta = 0$; indeed, we first note that the discriminant Δ does not change if we exchange the role of any two lines among $\overline{A}$, $\overline{B}$, $\overline{C}$ and $\overline{D}$, as shown in Remark 65. If the four lines have infinitely many common transversals, there must be at least one of these four lines, say $\overline{A}$ without loss of generality, such that every points on $\overline{A}$ belongs to common transversals. This implies that Equation (5) has infinitely many solutions and thus that its discriminant Δ is 0.

To sum things up, if $\Delta > 0$, there are two transversals, if $\Delta < 0$, there is no transversal and if $\Delta = 0$, there may be either a unique transversal or infinitely many transversals. Precisely, if $\Delta = 0$, there are infinitely many transversals if and only if lines $\overline{A}$ and $\overline{B}$ intersect at a point that is coplanar with lines $\overline{C}$ and $\overline{D}$ or at a point that lies on $\overline{C}$ or $\overline{D}$.

The discriminant Δ is a degree 16 irreducible polynomial in the homogeneous coordinates of the points defining the lines. If the points are given with affine coordinates, this degree drops to 12 (and this degree 12 polynomial is irreducible as well). Furthermore, if $\Delta = 0$, testing whether there are infinitely many transversals can trivially be done with a predicate of degree lower than 12. Hence, determining the number of transversals to four lines can be done with a predicate of degree 12.

□

8.4.2 Number of transversals to four line segments

We study here the degree of predicates for determining the number of transversals to segments. We study first the case of 1 segment and 3 lines and the case of 2 segments and two lines; the case of four segments will follow.

Lemma 72. *Given a line segment and three lines, there is a predicate of degree* 12 *in the Cartesian coordinates of their defining points to determine whether they admit 0, 1, 2, or infinitely many transversals.*

Proof. If Lemma 71 concludes that there is no stabbing line, then the result also applies for the line segment A. If it concludes that there are infinitely many stabbing lines, then the result also applies for the line segment A unless this segment is a single point (i.e. $\mathbf{a} = \mathbf{a}'$).

In the case of one or two line transversals to $\overline{A}$, $\overline{B}$, $\overline{C}$, and $\overline{D}$, it remains to determine whether they intersect the line segment A. This is equivalent to determine the sign of the solutions of Equation (5).

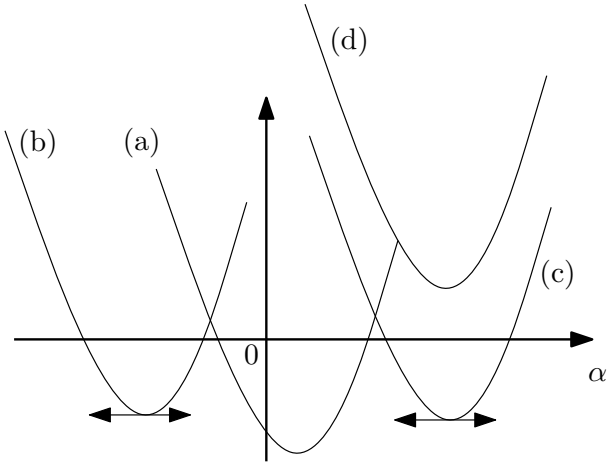


Figure 8.3: Graph of f depending on the signs of $f(0)$, $f'(0)$, and Δ.

In fact it is more efficient to first attempt to locate the roots of Equation (5), and then only resolve Lemma 71 (i.e., evaluate the discriminant Δ) when the root location is not complete. Precisely, as we show below, we need to evaluate Δ only in the case where both roots of Equation (5), if they are real, are non-negative.

If $F_\alpha = 0$ or $H_\alpha = 0$, the signs of the roots of Equation (5) are obvious, so we can restrict ourselves to the case where $F_\alpha \neq 0$ and $H_\alpha \neq 0$. We further assume that $F_\alpha > 0$ (the case $F_\alpha < 0$ is similar). We define $f(x) = F_\alpha x^2 + G_\alpha x + H_\alpha$ after Equation (5). Figure 8.3 shows possible graphs of f. If $f(0) = H_\alpha < 0$, f has two real roots, one is positive and the other one is negative (case (a) on the figure). Otherwise, if $f'(0) = G_\alpha \geqslant 0$ (case (b)), the minimum of f happens for a non-positive value of x and f has no non-negative root (it has either two negative roots or no root). Otherwise we need to evaluate the sign of Δ to distinguish between case (c) where f has two positive real roots and case (d) where it has none.

□

Lemma 73. *Given two line segments and two lines, there is a predicate of degree* 12 *in the Cartesian coordinates of their defining points to determine whether they admit 0, 1, 2, or infinitely many transversals.*

Proof. Applying Lemma 72 once for A, $\overline{B}$, $\overline{C}$ and $\overline{D}$ and once for $\overline{A}$, B, $\overline{C}$ and $\overline{D}$ gives the solution in most cases. One case where this does not give the answer is when $\overline{A}$, $\overline{B}$, $\overline{C}$ and $\overline{D}$ have infinitely many transversals. Then A, B, $\overline{C}$ and $\overline{D}$ may have 0, 1, 2 or infinitely many transversals.

The other case that remains is when $\overline{A}$, $\overline{B}$, $\overline{C}$ and $\overline{D}$ have two common transversals, one intersects A and one intersects B. Indeed, we need to check whether the transversal that intersects A is the same as the one that intersects B. For the sake of simplicity, we omit the case where a transversal goes through the endpoint of one of the segments. The values of (α, α') that correspond to the two transversals are $(\alpha_1, 1)$ and $(\alpha_2, 1)$ with $\alpha_1 < 0 < \alpha_2$. Similarly the values of (β, β') are $(\beta_1, 1)$ and $(\beta_2, 1)$ with $\beta_1 < 0 < \beta_2$. The two transversals then either match α_1 with β_1 and α_2 with β_2 or α_1 with β_2 and α_2 with

β_1. Equation 7 enables us to distinguish the two cases. Indeed, it describes a line that contains the two points (α_i, β_j) that are solution. If its slope is positive, we are in the first case, and otherwise we must be in the second case.

If Lemma 72 detects no stabbing line for A, $\overline{B}$, $\overline{C}$ and $\overline{D}$ there is no stabbing line for A, B, $\overline{C}$ and $\overline{D}$.

□

Lemma 74. *Given four line segments, there is a predicate of degree* 12 *in the Cartesian coordinates of their endpoints to determine whether they admit 0, 1, 2, 3, 4, or infinitely many transversals.*

Proof. Evaluate Lemma 72 for $(A, \overline{B}, \overline{C}, \overline{D})$, $(\overline{A}, B, \overline{C}, \overline{D})$, $(\overline{A}, \overline{B}, C, \overline{D})$ and $(\overline{A}, \overline{B}, \overline{C}, D)$. If we find two stabbing lines in all cases then there are two stabbing lines for (A, B, C, D). If we find no stabbing line in at least one case, then there is no stabbing line for (A, B, C, D). Otherwise assume without loss of generality that $(A, \overline{B}, \overline{C}, \overline{D})$ has only one stabbing line, then we will use Lemma 73 for $(A, B, \overline{C}, \overline{D})$, $(A, \overline{B}, C, \overline{D})$ and $(A, \overline{B}, \overline{C}, D)$ and the final answer is the minimum of the three answers. As for Lemma 71 the predicate has degree 12 (16 with homogeneous coordinates). □

8.4.3 More lines and segments

Lemma 75. *Given five lines, there is a predicate of degree* 15 *in the Cartesian coordinates of their defining points to determine whether they admit 0, 1, 2, or infinitely many transversals.*

Proof. We first write the Plücker coordinates of the five lines in a 6x5 matrix. Two lines intersect if and only if their side operator is 0. This means that a line T is a common transversal to all five lines if and only if the vector of its Plücker coordinates (properly reordered) is in the kernel of this matrix. The five lines might have linearly dependent Plücker coordinates, in which case we can forget about one of them, because intersecting the other four implies intersecting this one as well. If the lines are linearly independent (which is the generic case, even for five lines that admit a common transversal), the 5x5 minors of the matrix define a generator of the dimension 1 kernel of the matrix. All that remains to be seen is whether this vector corresponds to a line, that is whether its self side operator is 0. This gives a degree 15 irreducible polynomial in the Cartesian coordinates, or a degree 20 irreducible polynomial in the homogeneous coordinates. □

Remark 76. *Notice that in the case where the five lines have linearly independent Plücker coordinates and they have a common transversal, we managed to compute the Plücker coordinates of this transversal as a rational fraction in the input. So if all the input has rational coordinates and we can determine that 4 given lines have 2 transversals and they have irrational coordinates, then the only way a fifth line can intersect one of these transversals is if its Plücker coordinates are a linear combination of the Plücker coordinates of the 4 initial lines, and then it intersects both transversals.*

Lemma 77. *Given five line segments, there is a predicate of degree* 15 *in the Cartesian coordinates of their endpoints to determine whether they admit 0, 1, 2, 3, 4, 5, or infinitely many transversals.*

Proof. We start by presenting an alternative approach to the previous lemma.

If Equation (9) is a linear combination of Equations (3) and (4), that is all 4 of the 3×3 determinants constructed with the coefficients are zero, then one of the lines, say $\overline{E}$, is such that any line stabbing $\overline{A}$, $\overline{B}$, $\overline{C}$, $\overline{D}$ also stabs $\overline{E}$. Thus the answer is obtained by applying Lemma 71 to lines $\overline{A}$, $\overline{B}$, $\overline{C}$ and $\overline{D}$.

In $\alpha\beta\zeta$-space, the above three equations can be solved to compute the coordinates of the intersection point M as a quotient of 3×3 determinants built on the coefficients of these equations. Then, we substitute the coordinates of M in the quadric equation $\zeta = \alpha\beta$, if the equation is verified, there is one stabbing line, otherwise there is no stabbing line.

Using Cartesian coordinates, the degree of this predicate is 18. However, there is only one polynomial of degree 18 involved in this predicate (the next polynomial of highest degree has degree 15) and it is not irreducible. We first notice that the side operator of lines $\overline{A}$ and $\overline{B}$ is an irreducible factor of degree 3 of this polynomial. This is quite natural because, if lines $\overline{A}$ and $\overline{B}$ intersect, chosing (α, α') and (β, β') so that the two parametric points are the same solves all equations like 3, 4 and 9 whereas there may not be any common transversal. The remaining degree 15 factor is exactly the polynomial computed in the previous approach. Notice that we cannot prove that the degree 18 polynomial is really the product of the degree 15 and the degree 3 polynomials. Indeed, their developed forms have too many terms for a software like Maple running on a real computer to compute all of them. However, they appear to give the same value when evaluated on a large set of random input, which makes us pretty confident.

Coming back to the proof of Lemma 77, we first note that by proceeding as before, and by verifying that the coordinate (α, α') of M satisfies $\alpha\alpha' \geqslant 0$, we can determine, with a predicate of degree 15, the number of line transversals to one segment and four lines.

Assuming we are in the case where the 5 lines supporting the segments have a single common transversal, we now evaluate this predicate for $(A, \overline{B}, \overline{C}, \overline{D}, \overline{E})$, $(\overline{A}, B, \overline{C}, \overline{D}, \overline{E})$, $(\overline{A}, \overline{B}, C, \overline{D}, \overline{E})$, $(\overline{A}, \overline{B}, \overline{C}, D, \overline{E})$ and $(\overline{A}, \overline{B}, \overline{C}, \overline{D}, E)$. If we find one stabbing line in all cases then there is exactly one stabbing line for (A, B, C, D, E). If we find no stabbing line in at least one case, then there is no stabbing line for (A, B, C, D, E). As for Lemma 75 the predicate has degree 15 if we never compute the degree 18 polynomial and only deduce its sign from the signs of the degree 3 and 15 polynomials.

Degenerate cases need a more complicated handling. When one line is "useless", meaning that intersecting it is a mere consequence of intersecting the others, we cannot simply ignore the segment as we were ignoring the line. Instead, we use an equation like Equation 3 where $\overline{B}$ is the "useless" line in order to translate the restriction on the values of (β, β') in terms of (α, α'). We can then ignore this segment and only remember that the possible values of (α, α') are less than the initial set with $\alpha\alpha' \geqslant 0$. □

Lemma 78. *Given six or more lines or line segments, there is a predicate of degree* 15 *in the Cartesian coordinates of their defining points to determine how many transversals they admit.*

Proof. If the input only includes lines, we can write all their Plücker coordinates in a matrix. If the rank is at most 3 then there are infinitely many common transversals. Otherwise we keep four independent lines and check that they have two transversals (not zero).

For each other line we check which of the two transversals it intersects using Lemma 75. The solution is then composed of those of the 2 transversals that were intersected by all the lines.

For segments, there are two cases. If the lines have finitely many transversals, combining the previous technique with Lemma 77 solves the problem. Otherwise the lines have infinitely many common tangents. This case is mostly handled in [29]. □

8.5 Transversals to four segments and a triangle

Lemma 79. *Let ℓ be a line transversal to four lines that admit finitely many transversals and let T be a triangle. There is a predicate of degree* 15 *in the Cartesian coordinates of the points defining the lines and the triangle to determine whether ℓ intersects (properly or not) T.*

Proof. This is similar to Lemma 75 but we need to know if a point $\alpha_i, \beta_i, \alpha_i\beta_i$ corresponding to a stabbing line is above or below the plane of Equation (9). This can be deduced by testing if M is above or below the quadric $\zeta = \alpha\beta$. As for Lemma 75 the predicate has degree 15. □

Given a line transversal ℓ to a set S of segments, a triangle T *occludes* ℓ if ℓ intersects T and if there exist two segments in S whose intersections with ℓ lie on opposite sides of T.

Lemma 80. *Let ℓ be a line transversal to four line segments that admit finitely many transversals and let T be a triangle. There is a predicate of degree* 15 *in the coordinates of the points defining the segments and the triangle to determine whether T occludes ℓ.*

Proof. This test is performed in two steps. First we check whether the transversal intersects the triangle as in Lemma 79. Then, assuming the transversal does intersect the triangle, for each of the four supporting lines we check where its intersection point with the transversal is with respect to the plane supporting the triangle. If all four intersection points are on the same side of the plane, the transversal is not occluded. If there are points on both sides of the plane, the transversal is occluded.

In order to check on which side of the plane supporting T the point of intersection of $\overline{A}$ and ℓ is, we compute the sign of the determinant $|T_1\ T_2\ T_3\ \alpha\mathbf{a} + \alpha'\mathbf{a}'|$ where T_i are the 3 vertices of T. This can be rewritten as:

$$\alpha|T_1\ T_2\ T_3\ \mathbf{a}| + \alpha'|T_1\ T_2\ T_3\ \mathbf{a}'|$$

The two coefficients of this polynomial in (α, α') are degree 3 polynomials in the input. Remember that (α, α') is a solution of Equation 5. Evaluating the sign of a polynomial at a root of an other polynomial amounts to sorting their respective roots, which means computing their resultant, that is a degree 12 polynomial in the input. The computation is exactly the same for the intersection points of ℓ with $\overline{B}$, $\overline{C}$ and $\overline{D}$.

□

8.6 Sorting transversals around a line

Lemma 81. *Given seven lines $\overline{A}$, $\overline{B}$, $\overline{C}$, $\overline{D}$, $\overline{B}^*$, $\overline{C}^*$, $\overline{D}^*$ with two stabbing lines S_1 and S_2 to $\overline{A}$, $\overline{B}$, $\overline{C}$, $\overline{D}$, and two stabbing lines S_3 and S_4 to $\overline{A}$, $\overline{B}^*$, $\overline{C}^*$, $\overline{D}^*$, there is a predicate of degree 36 in the Cartesian coordinates of the points defining the lines for ordering S_1, S_2, S_3 and S_4 around $\overline{A}$.*

Proof. Writing that $|\mathbf{a}\,\mathbf{a}'\ \beta\mathbf{b}+\beta'\mathbf{b}'\ \beta^*\mathbf{b}^*+\beta^{*\prime}\mathbf{b}^{*\prime}| = 0$ describes how β and β^* corresponding to the same sweeping plane correspond together. This equation is similar to Equation (3). Writing β in terms of β^* and substituting in Equation (6) we get β_1^* and β_2^* as solutions of a degree 2 equation with degree 12 coefficients that we have to compare with β_3^* and β_4^* which are solutions of a degree 2 equation with degree 6 coefficients. This yields a degree 36 predicate (degree 48 with homogeneous coordinates). □

Remark 82. *The degree 36 in the above predicate is likely not optimal. Indeed, after replacing the coordinates of the points by values of the type $xt + y$ where x and y are pseudo-random integer values, the resulting polynomial in t can be factored in $\mathbb{Q}[t]$ into one irreducible term of degree 24 and two squared terms of degree three (which are the side operator of $\overline{A}$ with $\overline{B}$ and $\overline{B}^*$, respectively). This implies that with high probability, the polynomial of degree 36 is not irreducible. Note that the polynomial of degree 36 cannot successfully be factored in Maple because it is too large.*

Remark 83. *Note also that if lines B and B^* coincide then it is straightforward to show that the degree of the predicate drops to 24 (which can be presumably factored, as above, into two polynomials with degrees 21 and 3).*

Part IV

Visibilité en présence de sphères

Chapter 9

Lower bounds on free lines and maximal free line segments tangent to balls

9.1 Introduction

Consider a set of objects in $\mathbb{R}^3$. A line in $\mathbb{R}^3$ is said to be free if it does not intersect the interior of any object[15]. A maximal free line segment is a (possibly infinite) segment that does not intersect the interior of any object and is not contained in any other segment satisfying the same property. We are interested here in the combinatorial complexity of the space of free lines and the set of maximal free line segments.

For scenes where the objects are k polytopes of total complexity n, this has been studied in [28] where the authors give a bound of $\Theta(n^2k^2)$ in the worst case on the number of isolated maximal free line segments tangent to 4 objects (which is closely related to the complexity of the space of maximal free line segments). When the objects are unit balls, an upper bound of $O(n^{3+\epsilon})$ for all $\epsilon > 0$ was proved in [4] on the complexity of the space of free lines.

Our contribution here is a worst-case lower bound of $\Omega(n^3)$ on the complexity of the space of free lines among balls (not of the same radius) and a worst-case lower bound of $\Omega(n^4)$ on the space of maximal free line segments among unit balls. Note that the second bound is trivially tight. Although the first one almost matches the upper-bound of $O(n^{3+\epsilon})$ from [4], the balls have different radii here instead of having the same radius as in [4], so both problems remain open for the worst case complexity of the space of free lines: in the presence of unit balls the complexity is $\Omega(n^2)$ and $O(n^{3+\epsilon})$ [44] while for balls of arbitrary radii it is $\Omega(n^3)$ and $O(n^4)$. These complexities are summarized in Table 9.1.

We will describe our examples using a Cartesian coordinate system (x, y, z). Denote by M_x, M_y and M_z the coordinates of a point M (or of the center of a ball M) in this coordinate system.

[15] We can assume here that all objects have a non-empty interior.

	Free lines		Maximal free line segments	
Unit balls	$\Omega(n^2)$	trivial	$\Omega(n^4)$	Theorem 89
	$O(n^{3+\varepsilon})$	[4]	$O(n^4)$	trivial
Any balls	$\Omega(n^3)$	Theorem 84	$\Omega(n^4)$	Theorem 89
	$O(n^4)$	trivial	$O(n^4)$	trivial

Table 9.1: Know complexities of free spaces.

9.2 $\Omega(n^3)$ free lines tangent to balls

Theorem 84. *The complexity of the space of free lines among n balls is $\Omega(n^3)$ in the worst case.*

We prove this by giving an example. For convenience, our example involves $3n+3$ balls instead of just n, which does not affect the result.

Consider a set of unit balls $\mathcal{B} = \{B_0 \dots B_n\}$ whose centers have coordinates $(3(i - n/2), 0, -1)$. Now add two sets of balls of radius ε: $\mathcal{A}^- = \{A_0^- \dots A_n^-\}$ and $\mathcal{A}^+ = \{A_0^+ \dots A_n^+\}$. The center of A_i^- has coordinates $(-3n, 3(i - n/2)\varepsilon, 0)$, and A_i^+ is its symmetric with respect to the $yz-$plane. Figure 9.1 shows this set of balls seen from $z = +\infty$, from $y = -\infty$ and from $x = -\infty$. We additionally assume that ε is small enough (more precisely we will see that small enough means $\varepsilon < \frac{1}{5400n^2}$).

We prove Theorem 84 by proving the following bound.

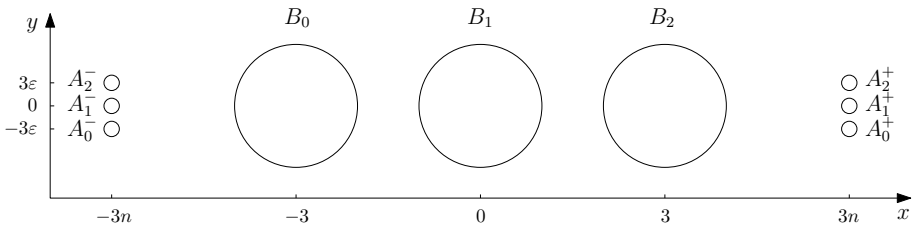
Theorem 85. *There are $\Omega(n^3)$ isolated[16] free lines tangent to four of the balls of $\mathcal{A}^\pm$ and $\mathcal{B}$.*

We first describe intuitively why our construction gives this $\Omega(n^3)$ bound. First, consider only two consecutive balls B_i and B_{i+1}. We study the lines that are tangent to them close to the highest points of the balls. These lines are almost in the horizontal plane $z = 0$. Now, in this plane, the balls in $\mathcal{A}^-$ and $\mathcal{A}^+$ form two sets of gates which decompose the set of free lines in $\Omega(n^2)$ connected components defined by the gates the line goes through. On the boundary of each component there are lines tangent to the gates, and in particular there are $\Omega(n^2)$ free lines tangent to a ball of $\mathcal{A}^-$ and one of $\mathcal{A}^+$ (and to the 2 balls of $\mathcal{B}$). Since this can be done for any two consecutive balls of $\mathcal{B}$, there are $\Omega(n^3)$ free lines tangent to four balls, and since the centers of these balls are not aligned, these tangents are isolated (see [25]).

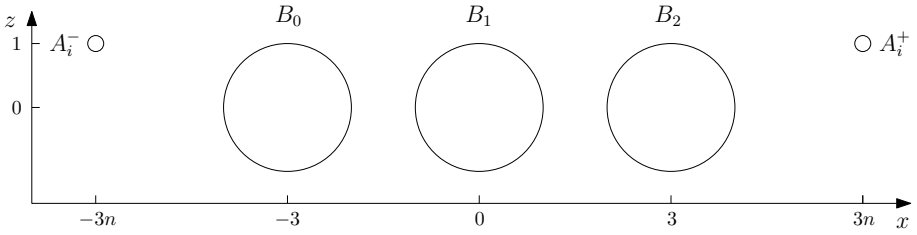
Our first step is to formalize the fact that we can pretend that the tangents to 2 consecutive balls in $\mathcal{B}$ lie in the plane $z = 0$. We start with the following technical lemma.

Lemma 86. *Let $\tilde{B}_0$ and $\tilde{B}_1$ be 2 unit balls centered at $(0,0,0)$ and $(3,0,0)$. Let L be a line tangent to $\tilde{B}_0$ and $\tilde{B}_1$ at $M_0(x_0, y_0, z_0)$ and $M_1(x_1, y_1, z_1)$ respectively. Assume that z_0 and z_1 are positive, $|y_0|$ and $|y_1|$ are smaller than a constant $m < 1/25$ and $|y_1 - y_0|$ is smaller than some constant $\alpha \leqslant 2m$. Then $|z_1 - z_0|$ is smaller than $110m\alpha$ and both z_0 and z_1 are larger than $1 - 100m^2$.*

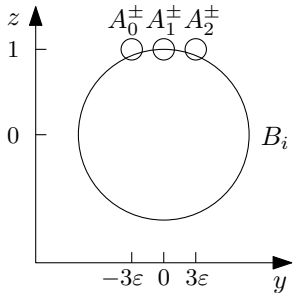
[16] A line is isolated with repect to a set of tangent balls if the line cannot be moved continuously while remaining tangent to these balls.



(a) Projection on the xy-plane.



(b) Projection on the xz-plane.



(c) Projection on the yz-plane.

Figure 9.1: Illustration of our construction for Theorem 84.

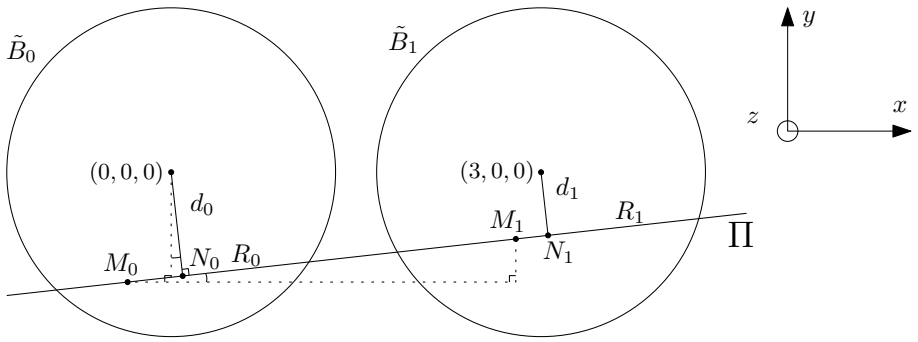


Figure 9.2: For the proof of Lemma 86: balls $\tilde{B}_0$ and $\tilde{B}_1$ viewed from above.

Proof. Notice that the result is intuitively clear. Writing that M_0 is on $\tilde{B}_0$ gives: $x_0^2+y_0^2+z_0^2=1$. Similarly, $(x_1-3)^2+y_1^2+z_1^2=1$. The difference is $(x_1-3)^2-x_0^2+y_1^2-y_0^2+z_1^2-z_0^2=0$. Now, if instead of balls we had disks (flat in the x-direction), we would have $x_0=0$, $x_1=3$, $|z_i|=\sqrt{1-y_i^2}\geqslant\sqrt{1-m^2}\geqslant 1-m^2$ and $|z_1-z_0|=\frac{|y_1-y_0||y_1+y_0|}{|z_1+z_0|}<\frac{2m\alpha}{2(1-m^2)}<2m\alpha$.

Since the balls are not disks, we need a few more steps. We consider the vertical plane Π that contains L (see Figure 9.2). It cuts the 2 spheres in two circles of centers N_0 and N_1 and radii R_0 and R_1. First, we prove that $|N_{0y}|$ and $|N_{1y}|$ are smaller than $5m$. In projection on the xy-plane, since M_0 and M_1 are on L, the absolute value of the slope of L is $\frac{|y_1-y_0|}{|x_1-x_0|}\leqslant 2m$ since $|y_1-y_0|\leqslant 2m$ and $|x_1-x_0|\geqslant 1$. Now N_i is in Π so its projection on the xy-plane is on the projection of L. Since $|N_{ix}-x_i|\leqslant 2$ (M_i and N_i are in the same unit ball), $|N_{iy}-y_i|\leqslant 2\cdot 2m$ and thus $|N_{iy}|\leqslant|y_i|+4m\leqslant 5m$.

Let d_i denote the signed distance from the center of $\tilde{B}_i$ to Π such that d_i has the same sign as N_{iy}. Notice that because the two angles shown on Figure 9.2 are equal, they have the same cosine, that is:

$$N_{iy}/d_i=(x_1-x_0)/\sqrt{(x_1-x_0)^2+(y_1-y_0)^2}.$$

The right-hand expression can be rewritten as:

$$\frac{1}{\sqrt{1+\left(\frac{y_1-y_0}{x_1-x_0}\right)^2}}\geqslant\frac{1}{\sqrt{1+4m^2}}>\frac{1}{2}.$$

We then have $d_i=\chi N_{iy}$ where $0<\chi<2$. This implies that $d_i<2N_{iy}\leqslant 10m$ and $|d_1-d_0|=\chi|N_{1y}-N_{0y}|<2|N_{1y}-N_{0y}|$. Once again, the projections of M_0, M_1, N_0 and N_1 on the xy- plane are aligned, so the slope of the projection of L is $(N_{1y}-N_{0y})/(N_{1x}-N_{0x})=(y_1-y_0)/(x_1-x_0)$. Since $|N_{1x}-N_{0x}|\leqslant 5$, $|d_1-d_0|<10\alpha$.

The radii of the intersection circles satisfy $d_i^2+R_i^2=1$. This implies that $R_i\geqslant\sqrt{1-(10m)^2}$. Also, $(R_1-R_0)(R_1+R_0)=-(d_1-d_0)(d_1+d_0)$, so $|R_1-R_0|\leqslant 10\alpha\frac{20m}{2\sqrt{1-(10m)^2}}$ $<110m\alpha$ because, since $m<1/25$, $1/\sqrt{1-(10m)^2}<1.1$.

We now work in the plane Π, using a Cartesian coordinate system (w,z) (see Figure 9.3). Let θ be the (unsigned) angle between L and the w-axis. We have $z_i=R_i\cos\theta$.

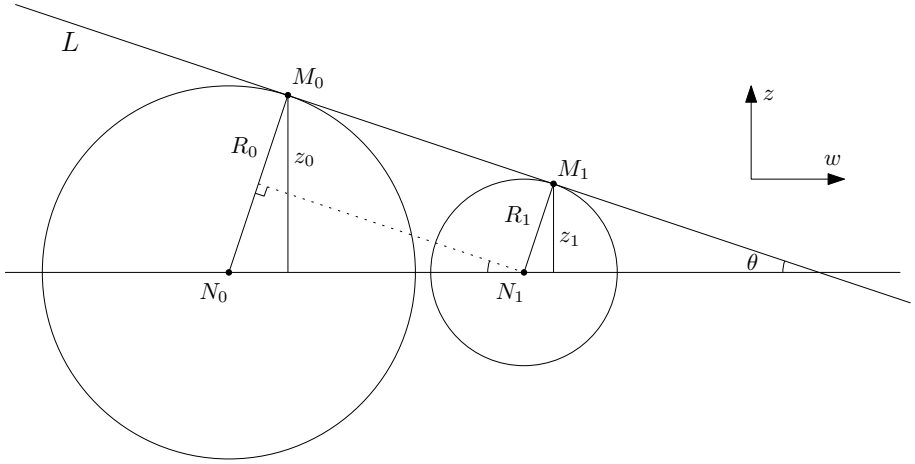


Figure 9.3: For the proof of Lemma 86: intersection of balls $\tilde{B}_0$ with plane Π.

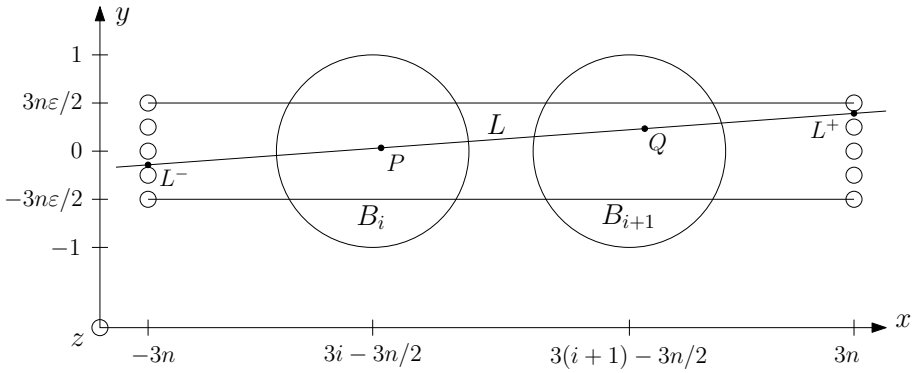


Figure 9.4: A line L for Lemma 87.

Therefore, $z_1 - z_0 = (R_1 - R_0)\cos\theta$ and $|z_1 - z_0| \leqslant |R_1 - R_0| < 110m\alpha$, which is the first inequality of the lemma.

Consider now the line in Π parallel to L through N_1 if $R_1 \leqslant R_0$ and through N_0 otherwise, as shown on Figure 9.3, and remember that the distance between N_0 and N_1 is at least 1. We have that $\sin\theta = |R_1 - R_0|/||N_1 - N_0|| < 110m\alpha \leqslant 220m^2 < 10m$. Hence $\cos\theta > \sqrt{1-(10m)^2}$. We have already proved that $R_i \geqslant \sqrt{1-(10m)^2}$. Therefore $z_i = R_i\cos\theta > 1 - 100m^2$ which concludes the proof. Note that all the constants could easily be improved. □

We now prove that roughly speaking a line tangent to two consecutive balls of $\mathcal{B}$ near their pole (their point in the plane $z = 0$) intersects the convex hull of $\mathcal{A}^-$ and $\mathcal{A}^+$ respectively and thus that the balls of $\mathcal{A}^\pm$ play the role of gates as discussed earlier.

Lemma 87. *Let $i \in 0, \ldots, n-1$ and let L be a line tangent to B_i and B_{i+1} at some points with positive z-coordinate. Assume that L intersects the plane $x = 3n$ in a point L^+ and the plane $x = -3n$ in a point L^- such that $|L^+_y|$ and $|L^-_y|$ are both smaller than $3n\varepsilon/2$ with $\varepsilon < \frac{1}{5400n^2}$. Then $|L^+_z|$ and $|L^-_z|$ are both smaller than $\varepsilon/2$.*

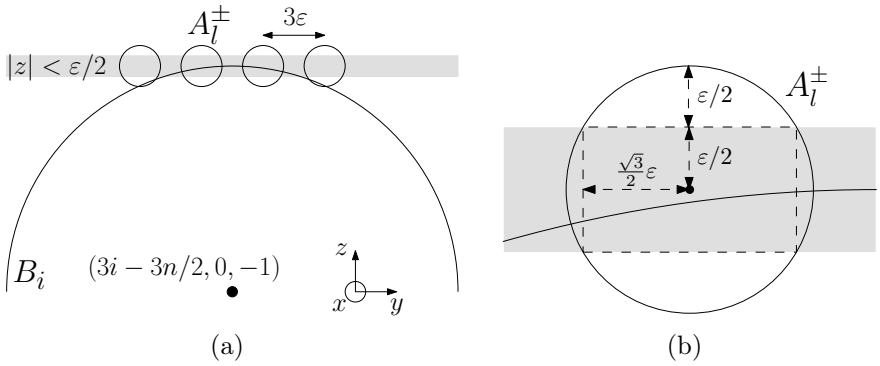


Figure 9.5: For the proof of Lemma 88: lines L intersect planes $x = \pm 3n$ in the shaded region.

Proof. Let P and Q denote the tangency points of L on B_i and B_{i+1} (refer to Figure 9.4). L^-, P, Q and L^+ are aligned in this order on L, and $|L_y^+|$ and $|L_y^-|$ are both smaller than $3n\varepsilon/2$, so $|P_y|$ and $|Q_y|$ are smaller than $3n\varepsilon/2$. Furthermore, the slope of the projection of L in the xy-plane is $\frac{L_y^+ - L_y^-}{L_x^+ - L_x^-} = \frac{Q_y - P_y}{Q_x - P_x}$. Notice that by hypothesis $|L_y^+ - L_y^-| \leqslant 3n\varepsilon$, $L_x^+ - L_x^- = 6n$ and $Q_x - P_x \leqslant 5$ so $|Q_y - P_y| \leqslant 5\varepsilon/2$. We can now apply Lemma 86 because $|P_y|$ and $|Q_y|$ are both smaller than $m = 3n\varepsilon/2$ which is smaller than $1/25$ since $\varepsilon < 1/5400n^2$ and $|Q_y - P_y| \leqslant 5\varepsilon/2$ which is smaller than $2m = 3n\varepsilon$ ($n \geqslant 2$). We thus get $|Q_z - P_z| \leqslant 110\frac{3n\varepsilon}{2}\frac{5\varepsilon}{2}$ and $|Q_z| \leqslant 100(\frac{3n\varepsilon}{2})^2$.

L^-, P, Q and L^+ are still aligned on L, and we now compute the slope of the projection of L on the xz-plane: $\frac{L_z^+ - Q_z}{L_x^+ - Q_x} = \frac{Q_z - P_z}{Q_x - P_x}$. By definition, $|L_x^+ - Q_x| < 6n$ and $Q_x - P_x \geqslant 1$ so $|L_z^+| \leqslant |Q_z| + 6n|Q_z - P_z| \leqslant (\frac{9}{4}100 + 6\frac{15}{4}110)n^2\varepsilon^2 = 2700n^2\varepsilon^2 < \varepsilon/2$ since $\varepsilon < \frac{1}{5400n^2}$. The same holds for $|L_z^-|$ □

Lemma 88. *For any (i,j,k) (except possibly the extremal ones), there are four free lines tangent to B_i, B_{i+1}, A_j^- and A_k^+.*

Proof. We consider the projection along the z-axis. In the xy-plane, we denote by S^+ and S^- the segments $x = \pm 3n$, $-3n\epsilon/2 < y < 3n\epsilon/2$. Any line L as in Lemma 87 projects to a line that intersects both segments. Reciprocally, consider a line in the xy-plane that intersects S^- and S^+. It defines a vertical plane that contains it. This plane intersects balls B_i and B_{i+1} in two disks. Their upper bitangent is the unique line L satisfying the hypotheses of Lemma 87 that projects to it. There is thus a bijection between the lines L of Lemma 87 and the lines in the projection plane that intersect both segments. This property of intersecting both segments gives a natural parameterization of this set of lines by the y-coordinates of the intersections, called u (for S^-) and v (for S^+). Notice that with the notations of Lemma 87, $u = L_y^-$ and $v = L_y^+$.

Using this parameterization, we consider this set of line (represented as a square in the (u, v)-parameter space) and the balls $A_l^{\pm}$ that each line intersects. Referring to Figure 9.5, notice that the set of values of u such that for all v the corresponding lines L intersect a given ball A_j^- contains some non-empty interval u_j (which depends only on j). More

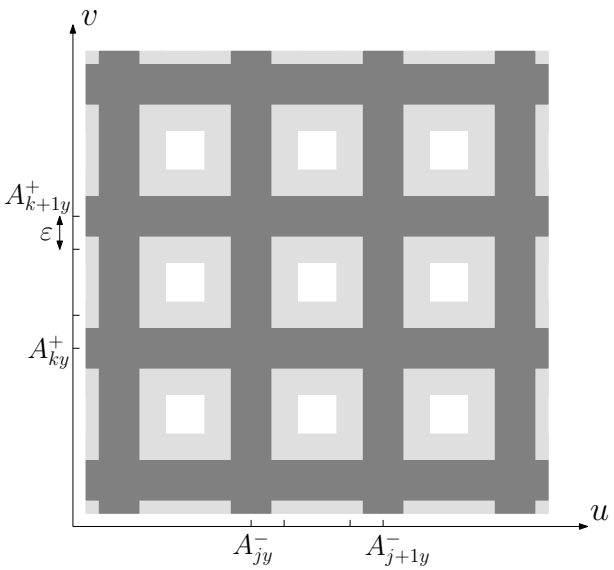


Figure 9.6: For the proof of Lemma 88: in the system (u, v), the lines in black intersect a ball in $\mathcal{A}^{\pm}$, the lines in white are free and the lines in grey are unknown.

precisely, if $|u - A^-_{jy}| < \frac{\sqrt{3}}{2}\varepsilon$, L^- is inside A^-_j (see Lemma 87) so L obviously intersects A^-_j.

On the other hand, the set of values of u such that there exists a v for which the corresponding line L intersects a given ball A^-_j is contained in an interval u'_j such that $u'_j \cap u'_l = \emptyset$ if $l \neq j$. Precisely, if $A^-_{jy} + \frac{5}{4}\varepsilon < u < A^-_{j+1y} - \frac{5}{4}\varepsilon$, then for all points on L that have a x-coordinate between $-3n - \varepsilon$ and $-3n + \varepsilon$, the y-coordinates differ from $u = L^-_y$ by at most $\frac{|L^+_y - L^-_y|}{L^+_x - L^-_x}\varepsilon$ (the slope of the projection of L onto the xy-plane times the difference of the x-coordinates). By definition this is smaller than $3n\varepsilon^2/6n < \varepsilon/4$, which means that the y-coordinate stays between $A^-_{jy} + \varepsilon$ and $A^-_{j+1y} - \varepsilon$. Therefore L cannot intersect A^-_j or A^-_{j+1} (and much less the other balls $\mathcal{A}^-$).

We have thus proved that for each ball A^-_j there is an interval u'_j of values of u outside of which for all v the line L (parameterized by (u, v)) doesn't intersect A^-_j and that each u'_j contains a non-empty interval u_j of values of u such that for any v the line L intersects A^-_j. Moreover the intervals u'_j are pairwise disjoint. Similarly, we consider the intervals v_k and v'_k.

We now partition the set space of lines L in paramater space (u, v) as follows (see Figure 9.6): the black region is the set of $(u, v) \in [\frac{-3n\epsilon}{2}, \frac{3n\epsilon}{2}]^2$ such that $u \in u_l$ or $v \in v_l$ for some l; the white region is the set of (u, v) such that $u \notin \bigcup_l u_l$ and $v \notin \bigcup_l v_l$; the grey region is the complement of the black and white regions.

Consider a line L that corresponds to a point in a white region. If we decrease u, at some point the line becomes tangent to some A^-_j. Similarly, if we increase u, decrease v or increase v, at some point the line becomes tangent to A^-_{j+1}, A^+_k or A^+_{k+1} respectively. There are then 4 balls whose tangents define the boundaries of a connected component of free lines tangent to B_i and B_{i+1}. The points where these boundary pieces meet correspond to lines tangent to 2 consecutive balls of $\mathcal{B}$, one of $\mathcal{A}^-$ and one of $\mathcal{A}^+$. □

Using the result of [25] that shows that four distinct balls can have non-isolated common tangents only if their centers are aligned, Theorem 85 is a direct consequence of Lemma 88.

Remark Notice that in our construction several balls are aligned and have infinitely many common tangents. However, these are not lines we are interested in. By substracting $(i-n/2)^2\varepsilon_2$ (where ε_2 is very small compared to any other length involved here) from the z-coordinate of the center of every ball the scene can be made non-degenerate without changing the number of isolated tangents we consider in the proof. This means that n balls that admit finitely many free lines tangent to 4 of them may admit $\Omega(n^3)$ of them in the worst case.

9.3 $\Omega(n^4)$ line segments tangent to unit balls

9.3.1 Construction

Theorem 89. *The complexity of the space of maximal free line segments among n balls is $\Omega(n^4)$ in the worst case.*

Once again, we prove this theorem by giving an exemple.

Consider two sets of unit balls $\mathcal{A} = \{A_1 \dots A_n\}$ and $\mathcal{B} = \{B_0 \dots B_n\}$. The center of A_i has coordinates $(-L - 3i, i \cdot \varepsilon, 0)$. The center of B_i has coordinates $(-L + 3i, (2 - \eta)\cos(\alpha_i), (2 - \eta)\sin(\alpha_i))$ where $\alpha_i = -\frac{\alpha}{2} + \frac{\alpha}{n}i$. Figure 9.7 shows an orthogonal view of the sets $\mathcal{A}$ and $\mathcal{B}$ along the $x-$direction.

Now consider the sets $\mathcal{C} = \{C_1 \dots C_n\}$ and $\mathcal{D} = \{D_0 \dots G_n\}$ obtained by symmetry with respect to the $yz-$plane.

Lemma 90. *We can choose (in this order) the values $\alpha < \pi$, η small enough, ε small enough, L large enough, so that there are $\Omega(n^4)$ $4-$tuples (i, j, k, l) such that A_i, B_j, C_k and D_l admit a common non-obstructed tangent line segment.*

Proof. Consider the affine transformation changing x into x/L. It flattens the spheres into ellipsoids. When L tends to infinity, the transformed scene tends to two flat versions of Figure 9.7 on the planes $x = \pm 1$, facing each other. Joining the $\Theta(n^2)$ intersections on each side defines $\Theta(n^4)$ free line segments tangent to 4 of the disks. This means that for L close enough to infinity, these free line segments still exist. Moreover, each of the free line segments tangent to four ellipsoids remains tangent by the inverse affine transformation. □

9.3.2 Discussion

It is also possible to give an alternative proof that looks like the proof of Lemma 85. This other proof has the advantage that it shows how large or small the constants α, η, ε and L need to be. However, it is very tedious, so we only sketch it here. We consider any pair of balls, one from $\mathcal{A}$ and one from $\mathcal{C}$. First we want to make sure that other balls

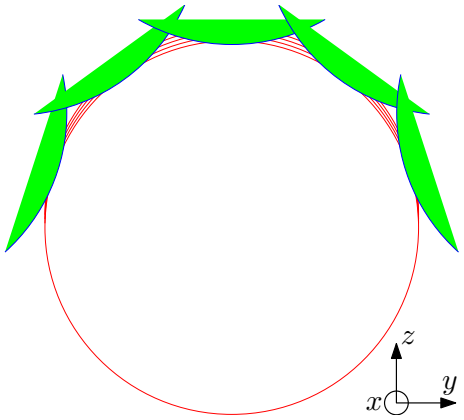


Figure 9.7: View of A_i and B_j.

from $\mathcal{A}$ or $\mathcal{C}$ don't occlude their bitangents. If we set $\alpha = \pi/6$ for instance, this gives a condition $\varepsilon > C_0/L$ (the C_i denote various positive constants that will not be made explicit). Then we want the balls from $\mathcal{B}$ and $\mathcal{D}$ to intersect these sets of bitangents: $\eta > C_1 n/L$. We also need some of those bitangents not to intersect the balls from $\mathcal{B}$ and $\mathcal{D}$: $(\alpha/n)^2 - C_2\eta - C_3 n\varepsilon > C_4 n/L$. When all those conditions are satisfied, the lines tangent to a ball from $\mathcal{A}$ and one from $\mathcal{C}$ can be represented in a grid similar to Figure 9.6 and the same arguments makes it possible to conclude the proof. Now we only have to chose the values of η, ε and L. We can chose them in this order and end up with $\eta \sim \frac{1}{n^2}$, $\varepsilon \sim \frac{1}{n^3}$ and $L \sim n^3$. We do not know if it is possible to have as many tangents in a scene of smaller diameter.

Chapter 10

Visibility in an environment of spheres

Ce travail a été réalisé en collaboration avec H. Alt.

10.1 Introduction

We have seen in previous chapters a number of interesting results about worst-case scenes, where in particular the visibility graph (the graph with vertices representing the objects and edges between the objects that can see each other) can easily have a quadratic size.

However, it is also interesting to know what happens in general, for nicely behaved scenes. For instance, when walking in the woods, one usually cannot see more than a hundred meters away, because the trunks of the trees block all the horizontal visibility. This means that the number of trees one can see from a point is bounded and the visibility graph has linear size. This result has been known for a long time, in 2D (the case of the forest) as well as in 3D where it gives rise to Olbers' famous paradox. This paradox is the simple observation that in an infinite universe that contains uniformly distributed luminous stars, the sky at night should be as bright as the sun. Explaining this phenomenon was an important part of elaborating a consistant model of the universe.

A study in [46] makes it possible to give a linear bound on the number of lines tangent to 4 spheres in a scene of unit spheres randomly distributed in a universe that is a 3D ball, or according to a Poisson distribution.

Our aim here is to generalize those results by giving a strong bound on the number of balls a given ball can see, in any dimension. The bound we present also makes it possible to bound the average of the square (or any power) of the number of balls a ball can see. This allows in particular to bound the number of triplets of balls one ball can see, and thus the number of lines tangent to four balls.

10.2 Basic idea

Let S be a set of randomly distributed points in $\mathbb{R}^d$. Precisely the distribution we consider is a Poisson measure with intensity λ. This means that the probability to have k points in a region of volume V is $\frac{e^{-\lambda V}(\lambda V)^k}{k!}$ (so the expected number of points is λV) and that

the distributions of any two disjoint regions are independent. The vocabulary used will correspond to the case where $d = 3$, and d is considered a constant. The points of S are the centers of unit balls, and by extension S may also be used to designate this set of balls. We first prove here that the probability for a ball to see more than n balls decreases very fast.

Theorem 91. *The probability for a ball B to see more than n balls of S is*

$$O\left(\left(\frac{n}{\lambda}\right)^{1-1/d} e^{-Cn^{1/d}\lambda^{1-1/d}}\right).$$

Proof. Let B be any unit ball. The purpose is to give an upper-bound on the number of balls visible from B. Let E be a sphere with the same center as B and radius R. We want to show that for a large enough value of R, no point of E is visible from B.

Consider B as a light source. A ball of S whose center lies inside E generates on E an umbra (complete shadow) delimited by a circle of radius 1. Actually, the ball has to be inside a ball slightly smaller than E, or we have to consider circles of radius slightly smaller than 1, but the difference is not significant.

Decompose E into small pieces of roughly the same size. Consider the circles of radius 1 that can be drawn on E and call center of such a circle the projection of its real center onto E (for a large R, those two centers are very close). We want the pieces of E to be small enough that every unit circle encloses at least one of them. For each piece, consider the set of centers of unit circles that enclose it (call it the covering zone of this piece). Since every circle encloses at least one piece, the covering zones cover E. If things are done properly, all covering zones have roughly the same size. This can be achieved with $O(R^{d-1})$ pieces. The zones then have an area $\Omega(1)$.

A zone is a set of centers of circles on E, which correspond to unit balls in E whose centers lie inside a conical region whose appex is the center of E. This conical region has volume $\Omega(R)$. The probability that such a region does not contain any point of S is $O(e^{-C\lambda R})$ for some positive constant C that depends only on d. The probability that at least one region is empty is $O(R^{d-1}e^{-C\lambda R})$. As expected, this tends to 0 when R tends to $+\infty$.

Let n be a positive integer. We are going to give an upper bound on the probability for B to see more than n balls. For any value of R, if B can see more than n balls, then either E contains more than n points of S, or some points of E are visible from B (the two are not exclusive). All we need now is to chose an appropriate value for R, say for instance such that $n = 2 \cdot \lambda V_d(R)$ where $V_d(R)$ is the volume of a d-dimensional ball of radius R. The probability for B to see more than n balls is then $O\left(\left(\frac{n}{\lambda}\right)^{1-1/d} e^{-Cn^{1/d}\lambda^{1-1/d}}\right) + O\left(e^{-n/2}\frac{(n/2)^n}{n!}\right)$. The second term is negligible compared to the first one. □

10.3 Consequences

This exponential bound on the probability to see many balls implies that the average number of balls one ball can see is bounded. Morever, for any positive integer k, the

average of the kth-power of the number of balls one ball can see is bounded (this bound still depends on k of course). Using this observation, for any data structure that involves sets of balls that can see each other for which a polynomial upper bound is known, we know the average case is locally bounded, or in other words the structure has linear size on average. This includes the visibility graph, skeleton and complex for instance.

The expected distance where visibility ends is $O(\lambda^{-d})$. Indeed, it is upper-bounded by a constant plus a constant times $\int_0^\infty R^{d-1}e^{-C\lambda R} = \frac{(d-1)!}{(C\lambda)^d}$.

10.4 Algorithmic point of view

This study can be seen from an algorithmic point of view. In order to compute all the balls that a given ball can see, we start with a sphere E that is large enough to contain on average 20 balls. We then check whether the balls inside E completely occlude E from the central ball (for instance using the bucket technique presented above for enumeration purposes). If the balls do occlude E then we stop. Otherwise we try again with a sphere E twice as large. If we are really interested only in the balls that can see the central ball and not only a superset of these, we can then check for visibility among the set of balls inside E, which has a bounded expected size.

We need to be able to determine quickly which balls are inside E. For this, we can for instance store the balls in a grid with $\Theta(1)$ points per cell.

10.5 Other objects

Objects and obstacles are not the same

In the previous analysis, we do not use the fact that the objects we see and the obstacles are the same, we only require that they have the same density. It would actually work just as well if the two densities differed only by a constant factor.

Various shapes, same size

Instead of balls, we now have fat objects. We assume that there exist two constants A and B such that for every object there is a point (called center of the object) for which the ball of radius A (resp. B) centered at this point is contained in (resp. contains) the object. We can then replace an object by either its interior ball, when it is considered as an obstacle, or a union of balls of radius A that cover its enclosing ball but are transparent to each other, when we are interested in the visibility from this object. All that changes in the analysis (besides a scaling of factor $1/A$ is that we now need the visibility from each of the covering balls to be blocked, so this introduces a constant factor in the probability, corresponding to the number of balls of radius A that are necessary to cover a ball of radius B.

Balls of random size

The objects are still balls with centers uniformly distributed, but now the radius of each ball is chosen uniformly and independently at random between 0 and 1. We consider as obstacles the objects of size larger than 1/2. The objects larger than 1/2 are covered with balls of size 1/2 as in the previous paragraph. If we discard each object with probability α, we get a similar distribution of density $\alpha\lambda$ (easy computation). The first remark of this section, that says that obstacles and objects that can see each other can have densities that different by a constant factor, is sufficient to prove that the main result extends to this case.

2D-fat objects with random orientation

We now consider objects that are contained in a ball of radius B and contain a disk of radius A. We further impose that the constants A and B be the same for all objects. For each object, the orientation is chosen uniformly and independently at random. The main difference with the previous case is that the objects that are discarded are not always the same, they depend on the viewpoint. But simply using the idea that an object has a certain probability to block visibility as well as a ball of small fixed radius makes this case similar to the previous one (there are just a few more details to take into account).

10.6 Bounded universe

The study above about an infinite universe is interesting, but it would be even more so it if was extended to another universe than $\mathbb{R}^d$. However, as soon as the universe has a boundary, things are not so nice any more. For instance, in a universe that is a half-space delimited by a hyperplane, a ball that is exactly on the boundary is expected to see an infinity of other balls, in contrast to the case studied here where each ball could only see a finite number of other balls with probability 1. Still, the expected number of balls a ball can see only gets large very close to the boundary of the universe, and the probability for a ball to be that close to the boundary is correspondingly small, so some bounds could be obtained with more work. Other types of boundaries for the universe (say a sphere) have different properties.

Another possible extension would be to study, in a finite universe, how the result is affected by fixing the total number of balls.

Chapter 11

Helly-type theorems for approximate covering

Ce travail, réalisé avec J. Demouth, O. Devillers et X. Goaoc, a été soumis au Symposium on Discrete Algorithms.

Let $\mathcal{F}$ be a covering of a unit ball $U \subset \mathbb{R}^d$ by unit balls. We prove that for any $\epsilon > 0$, the smallest subset of $\mathcal{F}$ leaving at most a volume ϵ of U uncovered has size $O\left(\epsilon^{\frac{1-d}{2}} \operatorname{polylog} \frac{1}{\epsilon}\right)$. We give an example showing that this bound is tight in the worst-case, up to a logarithmic factor, and deduce an algorithm to compute such a small subset of $\mathcal{F}$ in time $O\left(|\mathcal{F}|\epsilon^{\frac{1-d}{2}} \operatorname{polylog} \frac{1}{\epsilon}\right)$. We then extend these results in several directions, including covering boxes by boxes and visibility among disjoint unit balls in $\mathbb{R}^3$.

11.1 Introduction

A family $\mathcal{F}$ of sets covers a set U if the union of the elements of $\mathcal{F}$ contains U. In the geometric setting, both U and the elements of $\mathcal{F}$ are subsets of a geometric space, for example points, hyperplanes or balls in $\mathbb{R}^d$. Many geometric problems, for example visibility problems, reduce to deciding if a given family $\mathcal{F}$ covers an object U.

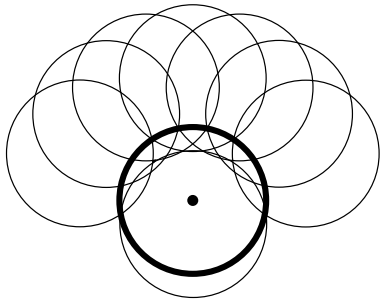


Figure 11.1: Arbitrarily many disks can be necessary to cover another disk.

A family $\mathcal{F}$ covers a set U if and only if the intersection of U with the complement of the elements of $\mathcal{F}$ is empty. Thus, it follows from Helly's theorem [71] that if U is convex and the family $\mathcal{F}$ consists of complements of convex sets, the smallest subset of $\mathcal{F}$ covering U has size at most $d+1$. The cases such that this size is bounded are in fact rather exceptional: for most classes of objects there exist arbitrarily large minimal covering families (Figure 11.1 shows an example of a minimal covering of a unit disk by 8 unit disks in the plane).

Yet, it seems clear that if a family $\mathcal{F}$ of unit disks covers a unit disk U, then a *small* subset of $\mathcal{F}$ is sufficient to cover *most* of U. Let $\mathcal{G} \cup \{U\}$ be a collection of measurable sets in some measurable space, typically $\mathbb{R}^d$. Given a real $\epsilon > 0$, we say that $\mathcal{G}$ is an *ϵ-covering* of U if at most a measure ϵ of U is not covered by $\mathcal{G}$. In this article, we study the size of the smallest ϵ-covering contained in a covering. We show that, in several cases, for any $\epsilon > 0$ this size can be bounded independently of the size of the covering. In other words, we prove Helly-type theorems for approximate covering.

Our results. Our main results are the following:

- We prove that the size of the smallest ϵ-covering contained in a covering of a unit ball in $\mathbb{R}^d$ by n unit balls can be bounded independently of n (Theorem 92). We also give examples of covering in which our bound is tight up to a polylogarithmic factor (Lemma 97).

- We extend this result to two other covering problems: the covering of a unit box by unit boxes (Theorems 100 and 101) and visibility occlusion among disjoint unit balls in $\mathbb{R}^3$ (Theorem 103). In the particular case of the covering of a unit square by unit squares in the plane, we also give examples where our bound is tight. In all cases, our bounds are independent of the size of the covering family.

- For each of the above situations, we obtain algorithms (Theorems 98, 102 and 105) that, given a family of objects $\mathcal{F} \cup \{U\}$, decide in $O(|\mathcal{F}| f(\epsilon))$-time that $\mathcal{F}$ does not cover U or that $\mathcal{F}$ is an ϵ-cover of U – with $f(\epsilon)$ denoting our bound on the size of the smallest ϵ-covering subset for that situation. More precisely, our algorithms return either a point in U not covered by $\mathcal{F}$ or an ϵ-cover of U of size $f(\epsilon)$ contained in $\mathcal{F}$.

Related work. The classical problem of set cover asks, given a covering $\mathcal{F}$ of a set U, for the smallest subset of $\mathcal{F}$ that covers U. Since most versions are **NP**-hard, approximation algorithms have been largely investigated; namely, one looks for a subset of $\mathcal{F}$ that covers U and whose size is within some multiplicative factor from the optimal. Approximation in this setting is also difficult as in many cases, no polynomial-time algorithm yielding constant factor approximations exists unless **P**=**NP** [78]. In the geometric setting, the situation is essentially the same [32, 37]. In this paper, we relax the covering problem in a different direction by allowing the covering to be imperfect. We are not aware of previous results in this direction.

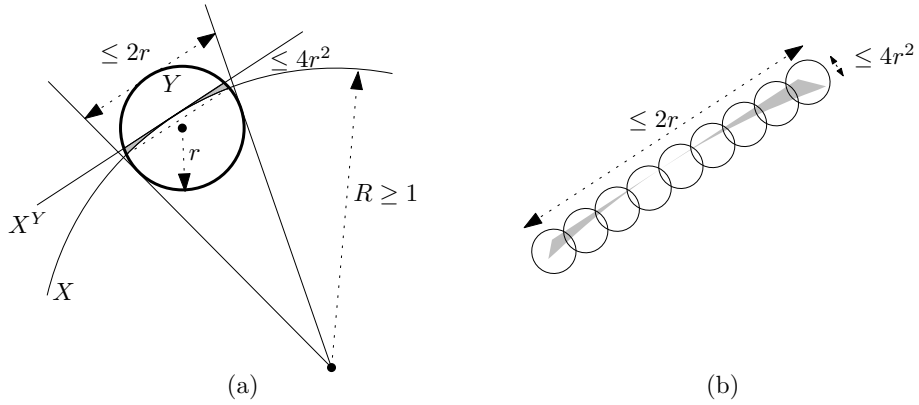


Figure 11.2: (a) $\left(X^Y \cap Y\right) \setminus (X \cap Y)$ is inscribed in a rectangle of sides at most $2r$ and $1 - \cos(\sin^{-1} r) \leqslant 4r^2$. (b) A covering with disks of radius $4r^2$ of $\left(X^Y \cap Y\right) \setminus (X \cap Y)$.

11.2 Covering a unit ball by unit balls

In this section, we study the case where all objects are unit balls in $\mathbb{R}^d$. We prove:

Theorem 92. *Let U be a unit ball in $\mathbb{R}^d$ and $\mathcal{F}$ a covering of U by unit balls. For any $\epsilon > 0$, the smallest subset of $\mathcal{F}$, that is an ϵ-covering of U, has size $O\left(\epsilon^{\frac{1-d}{2}} \operatorname{polylog} \frac{1}{\epsilon}\right)$.*

We also give examples for which this bound is tight up to the logarithmic factor (Theorem 97). For the clarity of the exposition, we prove the result in two dimensions (Section 11.2.1) before discussing the general case (Section 11.2.2). Our proof leads to an algorithm (discussed in Section 11.2.3) for computing, given $\mathcal{F}$, U and ϵ, either a point in U not covered by $\mathcal{F}$ or a small ϵ-covering of U contained in $\mathcal{F}$.

11.2.1 The planar case

Upper bound. For two disks X and Y, we denote by X^Y the half-plane containing X and bounded by the tangent to X at the projection[17] of the center of Y on the boundary of X (see Figure 11.2). We denote by $\mathcal{F}^Y$ the collection $\left\{X^Y \mid X \in \mathcal{F}\right\}$. We first start by a technical lemma:

Lemma 93. *Let Y be a disk of radius $r < 1$ and $\mathcal{F}$ a covering of a unit disk U by unit disks. Then, $U \cap Y$ can be covered by a triple $C(Y) \subset \mathcal{F}$ and a collection $R(Y)$, of at most $\frac{3}{r}$ disks of radius $4r^2$.*

Proof. Since the collection $\mathcal{F}^Y$ covers U, it also covers $U \cap Y$ and, by Helly's theorem, three of these half-planes must cover $U \cap Y$. We denote by $C(Y)$ the corresponding balls in $\mathcal{F}$. For any disk $X \in \mathcal{F}$, the area $\left(X^Y \cap Y\right) \setminus (X \cap Y)$ is inscribed in a rectangle (see Figure 11.2) with sides respectively smaller than $2r$ and $4r^2$. This rectangle can thus be covered by overlapping disks of radius $4r^2$ centered on its larger axis (Figure 11.2(b)). By

[17] If the two disks have the same center, we can choose any tangent to X.

choosing the disks so that the height covered at the intersection between two disks is, at least, $4r^2$, we need only $\frac{1}{r}$ disks. □

We can now prove Theorem 92 for the case $d = 2$:

Proof of Theorem 92 for $d = 2$. Let r be a real in the open interval $\left(\frac{1}{100}, \frac{1}{10}\right)$, let $\mathcal{R}_1$ be a covering of U by μr^{-2} disks of radius r, for some constant μ, and let $\mathcal{C}_1$ denote the empty set. We construct an ϵ-cover by iterating the following operation:

$$\mathcal{C}_{i+1} \leftarrow \mathcal{C}_i \cup \left(\bigcup_{Y \in \mathcal{R}_i} C(Y) \right) \quad \text{and} \quad \mathcal{R}_{i+1} \leftarrow \bigcup_{Y \in \mathcal{R}_i} R(Y),$$

where $C(\cdot)$ and $R(\cdot)$ denote the sets defined in Lemma 93. By induction, we have that, for any $i \geqslant 0$, $\mathcal{C}_i \cup \mathcal{R}_i$ covers U. Let α_i denote the area covered by disks in $\mathcal{R}_i$; $\mathcal{C}_i$ is an α_i-cover of U. The disks in $\mathcal{R}_i$ have radius r_i satisfying the recurrence relation

$$r_i = 4r_{i-1}^2, \quad \text{with} \quad r_1 = r,$$

and thus $r_i = \frac{1}{4}(4r)^{2^{i-1}}$. The number of disks in $\mathcal{R}_i$ is governed by the relation

$$|\mathcal{R}_i| \leqslant \frac{3}{r_{i-1}} |\mathcal{R}_{i-1}|,$$

which gives:

$$|\mathcal{R}_i| \leqslant \left(\prod_{k=1}^{i-1} 12(4r)^{-2^{k-1}} \right) |\mathcal{R}_1| \leqslant 12^{i-1}(4r)^{1-2^{i-1}} \mu r^{-2},$$

and:

$$\alpha_i \leqslant \pi r_i^2 |\mathcal{R}_i| = 12^{i-1} \mu \pi (4r)^{2^{i-1}-1}.$$

Moreover, for each element in $\mathcal{R}_{i-1}$, we add three disks from $\mathcal{F}$ to $\mathcal{C}_i$. Thus, the size of $\mathcal{C}_i$ is given by

$$|\mathcal{C}_i| \leqslant 3 \sum_{k=1}^{i-1} |\mathcal{R}_k| \leqslant 3\mu r^{-2} 12^{i-1} \sum_{k=1}^{i-1} (4r)^{1-2^{k-1}} = O(12^i i (4r)^{1-2^{i-2}}),$$

for $r \in \left(\frac{1}{100}, \frac{1}{10}\right)$. Any sufficiently small $\epsilon > 0$ can be expressed as $\epsilon = 12^{k-1}\mu\pi(4r)^{2^{k-1}-1}$ for some integer k and real $r \in \left(\frac{1}{100}, \frac{1}{10}\right)$. The previous inequalities then give that $\alpha_i \leqslant \epsilon$ and $|\mathcal{C}_k| = O\left(\epsilon^{-\frac{1}{2}} \log^6 \frac{1}{\epsilon}\right)$. □

Lower bound. The following construction shows that the upper bound in Theorem 92 is optimal for $d = 2$ up to the logarithmic factor.

Lemma 94. *There exists a family $\mathcal{F}$ of unit disks in $\mathbb{R}^2$ covering a unit disk $U \subset \mathbb{R}^2$ such that, for arbitrary small $\epsilon > 0$, any ϵ-covering of U contained in $\mathcal{F}$ has size $\Omega(\epsilon^{-\frac{1}{2}})$.*

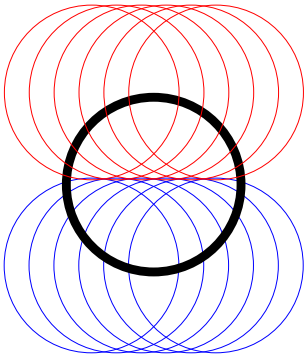


Figure 11.3: A family whose smallest ϵ-cover has size $\Omega\left(\frac{1}{\sqrt{\epsilon}}\right)$.

Proof. We equip the plane $\mathbb{R}^2$ with a frame (O, x, y) where O denotes the center of U. Let $\mathcal{F}$ be the (infinite) family of all unit disks tangent to the x-axis inside U (see Figure 11.3) and let $\mathcal{G}$ be a finite subset of $\mathcal{F}$ that covers U except for an area of at most ϵ. Consider the subset $\mathcal{G}^+ \subset \mathcal{G}$ of disks whose centers are above the x-axis and let $x_1, \ldots, x_k$ denote the abscissae of the tangency points of the disks in $\mathcal{G}^+$, sorted increasingly.

Let $\alpha_i = x_i - x_{i-1}$. For ϵ small enough, since $\mathcal{G}$ is an ϵ-cover we have:

$$x_k - x_1 \geqslant 1 \quad \Rightarrow \quad \sum_{i=2}^{k} \alpha_i \geqslant 1.$$

The uncovered area of U above the x-axis and between the $(i-1)^{th}$ and the i^{th} disks is at least $\frac{\alpha_i^3}{24}$ since this area is bounded from below by:

$$2\int_0^{\frac{\alpha_i}{2}} \frac{1}{2}x^2 dx = \frac{\alpha_i^3}{24}.$$

Thus, $\sum_{i=2}^{k} \alpha_i^3 \leqslant 24\epsilon$ and Hölder's inequality yields:

$$1 \leqslant \sum_{i=2}^{k} \alpha_i \leqslant \left(\sum_{i=2}^{k} \alpha_i^3\right)^{\frac{1}{3}} \left(\sum_{i=2}^{k} 1^{\frac{3}{2}}\right)^{\frac{2}{3}} \leqslant (24\epsilon)^{\frac{1}{3}}(k-1)^{\frac{2}{3}}.$$

The statement follows. □

Remark. This example involves an infinite covering family, but if we let the disks intersect the x-axis on arbitrarily small lengths, the same can easily be achieved with a finite family.

11.2.2 Arbitrary dimension

Upper bound. We start with a generalization of Lemma 93.

Lemma 95. *Let Y be a ball of radius $r < 1$ and $\mathcal{F}$ a covering of a unit ball $U \subset \mathbb{R}^d$ by unit balls. Then $U \cap Y$ can be covered by a $(d+1)$-tuple $C(Y) \subset \mathcal{F}$ and a collection $R(Y)$ of $O(r^{1-d})$ balls of radius $\rho(r) = O(r^2)$.*

Proof. Given two balls X and Y, we denote by X^Y the half-space containing X and bounded by the hyperplane tangent to X at the projection of the center of Y on ∂X. Notice that this is well defined whenever X and Y have distinct centers. We call $\mathcal{F}^Y$ the collection of all X^Y for X in $\mathcal{F}$.

Let Y be some ball. If a ball of $\mathcal{F}$ has the same center as Y then it covers Y and we are done. We can then assume that it is not the case. Since $\mathcal{F}$ covers U, $\mathcal{F}^Y$ also covers U and in particular it covers $Y \cap U$ and by Helly's theorem there are $d+1$ elements in $\mathcal{F}^Y$ that cover $Y \cap U$; we denote by $C(Y)$ the corresponding $d+1$ balls in $\mathcal{F}$. If the radius of Y is $r < 1$ then the region $\left(X^Y \cap Y\right) \setminus (X \cap Y)$ is included in a cylinder defined by a $(d-1)-$dimensional ball of radius r and an orthogonal segment of length $O(r^2)$. This region can thus be covered by a collection $R_X(Y)$ of $O(r^{1-d})$ balls of radius $\rho(r) = O(r^2)$. Covering the $d+1$ regions corresponding to the $d+1$ balls $X \in C(Y)$ gives a collection $R(Y) = \bigcup_{X \in C(Y)} R_X(Y)$ which concludes the proof. □

The proof of Theorem 92 for a general d follows the same approach as in the case $d = 2$ so we omit the details of the computations.

Proof of Theorem 92. Let $\mathcal{F}$ be a covering of a unit ball U by unit balls in $\mathbb{R}^d$. Let $0 < r_0 < 1$ be small enough so that there exists a constant $K > 0$ such that for any $0 < r \leqslant r_0$, $\rho(r) \leqslant K \cdot r^2$ for the function ρ introduced in Lemma 95. Call $C = Kr_0$. We further assume that r_0 is small enough so that $0 < C < 1$. Again, we construct a small ϵ-covering from $\mathcal{F}$ by starting with a covering $\mathcal{R}_0$ of U by $O(r_0^{-d})$ balls of radius r_0, setting $\mathcal{C}_0 = \emptyset$ and iterating:

$$\mathcal{C}_{i+1} \leftarrow \mathcal{C}_i \cup \left(\bigcup_{Y \in \mathcal{R}_i} C(Y) \right) \quad \text{and} \quad \mathcal{R}_{i+1} \leftarrow \bigcup_{Y \in \mathcal{R}_i} R(Y),$$

After k iterations, $\mathcal{C}_k$ has size at most $O(C^{(1-d)2^{k-1}} D^{k-1})$ (where D is a positive constant) and covers U except for the region covered by the balls in $\mathcal{R}_k$, which consists of $O(C^{(1-d)2^k} D^k)$ balls of radius $O(C^{2^k})$. The volume possibly not covered by the balls in $\mathcal{C}_k$ is thus $O(C^{(1-d)2^k + d2^k} D^k) = O(C^{2^k} D^k))$. By choosing k such that $\epsilon = C^{2^k} D^k$, we get an $O(\epsilon)$-cover of U of size $O(\epsilon^{\frac{1-d}{2}} \operatorname{polylog}(\frac{1}{\epsilon}))$. □

Note that the constant hidden in the $O()$ notation depends on d.

Lower bound. To generalize the lower bound we use the following lemma:

Lemma 96. *Let p be a point and R a convex region of volume v of $\mathbb{R}^d$. Let δ be the distance from p to its furthest point in R. The part of R at distance larger than $\delta/2$ from p has volume $\Omega(v)$.*

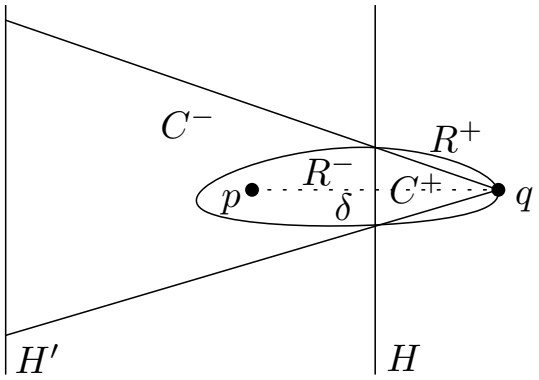


Figure 11.4: For the proof of Lemma 96.

Proof. We refer to Figure 96. We call $q \in R$ the furthest point from p (or one of them). Let H be the hyperplane that consists of points equidistant from p and q and let H' be the hyperplane parallel to H at distance 2δ from q and δ from p.

H intersects R in a convex set U. We draw the half-cone C centered at q and whose intersection with H is U. The part of R at distance larger than $\delta/2$ from p contains the part R^+ of R that is on the same side of H as q. Furthermore, R^+ contains the part C^+ of C on the same side of H as q. The part R^- of R on the other side of H is contained in the region $\mathcal{C}^-$ delimited by $\mathcal{C}$, H and H'. Since the volumes of C^+ is equal to $4^d - 1$ time that of C^-, the statement follows. □

We can now prove that the bound of Theorem 92 is optimal up to the logarithmic factor.

Theorem 97. *There exists a family $\mathcal{F}$ of unit balls in $\mathbb{R}^d$ covering a unit ball $U \subset \mathbb{R}^d$ such that for arbitrarily small $\epsilon > 0$ any subset of $\mathcal{F}$ that is an ϵ-covering of U has size $\Omega(\epsilon^{-\frac{d-1}{2}})$.*

Proof. Let H be an hyperplane through the center of U, let B be the unit ball $U \cap H$ and let $\mathcal{F}$ denote the set of all unit balls tangent to H in a point of B. Observe that $\mathcal{F}$ covers U. We assume that H is given by $x_d = 0$ and, to simplify the description, consider it to be horizontal.

The portion of U on one side of H is covered by the balls of $\mathcal{F}$ that are on that side of H. We thus only argue about the portion $U^+$of U above H and covered by the set $\mathcal{F}^+$ of balls in $\mathcal{F}$ above H. We denote by ∂U^+ the part of the boundary of U above H.

Let $\mathcal{G}^+ \subset \mathcal{F}^+$ be a family of k balls. For each ball $X \in \mathcal{G}^+$, let P_X denote the parabola with equation $2 * x_d = \sum_{i=1}^{d-1}(x_i - t_i)^2$ where $(t_1, ..., t_{d-1}, 0)$ is the tangency point of X with H. Since X is completely above the parabola P_X (see Figure 11.5), the volume of U not covered by $\mathcal{G}^+$ is bounded from below by the volume of the region above B and under the parabolas and ∂U^+.

Let $\mathcal{T}^+$ denote the set of tangency points of G^+ on H. The height of the lowest parabola above a point p in B is proportional to the square of the distance from p to the

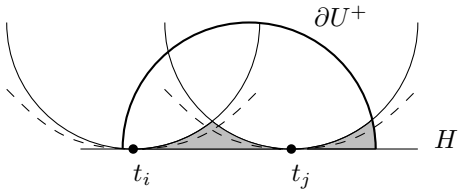


Figure 11.5: Approximation of a sphere by a parabola.

closest point in $\mathcal{T}^+$. Let C be a cell of the Voronoï diagram of $\mathcal{T}^+$ restricted to B and let v denote its volume. The diameter of C is $\Omega(v^{\frac{1}{d-1}})$ and, by Lemma 96, a subset of C of volume $\Omega(v)$ is at distance $\Omega(v^{\frac{1}{d-1}})$ from its center in $\mathcal{T}^+$. The volume between this cell and the parabola above it is thus $\Omega(v^{1+\frac{2}{d-1}})$. Since the cells partition B, the sum of their volumes is $\Omega(1)$ and Hölder's inequality yields:

$$\Omega(1) = \sum_{i=1}^{k} v_i \leqslant \left(\sum_{i=1}^{k} v_i^{1+\frac{2}{d-1}}\right)^{\frac{d-1}{d+1}} \left(\sum_{i=1}^{k} 1^{\frac{d+1}{2}}\right)^{\frac{2}{d+1}} = \left(\sum_{i=1}^{k} v_i^{1+\frac{2}{d-1}}\right)^{\frac{d-1}{d+1}} k^{\frac{2}{d+1}}.$$

Hence, the volume below the parabolas is $\Omega(k^{-\frac{2}{d-1}})$. To take ∂U^+ into account, we consider the ball B' obtained by scaling B by a factor $\frac{1}{2}$. The previous argument still yields that the volume between B' and the parabolas is $\Omega(k^{-\frac{2}{d-1}})$. Also, above any point in B', the ratio of the height of the lowest parabola to that of ∂U^+ is bounded. Thus, the volume above B' and below the parabolas and ∂U^+ is $\Omega(k^{-\frac{2}{d-1}})$. It follows that the volume not covered by $\mathcal{G}^+$ is $\Omega(k^{-\frac{2}{d-1}})$. Equivalently, any subset of $\mathcal{F}^+$ leaving a volume at most ϵ of U^+ uncovered has size $\Omega(\epsilon^{\frac{1-d}{2}})$. □

11.2.3 Algorithm

To turn the proof of Theorem 92 into an algorithm, it suffices to describe how, given a ball Y, the sets $C(Y)$ and $R(Y)$ defined in Lemmas 93 and 95 can be computed. We consider d as a constant.

Recall that $\mathcal{F}$ is a covering of a unit ball U by unit balls, and that Y is some ball with radius $r < 1$. After computing $\mathcal{F}^Y$, in $O(|\mathcal{F}|)$ time, the problem becomes to find a $(d+1)$-tuple in $\mathcal{F}^Y$ that covers $U \cap Y$. This can be done by turning it into a LP-type problem, a special class of optimization problems introduced by Sharir and Welzl [100].

Specifically, let $\phi : 2^{\mathcal{F}^Y} \to \mathbb{R}$ be the map that associates to $\mathcal{G} \subset \mathcal{F}^Y$ the real

$$\phi(\mathcal{G}) = \min\left\{t \in [0, +\infty) \middle| \cup_{x \in \mathcal{G}} x \oplus D(t) \text{ covers } Y \cap U\right\}$$

where $\oplus$ and $D(t)$ denote respectively the Minkowski sum operator and the disk of radius t centered at the origin. $(\mathcal{F}^Y, \phi)$ is a LP-type problem of dimension $d+1$, and thus there exists a subset $\mathcal{H} \subset \mathcal{F}^Y$ of size $d+1$ such that $\phi(\mathcal{H}) = \phi(\mathcal{F}^Y)$; thus, if $\mathcal{F}^Y$ covers $U \cap Y$ so does $\mathcal{H}$. Moreover, $\mathcal{H}$ can be computed in time $O(|\mathcal{F}^Y|)$ (using e.g. the algorithm of Seidel [97]) and if $\mathcal{F}^Y$ does not cover $U \cap Y$, then the hyperplanes bounding the half-spaces

$$\left\{x \oplus D(\phi(\mathcal{H})) \middle| x \in \mathcal{H}\right\}$$

intersect in a point not covered by $\mathcal{F}^Y$. As a result, we obtain:

Theorem 98. *Let $\mathcal{F}$ be a covering of a unit ball $U \subset \mathbb{R}^d$ by unit balls. We can compute a point in U not covered by $\mathcal{F}$ or an ϵ-cover of U of size $O\left(\epsilon^{\frac{1-d}{2}} \operatorname{polylog} \frac{1}{\epsilon}\right)$ contained in $\mathcal{F}$ in time $O\left(|\mathcal{F}|\epsilon^{\frac{1-d}{2}} \operatorname{polylog}(\frac{1}{\epsilon})\right)$.*

This (standard) formulation was obtained using Amenta's technique [9, 10] for relating Helly-type theorems and LP-type problems. For the sake of completeness, the details are presented in Appendix 11.5.1.

11.3 Extensions

11.3.1 Covering by other shapes

11.3.1.1 Smooth convex sets

Lemma 93 requires that (i) given a disk Y, the set $U \cap Y$ be convex and that (ii) the difference between $X^Y \cap Y$ and $X \cap Y$ can be covered by $O(\frac{1}{r})$ balls of radius $O(r^2)$. If an object is convex and its boundary has a curvature of bounded norm, then for any point M on this boundary the object contains a ball (of radius bounded away from 0) and is contained in a half-space delimited by a hyperplane tangent to both the object and the ball in M; this means that covering the region between the ball and the hyperplane is enough to cover the region between the object and the hyperplane. Theorem 92 thus extends to:

Corollary 99. *Let U be a bounded convex set in $\mathbb{R}^d$ and $\mathcal{F}$ a covering of U by smooth convex sets whose curvatures have a norm at most γ. For any $\epsilon > 0$, the smallest subset of $\mathcal{F}$ that is an ϵ-covering of U has size $O\left(\epsilon^{\frac{1-d}{2}} \operatorname{polylog} \frac{1}{\epsilon}\right)$.*

Note that the constant in the $O()$ depends on γ and d.

11.3.1.2 Boxes

While our proof technique requires the covering objects to be smooth, similar results hold in the non-smooth case. We illustrate this with the example of the covering of a unit box (i.e. a cube with axis-aligned edges) by unit boxes in $\mathbb{R}^d$.

Upper bounds. First, let us observe that a result similar to Theorem 92 holds in arbitrary dimension.

Theorem 100. *Let U be a unit box in $\mathbb{R}^d$ and $\mathcal{F}$ a covering of U by unit boxes. For any $\epsilon > 0$, the smallest subset of $\mathcal{F}$ that is an ϵ-covering of U has size $O\left(\epsilon^{-O(d2^d)}\right)$.*

Sketch of proof. Let $\mathcal{F}$ be a family of unit boxes that covers a unit box U, in $\mathbb{R}^d$. Some box $X \in \mathcal{F}$ covers the center c of U and thus completely covers one of the 2^d half-sized boxes of $\mathbb{R}^d$ defined by U and the axis-parallel hyperplanes through c. By applying this idea recursively, we get after k iterations a set of m elements of $\mathcal{F}$ that cover all but possibly s boxes of volume $\left(2^{-k}\right)^d$, with

$$m = \frac{\left(2^d-1\right)^k - 1}{2^d - 2} \quad \text{and} \quad s = \left(2^d-1\right)^k.$$

Thus, for

$$k \geqslant \log_{\frac{2^d}{2^d-1}} \frac{1}{\epsilon},$$

we obtain a family of size $O\left(\epsilon^{-O(d2^d)}\right)$ that leaves uncovered a volume at most ϵ of U. □

In the planar case, a better bound can be obtained as follows:

Theorem 101. *Let U be a unit square in the plane and $\mathcal{F}$ a finite covering of U by unit squares. For any $\epsilon > 0$, the smallest subset of $\mathcal{F}$ that is an ϵ-covering of U has size $O\left(\frac{1}{\epsilon}\right)$.*

Proof. Let $Y \subset U$ be an axis-aligned rectangle. Since $\mathcal{F}$ covers Y, there is a box X_1 that covers the center of Y. Since $\mathcal{F}$ is finite, we can chose X_1 such that $Y \cap X_1$ is maximal for the inclusion. If X_1 does not have a corner inside Y, X_1 covers at least half of Y. Otherwise, let p be the corner of X_1 that is inside Y. Since $\mathcal{F}$ is finite and completely covers Y, some other square X_2 covers p and together, X_1 and X_2 cover at least half of Y. Let $C(Y)$ denote these two squares (or just X_1 in the first case) and $R(Y)$ consist of the 2 rectangular pieces (at most) of Y not covered by $C(Y)$.

We define $\mathcal{R}_0 = \{U\}$ and $\mathcal{C}_0 = \emptyset$ and iterate:

$$\mathcal{C}_{i+1} \leftarrow \mathcal{C}_i \cup \left(\bigcup_{Y \in \mathcal{R}_i} C(Y) \right) \quad \text{and} \quad \mathcal{R}_{i+1} \leftarrow \bigcup_{Y \in \mathcal{R}_i} R(Y).$$

Since the area not covered by $\mathcal{C}_i$ is halved at every iteration, we get that $\mathcal{C}_i$ is a 2^{-i}-covering of U. Furthermore,

$$|\mathcal{C}_i| \leqslant 2 \sum_{k=0}^{i-1} |\mathcal{R}_k| \leqslant 2 \sum_{k=0}^{i-1} 2^k = 2^{i+1} - 2.$$

The statement follows. □

Lower bound. Applying the same technique as in Lemma 97 to the (infinite) family of squares tangent to one of the diagonals of U (see Figure 11.6) shows that the bound of Theorem 101 is tight. This construction can be generalized to dimension d using the same arguments as for balls, and yields a lower bound of $\Omega(\epsilon^{1-d})$ which does not match our upper bound for $d > 2$.

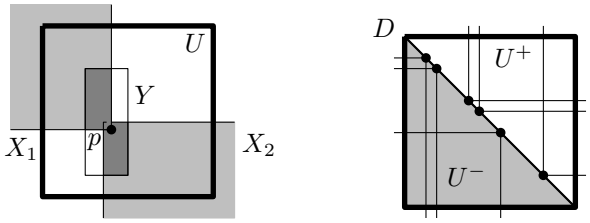


Figure 11.6: Left: For the proof of Theorem 101. Right: Lower bound of $\Omega(\frac{1}{\epsilon})$ for the size of the smallest ϵ-cover of U in $\mathcal{F}$.

Algorithms. The above proof, combined with the technique described in Section 11.2.3, leads to the following algorithm:

Theorem 102. *Given a covering $\mathcal{F}$ of a unit square U by unit squares, we can compute in $O\left(\frac{|\mathcal{F}|}{\epsilon}\right)$-time a point in U not covered by $\mathcal{F}$ or an ϵ-cover of U of size $O\left(\frac{1}{\epsilon}\right)$ contained in $\mathcal{F}$.*

11.3.2 Approximated visibility among 3D unit balls

Two among n objects are *visible* if they support the endpoint of a segment that intersects no other object, and such a segment is called a *visibility segment*. Visibility between objects can be recast as a covering problem by observing that two objects are mutually visible if and only if the set of segments they support is not covered by the set of segments supported by these two objects and intersecting some other object.

A natural "volume" to quantify approximate visibility between two objects – similarly to the ϵ-coverings discussed so far – is given by the measure of the set of lines supporting visibility segments between these two objects. In fact, this corresponds, up to normalization, to the *form factor* used in computer graphics (when constant basis functions are used) to quantify visibility for simulating illumination. We call this measure the *amount of visibility* between the two objects. Building on Theorem 92, we prove:

Theorem 103. *Let $\mathcal{F} \cup \{A, B\}$ be a collection of disjoint unit balls in $\mathbb{R}^3$ such that A and B are mutually invisible. For any $\epsilon > 0$, there exists a subset $\mathcal{G}_\epsilon \subset \mathcal{F}$, of size $O\left(\epsilon^{-\frac{7}{2}} \operatorname{polylog} \frac{1}{\epsilon}\right)$, such that the amount of visibility between A and B in $\mathcal{G}_\epsilon \cup \{A, B\}$ is $O(\epsilon)$.*

Measure in line space. Recall that there exists, up to scaling by some constant, a unique measure over lines in $\mathbb{R}^3$ that is invariant under rigid motions [95]. We choose the constant such that the set of lines intersecting a unit ball has measure $4\pi^2$.

Let S be a measurable set of lines, let $\vec{S}$ denote its set of directions and, for $u \in \mathbb{S}^2$, let $S(u)$ be the set of lines in S with direction u. Finally, let $|\vec{S}|$ denote the area of $\vec{S}$ (on the unit sphere of directions) and let $|S(u)|$ be the measure of $S(u)$, *i.e.* the area of the intercept of $S(u)$ with a plane orthogonal to u.

Lemma 104. *The measure of a set of lines S is bounded from above by $|\vec{S}| \times \max_u |S(u)|$.*

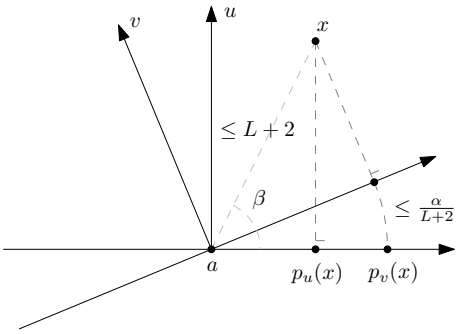


Figure 11.7: How a change of projection affects the projection.

Proof. Let us represent a line by its direction, given in spherical coordinates $(\theta, \phi) \in [0, 2\pi) \times [0, \pi]$, and a point (x, y) in the plane orthogonal to its direction through the origin. With our choice of constant, the density of the measure on the space of lines is then

$$dG = dxdy \sin\theta d\theta d\phi$$

and the statement follows from integrating separately along the couples (x, y) and (θ, ϕ). □

We now prove Theorem 103:

Proof of Theorem 103. Let us fix $\epsilon_0 > 0$. Let a and b be the respective centers of balls A and B. Given $u = (\theta, \phi) \in \mathbb{S}^2$, we denote by $p_u(\cdot)$ the projection on the plane through a with normal u, equipped with a frame with origin at a and with $p_u([0, 2\pi), \phi)$ as x-axis (in the sequel, points $p_u(\cdot)$ are considered in the two-dimensional affine space). The proof consists of four steps:

Step 1. We first find a small subset of $\mathcal{F}$ that blocks visibility between A and B for some given direction $u \in \mathbb{S}^2$. Let $\mathcal{F}_u$ denote the collection of balls that block visibility between A and B along u (*i.e.* a ball X belongs to $\mathcal{F}_u$ if some oriented line with direction u intersects X in-between A and B). Since A and B are mutually invisible, $p_u(\mathcal{F}_u)$ is a collection of unit discs that covers $p_u(A) \cap p_u(B)$. Furthermore, $p_u(A) \cap p_u(B)$ is a bounded convex set. Hence, Corollary 99 yields that for any $\epsilon_0 > 0$, there exists a subset $\mathcal{H}_u \subset \mathcal{F}_u$ of size at most

$$|\mathcal{H}_u| = O\left(\epsilon_0^{-\frac{1}{2}} \operatorname{polylog} \frac{1}{\epsilon_0}\right)$$

such that $p_u(\mathcal{H}_u)$ is an ϵ_0-covering of $p_u(A) \cap p_u(B)$.

Step 2. We now argue that a subset that almost blocks visibility in direction u still almost blocks visibility in any direction v close enough to u. Let $\alpha > 0$ be some constant and $v \in \mathbb{S}^2$ be a vector making, with u, an angle of at most $\frac{\alpha}{L+2}$ where L is the distance

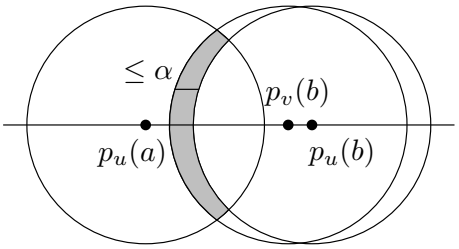


Figure 11.8: Area of the difference between the projections.

between a and b. For any ball $X \in \mathcal{F}_u$, with center x, we have (see Figure 11.7)

$$p_u(x)p_v(x) \leqslant (L+2)\left(\cos\left(\beta - \frac{\alpha}{L+2}\right) - \cos\beta\right) \tag{10}$$

$$\leqslant 2(L+2)\sin\left(\beta - \frac{\alpha}{2(L+2)}\right)\sin\left(\frac{\alpha}{2(L+2)}\right) \leqslant \alpha, \tag{11}$$

since $\sin x \leqslant x$ for $x \leqslant 1$. So, the disk with center $p_u(x)$ and radius $1-\alpha$ is contained in $p_v(X)$. It follows that, for any vector v making angle at most $\frac{\alpha}{L+2}$ with u, $p_v(\mathcal{H}_u)$ covers $p_u(A) \cap p_u(B)$ but an area of at most:

$$\epsilon_0 + 2\pi\alpha|\mathcal{H}_u|.$$

By definition of p_u, we have $p_u(A) = p_v(A)$ and, for the same reason as above, $p_u(b)p_v(b) \leqslant \alpha$. Thus, the area of the difference

$$(p_v(A) \cap p_v(B)) \setminus (p_u(A) \cap p_u(B))$$

is bounded from above by 2α (see Figure 11.8). Hence, $p_v(\mathcal{H}_u)$ covers $p_v(A) \cap p_v(B)$ but an area of at most:

$$\epsilon_0 + 2\alpha + 2\pi\alpha|\mathcal{H}_u|.$$

Note that for a ball $X \in \mathcal{H}_u$, having a non-empty intersection $p_v(A) \cap p_v(B) \cap p_v(X)$ does not guarantee that X blocks visibility between A and B: lines with directions u and v may intersect the three balls in different orders. It thus remains to remove the area covered by $p_v(\mathcal{H}_u \setminus \mathcal{F}_v)$; we claim that this area is $O(\alpha)$ and refer to Appendix 11.5.2 for the details.

Step 3. We now almost block visibility between A and B by applying the previous construction to a sample of $\mathbb{S}^2$. The directions T of common line transversals to A and B make up a disc of radius $\arcsin\left(\frac{2}{L}\right)$ on $\mathbb{S}^2$. We can thus choose a collection D of $O\left(\alpha^{-2}\right)$ directions such that the discs of radii $\frac{\alpha}{L+2}$ centered on these directions completely cover T. Let $\mathcal{H}$ and h denote respectively:

$$\mathcal{H} = \bigcup_{u \in D} \mathcal{H}_u \quad \text{and} \quad h = \max_{u \in \mathbb{S}^2} |\mathcal{H}_u| = O\left(\epsilon_0^{-\frac{1}{2}} \operatorname{polylog} \frac{1}{\epsilon_0}\right).$$

$\mathcal{H}$ has size $O(\alpha^{-2}h)$ and, for any $u \in \mathbb{S}^2$, $p_u(\mathcal{H} \cap \mathcal{F}_u)$ covers $p_u(A) \cap p_u(B)$ except an area of at most:

$$\epsilon_0 + O(\alpha) + 2\pi\alpha h.$$

Let V denote the set of lines intersecting A and B and no ball in $\mathcal{H}$ between A and B. Lemma 104 yields that the measure of V is bounded from above by:

$$(\epsilon_0 + O(\alpha) + 2\pi\alpha h)\,\pi \arcsin^2\left(\frac{2}{L}\right) = O(\epsilon_0 + \alpha h).$$

Step 4. We now have a set $\mathcal{H}$ of size $O\left(\alpha^{-2}\epsilon_0^{-\frac{1}{2}} \operatorname{polylog} \frac{1}{\epsilon_0}\right)$ that blocks visibility between A and B up to a set of lines of measure $V = O(\epsilon_0 + \alpha h)$. By choosing $\alpha = \epsilon_0^{\frac{3}{2}}$, we get

$$|\mathcal{H}| = O\left(\epsilon_0^{-\frac{7}{2}} \operatorname{polylog} \frac{1}{\epsilon_0}\right) \quad \text{and} \quad V = O(\epsilon_0 + \alpha h) = O\left(\epsilon_0 \operatorname{polylog} \frac{1}{\epsilon_0}\right).$$

Finally, setting ϵ_0 such that $\epsilon = \epsilon_0 \operatorname{polylog} \frac{1}{\epsilon_0}$, $\mathcal{H}$ is a subset of $\mathcal{F}$ of size $O\left(\epsilon^{-\frac{7}{2}} \operatorname{polylog} \frac{1}{\epsilon}\right)$, such that the amount of visibility between A and B in $\mathcal{H} \cup \{A, B\}$ is $O(\epsilon)$. □

Algorithm. Combining the proofs of Theorem 98 and Theorem 103, we obtain:

Theorem 105. *Let $\mathcal{F}$ be a collection of disjoint unit balls in $\mathbb{R}^3$ and let A and B be two unit balls. We can compute in $O\left(|\mathcal{F}|\epsilon^{-\frac{7}{2}} \operatorname{polylog}(\frac{1}{\epsilon})\right)$-time a visibility segment between A and B or a subset $\mathcal{G}_\epsilon \subset \mathcal{F}$, of size $O\left(\epsilon^{-\frac{7}{2}} \operatorname{polylog} \frac{1}{\epsilon}\right)$, such that the amount of visibility between A and B in $\mathcal{G}_\epsilon \cup \{A, B\}$ is $O(\epsilon)$.*

11.4 Conclusion

In this paper, we considered the problem of, given a covering $\mathcal{F}$ of a set U, bounding the size of the smallest ϵ-covering $\mathcal{G} \subset \mathcal{F}$ of U. We showed that this size can be bounded independently of $|\mathcal{F}|$ for coverings of a unit ball by unit balls, for coverings of a unit box by unit boxes and for 3-dimensional visibility among disjoint unit balls. We also presented simple algorithms that, given a family $\mathcal{F}$ and a set U, certify either that $\mathcal{F}$ does not cover U or that $\mathcal{F}$ misses at most a volume ϵ of U.

Our bounds show that the size of the smallest ϵ-covering depends on the geometry of the covering sets: our (almost) tight bounds for disks and squares differ by $\Theta(\sqrt{\epsilon})$. Do other simple shapes lead to different bounds? Another obvious open problem is closing the gap between our lower and upper bounds for covering by boxes in dimension $d \geqslant 3$.

11.5 Appendix

11.5.1 For the proof of Theorem 98

We first recall some basic facts on the class of optimization problems called LP-type problems (or *generalized linear programming*). Let $\mathcal{H}$ be a set and ϕ a map $\phi : 2^{\mathcal{H}} \to \Omega$

from the family of subsets of $\mathcal{H}$ to some completely ordered set Ω. The pair $(\mathcal{H}, \phi)$ is a LP-type problem if it satisfies two properties:

Monotonicity: if $F \subset G \subset \mathcal{H}$ then $\phi(F) \geqslant \phi(G)$.

Locality: if $F \subset G \subset \mathcal{H}$ and $\phi(F) = \phi(G)$ then for any $x \in \mathcal{H}$:

$$\phi(F \cup \{x\}) \neq \phi(F) \Leftrightarrow \phi(G \cup \{x\}) \neq \phi(G).$$

A subset $B \subset F$, such that $\phi(B) = \phi(F)$, which is minimal for this property is a *basis* of F. The *combinatorial dimension* of a LP-type problem is the maximal cardinality of a basis, possibly $+\infty$. Recall that for LP-type problem with constant combinatorial dimension, a basis B of $\mathcal{H}$ can be computed in $O(|\mathcal{H}|)$-time (see [9, Chapter 7]).

We now prove the case $d = 2$ of Theorem 98. We start by a simple lemma:

Lemma 106. *Let $\mathcal{H}$ be a family of half-spaces in $\mathbb{R}^d$ and Y a ball. We can compute, in $O(|\mathcal{H}|)$-time, either a $(d+1)$-tuple in $\mathcal{H}$ that covers Y or a point in Y not covered by $\mathcal{H}$.*

Proof. Let $\phi : 2^{\mathcal{H}} \to \mathbb{R}$ be the map that associates to $\mathcal{G} \subset \mathcal{H}$ the real

$$\phi(\mathcal{G}) = \min \left\{ t \in [0, +\infty) \middle| \cup_{x \in \mathcal{G}} x \oplus D(t) \text{ covers } Y \right\}$$

where $\oplus$ and $D(t)$ denote respectively the Minkowski sum operator and the disk of radius t centered at the origin. The problem $(\mathcal{H}, \phi)$ is clearly a LP-type problem. Furthermore, Helly's theorem implies that its combinatorial dimension is bounded, more precisely by $d + 1$, and a basis B of $\mathcal{H}$ can be computed in $O(|\mathcal{H}|)$ time. If $\phi(B) = 0$ then B is a $(d+1)$-tuple in $\mathcal{H}$ that covers Y, otherwise $\mathcal{H}$ does not cover Y. In the latter case, observe that the boundaries of the half-spaces $x \oplus D(\phi(B))$ intersect in a point that is not covered by $\cup_{x \in \mathcal{H}} x$. □

From there, the proof of Theorem 98 is almost immediate:

Proof of Theorem 98. We construct the sets $\mathcal{C}_i$ and $\mathcal{R}_i$ by repeating, as indicated in the proof of Theorem 92, the operation:

$$\mathcal{C}_{i+1} \leftarrow \mathcal{C}_i \cup \left(\bigcup_{Y \in \mathcal{R}_i} C(Y) \right) \quad \text{and} \quad \mathcal{R}_{i+1} \leftarrow \bigcup_{Y \in \mathcal{R}_i} R(Y).$$

Assume we are given $\mathcal{C}_i$ and $\mathcal{R}_i$. For every ball $Y \in \mathcal{R}_i$ we run the algorithm described in Lemma 106 and obtain either a point in $Y \cap U$ not covered by $\mathcal{F}$ or a family $C(Y)$; in the former case we stop and return that U is not covered and in the latter, we compute $R(Y)$. Overall, the time spent for computing $\mathcal{C}_i$ and $\mathcal{R}_i$ is respectively $O\left(|\mathcal{C}_i| * |\mathcal{F}|\right)$ and $O(|\mathcal{R}_i|)$. Let k denote the number of iterations performed. Since we need not compute $\mathcal{R}_k$, the complexity of the algorithm is $O\left(|\mathcal{C}_k| * |\mathcal{F}| + |\mathcal{R}_{k-1}|\right)$; with the same convention as in the proof of Theorem 92, $|\mathcal{C}_k|$ is $O(\epsilon^{\frac{1-d}{2}} \operatorname{polylog}(\frac{1}{\epsilon}))$ and $|\mathcal{R}_{k-1}|$ is $O\left(\epsilon^{\frac{1-d}{2}}\right)$ so the time complexity of the algorithm is $O\left(|\mathcal{F}|\epsilon^{\frac{1-d}{2}} \operatorname{polylog}(\frac{1}{\epsilon})\right)$. □

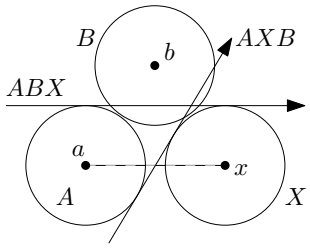


Figure 11.9: Two different geometric permutations.

11.5.2 For the proof of Theorem 103

In this appendix we prove the following claim used in the proof of Theorem 103.

Claim. The area covered by $p_v(\mathcal{H}_u \setminus \mathcal{F}_v)$ is $O(\alpha)$.

Proof: First, observe that if a ball X is in $\mathcal{F}_u \setminus \mathcal{F}_v$ and is such that $p_v(X) \cap p_v(A) \cap p_v(B) \neq \emptyset$ then the balls $\{A, B, X\}$ have two distinct geometric permutations (along direction u we have AXB whereas along direction v the permutation is ABX or XAB). Since these are disjoint unit balls, the centers of two of them are separated by a distance of at most $2\sqrt{2}$ (see Figure 11.9). If these two balls are A and B then the theorem holds since they have at most a constant number of blockers. Otherwise, an immediate packing argument yields that at most a constant number, say c_1, of balls in $\mathcal{F}_u \setminus \mathcal{F}_v$ contribute to cover $p_v(A) \cap p_v(B)$. Also, there is some direction w in the interval $[u, v]$ such that $p_w(X)$ is tangent to $p_w(Y)$ with $Y \in \{A, B\}$. Since

$$p_v(x)p_v(y) \geqslant p_w(x)p_w(y) - p_w(x)p_v(x) - p_w(y)p_v(y) \geqslant 2 - 2\alpha$$

the area $p_v(X) \cap p_v(Y)$ is bounded from above by

$$2\int_{1-\alpha}^{1} \sqrt{1-x^2}dx,$$

which is, at most, 2α (since $\sqrt{1-x^2} \leqslant 1$ on $[1-\alpha, 1]$). This also bounds the contribution of $p_v(X)$ in covering $p_v(A) \cap p_v(B)$ and the claim follows. □

Part V

Travaux annexes

Chapter 12

Octrees with near optimal cost for ray-shooting

Ce travail a été réalisé avec H. Brönnimann. Des versions préliminaires ont été présentées à la Canadian Conference on Computational Geometry [31] et au Latin American Theoretical Informatics Symposium [30]. La version complète est publiée dans Computational Geometry: Theory and Applications [64].

Predicting and optimizing the performance of ray shooting is a very important problem in computer graphics due to the severe computational demands of ray tracing and other applications, e.g., radio propagation simulation. Aronov and Fortune were the first to guarantee an overall performance within a constant factor of optimal in the following model of computation: build a triangulation compatible with the scene, and shoot rays by locating origin and traversing until hit is found. Triangulations are not a very popular model in computer graphics, but space decompositions like kd-trees and octrees are used routinely. Aronov and coll.[13] developed a cost measure for such decompositions, and proved it to reliably predict the average cost of ray shooting.

In this paper, we address the corresponding optimization problem on octrees with the same cost measure as the optimizing criterion. More generally, we solve the generalization for generalized octrees in any d dimensions with scenes made up of $(d-1)$-dimensional simplices. We give a construction of trees which yields cost $O(M)$, where M is the infimum of the cost measure on all trees. Sometimes, a balance condition is important (informally, balanced trees ensures that adjacent leaves have similar size): we also show that rebalancing does not affect the cost by more than a constant multiplicative factor. These are the first and only known results that provide performance guarantees on the approximation factor for 3-dimensional ray shooting with this realistic model of computation. Our results have been validated experimentally by Aronov and coll.[11].

12.1 Introduction

Given a set S of objects, called a *scene*, the ray-shooting problem asks, given a ray, what is the first object in S intersected by this ray. Solving this problem is essential in answering visibility queries. Such queries are used in computer graphics (e.g., ray

tracing and radiosity techniques for photo-realistic 3D rendering[62]), radio- and wave-propagation simulation[22], and a host of other practical problems.

A popular approach to speed up ray-shooting queries is to construct a space decomposition such as a quadtree in 2D or an octree in 3D. The query is then answered by traversing the leaves of the tree as they are intersected by the ray, and for each cell in turn, testing for an intersection between the ray and the subset of objects intersecting that cell. The performance of such an approach greatly depends on the quality of that space decomposition.

Unfortunately, not much is understood about how to measure this quality. Practioners use a host of heuristics and parameters of the scene, of which the object count is less important than, e.g., the size of the objects in the scene, and other properties of the object distribution (density, depth complexity, surface area of the subdivision). Those parameters are used to develop automatic termination criteria for recursively constructing the decompositions (see Section 12.3). While they perform acceptably well most of the time, none of these heuristics performs better than the brute-force method in the worst case. More importantly, occasionally the termination criteria will produce a bad decomposition, and in any case there is no way to know the quality of the decomposition because lower bounds are hard to come by.

In[13], we proposed a measure for bounded-degree space decompositions, based on the surface area heuristic. This cost measure is a simplification of a more complicated but theoretically sound cost measure: under certain assumptions on the ray distribution, the more complicated cost measure provably reflects the cost of shooting an average ray using the space decomposition. The simplified cost measure has a very similar predicting power for many scenes encountered in computer graphics. This has been experimentally verified[13, 11]. We conjecture it would take a very artificially constructed and unrealistic scene to bring forth a discrepancy. It thus makes sense to try and optimize the data structure with respect to this simplified cost measure.

In this paper, we are interested in constructing trees with cost as low as possible, with a guaranteed approximation ratio. First, we observe that tree construction heuristics used in computer graphics do not have a bounded approximation ratio. We give and analyze algorithms that produce trees with cost $O(M)$, where M is a lower bound on the cost of any tree for the given scene. We also examine the effect of rebalancing the tree on the cost measure, and prove that rebalancing only increases the cost by a constant multiplicative factor. The only objects we consider are simplices: points and segments inside the unit square $[0,1]^2$ in $\mathbb{R}^2$, or points, segments and triangles inside the unit cube $[0,1]^3$ in $\mathbb{R}^3$, and $(d-1)$-simplices in $\mathbb{R}^d$. We however assume the real-RAM model so as to avoid a discussion on the bit-length of the coordinates. (This is also justified by our application.)

In a follow-up paper[11] to[13], we evaluate empirically several heuristics, including those presented here, to optimize the cost value of an octree for a given scene. Both our algorithm (*3-greedy* in this paper's terminology) and a simpler heuristic (which we call *1-greedy* or *greedy without lookahead*) give the best cost. In fact, we find that all the reasonable variants end up with approximately the same cost which must be within a factor $O(1)$ of optimal. There is no guarantee that this cost is within $(1+\varepsilon)$ of optimal, though. The polynomial-time algorithm suggested in the remark of Section 12.3.4 (full subdivision to depth $\log_2 n + C$ followed by dynamic programming) would be guaranteed

to give a $O(1)$ approximation factor as well, and we'd even expect it to be very close to optimal, but it was too time- and space-consuming for us to try in[11]. Nevertheless, we find it hard to believe that the cost to which all these variants converge could be sub-optimal, and it thus appears experimentally that our procedure yields a near-optimal tree in practice.

Related work. There has been a lot of work on quadtrees and octrees in the mesh generation and graphics community (see the book by Samet[93], the thesis of Moore[84], or the survey by Bern and Eppstein[21] for references). Since they are used for discretizing the underlying space, usual considerations include the tradeoff between the size of the tree and their accuracy with respect to a certain measure (that usually evaluates a maximum approximation error with respect to some surface). These are not usually relevant for ray shooting.

There is, however, a rich history of data-structure optimization for ray shooting in computer graphics. (See the surveys in[13, 11, 34] and refs. therein.) Cost measures have been proposed for ray shooting in octrees by MacDonald and Booth[79], Reinhard and coll.[91], Whang and coll.[108], and for other structures, such as bounding volume hierarchies[66, 103, 107], bintrees[79], BSP-trees[86], uniform grids[38] and hierarchical uniform grids[33]. We should also mention the work of Havran and coll.[69], who propose a method to determine experimentally the most efficient space subdivision for a given scene, by picking a similar scene (according to certain characteristics such as size, number of objects, densities, etc.) in a database of scenes for which the most efficient scheme has been determined. All of these approaches use heuristic criteria (sometimes very effectively) but none offer theoretical guarantees. In particular, the cost function c of MacDonald and Booth is the same as ours (see Section 12.2.1 below) and their greedy procedure almost identical to our 1-greedy heuristic. They do not however derive the theoretically correct formulation c^* nor establish the connection to c, but they do explore a wider parameter space by allowing the subdivision of their bintree to be different than the spatial median (which we do not in this work).

12.2 General cost measure results

12.2.1 Motivation, definitions, and problem statement

In this paper we consider the problem of shooting rays into a scene consisting of solids, represented by their boundaries. In particular, we assume that these boundaries have already been subdivided into elements of constant combinatorial complexity, so that the only objects we consider are simplices of dimension at most $d-1$ in $\mathbb{R}^d$: points and segments inside the unit square $[0,1]^2$ in $\mathbb{R}^2$, points, segments and triangles inside the unit cube $[0,1]^3$ in $\mathbb{R}^3$, and $(d-1)$-simplices in $\mathbb{R}^d$.

The algorithm generally used for ray shooting with the help of a (convex) spatial subdivision $\mathcal{T}$ such as a quadtree (in 2D) or an octree (in 3D) is the following: when shooting a ray, the cells of that subdivision (in our context, the leaves of the octree) are traversed in the order they are encountered by the ray, starting with the cell that contains

the origin of the ray. For each cell traversed, all the objects intersecting that cell are tested against the ray to find the first hit. If a hit is found inside the cell, then the traversal is stopped and the first such hit is returned, otherwise the algorithm proceeds to the next cell in the traversal. If it reaches the outside boundary of the enclosing volume, the ray is declared unoccluded and the algorithm returns "no hit".

The worst-case cost of this algorithm depends on the stabbing number of the subdivision. While it is not hard to fabricate scenes with a worst-case linear time ray, in a typical application of ray shooting such as ray tracing or wave propagation simulation, many rays are shot. The quantity of interest is really the average cost for a random ray following a certain distribution of rays. A common heuristic in computer graphics consists of approaching the density of rays hitting a convex cell by its perimeter area, as motivated by Crofton's formula in integral geometry[95]. In fact, this is rigorous for line traversal if the line distribution is the rigid-motion invariant distribution, but this is ill-defined for rays. In[13], we observed that for a uniform distribution of rays originating either outside the subdivision or on the surface of objects, the density of rays traversing a cell was proportional to the area of the perimeter plus the total area of the perimeter of the objects contained in the cell. The average cost of traversing a cell is proportional to the number of intersection tests performed multiplied by the density, plus some overhead γ for finding the next cell in the traversal,[18] and thus the following cost measure should be considered for modeling the total traversal costs[13]:[19]

$$c_S^*(\mathcal{T}) := \sum_{\sigma \in \mathcal{L}(\mathcal{T})} \big(\gamma + |S_\sigma|\big) \times \big(\lambda_{d-1}(\sigma) + \lambda_{d-1}(S_\sigma \cap \sigma)\big), \tag{12}$$

where $\mathcal{L}(\mathcal{T})$ is the set of leaves of the tree, S_σ is the set of scene objects intersecting a leaf σ, $\lambda_{d-1}(\sigma)$ is the perimeter length (if $d = 2$) or surface area (if $d = 3$) of σ, and $\lambda_{d-1}(S_\sigma \cap \sigma)$ is the perimeter or area of the portion contained in σ of the objects in S_σ. The last term $\lambda_{d-1}(S_\sigma \cap \sigma)$ is somewhat complicated to evaluate, and so the following simplification was introduced in[13]:

$$c_S(\mathcal{T}) := \sum_{\sigma \in \mathcal{L}(\mathcal{T})} \big(\gamma + |S_\sigma|\big) \times \lambda_{d-1}(\sigma). \tag{13}$$

This simplified cost function provably models the cost of finding all the objects intersected by a random line (with respect to the rigid-motion invariant distribution of lines), and was already proposed earlier by MacDonald and Booth[79] but without the connection to c_S^*. MacDonald and Booth[79] (for bintrees) and Aronov and coll.[13, 11] (for bounded-degree spatial subdivisions, with extensive experimentation for octree) argue and provide ample

[18] Technically speaking, the cost functions make sense only for spatial subdivisions with bounded degree, so that the cost of finding the next cell in the traversal can be bounded by a constant γ. But this is also true in an amortized sense for octrees, see[13] for the technical details.

[19] Strictly speaking, the average cost of shooting a ray is $c_S^*(\mathcal{T})/\mu$ where μ is the total density of the ray distribution, i.e. the surface area of the outer perimeter of $\mathcal{T}$ plus the total surface area of the objects in S. For the purpose of optimizing ray shooting costs, however, if the outer bounding box and the scene are given, the denominator is constant and we need only concentrate on finding the best $\mathcal{T}$ optimizing $c_S^*(\mathcal{T})$ or $c_S(\mathcal{T})$.

experimental evidence that this cost function also reflects the average cost of ray shooting using the spatial subdivision induced by the leaves of $\mathcal{T}$.

The problem we then consider is, given a scene S, an overall bounding box $\mathcal{B}$, and a cost function such as Eq. (13), to construct an octree $\mathcal{T}$ subdividing $\mathcal{B}$ which minimizes $c_S(\mathcal{T})$. Note that all the definitions as phrased above are valid for $d \geqslant 4$ as well as for 2D and 3D, and for the sake of generality we consider below the problem for any *fixed* $d \geqslant 2$. Since S and $\mathcal{B}$ are also fixed, we will often omit them in the notation and use simply $c(\mathcal{T})$. In order to obtain theoretical guarantees, we restrain ourselves to split the octree nodes at their spatial median, although others have investigated other choices to heuristically improve costs. Even with this, we do not know if our problem is NP-complete or not. Nevertheless, we are not aware of any work providing optimality or even any sublinear bound on the approximation ratio. Namely, all the known termination criteria for subdividing an octree or other constructions mentioned earlier have no known sublinear bound on their approximation ratio.

12.2.2 Tree and object costs

Observe that the cost measure expressed in Eq. (13) can be decomposed into two terms: $c(\mathcal{T}) = c_t(\mathcal{T}) + c_o(\mathcal{T})$. The first term,

$$c_t(\mathcal{T}) := \gamma \sum_{\sigma \in \mathcal{L}(\mathcal{T})} \lambda_{d-1}(\sigma) = \gamma \lambda_{d-1}(\mathcal{L}(\mathcal{T})), \tag{14}$$

where $\mathcal{L}(\mathcal{T})$ denotes the set of leaves of $\mathcal{T}$ and λ_{d-1} is extended to sets of leaves by summation, is simply the tree total area (which is more than the surface area of the bounding box). We call c_t the *tree cost*. The second term $\sum_{\sigma \in \mathcal{L}(\mathcal{T})} |S_\sigma| \lambda_{d-1}(\sigma)$, can be computed object by object and this leads to

$$c_o(\mathcal{T}) := \sum_{s \in S} \lambda_{d-1}(\mathcal{L}_s(\mathcal{T})), \tag{15}$$

where $\mathcal{L}_s(\mathcal{T})$ denote the set of leaves of $\mathcal{T}$ intersected by s, and again λ_{d-1} is extended to sets of leaves by summation. We call c_o the *object cost*. Again, note that $\lambda_{d-1}(\mathcal{L}_s(\mathcal{T}))$ is more than the perimeter of the union $\cup_{s \in S} \mathcal{L}_s(\mathcal{T})$, as it measures the sum of the perimeters of all the leaves intersected by the object. (In particular, both sides of a face contribute if the face belongs to the interior of $\cup_{s \in S} \mathcal{L}_s(\mathcal{T})$.) It is useful to keep in mind the following simple observations: when subdividing a leaf σ, the total tree cost of its children is twice the tree cost of σ, and the contribution within σ to the object cost of an object is multiplied by $m/2^{d-1}$ where $m \in [1 \dots 2^d]$ is the number of children intersected by the object. Note that $m \leqslant 3$ for a segment in 2D and $m \leqslant 7$ for a triangle in 3D (unless they pass through the center of the cell). As the tree grows finer, the tree cost increases while the object cost presumably decreases.

The following lemma was given in[31] for quadtrees, and rephrased here to encompass any dimension as well.

Lemma 107. *For any $d \geqslant 2$ and any set S of simplices of dimension at most $d-1$ in $[0,1]^d$, $c(\mathcal{T}) \geqslant 2d\gamma + d\sqrt{2} \sum_{s \in S} \lambda_{d-1}(s)$.*

Proof. The tree cost cannot be less than $\lambda_{d-1}([0,1]^d)\gamma = 2d\gamma$, and the object cost cannot be less than $\sum_{s\in S}\lambda_{d-1}(s)$. We can improve this lower bound further by noting that any leaf σ that is intersected by an object s has area at least $d\sqrt{2}$ times $\lambda_{d-1}(s\cap\sigma)$. Indeed, the smallest ratio $\lambda_{d-1}(\sigma)/\lambda_{d-1}(s\cap\sigma)$ happens when s maximizes $\lambda_{d-1}(s\cap\sigma)$; this happens for a diagonal segment of length $\sqrt{2}$ for the unit square (of perimeter 4), and for a maximal rectangular section of area $\sqrt{2}$ for the unit cube (of area 6). In fact, the maximal section of the unit d-cube is $\sqrt{2}$[19], hence the ratio is at least $2d/\sqrt{2} = d\sqrt{2}$ in any dimension. □

12.3 Tree construction schemes

All we have said so far is independent of the particular termination criterion or algorithm used to construct the tree. In this section, we introduce several construction schemes and explore their basic properties.

12.3.1 Terminology and notation

We follow the same terminology as[31], and generalize it to encompass any dimension. For the d-cube $[0,1]^d$ and the cells of the decomposition, we borrow the usual terminology of polytopes (vertex, facet, h-face, etc.). The *square* is a quadtree that has a single leaf (no subdivision), the *cube* is an octree with a single leaf, and the *d-cube* is a single-leaf tree (for any d). We call this tree *unit* and denote it by $\mathcal{T}^{(\text{unit})}$. If we subdivide this leaf recursively until depth k, we get a *complete tree (of depth k)*, denoted by $\mathcal{T}_k^{(\text{complete})}$, and its leaves form a regular d-dimensional grid with 2^k intervals along each direction. Note that the root has depth 0 and its children have depth 1. In a tree, if only the cells incident to one facet (resp. d facets sharing a vertex, or touching any of the $2d$ facets) of a cell are subdivided, and this recursively until depth k, the subtree rooted at that cell is called a *k-side* (resp. *k-corner* and *k-border*) tree, and denoted by $\mathcal{T}_k^{(\text{side})}$ (resp. $\mathcal{T}_k^{(\text{corner})}$ and $\mathcal{T}_k^{(\text{border})}$); see Figure 12.3.1 for an illustration of the 2D case. In higher dimensions, there are other cases (one for each dimension between 1 and $d-2$). Generally speaking, we call the tree obtained by recursively subdividing until depth k all the cells touching the facets adjacent to given h-face a *k-corner of order h* (in dimension d), and denote it $\mathcal{T}_k^{(h,\,d\text{-corner})}$. A k-corner is of a corner of order 0, and a k-side is a corner of order $d-1$. All this notation is extended to starting from a cell σ instead of a unit tree, by substituting σ for $\mathcal{T}$: for instance, the complete subtree of depth k subdividing σ is denoted by $\sigma_k^{(\text{complete})}$.

The subdivision operation induces a partial ordering $\prec$ on trees, whose minimum is the unit tree. Again, this partial ordering is extended to subtrees of a fixed cell σ.

We consider algorithms for computing a tree for a given set S of objects, which subdivide each cell recursively until some given termination criterion is satisfied. In particular, we may recursively subdivide the unit cube until each leaf meets at most one object (or any fixed C). We call this the *separation criterion*, and the resulting tree the *minimum separating tree*, denoted $\mathcal{T}^{(\text{sep})}(S)$, with variants where the recursion stops at depth k, denoted $\mathcal{T}_k^{(\text{sep})}(S)$, or $C>1$ denoted $\mathcal{T}_C^{(\text{sep})}(S)$. (Note that the depth of $\mathcal{T}_C^{(\text{sep})}$ is always infinite if more than C simplices intersect.) In 3D, for non-intersecting triangles, a variant

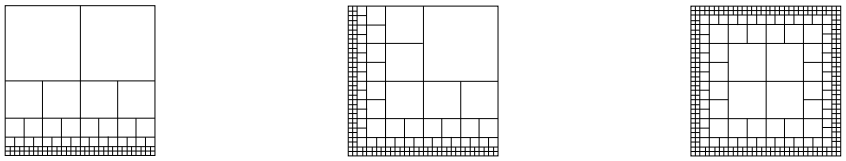


Figure 12.1: The k-side quadtree $\mathcal{Q}_k^{(\text{side})}$ (left), a corner $\mathcal{Q}_k^{(\text{corner})}$ (center), and a border $\mathcal{Q}_k^{(\text{border})}$ (right).

of[14] stops the recursive subdivision also when no triangle edge intersects the leaf (but any number of non-intersecting triangle interiors may slice the leaf). We will not analyze this variant in this paper.

12.3.2 Examples of cost computation

As examples and for completeness, we compute the costs of some of the configurations given above.

With no subdivision, the cost of the unit tree $\mathcal{T}^{(\text{unit})}$ is $c(\mathcal{T}^{(\text{unit})}) = 2d(\gamma + n)$.

For points, the cost of a full subdivision at depth k, $\mathcal{T}_k^{(\text{complete})}$, is at most $2d(2^k\gamma + \frac{2^d n}{2^{k(d-1)}})$ because each point belongs to at most 2^d leaves, and at least $2d(2^k\gamma + \frac{n}{2^{k(d-1)}})$. The latter is the exact value if all the points fall in a single leaf of $\mathcal{T}_k^{(\text{complete})}$, and it happens if the binary expansions of all point coordinates always have at least $k+1$ bits.

The tree costs of the k-side, k-corner, and k-border, and more generally of the k-corner of order h can be computed readily by noting that a k-side is a half-scaled copy of 2^{d-1} empty trees and 2^{d-1} $(k-1)$-sides, that a corner is made of $\binom{d}{i}$ half-scaled k-corners of order i, for every $0 \leqslant i \leqslant d-1$, and that a k-border is made of 2^d half-scaled $(k-1)$-corners:

$$\begin{aligned}
c_t(\mathcal{T}_k^{(\text{side})}) &= 2d\gamma(k+1), \\
c_t(\mathcal{T}_k^{(\text{corner})}) &= \frac{2d\gamma}{(2^{d-1}-1)^2} \times \left(k\left(2^{2d-1} - 2^{1+d} + 2\right) - 2^{d-1} + 1 + \frac{2^{d-1}-1}{2^{k(d-1)}} \right), \\
c_t(\mathcal{T}_k^{(\text{border})}) &= 2c_t(\mathcal{T}_{k-1}^{(\text{corner})})
\end{aligned}$$

Let S_n denote n distinct points very close to one corner of the unit cell, let's say the origin. Here, 'very close' means always within the cell incident to that vertex. It's then pointless to subdivide the other cells: since they do not contain points of S, their cost would be doubled. After k levels of recursively subdividing the incident cell, we obtain the tree $\mathcal{T}_k^{(\text{sep})}(S_n)$ whose cost is $c(\mathcal{T}_k^{(\text{sep})}(S_n)) = 12\gamma + (n - 2\gamma)2^{2-k}$ for $d = 2$; for any d, the cost is

$$c(\mathcal{T}_k^{(\text{sep})}(S_n)) = \frac{2dn}{2^{(d-1)k}} + 2d\gamma\left(1 + \frac{1 - 2^{(1-d)k}}{1 - 2^{1-d}}\right)$$

Whether this is an improvement over $\mathcal{T}^{(\text{unit})}$ for any value of k depends on n and γ. In particular, $\mathcal{T}_k^{(\text{sep})}(S_n)$ has cost lower than the unit tree for large values of k only if

$n > 2\gamma$ in 2D, and $n > \frac{4}{3}\gamma$ in 3D ($n > \frac{2^{d-1}}{2^{d-1}-1}\gamma$ in general). This example tells us that whether subdivision strategies based on the number of objects in a cell—like the separation criterion—produce optimal or near-optimal trees, depends strongly on the value of γ.

12.3.3 Dynamic programming and greedy strategies

As introduced in[31], the *dynamic programming algorithm* finds the tree that minimizes the cost over all trees with depth at most k, which we denote by $\mathcal{T}_k^{(\text{opt})}(S)$ (or $\sigma_k^{(\text{opt})}(S)$ if we start from a cell σ instead of the unit cell): the algorithm starts with the complete tree $\mathcal{T}_k^{(\text{complete})}$, and simply performs a bottom-up traversal of all the nodes, while maintaining the optimum cost of a tree rooted at that node. The decision whether to keep the subtree of a cell or prune it is based on the cost of the cell vs. the sum of the optimum costs of the subtrees rooted at its 2^d children.

Unfortunately, the memory requirements of this algorithm are huge for large values of k (although they remain polynomial if $k = \Theta(\log n)$; see next section). Therefore we also propose a greedy strategy with bounded lookahead: the algorithm proceeds by recursively subdiving the nodes with a greedy termination criterion: when examining a cell σ, we run the dynamic programming within σ with depth p (p is a parameter called *lookahead*). If the best subtree $\sigma_p^{(\text{opt})}(S)$ does not improve the cost of the unsubdivided node σ, then the recursion terminates. Otherwise, we replace σ by the subtree $\sigma_p^{(\text{opt})}(S)$ and recursively evaluate the criterion for the leaves of $\sigma_p^{(\text{opt})}(S)$. We call this the p-greedy strategy and denote the resulting tree by $\mathcal{T}^{(p\text{-greedy})}(S)$ (or $\sigma^{(p\text{-greedy})}(S)$ if we start from a cell σ instead of the unit cube). Note that unlike all the other trees constructed up to now, that tree could be infinite. We use the notation $\mathcal{T}_k^{(p\text{-greedy})}$ to denote the tree constructed with the p-greedy lookahead criterion combined with a maximum depth of k.

With one level of lookahead ($p = 1$), the *greedy strategy* simply examines whether subdivision at one level decreases the cost measure. Below, we show that this does not always yield good trees. We will analyze the greedy strategies with one or more levels of lookahead, first for points, then for simplices. But first, we must grapple with the issue of infinite depth.

12.3.4 Pruning beyond a given depth

The "optimal" tree may not have finite depth: it is conceivably possible to decrease the cost by subdividing ad infinitum. Indeed, this is the case for $n > 2\gamma$ in the example of $\mathcal{T}_k^{(\text{sep})}(S_n)$ given at the end of section 12.3.2. So we let M denote the infimum of $c(\mathcal{T})$ over all trees $\mathcal{T}$ for a scene S. (As a consequence of Lemma 107, $M \geqslant 2d\gamma$.) In order to have an algorithm that terminates, we usually add an extra termination criterion such as a maximum depth k.

We now show that pruning a tree beyond depth k for some choice of k increases the cost at most by a constant factor. We first show it for arbitrary convex obstacles (simplices in particular). Then we improve on the result for the case of points.

Lemma 108. *Let $\mathcal{T}$ be a d-dimensional tree which stores a set S of n convex objects of dimensions at most $d-1$. For any fixed constant C and $k = \log_2 n + C$, let $\mathcal{T}_k$ be the tree*

obtained from $\mathcal{T}$ by removing every cell of depth greater than k. Then $c(\mathcal{T}_k) = O(c(\mathcal{T}))$ with a constant depending on d, C and γ (not on S nor n).

Proof. First, the tree cost only increases when subdividing, so that $c_t(\mathcal{T}_k) \leqslant c_t(\mathcal{T})$. The cells of depth less than k are the same in $\mathcal{T}$ and $\mathcal{T}_k$, hence the object cost of those cells of $\mathcal{T}_k$ is at most the object cost of $\mathcal{T}$. It remains to bound the object cost of the leaves of $\mathcal{T}_k$ that have depth k.

Let us consider an arbitrary object, and since we are only concerned with depth k, let us consider the full subdivision of depth k instead of $\mathcal{T}_k$. We let K be 2^k. This is a d-dimensional grid $\mathcal{G}$ of side K, with K^d cells. Our first purpose is to give an upper bound on the number of cells of this grid that the object can intersect. We could not find a bound in the literature, so we include it here for completeness:

Lemma 109. *Let $\mathcal{G}$ be a $K \times \cdots \times K$ d-dimensional grid subdividing the unit cube, with K^d cells. Any convex object s of dimension at most $d-1$ intersects at most $(d+1) \cdot \left(4dK^{d-2} + K^{d-1}\lambda_{d-1}(s)\right)$ cells of $\mathcal{G}$.*

Proof of Lemma 109. Since the dimension of s is at most $d-1$, it belongs to a hyperplane π: consider the coordinate n_i of largest absolute value of a vector $\mathbf{n}$ normal to π, and project the object s, the vector $\mathbf{n}$, and the grid $\mathcal{G}$ onto a $(d-1)$-dimensional object s', vector $\mathbf{n}'$, and grid $\mathcal{G}'$ in the hyperplane $x_i = 0$. Note that $\lambda_{d-1}(s') \leqslant \lambda_{d-1}(s)$. By our choice of projection, each cell of $\mathcal{G}'$ intersected by s' is the projection of at most $d+1$ cells of $\mathcal{G}$ intersected by s. This is most easily seen by rewriting the equation of π as $x_i n_i + \mathbf{x}'\mathbf{n}' = \pi_0$, or $\Delta x_i n_i + \Delta\mathbf{x}'\mathbf{n}' = 0$, which means that $|\Delta\mathbf{x_i}| \cdot |n_i| \leqslant |\Delta\mathbf{x}' \cdot \mathbf{n}'| \leqslant \sum_{j\neq i} |\Delta x_j| \cdot |n_j| \leqslant |n_i| \sum_{j\neq i} |\Delta x_j|$, and thus when $\Delta\mathbf{x}'$ is in a cell of $\mathcal{G}'$, $|\Delta x_i|$ is at most $d-1$; in the worst case, this can span $d+1$ cells (the bottom and top cells being only touched by a single point). The $(d-2)$-dimensional boundary of s' is also convex, and may intersect at most $K^{d-1} - (K-4)^{d-1} \leqslant 4dK^{d-2}$ cells (by convexity, the worst case occurs when the boundary is largest possible, hence intersects all the border cells of $\mathcal{G}'$ and their neighbors). The relative interior of s' cannot contain more than $K^{d-1} \cdot \lambda_{d-1}(s')$ cells of $\mathcal{G}'$, and the lemma follows. □

Proof of Lemma 108 (cont.) The area of a cell of $\mathcal{G}$ is $2dK^{1-d}$. The cost of the object associated to depth k cells is then at most $2d(d+1)\,(4d/K + \lambda_{d-1}(s))$. Summing over all objects gives a total of at most $2d(d+1)\,(4dn/K + \sum_s \lambda_{d-1}(s))$. Substituting $2^{\log_2 n + C}$ for K shows that the combined object costs of the leaves of $\mathcal{T}_k$ at depth k is at most $2d(d+1)\left(d \cdot 2^{2-C} + \sum_s \lambda_{d-1}(s)\right)$. By Lemma 107, this is at most $\max\left\{\sqrt{2}(d+1), d(d+1)2^{2-C}/\gamma\right\}$ times the cost of $\mathcal{T}$. Putting everything together, $c(\mathcal{T}_k) \leqslant \left(1 + \sqrt{2}(d+1) + d(d+1)2^{2-C}/\gamma\right) c(\mathcal{T})$. □

Remark. A choice of $k = \log_2 n + C$ ensures that $\mathcal{T}_k$ has at most $(2^k)^d = O(n^d)$ leaves, for any fixed d. Hence the algorithm which computes the full subdivision at depth k and then applies the dynamic programming heuristic provably computes a tree whose cost is $O(M)$ in polynomial time, as a consequence of Lemma 108.

Unfortunately, this cannot be extended to yield approximations within $(1+\varepsilon)$. But as a side note, for scenes consisting of points only and with slightly more restrictive

hypotheses on $\mathcal{T}$, the depth k can be chosen so that $\mathcal{T}_k$ has size $O(n^{1+\frac{1}{d-1}}) = O(n^2)$ (for any $d \geqslant 2$) and cost as close as desired to that of $\mathcal{T}$.

Lemma 110. *Let $\mathcal{T}$ be a d-dimensional tree, which stores a set S of n points. Assume that $\mathcal{T}$ does not contain empty internal nodes (i.e. that are subdvided but do not contain any object). Let $\mathcal{T}_k$ be the tree obtained from $\mathcal{T}$ by removing every cell of depth greater than k. Then, for every $\varepsilon > 0$ there exists a C (that depends only on d, ε and γ but not on S nor n) such that, for $k = \frac{1}{d-1}\log_2 n + C$, we have*

$$c(\mathcal{T}_k) \leqslant (1+\varepsilon)c(\mathcal{T}).$$

Proof. The cost of a cell σ which contains n_σ points and has depth $k = \frac{1}{d-1}(\log_2 n + C)$ is $c(\sigma) = (\gamma + n_\sigma)2d \cdot 2^{-k(d-1)} = 2d(\gamma + n_\sigma)/(2^C n)$. The proof for points hinges on the fact that, unlike an arbitrary simplex, a point belongs to at most 2^d leaves,[20] and that to be subdivided, a cell needs to contain at least one point. Hence, in $\mathcal{T}_k$ there are at most $2^d n$ leaves that contain a point, and $\sum_\sigma n_\sigma \leqslant 2^d n$. Summing over all leaves at depth k which still contain points, one gets

$$\sum_{\sigma \in \mathcal{L}_k(\mathcal{T})} c(\sigma) \leqslant \frac{2d}{n}(\gamma + n_\sigma) \cdot \frac{d2^{d+1}(\gamma+1)}{2^C},$$

which can be made as small as $\varepsilon 2d\gamma$ for an appropriate value of C. By Lemma 107, this implies that the cost of the non-empty leaves at depth k in $\mathcal{T}_k$ is at most $\varepsilon c(\mathcal{T})$. The leaves of $\mathcal{T}_k$ at depth less than k also belong to $\mathcal{T}$, and the same holds as well for the leaves at depth k which do not contain any point. Hence their total cost is at most $c(\mathcal{T})$ and the lemma follows. □

This result may seem somewhat anecdotic, but scenes consisting of points only have some relevance (see Section 12.4.2).

12.4 Constructing trees whose costs approach the optimal

12.4.1 General case: simplices

The following lemma was proven in[31] for the case $d = 2$. Its statement and proof extend straightforwardly to higher dimensions.

Lemma 111. *The lookahead greedy strategy does not always give (asymptotically) a cost-optimal tree. Specifically, for any k, there is a set S of n objects such that no tree of depth at most k has cost less than $2d(\gamma + n)$, but some tree of depth at least $k+1$ has cost less than $2d(\gamma + n)$.*

[20] If this seems like nitpicking, consider that "point" might mean a very small simplex. Thus this result for points applies equally to any collection of small objects, where "small" means to be contained in some cube of side 2^{-k}. See also Section 12.4.2.

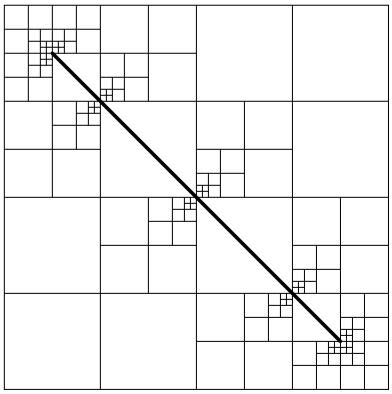


Figure 12.2: An example of quadtree $\mathcal{Q}^{(2)}_{k,m}$; here $m = 2$ and $d = 4$.

The counterexample for $d = 2$ consists[31] of n copies of the segment pq, where $p = (1-2^{-m-1}, 2^{-m-1})$, and $q = (2^{-m-1}, 1-2^{-m-1})$, and considering the cost-optimal quadtree $\mathcal{Q}^{(2)}_{k,m}$ of depth at most k. As long as $k \leqslant m+1$, the situation is similar to the case where pq is the whole diagonal and the cost-optimal quadtree of depth at most k is the square, as can easily be verified. When k becomes larger, however, it becomes more cost-effective to subdivide the corners, as shown by the quadtree in Figure 12.2 whose cost is less than $2d(\gamma + n)$.

Although the lookahead greedy strategy does not produce the optimal solution, in the counter-example above it does give a good approximation. In fact, this can be proven for all scenes.

Theorem 112. *Given a set S of convex objects of dimensions at most $d-1$ in the unit cube, let M be the infimum of $c(\mathcal{T})$ over all trees $\mathcal{T}$ storing S. There is an integer p which depends only on d ($p = 3$ for $d \leqslant 3$) such that the tree $\mathcal{T}^{(p\text{-greedy})}$ constructed by the p-greedy strategy has cost $c(\mathcal{T}^{(p\text{-greedy})}) = O(M)$.*

Proof. The intuition is that small objects behave well, and the cost of a big object is bounded below by a constant times its size so it cannot be reduced by very much. Let us look at a cell σ of the tree $\mathcal{T}^{(p\text{-greedy})}$: we are going to show that, when the optimal decomposition of depth at most p of a cell σ does not improve on the cost of σ, then the cost of σ is $O(M_\sigma)$ where M_σ is the infimum cost of all the possible tree subdivisions of σ. If this holds true for every leaf σ of the p-greedy strategy, then $c(\mathcal{Q}^{(p\text{-greedy})}) = O(M)$ as well. We will need a technical lemma:

Lemma 113. *Let $\mathcal{G}$ be a $K \times \cdots \times K$ d-dimensional grid subdividing the unit cube, with K^d cells. For every d and K, there exist constants $G_d(K)$ and $C_d(K) > 0$ such that for any convex object s of dimension at most $d-1$, either s intersects at most $G_d(K)$ cells of $\mathcal{G}$, or else $\lambda_{d-1}(s) \geqslant C_d(K)$.*

Proof of Lemma 113. In two dimensions, a point intersects at most 4 cells, and any segment intersecting at least five cells of $\mathcal{G}$ must have length at least $1/K$, hence $G_2 = 4$ and $C_2 = 1/K$ suffices. Intuitively, in three dimensions, the maximum a segment can intersect is by joining two opposite endpoints of the unit cube, leading to $G_3(K) = 7K-6$; in order

to intersect any more cells, it must subtend a triangle of height at least half a diagonal of the cells of $\mathcal{G}$, thus have an area of at least $C_3(K) = \sqrt{3} \times (\sqrt{3}/2K)/2 = 3/(4K)$. The same argument in higher dimension leads to $G_d(K) = 2^d(K-1)+1$ and some $C_d(K) > 0$. Unfortunately, a formal proof is much harder. We give an easier proof which leads to slightly worse constants but works in any dimension.

Using the same notation as for Lemma 108, s projects onto some hyperplane $x_i = 0$ into an object s', and s intersects at most $(d+1)$ cells of $\mathcal{G}$ for every cell of $\mathcal{G}'$ intersected by s'. We now argue that in $d-1$ dimensions, either s' entirely contains a cell of $\mathcal{G}'$ or else it is contained in a slab of width $w = 2d\sqrt{d}/K$. Recall that any convex set has a pair of ellipsoids (called the Löwner-John ellipsoids) $e' \subseteq s' \subseteq e''$ such that $s' \subseteq de'$. If the width of s' is at least $2d\sqrt{d}/K$ in any direction, then the width of e' is at least $2\sqrt{d}/K$ in any direction as well and thus e' contains a sphere of diameter $2\sqrt{d}/K$, which in turn must entirely contain a cell of $\mathcal{G}'$ (of diameter $\sqrt{d}/K$). In that case, $\lambda_{d-1}(s) \geqslant \lambda_{d-1}(s') \geqslant \lambda_{d-1}(e') \geqslant 1/K^{d-2}$. Otherwise, s' is contained in a slab of width w as claimed. Any of the N cells of $\mathcal{G}'$ intersected by this slab is entirely contained in a slab of width $w + 2\sqrt{d}$, whose intersection with the cube must be of $(d-1)$-volume at least N/K^{d-1} and at most $w\sqrt{2}$, since the maximum section of a cube is $\sqrt{2}$[19]. Thus $N \leqslant 2d\sqrt{d}K^{d-2}$. We may thus take $C_d(K) = 1/K^{d-2}$ and $G_d(K) = 2d(d+1)\sqrt{d}K^{d-2}$. □

Proof of Theorem 112 (cont.) We let $K = 2^p$ in the lemma, and choose p as the smallest integer such that $G_d(2^p) < (2^p)^{d-1}$. Assume there are a objects intersecting at most $G_d(2^p)$ cells, and b other objects. The cost of σ is $(\gamma + a + b)\lambda_1(\sigma)$. Since $\sigma_p^{(\text{complete})}$ has cost at most $\left(2^p\gamma + a\frac{G_d(2^p)}{2^{p(d-1)}} + 2^p b\right)\lambda_1(\sigma)$, which we assumed to be at least $c(\sigma)$, we have $c(\sigma) = (\gamma + a + b)\,\lambda_1(\sigma) \leqslant \left(2^p\gamma + a\frac{G_d(2^p)}{2^{p(d-1)}} + 2^p b\right)\lambda_1(\sigma)$, which implies that $a \leqslant (\gamma + b)(2^p - 1)\left(1 - \frac{G_d(2^p)}{2^{p(d-1)}}\right)^{-1}$.

By lemma 113, an object which belongs to more than $G_d(p)$ cells has measure at least $C_d(p)$ so its contribution to the cost is at least $(d\sqrt{2}C_d(2^p))\lambda_{d-1}(\sigma)$. The optimal cost M_σ is then greater than $(\gamma + bd\sqrt{2}C_d(2^p))\lambda_{d-1}(\sigma)$. We have then proved that M_σ is at least a fixed fraction of the cost of σ, and the theorem follows. □

Already in 2D, the separating quadtree strategy does not work as well for segments as for points, especially since it is not able to distinguish between a segment that barely intersects the corner of the square and the diagonal (in the first case it is usually good to subdivide, and in the second case it is not). The lookahead p-greedy strategy is thus a true improvement.

12.4.2 The case of points

Arguably, the case of points is of theoretical interest only, but has relevance since simplices are usually very small in a graphics scene (when they come from a subdivision surface), and can be thought of as points. This is lent credence by a recent trend: point cloud data (PCD) is becoming an important primitive in computer graphics, and several algorithms for rendering them have been given of late, which are amenable to our cost measure. (A

good introductory reference is the proceedings of the first Symposium on Point-Based Graphics[7].)

In the plane, the 1- and 2-greedy strategy may produce a quadtree of cost $\Theta(n)$ times the optimal cost, and so does 1-greedy in higher dimensions. Indeed, this can be seen by placing n points in the center of the square. With no subdivision, the cost is $4(\gamma + n)$. With full subdivision to depth $k \geqslant 1$, the cost is $24\gamma + 16(n - 2\gamma)/2^k$ which tends to 24γ when k tends to infinity. After one or two subdivisions, then, the cost is $8(n + \gamma)$ or $16\gamma + 4n$, both of which are higher than with no subdivision, hence the tree is not subdivided according to the 1-greedy or the 2-greedy heuristic. This shows that the approximation ratio is $(n+\gamma)/6 = \Theta(n)$. In any dimension, the same example shows that the 1-greedy strategy also has a bad approximation ratio of $\Theta(n)$. Nevertheless, 2-greedy (for $d \geqslant 3$) and 3-greedy (for $d = 2$) both work near-optimally.

Lemma 114. *Given a set S of n points in the unit d-cube, if M is the infimum of $c(\mathcal{T})$ over all trees $\mathcal{T}$, then $c(\mathcal{T}^{(\text{3-greedy})}) = O(M)$ for all $d \geqslant 2$, and $c(\mathcal{T}^{(\text{2-greedy})}) = O(M)$ for all $d \geqslant 3$.*

Proof. We prove that 3-greedy is near-optimal, for $d \geqslant 2$. Then we indicate what changes for 2-greedy when $d \geqslant 3$. The cost of a cell σ is $(\gamma + |S_\sigma|)\lambda_1(\sigma)$. The cost of a complete subdivision $\sigma_3^{(\text{complete})}$ of σ to depth 3 is $(8\gamma + \frac{1}{8^{d-1}} \sum_{p \in S_\sigma} n_p)\lambda_{d-1}(\sigma)$, where n_p is the number of leaves of that subdivision that contain p. Since n_p is at most 2^d, the latter cost is at most $(8\gamma + |S_\sigma|2^{3-2d})\lambda_{d-1}(\sigma)$. If σ is a leaf of $\mathcal{T}^{(\text{3-greedy})}$, this means that $\sigma_3^{(\text{opt})}$ is just the leaf σ, then certainly $c(\sigma_3^{(\text{complete})})$ is greater than $c(\sigma)$. Hence $8\gamma + |S_\sigma|2^{3-2d} \geqslant \gamma + |S_\sigma|$, i.e. $|S_\sigma| = \frac{7}{1-2^{3-2d}}\gamma = C_d$. The only case when the 3-greedy strategy fails, is when the lookahead says it is wrong to subdivide whereas it is not. In that case, the number of points is at most $C_d\gamma$. Then the cost is at most $(1 + C_d)\gamma\lambda_{d-1}(\sigma)$, and because of Lemma 107 this is $O(M_\sigma)$, where M_σ is the infimum cost of all the possible quadtree subdivisions of σ. Since lookahead only subdivides a cell if it has a subtree that actually improves the cost, necessarily $\mathcal{T}^{(\text{3-greedy})}$ is a subtree of $\mathcal{T}^{(\text{opt})}$ (where $\mathcal{T}^{(\text{opt})}$ is any optimal[21] tree). Thus every leaf σ of $\mathcal{T}^{(\text{3-greedy})}$ is a cell of $\mathcal{T}^{(\text{opt})}$, and since $c(\sigma) = O(M_\sigma)$ then $c(\mathcal{T}^{(\text{3-greedy})}) = O(\mathcal{T}^{(\text{opt})}) = O(M)$ as well.

Redoing the computation with lookahead $p = 2$ for $d \geqslant 3$, we find that $C_d = \frac{7}{1-2^{2-d}}$, and the rest of the proof follows similarly. □

On the example given in that section, the separation criterion is near optimal. In fact, Aronov and Schiffenbauer[15] have also and independently proven that $\mathcal{Q}^{(\text{AS-sep})}(S)$ is near-optimal, where $\mathcal{Q}^{(\text{AS-sep})}$ uses a slightly different separation criterion: subdivide a cell whenever it covers more than one point of S, and also if any of its orthogonal (resp. diagonal) neighbors has depth which differs by more than 1 (resp. 2). Note that their construction is valid only if all the points in S are distinct and produces smooth quadtrees (see next section).

[21] The cost M is an infimum and may not be achieved by any finite-depth tree. But for this proof, $\mathcal{T}^{(\text{opt})}$ can be allowed to have infinite depth. Alternately, one may consider an increasing sequence of trees $(\mathcal{T}_k^{(\text{opt})})_{k\to\infty}$ whose costs converge to M.

As for finding the optimal tree, the question is still open whether for given values of γ (and maybe of n) there exists a p such that the lookahead greedy strategy yields the *optimal* result, namely, the sequence $c(\mathcal{T}_k^{(p\text{-greedy})})$ converges towards M as k tends to infinity. All we know is that if $n = 5$ and $\gamma < 1$ tends to 1, the required p tends to infinity. We can also mention that if every point belongs to at most one cell, then $p = 1$ leads to the optimal tree.

12.5 Rebalancing quadtrees and octrees

In all the applications of quadtrees and octrees, it can be important to maintain aspect ratio (hence starting with a unit cube) and to ensure that two neighboring cells don't have wildly differing sizes. This has led several authors to propose balancing for trees. From our perspective, since the cost measure of[13] provably relates to the cost of traversal only for balanced trees, we are interested in balancing trees as well. Rebalancing is known to increase size by at most a constant factor. In this section, we prove that rebalancing also does not affect the cost by more than a multiplicative constant factor.

Two leaves are *k-adjacent* if they intersect in a convex portion of dimension k. A tree is called *k-balanced* if the depths of any two k-adjacent leaves differ by at most one. Notice that when considering two k-balanced trees, their intersection, constructed from the unit tree by subdividing all and only cells that are subdivided in both trees, is k-balanced. Thus for a tree $\mathcal{T}$, there is a unique balanced tree $\text{bal}_k(\mathcal{T}) = \min\{\mathcal{T}' : \mathcal{T} \prec \mathcal{T}' \text{ and } \mathcal{T}' \text{ is } k\text{-balanced}\}$, which is called the *k-rebalancing* of $\mathcal{T}$. For instance, 0-balanced quadtrees are what Moore called *smooth* quadtrees[84], and 1-balanced what he called *balanced* and others called 1-*irregular* or *restricted*.

Theorem 115. *Let $\mathcal{T}$ be a tree storing both points and/or simplices in the unit cube. Then for any k, $0 \leqslant k < d$, $c(bal_k(\mathcal{T})) = O(c(\mathcal{T}))$.*

Since the cost $c(\mathcal{T}) = c_t(\mathcal{T}) + c_o(\mathcal{T})$, the proof will be a simple consquence of the next two lemmas. The following lemma was first proven by Weiser in 2D, and by Moore for any dimension $d \geqslant 2$, for a cost function which is simply the number of leaves. Here, the cost function is the total surface area of the leaves but their proof applies as well.

Lemma 116. *Let $\mathcal{T}$ be a tree. Then for any $k \geqslant 0$, $c_t(bal_k(\mathcal{T})) \leqslant 3^d c_t(\mathcal{T})$.*

Proof sketch. The construction suggested by Moore[84] starts with $\mathcal{T}' \leftarrow \mathcal{T}$ and subdivides every leaf node of $\mathcal{T}'$ which has a 0-adjacent neighbor whose size is smaller than half, or equivalently whose depth is greater by more than one. The process iterates until no further subdivision is necessary. The process must terminate with $\mathcal{T}' = \text{bal}_0(\mathcal{T})$.

The basic observation is that, if all the neighbors of a leaf node σ have same or smaller depth, then σ never has to be subdivided[84]. Intuitively, σ will never force any further subdivision of these neighbors, and so they can only be subdivided as a side, corner, or border, on the opposite side of σ. Using this, Moore argues that every subdivision of an internal node σ in $\mathcal{T}'$ which is not an internal node of $\mathcal{T}$ can be "blamed" on a 0-adjacent internal node of $\mathcal{T}$ at the same depth as σ, and since a node in $\mathcal{T}$ can only be blamed by its 0-adjacent neighbors in $\mathcal{T}'$ at the same depth (at most 3^d of them), the total number m

of nodes of $\mathcal{T}'$ can only be 3^d times the total number of nodes of $\mathcal{T}$. This also holds for the number ℓ of leaves, since by induction it is not hard to see that $\ell = 1 + (2^d - 1)(m-1)/2^d$; indeed, subdividing a leaf σ adds 2^d extra leaves to m but σ is no longer counted in ℓ, so ℓ changes only by $2^d - 1$. Consult[84] for the details.

Regarding the cost, the above proof adapts naturally by monitoring, instead of the total number of nodes, the *total area* $c_a(\mathcal{T})$ of the tree (internal and external nodes). By induction, it is not hard to see that $c_a = 2c_t - 2d$; indeed, subdividing σ adds $2\lambda_{d-1}(\sigma)$ to c_a but c_t changes only by $\lambda_{d-1}(\sigma)$ since σ is no longer a leaf. When subdividing σ in $\mathcal{T}'$, the cost of σ's children can be "blamed" on a 0-adjacent internal node of $\mathcal{T}$ at the same depth as σ, and since a node in $\mathcal{T}$ can only be blamed by its 3^d neighbors in $\mathcal{T}'$ at the same depth, we easily have $c_a(\mathcal{T}') \leqslant 3^d c_a(\mathcal{T})$ which implies $c_t(\mathcal{T}') \leqslant 3^d c_t(\mathcal{T})$.

Since $\mathcal{T}' = \mathrm{bal}_0(\mathcal{T})$, which is a refinement of $\mathrm{bal}_k(\mathcal{T})$, for any $k > 0$, this implies that $c_t(\mathrm{bal}_k(\mathcal{T})) \leqslant c_t(\mathrm{bal}_0(\mathcal{T})) \leqslant 3^d c_t(\mathcal{T})$. □

Next we prove in Lemma 117 that the object cost of $\mathrm{bal}_k(\mathcal{T})$ is at most twice (for points) and some constant $B_d \leqslant 2d^2 4^d$ (for simplices) times that of $\mathcal{T}$. We prove this for a single object whose object cost is $\lambda_{d-1}(\mathcal{L}_s(\mathcal{T}))$, since the object cost is just the sum over all objects of that quantity. We conjecture that the best value is $B_d = 2^{d-1}$.

Lemma 117. *Let $\mathcal{T}$ be a tree, and consider the object cost of a single object $s \in S$ both in $\mathcal{T}$ and in $bal_k(\mathcal{T})$. If s is a point, then $\lambda_{d-1}(s \cap bal_k(\mathcal{T})) \leqslant 2\lambda_{d-1}(\mathcal{L}_s(\mathcal{T}))$. If s is a convex object of dimension at most $d-1$, then there is a constant B_d such that $\lambda_{d-1}(s \cap bal_k(\mathcal{T})) \leqslant B_d \lambda_{d-1}(\mathcal{L}_s(\mathcal{T}))$.*

Proof. For a point, the worst case is when it falls exactly on the center of a cell: it then belongs to all the children of the subdivided cell, and thus its contribution to the object cost is at most doubled. Note that further subdivision will never increase that contribution.

The situation is more complicated for simplices, but we can prove that the contribution of an object within a leaf σ of $\mathcal{T}$ is never increased by more than a constant factor. The proof relies in that, when rebalancing, σ is only subdivided along its boundary, because only its neighbors can force it to subdivide. Thus, the subtree σ' of σ in $\mathrm{bal}_k(\mathcal{T})$ must satisfy $\sigma' \prec \sigma_\infty$, where σ_∞ is the maximum subdivision of σ along its boundary (an infinite border). By the next lemma and appropriate scaling, the object cost $\lambda_{d-1}(s \cap \mathrm{bal}_k(\mathcal{T}))$ is at most $B_d \lambda_{d-1}(\mathcal{L}_s(\mathcal{T}))$. □

Lemma 118. *The object cost of a flat object in $\mathcal{T} \prec \mathcal{T}_\infty^{(\mathrm{side})}$, and in $\mathcal{T} \prec \mathcal{T}_\infty^{(\mathrm{border})}$ respectively, is bounded by constants $S_d = d4^d$ and $B_d = 2d^2 4^d$ respectively.*

Proof. We first prove the lemma for sides, then for borders. Let d be fixed, and let s_k (resp. c_k, b_k) be the maximum object cost of a simplex in $\mathcal{T}_k^{(\mathrm{side})}$ (resp. $\mathcal{T}_k^{(\mathrm{corner})}$, $\mathcal{T}_k^{(\mathrm{border})}$). By construction, a k-side is made up of 2^{d-1} half-scaled empty cells and 2^{d-1} half-scale $(k-1)$-sides. Going one level further, it is made of 2^{d-1} half-scaled empty cells, $2^{2(d-1)}$ quarter-scale empty cells, and $2^{2(d-1)}$ quarter-scale $(k-2)$-sides. If the simplex intersects all the $(k-2)$-sides, then it cannot intersect the half-scaled empty cells, and it thus intersect only 2^{d-1} half-scale $(k-1)$-sides, thus $s_k \leqslant s_{k-1}$ (the factor 2^{d-1} disappears

because of scaling). On the other hand, if it misses any of the $(k-2)$-sides, then the object cost decreases geometrically (but it may intersect all the empty cells, of total measure $4d$). More explicitly,

$$s_k \leqslant \max\left(s_{k-1}, 4d + \frac{2^{2d-2}-1}{2^{2d-2}} s_{k-2}\right).$$

This recurrence with $s_0 = 2d$ shows that s_k is bounded by $S_d = d4^d$ for all k. We can then consider a k-border as the union of $2d$ half-scaled k-sides (the superpositions only increase the cost), so that the cost is at most $2d^2 4^d$. □

Remark 1. The constant 3^d in Lemma 116 is tight for the number of leaves, but we do not know if it is tight for the cost of 0-balanced trees. When considering k-balanced trees, $k \geqslant 1$, it certainly is overly pessimistic. Indeed Moore shows that in the plane, 9 is tight for 0-balanced quadtrees, but 8 is tight for 1-balanced quadtrees.

Remark 2. Aronov and Schiffenbauer independently proved[15, Th. 2] that for any quadtree $\mathcal{Q}$ (not necessarily balanced), there is a balanced quadtree $\mathcal{Q}'$ such that $c(\mathcal{Q}') = O(c(\mathcal{Q}))$. Their construction may subdivide more than necessary.

Remark 3. It could also very well be that rebalancing actually decreases the cost. We don't know that, and we don't need it since we are mostly interested in trees for which $c(\mathcal{T}) = O(M)$. In any case, we can ask if there is a reverse theorem (lower bound on $c(\mathrm{bal}_k(\mathcal{T}))$ in terms of $c(\mathcal{T})$).

12.6 Conclusion

In this paper we have proved that instead of considering the optimal octree, and without increasing the cost too much, we may consider the octree given by the lookahead strategy. Still, this may yield an infinite subdivision. In order to have an effective algorithm, we need to add a termination criteria such as a depth limit of $\log_2 n$. As we have also proven, this increases again the cost at most by a constant factor. In practice, we find that greedy with or without lookahead yield the same approximation ratio.

All the results stated in this paper should extend easily to recursive grids and simplicial trees as well, in two and higher dimensions, with only small differences. However, the constants involved in the analysis would be even higher than they are here.

We conclude with a few open problems: first, is it true that by pruning at depth $k = \Theta(\log n)$, we can approach the cost to within $1+\varepsilon$ for simplices? Since the optimal tree might be infinite, there is little sense in asking for an algorithm that constructs the optimal tree. But if the answer to this question were true, it would be nice to have a PTAS with respect to the cost measure. We don't know if the greedy strategy for high enough lookahead would fit the bill.

Lastly, the cost measure considered here is $c(\mathcal{T})$ but the theoretically sound one that does model the average traversal cost during ray shooting is $c^*(\mathcal{T})$ (see Eq. 12). Our only result here is that the greedy strategy does not work.

Chapter 13

Farthest-Polygon Voronoi Diagrams

Ce travail, réalisé avec O. Cheong, H. Everett, J. Gudmundsson, S. Hornus, S. Lazard, M. Lee et H.-S. Na, a été présenté au European Symposium on Algorithms [35].

Given a family of k disjoint connected polygonal sites of total complexity n, we consider the farthest-site Voronoi diagram of these sites, where the distance to a site is the distance to a closest point on it. We show that the complexity of this diagram is $O(n)$, and give an $O(n \log^3 n)$ time algorithm to compute it.

13.1 Introduction

Consider a family $\mathcal{S}$ of geometric objects (called "sites") in the plane. The farthest-site Voronoi diagram of $\mathcal{S}$ subdivides the plane into regions, each region associated with one site $P \in \mathcal{S}$, and containing those points $x \in \mathbb{R}^2$ for which P is the farthest among the sites of $\mathcal{S}$.

While closest-site Voronoi diagrams have been studied extensively [17], their farthest-site cousins have received somewhat less attention. The case of (possibly intersecting) line segment sites was only solved recently by Aurenhammer et al. [18]; they gave an $O(n \log n)$ time algorithm to compute the diagram for n line segments.

Farthest-site Voronoi diagrams have a number of important applications. Perhaps the most well-known one is the problem of finding a smallest disk that intersects all the sites. This disk can be computed in linear time once the diagram is known, since its center is a vertex or lies on an edge of the diagram. Other applications are finding the largest gap to be bridged between sites, or building a data structure to quickly report the site farthest from a given query point.

We are here interested in the case of complex sites with non-constant description complexity. This setting was perhaps first considered by Abellanas et al. [1]: their sites are finite point sets, and so the distance to a site is the distance to the nearest point of that site. Put differently, they consider n points colored with k different colors, and their *farthest color Voronoi diagram* subdivides the plane depending on which color is farthest away. The motivation for this problem is the one mentioned above, namely to find a smallest disk that contains a point of each color—this is a facility location problem where

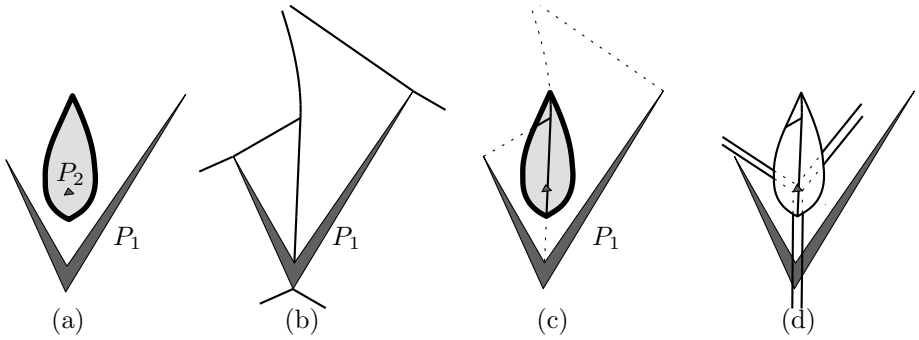


Figure 13.1: (a) The bisector of two polygons can be a closed curve. (b) The medial axis $\mathfrak{M}(P_1)$. (c) The Voronoi region $\mathfrak{R}(P_1)$. (d) The farthest-polygon Voronoi diagram $\mathfrak{F}(\{P_1, P_2\})$.

the goal is to find a position that is as close as possible to each of k different types of facilities (such as schools, post offices, supermarkets, etc.). In a companion paper [2] the authors study other color-spanning objects.

The farthest color Voronoi diagram is easily seen to be the projection of the upper envelope of the k Voronoi surfaces corresponding to the k color classes. Huttenlocher et al. [73] show that this upper envelope has complexity $\Theta(nk)$ for n points, and can be computed in time $O(nk \log n)$ (see also the book by Sharir and Agarwal [98, Section 8.7]).

Van Kreveld and Schlechter [105] consider the farthest-site Voronoi diagram for a family of disjoint simple polygons. Again, they are interested in finding the center of the smallest disk intersecting or touching all polygons, which they then apply to the cartographic problem of labeling groups of islands. Their algorithm is based on the claim that this *farthest-polygon Voronoi diagram* is an instance of the *abstract farthest-site Voronoi diagram* defined by Mehlhorn et al. [83]—but this claim is false, since the bisector of two disjoint simple polygons can be a closed curve, see Fig. 13.1(a). In particular, Voronoi regions can be bounded (which is impossible for regions in abstract farthest-site Voronoi diagrams).

Note that the farthest-polygon Voronoi diagram can again be expressed as the upper envelope of k Voronoi surfaces—but this does not seem to lead to anything stronger than near-quadratic complexity and time bounds. We show in this paper that in fact the complexity of the farthest-polygon Voronoi diagram of k disjoint simple polygons of total complexity n is $O(n)$, independent of the number of polygons.

We give a divide-and-conquer algorithm with running time $O(n \log^3 n)$ to compute the farthest-polygon Voronoi diagram. Our key idea is to build point location data structures for the partial diagrams already computed, and to use parametric search on these data structures to find suitable starting vertices for the merging step. This idea may find applications to the computation of other complicated Voronoi diagrams. Our algorithm implies an $O(n \log^3 n)$ algorithm to compute the smallest disk touching or intersecting all the input polygons.

We note that for a family of disjoint *convex* polygons, finding the smallest disk touching all of them is much easier, and can be solved in time $O(n)$ (where n is the total complexity of the polygons) [75].

13.2 Definition of Farthest-Polygon Voronoi Diagrams

We consider a family $\mathcal{S}$ of k pairwise-disjoint polygonal sites of total complexity n. Here, a *polygonal site* of complexity m is the union of m line segments, whose interiors are pairwise disjoint, but whose union is connected. (In other words, the corners[22] and edges of a polygonal site form a one-dimensional connected simplicial complex in the plane.) In particular, the boundary of a simple polygon is a polygonal site and we can model a polygon with h holes using its $h+1$ boundaries and connecting them with at most h additional edges. For a point $x \in \mathbb{R}^2$, the distance $d(x, P)$ between x and a site $P \in \mathcal{S}$ is the distance from x to the closest point on P.

We assume that the family $\mathcal{S}$ is in general position, that is, no disk touches four edges, and no two edges are parallel.

The *features* of a site P are its corners and edges. For a site $P \in \mathcal{S}$, we define the function $\Psi_P : \mathbb{R}^2 \mapsto \mathbb{R}$ as $\Psi_P(x) = d(x, P)$. The graph of Ψ_P is a *Voronoi surface*, it is the lower envelope of circular cones for each corner of P and of rectangular wedges for each edge of P. The orthogonal projection of this surface on the plane induces a subdivision of the plane, the *medial axis* $\mathfrak{M}(P)$ of P. Each cell C of this subdivision corresponds to a feature w of P, it is the set of all points $x \in \mathbb{R}^2$ such that w is or contains the unique closest point on P to x. Here, edges of P are considered relatively open, so the cell of a corner is disjoint from the cells of its incident edges. The boundaries between cells are the *arcs* and *vertices* of the medial axis. Since the medial axis of P is the closest-site Voronoi diagram of P's features, it can be computed in time $O(m \log m)$, where m is the complexity of P [61].

A pocket of P is a connected component of $\mathrm{CH}(P) \setminus P$, where $\mathrm{CH}(P)$ is the convex hull of P. The medial axis $\mathfrak{M}(P)$ contains exactly one tree for each pocket of P. If a pocket shares an edge with $\mathrm{CH}(P)$ that is not an edge of P, then its medial axis tree contains exactly one infinite arc. (There can be additional infinite arcs for each corner of an edge of P that appear on the convex hull.)

We now consider the function $\Phi : \mathbb{R}^2 \mapsto \mathbb{R}$ defined as $\Phi(x) = \max_{P \in \mathcal{S}} \Psi_P(x)$. The graph of Φ is the upper envelope of the surfaces Ψ_P, for $P \in \mathcal{S}$. The surface Φ consists of conical and planar patches from the Voronoi surfaces Ψ_P, and the arcs separating such patches are either arcs of a Voronoi surface Ψ_P (we call these *medial axis arcs*), or intersection curves of two Voronoi surfaces Ψ_P and Ψ_Q (we call these *pure arcs*). The vertices of Φ are of one of the following three types:

- Vertices of one Voronoi surface Ψ_P. We call these *medial axis vertices*.
- Intersections of an arc of Ψ_P with a patch of another surface Ψ_Q. We call these *mixed vertices*.
- Intersections of patches of three Voronoi surfaces Ψ_P, Ψ_Q, Ψ_R. We call these *pure vertices*.

The projection of the surface Φ onto the plane is the *farthest-polygon Voronoi diagram* $\mathfrak{F}(\mathcal{S})$ of $\mathcal{S}$. It is a subdivision of the plane into cells, arcs, and vertices. For a point $x \in \mathbb{R}^2$,

[22] We reserve the word "vertex" for vertices of the Voronoi diagram.

let us define $D(x) = D_{\mathcal{S}}(x)$ as the smallest disk centered at x that intersects all sites $P \in \mathcal{S}$. By definition, there is always at least one site that touches $D(x)$ without intersecting its interior, and the radius of $D(x)$ is equal to $\Phi(x)$. By our general position assumption, only the following five cases can occur:

- If $D(x)$ touches one site P in only one feature w, and all other sites intersect the interior of $D(x)$, then x lies in a cell of $\mathfrak{F}(\mathcal{S})$, and the cell belongs to the feature w of P.
- If $D(x)$ touches one site P in two or three features, and all other sites intersect the interior of $D(x)$, then x lies on a medial axis arc or medial axis vertex of $\mathfrak{F}(\mathcal{S})$, and is incident to cells belonging to different features of P.
- If $D(x)$ touches one feature w of site P, one feature u of site Q, and all other sites intersect the interior of $D(x)$, then x lies on a pure arc separating cells belonging to features w and u.
- If $D(x)$ touches two features of site P and one feature of site Q, and all other sites intersect the interior of $D(x)$, then x is a mixed vertex incident to a medial axis arc of P.
- If $D(x)$ touches one feature each of three sites P, Q, and R, and all other sites intersect the interior of $D(x)$, then x is a pure vertex.

Put differently, vertices of $\mathfrak{F}(\mathcal{S})$ are points $x \in \mathbb{R}^2$ where $D(x)$ touches three distinct features of sites. If all three features are on the same site, the vertex is a medial axis vertex. If the three features are on three distinct sites, then the vertex is a pure vertex. In the remaining case, if two features are on a site P, and the third feature is on a different site Q, the vertex is a mixed vertex.

Consider now an arc α of $\mathfrak{F}(\mathcal{S})$. If we let x move along α, then $\Phi(x)$—which is the radius of $D(x)$—changes continuously. Since the arc is defined by two features (corners or edges), $\Phi(x)$ cannot assume a local maximum in the interior of α, but it can assume a local minimum. We denote the location of such a local minimum a *pseudo-vertex* of $\mathfrak{F}(\mathcal{S})$. After introducing pseudo-vertices, $\Phi(x)$ is a monotone function on each arc, and we can thus orient all arcs in the direction of increasing $\Phi(x)$.

Fig. 13.2 illustrates all different vertex types, including the two types of pseudo-vertices of degree two (types (c) and (f)).

Since the local shape of $\mathfrak{F}(\mathcal{S})$ around a vertex v is determined solely by the features defining the vertex, in each configuration we can uniquely determine the orientation of the incident arcs. Fig. 13.2 shows the orientation of these arcs in each case. This orientation plays a crucial role in the next section, where we show that the complexity of the diagram, that is, the total number of vertices, is only $O(n)$.

In addition to the nine types of vertices discussed above, we need to consider *vertices at infinity*, that is, we consider the semi-infinite arcs of $\mathfrak{F}(\mathcal{S})$ to have a degree-one vertex at their end. For a vertex v at infinity, the "disk" $D(v)$ is a halfplane, and we have two cases:

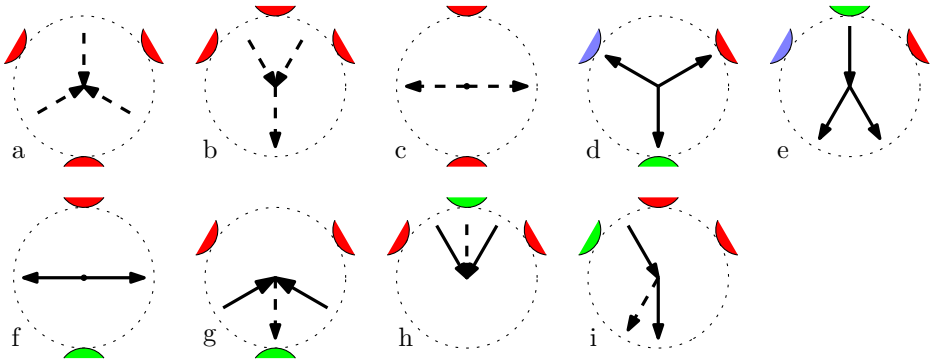


Figure 13.2: The different types of vertices in the farthest-polygon Voronoi diagram. Pure arcs are shown solid, medial axis arcs dashed. The arrows indicate the direction of increasing $\Phi(x)$.

- If $D(v)$ touches one site P in two features, and all other sites intersect the interior of $D(v)$, then $D(v)$ is the "infinite" endpoint of a medial axis arc, and we consider it a medial axis vertex at infinity.
- If $D(v)$ touches two distinct sites, and all other sites intersect its interior, then $D(v)$ is the "infinite" endpoint of a pure edge, and we consider it a pure vertex at infinity.

Finally, let us define the *Voronoi region* of a site $P \in \mathcal{S}$. The Voronoi region $\mathcal{R}(P)$ of P is simply the union of all cells, medial axis arcs, and medial axis vertices of $\mathfrak{F}(\mathcal{S})$ belonging to features of P. Voronoi regions are not necessarily connected, in fact, a single Voronoi region can have up to $k-1$ connected components.

13.3 Complexity Bound

Consider again Fig. 13.2. Let us call a vertex a *source* if it has more outgoing pure arcs than incoming pure arcs, and a *sink* if it has more incoming pure arcs than outgoing pure arcs. As shown in the figure, all finite pure vertices are sources. The pure vertices at infinity are sinks. The only other sinks are the mixed vertices of type (g) and (h). Note that mixed vertices of type (i) are neither sources nor sinks.

We can now partition the pure arcs of $\mathfrak{F}(\mathcal{S})$ into disjoint directed paths. Each such path starts at a source, and ends at a sink. This implies that we can bound the number of sources by twice the number of sinks, as at most two paths can end in a sink.

We show in the following that the number of pure vertices at infinity is at most $2k-2$, and that the number of mixed vertices is $O(n)$. Since the total number of vertices of all medial axes $\mathfrak{M}(P)$, for $P \in \mathcal{S}$, is $O(n)$, this implies the main theorem of this section:

Theorem 119. *The complexity of the farthest-polygon Voronoi diagram of a family of disjoint polygonal sites of total complexity n is $O(n)$.*

We first discuss the number of vertices at infinity.

Lemma 120. *The number of pure vertices at infinity of $\mathfrak{F}(\mathcal{S})$ is at most $2k-2$. The total number of vertices at infinity of $\mathfrak{F}(\mathcal{S})$ is $O(n)$.*

Proof. For two sites $P, Q \in \mathcal{S}$, consider the diagram $\mathfrak{F}(\{P, Q\})$. A pure vertex at infinity corresponds to an edge of $\mathrm{CH}(P \cup Q)$ supported by a corner of P and a corner of Q. But $\mathrm{CH}(P \cup Q)$ can have at most two such edges, since P and Q are disjoint and both are connected, and so $\mathfrak{F}(\{P, Q\})$ has at most two pure vertices at infinity.

Consider now again $\mathfrak{F}(\mathcal{S})$, and let $\sigma(\mathcal{S})$ denote the sequence of sites whose Voronoi regions appear at infinity in circular order, starting and ending at the same region. We claim that $\sigma(\mathcal{S})$ is a Davenport-Schinzel sequence of order 2, and has therefore length at most $2k-1$ [98]. Indeed, $\sigma(\mathcal{S})$ has by definition no two consecutive identical symbols. Assume now that there are two sites P and Q such that the subsequence $PQPQ$ appears in $\sigma(\mathcal{S})$. If we delete all other sites, then $\sigma(\{P, Q\})$ would still need to contain the subsequence $PQPQ$, and therefore $\mathfrak{F}(\{P, Q\})$ would contain at least three pure vertices at infinity, a contradiction to the observation above.

It now suffices to observe that the pure vertices at infinity are exactly the transitions between consecutive Voronoi regions, and their number is at most $2k-2$. All remaining vertices at infinity are medial axis vertices. Since the total complexity of all $\mathfrak{M}(P)$, for $P \in \mathcal{S}$, is $O(n)$, the bound follows. □

It remains to show that the total number of mixed vertices is $O(n)$. Consider a site $P \in \mathcal{S}$ of complexity m. Its medial axis $\mathfrak{M}(P)$ has complexity $O(m)$, and we can consider $\mathfrak{M}(P)$ as a graph embedded in $\mathbb{R}^2 \setminus P$. It consists of a collection of trees. In Lemma 122 below we show that for each tree $\mathcal{T}$ of $\mathfrak{M}(P)$ the intersection $\mathcal{T} \cap \mathcal{R}(P)$ is a connected subtree. Since the mixed vertices on $\mathcal{T}$ are exactly the finite leaves of this subtree, this implies that the number of mixed vertices on $\mathfrak{M}(P)$ is $O(m)$. Summing over all $P \in \mathcal{S}$ then proves that the number of mixed vertices of $\mathfrak{F}(\mathcal{S})$ is $O(n)$.

It remains to prove Lemma 122. We need another notation: for a point $x \in \mathbb{R}^2$ and a site P, let $D_P(x)$ denote the largest disk centered at x whose interior does not intersect P (and which is therefore touching P). Note that a point $x \in \mathfrak{M}(P)$ is in $\mathcal{R}(P)$ if and only if $D_{\mathcal{S}}(x) = D_P(x)$, which is true if and only if all other sites intersect the interior of $D_P(x)$.

Lemma 121. *Let γ be a path in $\mathfrak{M}(P)$, let $Q \in \mathcal{S} \setminus \{P\}$ be another site, and let γ_Q be the set of points $x \in \gamma$ where $D_P(x)$ intersects Q. Then γ_Q is a connected subset of γ, that is, a subpath.*

Proof. We can assume γ to be a maximal path in $\mathcal{T}$, connecting a corner w of P with another corner u (or possibly going to infinity). Assume for a contradiction that there are points x, y, z on γ in this order such that $x, z \in \gamma_Q$, but $y \notin \gamma_Q$.

The disk $D_P(y)$ separates one connected component of $\mathbb{R}^2 \setminus P$ into two components A and B. The two endpoints of γ must lie in different components, say $w \in A$ and $u \in B$.

We first argue that any disk D that does not intersect P cannot contain points in both A and B. Indeed, the boundary of D would have to intersect the boundary of $D_P(y)$ in four points, a contradiction.

The disk $D_P(z)$ touches P on the boundary of B, and so it contains a point in B—which implies that it cannot contain a point in A. Since $D_P(y)$ does not intersect Q and

Q is connected, this implies that Q is entirely contained in B. On the other hand, the disk $D_P(x)$ touches P on the boundary of A, and by our claim that means that it cannot intersect B. That implies that $D_P(x) \cap Q$ is empty, and the claim follows. □

The following lemma is an easy consequence of Lemma 121.

Lemma 122. *Let $\mathcal{T}$ be a tree of $\mathfrak{M}(P)$. Then $\mathcal{T} \cap \mathcal{R}(P)$ is a connected subtree of $\mathcal{T}$.*

13.4 Computing the Voronoi Diagram

The proof of Theorem 119 suggests an algorithm for computing the diagram by tracing the paths considered there. This is roughly equivalent to computing the surface Φ by sweeping a horizontal plane downwards, and maintaining the part of Φ above this plane. This is essentially the approach used by Aurenhammer et al. [18] for the computation of farthest-segment Voronoi diagrams. It does not seem to work for our diagram because of the mixed vertices of type (h), where Φ has a local maximum.

We instead offer a divide-and-conquer algorithm.

Theorem 123. *The farthest-polygon Voronoi diagram $\mathfrak{F}(\mathcal{S})$ of a family $\mathcal{S}$ of disjoint polygonal sites of total complexity n can be computed in expected time $O(n \log^3 n)$.*

Proof. Let $\mathcal{S} = \{P_1, \ldots, P_k\}$, and let n_i be the complexity of P_i. If $k = 1$, then $\mathfrak{F}(\mathcal{S})$ is simply the medial axis $\mathfrak{M}(P_1)$, which can be computed in time $O(n \log n)$ [61]. Otherwise, we split $\mathcal{S}$ into two disjoint families $\mathcal{S}_1, \mathcal{S}_2$ as follows:

- If there is a site P_i with complexity $n_i \geqslant n/2$, then $\mathcal{S}_1 = \{P_i\}$ and $\mathcal{S}_2 = \mathcal{S} \setminus \{P_i\}$.
- Otherwise there must be an index j such that $n/4 \leqslant \sum_{i=1}^{j} n_i \leqslant 3n/4$. We let $\mathcal{S}_1 = \{P_1, \ldots, P_j\}$ and $\mathcal{S}_2 = \{P_{j+1}, \ldots, P_k\}$.

We recursively compute $\mathfrak{F}(\mathcal{S}_1)$ and $\mathfrak{F}(\mathcal{S}_2)$. We show below that we can then merge these two diagrams to obtain $\mathfrak{F}(\mathcal{S})$ in time $O(n \log^2 n)$, proving the theorem. □

It remains to discuss the merging step. We are given a family $\mathcal{S}$ of disjoint polygonal sites of total complexity n, and we are given $\mathfrak{F}(\mathcal{S}_1)$ and $\mathfrak{F}(\mathcal{S}_2)$, where $\mathcal{S} = \mathcal{S}_1 \cup \mathcal{S}_2$ is a disjoint partition of $\mathcal{S}$.

Consider the diagram $\mathfrak{F}(\mathcal{S})$ to be computed. We color the Voronoi regions of $\mathfrak{F}(\mathcal{S})$ defined by sites in $\mathcal{S}_1$ red, and the Voronoi regions defined by sites in $\mathcal{S}_2$ blue. A pure arc of $\mathfrak{F}(\mathcal{S})$ is red if it separates two red regions, and blue if it separates two blue regions. The remaining pure arcs, which separate a red and a blue region, are called *purple*. A vertex of $\mathfrak{F}(\mathcal{S})$ is purple if it is incident to a purple arc. We observe that by our general position assumption, every finite purple vertex is incident to exactly two purple arcs, and so the purple arcs form a collection of open and closed chains, see Fig. 13.2.

Merging $\mathfrak{F}(\mathcal{S}_1)$ and $\mathfrak{F}(\mathcal{S}_2)$ can be done in linear time once all purple arcs are known. In fact, the diagram $\mathfrak{F}(\mathcal{S})$ consists of those portions of $\mathfrak{F}(\mathcal{S}_1)$ lying in the red regions of $\mathfrak{F}(\mathcal{S})$, and those portions of $\mathfrak{F}(\mathcal{S}_2)$ lying in the blue regions of $\mathfrak{F}(\mathcal{S})$.

We show below that if we have a starting vertex on every purple chain, then we can trace the purple chains in total time $O(n)$. For the open chains, we can use the purple vertices at infinity as starting vertices, as these are easy to compute. For the closed purple chains, we make use of the following lemma:

Lemma 124. *Any closed purple chain contains a mixed vertex.*

Proof. A closed purple chain is a compact set in the plane, and so $\Phi(x)$ assumes a maximum in some point x on this curve. But Φ can assume a local maximum only in mixed vertices of type (g) and (h), see Fig. 13.2. □

It remains to discuss the following three steps: (a) Computing the pure vertices at infinity; (b) Computing the mixed vertices; (c) Tracing the purple chains. Computing the pure vertices at infinity is relatively easy, making use of the same Davenport-Schinzel arguments used in Lemma 120, and we omit the details in this extended abstract.

13.4.1 Tracing the Purple Chains

We first need to discuss a monotonicity property of cells of $\mathfrak{F}(\mathcal{S})$. Let C be a cell of $\mathfrak{F}(\mathcal{S})$ belonging to feature w of site P. For a point $x \in C$, let x^* be the point on w closest to x. Let f_x be a directed line segment starting at x and extending in direction $\overrightarrow{x^*x}$ until we reach $\mathfrak{M}(P)$ (a semi- infinite segment if this does not happen). We call f_x the *fiber* of x. We note that if w is an edge, then all fibers of C are parallel and normal to w; if w is a corner then all fibers are supported by lines through w.

Lemma 125. *For any $x \in C$, the fiber f_x lies entirely in C (and therefore in $\mathcal{R}(P)$).*

Proof. The disk $D(x)$ touches P in x^* only, and its interior intersects all other sites. When we move a point y from x along f_x, the disk D centered at y through x^* keeps containing $D(x)$, and it therefore still intersects all other sites. This implies that $y \in C$ as long as D does not intersect P in another point. This does not happen until we reach $\mathfrak{M}(P)$. □

An immediate consequence is that cells are "monotone", this follows easily from Lemma 125.

Lemma 126. *Let C be a cell belonging to feature w. If w is a corner, then any line through w intersects C in a segment. If w is an edge, then any line normal to w intersects C in a segment.*

Lemma 126 implies that the boundary of a cell C belonging to a feature w consists of two chains monotone with respect to w (that is, monotone in the direction of an edge, and rotationally monotone around a corner). The *lower chain* is closer to the feature and consists of pure arcs only, the *upper chain* consists of medial axis arcs only.

An important consequence of Lemma 125 for tracing the purple chains is the following:

Lemma 127. *Let f_x be a fiber of a cell C in $\mathfrak{F}(\mathcal{S}_1)$ or $\mathfrak{F}(\mathcal{S}_2)$. Then f_x is intersected by at most one purple arc of $\mathfrak{F}(\mathcal{S})$.*

Proof. Assume for the contrary that two purple arcs intersect f_x in points p and q, where q lies on f_p. Then there is a point p' on f_x very close to p such that $p' \in \mathcal{R}(P)$ in $\mathfrak{F}(\mathcal{S})$, where C belongs to site P. But then the fiber $f_{p'}$ lies in $\mathcal{R}(P)$ in $\mathfrak{F}(\mathcal{S})$, and cannot intersect a purple arc at q. □

Lemma 127 allows us to trace purple chains very easily. Once we have located a starting vertex in both $\mathfrak{F}(\mathcal{S}_1)$ and $\mathfrak{F}(\mathcal{S}_2)$, we trace the chain through both diagrams at the same time. We observe that every time it intersects a cell boundary in $\mathfrak{F}(\mathcal{S}_1)$ or $\mathfrak{F}(\mathcal{S}_2)$, we have indeed found a vertex of the purple chain, so the total number of such intersections is only $O(n)$.

When we enter a cell C along a purple arc, we simply follow the arc through the cell as it intersects the fibers of C. We consider the upper and lower chain of C at the same time, and trace the cells of $\mathfrak{F}(\mathcal{S}_1)$ and $\mathfrak{F}(\mathcal{S}_2)$ in parallel. This allows us to charge the cost of tracing to those features of the two cells incident to the fibers the purple chain actually intersects. Since no fiber is intersected by more than one purple arc, the total tracing time is $O(n)$.

13.4.2 Computing the Mixed Vertices

This is the hardest part of the algorithm. We start by computing the randomized[23] point-location data structure of Mulmuley [85] (see also [43, Chapter 6]) for the two given Voronoi diagrams $\mathfrak{F}(\mathcal{S}_1)$ and $\mathfrak{F}(\mathcal{S}_2)$. This data structure only needs two primitive operations:

- For a given point p in the plane, determine whether the query point x lies left or right of p; and
- For an x-monotone line segment or parabola arc γ, determine whether the query point lies above or below γ.

Both cases can be summarized as follows: Given a *comparator* γ, determine on which side of γ the query point x lies. The comparator can be either a line or a parabola arc.

We compute the mixed vertices lying on each medial axis $\mathfrak{M}(P)$ separately. For each tree $\mathcal{T}$ of $\mathfrak{M}(P)$, the intersection $\mathcal{T} \cap \mathcal{R}(P)$ is a connected subtree by Lemma 122. We can determine the *inner* vertices of this subtree easily, by performing a point location operation for each vertex v of $\mathcal{T}$ in $\mathfrak{F}(\mathcal{S}_1)$ and $\mathfrak{F}(\mathcal{S}_2)$. This tells us which site is farthest from v—and v lies in $\mathcal{R}(P)$ if and only if this is the site P. Let I be the set of vertices of $\mathcal{T}$ that lie in $\mathcal{R}(P)$. We now need to consider two cases.

If I is non-empty, then every arc α of $\mathcal{T}$ incident to one vertex in I and one vertex not in I must contain exactly one mixed vertex x^* by Lemma 122. We locate this vertex x^* by using *parametric search* along the arc α [82]. The idea is to execute two point location queries in $\mathfrak{F}(\mathcal{S}_1)$ and $\mathfrak{F}(\mathcal{S}_2)$ using x^* as the query point. Each query executes a sequence of primitive operations, where we compare the (unknown) location of x^* with a comparator

[23] This is the only use of randomization in our algorithm. It can probably be avoided by using the point location data structure of Edelsbrunner et al. [58] instead.

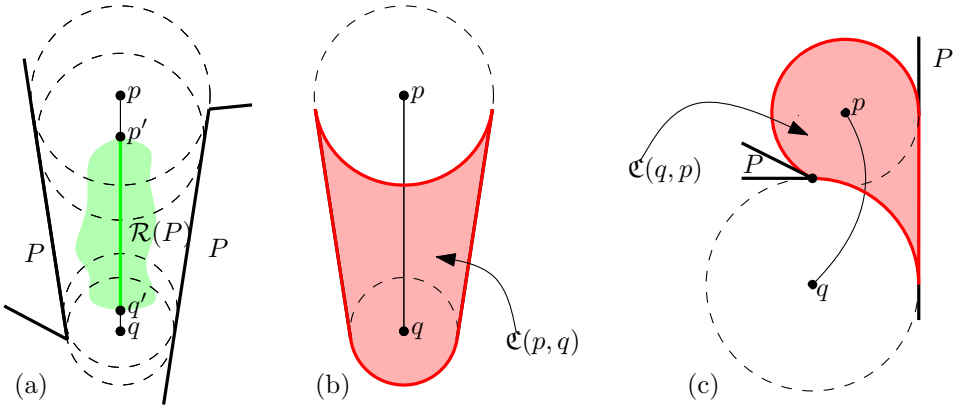


Figure 13.3: (a) pq is a sub-arc of $\mathfrak{M}(P)$. $p'q'$ is the intersection of pq with $\mathcal{R}(P)$. We also have $p'q' = \mathfrak{M}(P) \cap \mathcal{R}(P)$. (b) The cylinder $\mathfrak{C}(p,q)$ of the pair (p,q) in (a). (c) The cylinder $\mathfrak{C}(q,p)$ of the pair (q,p) for a parabola arc.

γ (a line or a parabola arc). This primitive operation can be implemented by intersecting α with γ, resulting in a set of at most four points. In $O(\log n)$ time, we can test for each of these points whether it lies in $\mathcal{R}(P)$. This tells us between which of these points the unknown mixed vertex x^* lies, and we can answer the primitive operation.

It follows that we can execute the two point location queries in time $O(\log^2 n)$, and we obtain the cells C_1 and C_2 of $\mathfrak{F}(\mathcal{S}_1)$ and $\mathfrak{F}(\mathcal{S}_2)$ containing x^*. The mixed vertex x^* lies on the bisector of the features w_1 and w_2 to which these two cells belong (one of them is necessarily a feature of P). We compute the intersection of this bisector with α to obtain x^*.

It remains to consider the case where I is empty, that is, no vertex of $\mathcal{T}$ lies in $\mathcal{R}(P)$. Nevertheless, the region $\mathcal{R}(P)$ may intersect a single arc α of $\mathcal{T}$, and there are then two mixed vertices on α that we need to find. We first need to identify the arcs of $\mathcal{T}$ where this could happen.

Let p and q be two points on the same arc α of $\mathfrak{M}(P)$. We define the *cylinder* $\mathfrak{C}(p,q)$ of the pair (p,q) as

$$\mathfrak{C}(p,q) = \bigcup_{x \in pq} D_P(x) \setminus D_P(p),$$

where the union is taken over all points x on the arc α between p and q, see Fig. 13.3. We define a condition $G(p,q)$ as follows: Let Q be a site farthest from p, and let w be a feature of Q closest to p. Then $G(p,q)$ is true if $w \in \mathfrak{C}(p,q)$ or if Q intersects $D_P(q)$. We omit the easy proof of the following lemma:

Lemma 128. *Let p, q be points on the same arc α of $\mathfrak{M}(P)$, such that neither p nor q lie in $\mathcal{R}(P)$. If α intersects $\mathcal{R}(P)$ between p and q, then $G(p,q)$ and $G(q,p)$ both hold.*

Let us call an arc α connecting vertices p and q of $\mathcal{T}$ a *candidate arc* if $G(p,q)$ and $G(q,p)$ both hold.

Lemma 129. *If α is a candidate arc of a tree $\mathcal{T}$ of $\mathfrak{M}(P)$, then all points in $\mathcal{T} \cap \mathcal{R}(P)$ lie on α.*

Proof. Let α connecting p and q be a candidate arc. Since $G(p,q)$ holds, there is a site $Q \neq P$ such that Q does not intersect $D_{\mathbb{S}}(p)$, but does intersect $\mathfrak{C}(p,q)$. In particular, there must be a point $x \in \alpha$ such that Q intersects $D_{\mathbb{S}}(x) \supset D_P(x)$. Let now y be a point on $\mathcal{T}$ such that the path from q to y goes through α. By Lemma 121, Q cannot intersect $D_P(y)$, and so $y \notin \mathcal{R}(P)$.

Symmetrically, since $G(q,p)$ holds, we find that any point $z \in \mathcal{T}$ such that the path from p to z goes through α cannot be in $\mathcal{R}(P)$, and so any point of $\mathcal{T} \cap \mathcal{R}(P)$ must lie on α. □

Lemma 129 implies immediately that if there are two candidate arcs, then $\mathcal{T} \cap \mathcal{R}(P)$ is empty, and there are no mixed vertices on $\mathcal{T}$.

Since we have point-location data structures for $\mathfrak{F}(\mathcal{S}_1)$ and $\mathfrak{F}(\mathcal{S}_2)$, we can test the condition $G(p,q)$ in time $O(\log n)$ for a given arc α in $\mathfrak{M}(P)$ and two points $p, q \in \alpha$. This allows to identify all candidate arcs in $O(m \log n)$ time, where m is the complexity of $\mathcal{T}$. If there is zero or more than one candidate arcs, we can stop immediately, as there are no mixed vertices on $\mathcal{T}$.

We briefly sketch how to handle the case of the presence of a single candidate arc α. We apply parametric search guided by our predicate $G(\cdot,\cdot)$. This leads either to the conclusion that $\mathcal{R}(P)$ is empty, or to the discovery of a point inside $\alpha \cap \mathcal{R}(P)$, which we can plug in the previous method (when I is not empty) to find the two mixed vertices on α.

13.5 Bonus : diagramme de Voronoï du polygone le plus éloigné sur une sphère

Nous considérons ici ce qui se passe sur une sphère au lieu du plan, et la situation est relativement plus simple.

La première chose que l'on peut remarquer est qu'avec trois polygones, le diagramme de Voronoï du polygone le plus proche a au plus deux sommets. En effet si les bissecteurs sont disjoints il n'y a aucun sommet et sinon les relations d'incidences sommet-arête, la relation d'Euler, et le nombre de faces (trois) implique qu'il y a deux sommets (dans des cas dégénérés les deux sommets pourraient d'en faire qu'un de degré 4, mais on ignore ici les cas dégénérés).

On fixe un polygone. Pour chaque autre polygone, on trace le bissecteur avec le polygone fixe. D'après la remarque précédente, deux de ces bissecteurs se coupent en au plus deux points. De plus, un bissecteur est un cercle topologique. La cellule du diagramme de Voronoï du polygone le plus éloigné du polygone fixe est ainsi une intersection de pseudo-disques. Cette cellule est donc connexe et simplement connexe.

Cette propriété fait du diagramme de Voronoï du polygone le plus éloigné (je parle ici du diagramme le plus dépouillé, qui partitionne juste le plan en régions de points ayant le même objet le plus éloigné) un graphe planaire dont le nombre de faces est majoré par le nombre de polygones. La taille du diagramme est donc linéaire en le nombre de polygones. Si on s'intéresse maintenant au diagramme plus complet qui encode également l'arête ou le sommet le plus proche sur le polygone le plus éloigné, il suffit de découper

chaque cellule du diagramme simple le long du diagramme de Voronoï des segments qui composent le polygone associé, et il est facile de voir que le découpage d'une cellule simple donne un nombre de cellules au plus linéaire en la taille du polygone.

Les propriétés de connexité et de simple connexité peuvent en partie être transposées dans le cas plan. Ainsi les cellules ayant une composante finie sont nécessairement connexes et simplement connexes. Une cellule peut ne pas être simplement connexe uniquement si elle est la seule à être infinie.

Chapitre 14

Conclusion

J'ai étudié dans cette thèse plusieurs questions géométriques et combinatoires liées à la visibilité 3D. Les résultats sur les silhouettes sont importants car ils donnent un fondement théorique à l'affirmation classique disant que la silhouette d'un polyèdre de taille n a en général une taille $\Theta(\sqrt{n})$ et montrent aussi ses limites. Il serait intéressant dans des travaux ultérieurs de montrer que l'on peut aussi, avec des hypothèses plus faibles, donner des bornes sur la complexité de la silhouette pour des polyèdres opaques. Ces travaux ne devraient pas avoir d'impact sur la conception des algorithmes de triangulation de surfaces ou de construction de silhouette. Leur intérêt principal est qu'ils justifient l'analyse de la complexité d'algorithmes géométriques dans le cadre d'objets polyédriques à petite silhouette, un cadre qui par certains aspects est plus réaliste que ceux de la scène la pire ou d'objets uniformément distribués.

J'ai aussi beaucoup étudié la complexité des espaces libres de droites et segments en 3D. S'il reste dans de nombreux cas des imprécisions dans notre connaissance de la complexité de l'espace des droites libres (qui s'est avéré surprenamment difficile à étudier), la complexité de l'espace des segments libres maximaux est aujourd'hui mieux connue. Elle n'est pas petite, mais dans de nombreux cas elle semble raisonnable, et on comprend mieux les cas où elle est grande.

Une grande étape qui reste à franchir est, à l'aide d'une implantation efficace, de confirmer, ou d'infirmer, que les structures globales de visibilité peuvent être réellement utiles en pratiques. Les bornes de complexité présentées ici nous font espérer qu'il s'agira d'une confirmation, et ce d'autant plus qu'il reste un certain nombre d'éléments favorables qui n'ont pas encore été étudiés. Par exemple, il est sans doute possible de ne pas calculer les structures entières, grâce à une approche hiérarchique ou peut-être en les calculant localement à la demande. Aussi, une des raisons premières qui ont poussé à étudier les structures de visibilité est leur capacité supposée à tirer parti de la cohérence spatiale et temporelle des scènes. Ainsi, même si les structures globales de visibilité n'arrivent pas à rivaliser avec les approches actuelles sur des scènes statiques, elles pourront encore s'avérer très utile sur des scènes dynamiques.

Bibliographie

[1] M. Abellanas, F. Hurtado, C. Icking, R. Klein, E. Langetepe, L. Ma, B. Palop, and V. Sacristán. The farthest color Voronoi diagram and related problems. In *Abstracts 17th European Workshop Comput. Geom.*, pages 113–116. Freie Universität Berlin, 2001.

[2] M. Abellanas, F. Hurtado, C. Icking, R. Klein, E. Langetepe, L. Ma, B. Palop, and V. Sacristán. Smallest color-spanning objects. In *Proc. 9th Annu. European Sympos. Algorithms*, volume 2161 of *Lecture Notes Comput. Sci.*, pages 278–289. Springer-Verlag, 2001.

[3] P. K. Agarwal. On stabbing lines for convex polyhedra in 3D. *Computational Geometry : Theory and Applications*, 4(4) :177–189, 1994.

[4] P. K. Agarwal, B. Aronov, V. Koltun, and M. Sharir. Lines avoiding unit balls in three dimensions. *Discrete and Computational Geometry*, 34(2) :231–250, 2005.

[5] P. K. Agarwal, B. Aronov, and M. Sharir. On lines avoiding unit balls in three dimensions. In Proceedings of the *20th ACM Annual Symposium on Computational Geometry (SoCG'04)*, pages 36–45, 2004.

[6] P. K. Agarwal and M. Sharir. Ray shooting amidst convex polyhedra and polyhedral terrains in three dimensions. *Siam Journal on Computing*, 25 :100–116, 1996.

[7] M. Alexa, M. Gross, H. Pfister, and S. Rusinkiewicz. Eurographics, symposium on point-based graphics 2004.

[8] H. Alt, M. Glisse, and X. Goaoc. On the worst-case complexity of the silhouette of a polytope. In *15th Canadian Conference on Computational Geometry - CCCG'03*, pages 51–55, 2003.

[9] N. Amenta. *Helly Theorems and Generalized Linear Programming.* Ph.D. thesis, U.C. Berkeley, 1993.

[10] N. Amenta. Helly-type theorems and generalized linear programming. *Discrete and Computational Geometry*, 12, 1994.

[11] B. Aronov, H. Brönnimann, A. Y. Chang, and Y.-J. Chiang. Cost-driven octree construction schemes : an experimental study. *Computational Geometry : Theory and Applications*, 21(1-2) :127–148, 2005. (Also in Proceedings of SoCG'03).

[12] B. Aronov, H. Brönnimann, A. Y. Chang, and Y.-J. Chiang. Cost prediction for ray shooting in octrees. *Computational Geometry : Theory and Applications*, 34(3) :159–181, 2006.

[13] B. Aronov, H. Brönnimann, A.Y. Chang, and Y.-J. Chiang. Cost prediction for ray shooting. In Proceedings of the *18th ACM Annual Symposium on Computational Geometry (SoCG'02)*, pages 293–302, 2002.

[14] B. Aronov and S. J. Fortune. Approximating minimum-weight triangulations in three dimensions. *Discrete and Computational Geometry*, 21(4) :527–549, 1999.

[15] B. Aronov and R. Schiffenbauer, 2001. unpublished manuscript.

[16] B. Aronov and M. Sharir. The common exterior of convex polygons in the plane. *Computational Geometry : Theory and Applications*, 8 :139–149, 1997.

[17] F. Aurenhammer and R. Klein. Voronoi diagrams. In Jörg-Rüdiger Sack and Jorge Urrutia, editors, *Handbook of Computational Geometry*, pages 201–290. Elsevier Science Publishers B.V. North-Holland, Amsterdam, 2000.

[18] F. Aurenhammer, R.L.S.Drysdale, and H. Krasser. Farthest line segment Voronoi diagrams. *Information Processing Letters*, 100 :220–225, 2006.

[19] K. Ball. Cube slicing in $\mathbb{R}^n$. In *Proc. Amer. Math. Soc. 97 (3) (AMS, Providence, 1986)*, pages 465–473, 1986.

[20] M. Bern, D. P. Dobkin, D. Eppstein, and R. Grossman. Visibility with a moving point of view. *Algorithmica*, 11 :360–378, 1994.

[21] M. Bern and D. Eppstein. Mesh generation and optimal triangulation. In Ding-Zhu Du and Frank Hwang, editors, *Computing in Euclidean Geometry*, volume 1. World Scientific, Lecture Notes Series on Computing, 1992.

[22] H. L. Bertoni. *Radio Propagation for Modern Wireless Systems*. Prentice Hall Professional Technical Reference, 1999.

[23] J.-D. Boissonnat and S. Oudot. Provably good sampling and meshing of surfaces. *Graphical Models, special issue of Solid Modeling*, 67(5) :405–451, 2005.

[24] J.-D. Boissonnat and M. Yvinec. *Algorithmic Geometry*. Cambridge University Press, 1998.

[25] C. Borcea, X. Goaoc, S. Lazard, and S. Petitjean. Common tangents to spheres in $\mathbb{R}^3$. *Discrete and Computational Geometry*, 35(2) :287–300, 2006.

[26] H. Brönnimann, O. Devillers, V. Dujmović, H. Everett, M. Glisse, X. Goaoc, S. Lazard, H.-S. Na, and S. Whitesides. On the number of lines tangent to four convex polyhedra. In Proceedings of the *14th Canadian Conference on Computational Geometry (CCCG'02)*, pages 113–117, Lethbridge, Canada, 2002.

[27] H. Brönnimann, O. Devillers, V. Dujmović, H. Everett, M. Glisse, X. Goaoc, S. Lazard, H.-S. Na, and S. Whitesides. The number of lines tangent to arbitrary polytopes in $\mathbb{R}^3$. In Proceedings of the *20th ACM Annual Symposium on Computational Geometry (SoCG'04)*, pages 46–55, Brooklyn, NY, 2004.

[28] H. Brönnimann, O. Devillers, V. Dujmović, H. Everett, M. Glisse, X. Goaoc, S. Lazard, H.-S. Na, and S. Whitesides. Lines and free line segments tangent to arbitrary three-dimensional convex polyhedra. *SIAM Journal on Computing*, 2006.

[29] H. Brönnimann, H. Everett, S. Lazard, F. Sottile, and S. Whitesides. Transversals to line segments in three-dimensional space. *Discrete and Computational Geometry*, 34(3) :381–390, 2005.

[30] H. Brönnimann and M. Glisse. Cost optimal trees for ray shooting. In *Proc. 6th Latin American Symposium on Theoretical Informatics*, 2004.

[31] H. Brönnimann, M. Glisse, and D. Wood. Cost-optimal quadtrees for ray shooting. In *Proc. 14th Canadian Conference on Computational Geometry*, 2002.

[32] H. Bronnimann and M.T. Goodrich. Almost optimal set covers in finite vc-dimension. *Discrete and Computational Geometry*, 14 :463–479, 1995.

[33] F. Cazals and C. Puech. Bucket-like space partitioning data structures with applications to ray-tracing. In Proceedings of the *13th ACM Annual Symposium on Computational Geometry (SoCG'97)*, pages 11–20, New York, NY, USA, 1997. ACM Press.

[34] A. Y. Chang. *Theoretical and experimental aspects of ray shooting.* PhD thesis, 2004.

[35] O. Cheong, H. Everett, M. Glisse, J. Gudmundsson, S. Hornus, S. Lazard, M. Lee, and H.-S. Na. Farthest-polygon voronoi diagrams. In Proceedings of the *15th Annual European Symposium on Algorithms (ESA'07)*, page 12 pages, Eilat, Israel, 2007.

[36] F. Cho and D. Forsyth. Interactive ray tracing with the visibility complex. *Computers and Graphics*, 23(5) :703–717, 1999. Special issue on Visibility - Techniques and Applications.

[37] K.L. Clarkson and K. Varadarajan. Improved approximation algorithms for geometric set cover. *Discrete and Computational Geometry*, 37 :43–58, 2007.

[38] J. G. Cleary and G. Wyvill. Analysis of an algorithm for fast ray tracing using uniform space subdivision. *The Visual Computer*, 4(2) :65–83, 1988.

[39] R. Cole and M. Sharir. Visibility problems for polyhedral terrains. *Journal of Symbolic Computation*, 7(1) :11–30, 1989.

[40] D. Cox, J. Little, and D. O'Shea. *Ideals, Varieties and Algorithms. An Introduction to Computational Algebraic Geometry and Commutative Algebra.* Undergraduate Texts in Mathematics. Springer-Verlag, 1992.

[41] M. de Berg, H. Everett, and L. Guibas. The union of moving polygonal pseudodiscs – combinatorial bounds and applications. *Computational Geometry : Theory and Applications*, 11 :69–82, 1998.

[42] M. de Berg, D. Halperin, M. Overmars, and M. van Kreveld. Sparse arrangements and the number of views of polyhedral scenes. *International Journal of Computational Geometry and Applications*, 7(3) :175–195, 1997.

[43] M. de Berg, M. van Kreveld, M. Overmars, and O. Schwarzkopf. *Computational Geometry : Algorithms and Applications.* Springer-Verlag, Berlin, Germany, 2nd edition, 2000.

[44] E. Demaine, J. Mitchell, and J. O'Rourke. The open problems project – problem 61.

[45] J. Demouth, O. Devillers, H. Everett, M. Glisse, S. Lazard, and R. Seidel. Between umbra and penumbra. In Proceedings of the *23rd ACM Annual Symposium on Computational Geometry (SoCG'07)*, 2007.

[46] O. Devillers, V. Dujmović, H. Everett, X. Goaoc, S. Lazard, H.-S. Na, and S. Petitjean. The expected number of 3D visibility events is linear. *SIAM Journal on Computing*, 32(6) :1586–1620, 2003.

[47] D. P. Dobkin and D. G. Kirkpatrick. Fast detection of polyhedral intersection. *Theoretical Computer Science*, 27(3) :241–253, 1983.

[48] S. E. Dorward. A survey of object-space hidden surface removal. *Int. J. Computational Geometry and Applications*, 4 :325–362, 1994.

[49] G. Drettakis and E. Fiume. A fast shadow algorithm for area light sources using backprojection. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 223–230, New York, 1994. ACM Press.

[50] F. Duguet. Implémentation robuste du squelette de visibilité. Master's thesis, INRIA Sophia-Antipolis, 2001.

[51] F. Duguet. Shadow computations using robust epsilon visibility. Research Report 5167, INRIA, 2004.

[52] F. Duguet and G. Drettakis. Robust epsilon visibility. In John Hughes, editor, *Proceedings of ACM SIGGRAPH 2002*, pages 567–575. ACM Press / ACM SIGGRAPH, July 2002.

[53] F. Durand. *Visibilité tridimensionnelle : étude analytique et applications.* PhD thesis, Université Joseph Fourier - Grenoble I, 1999.

[54] F. Durand. A multidisciplinary survey of visibility, 2000. ACM SIGGRAPH course notes, Visibility, Problems, Techniques, and Applications.

[55] F. Durand, G. Dretakkis, and C. Puech. Fast and accurate hierarchical radiosity using global visibility. *ACM Transactions on Graphics*, 18(2) :128–170, 1999.

[56] F. Durand, G. Dretakkis, and C. Puech. The 3D visibility complex. *ACM Transactions on Graphics*, 21(2) :176–206, 2002.

[57] F. Durand, G. Drettakis, and C. Puech. The visibility skeleton : a powerful and efficient multi-purpose global visibility tool. *Computer Graphics* Proceedings, 31 :89–100, 1997. Proceedings of Siggraph'97.

[58] H. Edelsbrunner, L. J. Guibas, and J. Stolfi. Optimal point location in a monotone subdivision. *SIAM Journal on Computing*, 15(2), 1986.

[59] A. Efrat, L. Guibas, O. Hall-Holt, and L. Zhang. On incremental rendering of silhouette maps of a polyhedral scene. *Computational Geometry : Theory and Applications*, 38 :129–138, 2007. (Also in Proceedings of SODA'00).

[60] H. Everett, S. Lazard, B. Lenhart, J. Redburn, and L. Zhang. Predicates for line transversals in 3D, 2006. Accepted in *Computational Geometry : Theory and Applications*. (Also in Proceedings of 18th Canadian Conference on Computational Geometry - CCCG'06.).

[61] S. J. Fortune. A sweepline algorithm for Voronoi diagrams. *Algorithmica*, 2 :153–174, 1987.

[62] A. S. Glassner. *An introduction to ray tracing.* Morgan Kaufmann Publishers, San Francisco, 1989.

[63] M. Glisse. An upper bound on the average size of silhouettes. In Proceedings of the *22nd ACM Annual Symposium on Computational Geometry (SoCG'06)*, pages 105–111, June 2006.

[64] M. Glisse and H. Brönnimann. Octrees with near optimal cost for ray-shooting. *Computational Geometry : Theory and Applications*, 34(3) :182–194, 2006.

[65] X. Goaoc. *Structures de visibilité globales : tailles, calculs et dégénérescences.* Thèse d'université, Université Nancy 2, May 2004.

[66] J. Goldsmith and J. Salmon. Automatic creation of object hierarchies for ray tracing. *IEEE Comput. Graph. Appl.*, 7(5) :14–20, 1987.

[67] D. Halperin and M. Sharir. New bounds for lower envelopes in three dimensions, with applications to visbility in terrains. *Discrete and Computational Geometry*, 12 :313–326, 1994.

[68] J-M. Hasenfratz, M. Lapierre, N. Holzschuch, and F. Sillion. A survey of real-time soft shadows algorithms. In *Eurographics*, 2003.

[69] V. Havran, J. Prikryl, and J. Bittner. The best efficiency scheme, 2000.

[70] P. S. Heckbert. Discontinuity meshing for radiosity. In *Proceedings of the Third Eurographics Workshop on Rendering*, pages 203–215, May 1992.

[71] E. Helly. Über Mengen konvexer Körper mit gemeinschaftlichen Punkten. *Jahresbericht Deutsch. Math. Verein.*, 32 :175–176, 1923.

[72] S. Hornus. *Maintenance de la visibilité depuis un point mobile, et applications.* PhD thesis, Université Joseph Fourier - Grenoble I, 2006.

[73] D. P. Huttenlocher, K. Kedem, and M. Sharir. The upper envelope of Voronoi surfaces and its applications. *Discrete Comput. Geom.*, 9 :267–291, 1993.

[74] T. Isenberg, B. Freudenberg, N. Halper, S. Schlechtweg, and T. Strothotte. A developer's guide to silhouette algorithms for polygonal models. *IEEE Comput. Graph. Appl.*, 23(4) :28–37, 2003.

[75] S. Jadhav, A. Mukhopadhyay, and B. K. Bhattacharya. An optimal algorithm for the intersection radius of a set of convex polygons. *J. Algorithms*, 20 :244–267, 1996.

[76] K. Kedem, R. Livne, J. Pach, and M. Sharir. On the union of Jordan regions and collision-free translational motion amidst polygonal obstacles. *Discrete and Computational Geometry*, 1 :59–71, 1986.

[77] L. Kettner and E. Welzl. Contour edge analysis for polyhedron projections. In W. Strasser, R. Klein, and R. Ra, editors, *Geometric Modeling : Theory and Practice*, pages 379–394. Springer, 1997.

[78] C. Lund and M. Yannakakis. On the hardness of approximating minimization problems. *Journal of the ACM*, 41(5) :960–981, 1994.

[79] D. J. MacDonald and K. S. Booth. Heuristics for ray tracing using space subdivision. *Vis. Comput.*, 6(3) :153–166, 1990.

[80] P. Mamassian, D. C. Knill, and D. Kersten. The perception of cast shadows. *Trends in Cognitive Sciences*, 2(8) :288–295, 1998.

[81] M. McGuire. Observations on silhouette sizes. *Journal of Graphics Tools*, 9(1) :1–12, 2004.

[82] N. Megiddo. Applying parallel computation algorithms in the design of serial algorithms. *J. ACM*, 30(4) :852–865, 1983.

[83] K. Mehlhorn, S. Meiser, and R. Rasch. Furthest site abstract Voronoi diagrams. *Int. J. Comput. Geom. & Appl.*, 11(6) :583–616, 2001.

[84] D. W. Moore. *Simplicial mesh generation with applications*. PhD thesis, Cornell University, Ithaca, NY, USA, 1992.

[85] K. Mulmuley. A fast planar partition algorithm, I. *J. Symbolic Comput.*, 10(3-4) :253–280, 1990.

[86] B. Naylor. Constructing good partitioning trees. In Proceedings of *Graphics Interface (GI'93)*, pages 181–191. Morgan Kaufmann Publishers, San Francisco, 1993.

[87] T. Nishita and E. Nakamae. Half-tone representation of 3D objects illuminated by area sources or polyhedron sources. In *IEEE Computer Society's 7th International Computer Software and Applications Conference (COMPSAC '83)*, pages 237–242, 1983.

[88] J. O'Rourke. *Computational Geometry in C*. Cambridge Univsersity Press, 2nd edition, 1998.

[89] M. Pellegrini. On lines missing polyhedral sets in 3-space. *Discrete and Computational Geometry*, 12 :203–221, 1994.

[90] M. Pocchiola and G. Vegter. The visibility complex. *International Journal of Computational Geometry and Applications*, 6(3) :279–308, 1996. Proceedings of the 9th ACM Annual Symposium on Computational Geometry (SoCG'93).

[91] E. Reinhard, A. J. F. Kok, and F. W. Jansen. Cost Prediction in Ray Tracing. In *Rendering Techniques '96 (Proceedings of the Seventh Eurographics Workshop on Rendering)*, pages 41–50, New York, NY, 1996. Springer-Verlag/Wien.

[92] G. Salmon. *Analytic Geometry of Three Dimensions*. Cambridge, 1978.

[93] H. Samet. *The design and analysis of spatial data structures*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1990.

[94] P. V. Sander, Xianfeng Gu, Steven J. Gortler, Hugues Hoppe, and John Snyder. Silhouette clipping. In *SIGGRAPH '00 : Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, pages 327–334, New York, NY, USA, 2000. ACM Press / Addison-Wesley Publishing Co.

[95] L. A. Santalo. *Integral Geometry and Geometric Probability*. Cambridge University Press, New York, NY, 2004. Second edition.

[96] H. A. Schwarz. Sur une définition erronée de l'aire d'une surface courbe. In *Gesammelte Mathematische Abhandlungen*, volume 1, pages 309–311. Springer-Verlag, 1890.

[97] R. Seidel. Small-dimensional linear programming and convex hulls made easy. *Discrete and Computational Geometry*, 6(5) :423–434, 1991.

[98] M. Sharir and P. K. Agarwal. *Davenport-Schinzel Sequences and Their Geometric Applications*. Cambridge University Press, New York, 1995.

[99] M. Sharir and P. K. Agarwal. *Davenport-Schinzel Sequences and their Geometric Applications*. Cambridge University Press, 1995.

[100] M. Sharir and E. Welzl. A combinatorial bound for linear programming and related problems. In *STACS '92 : Proceedings of the 9th Annual Symposium on Theoretical Aspects of Computer Science*, pages 569–579. Springer-Verlag, 1992.

[101] A. J. Stewart and S. Ghali. Fast computation of shadow boundaries using spatial coherence and backprojections. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 231–238. ACM Press, 1994.

[102] J. Stolfi. *Oriented Projective Geometry : A Framework for Geometric Computations*. Academic Press, 1991.

[103] K. R. Subramanian and D. S. Fussell. Automatic termination criteria for ray tracing hierarchies. In *Proceedings of Graphics Interface '91*, pages 93–100, Toronto, Ontario, 1991. Canadian Information Processing Society.

[104] S. J. Teller. Computing the antipenumbra of an area light source. In *Computer Graphics* Proceedings, Annual Conference Series, ACM SIGGRAPH, pages 139–148. ACM Press, 1992.

[105] M. van Kreveld and T. Schlechter. Automated label placement for groups of islands. In *Proc. of the 22nd International Cartographic Conference*, 2005.

[106] L. Wanger. The effect of shadow quality on the perception of spatial relationships in computer generated images. *Computer Graphics*, 25(2) :39–42, 1992.

[107] H. Weghorst, G. Hooper, and D. P. Greenberg. Improved computational methods for ray tracing. *ACM Trans. Graph.*, 3(1) :52–69, 1984.

[108] K.-Y. Whang, J.-W. Song, J.-W. Chang, J.-Y. Kim, W.-S. Cho, C.-M. Park, and I.-Y. Song. Octree-R : An adaptive octree for efficient ray tracing. *IEEE Transactions on Visualization and Computer Graphics*, 1(4) :343–349, 1995.

[109] A. Wiernik and M. Sharir. Planar realizations of nonlinear Davenport-Schinzel sequences by segments. *Discrete and Computational Geometry*, 3 :15–47, 1988.

[110] A. Woo, P. Poulin, and A. Fournier. A survey of shadow algorithms. *IEEE Computer Graphics and Applications*, 10(6) :13–32, 1990.

[111] C. K Yap. Towards exact geometric computation. *Computational Geometry : Theory and Applications*, 7(1-2) :3–23, 1997.