



HAL
open science

Gestion de grands catalogues et application de relevés infrarouges à l'étude de la structure galactique

Sebastien Derriere

► **To cite this version:**

Sebastien Derriere. Gestion de grands catalogues et application de relevés infrarouges à l'étude de la structure galactique. Astrophysique [astro-ph]. Université Louis Pasteur - Strasbourg I, 2001. Français. NNT: . tel-00004024

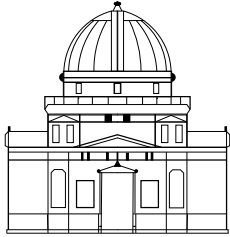
HAL Id: tel-00004024

<https://theses.hal.science/tel-00004024>

Submitted on 18 Dec 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



*Observatoire Astronomique
de Strasbourg*



THÈSE DE DOCTORAT

présentée par

Sébastien Derriere

Gestion de grands catalogues et application de relevés infrarouges à l'étude de la structure galactique

Soutenue le 15 juin 2001, devant la commission d'examen :

Présidente du jury	Mme Agnès Acker, Professeur
Rapporteurs	M. Olivier Bienaymé, Astronome M. James Lequeux, Astronome honoraire M. Mehrdad Moshir, Directeur de recherche
Directeurs de thèse	M. Daniel Egret, Astronome Mme Annie Robin, DR CNRS
Invité	M. Nicolas Epchtein, DR CNRS

À mon grand-père.

Remerciements

Je tiens tout d'abord à remercier chaleureusement Daniel Egret et Annie Robin pour m'avoir encadré au cours de cette thèse. J'ai particulièrement apprécié leur gentillesse et leur disponibilité (malgré les responsabilités de directeur de Daniel, et la distance me séparant de Besançon), tout comme leur compétence et leur rigueur.

Je remercie également les rapporteurs Olivier Bienaymé, James Lequeux et Mehrdad Moshir, pour avoir pris le temps de lire cette thèse, et m'avoir permis d'améliorer le manuscrit par leurs remarques, ainsi qu'Agnès Acker pour avoir accepté de présider le Jury.

Je remercie à travers Nicolas Epchtein toute l'équipe du projet DENIS, pour m'avoir permis de collaborer à l'aventure.

Je ne voudrais pas oublier les personnes de l'Observatoire de Strasbourg, notamment François Ochsenbein qui à toujours réponse à tout avec le sourire, Françoise Genova, directrice d'équipe hors pair, et l'équipe informatique qui assure quotidiennement derrière Jean-Yves Hangouët la bonne marche des machines.

Je remercie aussi ma famille pour leur soutien constant. Vous pouvez enfin arrêter de me demander combien d'années d'études il me reste à faire !

Difficile épreuve que de résumer en quelques lignes des années passées aux cotés de tous ceux qui donnent un sens au mot amitié. À maître Nico pour les parties de Blast endiablées, à Sophie qui peut encore me battre à bd, à Nathalie pour son énergie et son sourire si communicatifs, à Patrick pour les années 80 (j'y arriverai, un jour), à Laurent pour (ou malgré?) son humour, à Arnaud pour ses 62 recettes de saumon (on remet ca quand?), à Sandrine pour sa tenacité à Kahuna, à Jérôme le karateka (zen, zen, zen), à Nausicaa la grenouille, à Philippe pour les leçons de photo, de guitare et son talent à finir les plats, mais aussi à Lysou pour sa hot-line, à Eric pour être toujours là depuis des années, et à tous ceux que (honte à moi) j'oublie dans ma précipitation, j'adresse un énorme merci.

Enfin merci à celle dont la présence à mes côtés jour après jour est le plus beau des cadeaux.

Table des matières

Introduction	1
I Gestion des grands catalogues	3
1 Grands catalogues et relevés infrarouges	5
1.1 Introduction	5
1.1.1 Cataloguer le ciel	5
1.1.2 L’astronomie infrarouge	9
1.2 Le projet DENIS	13
1.2.1 Caractéristiques de l’instrument	13
1.2.2 Les observations	16
1.2.3 La chaîne de traitement des données	17
1.2.4 Objectifs scientifiques et premiers résultats	19
1.3 Le projet 2MASS	21
1.3.1 Fiche technique	21
1.3.2 Stratégie d’observation	22
1.3.3 Traitement des données	25
1.3.4 Objectifs scientifiques et résultats	27
1.4 Les catalogues de sources ponctuelles	28
1.4.1 Caractéristiques	28
1.4.2 Le catalogue DENIS	28
1.4.3 Le catalogue 2MASS	29
2 Compression et indexation	31
2.1 Manipulation de grands volumes de données	31
2.1.1 Position du problème	31
2.1.2 Techniques mises en œuvre	32
2.1.3 Un exemple simple de codage	34
2.2 Codage des données DENIS	35
2.2.1 Catalogues LDAC et PDAC	35
2.2.2 Aspects communs pour le codage	37
2.2.3 Astuces spécifiques et performances	40
2.3 Codage des données 2MASS	41

2.3.1	Introduction	41
2.3.2	Découpage du ciel	42
2.3.3	Codage	43
2.3.4	Résultats	47
3	Validation des données DENIS	49
3.1	Une étape essentielle	49
3.2	Quelques notions d'astrométrie	50
3.2.1	Systèmes de référence : FK5, ICRS	50
3.2.2	Catalogues astrométriques	51
3.3	Les données LDAC	53
3.3.1	Validation de l'astrométrie	53
3.3.2	Photométrie et filtrage des parasites	60
3.4	Les données PDAC	62
3.4.1	Astrométrie	62
3.4.2	Photométrie et filtrage des parasites	63
3.5	Comparaisons LDAC–PDAC	78
3.5.1	Astrométrie	78
3.5.2	Photométrie	79
3.5.3	Conclusion	80
4	Accès en ligne	83
4.1	L'astronomie en ligne	83
4.1.1	Vers le « tout-numérique »	83
4.1.2	Le rôle du CDS	84
4.2	Les catalogues DENIS et 2MASS	85
4.2.1	Accès au CDS	85
4.2.2	Les pages Web DENIS	87
4.3	Interfaces dédiées	88
II	Structure galactique	93
5	La Voie Lactée	95
5.1	Historique	96
5.1.1	Un univers-île	96
5.1.2	Nébuleuses et Galaxie	96
5.1.3	Le Grand Débat	98
5.2	La structure de la Galaxie	101
5.2.1	Une galaxie spirale ordinaire	101
5.2.2	Le bulbe	105
5.2.3	Le halo	107
5.2.4	Le disque	109

6	Modèles de la Galaxie	115
6.1	Quelques exemples	115
6.1.1	Modèles de masse	115
6.1.2	Modèles de première génération	116
6.1.3	Les années 1990	117
6.2	Le modèle de Besançon	118
6.2.1	Caractéristiques du modèle	118
6.2.2	Fonctions de luminosité	119
6.2.3	Les lois de densité	123
6.2.4	Cohérence dynamique	124
6.2.5	Catalogues de pseudo-étoiles	126
6.3	Gauchissement et évasement	128
6.3.1	Description	128
6.3.2	Prise en compte dans le modèle	130
7	Relevés IR et structure galactique	135
7.1	Sélection des données	135
7.1.1	Les catalogues PDAC	136
7.1.2	Zones sélectionnées	136
7.2	Premiers ajustements	141
7.2.1	Observables et estimateurs	142
7.2.2	Paramètres de densité	144
7.3	Effets de l'extinction	146
7.4	Gauchissement et évasement	149
	Conclusions et perspectives	153
	Bibliographie	155
A	Notations	171
B	Identification croisée	173
B.1	Rappels de Statistique	173
B.1.1	Flux de Poisson 2D	173
B.1.2	Flux Normal 2D	174
B.2	Principe de l'identification croisée	175
B.2.1	Introduction au problème	175
B.2.2	Difficultés rencontrées	175
B.3	Le cas du plus proche voisin	176
B.3.1	Positions exactes	176
B.3.2	Incertitude sur les positions	177
B.3.3	Formalisme	178
B.3.4	Vraisemblance des identifications	179
B.4	Simulation numérique	180
B.5	Données réelles	182

B.6	Vers une approche plus générale	182
C	Publications	187
C.1	Journaux à comité de lecture	187
C.2	Autres publications	187

Table des figures

1.1	<i>Almageste</i> de Ptolémée	6
1.2	Atlas céleste de Fortin	8
1.3	Transmittivité de l’atmosphère dans l’infrarouge proche	11
1.4	L’instrument DENIS sur site	14
1.5	Montage optique de l’instrument DENIS	15
1.6	Transmission des filtres DENIS	15
1.7	Stratégie observationnelle du projet DENIS	16
1.8	Extinction dans les régions centrales du bulbe	20
1.9	Chemin optique de la caméra à 3 canaux de 2MASS	22
1.10	Stratégie observationnelle pour 2MASS	23
1.11	Création des images combinées 2MASS	24
1.12	Organigramme du 2MAPPS	26
2.1	Exemple simple de codage binaire	33
2.2	Codage en binaire des catalogues DENIS issus du LDAC	38
2.3	Structure d’un fichier binaire pour un strip DENIS	40
2.4	Zones couvertes par la deuxième publication de données 2MASS	42
2.5	Optimisation du découpage des fichiers pour 2MASS	44
2.6	Codage en binaire du catalogue 2MASS	45
2.7	Codage binaire à deux modes	46
3.1	Recherche de décalage astrométrique	55
3.2	Contrôle de l’astrométrie : exemple de bonne calibration	56
3.3	Exemples de mauvaise calibration astrométrique	58
3.4	« Trous » de l’USNO A2.0 dans la barre du LMC	59
3.5	Validation photométrique au LDAC	61
3.6	Diagrammes couleur-couleur des catalogues LDAC	61
3.7	Artéfacts dans la base PDAC	65
3.8	Validation et filtrage PDAC	67
3.9	Exemple de filtrage au PDAC	77
3.10	Comparaison astrométrique LDAC-PDAC	78
3.11	Diagrammes de complétude PDAC et LDAC	79
3.12	Diagrammes couleur-magnitude PDAC et LDAC	81
4.1	Requête sur le catalogue DENIS dans Vizier	87

4.2	Utilisation du catalogue DENIS dans ALADIN	90
4.3	Exemple d'informations sur un strip DENIS	91
5.1	La Galaxie selon Kapteyn et Shapley	99
5.2	La Voie Lactée dans différentes longueurs d'onde	104
5.3	Courbe de rotation pour la Voie Lactée	109
5.4	NGC3198	110
6.1	Les 4 populations du modèle de Besançon	121
6.2	Production d'un catalogue de pseudo-étoiles	132
6.3	ESO 510-13 : une galaxie gauchie	133
6.4	Gauchissement et évasement	134
7.1	Couverture du relevé DENIS	136
7.2	Zones et strips sélectionnés	141
7.3	Limite de complétude	143
7.4	Profil d'extinction : données simulées	148
7.5	Profil d'extinction : application	149
7.6	Effet du gauchissement et de l'évasement	150
7.7	Isocontours $R_{\text{warp}} - R_{\text{flare}}$	151
B.1	Distance au ppv pour un flux de Poisson 2D	174
B.2	Distribution $d(\text{ppv})$ pour des catalogues simulés	181
B.3	Vraisemblance pour des catalogues simulés	181
B.4	Distribution $d(\text{ppv})$ entre les catalogues DENIS et USNO	184
B.5	Vraisemblance de l'identification croisée des catalogues DENIS et USNO	185

Introduction

CETTE thèse coïncide avec l'arrivée à maturité de deux grands projets de cartographie du ciel dans l'infrarouge proche : DENIS et 2MASS. Ces programmes, avec d'autres relevés dans le domaine optique, marquent un progrès décisif de l'astronomie à l'aube du XXI^e siècle : les catalogues d'objets produits par ces grands relevés couvrent une grande partie du ciel, et fournissent des informations pour plusieurs dizaines de millions de sources, ce qui représente un gain de plusieurs ordres de grandeur par rapport aux catalogues antérieurs.

Les volumes de données considérables mis en jeu par les catalogues de sources ponctuelles des projets DENIS et 2MASS (un vaste ensemble de paramètres pour des dizaines de millions d'objets) nécessitent des outils et des méthodes de gestion adaptés, que ce soit pour une simple requête autour d'une position céleste, ou pour la recherche de catégories d'objets particuliers.

De plus, en combinant une grande sensibilité et des surfaces d'observation très grandes, ces projets dépassent de très loin les données dans l'infrarouge proche disponibles précédemment. La faible extinction interstellaire à ces longueurs d'onde fait de ces catalogues un support de choix pour l'étude de la structure de notre Galaxie.

Le travail mené au cours de cette thèse comporte donc deux axes majeurs, qui seront développés dans deux parties distinctes :

1. le développement de méthodes de gestion pour ces grands relevés ;
2. leur utilisation pour l'étude de la structure galactique.

Gestion des grands catalogues

Les grands relevés astronomiques modernes, tels DENIS et 2MASS, ainsi que les programmes de numérisation de plaques photographiques (USNO), produisent des catalogues de sources dont la taille peut atteindre quelques dizaines de gigaoctets. Les bases de données classiques ne sont pas adaptées pour gérer de façon optimale de tels volumes de données.

Je relate rapidement dans le premier chapitre l'évolution historique des catalogues d'objets en astronomie, puis je présente les projets DENIS et 2MASS, dont les catalogues de sources ponctuelles seront exploités dans les chapitres suivants.

Afin d'avoir un accès rapide au contenu de ces catalogues, nous avons utilisé des méthodes de compression binaires optimisées, mais sans perte d'information, pour créer une base de données dédiée, tout en gardant une indexation de l'ensemble de

la base en positions célestes. On réduit ainsi au maximum les temps de lecture, et on facilite les requêtes par coordonnées. Ces méthodes de compression et d'indexation font l'objet du deuxième chapitre.

Le chapitre 3 décrit le travail de validation des données DENIS que j'ai effectué, afin de contrôler la qualité de l'astrométrie et de la photométrie.

Les catalogues, codés en binaire, sont accessibles et interrogeables depuis le Web. Les modalités d'interrogation et les interfaces sont exposées dans le chapitre 4.

Structure galactique

La compréhension de la nature et de structure de la Voie Lactée s'est faite pour l'essentiel au cours des 100 dernières années. Je rappelle dans le chapitre 5 les principales étapes historiques dans l'étude de la structure galactique, ainsi que la vision moderne des différents constituants de la Voie Lactée.

La modélisation de la Galaxie joue un rôle fondamental dans l'étude de la structure galactique. Le chapitre 6 décrit quelques modèles galactiques, et je m'attarde particulièrement sur le modèle de Besançon que j'ai utilisé par la suite : ce modèle, qui tient compte à la fois de l'évolution stellaire, de l'évolution galactique et de la dynamique galactique, permet de créer des catalogues synthétiques d'étoiles ayant les mêmes propriétés que les catalogues observés, pour n'importe quelle direction. Le modèle utilise différentes lois de densités pour simuler les différentes populations galactiques, et reproduit les erreurs observationnelles. On joue sur les paramètres des lois de densités pour ajuster le modèle aux observations.

J'ai apporté une modification au modèle de Besançon, afin de rendre compte de deux effets non pris en compte auparavant : le gauchissement et l'évasement du disque galactique. Ces déformations du disque de la Voie Lactée sont connues depuis longtemps dans la couche d'hydrogène neutre, mais n'avaient été que peu étudiées dans le disque stellaire.

Le dernier chapitre est consacré à la confrontation du modèle de Besançon avec des données DENIS. Le principal avantage de l'infrarouge proche est que l'extinction y est bien moindre que dans le visible (on a un facteur 10 entre les bandes V et K_s). On peut donc détecter les étoiles jusqu'à des distances de plusieurs kpc, y compris à basse latitude galactique, beaucoup plus facilement que dans le domaine optique.

Les données DENIS permettent donc de contraindre les paramètres structurels, et en particulier de déterminer la forme du gauchissement et de l'évasement du disque galactique.

Première partie

Gestion des grands catalogues

Astronomers decay faster than plates.

D. Monet

Chapitre 1

Grands catalogues et relevés infrarouges

1.1 Introduction

1.1.1 Cataloguer le ciel

DÉPUIS l'Antiquité, les hommes ont dressé des listes des objets qu'ils aperçoivent dans le ciel nocturne. On a trouvé la trace d'observations astronomiques au temple de Baal (Babylone) dès 2500 av. J.C., et en Chine vers 2300 av. J.C. (sous l'empereur Yao). Les astronomes grecs et romains ont attribué aux étoiles les plus brillantes des noms¹ encore utilisés actuellement, et les ont regroupées en constellations qui figurent les grands événements de la mythologie.

A l'œil nu

Un des premiers catalogues dont on ait une trace, avec une liste d'objets et leur position dans le ciel, est celui d'Aristillus et Timocharis (-284). Ils avaient mesuré les positions des étoiles les plus brillantes à l'aide d'une arbalète (ou bâton de Jacob).

En -204, Eratosthène répertorie plus de 700 étoiles en utilisant des sphères armillaires pour mesurer les positions. Hipparque établit vers 128 av. J.C. un nouveau catalogue de plus de 850 étoiles, et il met en évidence la précession des équinoxes en comparant les positions des étoiles à celles de catalogues plus anciens. Ptolémée reprend dans l'*Almageste* (vers 150 après J.C.) les travaux d'Hipparque (il y a polémique pour savoir s'il a vraiment observé lui-même ou s'il a juste corrigé les positions de la précession). On trouve dans l'*Almageste* les positions et une indication de magnitude pour un peu plus d'un millier d'étoiles (Fig. 1.1).

La situation n'évolue guère tant que l'œil reste le seul outil d'observation ; de

¹Répertorier et nommer sont deux activités fortement liées : tout catalogue astronomique répertoriant des objets les désigne par un identificateur unique (son nom). Un même objet détecté dans plusieurs catalogues pourra ainsi être désigné par plusieurs identificateurs. On recense par exemple, pour Véga, 44 « noms » différents.

Octava 86

Longitudo et Latitudo ac Magnitudo stellarum fixarum

Forme et Stelle natur Tenze.	Imago Trigesimaquinta	Longitudo			Latitudo			Magnitudo
		g	m	sec	g	m	sec	
Septentrionalis que est in capite sublimati sine audacia		1	27	0	M	18	50	nebulosa
Lucida que est super humerum dextrum: et ipsa redit ad rapinam que appropinquat ad terram in humero orionis		2	2	0	M	17	0	1 .e.l.
Que est super humerum sinistrum		1	20	20	M	17	30	2 .e.m.
Sequens que est sub istis duabus		1	25	0	M	18	0	4 .e.l.
Que est super cubitum dextrum		2	4	20	M	14	30	4
Que est super brachium dextrum		2	6	20	M	11	50	6
Sequens duplex meridionalis quadrilateri quod est in palma dextra		2	6	30	M	10	40	4
Antecedens lateris meridionalis		2	6	0	M	9	45	4
Sequens lateris septentrionalis		2	7	20	M	8	15	6
Antecedens lateris septentrionalis		2	6	40	M	8	15	6
Antecedens duarum que sunt in figura pineali		2	1	40	M	3	45	5
Sequens earum		2	4	20	M	3	15	5
Sequens quatuor que sunt quasi super lineam rectam super dorsum		1	27	30	M	19	40	4
Antecedens banc		1	26	20	M	20	0	6
Antecedens etiam banc		1	25	20	M	20	20	6
Reliqua et est antecedens quatuor		1	24	10	M	20	40	5
Longior novem que sunt in dorso manus sinistre in septentrione		1	20	30	M	8	0	4
Secunda post istam in septentrione		1	19	20	M	8	10	4
Tertia post eam in septentrione		1	18	0	M	10	15	4
Quarta post eam in septentrione		1	16	20	M	12	50	4
Quinta post eam in septentrione		1	15	10	M	14	15	4
Sexta post eam in septentrione		1	14	30	M	15	50	3
Septima post eam in septentrione		1	14	50	M	17	10	3
Octava post eam etiam in septentrione		1	15	20	M	20	20	3
Reliqua et novem ultima a meridie		1	16	20	M	21	30	3
Antecedens trium que sunt super cingulam		1	25	20	M	24	10	2
Media earum		1	27	20	M	24	50	2
Sequens trium		1	28	10	M	25	40	2
Que est apud capulum ensis		1	23	50	M	25	50	3
Septentrionalis trium coniunctarum cum capite ensis		1	26	50	M	28	40	4 .e.l.
Media earum		1	26	40	M	29	40	3
Meridionalis trium		1	27	0	M	29	50	3
Sequens duarum que sunt sub extremitate ensis		1	27	40	M	30	40	4
Antecedens earum		1	26	10	M	30	50	4
Lucida que est in pede sinistro: et est communis ei et aque		1	19	10	M	31	30	1
Que est super declinationem ea ad septentrionem: et est super calcaneum		1	21	0	M	30	15	4 .e.m.
Que est super calcaneum sinistrum exterius		1	23	20	M	31	10	4
Que est super genu dextrum septentrionale		1	0	10	M	33	30	3
Illarum triginta octo stellarum in magnitudine prima sunt due. in secunda quatuor. in tertia octo. in quarta quindecim. in quinta tres. in sexta quinq. et nebulosa una.								
Stellatio Sexu.		Imago Trigesimasexta						
Que est post illam que est in pede sublimati sine audacia super principium fluminis		1	18	20	M	31	50	4
Que est declinatio hac ad septem. et est in tortuositate apud opphedeete crus sublimati sine audacia		1	18	50	M	28	15	4
Sequens duarum continuarum que sunt post banc		1	18	0	M	29	50	4
Antecedens earum		1	14	40	M	28	15	4
Sequens duarum continuarum etiam		1	13	10	M	29	15	4
Antecedens earum		1	20	10	M	25	20	4
Sequens trium que sunt post istam		1	6	20	M	26	0	4
Media earum		1	5	30	M	27	0	4
		1	2	50	M	27	50	4

Fig. 1.1: Extrait de l'Almageste de Ptolémée. Cette page correspond aux étoiles de la constellation d'Orion. Ptolémée donne une indication de la brillance des étoiles : les plus brillantes ont une magnitude de 1. Le catalogue peut être consulté en ligne : <http://vizier.u-strasbg.fr/cgi-bin/VizieR?source=V/61> .

plus, l'idée de *sphère des fixes*, commune aux systèmes d'Aristote, d'Hipparque ou de Ptolémée, et qui impose un univers fini et immuable, ne sera remise en cause qu'après les travaux de Copernic, par Thomas Digges (1576) et Giordano Bruno (1584). L'*Almageste*, malgré une précision sur les positions de l'ordre du degré, reste donc la référence pendant près de 15 siècles.

Parmi les astronomes orientaux, on peut citer le prince Ulugh Begh, qui aménage à Samarkand au XV^e siècle un observatoire où il effectue des mesures d'une grande précision (quelques minutes d'arc), grâce à un immense quart-de-cercle méridien (40,2 m de rayon). Malheureusement, ses travaux restent inconnus des astronomes occidentaux.

Avec les grands navigateurs, comme Magellan, les objets les plus brillants de l'hémisphère sud vont être découverts. Mais la plus grande avancée avant l'apparition de la lunette sera la publication par Tycho Brahe (1602) d'un catalogue d'environ 1200 étoiles, avec une précision nettement améliorée par rapport à Ptolémée (on atteint la minute d'arc). C'est sur la base de ce catalogue que Johann Bayer publie en 1603 dans *Uranometria* le premier véritable atlas du ciel : on y retrouve les 48 constellations de Ptolémée, et chaque étoile se voit attribuer une lettre grecque à l'intérieur de chaque constellation, par ordre de magnitude décroissante².

Lunettes et télescopes

L'utilisation d'instruments optiques va provoquer un bond en avant de l'astronomie. En 1609, Galilée utilise pour la première fois une lunette pour observer le ciel. Il parvient ainsi à distinguer dans la Voie Lactée de nombreuses étoiles ne pouvant pas être résolues à l'œil nu.

C'est la grande époque des atlas du ciel : les étoiles connues sont représentées sur des cartes illustrées par les figures mythologiques. L'atlas de Johannes Hevelius, *Firmamentum Sobiescianum sive Uranographia*, publié en 1690, contient non seulement ses propres observations dans l'hémisphère nord, mais aussi des cartes de l'hémisphère sud répertoriant les 341 étoiles observées par Halley depuis St. Hélène en 1676. L'*Atlas coelestis* de John Flamsteed, publié en 1729, est un des plus connus, et a été repris et amélioré par Fortin dans l'édition française de 1776 (Fig. 1.2). La précision sur les positions est d'environ 10'' dans cet atlas.

La puissance croissante des lunettes, puis des télescopes qui les supplantent vers 1730 permet d'obtenir des catalogues de plus en plus volumineux. William Herschel recense 2478 objets nébuleux à la fin des années 1700. Johann Bode publie en 1801 *Uranographia* : cet ouvrage comprenant listes d'objets et cartes illustrées est alors l'atlas le plus complet jamais paru, car il reprend les objets du catalogue de Herschel, obtenant en tout plus de 17 200 étoiles. Ce travail va également marquer la fin des atlas célestes : les figures mythologiques vont disparaître des recueils d'astronomie pour laisser la place à des catalogues et à des cartes moins imagées, mais toujours plus précises.

²Voir <http://www.virtualsky.org/uranometria/>.

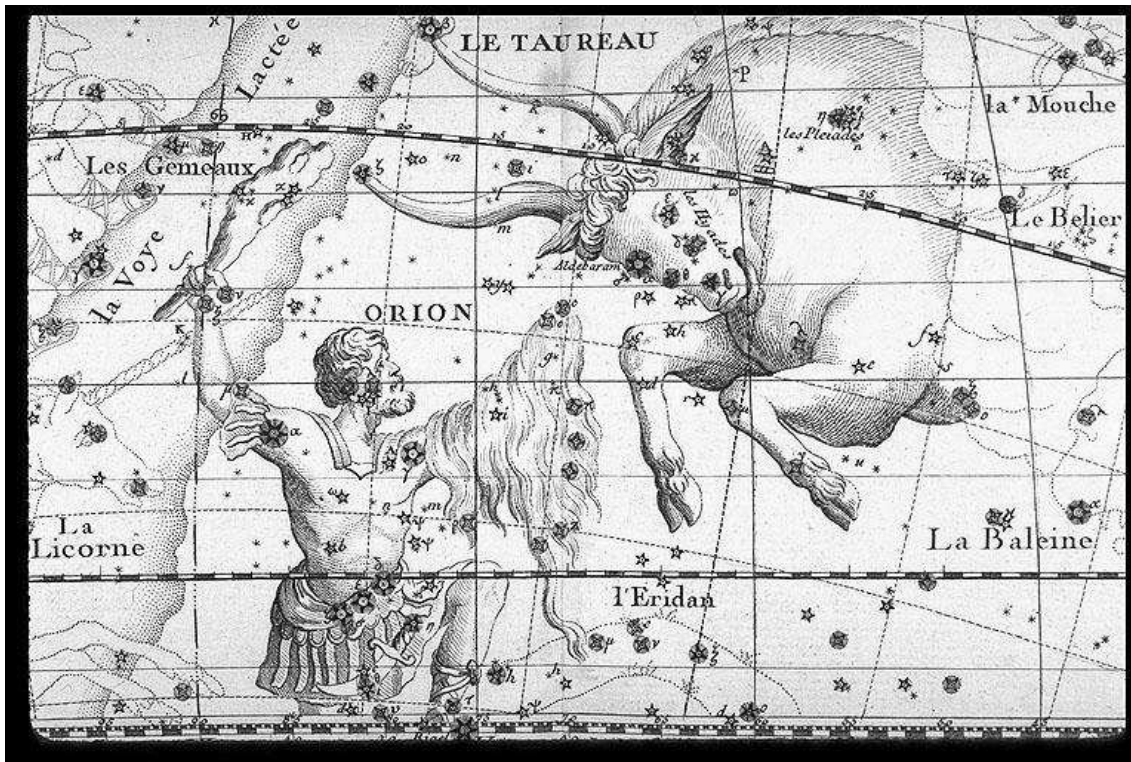


Fig. 1.2: La région du Taureau dans l'atlas de Fortin (1776). Fortin a repris l'Atlas Céleste de Flamsteed, en améliorant sensiblement la qualité artistique des illustrations.

Les télescopes de diamètre croissant permettent en effet non seulement de découvrir de nouveaux objets, mais aussi d'améliorer toujours la précision des mesures.

Les plaques photographiques

Vers 1880, des pionniers comme Henry Draper utilisent les récentes techniques photographiques pour obtenir des images d'objets astronomiques. C'est une révolution aussi importante que l'invention de la lunette, car il n'est plus nécessaire de tracer à la main des dessins de ce que l'on observe, et en faisant des poses assez longues, on peut accumuler les photons venant de sources faibles, ce qui augmente considérablement le nombre d'objets détectables.

On entre ainsi dans l'ère des catalogues modernes, qui répertorient de plus en plus d'objets. On peut citer parmi ces nouveaux catalogues :

- le *New General Catalogue* (NGC) de Dreyer (1888), qui recense 7 840 nébuleuses et amas d'étoiles. Avec l'*Index Catalogue* (IC), et le *Second Index Catalogue* (Dreyer, 1895; Dreyer, 1908), le nombre d'objets sera porté à 13 226 ;
- la série des *Durchmusterung Catalogues* : ces relevés visuels, puis photographiques du ciel, entrepris de 1859 à 1932, cataloguent plusieurs centaines de milliers d'étoiles ;

- le *Bright Star Catalogue* (1908), qui collecte des informations détaillées sur 9110 objets plus brillants que la magnitude 6.5 (identifiés par un numéro HR – *Harvard Revised Number*);
- le *Henry Draper catalogue* (Cannon & Pickering, 1924), qui contient avec son extension (Cannon, 1936) 272,150 étoiles ;
- l’ambitieux projet de Carte du Ciel (*Astrographic Catalogue*), dont l’objectif était de photographier et de déterminer les positions de toutes les étoiles plus brillantes que la magnitude 11. Le projet, lancé en 1891, a mobilisé 20 observatoires, chacun étant chargé de couvrir une zone de déclinaison. Au total, plus de 4.6 millions d’étoiles ont été observées, parfois jusqu’à la magnitude 13.0. Ces observations relativement anciennes (avant les années 1950) permettent aujourd’hui de déterminer les mouvement propres des étoiles brillantes en disposant d’une base de temps relativement importante.

Électronique et informatique : les grands catalogues

La fin du XX^e siècle voit un nouveau bond en avant : les progrès de l’informatique permettent de gérer d’énormes volumes de données, et les capteurs CCD (*Charge-Coupled Device*) apportent un gain considérable en sensibilité. Ces détecteurs électroniques peuvent être embarqués dans les missions spatiales, autre progrès décisif de cette fin de millénaire : en s’affranchissant de l’atmosphère, on peut désormais obtenir des catalogues à n’importe quelle longueur d’onde (rayons gamma, rayons X, UV, infrarouge, ondes millimétriques).

Les nouvelles technologies permettent aussi de retraiter des observations anciennes : la numérisation de plaques photographiques vieilles de plusieurs dizaines d’années (et en particulier des grands relevés du ciel effectués par divers observatoires) permet de créer des catalogues contenant des centaines de millions d’étoiles.

La table 1.1 présente quelques uns des grands catalogues astronomiques modernes. On voit que ces catalogues contiennent souvent 10^6 à 10^9 objets. Par la suite, on parlera de grands catalogues quand le nombre d’objets est supérieur à 10^7 .

1.1.2 L’astronomie infrarouge

Quelque soit le domaine spectral, deux points fondamentaux conditionnent les observations astronomiques au sol :

- la transparence de l’atmosphère à la longueur d’onde considérée ;
- l’existence d’un détecteur sensible à cette longueur d’onde.

Le domaine visible est évidemment le plus aisément accessible, parce que l’atmosphère y est transparente et que les détecteurs adaptés sont simples à mettre en œuvre. Si le domaine visible se termine pour l’œil humain dans le rouge vers 7000 \AA , on parle en général en astronomie de domaine optique jusqu’à des longueurs d’onde de $1 \text{ }\mu\text{m}$, et l’infrarouge s’étend de 1 à $350 \text{ }\mu\text{m}$. Cette séparation correspond à la fois à un changement du type de détecteurs utilisés et à un changement des propriétés de l’atmosphère. Voyons ce qui se passe à ces grandes longueurs d’onde.

Année	Catalogue	Nombre d'objets	Caractéristiques
1988	PPM nord	181 731	Positions et mouvements propres, (précision 0.27" et 0.43"/siècle)
1993	PPM sud	197 179	(complément pour l'hémisphère sud)
1989	IRAS	≈500 000	Infrarouge lointain (satellite)
1992	GSC 1.1	≈15 000 000	25 258 765 enregistrements, mais avec des détections multiples
1997	Tycho	1 058 332	Etoiles observées lors de la mission spatiale Hipparcos
1997	USNO A1.0	488 006 860	Positions, magnitudes R et B
1997	ACT	988 758	Combinaison de Tycho et de la Carte du Ciel (mouvements propres)
1998	USNO A2.0	526 280 881	Calibration sur l'ACT
2000	Tycho-2	2 539 913	Positions et mouvements propres (précision 60 mas et 2.5 mas/an)
2000	2MASS(*)	162 213 354	47% du ciel, infrarouge proche : J , H , K_s
2001	GSC-II(*)	≈2 000 000 000	2 magnitudes et mouvements propres

Tab. 1.1: *Caractéristiques de quelques catalogues récents ou en projet (les catalogues marqués d'une astérisque sont provisoires ou en projet). Les catalogues qui contiennent un nombre d'objets plus restreint sont en général bien plus précis.*

Observations au sol

Si l'atmosphère est bien transparente jusque 0.7 μm , sa transmittivité se dégrade dans l'infrarouge proche pour des longueurs d'onde croissantes (Fig 1.3).

Des bandes d'absorption de plus en plus larges, principalement dues à l'eau et au CO_2 , apparaissent. Les filtres larges couramment utilisés pour les observations astronomiques à ces longueurs d'onde (J , H , K , L) ont des bandes passantes se situant dans les fenêtres de bonne transmittivité entre les bandes d'absorption (les longueurs d'onde centrales de ces filtres sont respectivement 1.25, 1.65, 2.2 et 3.5 μm).

La transmission de l'atmosphère dépend en pratique fortement du lieu d'observation, en particulier de l'altitude et de l'humidité. C'est pourquoi il existe différents types de filtres selon les observatoires, et donc des systèmes photométriques légèrement différents³.

Un autre problème intervenant dans les observations au sol dans l'infrarouge est l'émissivité de l'atmosphère (*airglow*). En dessous de 2 μm , c'est surtout l'émission de radicaux OH^- qui domine, et pour de plus grandes longueurs d'onde, l'émission thermique de l'atmosphère devient dominante (on atteint la queue de distribution de Rayleigh-Jeans d'un corps noir à température de l'atmosphère).

Il faut également tenir compte de l'émission thermique de l'environnement immé-

³Dans le cadre du projet de télescope infrarouge Gemini, un important travail de normalisation des filtres infrarouges a été entrepris.

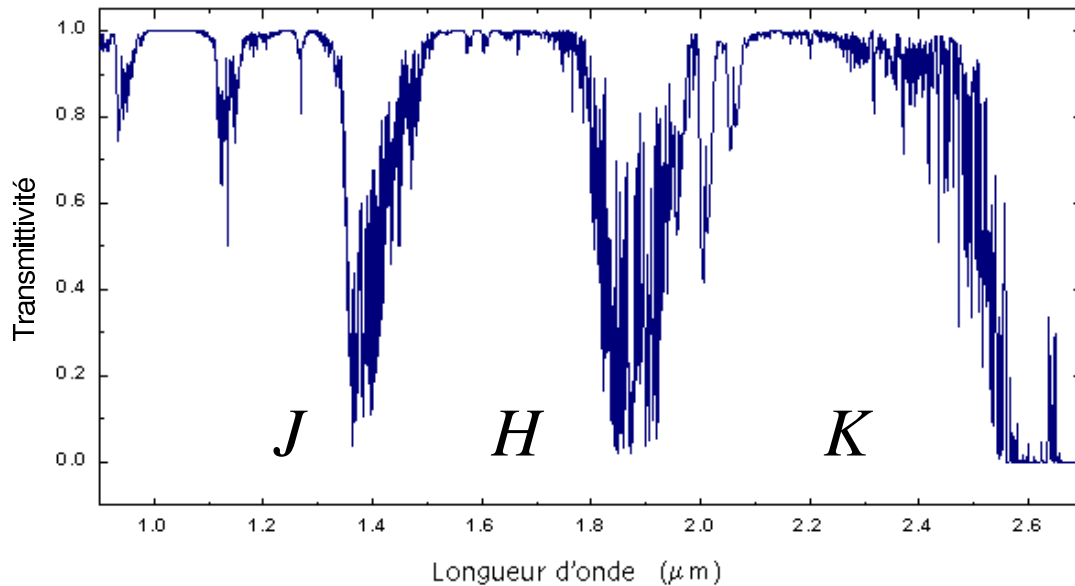


Fig. 1.3: Transmittivité de l'atmosphère dans l'infrarouge proche pour des conditions climatiques moyennes à Mauna Kea (Hawaï). L'atmosphère est transparente pour une transmittivité de 1, et opaque pour une transmittivité de 0. On a reporté les noms des filtres correspondant aux fenêtres d'observation depuis le sol.

diat du détecteur : on tente de le réduire le plus possible en utilisant des techniques de refroidissement (cryostats, enceintes de thermalisation).

Petit historique de l'astronomie infrarouge

Dès sa découverte, le rayonnement infrarouge est lié à l'astronomie, puisque c'est Sir William Herschel qui le met le premier en évidence en étudiant le rayonnement solaire, au début des années 1800.

Les astronomes américains Coblentz, Abbot, Pettit et Nicholson sont, dans les années 1920, les premiers à conduire des observations systématiques d'étoiles dans l'infrarouge. Les détecteurs utilisés sont des bolomètres, des radiomètres et des thermopiles, qui mesurent les différences d'échauffement avec ou sans la présence de filtres comme des lames de verre, de mica ou des cellules à vapeur d'eau.

Kuiper est le premier à utiliser des photomètres (PbS) à la fin des années 40, mais il faudra attendre les années 1960 pour disposer de calibrations assez précises, et de photoconducteurs équipés de systèmes de refroidissement.

La première véritable cartographie du ciel dans l'infrarouge proche fut entreprise à la fin des années 1960 par Neugebauer & Leighton (1969) avec le *Two Micron Sky Survey* (TMSS). Malgré une magnitude limite de 3.1 dans la bande K (2.2 μm), plus de 5000 objets de l'hémisphère nord, dont beaucoup d'objets rouges inconnus auparavant, furent détectés avec un télescope de 62" au Mont Wilson.

Le catalogue de l'AFGL (*Air Force Geophysics Laboratory*, Price & Walker (1976)) est le résultat d'un relevé du ciel à 4.2, 11.0, 19.8 et 27.4 μm (mais il contient seulement 2363 sources) .

Les quelques relevés plus profonds entrepris pour l'étude du disque galactique (Eaton et al. (1984), Kawara et al. (1982)) ne couvrent que de très petites zones (typiquement quelques centaines de minutes carrées).

Jusqu'à la fin des années 80, les progrès des détecteurs ne seront pas suffisants pour motiver un nouveau relevé sur une grande partie du ciel dans l'infrarouge proche (voir Price (1988) pour une revue des relevés infrarouges jusque 1988).

A plus grande longueur d'onde

L'émission atmosphérique interdisant les observations depuis le sol dans l'infrarouge lointain, il a fallu attendre l'ère des missions spatiales⁴ pour faire des observations à ces grandes longueurs d'onde. Les principales missions effectuées dans ce domaine sont :

- IRAS (*Infrared Astronomical Satellite*, lancé en 1983), un relevé du ciel à 12, 25, 60 et 100 μm ;
- COBE (*Cosmic Background Explorer*, lancé fin 1989), pour étudier l'émission diffuse infrarouge et le rayonnement micro-ondes de l'univers primordial ;
- ISO (*Infrared Space Observatory*, lancé en 1995), le successeur d'IRAS, observant de 2.5 à 240 microns avec un gain d'un facteur 1000 en sensibilité et 100 en résolution angulaire à 12 μm ;
- MSX (*Midcourse Space Experiment*, lancé en 1996), satellite militaire comprenant un télescope observant dans 6 bandes photométriques infrarouges entre 6.8 et 25.1 μm .

Nouveaux détecteurs, nouveaux relevés

La technologie des capteurs infrarouges a été stimulée par les besoins des militaires⁵, et des cibles de grandes dimensions (256×256 pixels, voire plus), sensibles dans l'infrarouge proche, avec de bons rendements quantiques (et une efficacité supérieure de plus de 4 ordres de grandeur aux meilleurs photodétecteurs), sont devenus accessibles à des prix raisonnables. Ces progrès techniques rendaient envisageables des relevés du ciel à très grande échelle dans l'infrarouge proche, avec un gain considérable par rapport au TMSS.

Le *Two-Micron Galactic Survey* (TMGS, Garzón et al. (1993)), utilisant une caméra infrarouge montée sur le télescope Carlos Sanchez (1.5 m) aux Canaries, a ainsi permis de couvrir le plan galactique (entre -15° et $+15^\circ$ de latitude pour des

⁴On peut toutefois citer des observations à bord d'avions (10 km d'altitude) ou de ballons (jusque 40 km), permettant de réduire d'un facteur 100 à 1000 cette émission. Voir Price (1988) pour une liste de ces travaux pionniers.

⁵En raison de l'importance stratégique de la détection du rayonnement thermique, qui, pour des objets entre 250 et 330 K, se situe entre 8 et 30 μm .

longitudes de -5° à 30° , et entre -5° et $+5^\circ$ de latitude entre 30° et 180° de longitude), avec une magnitude limite $m_K = 10$.

Deux projets de grande ampleur ont aussi émergé au début des années 1990 : DENIS (*DEep Near Infrared Survey of the Southern Sky*), et 2MASS (*Two Micron All Sky Survey*), avec comme objectifs respectifs la cartographie de l'hémisphère sud dans les bandes I , J , et K_s (pour DENIS), et de tout le ciel en J , H , et K_s (pour 2MASS).

Les observations de ces deux relevés étaient en cours pendant ma thèse. La suite de ce chapitre est consacrée à une description rapide de chacun des projets.

1.2 Le projet DENIS

Le projet DENIS (Epchtein, 1997) est piloté par un consortium regroupant de nombreux laboratoires européens et un institut brésilien. L'objectif est une cartographie complète de l'hémisphère austral dans une bande optique et deux bandes du proche infrarouge.

1.2.1 Caractéristiques de l'instrument

L'instrument DENIS (Copet et al., 2001) est installé au foyer Cassegrain du télescope de 1 m de l'ESO, à La Silla (Chili) (voir Fig. 1.4).

Il s'agit d'une caméra à trois voies, construite à l'Observatoire de Paris avec la contribution de divers instituts européens, permettant d'observer une même région du ciel simultanément dans trois bandes photométriques :

- I : filtre Gunn- i , centré sur $0.82 \mu\text{m}$;
- J : filtre J -Johnson, centré sur $1.25 \mu\text{m}$;
- K_s : filtre K_{short} , centré sur $2.15 \mu\text{m}$.

La lumière incidente est séparée en trois faisceaux par deux lames dichroïques (Fig. 1.5).

La figure 1.6 donne les courbes de transmission des filtres et des lames dichroïques. Le filtre I , intermédiaire entre les domaines visible et infrarouge, distingue DENIS des relevés infrarouges classiques. Il permet en effet une meilleure jonction avec les catalogues dans le domaine optique, ce qui facilite les identifications croisées. La bande I facilite aussi la recherche de naines brunes (Delfosse et al., 1997), car son association aux deux bandes infrarouges permet de les distinguer des autres objets.

Le filtre K_s , plus étroit et centré à plus courte longueur d'onde que le K -Johnson standard, ce qui permet de mieux éviter une bande de l'eau, et de réduire ainsi la brillance du fond de ciel. L'ensemble de l'instrument est placé dans une enveloppe de thermalisation pour réduire le bruit de fond thermique.

Pour chaque voie optique, un réducteur de focale permet d'avoir un champ de $12' \times 12'$ sur le détecteur (voir table 1.2 pour les caractéristiques techniques des trois caméras). En raison des tailles différentes des cibles, l'échelle des images est de $1''$ par

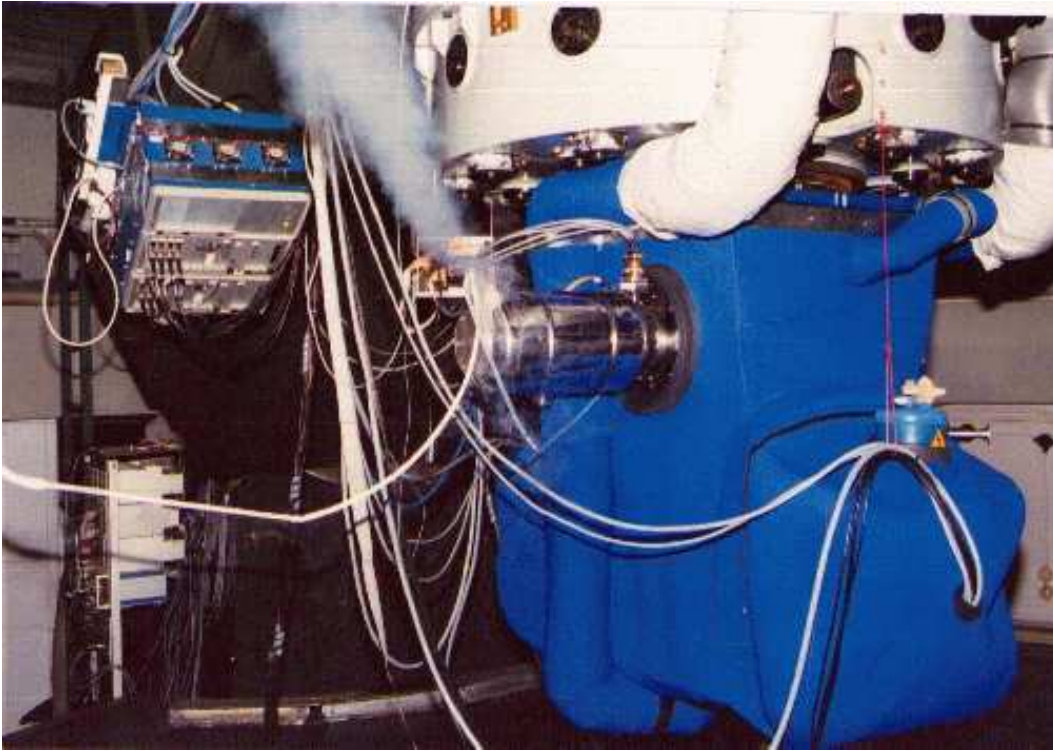


Fig. 1.4: L'instrument DENIS, au foyer du télescope de 1 m de La Silla, dans son enveloppe de thermalisation.

pixel en I , mais de seulement $3''$ par pixel en J et K_s . Afin d'améliorer l'échantillonnage dans les voies J et K_s , un système de micro-balayage (*micro-scanning*) a été mis au point : un miroir vibrant décompose chaque pose en 9 sous-images décalées d' $1/3$ de pixel en ascension droite et de $7/3$ de pixel en déclinaison. On obtient ainsi des images ré-échantillonnées à $1''$ par pixel.

Bande	I	J	K_s
Longueur d'onde centrale (μm)	0.82	1.25	2.15
Cible	Tektronix	NICMOS-3	NICMOS-3
Taille (pixels)	1024×1024	256×256	256×256
Temps de pose	8.998 s	8.809 s	8.809 s
Magnitude limite (3σ)	18.5	16.5	14.0
Magnitude de saturation	10.0	8.0	6.0

Tab. 1.2: Caractéristiques des 3 caméras DENIS. Les cibles sont des capteurs disponibles dans le commerce.

La calibration complète de l'instrument DENIS, tenant compte de tous les éléments (atmosphère, miroirs, dichroïques, objectifs, détecteurs) est décrite par Fouqué et al. (2000).

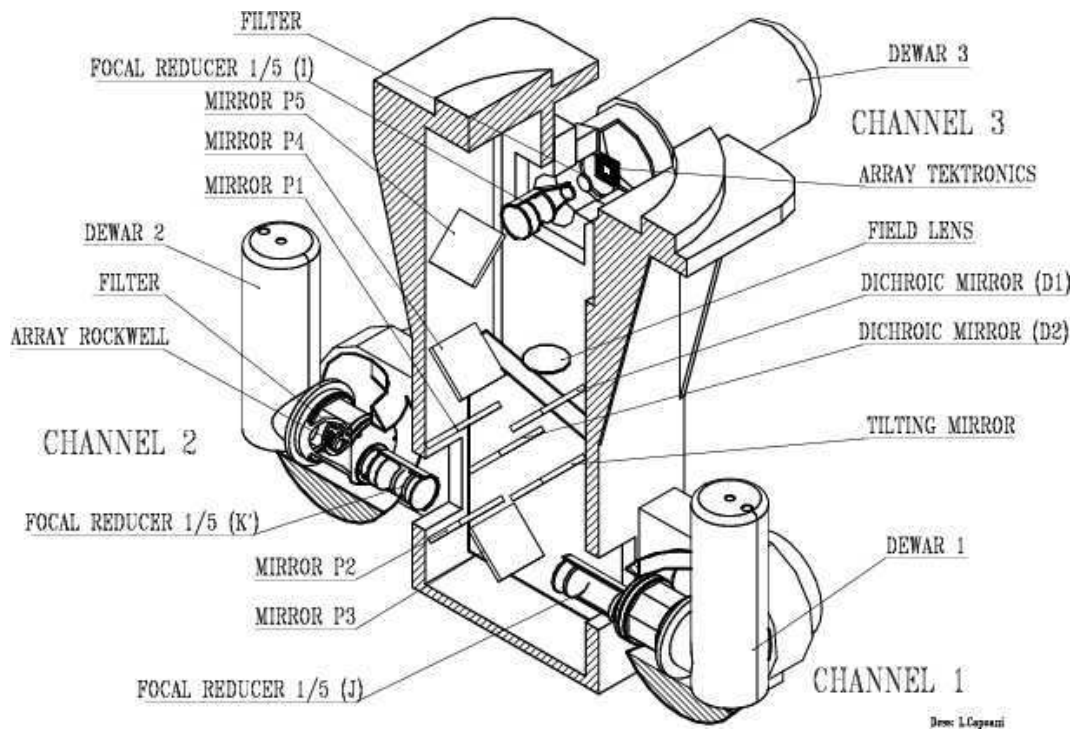


Fig. 1.5: Montage optique de l'instrument DENIS (schéma de L. Capozzi).

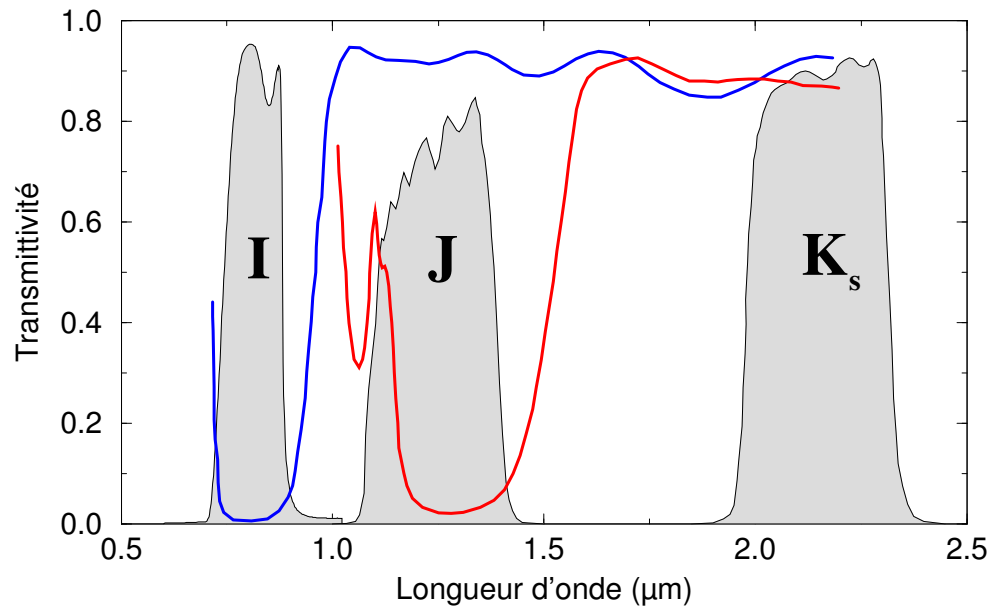


Fig. 1.6: Courbes de transmission des filtres DENIS (grisé), et transmittivité des deux miroirs dichroïques (traits pleins).

1.2.2 Les observations

Les observations ont débuté en décembre 1995 à La Silla. L'hémisphère austral a été divisé en trois zones de déclinaison de 30 degrés :

- zone équatoriale : de $+2^\circ$ à -28° ;
- zone intermédiaire : de $-27^\circ 50'$ à $-57^\circ 50'$;
- zone polaire : de $-57^\circ 40'$ à $-87^\circ 40'$.

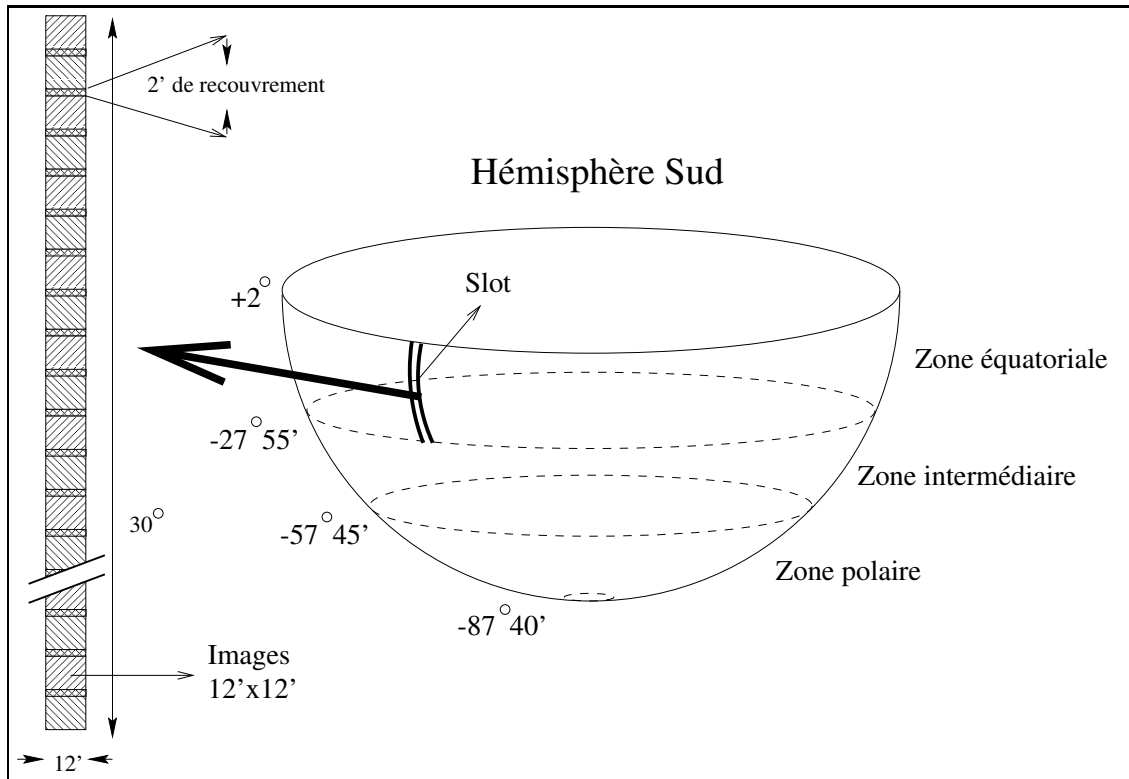


Fig. 1.7: Stratégie observationnelle pour DENIS. On observe à ascension droite constante des **strips** de 30° de long en déclinaison. Chaque **strip**, composé de 180 images, correspond à un emplacement numéroté (le **slot**) dans l'une des trois zones de déclinaison.

A l'intérieur de chaque zone, on a défini des *slots* (Fig. 1.7), correspondant à des régions de 30° de long en déclinaison et $12'$ de large en ascension droite. On a entre deux *slots* adjacents un chevauchement minimum de $2'$. L'hémisphère complet contient 5112 *slots* numérotés de 0 à 5111.

Un *strip* DENIS correspond à une séquence d'images de $12' \times 12'$ couvrant un *slot*. Deux images successives sont séparées de $10'$ en déclinaison pour assurer $2'$ de recouvrement. Un *strip* est donc composé de 180 images.

Notons qu'un même *slot* peut être observé plusieurs fois (en raison de mauvaises conditions météorologiques lors de la première observation, par exemple). Par conséquent, deux *strips* différents peuvent correspondre à un même numéro de *slot* (les *strips* sont, eux, numérotés par ordre chronologique d'observation).

Chaque nuit, une procédure d'optimisation des observations détermine les slots à observer en fonction de divers paramètres : minimisation de la masse d'air, éloignement du slot par rapport à la lune, priorités scientifiques éventuelles, observations de strips adjacents pour obtenir une couverture complète de certaines zones, etc.

Avant et après chaque strip (séquence de 180 images), on effectue une séquence astrométrique (observation d'une zone contenant des étoiles du GSC pour assurer un bon pointage) et une séquence photométrique (8 observations d'une même étoile standard, proche du début du strip). L'ensemble de la séquence d'observation dure environ 1h30, ce qui permet d'observer entre 6 et 8 strips par nuit.

En marge du relevé complet de l'hémisphère, constitué uniquement de strips, un autre mode d'utilisation de l'instrument DENIS a été prévu : les *batches*. Un *batch* est une séquence de 100 images seulement, mais couvrant l'équivalent de 4 strips de large (et donc beaucoup plus court qu'un strip en déclinaison). On peut utiliser des *batches*, par exemple, pour couvrir rapidement une zone du ciel présentant un grand intérêt, mais leur usage reste limité car ils ne contribuent pas à l'avancement de la couverture du relevé.

A la fin de l'année 2000, la couverture du ciel sud devrait être de 92%, et la fin des observations est prévue pour le courant de l'année 2001.

1.2.3 La chaîne de traitement des données

La Silla

Au fur et à mesure de l'acquisition des données, les neuf sous-images infrarouges (obtenues par la procédure de micro-balayage en J et K_s) sont combinées en une image de 768×768 pixels, le fond de ciel est soustrait, et les images sont compressées puis stockées sur DAT.

Une correction du champ plat (*flat field*) est aussi effectuée en temps réel afin d'afficher les images sur les écrans de contrôle.

PDAC

Les DAT sont expédiées à l'IAP⁶, où se trouve le *Paris Data Analysis Center* (PDAC), qui est responsable de l'archivage et du traitement des images (Borsenberger, 1997).

Les images sont corrigées de l'émission du ciel, du biais (émission de l'instrument lui-même), et une correction du champ plat est appliquée. Le PDAC établit une carte des mauvais pixels, et étudie les variations de la PSF (*Point Spread Function*, ou réponse impulsionnelle) pour mettre en évidence des dégradations éventuelles de la qualité des images.

⁶<http://pcnt1.iap.fr/Denis/>

LDAC

Les images traitées au PDAC sont envoyées sur support DLT aux Pays-Bas, au Leiden Data Analysis Center (LDAC). Le LDAC effectue l'extraction des sources ponctuelles, ainsi que les calibrations astrométrique et photométrique. L'extraction de sources se fait à l'aide du logiciel SExtractor (Bertin & Arnouts, 1996).

La calibration astrométrique se fait par identification croisée avec un catalogue de référence. Le catalogue USNO A1.0 (Monet et al., 1997) était utilisé jusqu'à la parution de la deuxième version (USNO A2.0, plus de $520 \cdot 10^6$ sources) qui l'a remplacé.

Le LDAC est chargé du traitement des calibrations photométriques (voir p. 17), qui permettent de fixer la photométrie absolue des sources extraites.

Le contenu des catalogues de sources ponctuelles produits au LDAC est décrit plus en détails dans la section 2.2.

Les bases de données

C'est ensuite le PDAC qui est responsable du développement des bases de données de sources ponctuelles (SODA) et des sources étendues (LODA), à partir des données traitées revenant du LDAC.

En mars 1997, le PDAC a commencé à effectuer en parallèle le calcul de la solution astrométrique pour les images, et depuis 1999, une nouvelle chaîne de traitement complète fonctionne au PDAC :

- l'extraction de sources est faite à l'aide d'un logiciel optimisé pour les champs encombrés (avec la méthode de reconstruction de la PSF décrite par Alard & Lupton (1998));
- l'astrométrie est calibrée sur celle du catalogue USNO A2.0 ;
- les séquences de calibration photométriques sont traitées et stockées dans une base de données ;
- la photométrie des sources est calculée en utilisant un ajustement polynomial de la PSF (*PSF fitting*) dans l'image, et le long du strip (Alard, 2000).

Le PDAC gère la base de données de sources ponctuelles ainsi obtenue (qui contient actuellement le produit de la réduction des strips observés après février 1999).

Le détail du contenu de cette base est présenté dans la section 2.2.

La description du fonctionnement des chaînes de traitement des données du LDAC et du PDAC sera à nouveau abordée au chapitre 3, quand nous discuterons la validation des données DENIS.

Accès public aux données

C'est le *Centre de Données Astronomiques de Strasbourg* (CDS) qui a pour mission de mettre à disposition de l'ensemble de la communauté scientifique les données

traitées et calibrées⁷.

Un premier échantillon de données traitées au LDAC a été rendu public en décembre 1998 via le CDS (Epchtein et al., 1999).

Les chapitres 2, 3 et 4 détaillent le travail que j'ai effectué pendant ma thèse autour de l'accès aux données DENIS.

Performances

Les précisions obtenues sur les données sont de l'ordre de :

- 0.5" pour l'incertitude astrométrique ;
- 0.05 mag pour les étoiles brillantes, et 0.2 mag pour les plus faibles.

1.2.4 Objectifs scientifiques et premiers résultats

Les applications potentielles d'un grand relevé infrarouge couvrent de nombreux domaines de recherche. DENIS est bien sûr particulièrement adapté à l'étude des objets émettant dans le proche infrarouge (étoiles évoluées, objets de faible masse, étoiles dans des nuages moléculaires, etc.), mais la faible extinction interstellaire à ces longueurs d'onde⁸ peut aussi servir de nombreux autres projets.

Objets faibles du voisinage solaire

L'étendue du domaine spectral couvert par DENIS rend ce relevé particulièrement bien adapté à la détection des naines rouges des types les plus tardifs (au-delà de M7). L'analyse de 220 degrés carrés de données à haute latitude galactique a permis d'identifier 3 naines brunes (Delfosse et al., 1997), et 25 naines plus tardives que M7, ce qui multipliait à l'époque par deux la liste de ces objets très froids (Delfosse et al., 1998).

Le relevé complet devrait contenir suffisamment de naines froides pour déterminer de façon robuste la fonction de luminosité locale du disque pour ces objets, et pour accroître considérablement notre connaissance sur la physique des naines brunes (avec un suivi spectroscopique des objets détectés).

L'extinction, le milieu interstellaire

En utilisant des comptages d'étoiles dans le relevé DENIS, on peut construire des cartes d'extinction détaillées pour les nuages moléculaires proches. Cambrésy et al. (1997) ont ainsi obtenu une carte d'extinction du nuage Caméléon I avec une résolution sans précédent. On peut sonder des régions obscurcies de 30 magnitudes dans le visible.

En combinant ces cartes avec les données IRAS à 60 et 100 μm , et avec des cartes de l'émission millimétrique du CO, on pourra connaître en détail la température et la distribution du gaz et des poussières dans ces régions.

⁷<http://cdsweb.u-strasbg.fr/denis-public/>

⁸ $A_I/A_V = 0.60$, $A_J/A_V = 0.29$, et $A_{K_s}/A_V = 0.11$ (Bessell & Brett, 1988).

Étoiles jeunes dans les nuages moléculaires

On peut détecter avec DENIS de nombreuses étoiles en cours de formation (*Young Stellar Objects* – YSOs, ou proto-étoiles) dans les nuages moléculaires. Copet (1996) a détecté dans le nuage moléculaire d'Orion de nombreux candidats T Tauri.

Cambrésy et al. (1998) ont détecté des YSOs qui forment probablement une nouvelle population d'étoiles T Tauri de faible luminosité. Ces objets jeunes émettent parfois fortement en rayons X, et les données DENIS pourront donc être comparées à celles des satellites ROSAT ou XMM.

La structure galactique

Il est quasiment impossible dans le domaine optique de sonder le disque à plus de 2 kpc. Mais en raison de la faible absorption par le milieu interstellaire dans la bande K_s , les géantes rouges et les supergéantes peuvent être détectées avec DENIS jusque dans le centre galactique. Les supergéantes indiquent les régions de formation d'étoiles, et l'étude des géantes permet de tracer la distribution et l'évolution des populations plus vieilles.

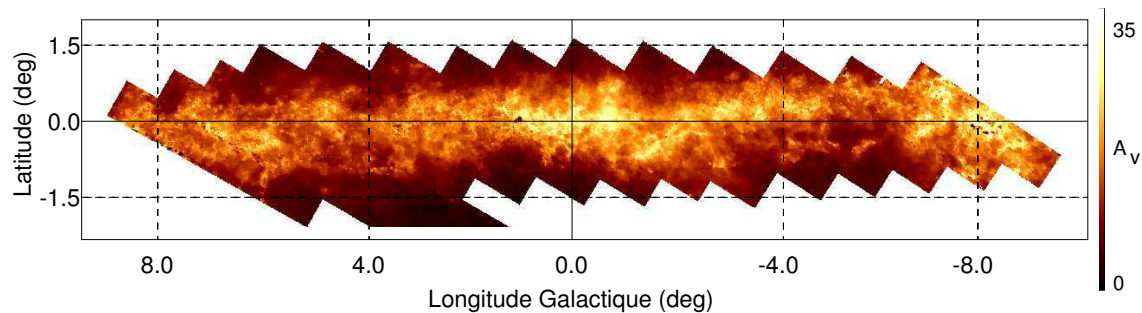


Fig. 1.8: Carte d'extinction pour les régions centrales du bulbe galactique (Schultheis et al., 1999). On atteint plus de 30 magnitudes visibles d'extinction dans les zones les plus obscurcies.

Schultheis et al. (1999) ont réalisé une carte d'extinction des régions centrales du bulbe galactique (Fig. 1.8). L'identification croisée des données DENIS et ISOGAL (observations à 7 et 15 μm) devrait permettre de modéliser le bulbe et la barre galactique. Des candidats d'étoiles variables ont déjà été sélectionnés avec ces données (Schultheis et al., 2000).

Un premier travail de thèse avait été réalisé par Ruphy (1996) sur la structure galactique, avec les toutes premières données DENIS. Nous verrons au chapitre 7 les résultats que j'ai obtenus au cours de cette thèse dans ce domaine.

Les Nuages de Magellan

Un catalogue des deux Nuages a été publié (Cioni et al., 2000b), à partir duquel, une étude de la morphologie des Nuages a déjà été faite par Cioni et al. (2000a).

Cioni et al. (2000c) ont trouvé des modules de distance $m - M$ de 18.55 pour le Grand Nuage de Magellan (LMC) et de 18.99 pour le Petit Nuage (SMC). Les étoiles AGB ont des magnitudes bolométriques inférieures à -3.6 (plus de 1700 L_{\odot}). Les corrections bolométriques de ces objets dépendent de leur couleur (Groenewegen, 1997). Mais avec une limite de 14.0 en K_s , DENIS doit détecter toutes les AGB des Nuages ayant une correction bolométrique $m_{bol} - K_s$ supérieure à 1.4, ce qui représente la quasi-totalité de ces objets (seules les quelques AGB très rouges, obscurcies avec un fort taux de perte de masse ne seront pas visibles).

Galaxies et cosmologie

Le relevé DENIS permet de construire un grand catalogue de galaxies en s'affranchissant des problèmes associés à l'extinction interstellaire, qui empêche en particulier de déterminer précisément une magnitude limite. De plus, les couleurs infrarouges sont moins sensibles aux épisodes récents de formation stellaire (Mamon et al., 1997), et permettent donc une meilleure estimation de la masse de la composante stellaire.

Vauglin et al. (1999) ont publié un premier catalogue des 20 260 galaxies les plus brillantes détectées dans la bande I dans un quart du ciel sud. Plus de 70% de ces galaxies sont nouvelles.

Les données DENIS permettent surtout de détecter de nouvelles galaxies dans la « zone d'exclusion » (*zone of avoidance*), à basse latitude galactique. On s'attend à recenser 900 000 galaxies dans la bande I , 500 000 dans la bande J , et 50 000 dans la bande K_s .

Un programme (KLUN+) est en cours pour observer l'émission HI (hydrogène neutre) de 8000 galaxies brillantes détectées par DENIS, au radiotélescope de Nançay.

1.3 Le projet 2MASS

2MASS (Skrutskie et al., 1997) est le produit d'une collaboration entre l'Université du Massachusetts (UMASS) et l'IPAC⁹ (*Infrared Processing and Analysis Center*, (JPL/Caltech)) pour observer le ciel complet dans trois bandes photométriques dans l'infrarouge proche.

1.3.1 Fiche technique

Les observations sont faites à l'aide de deux télescopes de 1.3 m situés respectivement au Mont Hopkins (Arizona) pour l'hémisphère nord, et à Cerro Tololo (Chili) pour le sud. Ces télescopes sont équipés de deux caméras identiques à trois canaux (Fig. 1.9), permettant des observations simultanées dans les bandes J (1.25 μm), H (1.65 μm), et K_s (2.17 μm).

⁹<http://www.ipac.caltech.edu/2mass/>

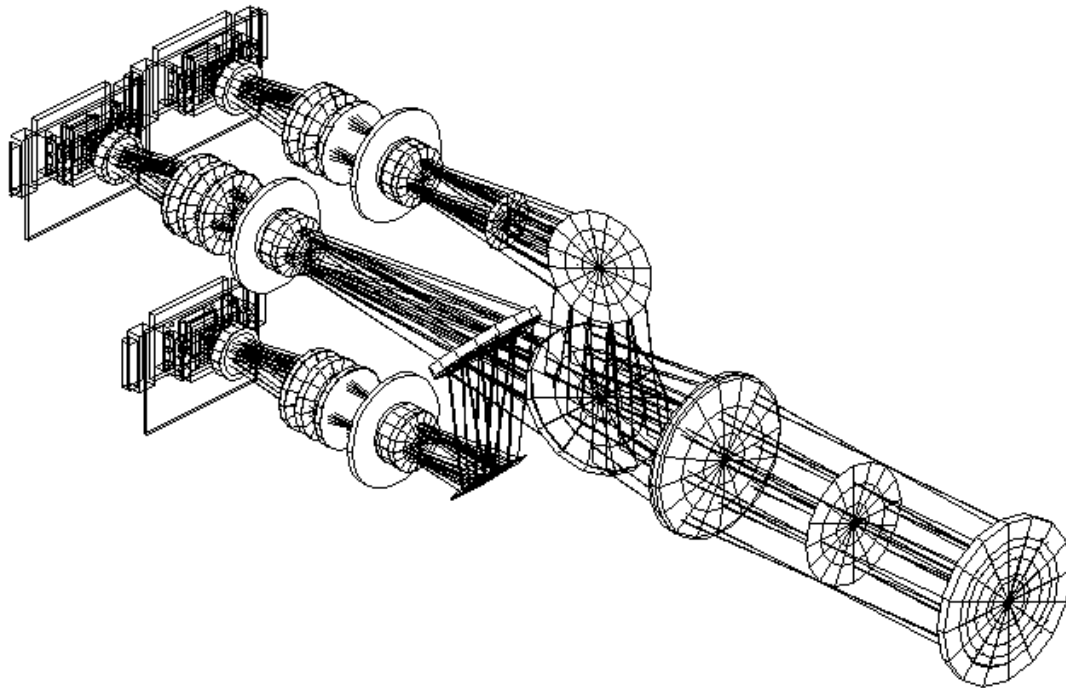


Fig. 1.9: Chemin optique pour la caméra à 3 canaux de 2MASS.

Chaque canal est équipé d'un détecteur 256×256 , permettant une résolution de $2''$ par pixel pour un champ de $8.5' \times 8.5'$ (voir Tab. 1.3).

1.3.2 Stratégie d'observation

Afin d'assurer une couverture complète, le ciel a été divisé en 59 650 **tuiles** de $8.5'$ de large en ascension droite, et 6° en déclinaison. Une tuile est composée de 272 images de $8.5' \times 8.5'$, avec un recouvrement de $5/6$ d'une image à l'autre (Fig. 1.10).

On a un tramage (*dithering*) induit par les recouvrements importants entre deux images successives : l'image finale est une combinaison de 6 images individuelles (Fig. 1.11), chaque point du ciel étant observé 6 fois, en différents endroits par rapport au centre d'un pixel (le pas en déclinaison induit un déplacement d'une fraction de pixel, et une légère rotation de la caméra provoque des décalages en ascension droite). Cette procédure de tramage présente plusieurs avantages :

- la PSF est « aiguisée » : à une dimension, on peut voir cela comme le passage d'une PSF en créneau de $2''$ de large, en une PSF de même largeur à mi hauteur, mais de forme triangulaire. La qualité d'image est améliorée, ce qui permet une

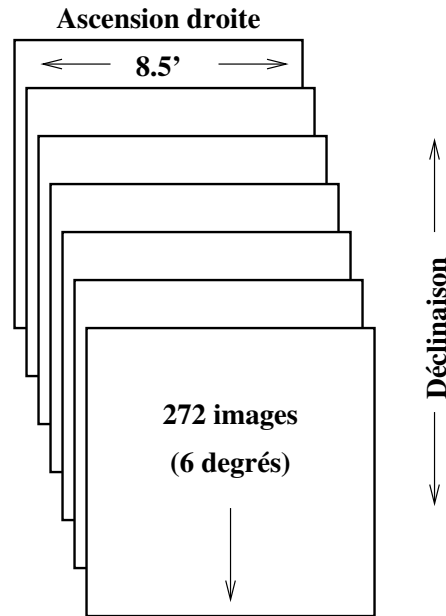


Fig. 1.10: Stratégie observationnelle pour 2MASS. Chaque tuile de $8.5' \times 6^\circ$ est composée de 272 images de $8.5' \times 8.5'$, chaque point d'une tuile étant observé six fois.

meilleure séparation des sources très serrées, et une meilleure discrimination étoiles/galaxies ;

- l'uniformité de la photométrie, de l'astrométrie et des images est améliorée, car les biais dus à l'échantillonnage à l'échelle d'un pixel sont supprimés ;
- l'astrométrie et la photométrie sont améliorées par le lissage des variations de la réponse sur un pixel.

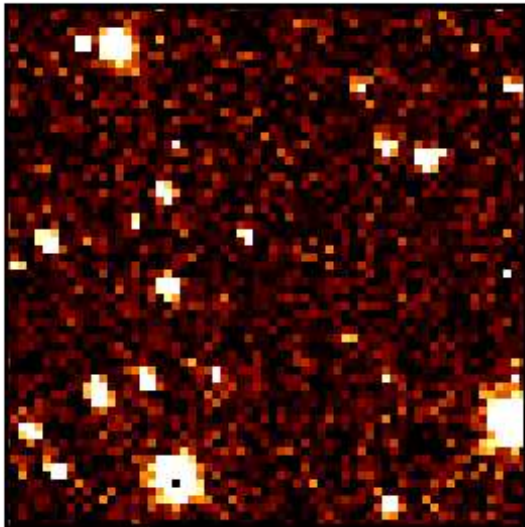
Au cours de l'observation d'une tuile, le télescope balaye le ciel en déclinaison à vitesse constante, pendant que le miroir secondaire tourne en sens inverse afin de figer l'image d'une position fixe sur le ciel. Entre deux images, le secondaire bascule rapidement (50 ms) pour se positionner sur l'image suivante, puis reprend

Cible infrarouge	256 × 256 NICMOS3 (HgCdTe)
Taille d'un pixel	2.0"
Bandes photométriques	J, H, K_s
Télescopes	1.3 m, Cassegrain, monture équatoriale
Temps d'intégration	$6 \times 1.3 \text{ s/image} = 7.8 \text{ s}$
Sensibilité (3σ)	17.1, 16.4, 15.6 en $J, H,$ et K_s
Précision photométrique	5% pour les sources brillantes
Uniformité de la photométrie	4% sur tout le ciel
Précision sur l'astrométrie	0.5"

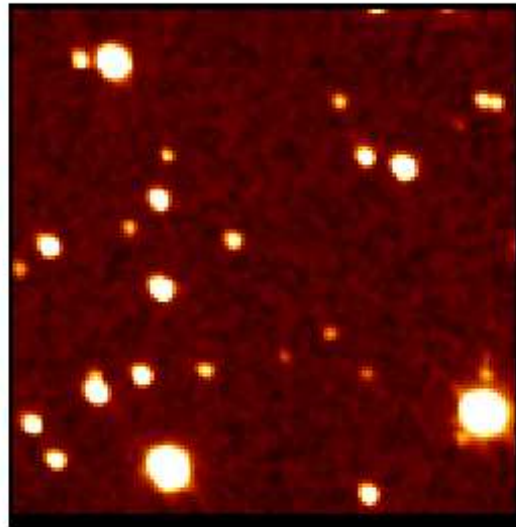
Tab. 1.3: Caractéristiques techniques de 2MASS, et performances attendues.

son mouvement de compensation du balayage en déclinaison. On a deux lectures des cibles pour chaque position. La première lecture (READ1 ou R1), effectuée juste après le repositionnement du miroir, correspond à un temps d'intégration très court (0.05 s). La deuxième lecture (READ2 ou R2) s'effectue après une pose de 1.3 s. L'intérêt de la première lecture est de pouvoir déterminer une magnitude pour les étoiles très brillantes qui sont saturées en R2.

En tenant compte des temps morts (*reset*, repositionnement du secondaire), il faut environ 7 minutes pour observer une tuile complète.



Une seule exposition de 1.3s
2.0'' / pixel



Combinaison de 6 expositions de 1.3s
1.0'' / pixel

Fig. 1.11: Exemple de combinaison de 6 observations individuelles en une seule image. On peut voir le gain important en résolution obtenu entre une observation simple (à gauche) et la combinaison (à droite), ainsi que le gain pour les mauvais pixels (dans l'étoile brillante en bas, par exemple).

Il y a entre les tuiles un recouvrement de 0.95' en ascension droite et de 8.5' en déclinaison. L'hémisphère nord est restreint aux déclinaisons supérieures à +12°, car le rendement attendu (en nuits de bonne qualité) est supérieur pour le télescope installé au Chili.

Des *tuiles de calibration* sont observées à des intervalles d'une heure tout au long de chaque nuit. Une tuile de calibration ne fait qu'un degré de long, mais elle est ré-observée 6 fois au cours de la même séance de calibration. Comme chaque point est toujours observé dans 6 images, chaque étoile est donc mesurée 36 fois dans une même séance de calibration.

Les observations ont débuté au printemps 1997 dans l'hémisphère nord, et un an plus tard dans le sud. L'ensemble du ciel est couvert depuis fin octobre 2000.

1.3.3 Traitement des données

C'est l'IPAC qui est en charge du traitement des quelque 20 TB de données brutes issues des observations.

Une chaîne de traitement complète a été mise au point, le *2MASS Production Processing System* (2MAPPS¹⁰, Cutri (1997)). Cette chaîne effectue l'ensemble des opérations, depuis la réception des données, jusqu'à l'intégration des sources et des images dans une base de données de travail (WSDB, *Working Survey Data Base*). La figure 1.12 donne un aperçu de l'organigramme du 2MAPPS.

Les données venant des télescopes sont acheminées à l'IPAC sur bandes DLT. Les principales étapes du traitement sont :

- un premier traitement des images, pour calculer les décalages entre les différentes images, dans chaque bande :
 - correction du champ plat et du biais ;
 - estimation du *seeing* ;
 - calcul du fond de ciel, du bruit moyen ;
 - extraction sommaire des sources ponctuelles.
- extraction des sources ponctuelles dans chaque bande, et construction des images combinées :
 - combinaison des sous-images (avec les décalages calculés à l'étape précédente) ;
 - extraction des magnitudes R2-R1 sur les images combinées ;
 - extraction des magnitudes R1 sur les images individuelles ;
 - fusion des listes de sources R2-R1 et R1 ;
- jonction des sources ponctuelles entre les 3 bandes ;
- reconstruction des positions des sources ;
- extraction des galaxies et objets à faible brillance de surface sur les images combinées ;
- traitement des séquences de calibration, détermination des masses d'air et point zéro des observations ;
- stockage des images et des listes de sources calibrées dans la WSDB.

L'IPAC est chargé de délivrer à la communauté astronomique :

- un atlas d'images ;
- un catalogue de sources ponctuelles ;
- un catalogue de sources étendues ;
- les documentations associées.

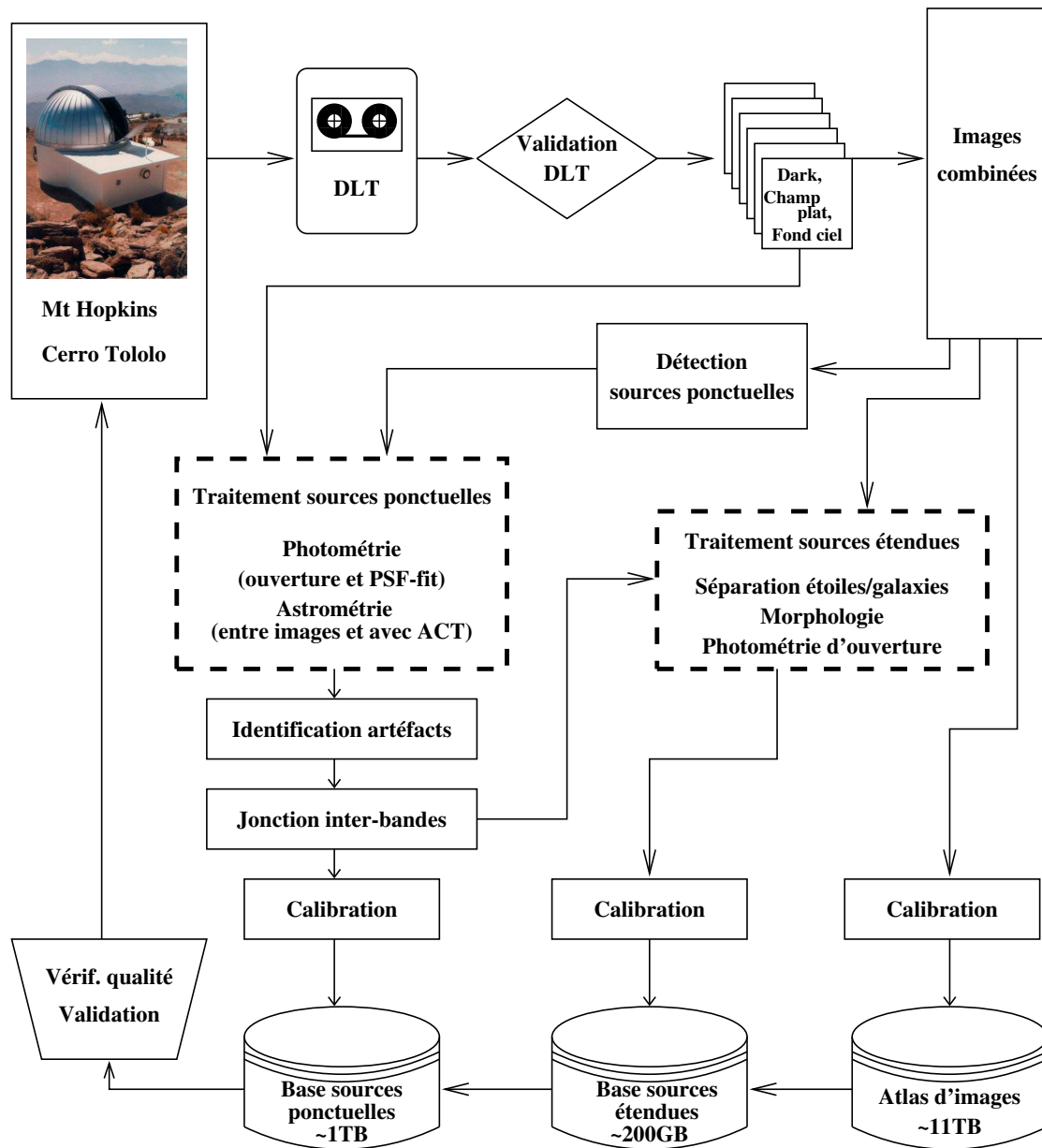


Fig. 1.12: Organigramme du 2MAPPS. Les DLT contenant les observations brutes sont vérifiées. Les images sont nettoyées (correction de champ plat, dark, fond du ciel). Une première extraction des sources ponctuelles permet de déterminer précisément les décalages entre les images et de construire les images combinées, pour chaque bande. Les extractions de sources ponctuelles sont faites à partir des images combinées, et les différentes bandes photométriques sont jointes. Le traitement des sources étendues utilise à la fois les images (surtout pour les objets à faible brillance de surface) et les résultats de l'extraction de sources ponctuelles. Le traitement des séquences de calibration permet d'effectuer les différentes calibrations avant stockage dans la WSDB. Après validation des données, l'information est transmise aux sites d'observation (pour signaler, par exemple, les zones à ré-observer pour cause de mauvaise qualité).

L'accès public aux données est assuré par l'IRSA (*InfraRed Science Archive*)¹¹, et une copie est disponible au CDS.

Une copie des données publiques est également accessible au CDS¹², comme on le verra au chapitre 4.

Il est également prévu que l'atlas du ciel ALADIN (Bonnarel et al., 2000) puisse proposer les images 2MASS.

1.3.4 Objectifs scientifiques et résultats

Les objectifs scientifiques de 2MASS sont assez similaires à ceux du projet DENIS. Les principales différences entre les deux relevés sont la présence de la bande *I* dans DENIS, contre la bande *H* pour 2MASS, et le fait que 2MASS soit un peu plus profond dans la bande *K_s* que DENIS.

Les observations dans l'infrarouge proche, moins sensible à la présence des poussières du milieu interstellaire, permettent d'étudier les structures à grande échelle de la Galaxie. Le catalogue final de sources étendues sera le premier recensement des galaxies plus brillantes que la magnitude 13.5 en *K_s*, complet pour tout le ciel, y compris les régions à basse latitude galactique (« zone d'exclusion »). Le catalogue fournira également une base statistique pour la recherche d'objets rares (étoiles de très faible masse, naines brunes, noyaux actifs de galaxies (AGN) très rouges).

Étoiles de très faible masse

Kirkpatrick et al. (1999) ont trouvé de nouveaux candidats de naines L, cette classe d'objets de faible masse, plus froids que les naines M, définie par Delfosse à partir des données DENIS. Cette définition repose sur la découverte dans les données 2MASS de 20 nouveaux objets de type plus tardif que M9.5 (on en connaissait seulement 9 auparavant).

Quelques objets très froids, des naines de type T (présentant des bandes d'absorption du méthane dans leur spectre), ont également été détectés (Burgasser et al., 1999). La température effective de ces objets dont on connaît très peu de représentants est de l'ordre de 1300 K.

La découverte de ces objets extrêmes, mais aussi de nombreuses étoiles de type tardif, permet d'améliorer notre connaissance de la fonction de luminosité des objets de faible masse (Liebert et al., 2000).

Morphologie des galaxies dans l'infrarouge proche

De nombreuses galaxies nouvelles ont été détectées par 2MASS, en particulier dans les régions obscurcies par le disque de notre propre Galaxie, à basse latitude (Jarrett et al., 2000). A partir d'un échantillon de galaxies, un atlas morphologique

¹⁰http://www.ipac.caltech.edu/2mass/data_processing/index.html

¹¹<http://irsa.ipac.caltech.edu/>

¹²<http://vizier.u-strasbg.fr/cgi-bin/VizieR?-source=B/2mass>

des galaxies dans l'infrarouge proche a été construit, pour présenter la forme des galaxies classées dans la séquence de Hubble, à des longueurs d'onde de 1–2 μ m (Jarrett, 2000).

AGN rouges

La vision des noyaux actifs de galaxies dans le visibles est fortement sensible à la quantité de poussière présente. Dans l'infrarouge, de nombreux AGN invisibles aux longueurs d'onde optiques peuvent être mis en évidence. Une recherche d'objets rouges dans les données 2MASS permet de sélectionner des AGN très rouges, auparavant inconnus (Nelson et al., 1999).

1.4 Les catalogues de sources ponctuelles

Les autres chapitres de la première partie de cette thèse sont consacrés à l'analyse, à la manipulation et à la mise en ligne des catalogues de sources ponctuelles produits par DENIS et 2MASS. Cette section décrit brièvement la nature de ces données.

1.4.1 Caractéristiques

Les projets DENIS et 2MASS sont caractéristiques de l'astronomie moderne, qui est devenue une science du tout-numérique. Depuis les capteurs numériques, jusqu'aux catalogues interrogeables en ligne, tout le traitement est informatisé.

Le nombre de sources présent dans les catalogues est très grand, mais on a, de plus, énormément de paramètres pour chaque source. Ces paramètres sont le fruit de chaînes de traitement de plus en plus sophistiquées, où l'on veut garder un maximum d'information sur les processus d'extraction, de calibration et de caractérisation des sources.

Il faut aussi garder à l'esprit que les techniques modernes de traitement de l'information rendent envisageable la gestion de tels volumes de données.

1.4.2 Le catalogue DENIS

Le catalogue de sources ponctuelles DENIS devrait contenir plus de $100 \cdot 10^6$ sources.

Les données sont traitées et extraites au LDAC strip par strip. En décembre 2000, le nombre de strips traités au LDAC et au PDAC étaient respectivement supérieurs à 1000 et 2000. Un ensemble de 102 strips traités au LDAC constitue le premier échantillon de données ouvert au public (un peu plus de 17 millions de sources).

Dans chacun des deux catalogues, on peut avoir jusqu'à plus de 70 paramètres pour chaque source.

Le catalogue final sera le produit de la combinaison de tous les strips traités, et imposera donc la gestion de toutes les zones de recouvrement, pour que chaque source n'apparaisse qu'une seule fois.

1.4.3 Le catalogue 2MASS

On a vu que tout le traitement des données 2MASS était effectué par l'IPAC. Il y a eu au cours de ma thèse 3 phases de diffusion de données sur les sources ponctuelles, chacune correspondant à un catalogue remplaçant le précédent.

- le *2MASS Sampler*, publié en décembre 1998, était un premier échantillon de 227 197 objets ;
- la première grande diffusion (IDR1), en mai 1999, contenait 20 204 378 sources ;
- l'IDR2, en mars 2000, couvre 47% du ciel et contient 162 213 354 sources.

Le format des catalogues a très légèrement évolué, mais on a toujours une cinquantaine de paramètres pour chaque objet¹³. Le catalogue final 2MASS, dont la publication est prévue en 2002, devrait contenir environ $350 \cdot 10^6$ sources ponctuelles.

¹³Dans le catalogue de sources étendues, on a jusqu'à 400 paramètres différents pour chaque source !

Chapitre 2

Compression et indexation

2.1 Manipulation de grands volumes de données

2.1.1 Position du problème

ON A VU au chapitre précédent que les projets DENIS et 2MASS vont détecter un très grand nombre de sources : les catalogues de sources ponctuelles résultants contiennent des dizaines de paramètres pour des centaines de millions de sources. Nous devons alors faire face à plusieurs problèmes :

- la **taille des catalogues** devient énorme. Sous forme de fichiers textes, ces catalogues atteignent des volumes de quelques centaines de gigaoctets (Go). Les systèmes classiques de gestion de bases de données sont inadaptés pour traiter de tels volumes. Il est nécessaire de **compresser** au mieux les catalogues pour pouvoir les manipuler plus facilement ;
- il est difficile de garder un **accès rapide** aux données. Il faut optimiser les recherches dans le catalogue de manière à avoir les temps de réponse les plus courts possibles. La compression permet déjà de réduire la durée des opérations d'entrée/sortie (en diminuant les volumes de données à lire), mais on **indexe** également le catalogue, de manière à pouvoir accéder plus rapidement à l'information désirée ;
- on doit faire face à des **requêtes multicritères**. Le nombre de paramètres disponibles pour chaque objet étant très grand, les requêtes peuvent porter simultanément sur plusieurs contraintes (ex : quels sont les objets dans telle région ayant une magnitude inférieure à m , avec une couleur supérieure à c . . .).

Nous allons voir la façon dont nous avons géré les problèmes de **compression** et d'**indexation** pour la gestion des catalogues de sources ponctuelles de DENIS et de 2MASS.

2.1.2 Techniques mises en œuvre

Nous avons choisi d'appliquer aux catalogues de sources ponctuelles une compression sans perte d'information, en les codant sous forme de fichiers binaires, en adoptant une indexation sur la position céleste des sources (Derriere et al., 2000)¹.

Avec une méthode similaire, les 6 Go du catalogue USNO A1.0 (Monet et al., 1997) peuvent être compressés en 3.4 Go (Ochsenbein, 1998).

Pourquoi coder en binaire ?

L'unité d'information élémentaire en informatique est le **bit** (abréviation de *binary digit*). Les ordinateurs fonctionnent en base 2, et un bit ne peut prendre que les valeurs 0 ou 1.

Un **octet** est une séquence de 8 bits consécutifs. On parle le plus souvent en octets (ou en *bytes*) pour la transmission d'information. En effet, dans la norme ASCII (*American Standard Code for Information Interchange*), les principaux caractères typographiques (lettres, chiffres, ponctuation) sont codés sur 7 bits (ce qui donne $2^7 = 128$ caractères). Avec un huitième bit servant de validation ou de signe, chaque caractère de la table ASCII tient sur un octet : l'ASCII peut donc être considéré comme un alphabet machine, chaque lettre occupant un octet.

L'octet est donc en pratique l'unité de base pour transmettre de l'information.

Pour transmettre en ASCII une chaîne de x caractères, il faut utiliser x octets.

Voyons le cas où la chaîne en question est un nombre, par exemple « 29 ». Il faut en ASCII 2 octets (16 bits) pour transmettre cette chaîne (puisqu'on ne considère que le nombre de caractères typographiques utilisés, sans se soucier du sens véhiculé).

D'un point de vue numérique, 29 s'écrit en base 2 : 11101. 5 bits suffisent à coder cette valeur. Puisque les machines échangent ou lisent par défaut des octets, il reste 3 bits inutilisés sur l'octet où l'on a écrit '29'. Ces 3 bits peuvent servir à coder une autre valeur.

On voit donc l'intérêt de transformer les données du texte (ASCII) au binaire, pour la compression : dans l'exemple ci-dessus, on a réduit le nombre de bits nécessaires au codage de 16 à 5. Les catalogues de sources contiennent des ensembles de valeurs que l'on va s'efforcer de coder au mieux, en binaire.

Codage binaire

Le codage binaire repose sur le principe suivant : on transforme pour chaque source, chaque paramètre² p_i en un entier n_i , qui peut s'écrire en binaire sur un nombre de bits B_i . Les blocs de B_i bits sont ensuite concaténés pour former une chaîne de caractères binaire (Fig. 2.1).

¹Cette communication au colloque ADASS IX se trouve dans l'annexe C.2, p. 206.

²L'indice i distingue les différents paramètres présents pour chaque source.

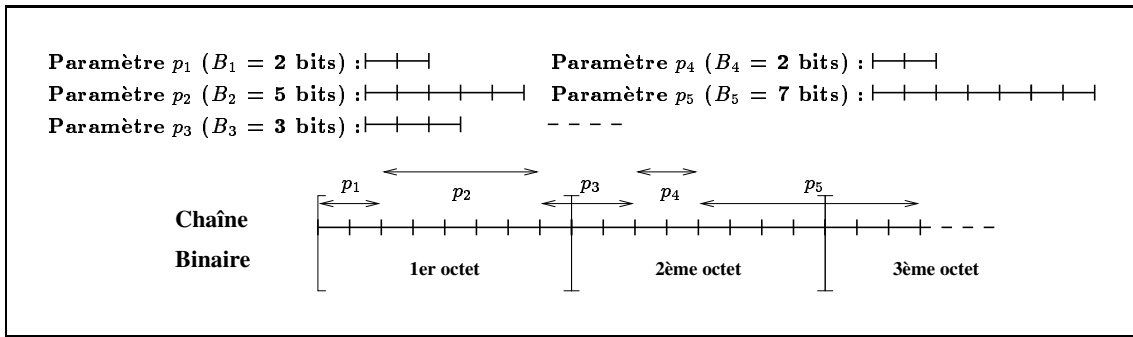


Fig. 2.1: Exemple simple de codage binaire d'un ensemble de paramètres p_i dans une chaîne de caractères.

Pour que le codage soit optimal, on doit s'efforcer de choisir les B_i les plus petits possibles.

Les valeurs individuelles³ $p_{i,j}$ de chaque paramètre p_i sont converties simplement en entiers positifs $n_{i,j}$ par des formules de la forme

$$n_{i,j} = \frac{p_{i,j} - \omega_i}{\varepsilon_i} \quad (2.1)$$

où ω_i est un seuil (la valeur minimale du paramètre, $\min_j(p_{i,j})$, par exemple), et ε_i est le pas élémentaire de variation de p_i .

Sur x bits, le plus grand entier pouvant être codé est $2^x - 1$. On trouve donc simplement le nombre de bits B_i nécessaire au codage du paramètre p_i à partir de la valeur maximale prise par n_i :

$$B_i = \lceil \log_2(\max_j(n_{i,j}) + 1) \rceil. \quad (2.2)$$

Dans l'équation ci-dessus, la notation $\lceil x \rceil$ représente l'arrondi à la valeur entière supérieure d'un nombre décimal x (fonction $\text{ceil}(x)$ en programmation).

Indexation

L'indexation sur les positions célestes facilite les requêtes en coordonnées, qui sont le type d'interrogation le plus courant (pour chercher si une source connue est détectée dans le catalogue, à une certaine position).

En pratique, on adopte un découpage hiérarchique de la région du ciel couverte par le catalogue en petites zones. Lors d'une requête autour d'une position, on peut accéder sélectivement aux zones concernées par la requête, et ne lire que les données s'y trouvant.

On doit donc conserver, en plus des données de chaque source, un certain nombre d'informations supplémentaires (positionnement des zones, nombre de sources par

³L'indice j distingue les différentes sources dans le catalogue.

zone, type de codage, paramètres pour le décodage (ω_i, ε_i ...). Ces informations peuvent être stockées dans un en-tête, avec les données binaires.

2.1.3 Un exemple simple de codage

Supposons que nous devons appliquer notre méthode de codage binaire aux données du tableau 2.1.

Ascension droite	Déclinaison	mag	Δ mag	f_1	f_2
134.186270	-57.559999	13.686	0.27	100	0
134.023046	-57.560433	7.420	0.06	99	1
134.262433	-57.560419	13.906	0.31	21	1
134.006997	-57.560934	13.331	0.22	100	0
134.143560	-57.560994	12.443	0.14	45	1
134.093502	-57.561394	14.150	0.36	100	0
134.266528	-57.560572	32.000	9.00	0	1
134.352799	-57.561070	32.000	9.00	0	0
134.048300	-57.562101	13.736	0.28	80	0
134.423419	-57.562048	32.000	9.00	0	1
134.211157	-57.560932	13.249	0.21	87	1
134.155622	-57.561257	10.104	0.10	100	0
134.083204	-57.562583	13.375	0.22	100	0

Tab. 2.1: Exemple simple de données à coder.

Il faut d'abord trouver le format optimal pour le codage, c'est à dire les valeurs des B_i , pour nos 6 paramètres.

Prenons l'ascension droite. Les valeurs extrêmes sont 134.006997 et 134.423419, et la précision de 6 décimales. En adoptant, dans l'équation 2.1, $\omega_{RA} = 134.006997$ et $\varepsilon_{RA} = 10^{-6}$, on peut convertir toutes les ascensions droites en entiers compris entre 0 et 416 422. D'après l'équation 2.2, il faut donc au minimum $B_{RA} = 19$ bits pour coder les valeurs de l'ascension droite.

Pour f_1 , qui est un entier compris entre 0 et 100, on obtient simplement avec $\omega_{f_1} = 0$ et $\varepsilon_{f_1} = 1$, $B_{f_1} = 7$ bits.

En appliquant le même genre de traitement à tous les paramètres de l'exemple, on peut en déduire le codage décrit dans le tableau 2.2.

Avec ces règles de codage, on peut convertir chaque valeur $p_{i,j}$ du catalogue de départ en un entier $n_{i,j}$ codé en binaire sur B_i bits. Pour chaque enregistrement j , la concaténation des valeurs binaires (comme sur la Fig. 2.1) forme une chaîne de caractères. Dans notre exemple, l'ensemble des B_i tient sur 64 bits, ce qui fait une chaîne de 8 octets, à comparer à la longueur des chaînes de caractères pour écrire complètement en ASCII les valeurs numériques (environ 40 octets par enregistrement).

Paramètre	ω_i	ε_i	B_i
Ascension droite	134.006997	10^{-6}	19
Déclinaison	-57.562583	10^{-6}	12
mag	0	10^{-3}	15
Δmag	0	10^{-2}	10
f_1	0	1	7
f_2	0	1	1
Total			64

Tab. 2.2: Exemple de codage binaire pour les données du tableau 2.1.

Valeurs à stocker pour le décodage

On voit que pour retrouver les valeurs de départ à partir des valeurs entières codées en binaire, il faut inverser l'équation 2.1.

Il faut donc stocker les ω_i et ε_i (dans un en-tête, par exemple) pour pouvoir effectuer le décodage par la suite.

2.2 Codage des données DENIS

L'unité de base des observations DENIS est le **strip**, une bande de 30° de long et $12'$ de large (voir section 1.2). Les deux centres de traitement de données (LDAC et PDAC) produisent, strip par strip, des catalogues de sources ponctuelles à partir des images.

Nous avons adopté pour le codage une architecture adaptée au traitement des données au fur et à mesure de leur disponibilité, en créant un fichier de données binaire par strip. On ne gère donc pas, dans un premier temps, les associations de sources se trouvant dans une zone de recouvrement entre deux strips adjacents. Ces sources seront présentes en plusieurs exemplaires (un dans chaque strip) dans la base.

2.2.1 Catalogues LDAC et PDAC

Les catalogues issus des traitements au LDAC et au PDAC ne sont pas identiques, que ce soit pour la façon de les produire ou pour leur format.

LDAC

Le LDAC utilise le logiciel SExtractor (Bertin & Arnouts, 1996) pour effectuer l'extraction des sources ponctuelles. La chaîne de traitement des données a évolué au cours de ma thèse, et le nombre de paramètres pour chaque source a augmenté. Les premières versions des catalogues de strips contenaient les paramètres suivants :

- la position équatoriale (α, δ) , l'incertitude sur la position et un identificateur ;

- pour chacune des trois bandes photométriques :
 - deux magnitudes calculées par photométrie d'ouverture (en mesurant le flux dans des cercles centrés sur la source, de rayons 7'' et 15''), et les erreurs associées ;
 - les paramètres d'une ellipse circonscrite à la source ;
 - un indicateur de stellarité (donné par SExtractor) ;
 - divers *flags* concernant l'extraction.

En mai 1998, une magnitude de type Kron (Kron, 1980) a été ajoutée dans chaque bande, ainsi qu'une indication sur la présence d'autres objets dans le voisinage immédiat de chaque source, et le numéro de l'image d'où est extraite la source.

Pour les données réduites en 2000, un quatrième système photométrique a été introduit, utilisant un ajustement de la PSF, et les contreparties optiques (magnitudes *B* et *R* de l'USNO A2.0) sont également données, si elles existent.

On a finalement 71 paramètres par source, dans la dernière version du traitement LDAC (Fig. 2.2).

Il faut noter que les zones de recouvrement entre deux images successives dans un même strip sont traitées au LDAC, de façon à ce qu'il n'apparaisse dans le catalogue final qu'une seule source. Les deux extractions d'un même objet dans les deux images sont fusionnées en un seul enregistrement dans le catalogue de strip résultant (en prenant la moyenne des magnitudes, et une moyenne pondérée des coordonnées).

PDAC

L'extraction de sources au PDAC utilise les algorithmes développés par Alard & Lupton (1998). Pour chaque source, les magnitudes sont calculées pour chaque bande dans 7 systèmes photométriques, dont l'un utilise un ajustement polynomial de la PSF (Alard, 2000).

On a pour chaque source 72 paramètres, qui diffèrent en partie de ceux du LDAC, parmi lesquels :

- la position équatoriale (α, δ), le numéro d'image et de strip, et le jour julien ;
- pour chacune des trois bandes photométriques :
 - les 7 magnitudes et les erreurs associées ;
 - des indicateurs de la qualité de l'extraction et de la corrélation à la PSF ;
 - les positions en pixel dans l'image ;
 - un *flag* (pour indiquer si l'objet est saturé, s'il s'agit probablement d'un artéfact...).
- les magnitudes *B* et *R*, l'époque et la distance à la contrepartie optique (catalogue USNO A2.0).

Dans la base de données du PDAC, les zones de recouvrement entre deux images successives dans un même strip ne sont pas traitées. Si on a deux détections d'un

même objet présent sur deux images adjacentes, il y aura deux enregistrements distincts dans la base de données. La question de la combinaison des détections multiples pour la production d'un catalogue sera abordée au chapitre 3, avec la validation des données.

2.2.2 Aspects communs pour le codage

Il est évident qu'il y a une grande similitude entre les catalogues produits au LDAC et au PDAC, puisqu'ils utilisent, à la base, les mêmes données. Nous avons choisi de coder, dans les deux cas, les strips individuellement. Les grandes lignes du codage sont donc très semblables.

La taille des catalogues dépend de la zone du ciel traversée par le strip. Les étoiles sont en effet localisées en grande majorité dans le plan galactique. Les strips à haute latitude galactique contiennent en général quelques dizaines de milliers de sources, mais ceux traversant le bulbe peuvent contenir plus d'un demi-million d'objets.

Stockage des positions

Un strip correspond à une zone bien déterminée sur le ciel. Les positions (α, δ) dans les catalogues sont données comme des nombres décimaux, avec 6 décimales de précision.

Des sources voisines ont des valeurs de la position assez proches : plutôt que de coder la valeur complète des positions, on peut, après avoir stocké des valeurs (α_x, δ_x) de référence (les ω_i de la formule 2.1), ne coder les positions que comme des décalages (de petits entiers) par rapport à ces valeurs.

Nous avons tout d'abord trié les sources en déclinaison, puis nous les avons regroupées en paquets de 12' (0.2°). On stocke une position de référence pour chaque intervalle en déclinaison. Il suffit alors de 18 bits par source ($0.2/10^{-6} < 2^{18}$) pour coder les valeurs des décalages en déclinaison par rapport à la position de référence.

La situation est légèrement différente pour l'ascension droite. En effet, si la zone couverte par un strip fait toujours 0.2° (en distance angulaire sur le ciel), cela correspond à un intervalle d'autant plus grand que l'on se trouve à haute latitude ($\Delta\alpha = 0.2/\cos(\delta)$).

A la limite, au pôle, l'ascension droite peut prendre toutes les valeurs de 0.0 à 360.0 degrés. Pour coder en binaire toutes les valeurs possibles avec 6 décimales de précision, soit $3.6 \cdot 10^8$ valeurs, il faut 29 bits (on conserve ici toutes les décimales, même si elles n'ont pas un bien grand sens physique).

Nous avons introduit la possibilité de coder l'ascension droite sur 29 bits (près du pôle sud) ou 21 bits (un octet de moins), selon le cas. On peut gagner un octet pour les zones situées plus au nord que $\delta \approx -84.5^\circ$, ce qui représente la majorité des données. On stocke, dans un en-tête associé à chaque zone de déclinaison, la façon dont sont codées les données des ascensions droites pour cette zone (21 ou 29 bits).

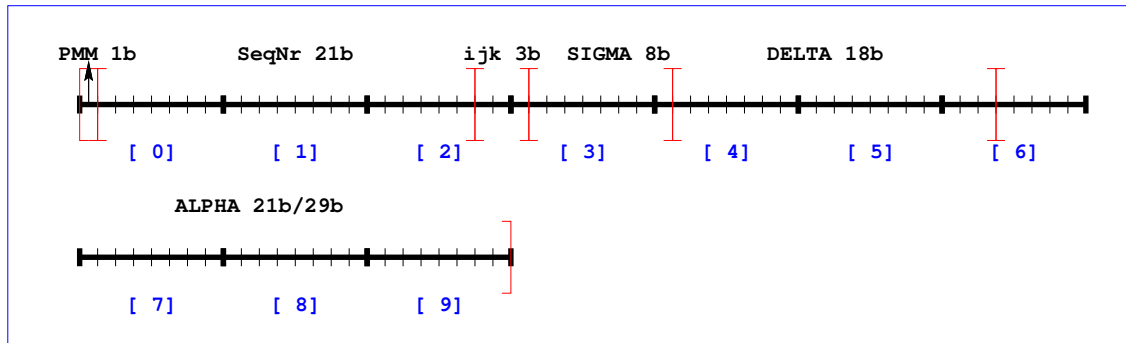
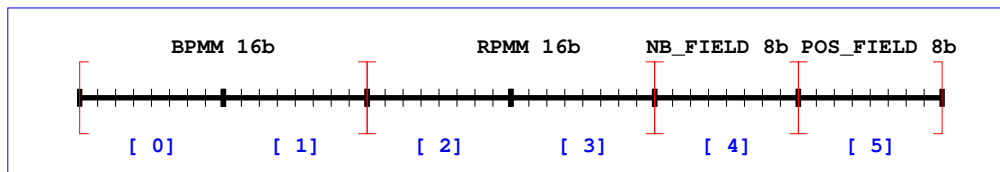
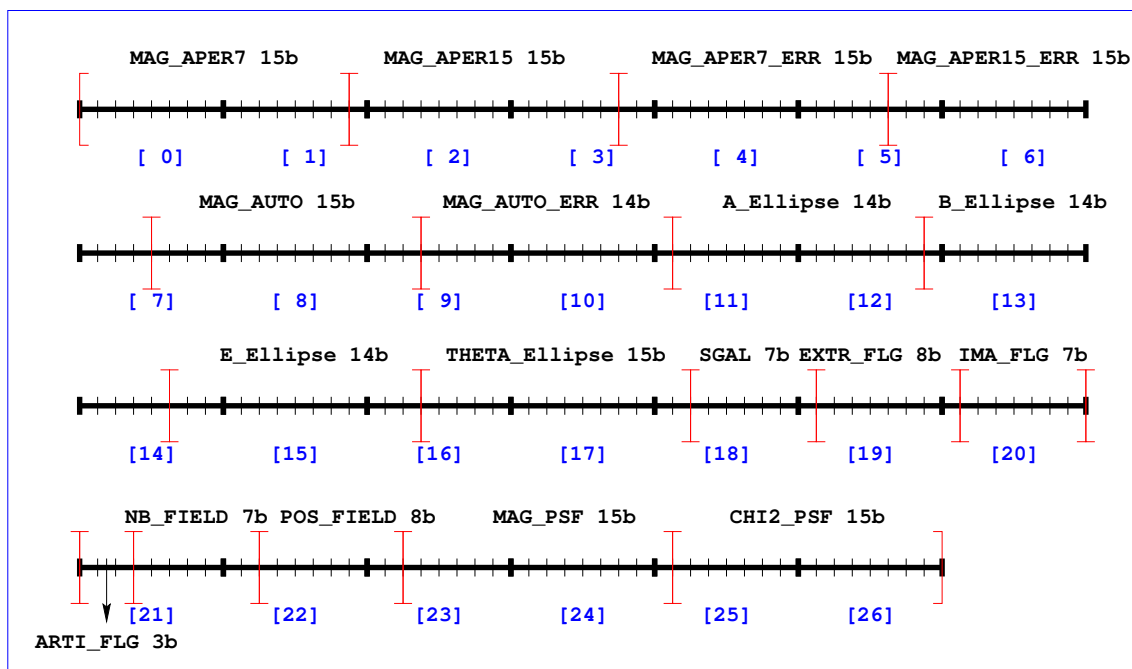
POSITION:**DONNÉES OPTIQUES:****DONNÉES PHOTOMÉTRIQUES:**

Fig. 2.2: Schéma du codage en binaire des catalogues DENIS issus du LDAC. On a reporté au dessus des chaînes les noms des paramètres et le nombre de bits nécessaire pour les coder, et en dessous, entre crochets, les numéros des octets dans la chaîne. On notera les trois catégories principales : l'information sur la position, toujours présente, et les informations sur la contrepartie optique (USNO) et les magnitudes (I , J , K_s), optionnelles.

Cas des données manquantes

En raison des limites de détection différentes des 3 caméras, mais aussi des couleurs intrinsèques des objets, un objet DENIS n'est pas forcément toujours détecté dans les 3 bandes photométriques. De même, un objet n'a pas forcément de contrepartie optique dans l'USNO.

Il est dans ce cas avantageux pour la compression de ne pas inclure du tout, dans le fichier binaire, de données pour ces informations absentes (plutôt que d'écrire des séries de 0 inutiles). Encore faut-il pouvoir retrouver, au décodage, les enregistrements où on a codé des données et ceux où on les a ignorées.

Nous avons défini trois types de données distinctes, que l'on code chacune sur une chaîne de caractères :

- chaîne 1 : les données toujours présentes pour un objet (ex : position) ;
- chaîne 2 (optionnelle) : les données photométriques (ex : magnitudes, erreurs). Cette chaîne a la même structure pour chacune des couleurs I , J , et K_s ;
- chaîne 3 (optionnelle) : les données optiques (contrepartie USNO).

Nous allons réserver dans la première chaîne, et pour chaque source, quatre bits dont la valeur (0 ou 1) indiquera l'absence ou la présence de la chaîne optionnelle correspondante (I , J , K_s , ou USNO).

Quel gain peut-on attendre ? Prenons le cas de la bande I , dont les données sont codées dans une chaîne de L_I caractères ($8L_I$ bits). On ajoute, pour **toutes** les sources, **1 bit** b_I indiquant si l'objet est détecté dans la bande I ou non. Si $b_I = 1$, on écrit $8L_I$ bits de données pour cette bande. Si $b_I = 0$, on n'écrit rien.

Soit f la proportion de sources non détectées en I . Par rapport à la situation où on code systématiquement $8L_I$ bits de données sans se préoccuper de savoir si il y a détection ou pas, on gagne de la place si :

$$1 \times f + (1 + 8L_I) \times (1 - f) < 8L_I ,$$

c'est à dire si $f > \frac{1}{8L_I}$.

En pratique (voir Fig. 2.2, $L_{\text{mag}} = 27$), cette condition est largement validée pour les données DENIS, et permet un gain de place important lors du codage (de l'ordre de 45%).

Structure d'un fichier binaire

Le découpage en position adopté se retrouve dans la structure du fichier binaire : il est divisé en blocs de données correspondant aux différentes zones. Le nombre de sources dans chaque bloc n'est bien sûr pas constant, et on vient de voir que la taille des données codées par source peut changer d'une source à l'autre (en fonction du nombre de chaînes optionnelles effectivement écrites).

La figure 2.3 schématise l'organisation d'un fichier binaire :

- en début de strip, on code quelques informations globales (numéro de strip, positions extrêmes, nombre total de sources, ...);
- on a ensuite une série d'en-têtes (un par zone de déclinaison) contenant des informations globales pour chaque zone (valeurs de référence pour la reconstruction des positions, nombre d'objets dans le bloc de données correspondant), ainsi que le nécessaire pour accéder au bloc de données binaires (position du début, longueur);
- enfin, on a les blocs de données binaires, composés de la juxtaposition de toutes les chaînes de caractères pour chaque source (le nombre de chaînes stockées varie, comme on peut le voir en bas de la figure 2.3).

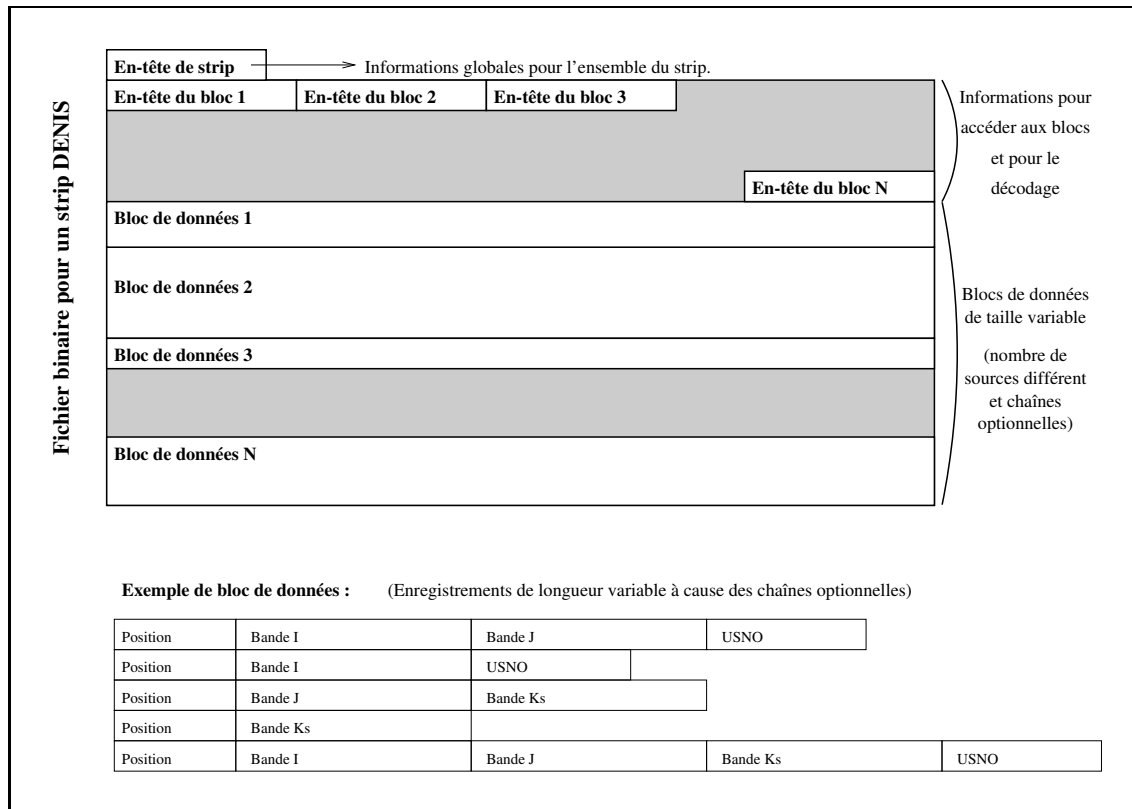


Fig. 2.3: Structure d'un fichier binaire de données pour un strip DENIS.

2.2.3 Astuces spécifiques et performances

Pour les données du PDAC, nous avons utilisé quelques astuces dans le but de réduire encore la taille des données à coder :

- parmi les paramètres disponibles pour chaque source, on a le numéro de l'image d'où elle est extraite. C'est un entier long (il y a plus d'un million d'images), mais qui ne prend que 180 valeurs différentes dans un strip. On stocke donc

- le numéro de la première image dans l'en-tête du strip, et on ne garde pour chaque source qu'un décalage entre 0 et 179 (sur 8 bits) ;
- pour les sources ayant une contrepartie optique, on a l'époque de la plaque photographique correspondante. C'est un nombre décimal (par exemple 1979.703). En pratique, le nombre de valeurs différentes de ce paramètre dans un strip est faible (car il y a un nombre restreint de plaques sur la zone couverte par un strip). On stocke donc directement dans l'en-tête du strip les x valeurs différentes, et on code pour chaque source la valeur par un indice entier entre 0 et $x - 1$, beaucoup plus court à écrire que la valeur complète.

Performances

Pour les strips du LDAC, les fichiers binaires sont en moyenne compressés d'un facteur 13 par rapport aux catalogues ASCII. La commande `gzip`, permet un taux de compression de seulement 6.5, en moyenne, sur ces mêmes fichiers ASCII. Notre méthode de compression est donc deux fois plus performante que l'algorithme LZ77 utilisé par la commande `gzip`, mais surtout nos fichiers compressés sont indexés en position, et l'accès aux données à l'intérieur est possible très rapidement, sans avoir besoin de décompresser le fichier complet.

Les strips traités au LDAC ont été recopiés au CDS et codés en binaire : plus de 500 strips sont disponibles dans une base de données accessible aux membres du consortium DENIS, et 102 strips parmi ceux-ci ont été sélectionnés pour une première distribution publique des données (Epchtein et al., 1999), après validation (voir chapitre 3).

Pour les strips du PDAC, le facteur de compression obtenu pour les fichiers binaires est de 9.3, à comparer avec un facteur d'environ 6.2 avec la commande `gzip`. La compression est un peu moins importante en raison d'une présentation plus compacte des fichiers ASCII, et d'un nombre de paramètres optionnels moins importants que pour le LDAC.

Les données du PDAC sont actuellement en cours de traitement au CDS.

Les performances en termes de rapidité d'accès aux données, ainsi que les interfaces pour l'accès sont discutées au chapitre 4.

2.3 Codage des données 2MASS

2.3.1 Introduction

La stratégie de codage pour les catalogues de sources ponctuelles de 2MASS est légèrement différente. En effet, nous ne recevons pas au CDS les données au fur et à mesure de leur traitement, mais nous recevons les catalogues complets lors des diffusions publiques successives (*2MASS sampler*, *Incremental Data Release 1 et 2 (IDR)*).

Les catalogues successifs correspondent à un ensemble de tuiles 2MASS dispersées sur le ciel complet (cf. Fig. 2.4).



Fig. 2.4: *Projection Aitoff en coordonnées galactiques des zones couvertes par la deuxième grande publication de données 2MASS. La densité de sources est figurée en niveaux de gris.*

2.3.2 Découpage du ciel

Puisque les données peuvent être réparties sur l'ensemble du ciel (qui sera complètement couvert, à terme), on doit adopter un mode de découpage en position pour construire la base de données.

Les données seront stockées dans des fichiers binaires. Le catalogue final de sources ponctuelles 2MASS devrait contenir plus de 300 millions de sources, ce qui correspond, comme on va le voir, à plus de 20 Go de données compressées.

On doit conserver une taille de fichiers raisonnable, mais également veiller à ne pas avoir trop de fichiers dans un même répertoire, pour des raisons de performance. Nous avons donc adopté le découpage en fichiers suivant pour construire la base :

- des répertoires correspondant à des bandes de 6° en déclinaison ;
- dans chaque répertoire, 72 fichiers correspondant à des tranches de 20 minutes en ascension droite.

On aura donc en moyenne 150 000 sources par fichier, ce qui représente une taille de l'ordre de 10 Mo^4 .

⁴Ceci est une approximation très grossière. En pratique, l'écrasante majorité des sources ponctuelles est localisée dans le plan galactique, et il y aura des différences de taille importantes d'un fichier à l'autre.

A l'intérieur d'un fichier, on adopte également une indexation en position : nous allons définir dans chaque fichier un certain nombre de zones, avec pour chacune une position de référence, et dans chaque zone les positions seront stockées sous forme d'un écart à la position de référence.

Comme on l'a déjà vu pour les catalogues DENIS, un en-tête de données sera associé à chaque zone.

Comment optimiser le choix du niveau de sous-découpage en zones dans un fichier ? Quand on divise par deux la taille des zones en déclinaison, par exemple, on divise également par deux la valeur maximale à coder pour la distance en déclinaison dans une zone. On peut donc écrire cette valeur sur un bit de moins pour toutes les sources. Mais puisqu'il y a deux fois plus de zones, le nombre d'en-têtes à stocker va doubler.

Il faut tenter de minimiser le volume total V occupé par le codage des positions (α, δ) et des en-têtes. On peut remarquer que si les positions sont données avec 6 décimales de précision (comme c'est le cas pour 2MASS), on peut coder toutes les valeurs d'un intervalle de $15'$ sur 18 bits ($0^\circ 15' \times 10^6 < 2^{18}$).

Si on note b_α et b_δ les nombres de bits utilisés pour coder l'ascension droite et la déclinaison, N_T le nombre total d'objets dans le catalogue, et L la taille d'un en-tête en octets, le volume V (en octets) à minimiser est

$$V = N_T \cdot \frac{b_\alpha + b_\delta}{8} + L \cdot \frac{360^\circ \cdot 2^{18}}{15' \cdot 2^{b_\alpha}} \cdot \frac{180^\circ \cdot 2^{18}}{15' \cdot 2^{b_\delta}}$$

ou bien encore en simplifiant

$$V = N_T \cdot \frac{b_\alpha + b_\delta}{8} + 2025L \cdot 2^{27-(b_\alpha+b_\delta)}.$$

Comme on peut le voir sur la figure 2.5, cette fonction présente bien un minimum qui est la valeur optimale pour le découpage en position dans un fichier.

2.3.3 Codage

La méthode employée pour le codage ressemble fortement à celle utilisée pour DENIS. On a regroupé en chaînes de caractères les différents types de données (position, données photométriques, données optiques, voir figure 2.6). On s'arrange dans chaque cas pour que l'ensemble des valeurs occupe au mieux un nombre entier d'octets. Avec cette contrainte, nous avons adopté pour le découpage en position un codage sur 16 bits de l'ascension droite et de la déclinaison, ce qui correspond à des petites zones de $3'45''$.

Une différence notable par rapport aux données DENIS est qu'il y a toujours un minimum de données dans chacune des trois bandes 2MASS. En effet, dès qu'une source est détectée dans une des trois bandes 2MASS, une mesure de flux est effectuée dans les autres bandes au même endroit, quitte à ne fournir qu'une limite supérieure sur la magnitude si la détection échoue. La méthode consistant à réserver un bit pour indiquer la présence optionnelle d'un bloc de données est donc pour 2MASS seulement applicable aux données optiques.

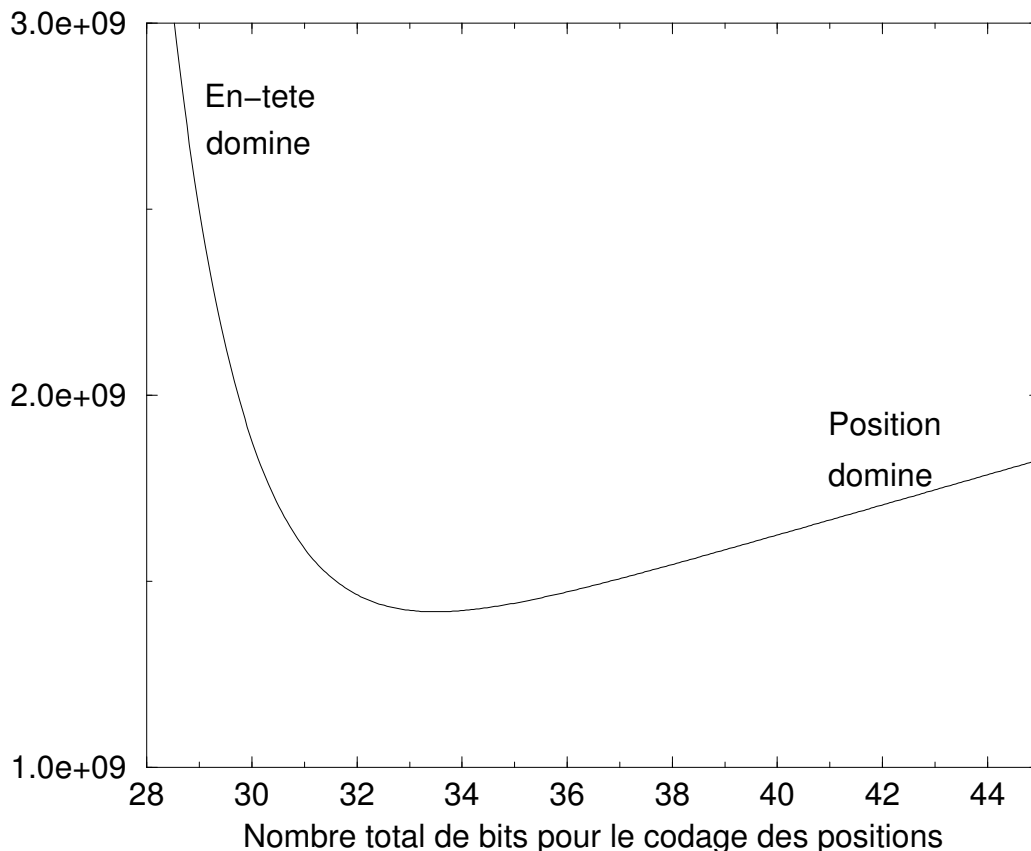


Fig. 2.5: Recherche du découpage en zones optimal dans les fichiers de données 2MASS. On a tracé ici la variation de l'espace total V (en octets) occupé par le codage des positions et des en-têtes, en fonction du nombre de bits ($b_\alpha + b_\delta$) choisi pour le codage des positions, pour 325 millions de sources, avec 12 octets par en-tête.

Quelques astuces

Date d'observation : Pour chaque source 2MASS, la date d'observation est fournie (sous la forme AAMMJJ, par exemple). Plutôt que de coder cette chaîne de caractères (6 octets, soit 48 bits), nous avons converti la date en un nombre de jours, à compter du début des observations 2MASS. On peut ainsi coder la valeur sur 11 bits (ce qui fait 2048 jours possibles, et couvre amplement la durée des observations).

χ_{PSF}^2 : La technique de codage présentée dans la section 2.1.2 est bien adaptée au cas où les valeurs à coder sont réparties de façon à peu près homogène. Il y a un paramètre dans les catalogues 2MASS pour lequel ce n'est pas du tout le cas. Il s'agit du χ^2 réduit de l'ajustement de la PSF. Dans l'immense majorité des cas, cette valeur est très faible, et on a donc des petits nombres à coder. Mais dans quelques

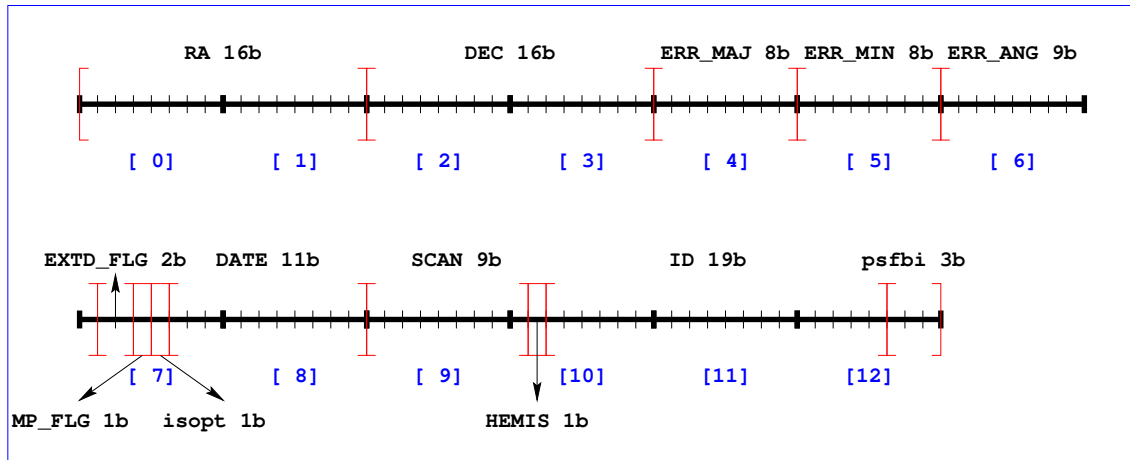
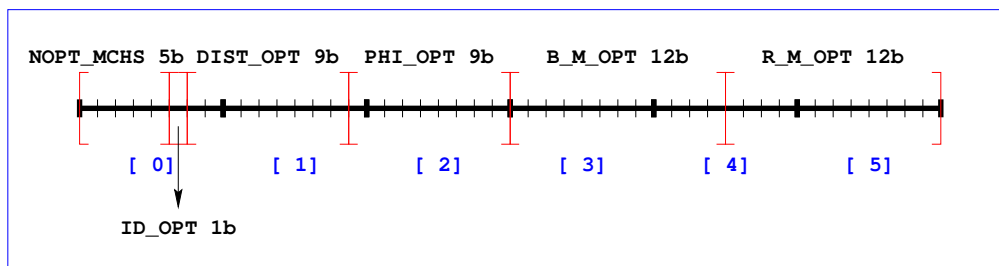
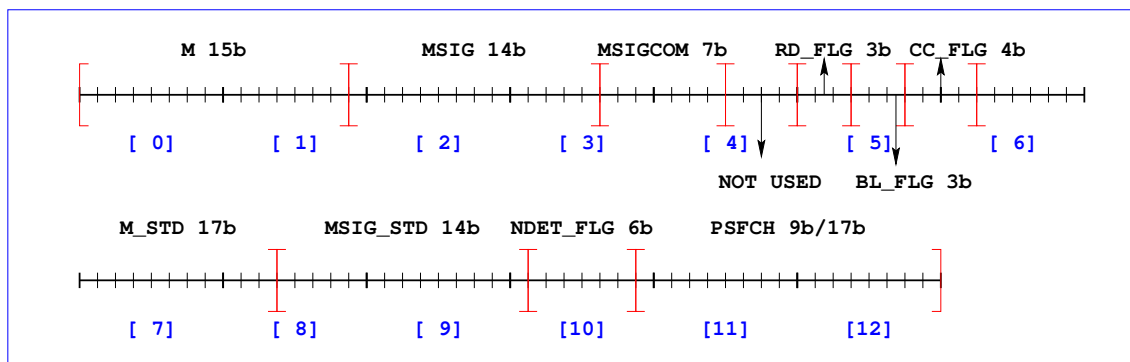
POSITION:**DONNÉES OPTIQUES:****DONNÉES PHOTOMÉTRIQUES:**

Fig. 2.6: Schéma du codage en binaire du catalogue 2MASS de l'IDR2. On a reporté au dessus des chaînes les noms des paramètres et le nombre de bits nécessaire pour les coder, et en dessous, entre crochets, les numéros des octets dans la chaîne.

rare cas, ce paramètre prend des valeurs très élevées. On a schématisé la situation sur la figure 2.7.

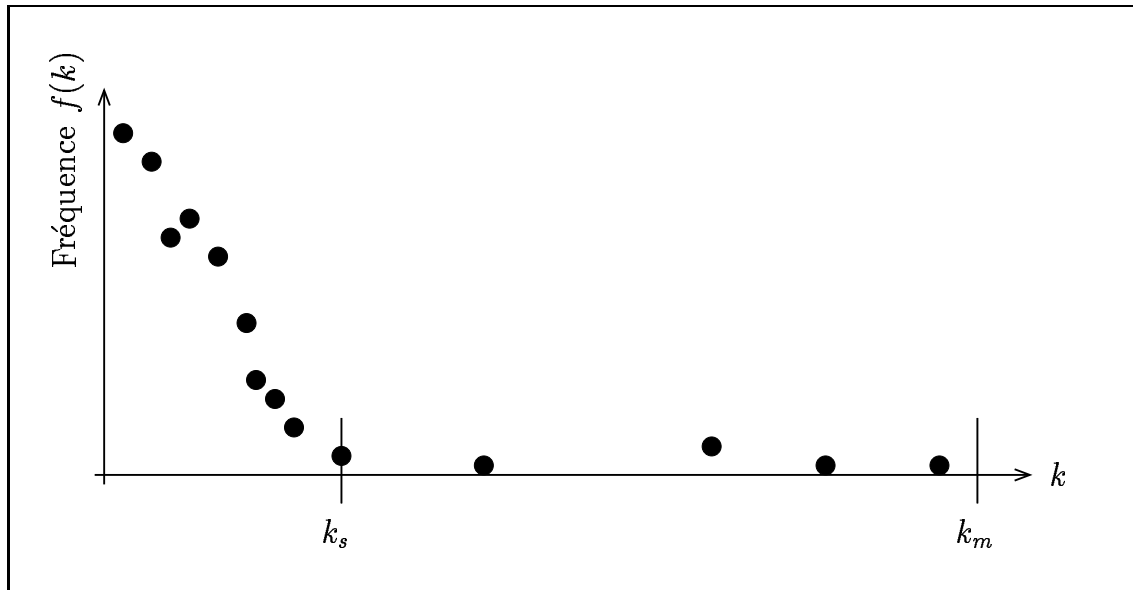


Fig. 2.7: Exemple de cas où un codage à deux modes peut s'avérer utile. Les valeurs à coder k ne dépassent jamais k_m . Le plus souvent, k est très petit (et tient donc sur peu de bits). Mais il y a quelques cas où k est très grand, et nécessite un nombre de bits bien plus important pour le codage.

Nous avons introduit pour ce paramètre la possibilité d'un codage à deux modes. Soit k la valeur entière à coder, avec $0 \leq k \leq k_m$. On définit une valeur seuil k_s :

- si $k \leq k_s$, on code le paramètre sur $B_s = \lceil \log_2(k_s + 1) \rceil$ bits ;
- si $k_s < k \leq k_m$, on code le paramètre sur $B_m = \lceil \log_2(k_m + 1) \rceil$ bits.

Si on se limitait à un codage simple (un seul mode), avec N valeurs à coder, la taille totale occupée (en bits) serait simplement

$$V_1 = N \cdot B_m$$

Si on nomme $f(k)$ la fréquence relative de la valeur k dans le catalogue, la taille occupée devient

$$V_2 = \left(\sum_{k=0}^{k_s} f(k) \right) \cdot B_s + \left(\sum_{k=s+1}^{k_m} f(k) \right) \cdot B_m + 1.$$

Le 1 dans l'expression ci-dessus vient du fait qu'il faut réserver ailleurs un bit pour préciser le mode de codage utilisé pour chaque valeur (codage court ou long). Le codage à deux modes apporte un gain si $V_2 < V_1$, c'est à dire s'il existe k_s tel que

$$\sum_{k=0}^{k_s} (B_m - B_s) f(k) > 1.$$

Attention, B_s dépend de k_s dans cette expression !

Pour les données 2MASS, l'utilisation d'un codage à deux modes ($B_s = 9$ bits, $B_m = 17$ bits) pour le χ^2 réduit nous a permis de gagner en moyenne 8 bits par valeur.

2.3.4 Résultats

Nous avons codé au CDS les données 2MASS après chaque diffusion publique (les dates et volumes sont donnés page 29). Le volume de données disponibles s'est accru progressivement, chaque diffusion incluant (avec quelques modifications de format, éventuellement) la diffusion précédente. Afin de garder la même logique, nous avons remplacé l'ensemble de la base 2MASS au CDS par les dernières données disponibles.

Sous forme binaire, l'IDR2 occupe un volume de 8.6 Go, soit environ 53 octets par source. La compression des catalogues ASCII avec `gzip` permet de gagner un facteur 3.3, mais on gagne encore en binaire un facteur 1.6 par rapport à `gzip`.

Mais là encore, les fichiers binaires sont indexés, et directement consultables sans qu'il soit nécessaire de les décompresser.

On voit le que taux de compression (≈ 5.4) est plus faible que celui atteint avec les catalogues de strip DENIS. Cela tient en grande partie au fait que l'on ne peut plus gagner de place avec l'absence de données photométriques comme c'était le cas avec DENIS.

Les performances et les méthodes pour l'accès aux données sont discutées au chapitre 4.

Chapitre 3

Validation des données DENIS

3.1 Une étape essentielle

LA VALIDATION des données est une étape essentielle dans la production d'un catalogue de qualité. J'ai été conduit à mener un certain nombre de tests sur les sources ponctuelles DENIS extraites par les chaînes de traitement du LDAC et du PDAC.

Ce travail de validation est bien sûr déjà en partie effectué par les centres de traitement eux-mêmes. Toutefois, les motivations pour effectuer une validation supplémentaire du contenu des catalogues DENIS sont multiples :

- on veut s'assurer que les catalogues produits sont de bonne qualité, tant au niveau de la forme (les données respectent bien le format annoncé) qu'au niveau du fond (les valeurs —positions, magnitudes— sont aussi exactes que possible). Ceci est vrai tant pour les données privées (accessibles seulement aux membres du consortium DENIS), que pour la production des catalogues publics, qui vont être distribués à la communauté astronomique internationale ;
- les deux centres de traitement utilisent des méthodologies légèrement différentes. On peut comparer leurs avantages et inconvénients respectifs. La mise en évidence de problèmes dans un catalogue peut éventuellement aboutir à des améliorations de la chaîne de traitement correspondante ;
- on peut identifier des données à retraiter, ou à réobserver (si les problèmes détectés viennent de mauvaises conditions météorologiques lors des observations, par exemple) ;
- cela nous a également servi pour sélectionner les données utilisées dans la deuxième partie de la thèse, pour l'étude de la structure galactique.

Cette validation des données est sensiblement plus facile à faire sur les catalogues, une fois que nous les avons codés en binaire, selon la méthode décrite au chapitre précédent.

Tout d'abord, le codage en soi permet de réaliser un certain nombre de tests sur le contenu des catalogues. En effet, la méthode de compression utilisée revient

à coder l'information dans un volume minimal. La moindre déviation du contenu des catalogues par rapport au format imposé pour le codage est immédiatement sanctionnée par une erreur. Par exemple, pour un paramètre variant de 0 à 100 (que l'on codera sur 7 bits), toute valeur négative, ou supérieure à 127 ($2^7 - 1$) sera mise en évidence simplement en testant la concordance avec le format imposé.

D'autre part, les données, une fois codées en binaire, sont beaucoup plus aisément manipulables, ce qui facilite grandement l'application de programmes de tests.

La suite de ce chapitre, après quelques rappels sur l'astrométrie, présente les tests effectués pour valider l'astrométrie et le filtrage des données DENIS, ainsi que quelques comparaisons entre des données réduites au PDAC et au LDAC.

3.2 Quelques notions d'astrométrie

L'astrométrie est la branche de l'astronomie consacrée à l'étude de la position des astres et de leurs mouvements.

Depuis les mesures de l'Antiquité avec un bâton de Jacob jusqu'aux satellites modernes, des progrès considérables ont été accomplis, mais on a dans tous les cas un système de référence, dans lequel on définit les positions d'une liste d'objets de référence.

3.2.1 Systèmes de référence : FK5, ICRS

C'est actuellement la division I de l'UAI (Union Astronomique Internationale) qui est en charge de définir et de faire respecter les systèmes de référence astrométriques. Un changement important a eu lieu le 1^{er} janvier 1998, avec l'adoption d'un nouveau système de référence céleste.

Auparavant, les systèmes en vigueur étaient fondés sur l'équateur et un équinoxe (déduits des observations du mouvement des planètes dans le système solaire), pour une certaine époque de référence, avec un ensemble de règles de transformation des positions.

Pour les applications pratiques, on utilisait des listes d'étoiles de référence (FK3, FK4, FK5), avec des positions exprimées dans le système ainsi défini.

Le FK5 J2000 (*Fundamental Katalog No. 5*, Fricke et al. (1988)), par exemple, est un catalogue de 1535 étoiles brillantes (jusqu'à la magnitude 7.5, avec une extension de 3117 étoiles jusqu'à la magnitude 9.5). L'incertitude sur les positions est de l'ordre de 30 à 40 millisecondes d'arc (mas) dans le FK5¹.

Depuis 1998, l'UAI a adopté comme système de référence l'ICRS (*International Celestial Reference System*, Feissel & Mignard (1998)).

¹Les erreurs systématiques y sont cependant plus grandes.

L'**ICRS** est un ensemble de prescriptions et de conventions, ainsi que la modélisation permettant de définir à tout moment un système d'axes de référence (trièdre).

L'innovation majeure est que le nouveau système n'est plus lié aux mouvements dans le système solaire, puisqu'il est centré sur le barycentre du système solaire et que ses axes sont orientés par rapport aux positions d'un ensemble de sources extragalactiques.

Ce changement a été motivé par les deux éléments suivants :

- les techniques d'interférométrie radio à très grande base (VLBI, *Very Long Baseline radio Interferometry*) permettent d'atteindre des précisions sur les positions de l'ordre d'une milliseconde d'arc ;
- de plus, les sources radio extragalactiques sont des objets très éloignés (essentiellement des quasars), et leur mouvement propre est trop faible pour être mesuré. Elles constituent donc de meilleures balises que des étoiles de notre Galaxie.

Bien qu'il y ait a priori aucune obligation en ce sens, pour des raisons de continuité, les axes de l'ICRS ont été choisis proches de ceux du FK5 J2000.

L'**ICRF** (*International Celestial Reference Frame*, Ma et al. (1998)) est une réalisation pratique de l'ICRS.

Il est constitué d'un ensemble de directions fiduciaires pour 608 sources radio extragalactiques, en accord avec les prescriptions de l'ICRS. La précision sur le système d'axes défini par l'ICRF est estimée à 0.02 mas.

On pourra se reporter à l'article de revue de Johnston & de Vegt (1999) pour plus de détails sur les systèmes de référence.

3.2.2 Catalogues astrométriques

Un système comme l'ICRS reste quelque chose d'abstrait. En pratique, on a besoin de catalogues de référence. L'ICRF défini par les sources radio n'est pas le mieux adapté pour les besoins courants dans les domaines radio ou infrarouge proche. On utilise donc dans ces domaines d'autres catalogues.

Hipparcos, Tycho, ACT

Les positions des quelque 120 000 étoiles du catalogue² de la mission astrométrique spatiale Hipparcos (ESA, 1997), ainsi que leurs mouvements propres, sont donnés dans le système de coordonnées de l'ICRS, pour l'époque J1991.25. Avec une

²On peut noter que toutes les étoiles du système FK5 figurent dans le catalogue Hipparcos.

incertitude sur l'alignement avec l'ICRS de 0.6 mas pour chacun des axes, **le catalogue Hipparcos est la meilleure réalisation de l'ICRF dans le domaine optique.**

Le catalogue Tycho (ESA, 1997) contient un peu plus d'un million d'étoiles observées par le repéreur d'étoiles (*star mapper*) du satellite Hipparcos. La précision sur les positions est environ 30 fois moins bonne pour ces sources, plus faibles que celles du catalogue Hipparcos.

Le catalogue ACT (Urban et al., 1998) fournit des mouvements propres avec une grande précision pour la plupart des étoiles du catalogue Tycho, en combinant les positions Tycho avec celles du Catalogue Astrographique (AC). Le TRC (*Tycho Reference Catalogue*, Høg et al. (1998)) est une autre combinaison de l'AC avec Tycho, réalisée indépendamment par l'équipe de Copenhague. La précision sur les positions J1991.25 est de l'ordre de 40 mas, et la précision sur les mouvements propres de 2.5 mas/an, pour chacun des deux catalogues.

La récente parution du catalogue Tycho-2 (Høg et al., 2000), qui contient les positions et mouvements propres de plus de 2.5 millions d'étoiles, marque un nouveau progrès. Il est issu des mêmes observations, mais grâce à un traitement plus sophistiqué, il est un peu plus précis et contient un nombre d'objets nettement plus grand que les catalogues Tycho, ACT et TRC.

GSC, USNO

Pour des applications pratiques de calibration astrométrique, on peut avoir besoin d'un nombre d'objets de référence bien plus grand que celui contenu dans des catalogues comme Hipparcos ou l'ACT. Les grands catalogues issus de la numérisation de plaques photographiques sont alors une bonne base de travail.

Ces catalogues présentent l'avantage de contenir un très grand nombre d'objets, mais en contrepartie la précision est nettement moins bonne que celle d'Hipparcos (on a affaire à plus d'objets faibles). Ce ne sont donc pas des catalogues astrométriques utilisables pour des mesures de très grande précision, mais ils sont suffisants pour des calibrations nécessitant une grande densité de sources de référence, avec des positions connues à 0.5–1" environ.

Parmi ces grands catalogues pouvant servir de référence, on peut citer :

- le GSC (*Guide Star Catalog*, Lasker et al. (1990)), qui contient plus de 25 millions de détections (mais pas autant de sources car il y a des détections multiples), extraites d'une collection de plaques de Schmidt d'une seule époque (≈ 1982) couvrant l'ensemble du ciel (Quick V Palomar pour l'hémisphère nord, et UK SERC J pour le sud et la zone équatoriale). Dans les versions 1.0, 1.1 et 1.2 de ce catalogue, l'astrométrie est calibrée sur des catalogues antérieurs (AGK3, SAOC, CPC). Les différences de position des sources GSC avec les catalogues de référence utilisés pour le calibrer sont de l'ordre de 0.5–0.9" (Russell et al., 1990). L'erreur externe, qui dépend de la position sur le

ciel (région observée) et sur la plaque photographique (centre ou bord), vaut de 0.2 à 0.8" (par coordonnée).

La version 1.3 du GSC (également appelée GSC-ACT) est une recalibration³ du GSC 1.1 utilisant le catalogue ACT⁴;

- l'USNO A1.0 (Monet et al., 1997) qui contient plus de 488 millions de sources. Ce catalogue résulte de la numérisation de plaques photographiques O et E du POSS I (*Palomar Observatory Sky Survey I*, couvrant l'hémisphère nord jusque -32°), et de plaques SRC-J et ESO-R dans l'hémisphère sud. L'USNO A1.0 utilise le GSC 1.1 comme catalogue de référence pour l'astrométrie ;
- l'USNO A2.0 (Monet et al., 1998), avec 526 millions de sources. La principale différence avec la version 1.0 est que l'astrométrie est calée sur l'ICRF défini par le catalogue ACT, et non plus sur le GSC. Le catalogue établissant l'ICRS dans le domaine optique est pourtant, comme on l'a vu, Hipparcos. Mais les étoiles du catalogue Hipparcos, ainsi que les plus brillantes du catalogue Tycho, sont saturées sur les plaques numérisées dans les catalogues USNO. Les étoiles les plus faibles de Tycho permettent cependant de lier directement la calibration des plaques à l'ICRS, et l'ACT a été préféré au catalogue Tycho en raison de la meilleure précision sur les mouvements propres.

3.3 Les données LDAC

3.3.1 Validation de l'astrométrie

Le but de la calibration astrométrique est de passer des coordonnées des sources en pixels (résultant du processus d'extraction de sources dans les images CCD), aux vraies positions sur le ciel.

Il ne suffit pas de connaître les coordonnées célestes du centre de l'image (direction pointée par le télescope) et la taille des pixels pour déterminer avec précision la position céleste de toutes les sources d'une image, car il y a une incertitude sur le pointage, et des distorsions sont possibles dans une image (et ces distorsions peuvent varier lors d'une séquence d'images : strip ou séquence de calibration).

La calibration astrométrique est une étape essentielle de la chaîne de traitement des données.

Traitement au LDAC

L'astrométrie des sources DENIS est déterminée à partir d'un **catalogue de référence** (CR). Puisque la calibration astrométrique est effectuée pour **chaque**

³Cette recalibration a été faite par le *Pluto Project*, http://www.projectpluto.com/gsc_act.htm.

⁴Le catalogue GSC-II, qui devrait paraître en 2001, sera le fruit de la numérisation de quelque 7000 plaques de Schmidt. Ce projet ambitieux doit conduire à un catalogue contenant tous les objets du ciel jusqu'à la magnitude $V = 18$, avec leurs mouvements propres, et des magnitudes dans au moins deux filtres. Le nombre d'objets attendu dans ce catalogue est de 2 milliards.

image DENIS, la densité de sources dans le CR doit être raisonnablement importante (typiquement, il faut au moins une source du CR par image DENIS). Le relevé complet étant composé de plus de 900 000 images, on ne peut pas utiliser le catalogue Hipparcos comme CR, car on aurait beaucoup trop d'images DENIS ne contenant aucune source Hipparcos.

Les premières versions de la chaîne de traitement de données du LDAC utilisaient comme CR le GSC 1.1. Avec 15 millions d'étoiles différentes environ, ce catalogue permet d'avoir dans la plupart des cas un nombre d'étoiles de référence suffisant dans chaque image⁵.

Depuis le début de l'année 1999, c'est l'ACT qui sert de CR pour le traitement au LDAC. En pratique, on utilise le catalogue USNO A2.0, lui-même calibré sur l'ACT, car il contient une densité de sources bien plus grande.

L'astrométrie est faite en deux temps au LDAC :

- la transformation affine décrivant la déformation des images est déterminée par un mécanisme de triangulation. Une fois les images corrigées de cette transformation, les décalages systématiques entre les différentes bandes sont calculés, et les objets extraits sont appariés entre les différentes bandes (I, J, K_s), dans les zones de recouvrement entre images consécutives d'un strip, et avec le CR ;
- ensuite, la calibration astrométrique résultant de cet appariement est effectuée, et une position est calculée pour chaque source.

Dans les catalogues de strips produits au LDAC, et sur lesquels nous avons effectué nos tests, chaque objet est présent une seule fois. L'appariement entre les détections d'un même objet dans plusieurs bandes est donc déjà effectué, ainsi que l'appariement de deux détections d'un même objet se trouvant sur la zone de recouvrement entre deux images.

Validation

Afin de valider l'astrométrie des sources contenues dans les catalogues produits au LDAC, nous avons effectué des comparaisons systématiques avec les catalogues USNO, qui nous permettent dans la plupart des cas d'avoir une base suffisante d'objets en commun.

La méthode que nous avons employée consiste à prendre toutes les sources USNO et toutes les sources DENIS sur une même région du ciel (par exemple, une image DENIS, $0.2^\circ \times 0.2^\circ$). On peut raisonnablement attendre qu'un certain nombre d'objets soient détectés à la fois par DENIS et par l'USNO (s'ils ne sont pas trop bleus ou trop rouges), avec des positions à peu près inchangées (si on ne considère que les petits mouvements propres) :

⁵Ce n'est cependant pas toujours vrai, surtout dans les zones de forte extinction. On s'attend à avoir, sur environ un million d'images DENIS, 8000 images avec une seule source de référence, et 4400 images sans aucune source de référence (A. Holl, communication privée).

- si l'astrométrie des deux catalogues est bonne, ces sources en commun se trouveront approximativement aux mêmes coordonnées (α, δ) ;
- si il y a un décalage systématique de l'astrométrie entre les deux catalogues $(\Delta\alpha, \Delta\delta)$, on va retrouver les sources communes à des positions (α, δ) et $(\alpha + \Delta\alpha, \delta + \Delta\delta)$.

Regardons pour **une** source USNO, **toutes** les sources DENIS se trouvant à une distance inférieure à un seuil R_c , et traçons tous les vecteurs de composantes $(\lambda_\alpha, \lambda_\delta)$ pointant de la source USNO vers chacune des sources DENIS (Fig. 3.1). On stocke pour chaque source USNO les composantes des vecteurs obtenus.

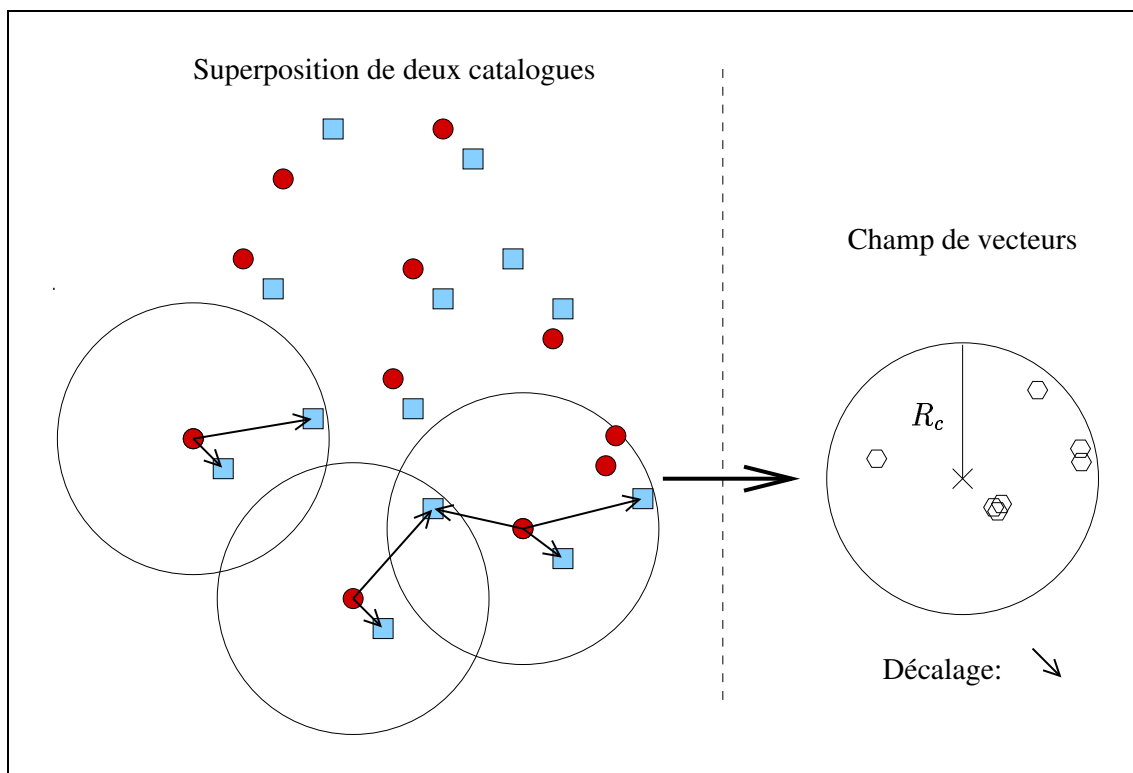


Fig. 3.1: Construction du champ de vecteurs des décalages astrométriques entre les sources de deux catalogues (symbolisées ici par les cercles et les carrés). On trace, pour chaque source du premier catalogue (cercles), des vecteurs vers toutes les sources du second catalogue (carrés) situées dans un rayon R_c . La superposition des vecteurs (dont on a ici indiqué l'extrémité par un hexagone) sur un même graphique fait apparaître une surdensité correspondant au décalage systématique entre les deux catalogues.

Si on répète cette opération pour toutes les sources USNO d'une région, et que l'on visualise la superposition de toutes les composantes des vecteurs, on observe :

- des vecteurs distribués au hasard dans le cercle de rayon R_c , correspondant à des associations aléatoires. Si les deux catalogues sont complètement indépen-

dants, on n'observe que ce bruit de fond uniforme⁶, qui dépend uniquement de la densité de sources dans les catalogues ;

- une surdensité de vecteurs localisée, qui correspond exactement au décalage astrométrique entre les deux catalogues⁷ (Fig. 3.1). S'il n'y a pas de décalage entre les deux catalogues, cette surdensité se trouve exactement au centre du cercle de rayon R_c , puisque les vecteurs associant les sources communes aux deux catalogues ont en moyenne des composantes nulles.

Cette méthode présente l'avantage d'être très simple et très rapide pour déterminer s'il existe un décalage entre deux catalogues. On calcule en effet une seule fois la distance aux sources voisines, alors qu'une méthode fondée sur une descente de gradient, par exemple, nécessiterait ce calcul à chaque itération.

Nous avons produit de façon systématique pour les catalogues du LDAC des graphiques comme celui de la figure 3.2 pour détecter des décalages éventuels de l'astrométrie. Comme les sources sont fusionnées dans les zones de recouvrement entre deux images consécutives, dans les catalogues LDAC, nous avons sélectionné uniquement des régions situées au centre de chaque image, pour produire un graphique par image, et ceci pour chaque strip.

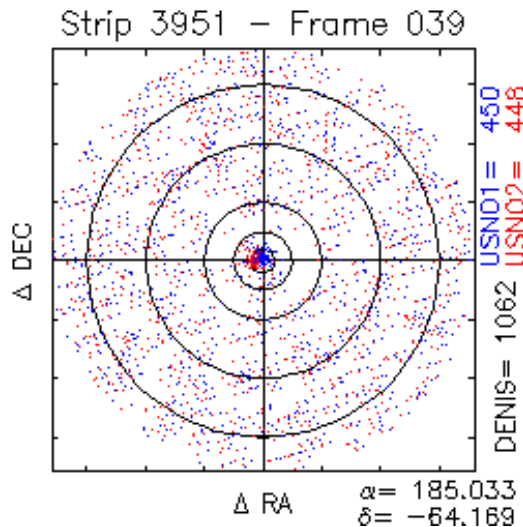


Fig. 3.2: Exemple de graphique produit pour valider l'astrométrie. Sur ce graphique, $R_c = 18''$, les cercles concentriques correspondent à des rayons de 1, 2.5, 5, 10 et 15". On a reporté en haut le numéro du strip et de l'image, en bas les positions du centre de la zone correspondante. Dans la marge de droite figurent les nombres de sources dans les différents catalogues. Chaque point représente l'extrémité d'un vecteur de composantes $x = (\alpha_{USNO} - \alpha_{DENIS}) \times \cos(\delta_{USNO})$, $y = \delta_{USNO} - \delta_{DENIS}$. Les points bleus et rouges correspondent respectivement à des comparaisons avec l'USNO A1.0 et l'USNO A2.0.

Résultats

La figure 3.2 est caractéristique d'une image pour laquelle l'astrométrie est bonne. On peut faire plusieurs remarques sur ce graphique :

- la surdensité de points pour laquelle le décalage des données DENIS est nul correspond à l'USNO A1.0 (points bleus). En effet, ce strip a été traité avec

⁶Ce point est démontré dans l'annexe B.6, page 183.

⁷Pour observer la surdensité, il faut bien sûr que R_c soit supérieur au décalage entre les deux catalogues !

- le GSC 1.1 comme catalogue de référence astrométrique, et on a vu que l'USNO A1.0 est, lui aussi, calibré sur le GSC 1.1 ;
- la dispersion de cette surdensité donne une idée de l'incertitude sur les positions, et sur la qualité de la transformation affine déterminée lors de la première étape de la calibration. Le nuage de points est ici bien concentré, dans le cercle de rayon $1''$. L'erreur sur les positions est typiquement de l'ordre de $0.5''$ pour les sources DENIS ;
 - on voit un léger décalage ($\approx 1''$) avec le catalogue USNO A2.0. Ce décalage est en fait uniquement dû à la différence d'astrométrie entre les deux versions de l'USNO. Dans certains cas, la différence entre les deux peut atteindre quelques secondes d'arc ;
 - le bruit de fond dû aux associations aléatoires de sources est visible sur l'ensemble du graphique. La densité de points, bien plus faible qu'à l'endroit du décalage, dépend uniquement de la densité de sources dans les deux catalogues⁸ (dont on peut se faire une idée à partir des nombres de sources donnés en marge du graphique).

Afin d'éviter un contrôle visuel de tous les graphiques (un par image, on le rappelle), nous avons mis au point une procédure de recherche automatique de la surdensité, consistant simplement à faire passer une fenêtre glissante dans le plan $(\Delta\alpha, \Delta\delta)$.

Cette procédure permet de vérifier que la surdensité se trouve bien au voisinage de $(0,0)$. Il faut noter que dans les cas où la densité d'objets dans les images est très faible (en raison d'une forte extinction, par exemple), il peut être très difficile de trouver une surdensité significative, et dans ce cas le résultat de la procédure de détection automatique est accompagné d'un avertissement.

Pour l'écrasante majorité des images, on obtient des graphiques semblables à la figure 3.2, preuve d'une astrométrie correcte. Cependant, lors de tests systématiques sur un grand nombre de strips, nous avons rencontré quelques images présentant des problèmes pour l'astrométrie :

- il peut y avoir un décalage systématique entre les positions des sources DENIS et celles de l'USNO. Ainsi, sur la figure 3.3-a, la surdensité est décalée d'environ $7.5''$. On peut noter au passage que le bruit de fond est important en raison d'une grande densité de sources dans cette région ;
- sur la figure 3.3-b, on voit deux zones de surdensité pour la même image, l'une bien au centre, et l'autre décalée de $12''$. Une analyse plus détaillée révèle que dans cette image, la calibration astrométrique est bonne pour les bandes I et K_s , mais pas pour la bande J : il y a eu un problème d'appariement des sources détectées en J , qui apparaissent systématiquement décalées.

⁸Voir discussion dans l'annexe B.6, page 183.

⁹On peut noter autour de la surdensité une zone où la densité du bruit de fond est plus faible. Cette caractéristique est expliquée en annexe, page 183.

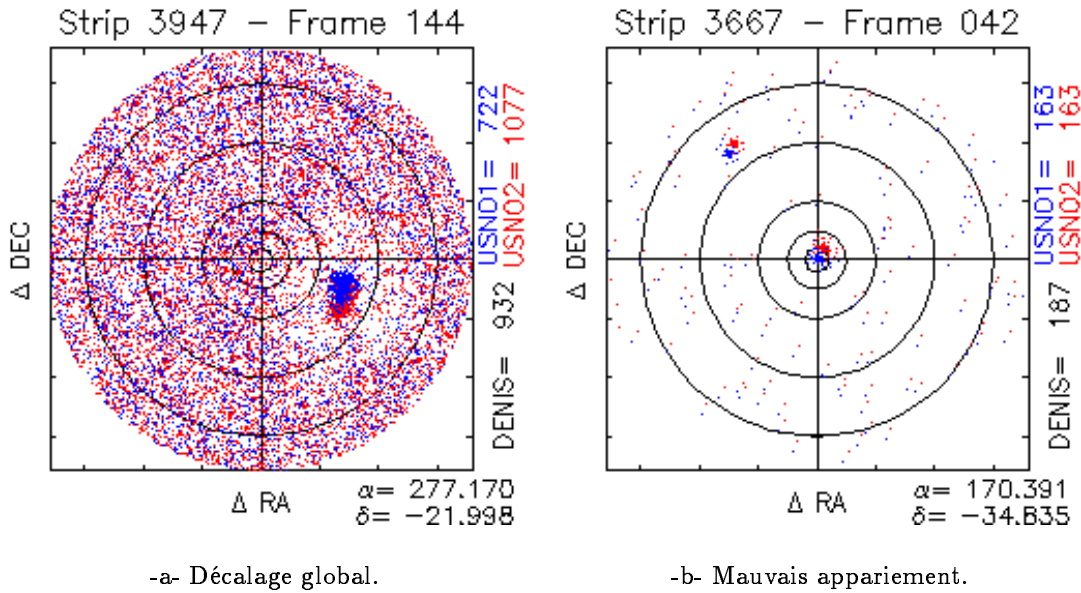


Fig. 3.3: Exemples de mauvaise calibration astrométrique. Dans le graphique de gauche, il y a un décalage systématique de l'astrométrie de 7.5". Notre procédure détecte la présence et la valeur de ce décalage.⁹ — Le graphique de droite est assez singulier car on a deux surdensités. En fait, un problème lors de l'appariement entre les bandes photométriques fait que les détections en J n'ont pas été associées aux autres bandes : les sources en I et/ou K_s ont une astrométrie correcte (surdensité centrale), et les sources en J sont décalées (surdensité à 12").

Ces problèmes de calibration astrométrique soulignent à quel point le traitement est sensible. La technique utilisée au LDAC pour déterminer la transformation affine (décrivant la déformation des images) repose sur une méthode de triangulation : tous les triangles entre les sources détectées dans une image devraient être calculés. Pour des raisons de performances (en temps de calcul), le nombre de sources utilisées reste limité (environ 40, mais c'est un paramètre configurable). Après correction de cette transformation affine, les décalages systématiques entre les différentes bandes sont déterminés.

Dans une zone de relativement forte densité, comme c'est le cas dans la figure 3.3-a, une mauvaise détermination de la transformation affine est peut-être en cause (ce qui expliquerait en outre l'étalement important de la surdensité). On peut supposer que la mauvaise correction du décalage systématique est en partie due à un pointage approximatif du télescope ayant conduit à de mauvaises coordonnées de départ pour la comparaison avec le catalogue de référence.

Il faut garder également à l'esprit que les catalogues de référence ne sont jamais parfaits, et peuvent induire des erreurs de calibration. On trouve par exemple dans le catalogue USNO de véritables « trous » (voir Fig. 3.4). La calibration dans ces zones est donc extrêmement délicate. D'une façon plus générale, les zones de forte

extinction sont difficiles à calibrer, car les objets présents dans les catalogues de référence (du domaine optique, en pratique) y sont sensiblement différents de ceux détectés dans l'infrarouge proche.

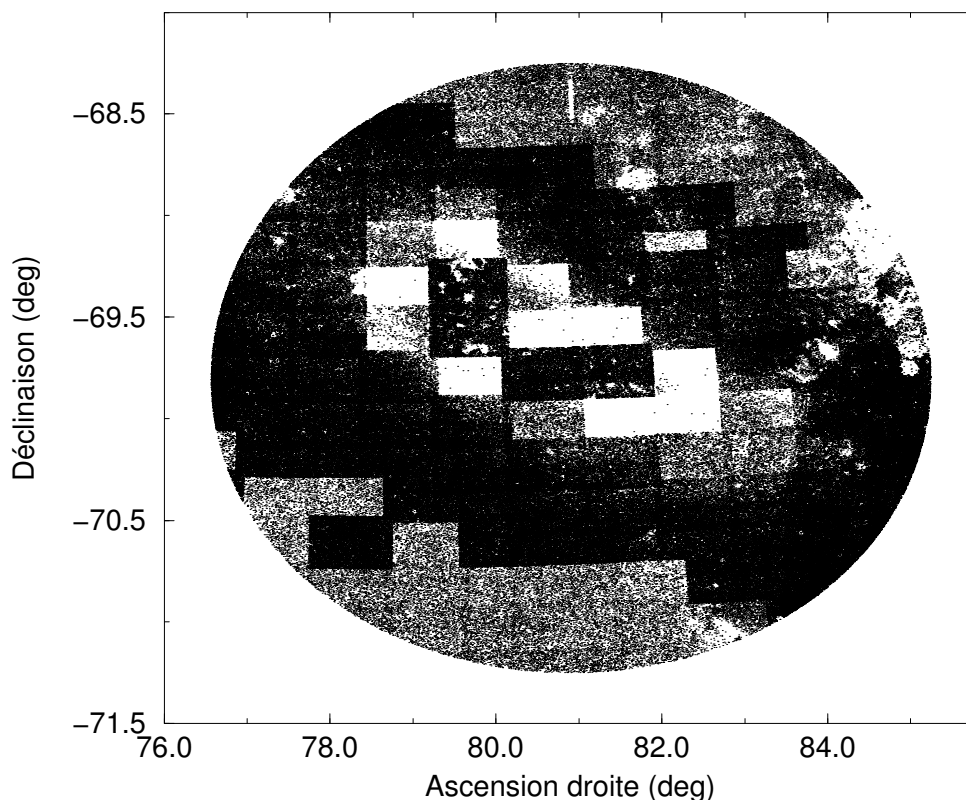


Fig. 3.4: Densité de sources du catalogue USNO A2.0 dans la direction du LMC. Les zones rectangulaires quasiment vides correspondent à la taille de la caméra de la machine (PMM) utilisée pour numériser les plaques photographiques. L'algorithme d'extraction de sources pour la création du catalogue USNO doit déterminer une valeur du fond de ciel sur un certain nombre de pixels, image par image. Dans les zones de très forte densité (comme c'est le cas ici dans la barre du Grand Nuage), on ne peut pas trouver un nombre de pixels de fond suffisant, et l'algorithme échoue (D. Monet, communication privée). Le même problème se pose autour des étoiles brillantes fortement saturées (d'où les zones circulaires apparaissant vides).

Les zones pour lesquelles un problème de calibration astrométrique a été mis en évidence doivent impérativement être retraitées. En effet, un décalage de plusieurs secondes d'arc risque de rendre impossible l'identification croisée des données avec un autre catalogue.

De plus, les décalages mis en évidence dans les données du LDAC sont en général présents dans plusieurs images consécutives, et peuvent varier d'une image à l'autre. Par conséquent, on a parfois des sources qui sont mal appariées sur les zones de

recouvrement, et apparaissent doublées dans le catalogue final (avec des positions erronées).

Cette validation de l'astrométrie du LDAC a été effectuée en particulier avant la publication de la base préliminaire des données DENIS, fin 1998 (Epchtein et al., 1999)¹⁰.

3.3.2 Photométrie et filtrage des parasites

Calibration photométrique

De même qu'on ne mesure pas directement les positions des sources sur le ciel, mais leur position sur le capteur CCD, on n'obtient pas directement avec un capteur numérique les magnitudes des objets, mais une mesure en ADU (*Analog-to-Digital Unit*). La relation de conversion entre les ADU et les magnitudes dépend des caractéristiques de l'instrument et de l'atmosphère, et varie avec le temps.

La calibration photométrique est obtenue pour chaque nuit à partir des observations d'étoiles standard (voir p. 17), objets dont les magnitudes sont bien connues. Chaque étoile standard est observée 8 fois au cours d'une séquence d'observation, mais en moyenne seules 6 mesures sont exploitables pour la calibration. Pour éviter les problèmes de saturation, seules les étoiles plus faibles que $I = 10.5$, $J = 8.0$ et $K_s = 6.5$ sont retenues.

Un coefficient d'extinction fixe est utilisé pour déterminer un point zéro pour la relation de calibration, pour chaque strip (Fouqué et al., 2000)¹¹.

Les points zéro sont donc déterminés de façon indépendante pour chaque nuit. La comparaison d'objets se trouvant sur les zones de recouvrement entre deux strips adjacents indique une précision de l'ordre de 0.05–0.1 magnitudes sur les points zéro photométriques. Dans le catalogue final DENIS, le traitement des zones de recouvrement permettra d'obtenir une photométrie globale de meilleure qualité, en mettant en évidence les variations des points zéro quotidiens.

Graphiques de qualité

Divers graphiques, dont un diagramme couleur-couleur et des histogrammes de comptage différentiel¹², ont été produits de façon systématique pour tous les strips traités au LDAC et archivés sous forme binaire au CDS (voir Fig. 3.5).

Ces diagrammes permettent de mettre en évidence certaines caractéristiques des données. Ainsi, le diagramme de gauche de la figure 3.6 correspond à un strip traversant une zone de forte extinction : on voit un fort rougissement de certains objets. On peut également, grâce à ces graphiques, détecter des erreurs importantes dans la calibration astrométrique. Sur le diagramme de droite de la figure 3.6, on voit

¹⁰Cet article figure en annexe, p. 188

¹¹Cet article figure en annexe, p. 195.

¹²On verra au chapitre 4 les autres graphiques produits.

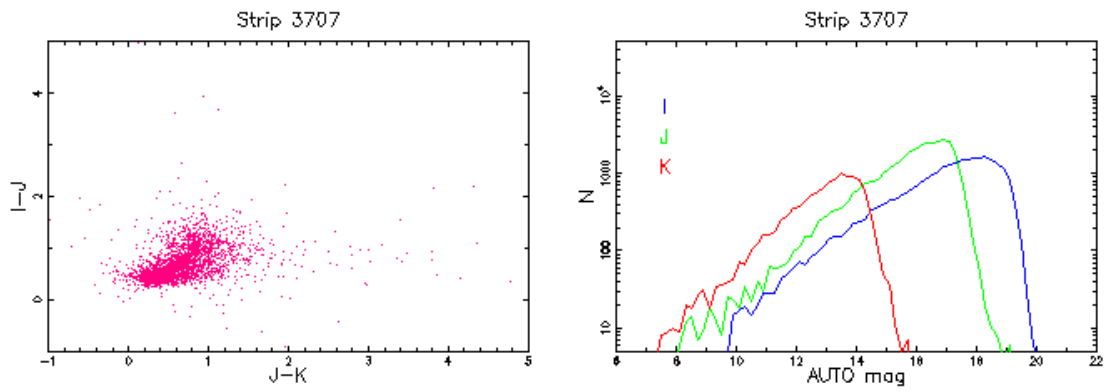


Fig. 3.5: Exemple de graphiques produits pour chaque strip traité au LDAC et archivé au CDS. A gauche, un diagramme couleur-couleur $[I - J], [J - K_s]$. A droite, le comptage différentiel du nombre d'objets dans chaque bande, par pas de 0.2 mag.

l'effet d'une mauvaise détermination du point zéro en K : les objets sont décalés de 2.4 mag vers la droite.

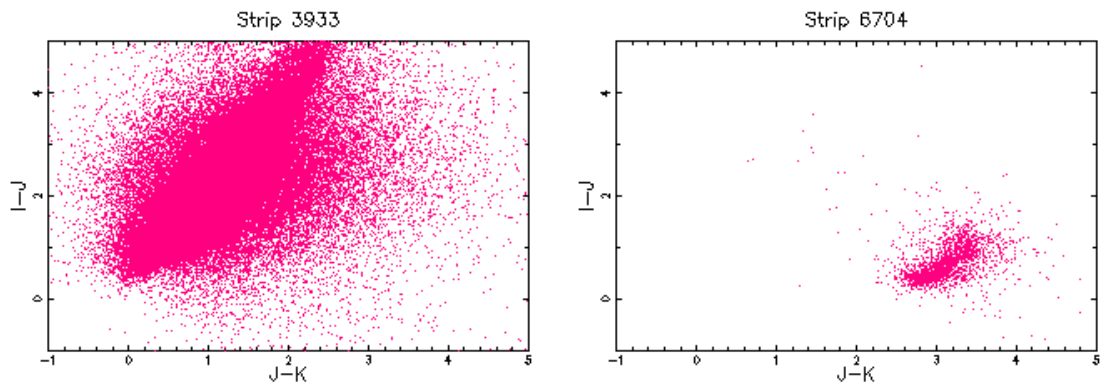


Fig. 3.6: Diagrammes couleur-couleur $[I - J], [J - K_s]$ produits pour les strips traités au LDAC. A gauche, le strip traverse une zone de forte extinction, d'où un grand nombre d'objets fortement rougis. A droite, une mauvaise détermination du point zéro provoque un décalage vers la droite du diagramme.

Artéfacts et parasites

L'extraction de sources permet de trouver les objets réels (étoiles, galaxies), mais elle détecte également un certain nombre de parasites et d'artéfacts, d'origines diverses : rayons cosmiques, pixels défectueux, figures de diffraction dues à l'instrument. . .

Ces mauvaises détections doivent autant que possible être éliminées du catalogue final. Les catalogues de sources ponctuelles produits pour chaque strip au LDAC ont déjà subi un filtrage pour enlever un certain nombre de ces artéfacts. Le logiciel

SExtractor, utilisé pour l'extraction de sources, donne en sortie, pour chaque objet, une indication de la qualité pouvant servir de critère de filtrage. D'autres critères interviennent également pour identifier les artéfacts.

Cependant, une analyse détaillée nous a permis de mettre en évidence un certain nombre d'artéfacts résiduels dans les catalogues produits :

- les *ghosts*. Dans les voies J et K_s de l'instrument DENIS, les étoiles très brillantes provoquent des réflexions parasites internes qui vont se traduire sur le détecteur par un objet fantôme, environ 6 magnitudes plus faible. Les *ghosts* sont localisés sur la cible à la verticale de l'étoile brillante, à une distance d'environ 86 pixels. Les dernières versions de la chaîne de traitement des données au LDAC incluent désormais un filtrage efficace des *ghosts*, ce qui n'était pas le cas pour les premiers catalogues produits ;
- les « triplets de micro-balayage ». On a vu (p. 14) qu'une procédure de micro-balayage permet d'améliorer la précision sur les positions en J et K_s . Toutefois, la présence de pixels défectueux sur les caméras NICMOS peut entraîner l'apparition d'artéfacts à la forme caractéristique : des groupes de 3 sources alignées verticalement, et séparées de $7/3$ de pixel sur les images ré-échantillonnées. Leur disposition caractéristique, et leur présence dans une seule couleur (J ou K_s) rend ces artéfacts facilement identifiables pour le filtrage ;
- des bandes de parasites. Dans quelques images en J ou en K_s , on observe une concentration très importante de pixels brillants sur une bande horizontale. Si ces parasites d'origine électronique sont extraits par SExtractor, ils peuvent être dix fois plus nombreux que les sources véritables dans une image ! Heureusement, leur présence dans une seule bande photométrique, et leur localisation spécifique facilite leur identification ;
- les poussières. La présence de poussières dans le télescope se traduit par des points brillants sur les images, qui peuvent être extraits comme des sources. D'une image à l'autre, les poussières restent présentes, mais leur position peut varier légèrement.

La première diffusion publique de données DENIS a été faite à partir d'une chaîne de traitement (au LDAC) présentant encore quelques imperfections. Le travail de validation a permis de choisir, pour la mise en ligne, les données les moins affectées par les problèmes mis en évidence, mais aussi d'améliorer le traitement des données et de signaler, le cas échéant, les strips devant subir un retraitement.

3.4 Les données PDAC

3.4.1 Astrométrie

Traitement au PDAC

Dans un premier temps, les relations de transformation des positions des objets dans les bandes I et K_s vers les positions dans la bande J sont déterminées. Les

coefficients de ces transformations ne varient pas de façon brusque à l'intérieur d'un strip, ce qui permet de traverser les régions quasiment vides de sources sans perdre le contact.

Ensuite, la bande I est ajustée sur le catalogue USNO A2.0 (qui est lui-même calibré sur l'ACT, comme on l'a vu), puis la transformation de la bande I vers la bande J est inversée, ce qui permet de calibrer les positions en J . La relation de transformation $K_s \mapsto J$ permet de calibrer les positions en K_s .

L'évolution des relations de transformation le long du strip est obtenue par une régression polynomiale du second ordre. Un contrôle est effectué pour vérifier qu'une solution correcte est obtenue. Si la procédure de calibration échoue (ceci peut se produire s'il manque des images dans le strip, ou si les conditions météorologiques étaient mauvaises pendant les observations), le strip n'est pas inclus dans la base de données.

Validation

Nous avons utilisé, pour valider la calibration astrométrique effectuée au PDAC, la technique déjà présentée dans la section 3.3.1 pour le LDAC.

Des tests systématiques effectués sur quelques strips traités au PDAC n'ont mis en évidence aucune image pour laquelle il y aurait un problème de calibration de l'astrométrie. Ceci confirme que les tests effectués au PDAC sur la qualité de la calibration astrométrique sont efficaces.

3.4.2 Photométrie et filtrage des parasites

Extraction des sources et calibration

Le traitement des images au PDAC a été décrit par Borsenberger (1997). Pour détecter les sources, tous les ensembles de pixels connexes ayant un niveau supérieur à un seuil ajustable (typiquement, 2σ au-dessus du bruit de fond) sont identifiés.

Le traitement photométrique repose sur une estimation de la PSF, qui varie à l'intérieur d'une image (à cause de l'optique de l'instrument), et d'une image à l'autre (à cause des variations atmosphériques).

La forme de la PSF est déduite du profil d'étoiles brillantes. Il faut donc, pour déterminer la forme de la PSF en divers points des images, un nombre suffisant d'étoiles : si ce nombre n'est pas disponible dans une seule image, les étoiles de plusieurs images consécutives sont utilisées.

Le modèle de PSF est une combinaison de 3 gaussiennes Γ_i de largeur fixe :

$$\Gamma_i = (A + BX + CY + DXY + EX^2 + FY^2) \exp \left[-\omega_i (X^2 + Y^2) \right] \quad (3.1)$$

où, si l'on note (x_0, y_0) la position dans l'image du centroïde sur lequel on centre la PSF, et x et y les coordonnées d'un point quelconque, on a $X = x - x_0$, $Y = y - y_0$,

et les coefficients A, B, C, \dots sont donnés par des polynômes de la forme :

$$\begin{aligned} A &= a_0 + a_1x + a_2y + \dots \\ B &= b_0 + b_1x + \dots \end{aligned}$$

Les coefficients $a_0, a_1, \dots, b_0, \dots$ sont déterminés par une méthode d'ajustement des moindres carrés décrite par Alard (2000).

Pour chaque source, une première valeur du flux est obtenue en calculant le produit scalaire de la PSF locale avec les pixels identifiés comme constituant la source, en ignorant les pixels défectueux, et en renormalisant. Six autres valeurs de flux sont également déterminées : les sommes des pixels dans des ouvertures circulaires de 6.5, 5.5 et 3.5" autour du pixel central, et les sommes des pixels dans des ouvertures de 3.5, 2.5 et 1.5", renormalisées par le modèle de PSF .

Ces 7 flux sont stockés pour chaque source. Les magnitudes, elles, sont calculées lors des requêtes, en utilisant les valeurs des calibrations associées. Ainsi, il n'est pas nécessaire d'effectuer à nouveau l'extraction des sources si les paramètres de calibration photométrique d'un strip doivent être modifiés.

Base de travail et catalogues

Il convient de bien distinguer la base de travail du PDAC des catalogues de sources ponctuelles.

La base de travail contient absolument toutes les extractions, quelle que soit leur pertinence (objets réels ou artéfacts, voir figure 3.7). Aucun filtrage n'a été appliqué, afin de ne perdre aucune information. De plus, chaque objet apparaît autant de fois qu'il est détecté (les différentes mesures d'un même objet sont toutes conservées, elles ne sont pas fusionnées).

La base contient, d'une part, des paramètres caractérisant chaque source, et d'autre part, à un niveau plus élevé, des valeurs pour calibrer les données (pour calculer les magnitudes à partir des flux, par exemple). A l'époque où j'ai été amené à manipuler les données du PDAC, un certain nombre de paramètres étaient encore indisponibles dans la base de données. Il s'agit essentiellement de *flags*, indiquant par exemple si une source est saturée, s'il s'agit d'un ghost, d'une poussière, si la source contient des pixels défectueux. . .

Pour produire des catalogues utilisables, il est nécessaire de filtrer au mieux les données contenues dans la base de travail. Ce filtrage nous permettra de produire les catalogues que nous allons utiliser au chapitre 7 pour l'étude de la structure galactique.

Critères de filtrage

En l'absence des *flags* identifiant certains artéfacts, nous avons dû définir d'autres critères de sélection. Pour chaque détection d'un objet dans une des 3 bandes photométriques, on connaît la position en pixels (x_M, y_M) de l'objet dans l'image corres-

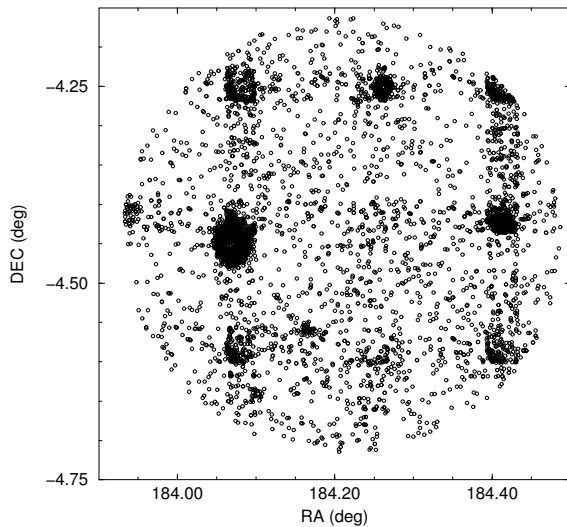


Fig. 3.7: Exemple de contenu de la base de travail du PDAC : sur cette petite zone du ciel, chaque symbole représente un enregistrement présent dans la base. On remarque des surdensités de détections de forme circulaire, correspondant à des artefacts dans les halos présents autour des étoiles brillantes en I . On note également des densités plus importantes dans les zones de recouvrement entre les strips adjacents, ainsi qu'entre les images consécutives, où les sources sont détectées plusieurs fois. Les bords des images sont aussi des régions où on trouve plus d'artefacts.

pondante. On dispose également de la valeur de la corrélation du profil de la source au modèle de PSF (valeur exprimée entre 0 et 100).

Ce coefficient de corrélation a tendance à diminuer pour les objets de magnitudes plus faibles, mais il permet surtout de mettre en évidence les parasites : par exemple, un rayon cosmique, même brillant, aura un profil très différent de la PSF, et donc un mauvais coefficient de corrélation.

Un autre critère utile pour le filtrage est le nombre de bandes dans lesquelles une source est détectée. Un objet détecté simultanément en I , J et K_s est a priori plus fiable qu'un objet détecté dans une seule bande.

Il faut garder à l'esprit qu'un relevé du ciel traverse des zones aux caractéristiques fort différentes : par exemple, à basse latitude galactique, dans les régions de forte extinction, on aura beaucoup plus d'objets non-détectés en I que vers le pôle galactique. Les critères de filtrage que nous voulons définir doivent être applicables à l'ensemble des données.

Pour définir des critères de filtrage efficaces, j'ai sélectionné, parmi les données disponibles à l'époque au PDAC, trois zones de test situées dans des directions galactiques correspondant à des densités de sources très différentes (voir Tab. 3.1).

La zone B est située à haute latitude, la zone C se trouve à proximité du centre galactique, et la zone A se situe dans le plan, mais loin du centre. Chaque zone couvre plusieurs strips en ascension droite.

Pour chacune de ces zones, j'ai produit des séries de graphiques similaires à ceux représentés dans la figure 3.8.

Ces graphiques permettent de visualiser la répartition des objets détectés dans les images, en fonction du coefficient de corrélation à la PSF et du nombre de bandes

Zone	l ($^{\circ}$)	b ($^{\circ}$)	R ($''$)	Nombre d'objets
A	250.3	+00.0	3600	176 297
B	286.8	+57.3	5000	84 351
C	5.3	+00.0	2500	211 797

Tab. 3.1: Zones de test pour le filtrage des données issues du PDAC. On a indiqué les coordonnées galactiques du centre de chaque zone circulaire, le rayon dans lequel on a extrait les objets, et le nombre d'objets présents dans la base de travail PDAC dans la zone.

dans lesquelles chaque objet est détecté¹³.

Voyons ce que contiennent ces graphiques, par exemple dans le cas de la figure 3.8-a :

- chaque **vignette** représente la densité de sources en fonction de la position en pixels dans l'image, cumulée pour toutes les images de la zone de test A, avec un pas d'intégration de 20×20 pixels ;
- chaque **rangée** de vignettes correspond à des sources détectées dans une certaine combinaison des bandes photométriques, et détectées uniquement dans ces bandes : 1^{re} rangée pour I seulement, 2^e pour I et K_s , 3^e pour I et J , et dernière rangée pour I , J et K_s ;
- chaque **colonne** correspond à un certain intervalle du coefficient de corrélation à la PSF (indiqué sous chaque vignette) ;
- toutes les vignettes de cette figure concernent la bande I : toutes les sources sont donc détectées au moins en I , les positions en pixels dans l'image sont les positions dans l'image I , et c'est le coefficient de corrélation à la PSF pour la magnitude I qui est utilisé.

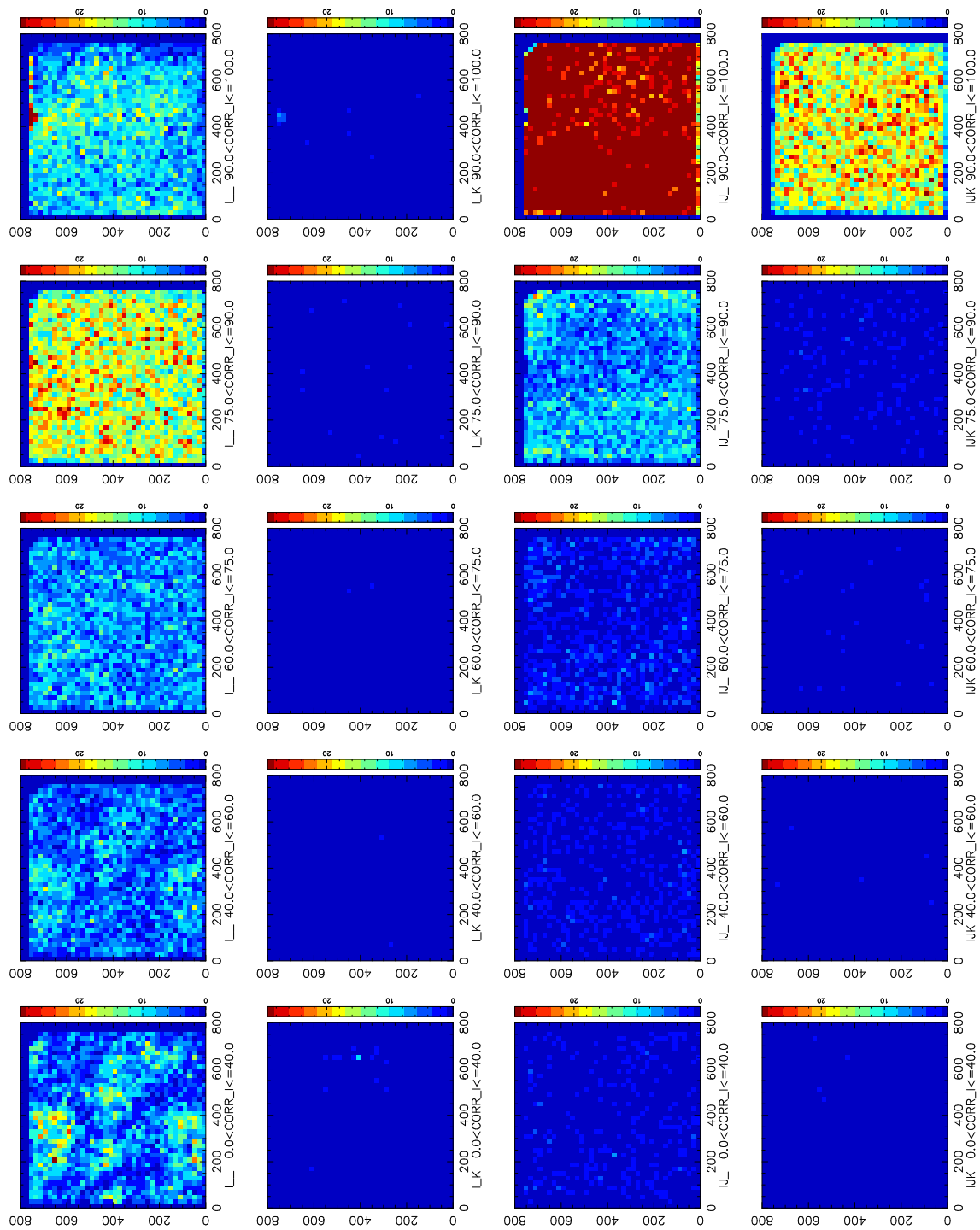
On peut noter que la taille des vignettes est de 800×800 pixels, alors que les images DENIS font 768×768 pixels. Le fait que l'extrême bord des images (où la qualité du capteur est moins bonne) n'est pas traité au PDAC explique aussi qu'il y ait une zone de faible densité en bordure des vignettes.

Si seuls des objets réels, répartis sur le ciel de façon homogène, étaient présents dans les données, ces objets devraient se répartir de façon uniforme dans les images, et donc la densité cumulée sur un grand nombre d'image devrait être uniforme, d'où une densité homogène dans chaque vignette.

A l'inverse, si l'on observe des structures, des inhomogénéités dans les vignettes, elles sont la signature de la présence de détections localisées préférentiellement à une certaine position dans les images, qui sont plus vraisemblablement des artéfacts que de vrais objets.

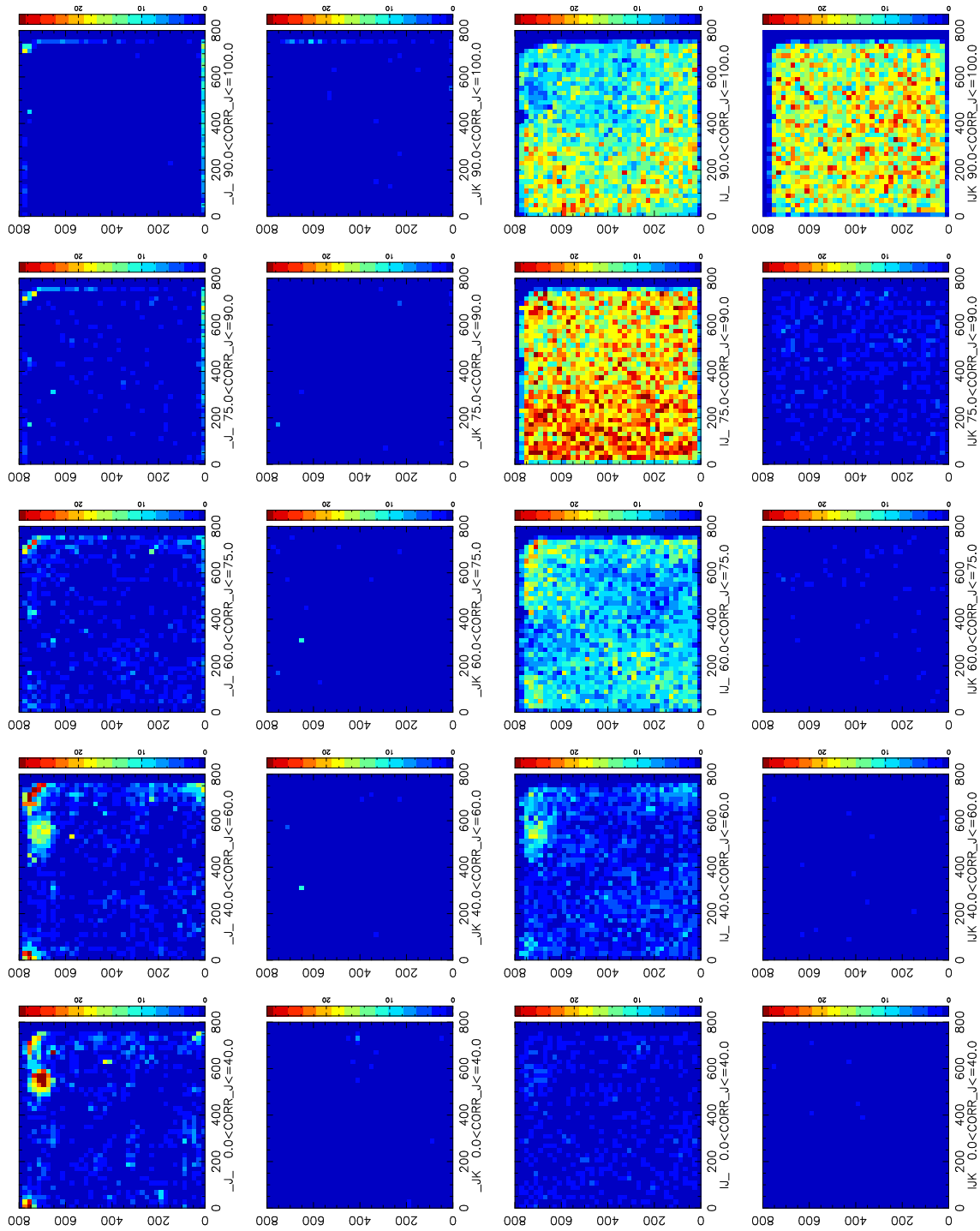
Sur l'ensemble des figures 3.8-a à 3.8-i, correspondant aux 3 zones de référence et aux 3 bandes photométriques, on note :

¹³Le seul critère de détection retenu pour un objet dans une bande est la présence de la magnitude obtenue à partir du modèle de PSF.



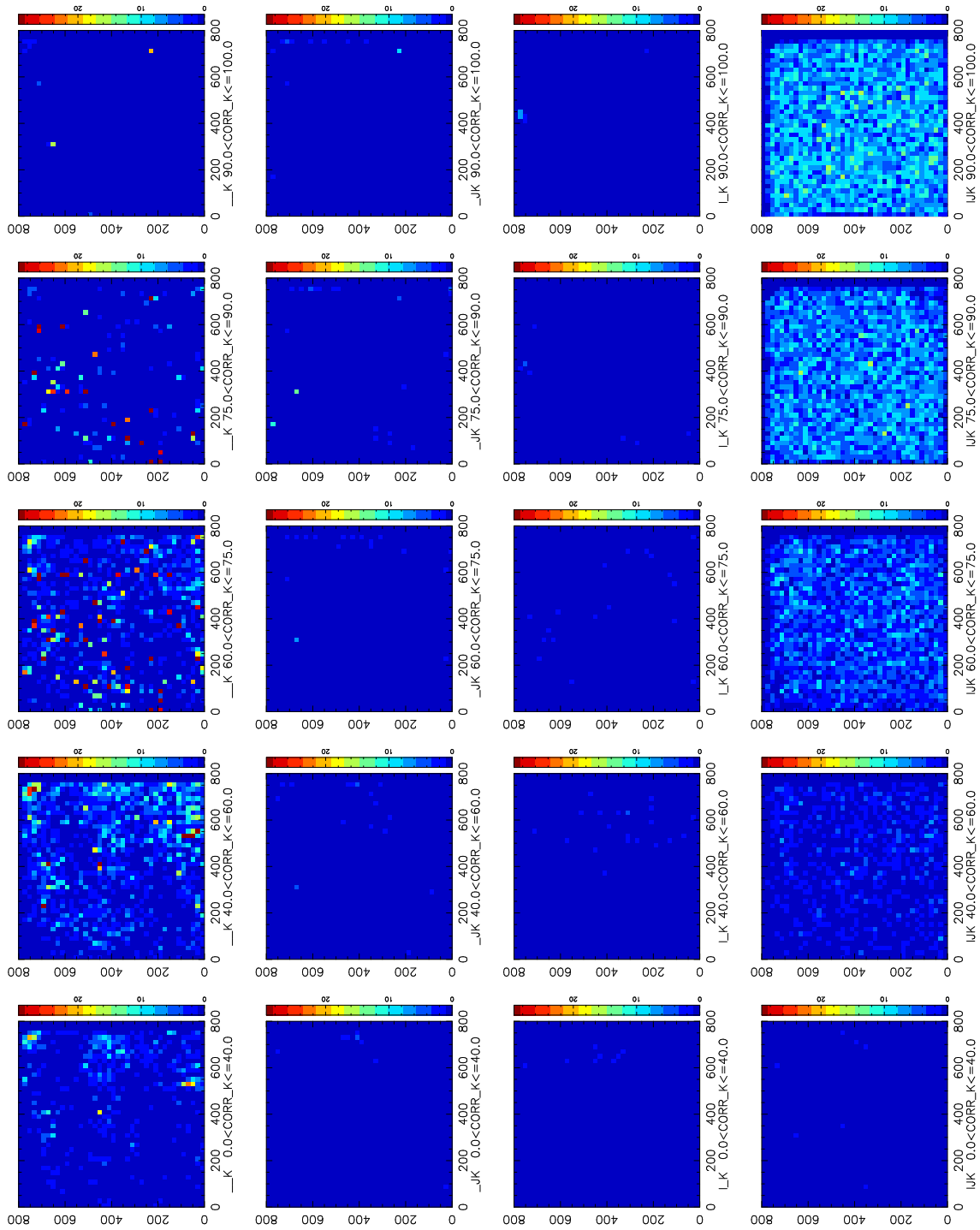
-a- Zone de test A, bande I

Fig. 3.8: Graphiques utilisés pour déterminer les critères de filtrage des données PDAC (voir p. 66 pour les explications).



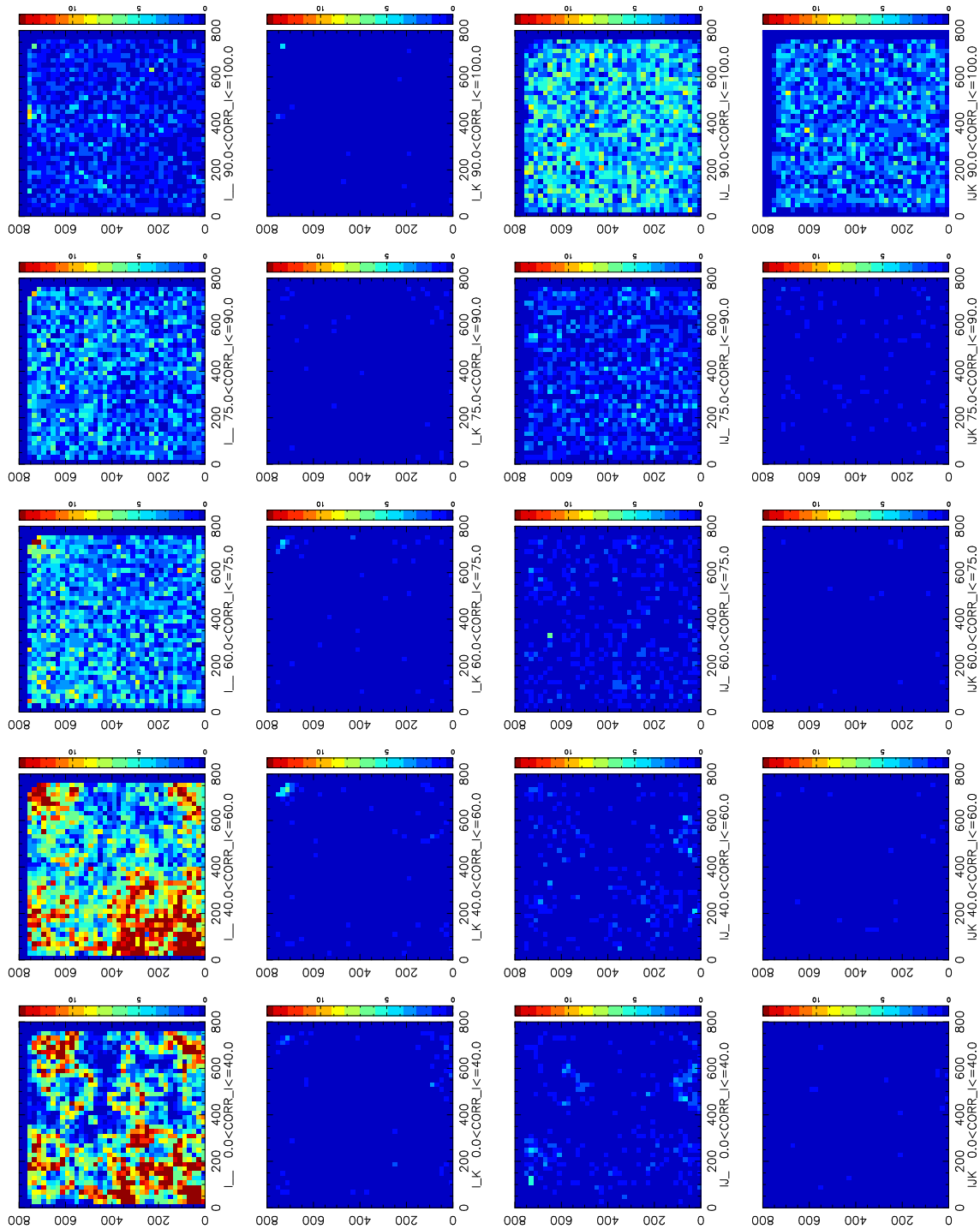
-b- Zone de test A, bande J

Fig. 3.8:(suite).



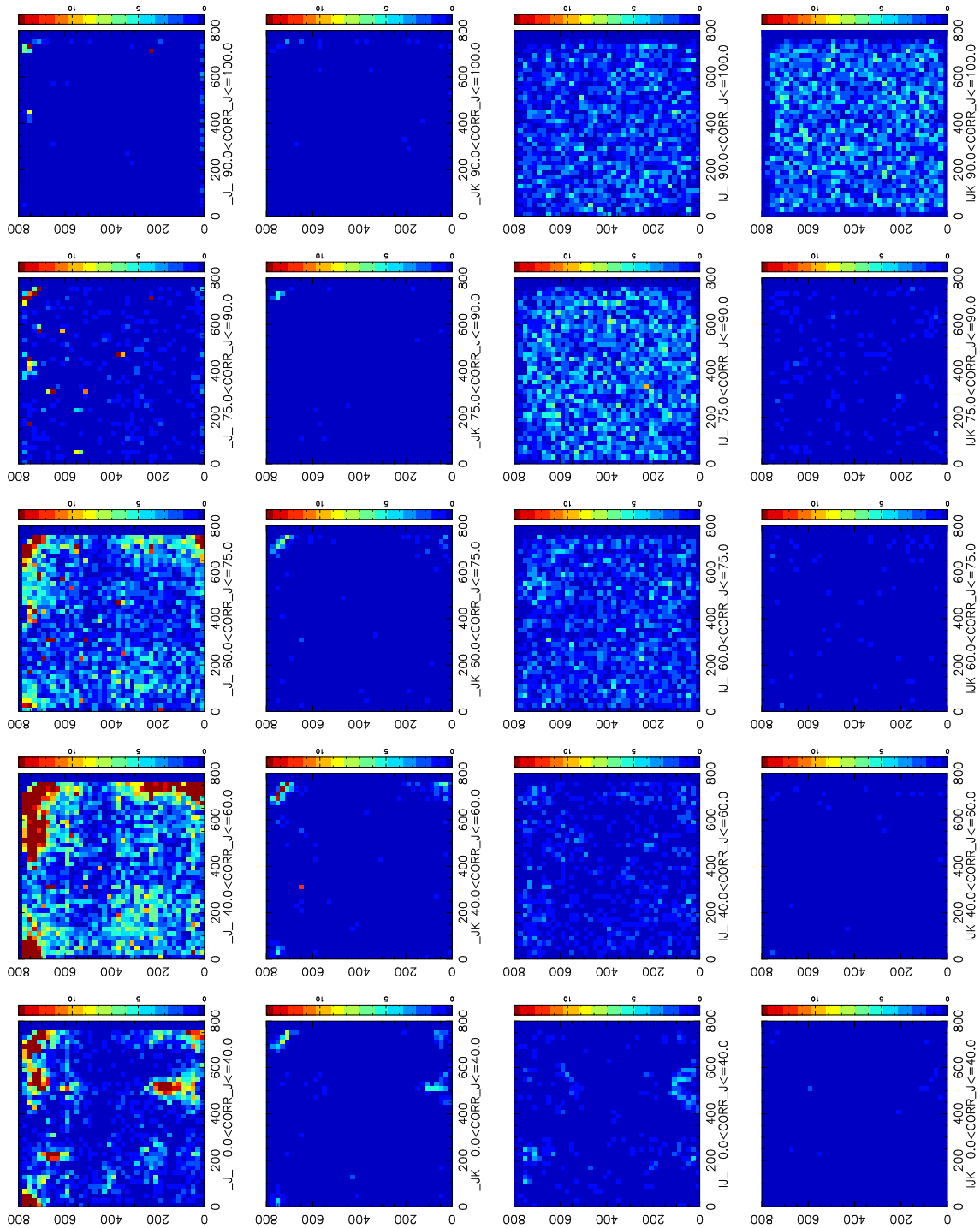
-c- Zone de test A, bande K_s

Fig. 3.8:(suite).



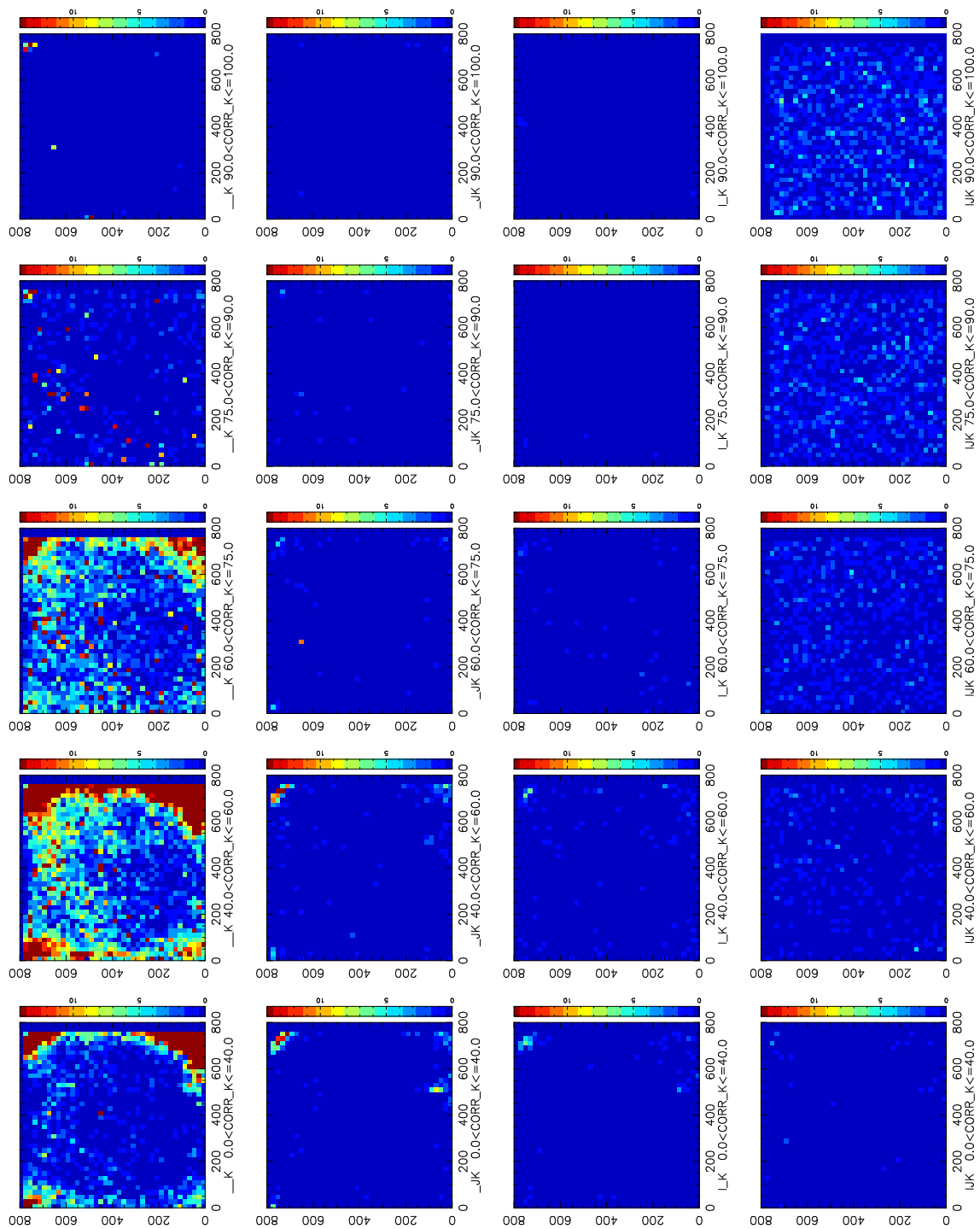
-d- Zone de test B, bande I

Fig. 3.8:(suite).



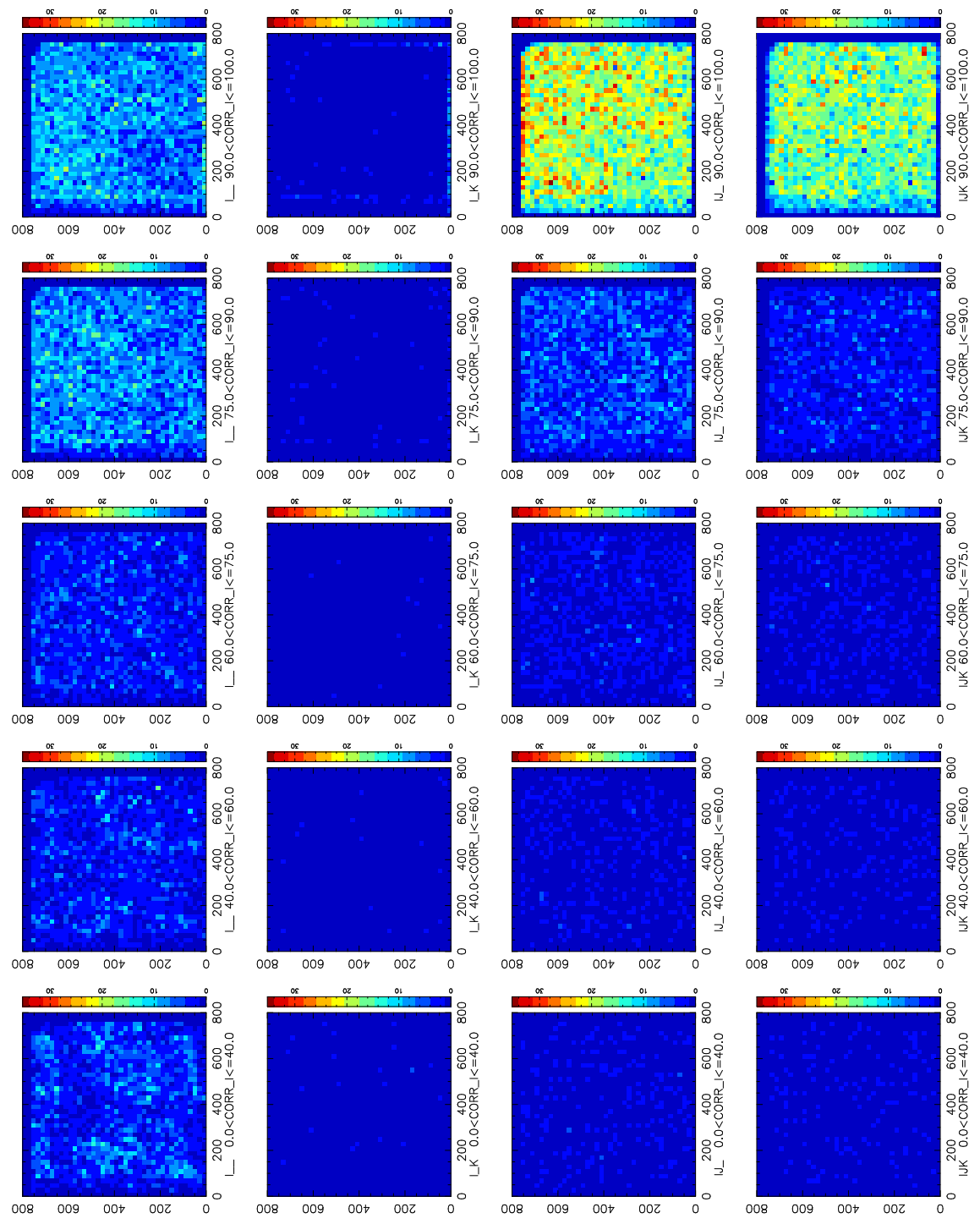
-e- Zone de test B, bande J

Fig. 3.8:(suite).



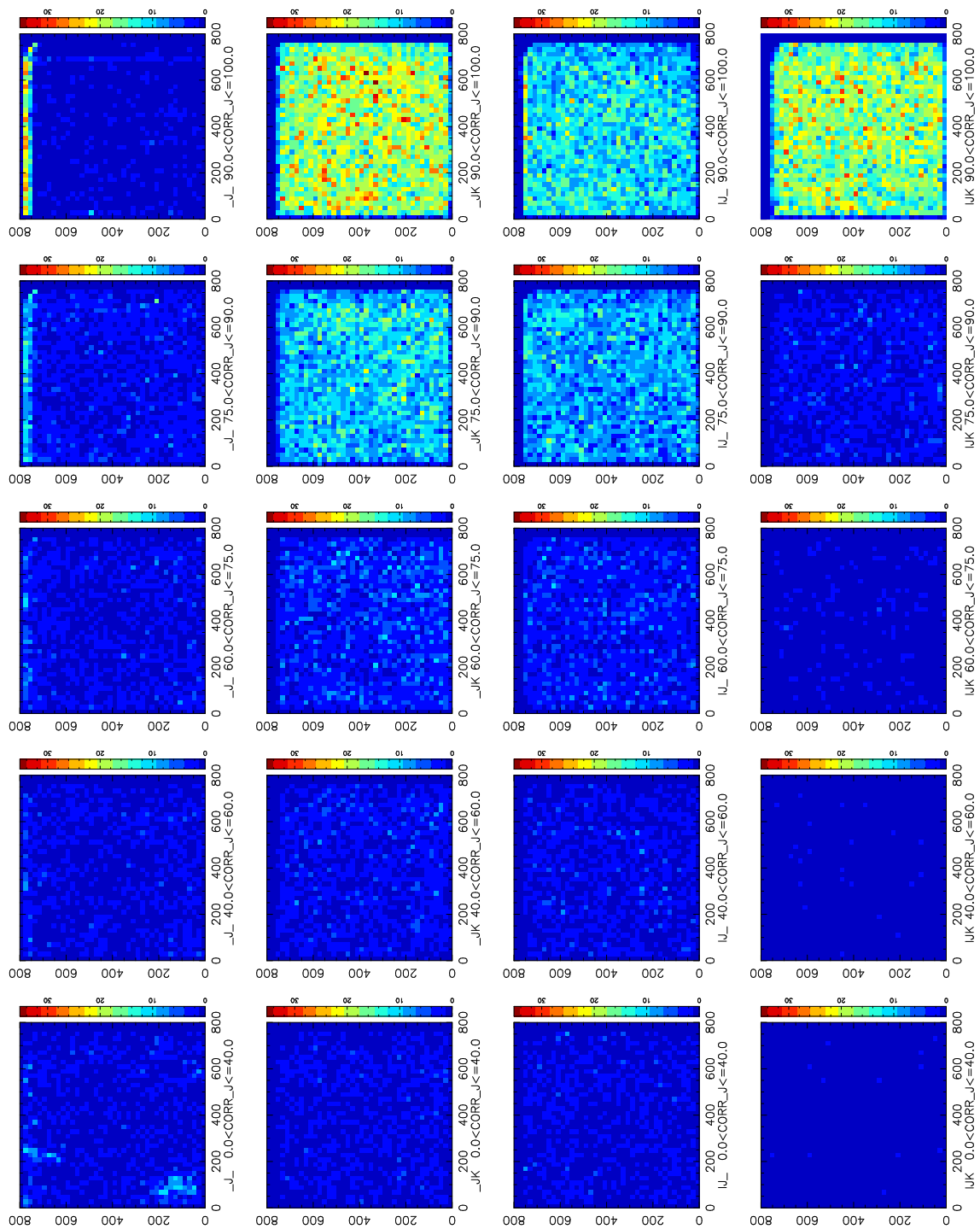
-f- Zone de test B, bande K_s

Fig. 3.8:(suite).



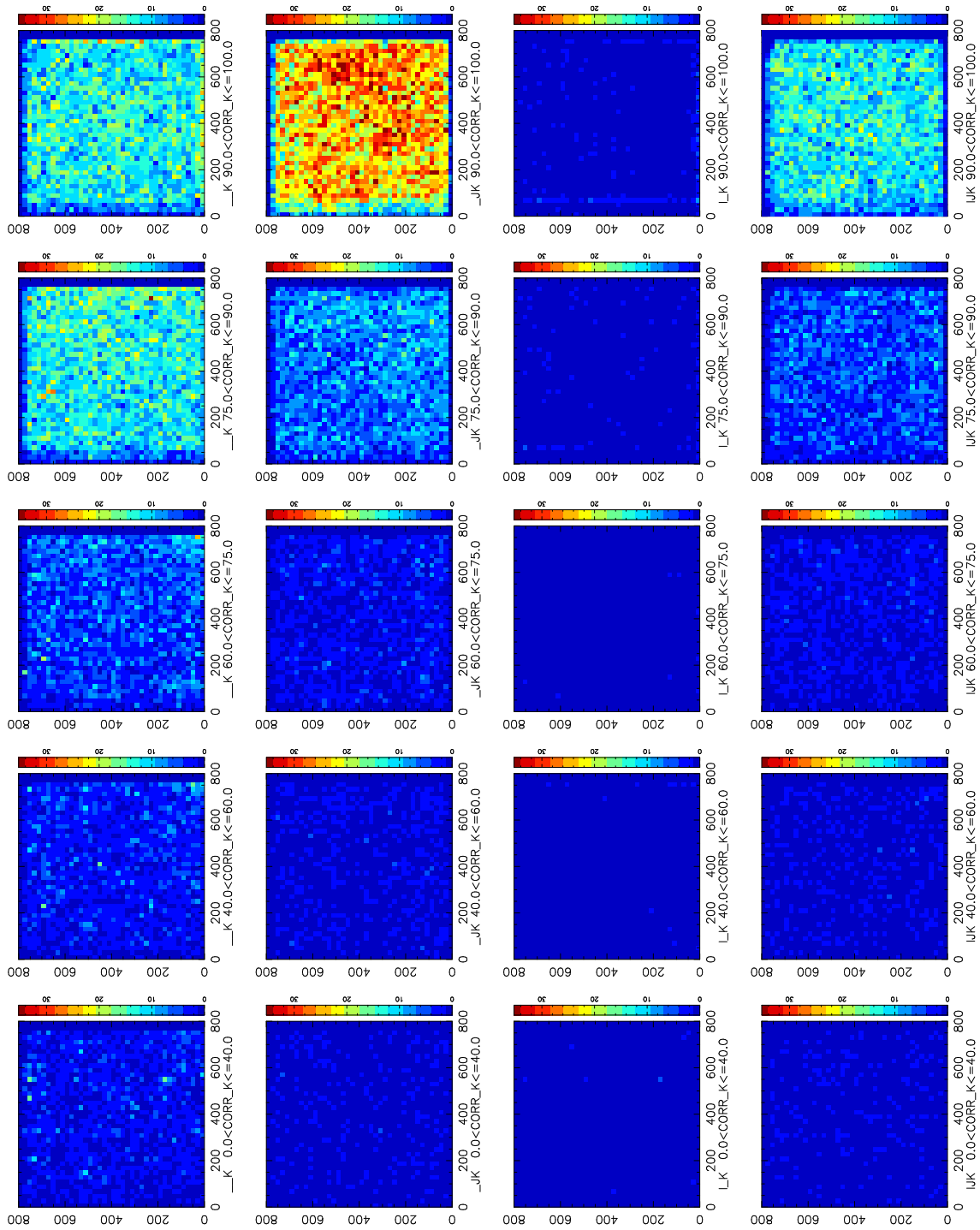
-g- Zone de test C, bande I

Fig. 3.8:(suite).



-h- Zone de test C, bande J

Fig. 3.8:(suite).



-i- Zone de test C, bande K_s

Fig. 3.8:(suite et fin).

- pour les sources détectées en I , J , et K_s , quelle que soit la zone ou le coefficient de corrélation, la densité est relativement homogène ;
- les objets détectés seulement en J sont localisés dans des structures le long des bords d'images : ce sont vraisemblablement majoritairement des artéfacts ;
- dans les zones A et B, on a peu de sources détectées seulement en J et K_s , et ce sont vraisemblablement des parasites sur les bords ou les coins des images. Par contre, dans la zone C, on a beaucoup d'objets réels détectés dans cette combinaison de bandes, et avec de bons coefficients de corrélation (voir Fig. 3.8-h et 3.8-i). En effet, de nombreuses sources ne sont pas détectées en I à cause de la forte extinction vers le centre galactique dans la zone C.
- de même, les objets détectés en K_s seulement sont presque uniquement des parasites dans les zones A et B (Fig. 3.8-c et 3.8-f), mais il peut s'agir de vrais objets, avec un bon coefficient de corrélation, dans la zone C (Fig. 3.8-i) ;
- la densité cumulée d'objets détectés seulement en I n'est homogène que pour des coefficients de corrélation élevés. Pour les mauvais coefficients de corrélation, les structures irrégulières sont dues à la superposition de sources détectées dans plusieurs images, dans les halos entourant les étoiles brillantes ;
- de même, la répartition des objets détectés en I et J n'est homogène que pour les bons coefficients de corrélation à la PSF.

Après avoir produit des graphiques comme ceux de la figure 3.8, en faisant varier les seuils de coefficient de corrélation, j'ai défini les critères de filtrage suivants :

- conserver toutes les sources détectées concomitamment en I , J et K_s ;
- rejeter toutes les sources détectées uniquement en I et K_s , ou seulement en J ;
- conserver les sources détectées en :
 - I et J si $\text{CORR}_I > 75$ et $\text{CORR}_J > 65$;
 - J et K_s si $\text{CORR}_J > 75$ et $\text{CORR}_{K_s} > 75$;
 - I si $\text{CORR}_I > 65$;
 - K_s si $\text{CORR}_{K_s} > 75$.

Appariement

Une fois le filtrage des artéfacts effectué, reste le problème des détections multiples des sources se trouvant dans les régions de recouvrement entre images et strips adjacents.

Deux sources sont appariées si :

- elles ne sont pas extraites de la même image ;

- elles se trouvent à moins de 2'' l'une de l'autre¹⁴.

L'appariement est fait de la manière suivante :

- à l'intérieur d'un strip :
 - si une source est détectée dans plus de bandes que l'autre, on rejette celle qui a le moins de détections ;
 - sinon, on fusionne les deux mesures (en prenant la moyenne des magnitudes) ;
- entre deux strips, seule la source la plus éloignée d'un bord d'image est conservée (car les bords peuvent encore être contaminés par des artefacts, malgré le filtrage).

Résultats

Ce sont ces critères que nous avons utilisés pour produire les catalogues employés au chapitre 7, pour l'étude de la structure galactique.

On peut voir sur la figure 3.9 le résultat du filtrage et de l'appariement sur l'échantillon de données qu'on avait présenté dans la figure 3.7.

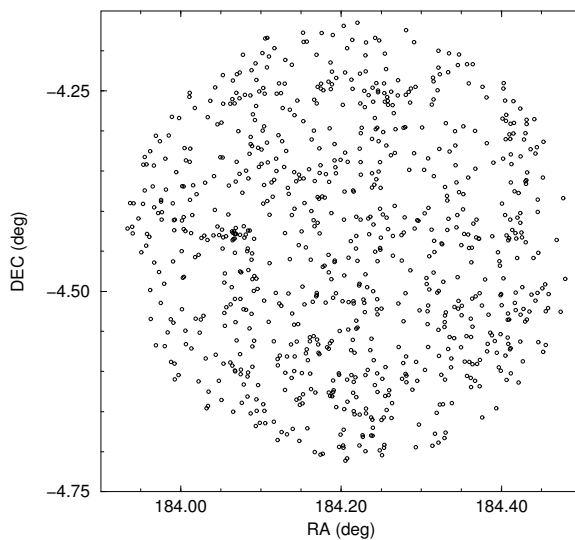


Fig. 3.9: Exemple de filtrage et d'appariement sur les données présentées dans la figure 3.7, page 65. On voit que la répartition des sources est beaucoup plus homogène. La quasi-totalité des artefacts dans les halos ont disparu, et on ne distingue plus les zones de recouvrement entre les images ou entre les strips.

Lorsque les *flags* associés aux sources seront disponibles, il sera possible d'améliorer sensiblement les critères de filtrage. Ceci permettra d'effectuer une diffusion importante de données DENIS issues du PDAC à l'ensemble de la communauté astronomique, vraisemblablement en 2001.

¹⁴Ce rayon permet de séparer de façon non ambiguë les vraies associations des associations aléatoires (la composante Poissonienne de l'annexe B).

3.5 Comparaisons LDAC–PDAC

En principe, les deux centres de traitement de données DENIS ne traitent pas les mêmes strips. Au PDAC, où le travail d'extraction des sources n'a démarré qu'en 1999, les strips traités sont ceux observés après février 1999. Au LDAC, ce sont les strips observés au début du projet (1996, 1997) qui sont traités en priorité.

Un petit nombre de strips identiques (correspondant à une nuit d'observation) ont cependant été traités indépendamment par les deux centres, afin de pouvoir comparer les performances respectives des deux chaînes de traitement.

Je présente dans la suite de cette section la comparaison des deux réductions sur le strip 3705. J'ai extrait toutes les sources contenues dans le catalogue du LDAC pour ce strip. D'autre part, après avoir extrait de la base de travail du PDAC toutes les détections dans ce strip, j'ai filtré et apparié ces données, selon les critères définis dans la section 3.4.2. Les résultats obtenus sur ce strip sont représentatifs de ceux obtenus sur l'ensemble des strips pour lesquels j'ai effectué des comparaisons.

3.5.1 Astrométrie

L'astrométrie des sources n'est pas directement comparable, puisque le strip a été réduit au LDAC avec le GSC 1.1 comme catalogue astrométrique de référence, alors qu'au PDAC, c'est l'USNO A2.0 qui a servi à la calibration.

On peut voir sur la figure 3.10 les différences de position entre les sources communes aux deux catalogues. Le critère d'association retenu pour la constitution d'un échantillon de sources communes est la présence de deux sources séparées de moins de 2" dans les deux catalogues.

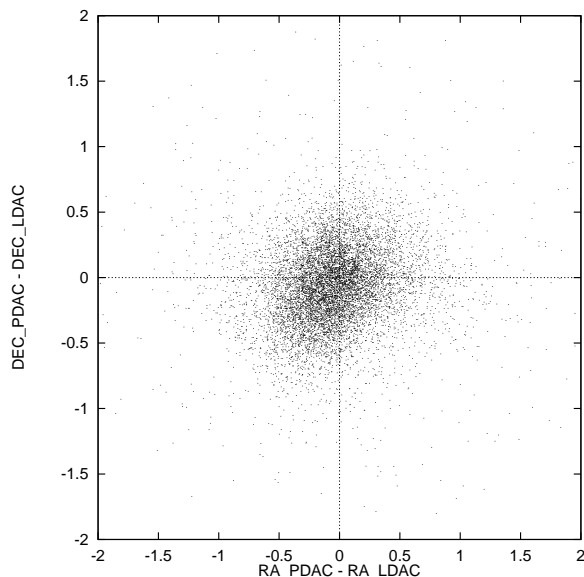


Fig. 3.10: Différences de position entre les sources communes identifiées dans les catalogues LDAC et PDAC, pour les premières images du strip 3705. La dispersion sur les valeurs est inférieure à 0.5". La figure n'est pas parfaitement symétrique car les décalages systématiques varient déjà sur ces quelques images.

Les différences systématiques de position entre le GSC 1.1 et l'USNO A2.0 vont entraîner entre les positions des sources au PDAC et au LDAC des décalages qui

varient le long d'un strip. Ces décalages pouvant être supérieurs à la dispersion des différences de positions entre les deux catalogues DENIS, la superposition d'un strip complet sur la figure 3.10 aurait nui à la lisibilité de la figure, c'est pourquoi on n'a retenu qu'une partie du strip.

On peut voir que la dispersion moyenne entre les positions dans les deux catalogues est inférieure à $0.5''$, ce qui donne une indication de la précision interne de l'astrométrie pour les sources DENIS.

3.5.2 Photométrie

La figure 3.11 présente l'allure des comptages différentiels d'objets dans les 3 bandes, pour le strip 3705, pour chacune des deux chaînes de traitement.

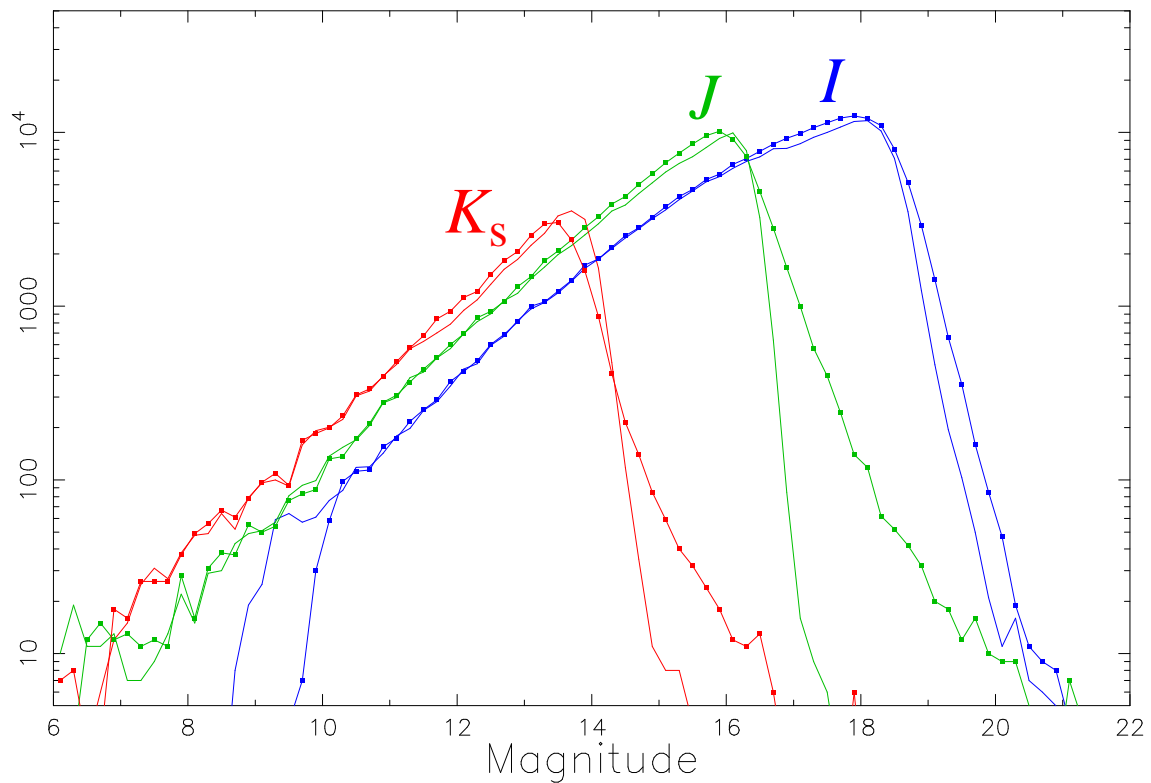


Fig. 3.11: Diagrammes de complétude pour le strip 3705, avec la réduction PDAC (magnitude déduite de l'ajustement au modèle de PSF, traits simples) et LDAC (magnitude pour une photométrie d'ouverture de $7''$, traits avec symboles). Le pas utilisé pour le comptage différentiel est de 0.2 magnitudes.

On remarque que des objets brillants en I sont extraits au PDAC, alors qu'ils ne figurent pas dans le catalogue du LDAC. Le traitement des objets saturés¹⁵ en I apparaît donc plus performant côté PDAC.

¹⁵On rappelle que la magnitude de saturation en I est de 10.0.

Le catalogue PDAC apparaît, en outre, plus profond dans la bande K_s que le catalogue LDAC.

Au-delà des limites de complétude en J et K_s , on note que la queue de distribution des objets faibles est beaucoup plus réduite pour le catalogue PDAC que pour le catalogue LDAC. Une analyse de ces sources faibles indique d'il s'agit majoritairement de parasites détectés seulement dans une bande photométrique, et mal filtrés au LDAC.

L'étude des différences de magnitude entre objets identiques dans les deux catalogues n'apporte pas beaucoup d'information supplémentaire. On note simplement que la dispersion moyenne entre les valeurs des magnitudes augmente pour les objets plus faibles, selon une loi concordant parfaitement avec les erreurs photométriques (similaires dans les deux catalogues).

Aucune différence systématique n'a été mise en évidence.

On peut cependant noter qu'un nombre plus important de sources au PDAC sont détectées dans un nombre de bandes plus important. On a ainsi plus de sources détectées dans les 3 bandes au PDAC qu'au LDAC. Ceci est essentiellement dû à la différence de limite de complétude en K_s .

Sur un diagramme couleur-magnitude (voir Fig. 3.12), on peut faire quelques remarques supplémentaires.

On retrouve, pour les objets brillants, l'absence de magnitude I mesurée au LDAC, alors que cette magnitude est présente au PDAC.

Pour les objets les plus faibles, on voit une différence notable, le catalogue PDAC étant nettement plus profond que le LDAC, qui semble tronqué au moins 0.5 magnitudes plus tôt, et semble souffrir d'un biais pour les magnitudes les plus faibles.

3.5.3 Conclusion

Vu les différences entre les catalogues du PDAC et du LDAC pour un même strip, j'ai choisi de n'utiliser pour l'étude de la structure galactique (Chapitre 7) que les données provenant du PDAC.

Ce choix a été essentiellement motivé par les meilleures performances du PDAC pour les objets faibles en K_s , car cette bande est de première importance pour l'étude des zones où l'on a une forte extinction interstellaire.

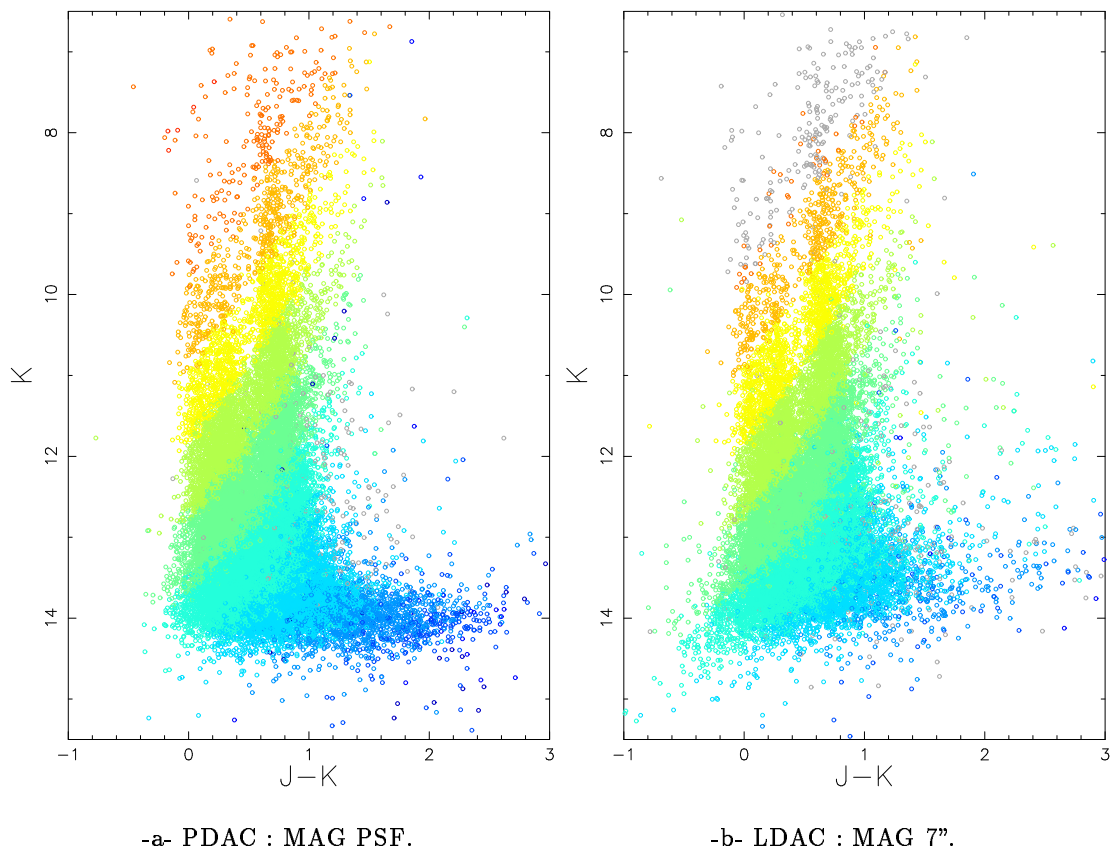


Fig. 3.12: Diagrammes couleur-magnitude pour les réductions PDAC (magnitudes dérivées de l'ajustement de PSF, à gauche) et LDAC (magnitudes pour une photométrie d'ouverture de 7", à droite) du strip 3705. Chaque symbole correspond à une détection dans un catalogue. Le code couleur adopté indique la magnitude I des objets, du rouge (objets les plus brillants en I) au bleu foncé (objets les plus faibles). Les symboles gris traduisent une non-détection en I .

Chapitre 4

Accès en ligne

4.1 L'astronomie en ligne

4.1.1 Vers le « tout-numérique »

L'ASTRONOMIE est en passe de devenir la première science « tout-numérique ». Ainsi, les détecteurs numériques (CCD), employés aussi bien pour les projets spatiaux que pour les télescopes au sol, produisent des flux de données dont le traitement est entièrement informatisé. L'essor des réseaux informatiques a considérablement facilité les échanges de données entre astronomes. On peut même aujourd'hui piloter à distance certains télescopes pour effectuer des observations : plusieurs télescopes situés à La Silla (Chili) peuvent être commandés depuis une station de contrôle située à Garching (Allemagne). Enfin, les articles publiés dans le domaine de l'astronomie sont également, dans leur immense majorité, intégralement consultables en ligne.

L'essor du Web, avec l'arrivée des premiers navigateurs hypertextes comme Mosaic en 1993, a énormément fait pour le développement massif de l'accès aux données en ligne. Auparavant, un certain nombre de données pouvaient déjà être récupérées par ftp, et quelques centres dans le monde offraient la possibilité de télécharger des catalogues.

L'aspect plus convivial et la souplesse d'utilisation du Web ont offert de nouvelles possibilités pour l'échange et la diffusion d'information en ligne. Avec l'ouverture au grand public d'Internet, la capacité des réseaux s'est accrue très rapidement, rendant possibles des échanges encore plus importants.

Les astronomes ont très vite perçu l'impact que ces nouvelles technologies pourraient avoir sur leur discipline. L'évolution des possibilités techniques a créé de nouveaux besoins, pour lesquels de nouveaux services ont dû être inventés. Face à la quantité sans cesse croissante des ressources astronomiques disponibles sur Internet, la mise au point de méthodes de recherche d'informations pertinentes, et de partage de données aux formats hétérogènes est plus que jamais d'actualité (Egret et al., 2000).

Quelques grands centres dans le monde sont spécialisés dans l'archivage et la diffusion de données astronomiques : images, catalogues, bibliographie. Les interactions possibles entre les différents types de données sont nombreuses.

Ainsi, la publication électronique des articles est couplée à l'accès en ligne aux catalogues : dès 1993, les tables de données des articles publiés dans le journal *Astronomy and Astrophysics* ont été traitées comme des catalogues de données accessibles en ligne (Ochsenbein & Lequeux, 1995).

On peut, à partir d'un objet trouvé dans un catalogue, rechercher automatiquement la liste des articles où cet objet est cité, et consulter ces articles en ligne. De là, on pourra interroger les bases de données observationnelles, pour récupérer des images ou des données spectroscopiques sur l'objet en question, rechercher dans divers catalogues les mesures disponibles, et finalement soumettre son propre article en ligne !

4.1.2 Le rôle du CDS

Le Centre de Données astronomiques de Strasbourg¹ (CDS), créé en 1972, est devenu un centre de référence pour la collecte, la gestion et la diffusion de données dans le domaine de l'astronomie (Genova et al. (1996; 1998; 2000)).

Le développement des liens entre les différents services du CDS sera une étape essentielle dans la construction d'un « observatoire virtuel » pour le XXI^e siècle.

Je vais passer en revue rapidement les principaux services offerts par le CDS.

SIMBAD

SIMBAD^{2,3} est la base de données de référence pour les identifications et la bibliographie d'objets astronomiques (Wenger et al., 2000).

En 2000, SIMBAD contient plus de 7.5 millions d'identificateurs pour environ 2.7 millions d'objets différents. Pour chaque objet figurent dans la base quelques mesures (position, magnitudes dans différents domaines de longueur d'onde), ainsi que les références bibliographiques où l'objet est cité (plus de 110 000 articles sont concernés).

Cet ensemble de données résulte d'un long travail d'identification croisée entre de nombreux catalogues, listes d'objets et articles de journaux, entrepris au début des années 1980, et constamment développé et mis à jour depuis.

VizieR

VizieR⁴ est une base de données rassemblant plusieurs milliers de catalogues astronomiques sous un format homogène.

¹<http://cdsweb.u-strasbg.fr/>

²*Set of Identifications, Measurements and Bibliography for Astronomical Data.*

³<http://simbad.u-strasbg.fr/Simbad>

⁴<http://vizier.u-strasbg.fr/>

Une description standardisée du contenu des catalogues permet leur inclusion dans un système de gestion de base de données (SGBD) relationnel. Un ensemble de liens, entre les tables de VizieR, et avec des services externes (bibliographiques, archives externes, serveurs d'images), permettent de naviguer entre les données des catalogues et d'autres données associées.

Il faut noter que les très grands catalogues (plus de 10^7 enregistrements) ne peuvent pas être gérés par un SGBD relationnel. Des outils spécifiques doivent être utilisés, comme on le verra dans la section 4.2.

ALADIN

ALADIN⁵ est un atlas interactif du ciel permettant d'accéder simultanément à des images numérisées du ciel, ainsi qu'à des catalogues et bases de données astronomiques (Bonnarel et al., 2000).

Cet outil permet de superposer, sur des images du ciel optique, les objets présents dans SIMBAD, des sources de catalogues contenus dans VizieR, mais aussi d'autres données, locales ou situées sur des serveurs distants (archive HST, ...).

Autres services

Parmi les autres services offerts par le CDS, on peut citer le dictionnaire de nomenclature (Lortet et al., 1994), des services bibliographiques, des services de pages jaunes (AstroWeb (Jackson et al., 1994), Star*s Family (Heck, 1997)), AstroGLU (Egret et al., 1998b).

Le CDS est également impliqué dans le développement d'outils et de méthodes d'intérêt général pour l'échange de données (comme le GLU, Fernique et al. (1998)) et la standardisation (avec l'introduction du *bibcode*, en collaboration avec NED⁶, Schmitz et al. (1995)).

Parmi les projets actuellement en développement, figurent un système de gestion des très grands catalogues reposant sur une base de données orientée objet, la mise au point de stratégies d'indexation multicritères des catalogues, et l'étude d'un système pour automatiser l'identification croisée de catalogues (Egret et al., 1998a; Egret et al., 1999)⁷.

4.2 Les catalogues DENIS et 2MASS

4.2.1 Accès au CDS⁸

Les catalogues DENIS et 2MASS font partie des très grands catalogues, comme les catalogues USNO ou GSC, et nécessitent un traitement spécial : ils sont trop

⁵<http://aladin.u-strasbg.fr/>

⁶NED : *NASA/IPAC Extragalactic Database*, <http://nedwww.ipac.caltech.edu/>

⁷Ces deux papiers figurent en annexe, p. 200 et 202

⁸Pour plus d'informations, on pourra consulter la communication au colloque ADASS IX (Derriere et al., 2000) reproduite p. 206.

volumineux pour être intégrés dans un SGBD relationnel.

Performances des requêtes

Nous avons vu au chapitre 2 comment ces deux catalogues ont été compressés, en conservant une indexation sur les positions. Pour chaque catalogue, j'ai mis au point un programme effectuant, pour une requête en position, la recherche, le décodage et l'affichage du résultat. Les performances pour l'exécution des requêtes sont données dans la table 4.1. On voit que dans le cas de 2MASS, les 162 millions d'objets disponibles peuvent être testés en moins d'une heure.

Catalogue	Temps/source (μ s)
DENIS (LDAC)	7.5
2MASS (IDR2)	21.0

Tab. 4.1: Performances des accès aux catalogues DENIS et 2MASS. Les temps indiqués correspondent à la durée moyenne de traitement d'une source lors d'une requête par position dans le catalogue. Les tests ont été effectués sur la machine vizier équipée d'un quadriprocesseur 400 MHz. La différence entre les deux catalogues vient principalement de la taille moyenne des enregistrements, plus petite pour DENIS.

Intégration aux services du CDS

La sortie de ces programmes de requête est standardisée, et complètement décrite par un fichier ReadMe, comme n'importe quel catalogue dans Vizier.

La totalité des données publiques de DENIS et de 2MASS est accessible depuis l'interface de Vizier⁹ : lors d'une requête (qui peut être multicritère), les contraintes sur la position sont transmises au programme de recherche et de décodage, qui va renvoyer un résultat pouvant être traité par Vizier (par exemple en ajoutant une sélection sur les magnitudes). On peut voir un exemple de résultat pour le catalogue DENIS sur la figure 4.1.

Puisque les catalogues DENIS et 2MASS sont accessibles dans Vizier, il sont utilisables dans l'interface d'ALADIN comme n'importe quel autre catalogue. La figure 4.2 présente un exemple d'utilisation pour les données DENIS.

Données accessibles

En décembre 2000, les données accessibles étaient composées de :

- pour 2MASS, l'IDR2¹⁰ (47% du ciel, plus de $162 \cdot 10^6$ sources), qui remplace l'IDR1 et le *2MASS sampler* ;

⁹Pour DENIS : <http://vizier.u-strasbg.fr/cgi-bin/Vizier?-source=B/denis> ; pour 2MASS : <http://vizier.u-strasbg.fr/cgi-bin/Vizier?-source=B/2mass>

¹⁰<http://www.ipac.caltech.edu/2mass/releases/second/doc/explsups.html>



VizieR Result Page



[CDS](#) · [Simbad](#) · [VizieR](#) · [Aladin](#) · [All Catalogues](#) · [Nomenclature](#) · [Biblio](#) · [VizieR-Help](#)

Result of VizieR Search within 10arcmin of 129.0-60 (J2000)
ordered by increasing *r*

B/denis/denis (First 20 million sources of DENIS Survey)

To get all columns for a row, just click on the row number in the leftmost 'Full' column.
The 4 columns in *italics* are computed by VizieR, and are *not part of the original data*.

Note: See also the [DENIS home page](#)

Full	<i>r</i>	RAJ2000	DEJ2000	RAJ2000	DEJ2000	ErrPos	Automag_I	e_Automag_I	Automag_J	e_Automag_J	Automag_K	e_Automag_K
	arcmin	"h:m:s"	"d:m:s"	deg	deg	deg	mag	mag	mag	mag	mag	mag
14	4.1376	08 36 31.58	-60 01 24.5	129.131602	-60.020696	0.000061			16.521		0.226	
15	4.1552	08 36 33.25	-60 00 09.6	129.138544	-60.002670	0.000037	17.749	0.186				
16	4.1937	08 36 33.31	-60 00 26.1	129.138786	-60.007243	0.000043	17.537	0.138				
17	4.2862	08 36 33.71	-59 59 13.1	129.140449	-59.986969	0.000059	18.138	0.204				
18	4.3084	08 36 34.27	-59 59 32.4	129.142774	-59.992325	0.000018	15.543	0.032	14.945		0.081	
19	4.4124	08 36 35.28	-59 59 52.5	129.147018	-59.997928	0.000038	17.436	0.151				
20	4.4529	08 36 35.14	-60 00 44.3	129.146406	-60.012299	0.000012	14.714	0.013	14.034		0.049	
21	4.4600	08 36 35.67	-60 00 07.3	129.148516	-60.002014	0.000010	13.670	0.008	12.774		0.013	12.153
22	4.4867	08 36 35.17	-59 59 06.4	129.146530	-59.985113	0.000023	16.952	0.083	15.472		0.113	
23	4.4897	08 36 32.75	-59 58 09.7	129.136464	-59.969350	0.000039	17.318	0.179	16.139		0.190	
24	4.4938	08 36 35.04	-60 01 00.7	129.145384	-60.016867	0.000039	17.727	0.172	16.763		0.238	
25	4.5282	08 36 33.37	-59 58 16.0	129.139879	-59.971672	0.000046	19.180	0.198				
26	4.5650	08 36 31.69	-60 02 16.3	129.132050	-60.037870	0.000052	18.237	0.205				
27	4.5967	08 36 36.28	-59 59 15.5	129.151187	-59.987635	0.000029	16.870	0.080				
28	4.6183	08 36 33.41	-59 58 02.0	129.139215	-59.967217	0.000047	17.876	0.185				
29	4.6196	08 36 33.65	-60 01 54.8	129.140222	-60.031892	0.000045	17.935	0.223				
30	4.6369	08 36 37.06	-59 59 48.0	129.154411	-59.996553	0.000053	18.508	0.233				
31	4.6551	08 36 36.50	-59 59 05.0	129.152100	-59.984727	0.000020	15.400	0.029	14.313		0.047	13.394
32	4.6671	08 36 36.06	-59 58 47.8	129.150260	-59.979384	0.000040	18.090	0.186	17.012		0.276	
33	4.6931	08 36 37.29	-59 59 27.8	129.155390	-59.991055	0.000044	18.121	0.175				
34	4.6940	08 36 34.58	-59 58 10.4	129.144063	-59.969543	0.000041	18.000	0.166				
35	4.7051	08 36 35.22	-59 58 20.5	129.148738	-59.972348	0.000030	17.020	0.106				
36	4.7335	08 36 36.57	-60 01 14.1	129.152365	-60.020588	0.000025	16.349	0.053	15.711		0.188	0.191

Fig. 4.1: Exemple du résultat d'une requête sur le catalogue DENIS dans VizieR.

- pour DENIS, 102 strips issus du LDAC (plus de $17 \cdot 10^6$ sources, Epchtein et al. (1999));
- toujours pour DENIS, plus de 400 strips supplémentaires du LDAC (soit un total de $86 \cdot 10^6$ sources), avec un accès restreint par mot de passe.

Les données traitées au PDAC sont en cours d'intégration.

4.2.2 Les pages Web DENIS

Pour le projet DENIS, des pages spéciales ont été créées¹¹ pour fournir des informations complémentaires sur le projet et la qualité des données (suite au travail de validation présenté au chapitre 3).

On trouvera sur ces pages :

- un suivi de l'avancement du projet ;
- une interface simple permettant d'interroger le catalogue de sources ponctuelles par position, ou par nom d'objet (on utilise une fonctionnalité de SIMBAD pour convertir le nom d'objet en une position sur le ciel) ;
- la possibilité de faire une recherche dans un strip avec des contraintes sur les couleurs ;

¹¹<http://cdsweb.u-strasbg.fr/denis-public/>

- l'ensemble des graphiques de qualité pour les strips, qui peuvent être consultés de différentes manières :
 - en effectuant une recherche sur une position ou un nom d'objet ;
 - par numéro de strip ou de slot ;
 - par une carte du ciel cliquable, comme celle présentée dans la figure 2 de Egret et al. (1999) (p. 205).
 - ou directement depuis le résultat d'une requête, en cliquant sur le numéro de strip.

La figure 4.3 présente un exemple des données fournies pour chaque strip. On y trouve :

- le numéro du strip ;
- le nombre total de sources ;
- le format du strip et le nombre de bandes (seuls les strips où les observations ont été faites dans les 3 couleurs sont mis dans la base) ;
- le numéro de slot et les coordonnées correspondantes ;
- les dates d'observation et de réduction ;
- les points zéro de la calibration photométrique ;
- un diagramme couleur-couleur ;
- les comptages différentiels d'objets dans les 3 bandes (le fichier de données correspondant est également accessible) ;
- le nombre d'objets en fonction de la déclinaison (le long du strip) ;
- le positionnement du strip par rapport au slot ;
- trois graphiques (un par bande) permettant de mettre en évidence la présence d'images fortement parasitées (c'est le cas ici dans la bande J vers -1.5°).

4.3 Interfaces dédiées

Toujours dans le cadre du projet DENIS, et avec les données réduites au LDAC, j'ai installé pour certains membres du consortium des interfaces dédiées à la recherche d'objets particuliers.

Naines brunes

La recherche de naines brunes est l'un des points forts parmi les retombées scientifiques de DENIS. La sélection de candidates naines brunes (ou étoiles de très faible masse) se fait principalement sur la base des couleurs de ces objets.

En adoptant les critères définis dans sa thèse par Delfosse (1997) pour la sélection d'objets candidats, j'ai mis au point une interface pour filtrer les strips DENIS.

Pour la liste d'objets sélectionnés, des liens permettent de rechercher directement les contreparties optiques (USNO), ou d'accéder aux images correspondantes (sur le serveur de l'IAP).

Etoiles AGB

Sous la responsabilité de Marie-Odile Mennessier, une base de données d'objets évolués (géantes rouges, AGB et post-AGB) a été développée à Montpellier : ASTRID¹².

Là encore, une sélection en couleurs permet de sélectionner dans les données DENIS des objets pouvant appartenir à ces classes. Ainsi, certains objets DENIS ont pu être incorporés à la base, après leur identification croisée avec les sources déjà présentes.

Astéroïdes

La recherche d'astéroïdes connus dans un relevé comme DENIS nécessite en premier lieu de combiner les éphémérides des objets avec dates d'observation des différents slots. Une fois ce travail effectué, on obtient une liste d'objets ayant pu être observés, avec leur position.

En lançant de façon systématique le programme d'interrogation, plus d'une centaine d'observations d'astéroïdes connus ont pu être extraites de la base privée de strips du LDAC.

¹²<http://162.38.131.33/AstridIndex.html>

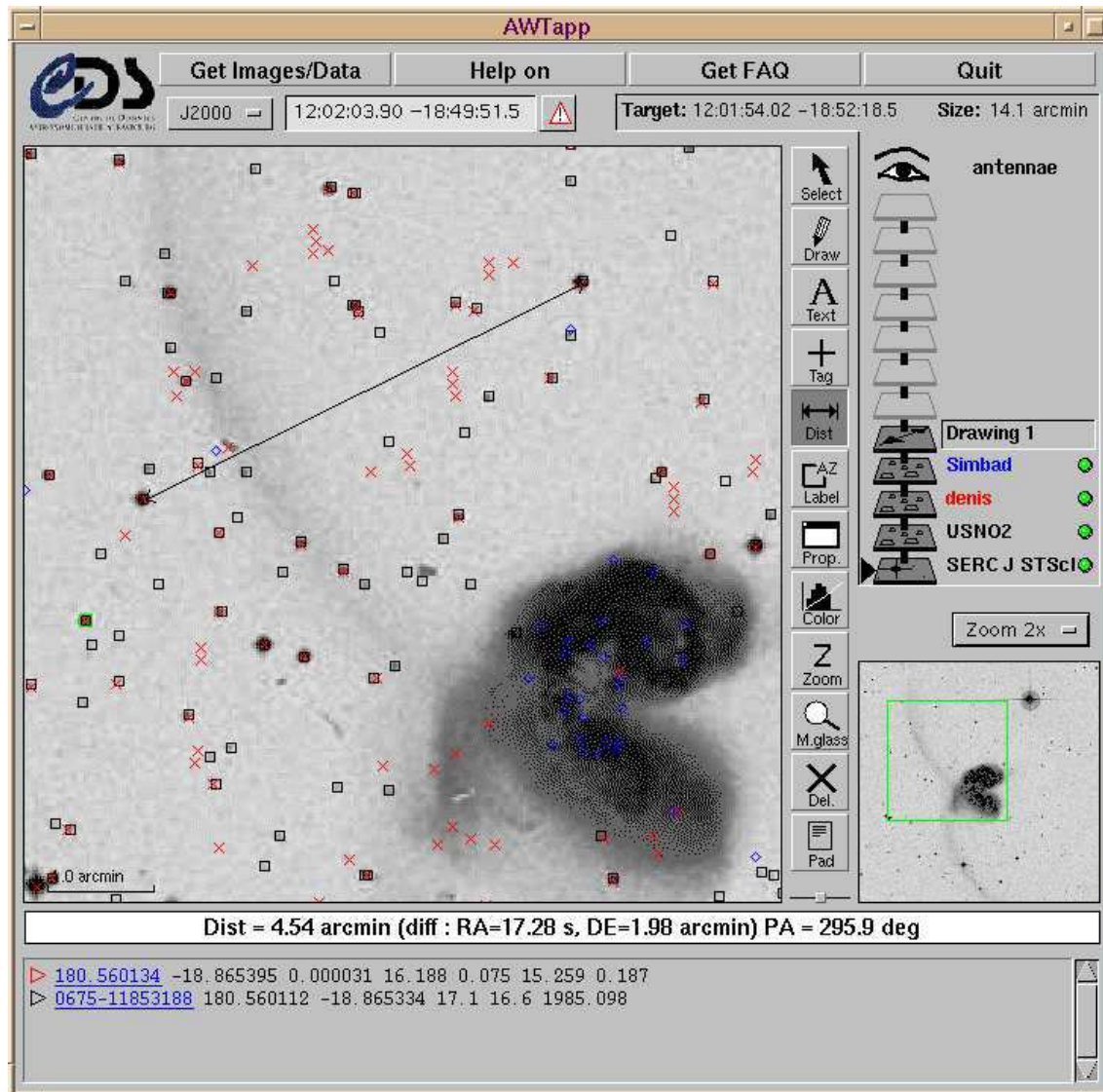


Fig. 4.2: Exemple d'utilisation du catalogue DENIS dans ALADIN, autour des galaxies des Antennes. Les sources présentes dans SIMBAD ainsi que celles du catalogue USNO ont été superposées.

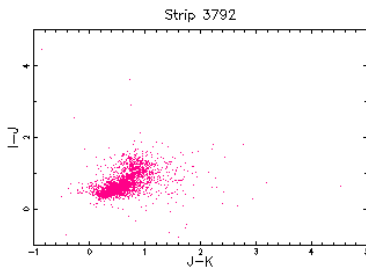
DENIS strip 3792

This strip contains **45546** records, **61** parameters in **IJK** . It covers the following region (slot **1083**):

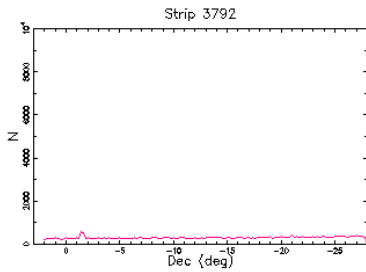
RA center	12 02 01.74	α	180.507239 deg
DEC north	+02 06 10.4	α	2.102893 deg
DEC south	-27 56 14.3	α	-27.937310 deg

Observation date: **29/03/96** --- Reduction date: **29-Oct-1998**.
 Zero points and systematic errors on magnitudes:

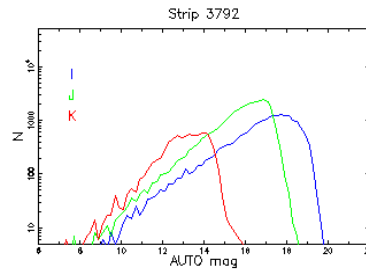
	Zero Point	Zero Point Sigma	MAG APER 7"	MAG APER 15"	MAG AUTO
I	23.607	0.0323	0.059	0.034	0.032
J	21.668	0.0775	0.082	0.048	0.054
K	19.836	0.0278	0.064	0.025	0.030



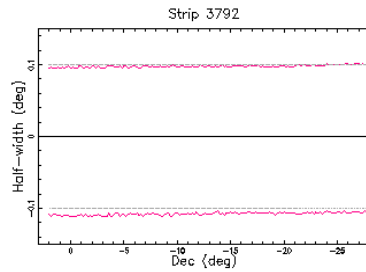
[I-J],[J-K] color-color diagram



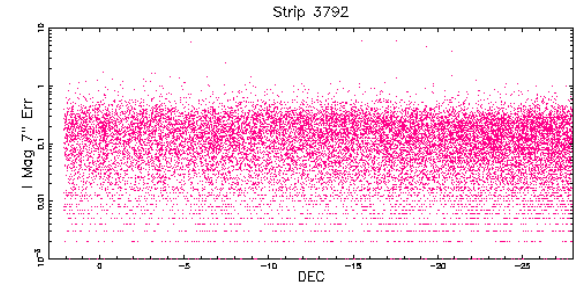
Number of sources per 12' dec. bins



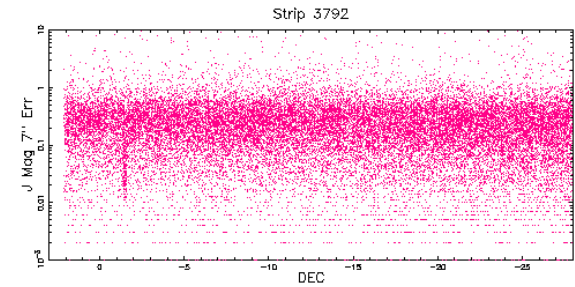
Raw star counts for I, J and K (0.2 mag bins)
 click on the graph to get the datafile.



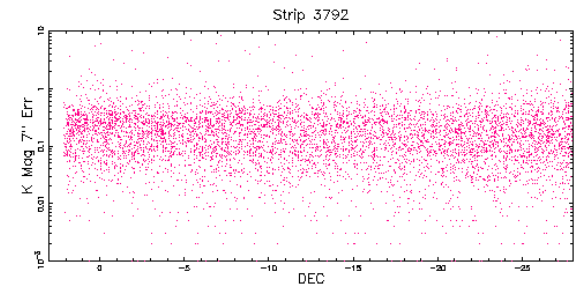
Actual range in RA with respect to average RA



Look for abnormal features in the I band



Look for abnormal features in the J band



Look for abnormal features in the K band

Fig. 4.3: Exemple de requête d'informations sur un strip DENIS. Ici, le numéro 3792
 (<http://cdsweb.u-strasbg.fr/DENIS-bin/getByNumP.pl?number=3792>).

Deuxième partie

Structure galactique

It isn't very bright to measure a blue magnitude for a red object.

Chapitre 5

La Voie Lactée

Préambule

LE SOLEIL n'est qu'une étoile parmi des milliards de voisines, toutes liées gravitationnellement dans un même ensemble formant un grand disque spiralant, notre Galaxie, la Voie Lactée. On trouve à une échelle bien plus grande d'autres galaxies, chacune formée de milliards d'étoiles.

À l'intérieur de la Galaxie, le soleil tourne avec le disque, à environ 28 000 années-lumière (8.5 kpc) du centre, et effectue un tour complet en environ 250 millions d'années.

Cette position du soleil est à la fois un avantage et un inconvénient pour étudier la Galaxie. Cela revient un peu à étudier une forêt depuis l'intérieur, mais sans pouvoir se déplacer, en restant à côté du même arbre (le Soleil). Il est pratique, pour étudier la structure à grande échelle d'une forêt, d'utiliser des photos satellites, même si on ne peut pas distinguer et identifier chaque arbre sur le cliché. On a une vue d'ensemble, un peu comme lorsque l'on regarde une galaxie extérieure à la notre : on ne discerne pas chaque étoile, mais on peut étudier sa morphologie à grande échelle, en raison du point de vue éloigné.

À l'inverse, imaginons-nous plantés dans un bois, sans possibilité de se déplacer ni même de changer de point de vue. On peut acquérir beaucoup d'informations sur les arbres les plus proches, les identifier, les mesurer, les classer, mais il devient extrêmement ardu de reconstruire une carte complète de la forêt, surtout si certaines zones éloignées sont denses ou brumeuses. L'astronome qui veut étudier la Voie Lactée est confronté à un problème similaire. Il peut obtenir une foule de renseignements sur les étoiles proches, et en ce sens la position du soleil, niché dans la Galaxie, est un avantage. Mais il est confronté aux pires difficultés pour dresser une carte à grande échelle de la Voie Lactée, à cause des nuages de poussières disséminés dans le disque, qui obscurcissent sa vue, et des régions centrales très denses qui masquent l'arrière-plan.

La suite de ce chapitre décrit les principales étapes historiques de la compréhension de la nature de notre Galaxie, et présente rapidement les caractéristiques majeures de la Voie Lactée.

5.1 Historique¹

Il a fallu attendre très longtemps avant de résoudre le problème du positionnement du Soleil dans l'Univers. La place de la Terre dans le système solaire est bien établie depuis le XVI^e siècle, avec les travaux de Copernic, Tycho Brahe et Kepler. Mais faute de pouvoir mesurer de façon fiable les distances et les mouvements du Soleil par rapport aux autres étoiles, la structure de l'Univers (à petite comme à grande échelle) est restée extrêmement spéculative pendant plusieurs siècles. Il a fallu attendre le XX^e siècle pour que les scientifiques arrivent à un consensus sur le sujet.

5.1.1 Un univers-île

La nature de la bande lumineuse diffuse barrant le ciel nocturne était inconnue des grecs et des romains de l'Antiquité : sa blancheur inspira des légendes où elle était constituée de lait, d'où le nom de Voie Lactée.

C'est Galilée qui, en pointant sa lunette dans cette direction en 1610, fit le premier la constatation que la clarté diffuse de la Voie Lactée pouvait se résoudre en une myriade d'étoiles individuelles indiscernables à l'œil nu, mettant ainsi en évidence sa nature stellaire.

Un siècle et demi plus tard, Emmanuel Kant échafaude dans son ouvrage *Allgemeine Naturgeschichte und Theorie des Himmels* plusieurs hypothèses en rapport avec la Galaxie. Par analogie avec la dynamique du système solaire, il explique qu'un système d'étoiles liées gravitationnellement devrait prendre la forme d'un disque, et que si le Soleil se trouve à l'intérieur de ce disque, l'ensemble des étoiles qui le composent doit avoir la même apparence que la Voie Lactée. Il assimile les étoiles se trouvant sur des orbites en dehors du disque aux comètes du système solaire.

Il va jusqu'à supposer qu'il pourrait exister dans l'Univers d'autres systèmes d'étoiles similaires, et que certaines des **nébuleuses** connues pourraient être des « univers-îles », de même nature que la Voie Lactée, mais situés à de très grandes distances.

5.1.2 Nébuleuses et Galaxie

Les nébuleuses furent aux XVIII^e et XIX^e siècles à l'origine de nombreux débats et polémiques. On appelait nébuleuse tout objet diffus n'étant pas une comète, mais la nature de ces objets n'était pas connue. On sait aujourd'hui que ce vocable groupait des objets de natures diverses : amas globulaires, galaxies, nébuleuses planétaires...

Un des thèmes très en vue en astronomie au XVIII^e siècle était la recherche de comètes (découvrir plusieurs comètes permettait en effet de se faire un nom dans la

¹Cette section reprend des éléments du chapitre 1.2 de *Galactic Astronomy* de Binney & Merrifield (1998).

communauté scientifique). Entre 1758 et 1782, le français Charles Messier recensa une liste² d'une centaine de « nébuleuses » fixes, à ne pas confondre avec des comètes.

Avec les progrès des télescopes, le nombre de nébuleuses connues s'est considérablement accru. William Herschel, rendu célèbre par la découverte d'Uranus en 1781, fut impressionné par le travail de Messier, et entreprit de cataloguer de façon systématique les objets du ciel profond de l'hémisphère nord. Avec l'aide de sa sœur Caroline, il recensa environ 2500 « nébuleuses », et tenta de les classer en huit groupes, d'après leur aspect. Son télescope lui permettait en effet de discerner des étoiles individuelles dans certains des objets, alors que d'autres semblaient formés de gaz uniquement. Son fils John poursuivit son travail, en observant aussi depuis l'hémisphère sud, portant le nombre total d'objets à 5000 au milieu du XIX^e siècle.

William Herschel avait également tenté de déterminer la forme de la Voie Lactée. A partir de comptages d'étoiles dans de nombreuses directions, et en supposant que toutes les étoiles aient la même brillance et qu'il puisse voir d'un bout à l'autre de la Galaxie, il trouva que celle-ci devait avoir la forme d'un disque, dont la hauteur faisait environ un sixième du diamètre, avec le Soleil situé près du centre. Mais ne disposant pas d'information sur la luminosité intrinsèque des étoiles, il ne pouvait pas fixer la taille absolue de la Galaxie.

Au début du XX^e siècle, Kapteyn entreprit une analyse systématique de la distribution des étoiles, grâce aux techniques photographiques, en mesurant leurs mouvements propres, ainsi que leurs vitesses radiales (par décalage Doppler dans les spectres). Les conclusions de cette étude étaient en bon accord avec l'image de Herschel. La Galaxie de Kapteyn avait la forme d'un sphéroïde aplati, avec un rapport d'axes de 1/5. La densité des étoiles diminuait de moitié tous les 800 pc dans le disque, et tous les 150 pc perpendiculairement au plan. Le Soleil était situé près du centre (650 pc), dans le plan. La taille de la Galaxie semblait donc de l'ordre de 10 kpc.

Deux points posaient cependant encore problème à l'époque :

- le modèle de la Galaxie de Kapteyn, avec le Soleil si près du centre galactique, semblait un peu trop héliocentrique. Une hypothèse supplémentaire permet d'expliquer cette situation privilégiée : faire intervenir un milieu interstellaire, avec lequel interagirait la lumière des étoiles lointaines. Les observations de la Voie Lactée montrent en effet des régions « vides » d'étoiles, que l'on peut expliquer en imaginant qu'un nuage de matériau absorbant se trouve sur la ligne de visée.

Dans un système d'étoiles de très grande taille, s'il existe entre les étoiles un milieu interstellaire absorbant, les étoiles situées au loin apparaîtront plus faibles : si on interprète comme un effet de distance la faiblesse de ces objets, on les croira plus lointaines, ce qui conduira à sous-estimer la densité réelle d'étoiles à grande distance. En outre, les étoiles les plus lointaines ne seront

²<http://www.seds.org/messier/>

pas détectées, donnant l'illusion que le Soleil se trouve au centre d'un système de taille réduite.

Kapteyn lui-même avait essayé d'estimer l'importance de l'extinction dans ses mesures. Supposant que le milieu interstellaire interagissait par diffusion Rayleigh sur le rayonnement des étoiles, il conclut (à tort) qu'il y avait très peu d'extinction dans ses données. On sait en fait aujourd'hui que c'est l'absorption par les poussières, et non la diffusion par le gaz, qui est la principale cause de l'extinction ;

- le deuxième problème toujours sujet à débat était la nature des nébuleuses spirales. On sait aujourd'hui que ces « nébuleuses » sont des galaxies extérieures à la nôtre. A l'époque, certains prétendaient que ces nébuleuses se trouvaient au sein de notre Galaxie (et pouvaient être des systèmes planétaires en formation, par exemple). En effet, en raison de leur petit diamètre apparent, imaginer que ces objets soient comparables en taille à la Voie Lactée impliquait qu'ils soient situés à des distances jugées absurdement grandes. D'autres astronomes, parmi lesquels Heber D. Curtis, soutenaient que ces objets étaient bien des systèmes de même nature que notre propre Galaxie, situés en dehors de celle-ci.

5.1.3 Le Grand Débat

Le modèle de Galaxie de Kapteyn fut remis en cause en 1917 par Harlow Shapley, suite à son étude de la distribution des amas globulaires. Les amas globulaires sont des groupements très compacts de 10^4 à 10^6 étoiles, mais les observations de l'époque permettaient déjà de distinguer des étoiles individuelles dans ces objets³. En utilisant des étoiles variables (Céphéides, RR-Lyrae) comme indicateurs de distance, Shapley découvrit une répartition des amas globulaires très différente de celle des étoiles : la majorité des amas se trouve en direction de la constellation du Sagittaire. L'ensemble des amas globulaires est réparti dans un immense nuage quasiment sphérique, dont le centre se trouvait selon Shapley à environ 15 kpc du Soleil en direction du Sagittaire, et dont il estima la dimension totale à 100 kpc.

Shapley supposa que le centre de la distribution des amas globulaires devait correspondre au centre de la Galaxie, et que celle-ci est donc environ dix fois plus grande que ne le suggérait le modèle de Kapteyn (voir Fig. 5.1).

Le modèle proposé par Shapley posait de nouvelles questions aux astronomes : si le centre de la Galaxie se trouve bien à 15 kpc en direction du Sagittaire, pourquoi n'y voit-on pas plus d'étoiles ? Pourquoi les comptages d'étoiles sont-ils en aussi grand désaccord avec la distribution des amas globulaires ?

La réponse définitive à ces questions ne sera donnée qu'en 1930, quand Robert Trumpler découvrira, en étudiant les amas ouverts, que la poussière présente dans le plan de la Galaxie obscurcit tous les objets situés à plus de quelques kiloparsecs.

En raison de sa découverte, Shapley va se tromper lourdement sur la nature des nébuleuses spirales. Pour lui, celles-ci appartiennent à la Voie Lactée. Il considère

³Certains amas globulaires brillants figurent dans la liste des nébuleuses de Messier.

en effet qu'il n'y a plus de problème pour les placer dans une Galaxie 10 fois plus grande que ce que l'on pensait jusqu'alors. Et l'Univers se limite pour lui à la seule Voie Lactée.

Curtis, lui, est intimement convaincu que les nébuleuses spirales sont les « univers-îles » décrits par Kant. Mais si ces systèmes sont de taille comparable à la Voie Lactée, la taille de la Galaxie de Shapley impose que les nébuleuses spirales se trouvent à des distances considérables, et jugées comme aberrantes à l'époque. C'est pourquoi Curtis privilégie le modèle plus compact de Kapteyn pour notre Galaxie.

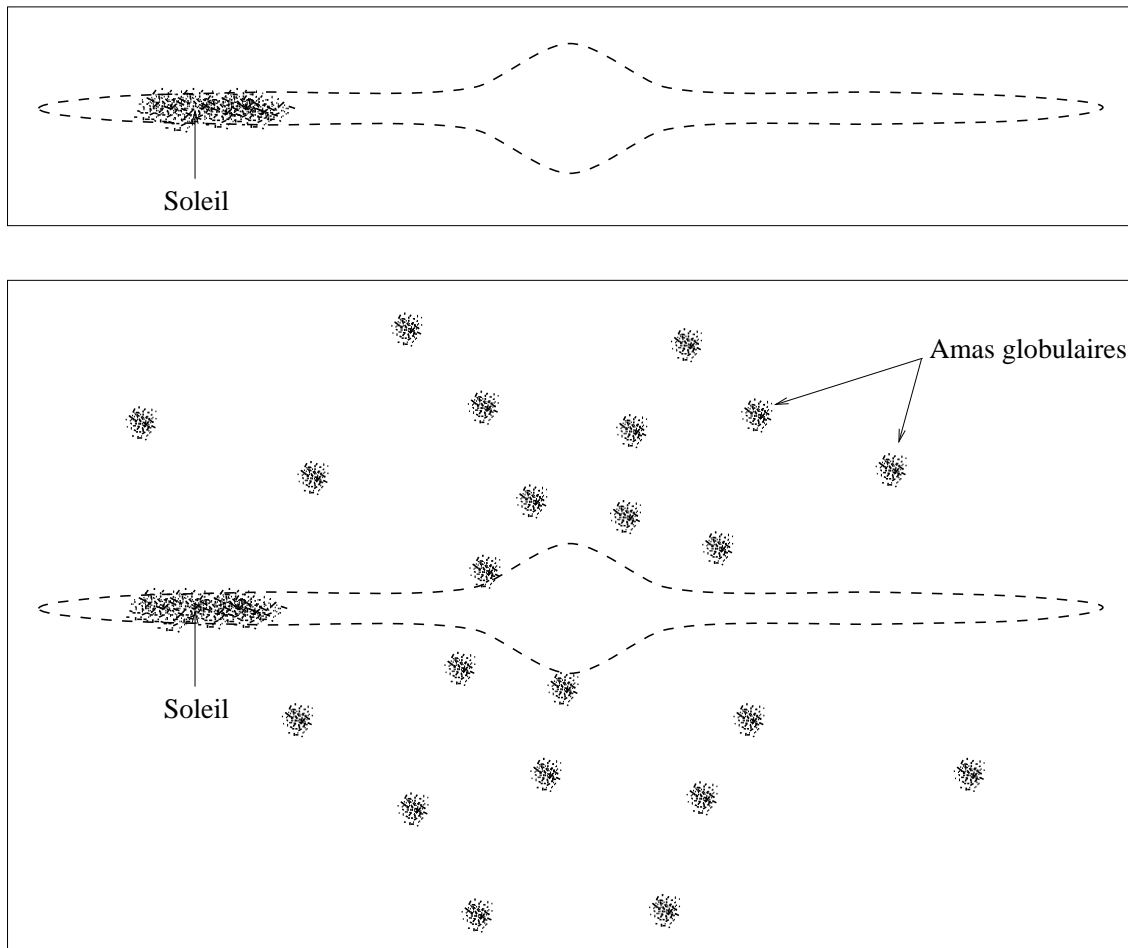


Fig. 5.1: Représentation schématique des visions de la Galaxie selon Kapteyn et Shapley. Pour Kapteyn (en haut), la distribution des étoiles indique que le Soleil se trouve proche du centre d'un système dont la taille est d'environ 10 kpc. En fait, l'extinction empêche de voir le reste de la Galaxie. La répartition des amas globulaires conduit Shapley à proposer une Galaxie beaucoup plus grande (en bas), avec le Soleil situé à 15 kpc du centre galactique (qui correspond au centre de la distribution des amas).

Ces deux astronomes se sont affrontés en avril 1920 lors d'une conférence qui est restée comme le « Grand Débat ». En fait, chacun y soutient une conclusion en

partie erronée, car aucun des deux n'accepte les distances gigantesques de l'Univers.

Chacun tente de trouver des arguments pour soutenir son point de vue. En 1885, une nouvelle étoile auparavant invisible fut découverte au cœur de la Nébuleuse d'Andromède (M31). Ce nouvel objet, nommé S Andromeda, fut tout d'abord assimilé aux novae⁴ déjà connues dans notre propre Galaxie. Mais à partir de 1910, des objets 10 000 fois plus faibles que S Andromeda sont aussi découverts dans M31, et ces objets font également penser à des novae. Pour Shapley, la magnitude de S Andromeda prouve que M31 se trouve dans notre Galaxie, car dans le cas contraire, cela impliquerait une luminosité beaucoup trop grande pour une nova. En réalité, S Andromeda était une supernova, mais ce type d'objets était encore inconnu à l'époque.

Shapley utilise aussi comme argument la **rotation de van Maanen**. Cet observateur respecté avait déduit de l'étude de plaques photographiques de nébuleuses spirales qu'on pouvait mesurer une vitesse de rotation pour ces objets : les périodes de rotation étaient voisines de 10^5 ans. Mais en supposant une taille de 5 kpc pour une galaxie, ceci impliquait des vitesses invraisemblables pour les étoiles situées à la périphérie de la galaxie. La seule conclusion possible était donc que les nébuleuses spirales étaient beaucoup plus petites que 5 kpc, et donc n'étaient pas des systèmes similaires à la Voie Lactée. La seule faille dans ce raisonnement était que van Maanen s'était trompé, et qu'aucun mouvement de rotation n'était visible, comme cela fut démontré quelques années plus tard.

Curtis, de son côté, avançait comme argument la diversité des tailles apparentes des nébuleuses spirales (de plus de 2 degrés pour M31 à quelques secondes d'arc pour les plus faibles). En considérant que ce sont tous des systèmes de même nature, et même si M31 appartient à la Voie Lactée, les plus faibles de ces objets se situent bien à l'extérieur de notre Galaxie.

D'autre part, il était connu à l'époque que certaines nébuleuses spirales vues par la tranche sont barrées par une bande obscure. Curtis fit le rapprochement entre ces bandes sombres et la « zone d'exclusion » de notre Galaxie : si on observe la distribution des nébuleuses spirales, on n'en trouve quasiment aucune à moins de 15 degrés de latitude par rapport au plan galactique. Curtis suggéra que les galaxies sont entourées d'un anneau de matière absorbante, et que cet anneau, dans le cas de la Voie Lactée, masque les galaxies extérieures à trop basse latitude.

La réconciliation entre les idées avancées par Shapley et Curtis se fit dans la décennie suivante. Edwin Hubble, grâce au télescope de 100 pouces du Mont Wilson, put étudier en détail quelques nébuleuses spirales. Il détecta dans certaines nébuleuses des étoiles faibles présentant les mêmes caractéristiques que les Céphéides dans notre Galaxie. En supposant l'universalité de la relation période-luminosité de ces étoiles variables, il put déterminer les distances de NGC 6822 (Hubble, 1925), M33 (Hubble, 1926) et M31 (Hubble, 1929), et prouver que ces objets étaient des

⁴Une nova est une étoile dont la luminosité présente de brusques sursauts, puis décroît progressivement.

systèmes stellaires à part entière, situés bien en dehors de la Voie Lactée (même en considérant la « grande » Galaxie de Shapley).

D'autre part, la reconnaissance de la présence de poussières absorbantes dans le milieu interstellaire (Trumpler, 1930) permit de réconcilier les vues apparemment discordantes de Shapley et Kapteyn sur la taille de notre Galaxie : la Voie Lactée est bien aussi étendue que Shapley l'avait trouvé à partir de l'étude des amas globulaires, mais la présence de poussière dans le disque masque les étoiles au-delà de quelques kiloparsecs, ce qui explique la faible dimension trouvée par Kapteyn.

Au début des années 1930, on a donc une vision assez juste de l'étendue de la Voie Lactée, et on admet que notre Galaxie n'est qu'un système stellaire parmi une multitude d'autres systèmes semblables.

Avec les progrès des instruments, les observations dans différents domaines de longueurs d'onde, mais aussi les développements des méthodes de calcul permettant de modéliser la Galaxie, la connaissance de celle-ci s'est progressivement affinée au cours du XX^e siècle. La section suivante présente rapidement les grandes lignes de la vision actuelle de la structure de la Voie Lactée.

5.2 La structure de la Galaxie

5.2.1 Une galaxie spirale ordinaire

Les astronomes s'accordent à penser que la Voie Lactée est une grande galaxie spirale, dont le disque mesure environ 30 kpc de diamètre⁵, et qui contient de l'ordre de quelques 10^{11} étoiles. Les instruments actuels ne nous permettent de voir qu'une infime fraction de ces étoiles, et c'est bien évidemment seulement dans le voisinage solaire que les observations permettent d'obtenir des échantillons complets pour tous les types d'objets.

Pour appréhender la structure à grande échelle de la Galaxie, les astronomes adoptent le principe de relativité, qui postule que **les lois de la nature doivent s'appliquer à tous les systèmes de référence, quel que soit leur état**. On suppose donc que les lois de la physique sont les mêmes en tout point de la Galaxie, mais aussi dans les galaxies extérieures.

Puisque la Voie Lactée n'est qu'une galaxie spirale comme les autres, l'étude de galaxies extérieures va nous donner des indications précieuses sur les phénomènes pouvant être présents au sein de notre propre Galaxie, et les caractéristiques structurelles à étudier.

En outre, la connaissance détaillée des positions et cinématiques d'un très grand nombre d'étoiles n'implique pas directement un progrès décisif de notre connaissance de la structure de la Galaxie. Pour reprendre une image utilisée par Gould (1999), la connaissance détaillée des positions et vitesses de 10^{34} molécules dans une pièce

⁵Le disque stellaire vieux fait 30 kpc de diamètre, mais le disque de gaz et le halo sont plus étendus.

n'apporte pas beaucoup plus d'information sur l'air de la pièce que la connaissance de la température, de la pression et de la composition chimique. Ainsi, pour la structure de la Galaxie, on groupe les étoiles en diverses **composantes** pouvant être décrites par des paramètres globaux (distribution spatiale, propriétés cinématiques, composition chimique, ...).

Populations stellaires

Profitant du black-out sur Los Angeles pendant la deuxième guerre mondiale, Walter Baade utilisa le plus puissant télescope de l'époque (toujours le 100 pouces du Mont Wilson) pour étudier des étoiles individuelles dans plusieurs galaxies extérieures (M32, NGC 205 et M31), avec différents filtres photométriques. En plaçant ces étoiles dans un diagramme couleur-magnitude, Baade (1944) distingua deux groupes d'objets (types I et II). Il fit l'analogie entre ces deux groupes et les types d'étoiles connus dans notre Galaxie, et suggéra que les galaxies, dans leur ensemble, sont composées de deux populations :

- type I : étoiles bleues (types O et B), et étoiles d'amas ouverts dans notre Galaxie, correspondant aux bras spiraux des galaxies externes ;
- type II : RR-Lyrae et étoiles des amas globulaires, présentes dans les galaxies elliptiques et les régions centrales (bulbes) des galaxies spirales.

On savait déjà que les bras spiraux des galaxies spirales, contrairement aux parties centrales, avaient des couleurs correspondant à des étoiles de types précoces (Seares, 1916; Carpenter, 1931), mais Baade va plus loin dans son étude, et surtout étend la notion de populations à notre propre Galaxie.

Cette notion de **populations stellaires** est toujours utilisée actuellement. On a découvert, avec les progrès de la spectroscopie, des différences de composition chimique entre les deux populations (Helfer et al., 1959), et avec la mise au point de modèles de structure et d'évolution stellaire (et les travaux de Sandage par exemple), des différences d'âge :

- la population II est composée d'étoiles vieilles, typiquement des géantes rouges de $0.8 M_{\odot}$. L'étude de la composition chimique de ces étoiles révèle qu'elles sont pauvres en éléments plus lourds que l'hélium : elles se sont formées au début de l'histoire des galaxies, dans un milieu encore peu enrichi en métaux⁶ ;
- les étoiles de la population I sont plus jeunes, et plus riches en éléments lourds. Apparues plus tardivement, elles ont pu incorporer, lors de leur formation, des métaux synthétisés et rejetés par les étoiles massives de la population II arrivées en fin de vie.

⁶Les éléments lourds ne proviennent pas de la nucléosynthèse primordiale (Big Bang), mais sont produits par les réactions nucléaires au cœur des étoiles. Lors des derniers stades de leur évolution –supernovae, nébuleuses planétaires–, les étoiles expulsent une partie de leur masse, et contribuent à l'enrichissement du milieu interstellaire en métaux.

Composantes cinématiques

Cette notion de populations stellaires se retrouve également dans les études cinématiques de la Galaxie. Bertil Lindblad avait présenté une théorie pour expliquer les mouvements des étoiles dans la Galaxie (1925; 1926b; 1926a; 1926c) : il suggéra que la Voie Lactée pouvait être décomposée en un ensemble de composantes possédant une symétrie par rapport au centre du système. Suivant Shapley, il situa le centre de la Galaxie à grande distance du Soleil, dans la direction du centre du système d'amas globulaires.

On savait alors qu'il existait une grande différence entre les cinématiques des amas globulaires et des RR-Lyrae d'une part, et les étoiles brillantes du voisinage solaire : les amas globulaires présentent des vitesses ($200 - 300 \text{ km}\cdot\text{s}^{-1}$) très importantes par rapport aux étoiles proches, entre lesquelles la dispersion des vitesses n'est que de quelques dizaines de kilomètres par seconde.

Lindblad et Oort vont développer une théorie cinématique de la Galaxie dans une série de papiers (Oort, 1927b; Lindblad, 1927; Oort, 1927a; Oort, 1928; Lindblad, 1930). L'image qui s'impose fait la distinction entre deux composantes. Les amas globulaires, ainsi que les étoiles d'une composante sphéroïdale, ont un mouvement de rotation très faible, et l'orientation de leurs orbites est aléatoire. Au contraire, le Soleil et la plupart des étoiles de son voisinage suivent un mouvement de rotation autour du centre galactique, à une vitesse de $200 \text{ à } 300 \text{ km}\cdot\text{s}^{-1}$. Ces étoiles possèdent une faible dispersion des vitesses, et forment le disque aplati de la Voie Lactée.

On retrouve donc une distinction cinématique entre les deux populations. Les objets de la population I se trouvent dans le disque, sur des orbites quasi circulaires, alors que les étoiles de la population II sont réparties dans un halo, où la dispersion des vitesses est grande, mais qui ne présente pas de mouvement de rotation.

Dans un article de revue, Sandage (1986) relate l'évolution des connaissances sur la notion de populations stellaires depuis les travaux de Baade, en passant par la conférence du Vatican (1957), où l'on distinguait cinq populations : le halo (population II), la population II intermédiaire (ou disque épais), un disque vieux, une population I vieille et une population I extrême). On glisse ainsi de l'idée de deux populations discrètes à une série de populations dont les paramètres évoluent progressivement, et dont la nature dépend intimement de l'évolution galactique.

Sandage met en relief les trois paramètres dominants dans le concept de population : **l'âge, la composition chimique et la cinématique.**

Les différents constituants

La figure 5.2 représente, dans différents domaines de longueurs d'onde (des ondes radio aux rayons X), les régions à basse latitude galactique ($\pm 10^\circ$) de la Voie Lactée.

Le développement au cours du XX^e siècle des techniques permettant d'observer dans d'autres domaines spectraux que le visible (en particulier les observations de la raie à 21 cm de l'hydrogène neutre et dans l'infrarouge proche) ont été décisives dans l'étude de la Galaxie.

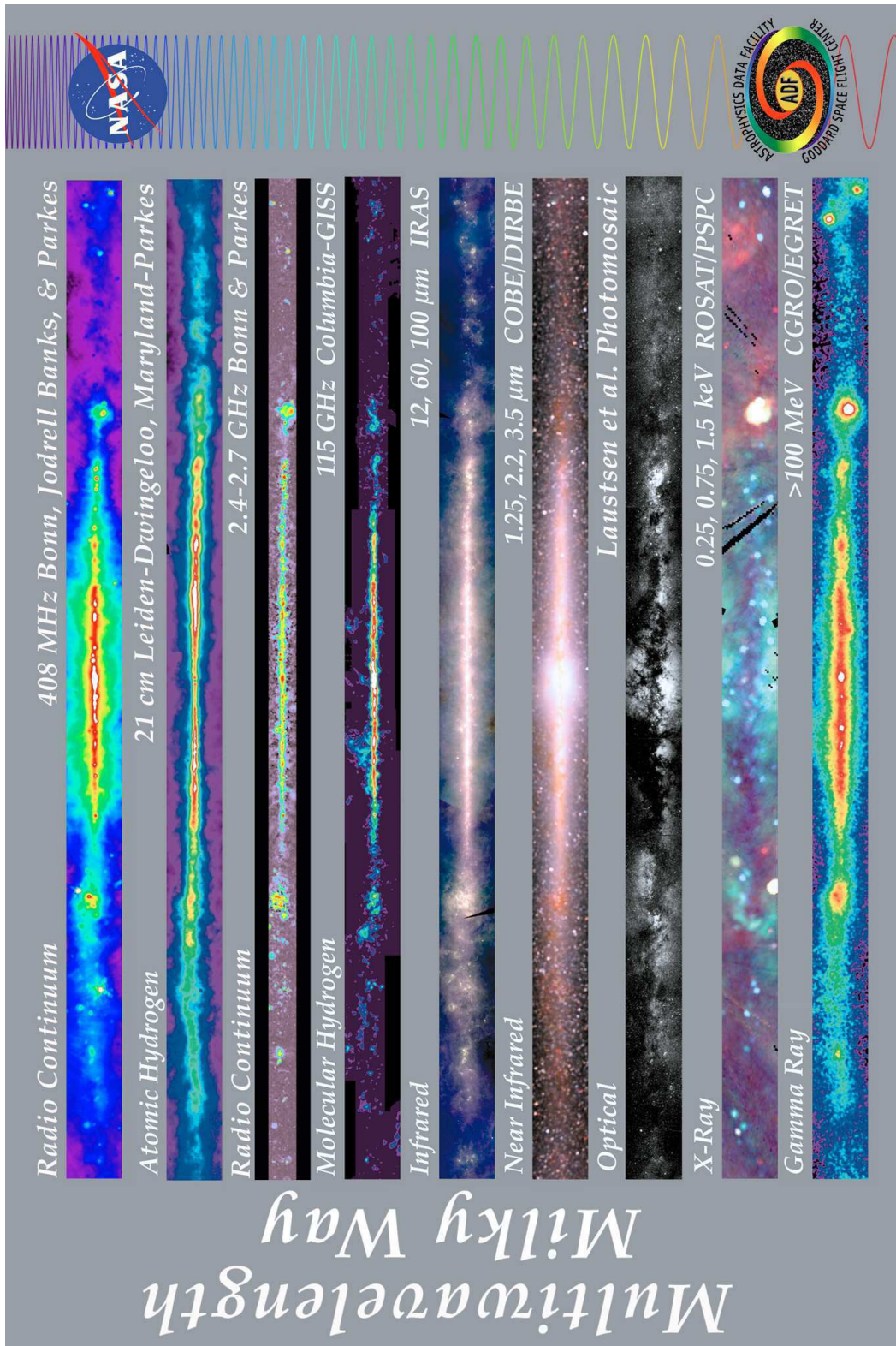


Fig. 5.2: La Voie Lactée dans différentes longueurs d'onde. ©Astrophysics Data Facility, <http://adc.gsfc.nasa.gov/mw/>

On peut distinguer trois composantes majeures dans la structure galactique :

- le bulbe, au centre de la Galaxie ;
- le halo, ou composante sphéroïdale ;
- le disque, très aplati, que l'on peut décomposer en un disque mince et un disque épais (population intermédiaire).

Ce découpage en composantes est bien sûr une simplification. Il n'y a pas, en réalité, de découpage aussi net dans la distribution des étoiles : on peut trouver, en un même endroit de la Galaxie, un mélange d'objets appartenant à plusieurs composantes. Les transitions entre les différentes composantes (bulbe – halo, bulbe – disque, disque mince – disque épais) sont encore mal connues et suscitent de nombreux travaux.

5.2.2 Le bulbe

Le bulbe est la partie centrale de la Galaxie. La plupart des étoiles du bulbe sont concentrées dans un rayon de 2 ou 3 kpc autour du centre galactique, situé approximativement à 8.5 kpc du Soleil dans la direction du Sagittaire.

Notre connaissance du bulbe galactique n'a vraiment progressé qu'avec les observations dans les domaines radio, infrarouge, et X. En effet, dans le visible, la présence de poussières dans le disque, sur la ligne de visée, obscurcit les étoiles du bulbe : cette extinction peut dépasser 30 magnitudes visibles par endroit (voir (Schultheis et al., 2000), et l'image optique de la figure 5.2).

Le noyau

La partie la plus interne du bulbe, le noyau, est une région de quelques parsecs renfermant une très grande densité de matière. Le centre de notre Galaxie est très peu actif, contrairement à certaines galaxies extérieures, dont le noyau est parfois plus lumineux que la Voie Lactée toute entière (ce sont les noyaux actifs de galaxie, AGN). On trouve dans le noyau de la Voie Lactée une source radio étendue, *Sagittarius A*, qui entoure une région très compacte, Sgr A*, considérée comme le centre galactique. L'émission intense de cette petite zone dans l'infrarouge et en rayons X, et l'observation de nuages de gaz se déplaçant rapidement indiquent qu'un trou noir de $3 \cdot 10^6 M_{\odot}$ se situe vraisemblablement au cœur du centre galactique. En étudiant le déplacement des étoiles situées à proximité de Sgr A*, Genzel et al. (2000) avancent également la présence d'un objet compact de $2.6 \cdot 10^6 M_{\odot}$ à $3.3 \cdot 10^6 M_{\odot}$ situé à 8 kpc du Soleil.

Métallicité

Frogel (1988) a publié une revue reprenant l'évolution de notre vision du reste du bulbe, depuis la définition du concept de populations stellaires par Baade (1944). L'idée d'un bulbe strictement composé d'étoiles de la population II a été progressivement abandonnée. Le bulbe de la Galaxie est dominé par des géantes de types

K et M. Arp (1959) fut le premier à mettre en évidence une différence importante entre des géantes du bulbe et les géantes d'un amas globulaire (M22, typique de la population II) : les étoiles du centre galactique étaient bien moins lumineuses que des étoiles de la population II de même couleur.

Cette différence fut associée à une différence de métallicité. Les étoiles du bulbe sont bien plus riches en éléments lourds que la population II. Rich (1988) trouve même pour un échantillon de géantes K du bulbe une métallicité moyenne deux fois supérieure à la métallicité solaire.

Les travaux plus récents (détaillés dans la revue de Wyse et al. (1997)) donnent une valeur légèrement plus basse de la métallicité moyenne, de l'ordre de -0.3 dex⁷, mais surtout ils confirment une très importante dispersion des valeurs. Les tentatives de mesurer un éventuel gradient de métallicité dans le bulbe sont compliquées par la présence d'étoiles du disque sur la ligne de visée, mais le fait que les étoiles du bulbe sont beaucoup plus métalliques que celles du halo n'est pas contesté.

Âge

Même si l'on trouve dans la région la plus centrale du bulbe des preuves d'une récente activité de formation d'étoiles (Gredel, 1996), le reste du bulbe est composé en majorité d'étoiles vieilles (au moins 10^{10} ans). Cette constatation repose sur l'observation d'un diagramme HR du bulbe, où les étoiles de type A–F ont déjà évolué, mais la détermination d'un âge précis pour les étoiles soulève deux problèmes :

- l'évolution des étoiles dépend fortement de leur métallicité ;
- il est difficile d'obtenir un échantillon d'étoiles qui ne soit pas biaisé ou contaminé par des objets situés à l'avant-plan.

Morphologie

La forme du bulbe ne peut pas être étudiée directement dans le visible, toujours à cause de l'extinction. Les observations dans l'infrarouge ou en radio fournissent par contre de précieux renseignements sur la répartition des étoiles ou du gaz vers le centre de la Galaxie.

On a de nombreux indices prouvant que **le bulbe n'est pas axisymétrique**. La dynamique des étoiles et des nuages de gaz vers le centre galactique, ainsi que le profil de luminosité dans l'infrarouge suggèrent une forme allongée du bulbe, indiquant la **présence d'une barre**. Cette découverte n'est pas exceptionnelle, car de nombreuses galaxies spirales extérieures possèdent en leur centre une barre (voire même plusieurs) plus ou moins importante.

Différents résultats indiquent que l'extrémité la plus proche de la barre se trouve dans le premier quadrant galactique, mais la morphologie exacte de la barre et du bulbe n'est pas encore connue précisément. Dwek et al. (1995) trouvent, à partir des données infrarouges DIRBE, un rapport d'axes (1:0.33:0.23) et un angle de 20° entre

⁷dex est une abréviation de *decimal exponent*. Les métallicités sont exprimées par rapport à la métallicité solaire : n dex correspondent à 10^n fois la valeur solaire.

la barre et la ligne de visée vers le centre galactique. À partir d'observations du plan galactique dans l'infrarouge, Hammersley et al. (2000) indiquent une extension de la barre à 4 kpc du centre galactique, avec un angle de 43° . Il pourrait donc y avoir, au centre de la Voie Lactée, un bulbe triaxial et une barre, ou bien une double barre.

5.2.3 Le halo

Le halo galactique présente la forme d'un sphéroïde légèrement aplati, s'étendant sur plus de 30 kpc de rayon. On y trouve des objets de la population II, caractérisés par un âge important et une métallicité nettement sous-solaire (-1.0 à -3.0 dex, avec une moyenne de -1.7 dex).

Les principaux traceurs du halo sont les amas globulaires et les variables de type RR Lyrae. On connaît environ 150 amas globulaires dans la Voie Lactée. Leur distribution suit une loi de densité décroissant radialement comme une loi de puissance d'indice $\alpha = -3.5$ (Harris, 1976; Djorgovski & Meylan, 1994). Ainsi, si on note R le rayon galactocentrique, z la hauteur au dessus du plan, et c l'aplatissement du halo :

$$\rho = \rho_0 \cdot (R^2 + z^2/c^2)^{-\alpha/2} \quad (5.1)$$

On trouve une loi de densité similaire à partir de la distribution des RR Lyrae : $\alpha = -3.1 \pm 0.2$ avec $c = 0.9$ pour Hawkins (1984), $\alpha = -3.024 \pm 0.077$ avec $c = 1$ ou $\alpha = -3.530 \pm 0.077$ avec $c = 0.5$ pour Wetterer & McGraw (1996). Les échantillons sont cependant petits, et les incertitudes non négligeables.

Le halo contient bien sûr des étoiles en dehors des amas globulaires. La distribution en densité de ces objets suit également une loi de puissance. À partir de comptages profonds dans plusieurs directions, Robin et al. (2000) trouvent un aplatissement de 0.76, et un indice plus petit que celui obtenu à partir des autres traceurs : $\alpha = -2.44$. Dans la même étude, Robin et al. (2000) contraignent également la fonction de masse initiale⁸ (IMF) pour le halo. Ils trouvent une pente de 1.9, ce qui donne une densité locale (au voisinage solaire) de $1.64 \cdot 10^{-4}$ étoiles $\cdot \text{pc}^{-3}$ pour le halo. La densité de masse locale du halo correspondante est $4.15 \cdot 10^{-5} M_\odot \cdot \text{pc}^{-3}$.

Les étoiles du halo circulent sur des orbites dont l'orientation semble aléatoire. Puisque le Soleil, et les étoiles du disque dans son voisinage, suivent un mouvement orbital dans le plan galactique à des vitesses voisines, les étoiles du halo dans le voisinage solaire se distinguent par leur grande vitesse **relative** (on peut les sélectionner en imposant $v > 220 \text{ km} \cdot \text{s}^{-1}$, par exemple).

Halo et formation galactique

L'âge important des étoiles du halo (plus de 10^{10} ans), l'absence de signes d'une récente formation d'étoiles dans cette partie de la Galaxie, et la faible métallicité de

⁸L'IMF donne la répartition en masse des étoiles lors de leur formation. On utilise classiquement une loi de puissance pour l'exprimer.

cette population indiquent que ces objets se sont formés au tout début de l'histoire de la Galaxie. Leur étude revêt donc un intérêt tout particulier pour contraindre l'histoire de la formation de la Galaxie.

Eggen, Lynden-Bell et Sandage (1962) ont avancé l'hypothèse d'une formation de la Galaxie par l'effondrement rapide (non dissipatif) d'un nuage de gaz homogène. Dans ce scénario (noté ELS), les étoiles du halo se forment toutes à la même époque (en moins de 200 millions d'années), sur des orbites fortement excentrées, tandis que le gaz s'effondre dans le plan et vers le centre galactique.

Searle & Zinn (1978) proposent un scénario plus turbulent, où les parties externes du halo galactique se forment par accréation de fragments de protogalaxies, qui continuent à évoluer pour atteindre l'équilibre dynamique, avec un temps de formation plus long, de l'ordre de 10^9 ans (scénario SZ). Chaque fragment évolue individuellement en termes de contraction, de dissipation, ou d'enrichissement en métaux.

Il existe des contraintes observationnelles pour et contre chacun de ces deux scénarii extrêmes, et les travaux les plus récents privilégient une évolution galactique hybride. La partie externe du halo se serait formée par agrégation de protogalaxies⁹ (Bekki & Chiba, 2000). On retrouve d'ailleurs dans la cinématique du halo des sous-structures interprétées comme des résidus des groupes ayant fusionné (Majewski, 1999).

A moins de 20 kpc du centre galactique, un effondrement **dissipatif** (Sandage & Fouts, 1987) aurait entraîné la formation d'un halo interne ayant un faible mouvement de rotation, puis la formation du disque (Chiba & Beers, 2000), en un temps plus long que proposé dans le scénario ELS, de l'ordre de 10^9 ans.

Masse cachée

Si l'on ne sait pas exactement comment s'est formée la Galaxie, il existe un problème non résolu encore plus intrigant : on ignore ce qui constitue 90% de la masse de la Galaxie. La courbe de rotation (Fig. 5.3) indique la présence d'une masse répartie dans un grand halo englobant la Voie Lactée, mais cette masse ne correspond à aucune matière lumineuse. On détecte sa présence uniquement par son action gravitationnelle.

Cette couronne de masse cachée (ou matière noire) englobe la Voie Lactée, mais on détecte un phénomène similaire dans la grande majorité des galaxies extérieures.

La nature de ce halo de matière noire reste hypothétique, et on avance comme possibles constituants :

- une matière noire **baryonique** :
 - les MACHOs (*Massive Compact Halo Object*), qui pourraient être des naines brunes (moins de $0.08 M_{\odot}$), des planètes joviennes ($0.001 M_{\odot}$), de vieilles naines blanches (Ibata et al., 1999) ou des trous noirs de $10^2 M_{\odot}$;

⁹La découverte de la galaxie naine du Sagittaire (Ibata et al., 1995), en orbite autour de la Voie Lactée et déformée par les forces de marée, vient renforcer cette version.

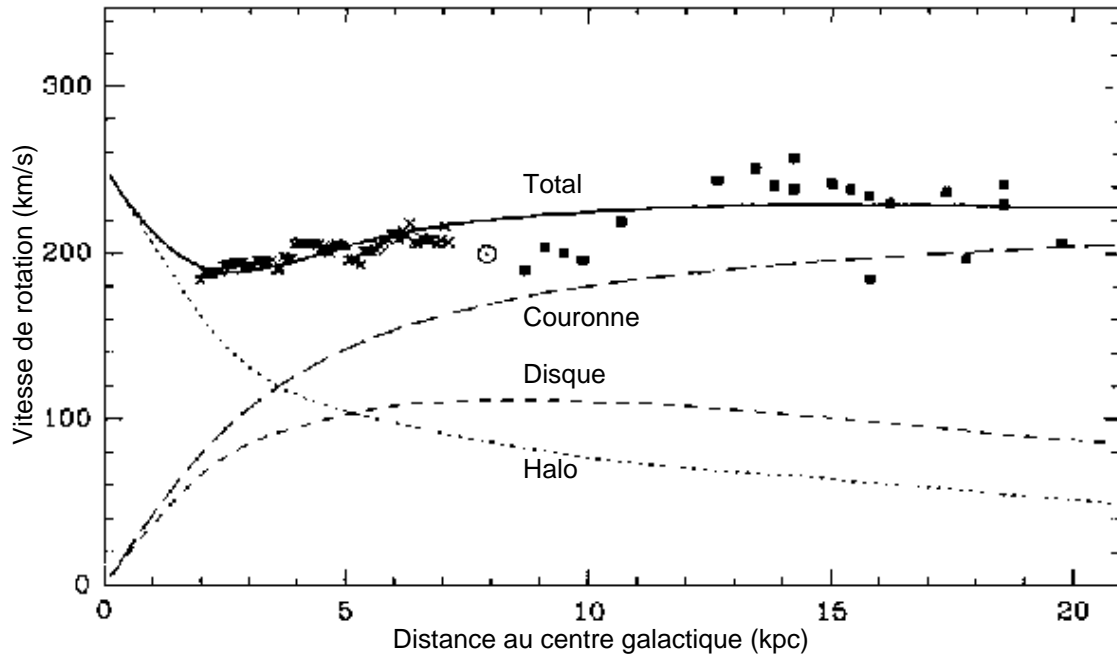


Fig. 5.3: La courbe de rotation de la Voie Lactée reste remarquablement plate jusqu'à plus de 20 kpc. La masse des composantes visibles (bulbe, disque, halo) ne permet pas d'expliquer ce comportement et il faut faire intervenir une couronne de masse cachée pour ajuster les observations. (Figure adaptée de la Fig. 9 de Merrifield (1992)).

- des nuages d'hydrogène moléculaire (Pfenniger et al., 1994) ;
- une matière noire **non baryonique**, formée de particules élémentaires encore inconnues, ou aux propriétés étranges. Parmi les nombreuses spéculations théoriques, on peut citer :
 - les WIMPs (*Weakly Interacting Massive Particle*), dont l'exemple le plus connu est le neutralino prédit par la supersymétrie ;
 - les neutrinos massifs. Si les neutrinos tauique ou muonique ont une masse non-nulle, même très faible, ils pourraient constituer la matière noire.

5.2.4 Le disque

Les disques sont les structures les plus proéminentes des galaxies spirales (voir Fig. 5.4). Leur forme très aplatie, la présence de gaz, d'étoiles jeunes et de bras spiraux leur donne une grande variété de formes, ce qui en fait des objets de choix dans les galeries d'images astronomiques.

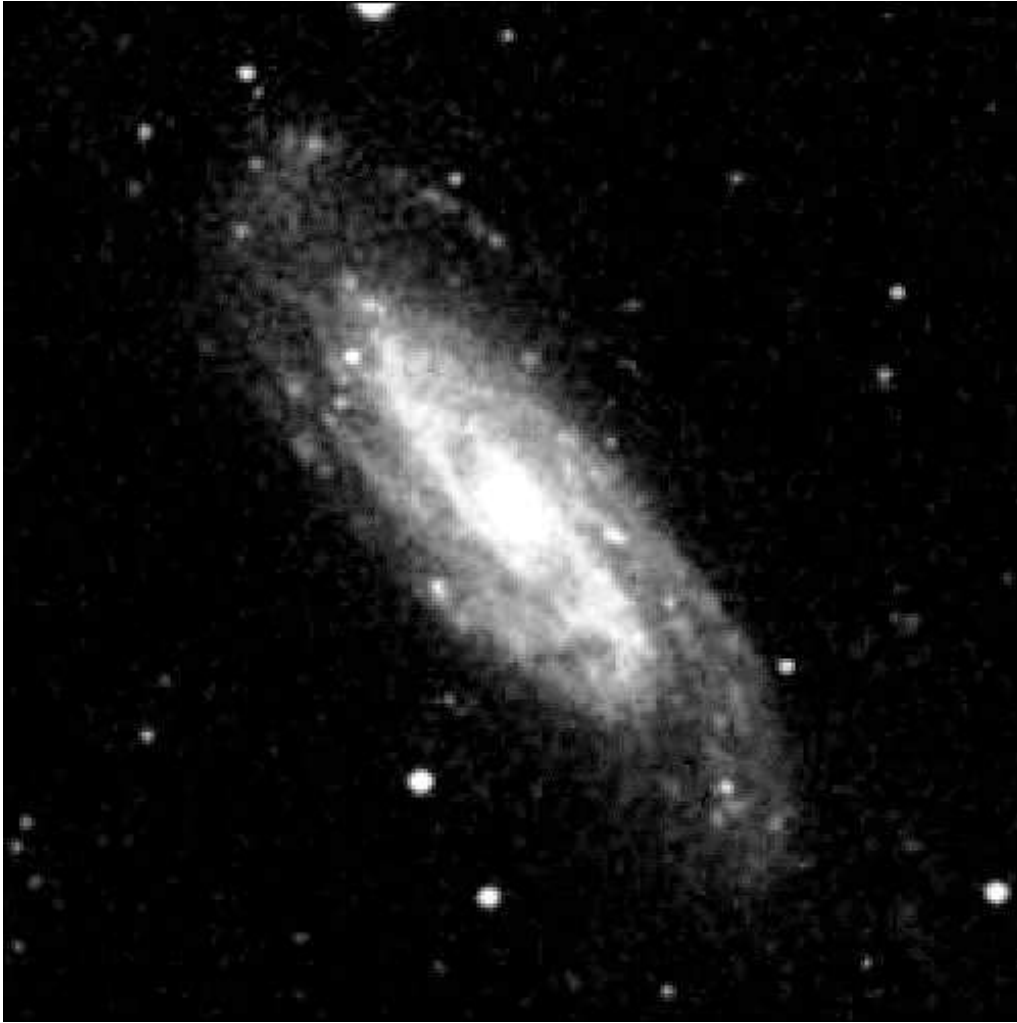


Fig. 5.4: La galaxie spirale NGC3198. On distingue dans le disque de cette galaxie des bras spiraux. (Cliché DSS1).

Le gaz

L'hydrogène est de loin l'élément le plus abondant dans l'Univers. Sous forme atomique (notée HI pour l'atome neutre), la transition hyperfine (due à l'interaction spin-spin) se traduit par l'émission d'une raie à 21.1 cm (domaine radio). L'émission à ces longueurs d'onde présente l'avantage considérable de ne pas être absorbée par les poussières.

L'intensité de l'émission à 21 cm est proportionnelle à la densité d'hydrogène atomique neutre sur la ligne de visée. En adoptant une courbe de rotation pour la Galaxie (en supposant que les nuages d'hydrogène se déplacent sur des orbites circulaires autour du centre galactique), on peut utiliser le décalage Doppler de la raie pour déterminer les distances des nuages, et donc cartographier la répartition de l'hydrogène neutre dans la Galaxie.

Les premières observations radio à 21 cm ont donné pour la première fois une vision à grande échelle de la structure de la Galaxie (Oort et al., 1958) :

- l’essentiel du gaz est confiné dans un disque plat (avec une échelle de hauteur de 100 à 220 pc selon le rayon galactique (Malhotra, 1995)) ;
- la distribution des nuages trahit la présence d’une structure spirale dans le plan, et permet de suivre le tracé de plusieurs bras ;
- le disque apparaît gauchi au-delà d’un certain rayon (le gauchissement et l’évasement du disque seront approfondis dans la section 6.3)

L’étude du gaz dans la Galaxie est importante pour comprendre l’évolution de celle-ci, car le taux de formation d’étoiles dans le disque dépend de la densité de masse surfacique dans le disque. La chute de gaz en provenance du halo vers le plan galactique (*infall*) permet de conserver une quantité de gaz suffisante pour entretenir la formation d’étoiles. Les estimations actuelles de la densité surfacique de gaz (HI et H₂) dans le voisinage solaire vont de 7 à 13 M_⊙ · pc⁻² (Romano et al., 2000).

Le disque mince

Le disque mince de la Galaxie contient des étoiles couvrant une très grande gamme d’âges, preuve d’une activité de formation d’étoiles sur une longue période. La métallicité moyenne dans le disque (population I de Baade (1944)) est plus importante que dans le halo.

La formation d’étoiles dans le disque se fait essentiellement au sein de nuages moléculaires géants. Beaucoup de ces nuages sont localisés dans les bras spiraux. Ceci explique que les étoiles massives de type O–B, à durée de vie courte (et donc formées récemment), soient localisées majoritairement dans la spirale, comme on le voit dans les galaxies extérieures.

Un ensemble d’étoiles formées dans une même région, issues d’un même nuage, peut constituer par la suite un amas ouvert. Les étoiles d’un amas ouvert ont une cinématique assez voisine, elles sont liées gravitationnellement et restent donc assez proches spatialement sur des durées très longues (plusieurs périodes de rotation galactique), avant que les interactions gravitationnelles ne dissipent progressivement l’amas.

Il existe pour les étoiles du disque une **relation âge–dispersion des vitesses** : les étoiles les plus vieilles ont une plus grande dispersion σ_w (composante perpendiculaire au plan galactique). En effet, les étoiles jeunes, formées dans le plan, possèdent à leur naissance un faible σ_w (celui du gaz). Ensuite, sur leurs orbites autour du centre galactique, les étoiles rencontrent des perturbations gravitationnelles dues à des nuages moléculaires géants et aux bras spiraux¹⁰ qui vont accroître la dispersion des vitesses perpendiculairement au plan (Spitzer & Schwarzschild, 1953)

¹⁰D’autres sources possibles de chauffage du disque ont été invoquées, par exemple la présence de trous noirs dans un halo sombre relativement aplati (Sanchez-Salcedo, 1999), ou l’accrétion de satellites (Velazquez & White, 1999).

Les dispersions σ_w pour le disque mince vont de $6 \text{ km}\cdot\text{s}^{-1}$ pour les étoiles les plus jeunes à $18 \text{ km}\cdot\text{s}^{-1}$ pour les plus âgées (voir Haywood et al. (1997a), Gómez et al. (1997) et le tableau 6.1, p. 125).

L'âge du disque n'est pas connu précisément. On peut l'estimer en ajustant des isochrones sur de vieux amas ouverts, ou bien en confrontant la fonction de luminosité des naines blanches du disque aux modèles de refroidissement de ces objets (voir Carraro (2001) pour une comparaison de ces méthodes). Les estimations donnent un âge d'environ 10 milliards d'années pour les étoiles les plus vieilles du disque.

L'origine du disque reste également assez mal définie. Dans un scénario de type ELS, le disque se forme rapidement suite à l'effondrement du gaz du halo. Les modèles récents d'évolution chimique du disque galactique (Portinari & Chiosi, 2000), qui tentent de reproduire le gradient radial de métallicité observé dans les étoiles du disque (Rolleston et al., 2000), semblent cependant plutôt indiquer une formation progressive du disque, comme proposé par Larson (1976). Dans ce genre de modèle, le disque se forme en plusieurs milliards d'années, du centre galactique vers l'extérieur. Les régions internes du disque sont donc plus vieilles, mais aussi plus riches en métaux que l'extrémité du disque, qui pourrait d'ailleurs toujours être en formation. On retrouve ces caractéristiques dans des galaxies spirales extérieures (Bell & de Jong, 2000).

La densité d'étoiles dans le disque mince est généralement décrite par une loi double exponentielle¹¹, radialement et perpendiculairement au plan. Ces lois reproduisent en effet de façon satisfaisante les profils de luminosité de galaxies extérieures (Misiriotis et al., 2000). Leur validité dans le cas de la Voie Lactée a cependant été remise en cause (Haywood et al., 1997a; Haywood et al., 1997b).

Dans le cadre des lois exponentielles, on a pour le disque mince une échelle de hauteur de 250 pc, et une échelle de longueur voisine de 2.5 kpc (Robin et al., 1992; Fux & Martinet, 1994; Ruphy et al., 1996; Porcel et al., 1998). On peut noter que, si les estimations de l'échelle de longueur variaient de 1 à 5.5 kpc il y a encore quelques années, les travaux les plus récents favorisent une valeur assez courte. L'étude de galaxies extérieures indique que l'échelle de longueur du disque mesurée dans l'infrarouge est plus courte que dans le visible (de Jong, 1996). Ce résultat peut être interprété comme un indice d'une formation du disque de l'intérieur vers l'extérieur.

Le disque épais

L'existence du disque épais, ou population intermédiaire, fut avancée lors de la célèbre conférence du Vatican (O'Connell, 1958). Comme son nom l'indique, cette population possède des propriétés intermédiaires entre le disque mince et le halo. La dispersion des vitesses ($\sigma_w \approx 40 \text{ km}\cdot\text{s}^{-1}$) y est plus grande que dans le disque mince, et la métallicité moyenne est de l'ordre de -0.7 dex.

¹¹On utilise aussi parfois une loi en sech^2 , correspondant à un disque isotherme auto-gravitant, pour le profil vertical (Camm, 1950).

La structure du disque épais est assez mal connue. Les estimations de la densité locale d'étoiles du disque épais vont de 2 à 15% de la densité du disque mince. L'échelle de hauteur, plus importante que pour le disque mince, varie de 750 à 1500 pc selon les études, et l'échelle de longueur semble légèrement plus grande que celle du disque mince.

Le disque épais est un élément essentiel dans la compréhension de l'évolution de la Galaxie (Bienaymé, 1999). De nombreuses hypothèses ont été avancées pour expliquer l'origine de cette population (voir Majewski (1993) pour une revue), mais il semble que l'accrétion d'un satellite soit la plus vraisemblable (Robin et al., 1996).

Chapitre 6

Modèles de la Galaxie

6.1 Quelques exemples

LA VOIE LACTÉE contient plusieurs centaines de milliards d'étoiles. Les modèles de la Galaxie n'ont bien évidemment pas la prétention de décrire exactement son contenu dans les moindres détails, mais plutôt de modéliser différentes composantes (bulbe, halo, disque, spirale. . .) par des lois générales permettant de reproduire de façon satisfaisante les diverses observations (distribution des étoiles en magnitudes, mouvements propres, dans différentes directions).

La mise au point de modèles de la Galaxie a véritablement pris son essor avec le développement de l'informatique et des méthodes numériques, qui ont permis de construire des modèles de plus en plus complexes et de les comparer aisément à des volumes de données toujours croissants.

L'utilisation d'un modèle permet de vérifier que l'on comprend bien les phénomènes. On peut tester un scénario, et vérifier les théories physiques.

6.1.1 Modèles de masse

Les premiers efforts de modélisation de la Voie Lactée portaient sur la répartition de la masse dans la Galaxie et l'étude du potentiel galactique, afin de mieux comprendre la courbe de rotation.

Des progrès importants ont été faits dans les années 1950. Schmidt (1956) utilise plusieurs ellipsoïdes, avec des fonctions de densité polynomiales, pour décrire différentes populations.

Einasto (1979) proposera par la suite de remplacer les polynômes par des exponentielles modifiées. Dans cet article, Einasto indique ce qu'il entend par modèle de la Galaxie : « *un ensemble de fonctions et de paramètres qui décrivent quantitativement les principales propriétés de la Galaxie et de ses populations* ».

Parmi les fonctions à décrire figurent le potentiel galactique ϕ et ses dérivées, K_R et K_z . Chaque population est décrite par une fonction de distribution dans l'espace des phases (positions et vitesses). En faisant l'hypothèse que les distributions des vitesses sont gaussiennes pour une population isotherme, et que les composantes de

vitesse perpendiculaires à la rotation sont nulles, on peut décrire une population (Einasto, 1979) par :

- sa densité spatiale ρ (et sa densité projetée P) ;
- les dispersions des vitesses : $\sigma_R, \sigma_\Theta, \sigma_z$;
- la vitesse de rotation du centroïde : V .

Les populations considérées par Einasto sont : le noyau, le bulbe, le halo, le disque, un ensemble gaz — étoiles jeunes (composante très aplatie), et une couronne massive (halo de matière noire).

6.1.2 Modèles de première génération

La méthode la plus répandue pour l'étude de la structure galactique repose sur les **comptages d'étoiles**. En effet, la distribution en magnitude absolue des étoiles (fonction de luminosité $\Phi(M)$) et la densité spatiale des objets (loi de densité ρ) permettent d'obtenir le nombre d'objets de magnitude apparente m observés dans une direction donnée :

$$N(m) = \sum_i \int_r \Phi_i(M) \rho_i(r) r^2 \omega dr . \quad (6.1)$$

L'équation 6.1 est appelée **équation de la statistique stellaire** ; elle est exprimée ici dans un cas général où plusieurs populations i interviennent. r est la distance héliocentrique, et ω l'angle solide¹.

Cette loi est en général non inversible : plusieurs combinaisons de $\Phi_i(M)$ et $\rho_i(r)$ peuvent conduire à des comptages d'étoiles identiques.

Pour contourner cette difficulté, on adopte en général, pour les lois de densité, des formes empiriques déduites de l'observation de galaxies spirales extérieures supposées semblables à la Voie Lactée. Les fonctions de luminosité sont obtenues à partir des observations au voisinage solaire. On peut ensuite faire varier les quelques paramètres qui contrôlent ces lois afin d'obtenir le meilleur ajustement aux observations.

Au cours des années 1980, de très nombreux travaux ont porté sur ce type d'ajustements, effectués principalement sur des régions à **haute latitude galactique** afin d'éviter les problèmes dus à l'extinction dans le plan galactique.

On peut en citer quelques exemples :

- Bahcall & Soneira (1980) ont proposé un modèle à deux composantes (disque et sphéroïde) pour prédire les comptages d'étoiles à plus de 20° de latitude, afin de préparer les observations du Télescope Spatial Hubble. Ils utilisent une loi de densité exponentielle pour le disque, un sphéroïde de de Vaucouleurs, et différentes fonctions de luminosité déterminées dans le voisinage solaire, qui permettent de reproduire grossièrement l'ensemble des observations alors disponibles.

¹L'annexe A, p. 171, regroupe l'ensemble des notations utilisées dans ce chapitre

Ils doivent toutefois introduire également un halo de matière noire pour obtenir une courbe de rotation plate pour la Galaxie ;

- Caldwell & Ostriker (1981) présentent un modèle à trois composantes : disque, sphéroïde et halo de matière noire (HMN). Ce modèle comporte peu de paramètres, et décrit grossièrement la répartition de la masse dans la Galaxie ;
- Gilmore & Reid (1983) font intervenir un disque épais pour tenir compte de la distribution des étoiles à grande distance au dessus du plan, vers le pôle galactique.

On trouvera dans la deuxième section de l'article de Gilmore et al. (1989) les références de nombreux modèles avancés dans les années 1980. J'aborderai en détail dans la section 6.2 la description du modèle de Besançon, également apparu à cette époque.

Le foisonnement de modèles pendant cette décennie s'explique par les progrès considérables des capacités de calcul, avec les performances croissantes des ordinateurs, mais aussi par les progrès des observations (notamment avec les capteurs CCD), qui offrent toujours plus de données pour contraindre et affiner les modèles.

6.1.3 Les années 1990

Les comptages d'étoiles de plus en plus profonds dans de nombreuses directions, mais aussi les progrès considérables dans de nouveaux domaines spectraux (en particulier l'infrarouge, moins sensible à l'extinction, avec des missions spatiales comme IRAS ou COBE), donnent une vision de plus en plus profonde et détaillée de la Galaxie.

Les modèles prennent en compte la complexité observée en faisant intervenir de plus en plus de structures non axisymétriques, en particulier à basse latitude galactique.

Ainsi, le modèle SKY, développé par Wainscoat et al. (1992) pour rendre compte des observations du satellite IRAS dans l'infrarouge moyen, fait intervenir un disque exponentiel, un bulbe et un halo stellaire, mais également des bras spiraux (y compris le bras local) et un anneau moléculaire.

Ce modèle a par la suite été étendu jusqu'au domaine ultraviolet (Cohen, 1994), avec l'ajout d'autres composantes (ceinture de Gould, ...).

Les données COBE/DIRBE ont été utilisées par Drimmel & Spergel (2001) pour construire un modèle de la distribution de la poussière et des étoiles dans le disque de la Galaxie. Les auteurs font intervenir une structure spirale et tiennent compte du gauchissement du disque².

Les modélisations récentes des parties les plus internes de la Galaxie adoptent également des formes non axisymétriques pour le bulbe (Dwek et al., 1995; Freudreich, 1998).

²Les phénomènes de gauchissement et d'évasement seront décrits dans la section 6.3.

6.2 Le modèle de Besançon

6.2.1 Caractéristiques du modèle

C'est Crézé (1979) qui est à l'origine du **modèle de synthèse de populations stellaires de la Galaxie** (ou, plus simplement, modèle de Besançon) que je vais décrire dans cette section. Bien qu'ayant bénéficié d'améliorations successives depuis la toute première version (Robin & Crézé, 1986a), le modèle de Besançon est resté fidèle au principe sur lequel il a été construit :

- Décrire le contenu stellaire de la Galaxie en utilisant une approche **globale**, tenant compte à la fois :
 - de l'évolution stellaire ;
 - de l'évolution galactique ;
 - de la dynamique galactique.
- Fournir en sortie, outre les paramètres fondamentaux des étoiles (masse, gravité, ...), des valeurs (magnitudes, couleurs, mouvements propres, ...) **directement comparables** aux observations.

Cette approche fait appel à des résultats théoriques obtenus dans des domaines variés de l'astrophysique. L'évolution stellaire décrit la structure interne des étoiles, principalement conditionnée par leur masse, leur âge et leur composition chimique.

L'évolution galactique repose sur un scénario de formation des étoiles en relation avec la matière interstellaire. On décrit généralement cette évolution en termes de **fonction de masse initiale** (IMF) (Salpeter, 1955) et de **taux de formation d'étoiles** (SFR) (Schmidt, 1959). Les étoiles enrichissent pendant les phases finales de leur évolution le milieu interstellaire en éléments lourds, qui vont pouvoir être incorporés par les générations suivantes d'étoiles. L'évolution chimique du milieu interstellaire induit la relation âge—métallicité des étoiles.

Enfin, les théories dynamiques décrivent le mouvement des objets sous l'influence du potentiel galactique, et l'évolution des orbites avec le temps.

Dans un premier temps, un modèle empirique simple a été construit : on part de la distribution en magnitude et type spectral des étoiles proches, et on fait l'hypothèse que la fonction de luminosité au voisinage solaire est représentative de l'ensemble des étoiles de la Galaxie. On peut ainsi, en extrapolant avec des lois de densité simples, ajuster la distribution des étoiles en magnitude et couleur.

Le modèle s'inspire à l'origine d'un code développé par Rocca-Volmerange et al. (1981) pour l'étude de l'évolution photométrique et spectroscopique des galaxies externes (et appliqué aux nuages de Magellan).

Avec les lois de densité trouvées, Robin & Crézé (1986a) ont introduit un potentiel galactique (de type Oort) pour modéliser en première approximation la dynamique

galactique (et relier les échelles de hauteur à l'âge des populations). Cette première version du modèle a été validée par comparaison avec des comptages d'étoiles dans différentes directions (Robin & Crézé, 1986b). Le modèle ayant passablement évolué depuis, je vais en présenter les principaux ingrédients dans la suite de ce chapitre.

J'aborde dans les sections 6.2.2 et 6.2.3 la description des fonctions de luminosité et des lois de densité utilisées pour les différentes populations. La section 6.2.4 décrit l'auto-cohérence dynamique du modèle. Enfin, la section 6.2.5 détaille le processus itératif de fabrication des catalogues de pseudo-étoiles.

6.2.2 Fonctions de luminosité

Les populations

On a vu, au chapitre 5, les principales composantes de la Galaxie. Pour en simuler le contenu en étoiles, on retient 4 *populations*, qui se distinguent notamment par leur âge :

- le disque mince ;
- le disque épais ;
- le halo ;
- le bulbe.

On considère dans le modèle de Besançon que chaque population est homogène, et que la répartition des étoiles dans le diagramme de Hertzsprung-Russel (HR) pour une population i est donnée par une fonction

$$F_i(M_v, \log T_{\text{eff}}, \text{Age}). \quad (6.2)$$

Cette fonction est calculée d'après le modèle d'évolution de (Haywood, 1994a) pour le disque mince, d'après Bergbusch & Vandenberg (1992) pour le disque épais et le halo, et Bruzual & Charlot (1993) pour le bulbe.

On verra plus loin que le disque mince est subdivisé en 7 composantes d'âges différents. Le disque épais, le halo et le bulbe sont considérées comme des populations ayant une seule génération d'étoiles, en première approximation.

On simule l'évolution chimique en attribuant à chaque âge une métallicité (avec une dispersion autour de cette valeur). On utilise la relation âge—métallicité empirique de Twarog (1980), notée $\psi(Z, \text{Age})$ dans la suite.

Lorsqu'on parle de fonction de luminosité, on donne en général la distribution $\Phi(M_v)$ à une dimension, qui s'exprime en étoiles par unité de magnitude absolue et par parsec cube. On voit qu'elle se déduit simplement de la fonction F_i ci-dessus, en général normalisée au voisinage solaire.

Contraintes observationnelles

La fonction de luminosité $\Phi(M_v)$ est surtout connue pour le voisinage solaire. Ainsi, Wielen et al. (1983) donnent la fonction de luminosité à partir de comptages

d'étoiles dans une sphère de 20 pc centrée sur le soleil. Jahreiß & Wielen (1997) utilisent les résultats du satellite HIPPARCOS pour donner une fonction de luminosité plus précise, toujours dans un rayon de 20 pc autour du soleil.

Les études d'amas dans notre Galaxie, mais aussi dans quelques galaxies proches laissent penser que l'IMF a une forme universelle (Hunter et al., 1997). La fonction de luminosité actuelle s'obtient à partir de l'IMF en tenant compte de la relation masse–luminosité et du taux de formation d'étoiles (SFR). La relation masse–luminosité dépend de l'âge et de la métallicité des étoiles (Kroupa & Tout, 1997).

Les fonctions $F_i(M_v, \log T_{\text{eff}}, \text{Age})$ pour chaque population dans le modèle sont contraintes par les fonctions de luminosités $\Phi_i(M_v)$ observées pour des échantillons représentatifs de ces populations.

Le disque mince

La détermination de l'IMF, qui va conditionner la forme de $F_{\text{disque mince}}$, est décrite dans l'article de Haywood (1994b). Je rappelle brièvement ici le principe de la méthode.

On se donne une IMF suivant classiquement une loi de puissance³

$$\varphi(m) \propto m^{-(1+x)}. \quad (6.3)$$

L'IMF est normalisée de sorte que

$$\int_{m_{\text{min}}}^{m_{\text{max}}} \varphi(m) dm = 1, \quad (6.4)$$

avec $m_{\text{min}} = 0.1 M_{\odot}$ et $m_{\text{max}} = 120 M_{\odot}$.

On se donne également une composition chimique initiale du milieu interstellaire, un SFR $\Psi(t)$, et un jeu de tracés évolutifs (grâce aux modèles d'évolution stellaire).

Le nombre d'étoiles nées dans un sous-intervalle de masse $[m_{\text{inf}}, m_{\text{sup}}]$ pendant un temps $\Delta\tau$ est

$$N = \int_{m_{\text{inf}}}^{m_{\text{sup}}} \varphi(m) dm \int_{\tau}^{\tau+\Delta\tau} \Psi(t) dt. \quad (6.5)$$

Haywood a utilisé 33 tracés évolutifs différents, issus respectivement de VandenBerg pour les étoiles de 0.12 à 1.00 M_{\odot} , et de Schaller et al. (1992) pour des masses de 1.00 à 120 M_{\odot} . En interpolant (avec des intervalles de temps et de masse adéquats) le long de chaque tracé, ainsi qu'entre les tracés de masses successives, il est possible de trouver la position dans le diagramme HR de n'importe quelle étoile (de masse initiale et composition chimique donnée), à n'importe quel stade de son évolution.

³Dans le cas d'une loi de Salpeter, l'indice x est constant, et vaut 1.35. On pourra, par la suite, être amené à utiliser une loi de puissance en plusieurs parties, avec différentes valeurs de x selon l'intervalle de masse considéré.

On peut noter que le simple fait d'utiliser des tracés évolutifs suppose que l'on accepte implicitement le postulat énoncé par Heinrich Vogt et Henry Norris Russell dans les années 1920, disant que la masse initiale et la composition chimique d'une étoile déterminent de façon unique sa structure interne et son évolution.

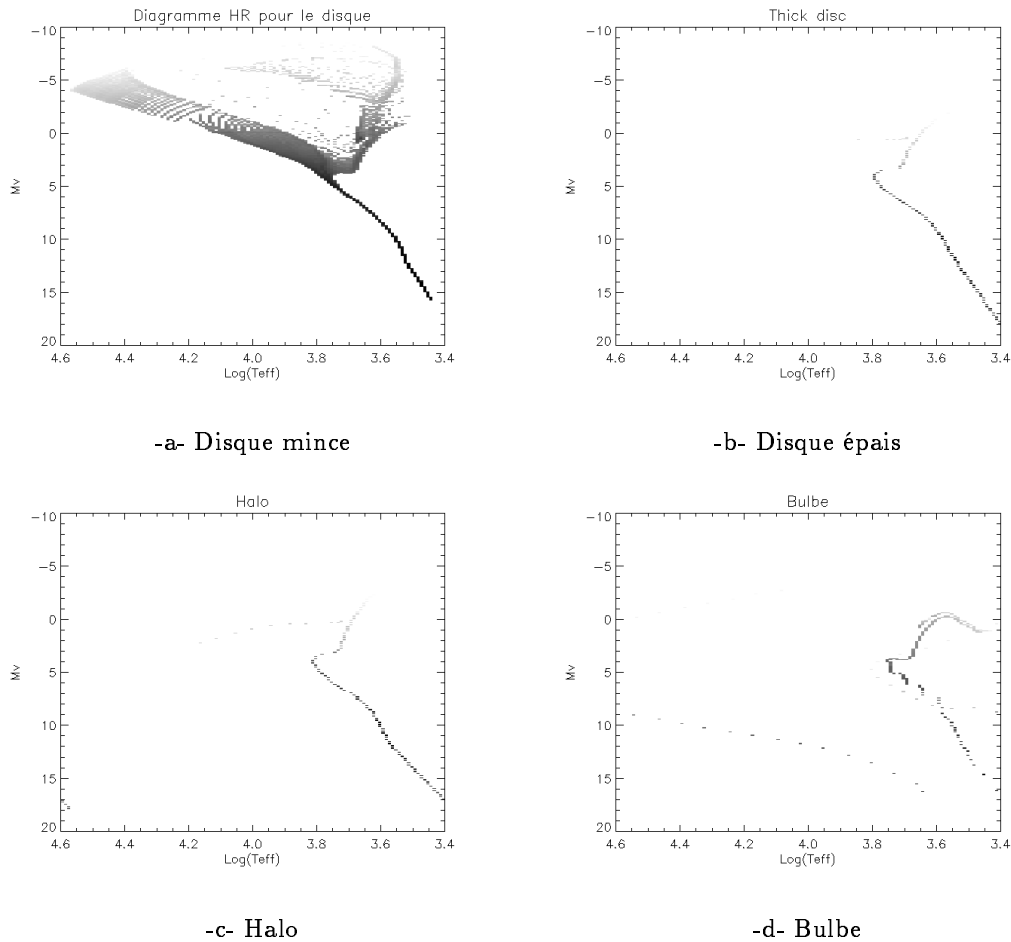


Fig. 6.1: Les différentes fonctions F_i utilisées dans le modèle pour les 4 populations stellaires : disque mince (a), disque épais (b), halo (c) et bulbe (d). Le niveau de gris est proportionnel au logarithme de la densité dans chaque cellule du diagramme.

Regardons maintenant l'évolution du disque. On tire à des intervalles de temps $\Delta\tau$ un nombre d'étoiles (équation 6.5) que l'on fait évoluer le long des tracés évolutifs. On peuple ainsi progressivement le diagramme HR. $\Delta\tau$ dépend de la masse des étoiles considérées, puisque les étoiles les moins massives évoluent beaucoup moins vite. Les étoiles en fin de vie donnent des résidus (naines blanches, étoiles à neutrons, trous noirs) et réinjectent du matériel dans le milieu interstellaire. Après un temps d'environ 10 milliards d'années pour le disque, on obtient un diagramme HR que l'on peut comparer aux observations.

Cette comparaison nécessite de convertir les températures effectives et luminosités (utilisées dans les modèles d'évolution stellaire) en types spectraux et magnitudes absolues. Haywood utilise pour cela les grilles de conversion de de Jager & Nieuwenhuijzen (1987) (type spectral – température), et les corrections bolométriques de Flower (1977) ou Flower (1975) selon le type spectral.

Haywood a utilisé comme contrainte observationnelle la fonction de luminosité de Wielen et al. (1983) pour les magnitudes absolues $4 \leq M_v \leq 12$, complétée pour les étoiles faibles par celle de Dahn et al. (1986).

Avec un SFR constant ($\Psi = 0.4(m_{\text{gaz}}/m_{\text{tot}})$), il parvient à ajuster les fonctions de luminosité observées avec une IMF d'indice $x \approx 0.7$ pour les objets de moins d' $1 M_{\odot}$. Pour les masses supérieures, on adopte un indice de 2.0 (Haywood et al., 1997a). On dispose ainsi d'une fonction $F_{\text{disque mince}}$ utilisable pour le modèle de Besançon (voir figure 6.1-a).

Autres populations

- On utilise pour décrire le **disque épais** les isochrones et fonctions de luminosité du modèle enrichi en oxygène de Bergbusch & Vandenberg (1992) pour les amas globulaires. Ce modèle est en effet utilisable pour l'amas 47 Tuc, considéré comme représentatif de la population intermédiaire.

On adopte un âge unique de 10 à $11 \cdot 10^9$ ans pour la population intermédiaire, et une métallicité de -0.7 dex, avec une dispersion de 0.25 dex.

La figure 6.1-b schématise la fonction de luminosité utilisée dans le modèle.

- Le **halo** est représenté par une population homogène d'âge $14 \cdot 10^9$ ans. L'évolution des étoiles retenue suit aussi le modèle de Bergbusch & Vandenberg (1992), avec une métallicité de -1.7 dex et une dispersion de 0.25 autour de cette valeur.

La fonction de luminosité est semblable à celle de l'amas globulaire M3 (voir par exemple les résultats de Rood et al. (1999) obtenus avec des observations HST), typique d'une population pauvre en métaux.

La figure 6.1-c schématise la fonction de luminosité utilisée dans le modèle.

- Le **bulbe** est représenté par une population de métallicité solaire de $10 \cdot 10^9$ ans, avec une IMF de Salpeter, suivant le modèle d'évolution de Bruzual & Charlot (1993) (voir figure 6.1-d).

Pour le disque épais et le halo, on ajoute une branche horizontale, qui ne figure pas dans le modèle de Bergbusch & Vandenberg (1992). On utilise celle décrite par Dorman (1992).

Naines blanches

Les naines blanches du disque ne sont pas calculées à partir du modèle d'évolution. On utilise une fonction de luminosité empirique (Sion & Liebert, 1977) avec

les modèles d'atmosphère de Bergeron et al. (1995), complétés par ceux de Hansen & Phinney (1998) pour la fin de la séquence de refroidissement des naines blanches.

Pour le disque épais, on utilise la fonction de luminosité calculée par Garcia-Berro (communication privée), tout comme Isern et al. (1998), en supposant une IMF de type Salpeter et un âge de 12 milliards d'années. Ces paramètres restent hautement spéculatifs car aucune naine blanche isolée du disque épais n'a encore été trouvée.

Des tests sont en cours pour mettre en évidence les naines blanches du halo, qui sont modélisées avec les mêmes modèles d'atmosphère, avec des hypothèses variables sur leur densité locale et sur leur fonction de luminosité.

6.2.3 Les lois de densité

Les lois de densité $\rho_i(r, \Theta, z, \text{Age})$ décrivent la répartition spatiale des différentes composantes dans la Galaxie. On peut en général choisir entre plusieurs formes dans le modèle de Besançon ; je ne présenterai ici que les lois que j'ai effectivement utilisées dans mes simulations.

On notera R le rayon galactocentrique et z la hauteur au dessus du plan galactique. On utilisera la variable a définie par $a^2 = R^2 + z^2/c^2$, où c est le rapport d'axes de l'ellipsoïde.

Le disque mince

Dans le scénario d'évolution dynamique de la Galaxie, les échelles de hauteur des étoiles dans le disque mince sont directement reliées à leur âge. Le découpage du disque mince en 7 composantes d'âges différents est décrit dans la section 6.2.4.

La composante la plus jeune (âge $\leq 0.15 \cdot 10^9$ ans) du disque mince suit une loi de la forme

$$\rho = \rho_0 \left[\exp\left(-\frac{a^2}{K_+^2}\right) - \exp\left(-\frac{a^2}{K_-^2}\right) \right]. \quad (6.6)$$

Cette composante est très aplatie puisque $c = 0.014$.

Les autres composantes du disque mince ($0.15 \cdot 10^9 < \text{âge} \leq 10^{10}$ ans) sont décrites par des exponentielles modifiées de type Einasto (1979)

$$\rho = \rho_0 \left\{ \exp\left[x - \left(x^{2N} + \frac{a^2}{K_+^2} \right)^{1/2N} \right] - \exp\left[x - \left(x^{2N} + \frac{a^2}{K_-^2} \right)^{1/2N} \right] \right\}. \quad (6.7)$$

x et N sont des paramètres structurels valant respectivement 0.5 et 1 dans notre cas. Le rapport d'axes c intervenant dans la variable a dépend de l'âge de la composante considérée, pour tenir compte de l'augmentation de l'échelle de hauteur pour les étoiles les plus vieilles. K_+ et K_- sont des échelles de longueur mesurées dans le plan.

Le disque épais

On adopte pour le disque épais une loi exponentielle radialement et verticalement, mais pour éviter une discontinuité de la pente à la traversée du plan, on complète le profil vertical par une parabole pour des hauteurs faibles ($|z| < h_c$), en imposant la continuité de ρ et de sa dérivée en $|z| = h_c$:

$$\rho = \rho_0 \exp\left(-\frac{R - R_\odot}{h_R}\right) \begin{cases} 1 - \frac{z^2}{2h_z h_c + h_c^2} & \text{pour } |z| < h_c \\ \frac{\exp(h_c/h_z)}{1+h_c/2h_z} \exp\left(-\frac{|z|}{h_z}\right) & \text{pour } |z| > h_c, \end{cases} \quad (6.8)$$

h_R et h_z étant respectivement les échelles de longueur et de hauteur. ρ_0 est la densité locale (l'intégrale de la fonction de luminosité).

Le halo

Le densité du halo décroît comme une loi de puissance, et on impose une densité constante pour les régions internes de la Galaxie (pour éviter qu'il n'y ait divergence) :

$$\begin{aligned} \rho &= \rho_0 \cdot a^{-\alpha} & \text{pour } a > 500 \text{ pc} \\ \rho &= \rho_0 \cdot 500^{-\alpha} & \text{pour } a < 500 \text{ pc} \end{aligned} \quad (6.9)$$

Les valeurs de l'aplatissement⁴ ($c = 0.7$) et de l'indice $\alpha = 2.6$ ont été déterminées par ajustement sur des comptages profonds dans plusieurs directions. Robin et al. (2000) ont récemment affiné la détermination de ces paramètres.

Le bulbe

On adopte pour le bulbe la loi de densité correspondant au modèle G2 proposé par Dwek et al. (1995).

Les régions qui seront considérées dans la suite de ce travail de thèse ont été choisies éloignées du centre galactique, car le modèle de Besançon est encore assez mal contraint pour le bulbe⁵.

Les paramètres du bulbe sont en cours d'ajustement, à l'aide des données DENIS (Chereul et al., en préparation).

6.2.4 Cohérence dynamique

Bienaymé et al. (1987) ont modifié la première version du modèle de Besançon (Robin & Crézé, 1986a) pour le rendre dynamiquement cohérent. Ils remplacent le potentiel de Oort initialement adopté par un potentiel obtenu par résolution de l'équation de Poisson avec le modèle de masse utilisé. Un des résultats de cet article

⁴On a toujours $a^2 = R^2 + z^2/c^2$.

⁵La présence d'une barre étendue (Hammersley et al., 2000), par exemple, n'est pas prise en compte.

a été de contraindre la densité de masse au voisinage solaire, et d'estimer ainsi la quantité de masse cachée présente localement dans le disque.

Le disque mince est décomposé en 7 composantes d'âges différents (Tab. 6.1), chaque âge étant associé à une dispersion des vitesses dérivée des travaux de Mayor (1974).

Age (10^9 ans)	σ_w ($\text{km}\cdot\text{s}^{-1}$)	
	Mayor	Haywood
0 — 0.15	6	6.0
0.15 — 1	10	8.0
1 — 2	14	10.0
2 — 3	18.5	12.1
3 — 5	23	14.9
5 — 7	25	17.9
7 — 10	25	21.0

Tab. 6.1: Les 7 composantes du disque mince.

Outre la composante stellaire des disques (mince et épais) et du halo, divers éléments sont pris en compte dans le modèle de masse de la Galaxie de Bienaymé et al. (1987) :

- un disque de matière interstellaire ;
- le bulbe, considéré comme une masse ponctuelle (seule son influence sur la dynamique au-delà d'un certain rayon est prise en compte) ;
- un halo de matière noire (HMN), nécessaire pour expliquer la forme de la courbe de rotation (voir p. 109) ;
- une composante hypothétique de masse cachée dans le disque (MCD).

Le procédé pour aboutir à un modèle dynamiquement cohérent est le suivant :

- 1— On part de lois de densité raisonnables pour le disque (on choisit une excentricité pour chaque composante) et la MCD, et on fixe une première estimation pour la masse du bulbe et la forme du HMN. Par résolution numérique de l'équation de Poisson

$$\nabla^2 \phi(r, z) = 4\pi G \sum_i \rho_i(r, z), \quad (6.10)$$

on obtient une première valeur du potentiel galactique $\phi(r, z)$. On peut déduire de ce potentiel la force radiale $K_r(r, z)$, et donc la vitesse circulaire à une distance r :

$$V_{\text{mod}}(r) = \sqrt{-K_r(r, z=0) \cdot r}. \quad (6.11)$$

- 2— On modifie ensuite la forme du HMN et la masse du bulbe de manière à ajuster la courbe de rotation galactique observée (Caldwell & Ostriker, 1981). Ceci conduit à un nouveau potentiel $\phi(r, z)$. On boucle sur les deux premières étapes pour obtenir une solution cohérente pour le bulbe et le HMN.

3— On injecte ensuite ce potentiel dans l'équation de Boltzmann sans collision

$$\sigma_w^2 \ln \frac{\rho(r, z)}{\rho(r, 0)} = -\phi(r, z) + \phi(r, 0) . \quad (6.12)$$

Cette équation n'est valable que pour une population stellaire isotherme et relaxée et n'est donc valable que pour les populations suffisamment vieilles. La composante la plus jeune du disque ne satisfait pas cette condition et le rapport d'axes a été fixé à $c = 0.014$. On joue sur les rapports d'axes des autres populations pour obtenir, avec le potentiel $\phi(r, z)$ trouvé à l'étape 2 et les relations âge – dispersion de la table 6.1, la validité de l'équation 6.12.

Ceci donne de nouvelles lois de densité, avec lesquelles on peut reprendre l'étape 2.

Le processus est répété jusqu'à convergence. On obtient ainsi un modèle *dynamiquement cohérent*, puisque le potentiel galactique obtenu est celui effectivement créé par les différentes composantes massives du modèle.

Bienaymé et al. (1987) ont étudié plusieurs modèles en faisant varier la composante de MCD. Ils sont arrivés à la conclusion que la densité locale de masse cachée dans le disque est, sinon nulle, très faible, inférieure à $0.03 M_\odot \cdot \text{pc}^{-3}$. Ils ont également déduit une densité de masse locale au voisinage solaire comprise entre 0.092 et $0.12 M_\odot \cdot \text{pc}^{-3}$.

Ces résultats peuvent être rapprochés de ceux de Crézé et al. (1998) qui trouvent, avec les données HIPPARCOS, une densité locale de $\rho_0 = 0.076 \pm 0.015 M_\odot \cdot \text{pc}^{-3}$, ne laissant pas de place à une masse cachée dans le disque⁶.

Haywood (1994a) a repris cette méthode avec le nouveau modèle d'évolution du disque, et revu à la baisse la relation âge–dispersion des vitesses (deuxième colonne du tableau 6.1).

C'est cette nouvelle relation qui sera utilisée par la suite.

6.2.5 Catalogues de pseudo-étoiles

Le modèle de Besançon ne simule bien évidemment pas avec exactitude chaque étoile de la Galaxie. Mais il peut produire en sortie des catalogues de pseudo-étoiles, c'est à dire des catalogues synthétiques d'étoiles dont les propriétés globales sont représentatives des objets dans une direction donnée, avec pour chaque étoile un ensemble de paramètres directement comparables aux observables (magnitudes, mouvement propres, ...).

L'organigramme de la figure 6.2 détaille le processus de fabrication d'un tel catalogue.

⁶Ce résultat est encore controversé : ils ont soustrait à la densité totale locale la contribution supposée des étoiles et de la matière interstellaire. Or il y a dans la Bulle Locale (100 pc autour du Soleil) un déficit de gaz, et la valeur utilisée pour la contribution de la matière interstellaire est une moyenne sur une portion du disque bien plus étendue. Il pourrait donc rester un peu de place pour de la matière noire dans le bilan local.

Les ingrédients

On définit avant tout les éléments suivants :

- la position du Soleil : R_{\odot}, z_{\odot} ;
- une direction d’observation (l, b) , ainsi qu’une surface (ou un angle solide ω) ;
- les lois de densité des différentes populations $\rho_i(R, \Theta, z, \text{Age})$, en accord avec les contraintes dynamiques définies à la section 6.2.4 ;
- la distribution moyenne de la métallicité pour chaque âge ainsi que les gradients galactiques ;
- les fonctions $F_i(M_V, \log T_{\text{eff}}, \text{Age})$, en accord avec les modèles d’évolution stellaire et galactique ;
- l’extinction interstellaire. On modélise cette extinction par un disque de matière absorbante, avec une loi de densité d’Einasto (1979), avec une échelle de hauteur de 140 pc, et une échelle de longueur de 5000 pc. On peut également introduire ponctuellement des nuages absorbants.

Le processus

On définit des éléments de distance δr (correspondant à des volumes δV), le long de la ligne de visée. Le barycentre de chaque case spatiale correspond à un point en coordonnées galactocentriques (R, Θ, z) .

Pour chaque pas en distance, on peut calculer un nombre théorique d’étoiles présent x , par population et par case du diagramme HR (et dans le cas du disque mince par tranche d’âge) dans cette case. On effectue un tirage aléatoire d’espérance x pour déterminer le nombre d’étoiles N de propriétés identiques qui vont être prises en compte⁷.

Le résultat

Pour chaque étoile, on tient compte de l’absorption calculée sur la ligne de visée, et on peut calculer magnitude, couleurs et métallicité⁸, en prenant bien soin d’introduire des dispersions pour compenser l’aspect discret du découpage en valeurs discrètes du diagramme HR.

Le catalogue de pseudo-étoiles peut contenir les propriétés physiques intrinsèques des objets (température effective, gravité, masse, etc. . .) et des valeurs directement comparables aux observables (magnitudes, couleurs, mouvements propres). Pour ces observables, le modèle simule donc les lois d’erreurs observées dans les mesures. Par exemple, les erreurs photométriques pour des catalogues issus d’observations avec

⁷Il convient de noter que x est un nombre décimal, qui est, dans l’immense majorité des cas, très inférieur à 1 (c’est en effet le nombre d’étoiles attendu dans *une* case du diagramme HR, pour *une* population, pour *un* pas en distance donné). On tirera donc le plus souvent 0 ou 1 étoile.

⁸La métallicité dépend de l’âge et de la position de l’étoile (pour tenir compte des gradients galactiques). Elle n’influe que sur la photométrie, par les effets de blanketing (Buser & Kurucz, 1992).

capteurs CCD sont bien approximées par des lois exponentielles. L'erreur sur la magnitude augmente exponentiellement pour les objets faibles. On en tient compte dans le modèle en introduisant une dispersion sur les magnitudes

$$\sigma_m = A + \exp(C \cdot m - B). \quad (6.13)$$

L'intervalle de valeurs possibles pour les observables est d'autre part nécessairement limité (limite de complétude pour les magnitudes, précision sur les mouvements propres). On ne conserve donc les étoiles simulées que si les observables se situent dans des limites prédéfinies.

Le catalogue produit est alors directement comparable aux observations.

6.3 Gauchissement et évasement

6.3.1 Description

Le gauchissement⁹

On entend par gauchissement (*warp*, en anglais) la déformation du disque d'une galaxie spirale, qui, au-delà d'un certain rayon galactocentrique, présente une torsion et s'éloigne du plan défini par les parties plus internes du disque, de façon symétrique (voir Fig. 6.3 et 6.4).

Briggs (1990) a tiré de l'analyse de 12 galaxies gauchies quelques caractéristiques décrivant le gauchissement :

- le gauchissement commence entre R_{25} (rayon galactocentrique où la magnitude B vaut 25 mag/arcsec²) et $R_{26,5}$ (rayon de Holmberg R_{H_0});
- la ligne de nœuds¹⁰ est à peu près droite jusque R_{H_0} ;
- la ligne de nœuds spirale légèrement dans le sens de la rotation au delà de R_{H_0} .

Les deux galaxies les plus proches, M31 et M33 sont gauchies. De nombreuses galaxies externes présentent également un disque gauchi (Bosma, 1991), et Sanchez-Saavedra et al. (1990), observant un gauchissement pour 42 des 86 galaxies de leur échantillon, en déduisent que le phénomène pourrait être commun à toutes les galaxies spirales (le gauchissement n'est en effet visible que sous certaines conditions d'inclinaison du disque).

Ce résultat penche en faveur d'une durée de vie longue pour le gauchissement d'un disque galactique. La question de savoir si le gauchissement est un phénomène transitoire ou de longue durée¹¹ reste cependant ouverte, et dépend du mécanisme à l'origine de la déformation. Plusieurs hypothèses pour expliquer le gauchissement sont en effet envisagées. On peut citer :

⁹On pourra se reporter à la revue de Binney (1992) pour plus de détails.

¹⁰On appelle ligne de nœuds l'axe de symétrie. Le long de cette ligne, le disque n'est pas déplacé (voir Fig. 6.4).

¹¹Le maintien d'un gauchissement sur de longues durées implique que le disque subisse un couple extérieur; de plus, comme le disque est en rotation, ce couple va entraîner une précession rigide du disque gauchi.

- la réponse du disque galactique à un champ magnétique extragalactique (Battaner & Jimenez-Vicente, 1998). Le mécanisme de gauchissement des disques d'accrétion en présence d'un champ magnétique est un phénomène bien étudié. Cette hypothèse pour expliquer le gauchissement des disques galactiques a été suggérée par Battaner et al. (1990) qui ont trouvé à grande échelle une corrélation entre les directions du gauchissement pour un échantillon de galaxies qu'ils attribuent à l'action des lignes du champ magnétique intergalactique¹² ;
- la rotation d'un disque « mal aligné » à l'intérieur d'un halo massif aplati (Toomre, 1983; Dekel & Shlosman, 1983; Sparke & Casertano, 1988). Dans ce cas, les axes de symétrie du disque et du halo ne sont pas alignés. Les parties internes du disque vont surtout être influencées par le disque lui-même, et donc présenter une géométrie coplanaire, alors que les parties externes du disque vont être sensibles au potentiel du halo, et donc avoir tendance à s'aligner dans son plan principal. Le gauchissement peut alors être vu comme une onde stationnaire dans le disque qui précesse autour de l'axe de symétrie du halo (Kuijken, 1991) ;
- un effet de marée, dû à des compagnons proches, et plus spécialement des galaxies satellites. Cet argument a été avancé très tôt (Burke, 1957; Kerr, 1957) pour expliquer le gauchissement de la Galaxie, avec comme principaux suspects les Nuages de Magellan. Mais les Nuages, seuls, sont trop lointains et trop peu massifs pour expliquer l'amplitude du gauchissement de la Voie Lactée (Hunter & Toomre, 1969). La prise en compte des perturbations du satellite sur un halo massif, qui peuvent amplifier l'effet du satellite seul, n'est pas convaincante dans le cas des Nuages de Magellan : García-Ruiz et al. (2000) trouvent que l'orientation de la ligne de nœuds du gauchissement induit serait décalée de 90 degrés par rapport à celle observée. La galaxie naine du Sagittaire apparaît alors comme un candidat potentiel pour un tel effet (Ibata & Razoumov, 1998).

L'évasement

L'évasement (*flare* en anglais) se traduit par une augmentation de l'échelle de hauteur du disque galactique à grand rayon galactocentrique (voir Fig. 6.4).

La densité surfacique étant plus faible à grande distance du centre galactique, la force de rappel K_z est moindre et le disque peut s'évaser. Mais l'échelle de hauteur peut être modifiée, sans que la loi de densité ne soit affectée (Schwarzkopf & Dettmar, 2000).

La Voie Lactée

On sait depuis relativement longtemps que la Galaxie présente un disque gauchi et évasé dans ses parties externes. Les observations dans la raie à 21 cm permettent

¹²L'observation d'une contrepartie stellaire rend cette hypothèse caduque car les étoiles ne sont pas sensibles à l'effet d'un tel champ magnétique, qui n'agit que sur le gaz.

d'observer en détails le gauchissement de la couche d'hydrogène neutre (Burton & Lintel Hekkert, 1986; Burton, 1988).

Le gauchissement commence approximativement à un rayon galactocentrique de 11 kpc. Le disque est incliné au dessus du plan moyen dans l'hémisphère nord, et en dessous dans l'hémisphère sud, la ligne de nœuds correspondant (à 10° près) à la direction Soleil – centre galactique (Diplas & Savage, 1991).

Le comportement est légèrement différent dans les deux hémisphères : au nord, l'amplitude du déplacement par rapport au plan augmente continûment, jusqu'à atteindre 3 kpc à une distance de 20 kpc. Dans l'hémisphère sud, la couche de gaz s'éloigne jusqu'à 1 kpc sous le plan avant de revenir quasiment dans le plan à 20 kpc (voir la Fig. 2 de Porcel et al. (1997), adaptée des données de Burton (1988)).

On observe le même genre de distribution pour la poussière interstellaire (Freudenreich et al., 1994), ainsi que pour le gaz ionisé (Azcarate et al., 1990).

La distribution des nuages moléculaires (May et al., 1997) met aussi en évidence un gauchissement et un évasement. A partir de sources IRAS nichées dans des nuages moléculaires, Wouterloot et al. (1990) ont quantifié l'évasement, qui commence à peu près au rayon (galactocentrique) du Soleil : l'échelle de hauteur augmente d'un facteur 2 entre 10 et 17 kpc. Cette valeur est comparable aux observations dans le gaz.

Les détections d'une contrepartie stellaire au gauchissement et à l'évasement observés dans la couche d'hydrogène neutre sont difficiles, car il faut trouver des traceurs du disque galactique sur de longues distances. Les étoiles de type O – B, très brillantes, ont permis de voir une composante stellaire (Miyamoto et al., 1988; Reed, 1996; Drimmel et al., 2000), mais sur des échantillons de taille limitée par la rareté de ces objets. On s'attend (a priori) à ce que les étoiles jeunes suivent le gauchissement du gaz dans lequel elles se forment. Ce n'est pas forcément le cas des étoiles vieilles : tout dépend de la stabilité dans le temps du gauchissement et du mécanisme qui le maintient.

On peut noter enfin le travail de Carney & Seitzer (1993), qui a utilisé la position décalée par rapport à $b = 0$ des étoiles du disque externe (supposé gauchi), pour étudier dans le domaine optique les propriétés du bord du disque.

6.3.2 Prise en compte dans le modèle

Dans sa version standard, le modèle de Besançon pouvait seulement simuler un disque plat et non évasé. J'ai donc été amené pendant cette thèse à introduire des modifications dans les lois de densité pour rendre compte du gauchissement et de l'évasement du disque.

Pour cela, je me suis inspiré des lois observées pour le gauchissement et l'évasement du gaz, et fait l'hypothèse que le disque stellaire (disque mince et disque épais) suit des lois de la même forme.

Le gauchissement

Pour le gauchissement, nous adoptons une représentation similaire à Porcel et al. (1997), où le disque est découpé en une série d'anneaux concentriques que l'on peut incliner par rapport à la ligne de nœuds (*tilted ring model*).

En adoptant le système de coordonnées galactocentriques (R, Θ, z) , le gauchissement intervient à partir d'un seuil R_{warp} , et se traduit par un déplacement vertical des anneaux

$$z_{\text{warp}} = z_c \cos(\Theta - \Theta_{\text{max}}) \quad (6.14)$$

où l'amplitude z_c du gauchissement augmente linéairement avec le rayon :

$$z_c = \gamma_{\text{warp}}(R - R_{\text{warp}}). \quad (6.15)$$

En accord avec les résultats de Burton (1988), on supposera que le Soleil se trouve sur la ligne de nœuds du gauchissement, et donc $\Theta_{\text{max}} = 90^\circ$.

L'évasement

Comme dans le travail de Gyuk et al. (1999), on rend compte de l'évasement du disque en introduisant un facteur k_{flare} dans les échelles de hauteur. Ce facteur permet de « diluer » le disque sur une plus grande hauteur, au-delà d'un certain rayon galactocentrique R_{flare} , l'évasement étant de plus en plus prononcé lorsque R augmente. On contrôle par un facteur γ_{flare} l'amplitude de l'évasement.

$$k_{\text{flare}} = \begin{cases} 1 & \text{pour } R < R_{\text{flare}} \\ 1 + \gamma_{\text{flare}}(R - R_{\text{flare}}) & \text{pour } R \geq R_{\text{flare}} \end{cases} \quad (6.16)$$

La modification des échelles de hauteur pour tenir compte de l'évasement (passage de h_z à $h_z \cdot k_{\text{flare}}$) implique un facteur correctif sur la loi de densité. En effet, il faut s'assurer que **la densité surfacique du disque n'est pas modifiée** par l'introduction de l'évasement. Une augmentation de l'échelle de hauteur de la loi de densité (pour reproduire l'évasement) sans correction de la densité en $z = 0$ se traduirait par une augmentation de la densité surfacique du disque. On impose donc que

$$\int_z \rho_{\text{sans évaselement}} = \int_z \rho_{\text{avec évaselement}} \quad (6.17)$$

Pour le disque mince, on reprend la loi de densité 6.7. Seul le terme en K_+ est important pour les rayons auxquels intervient l'évasement, et la condition 6.17 impose de multiplier la loi de densité par un facteur $f = 1/k_{\text{flare}}$ quand on introduit l'évasement.

La situation est un peu plus compliquée avec le disque épais. En imposant la condition 6.17 dans l'équation 6.8, le facteur correctif devient

$$f = \frac{(2k_{\text{flare}}h_z + h_c)(3h_z h_c + 3h_z^2 + h_c^2)}{(2h_z + h_c)(3k_{\text{flare}}h_z h_c + 3k_{\text{flare}}^2 h_z^2 + h_c^2)}.$$

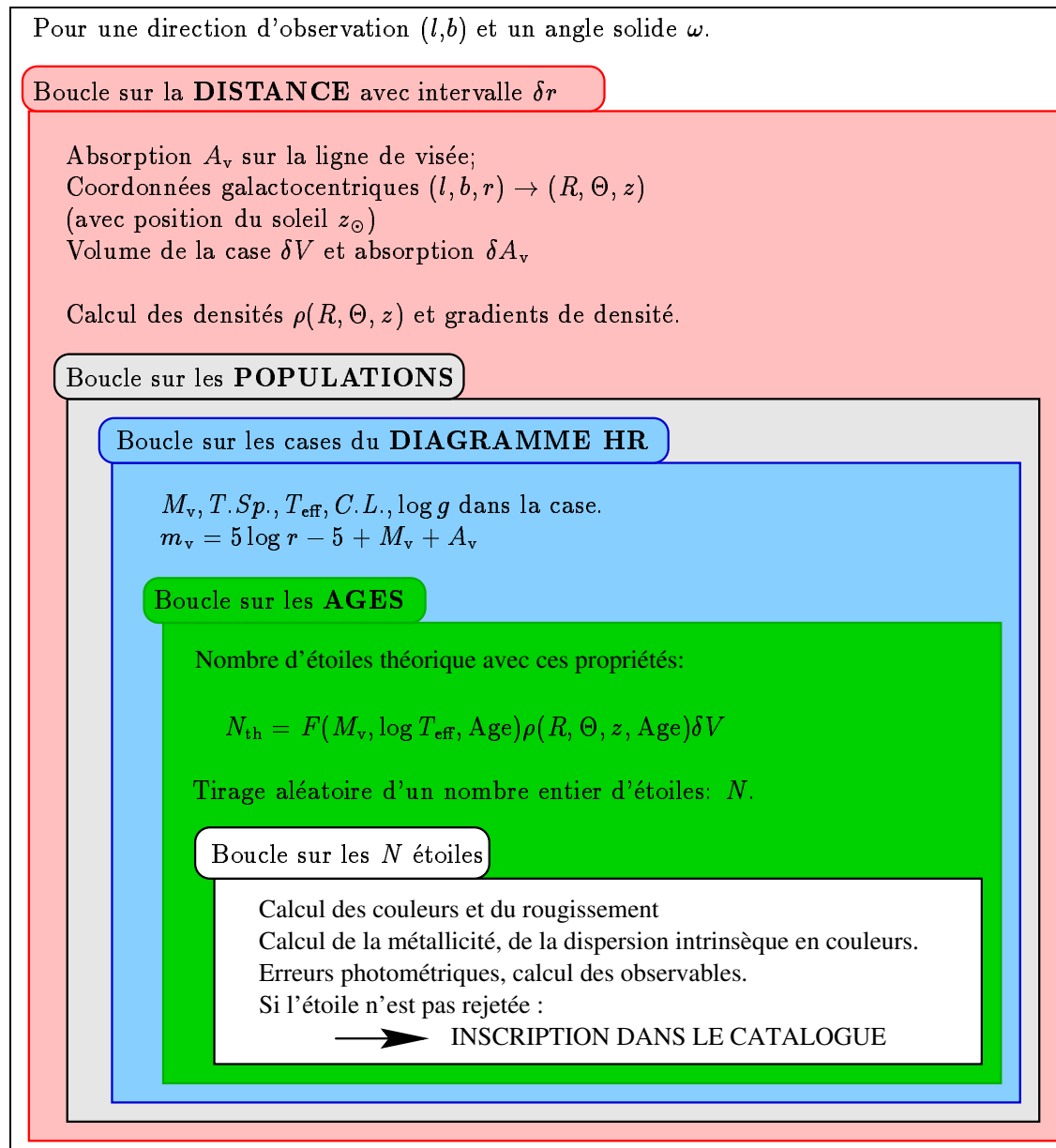


Fig. 6.2: Production d'un catalogue de pseudo-étoiles à l'aide du modèle de Besançon. Voir section 6.2.5 pour une description complète.

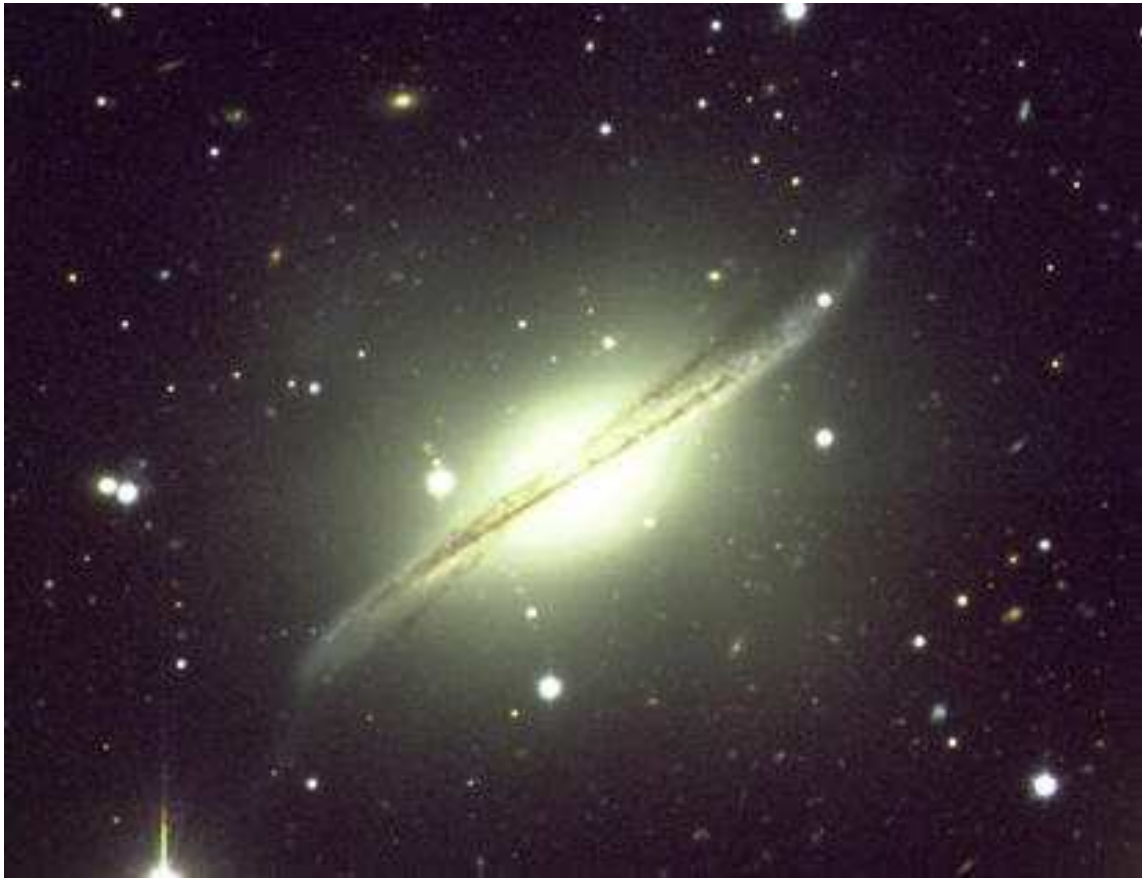


Fig. 6.3: *ESO 510-13 : une galaxie présentant un disque gauchi. Cliché FORS1/ESO.*

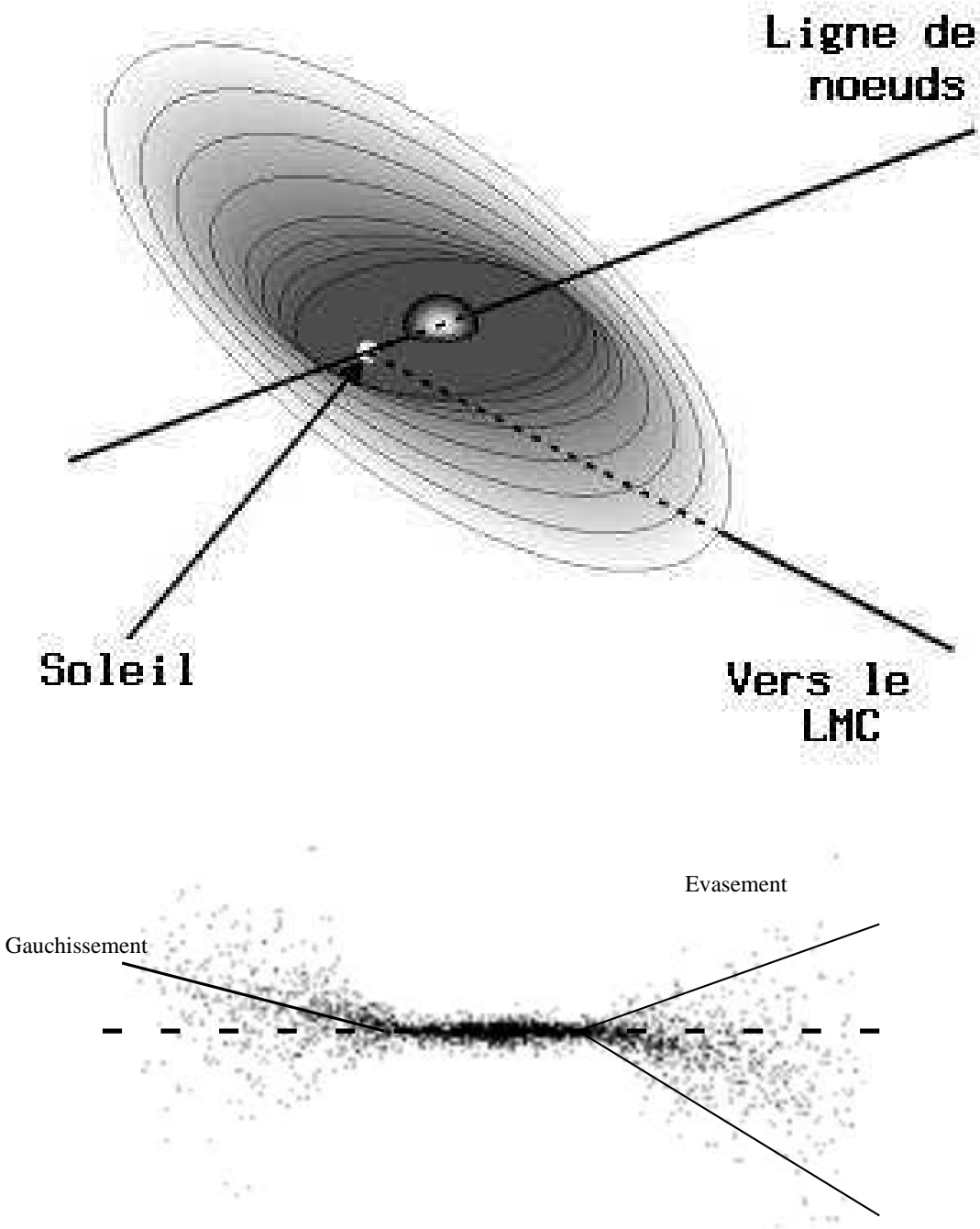


Fig. 6.4: La figure du haut montre l'effet d'un gauchissement du disque galactique (ici, pour un rayon $R_{\text{warp}} > R_{\odot}$). Le Soleil se situe sur la ligne de nœuds. Le plan moyen se trouve déplacé vers le haut ($z > 0$) pour les premier et deuxième quadrants, et vers le bas ($z < 0$) dans les troisième et quatrième. La figure du bas représente une coupe perpendiculairement au plan dans le cas d'un disque gauchi et évasé. Le gauchissement se traduit par un déplacement du plan moyen, et l'évasement par une dispersion sur des hauteurs plus grandes. Les figures sont adaptées de Evans et al. (1998).

Chapitre 7

Relevés IR et structure galactique

LES RELEVÉS dans l'infrarouge proche sont un support de premier choix pour l'étude de la structure galactique. Dans le domaine optique, les observations dans le plan galactique ou vers le bulbe sont limitées à des distances héliocentriques de 3 kpc environ, en raison de la forte extinction interstellaire à basse latitude (Schmidt-Kaler, 1975). Seules quelques rares régions, bénéficiant par chance d'une extinction moindre (la fenêtre de Baade, par exemple) permettent des sondages un peu plus profonds.

L'infrarouge proche présente deux avantages :

- l'extinction est considérablement plus faible que dans le visible. On gagne presque un facteur 10 à $2.2 \mu\text{m}$: $A_{K_s}/A_V = 0.11$ (Bessell & Brett, 1988) ;
- si on observe les distributions spectrales d'énergie des objets stellaires, on constate que le flux émis par des étoiles de types spectraux variés à ces longueurs d'onde reste important.

On peut donc détecter des étoiles jusqu'à de grandes distances, même à basse latitude galactique, où l'extinction pose problème dans le visible. Avec les sensibilités atteintes par les relevés DENIS et 2MASS (cf. Chapitre 1), les étoiles géantes peuvent être observées jusqu'aux frontières de la Galaxie là où l'extinction est faible, et jusqu'au-delà du bulbe vers le centre galactique.

Je présente dans ce chapitre les résultats d'ajustements du modèle de Besançon sur des données DENIS, et les résultats obtenus en particulier en termes de gauchissement et d'évasement du disque stellaire.

7.1 Sélection des données

Bénéficiant d'un accès privilégié aux catalogues du projet DENIS, suite au travail de codage et de mise en ligne, nous avons utilisé ces données pour faire les comparaisons avec le modèle, avant leur diffusion à l'ensemble de la communauté astronomique.

7.1.1 Les catalogues PDAC

Nous avons vu au chapitre 3 les caractéristiques des deux chaînes de traitement de données DENIS du LDAC et du PDAC.

Les comparaisons entre les réductions effectuées au PDAC et au LDAC (voir p. 78) indiquent que la qualité des catalogues produits au PDAC est supérieure à celle des catalogues LDAC. En particulier, la limite de complétude semble plus profonde en K_s au PDAC, ce qui est un argument de poids pour l'étude de la structure galactique. Nous avons donc utilisé exclusivement des données réduites au PDAC par la suite.

Ce choix est également motivé par la présence de problèmes occasionnels dans les catalogues du LDAC (comme les décalages de calibration astrométriques décrits p. 58), qui n'ont pas été observés dans les données PDAC. Enfin, la progression de la réduction de données se fait à un rythme plus soutenu au PDAC qu'au LDAC, ce qui devrait permettre d'obtenir plus rapidement une couverture du ciel plus large.

Filtrage

J'ai déjà décrit dans la section 3.4 (p. 62) les caractéristiques des données PDAC, et la nécessité d'appliquer un filtrage à la base de travail pour produire des catalogues exploitables.

Les critères de filtrage et d'appariement retenus sont ceux décrits p. 76.

7.1.2 Zones sélectionnées

Les observations DENIS couvrent uniquement la zone entre -88° et $+2^\circ$ en déclinaison. La figure 7.1 montre la surface correspondante en coordonnées galactiques.

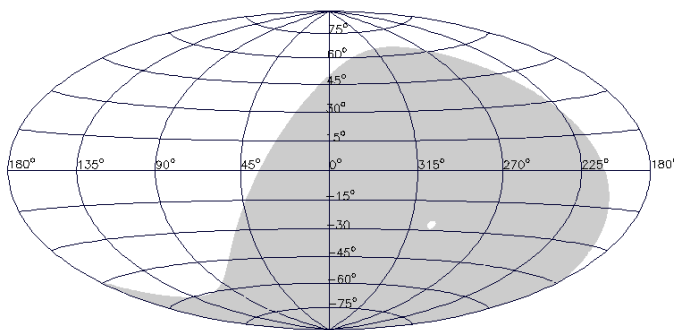


Fig. 7.1: Couverture du relevé DENIS en coordonnées galactiques (projection Aitoff). Les observations (zone grisée) couvrent notamment le centre galactique, le pôle galactique sud, et une grande partie des troisième et quatrième quadrants.

On notera parmi les régions couvertes tout le centre galactique et le pôle galactique sud. Le pôle galactique nord et l'anticentre, situés dans l'hémisphère nord, ne sont pas observables.

Afin d'étudier la structure galactique, on utilise classiquement un ensemble de zones distribuées à diverses longitudes et latitudes. Le tableau 7.1 contient la liste des zones que j'ai sélectionnées en tenant compte de plusieurs contraintes :

- couvrir un intervalle assez grand en longitude ;
- rester assez éloigné du centre galactique (car l’extinction peut être très importante, et la structure du bulbe est encore en cours d’ajustement dans le modèle) ;
- disposer d’un nombre de strips traités adjacents suffisamment grand pour extraire une zone circulaire continue qui contienne un nombre raisonnable de sources (quelques milliers). Ceci est surtout vrai à plus haute latitude ;
- vérifier que les strips concernés sont de bonne qualité ;
- enfin, vérifier qu’il n’y a pas de fluctuation de la densité de sources à l’intérieur de la zone. En effet, ceci peut indiquer des fluctuations de l’extinction à l’intérieur d’une zone, or nous voulons des zones homogènes.

Le choix des zones possibles étant également restreint par la quantité de strips observés et réduits au PDAC.

Zone	$l(^{\circ})$	$b(^{\circ})$	R ($''$)	Nb.	I	J	K_s
01A	293	+62	3000	6741	18.4	16.0	13.6
02A	325	-01	1500	48918	17.4	14.8	13.4
02B	327	+01	1500	36981	17.8	15.4	13.2
03C	274	-48	3600	11161	18.2	16.0	13.6
07A	270	-02	850	6552	18.0	15.8	13.4
07B	264	+02	900	5001	18.0	15.8	13.6
08A	240	-02	1100	10104	18.0	16.0	13.6
08B	231	+02	1100	9289	18.0	16.0	13.8
09C	270	-12	1150	6112	18.2	16.0	13.6
12A	295	-02	1000	15971	17.4	15.4	13.4
12B	294	+02	1100	17671	17.2	15.6	13.4
13C	272	+27	1500	4251	18.4	16.2	13.4

Tab. 7.1: Positions des zones sélectionnées, rayon et nombre total de sources retenues (toutes bandes confondues). Les limites de complétude adoptées dans les trois bandes figurent dans les trois dernières colonnes.

Un de nos objectifs est d’étudier le gauchissement et l’évasement du disque. Le tableau 7.2 résume les principales caractéristiques que l’on trouve dans la littérature pour ces deux effets, à partir d’observations dans différents domaines.

L’ensemble des études indique que le Soleil se trouve approximativement sur la ligne de nœuds ($\Theta_{\max} \approx 90^{\circ}$), et donc l’amplitude du gauchissement est négligeable vers $l = 0^{\circ}$ ou $l = 180^{\circ}$. C’est vers une longitude de 90° que l’amplitude du gauchissement est maximale dans l’hémisphère nord. Mais cette zone n’est pas observée par DENIS, et de plus la présence du bras spiral d’Orion-Cygnus sur la ligne de visée rendrait la détection du gauchissement difficile dans cette direction (Porcel et al., 1997).

Référence	Observations	Gauchissement	Évasement	Commentaires
Henderson et al. (1982)	HI (21 cm) $10^\circ < l < 350^\circ$ $-10^\circ < b < +10^\circ$	Maximum $z_c = -850$ pc à $R=17$ kpc vers $\Theta=260^\circ$ ($\gamma_{\text{warp}} \approx 0.12$)	h_z de 0.4 à 2 kpc	$R_\odot=10$ kpc, $V_\odot=250$ km·s ⁻¹
Burton & te Lintel Hekert (1986); Burton (1988)	HI (21 cm)	$\Theta_{\text{max}} \approx 80^\circ$	DéTECTÉ	$R_\odot=10$ kpc, $V_\odot=250$ km·s ⁻¹ Cartes de densité en HI
Grabelsky et al. (1987)	CO (traceur de H ₂) $270^\circ < l < 300^\circ$ $-1^\circ < b < +1^\circ$	$R=10.5$ kpc : $z_c = -46$ pc $R=12.5$ kpc : $z_c = -167$ pc ($\gamma_{\text{warp}} \approx 0.06$)	$R=10.5$ kpc : $h_z=112$ pc $R=12.5$ kpc : $h_z=182$ pc ($\gamma_{\text{flare}} \approx 0.31 \cdot 10^{-3}$)	$R_\odot=10$ kpc, $V_\odot=250$ km·s ⁻¹ Discussion sur le bras spiral de Carina
Miyamoto et al. (1988)	Nord, 518 étoiles OB	Plan incliné de 3° ($\gamma_{\text{warp}} \approx 0.05$), $\Theta_{\text{max}} \approx 90^\circ$, $R_{\text{warp}} > R_\odot = 8.5$ kpc	Non discuté	Aspect cinématique du gauchissement
Djorgovski & Sosin (1989)	Sources IRAS (étoiles vieilles)	Qualitativement identique à HI, mais amplitude plus faible	Non discuté	Soit le disque est tronqué vers $R=12$ kpc, soit la forme des gauchissements stellaire et gazeux est différente
Wouterloot et al. (1990)	≈ 1000 nuages moléculaires (IRAS)	$R_{\text{warp}}=11$ kpc ; sud : $z_c = -1$ kpc à $R=20$ kpc ($\gamma_{\text{warp}} \approx 0.11$)	Non linéaire ; h_z double entre $R=10$ et 17 kpc, mais $\times 9$ entre R_\odot et 25 kpc (γ_{flare} entre $0.1 \cdot 10^{-3}$ et $0.6 \cdot 10^{-3}$)	$R_\odot=8.5$ kpc, $V_\odot=220$ km·s ⁻¹ voir Fig. 9 pour la forme de l'évasement
Azcárate et al. (1990)	H166 α (H ionisé) $270^\circ < l < 300^\circ$ $-0.5^\circ < b < +0.5^\circ$	Vers $b < 0$ dans le sud	Non mentionné	

Tab. 7.2 (suite page suivante)

Référence	Observations	Gauchissement	Évasement	Commentaires
Diplas & Savage (1991)	HI (21 cm) $\Theta < 220^\circ$	Nord : $z_c = +4$ kpc pour $R = 24$ kpc	$h_z \approx 3$ kpc pour $R = 24$ kpc	$R_\odot = 10$ kpc, $V_\odot = 250$ km·s ⁻¹ Voir Fig. 7 pour les profils en R de z_c et h_z
Binney (1992)		$R_{\text{warp}} \approx R_\odot$, $\Theta_{\text{max}} = (90 \pm 10)^\circ$; nord : $z_c = +3$ kpc pour $R = 16$ kpc ; sud : z_c atteint -1 kpc, et redevient nul à $R = 15$ kpc	Non discuté	Revue des causes possibles
Freudenreich et al. (1994)	Poussières (DIRBE)	Qualitativement identique à HI, mais amplitude plus faible	Non discuté	$R_\odot = 8.5$ kpc, $V_\odot = 220$ km·s ⁻¹
Reed (1996)	1300 étoiles OB (hé- misphère sud)	Comparable à HI : $z_c = -1.5$ kpc pour $R = 15$ kpc ($\gamma_{\text{warp}} \approx 0.23$)	Non discuté	
May et al. (1997)	Nuages moléculaires (CO) $194^\circ < l < 270^\circ$	Maximum $z_c = -700$ pc	h_z de 200 à 800 pc entre $R = 9$ et 19 kpc	$R_\odot = 8.5$ kpc, $V_\odot = 220$ km·s ⁻¹
Porcel et al. (1997)	COBE, comparai- son avec HI de Burton	$\Theta_{\text{max}} = 90^\circ$; Sud : z_c de 200 à 800 pc entre $R = 10$ et 15 kpc ($\gamma_{\text{warp}} = 0.11$); Nord : de 200 à 2700 pc entre 12 et 20 kpc ($\gamma_{\text{warp}} = 0.31$); (voir Fig. 2)	Non pris en compte dans leur modèle	$R_\odot = 8.5$ kpc, $V_\odot = 220$ km·s ⁻¹ Un gauchissement identique au gaz est incompatible avec COBE : soit γ_{warp} est plus faible pour les étoiles, soit le disque stellaire est tronqué pour $R \approx 13$ kpc
Freudenreich (1998)	COBE (de 1.25 à 4.9 μm)	$\Theta_{\text{max}} = 90^\circ$, $R_{\text{warp}} = 4.3$ kpc ! Polynôme de degré 3 en R pour z_c (voir Fig. 12)	Pas d'évasement	$R_\odot = 8.5$ kpc ; $z_\odot = 16.5$ pc ; disque tronqué 4 kpc après R_\odot

Tab. 7.2 (suite page suivante)

Référence	Observations	Gauchissement	Évasement	Commentaires
Evans et al. (1998)	τ_{LMC} de MACHO	z_c de 0 à -1.5 kpc entre R_{\odot} et $R = 15$ kpc ($\gamma_{\text{warp}} \approx 0.21$)	Facteur 10 entre R_{\odot} et 25 kpc	$R_{\odot} = 8$ kpc ; les auteurs envisagent des h_R très élevés (jusque 5.5 kpc) pour augmenter la valeur de τ_{LMC}
Gyuk et al. (1999)	Comptages HST et τ_{LMC} $30^\circ < b < 40^\circ$	$R_{\text{warp}} = R_{\odot}$, $\Theta_{\text{max}} = 90^\circ$, $\gamma_{\text{warp}} \approx 0.21$	$\gamma_{\text{flare}} \approx 0.23 \cdot 10^{-3}$ à $0.54 \cdot 10^{-3}$; $R_{\text{flare}} = R_{\odot}$	$R_{\odot} = 8$ kpc, disque tronqué à 25 kpc ; même avec 25% de disque épais et $h_r = 4.5$ kpc, le disque gauchi et évasé ne peut pas expliquer seul la valeur de τ_{LMC}
Drimmel et al. (2000)	10544 étoiles OB de HIPPARCOS	$\Theta_{\text{max}} = 90^\circ$, $R_{\text{warp}} = 6.5$ kpc ; amplitude quadratique en R	Non discuté	La signature cinématique de ces étoiles n'est pas compatible avec un gauchissement à longue durée de vie
Drimmel & Spergel (2001)	COBE/DIRBE (étoiles/poussières)	Amplitude quadratique en R ; $R_{\text{warp}} = 0.85 R_{\odot}$ Étoiles : $z_c = 0.25$ kpc pour $R = 10$ kpc et 2.7 fois plus pour la poussière	h_z de 134 à 188 pc entre $R = 0.5$ et $1.0 R_{\odot}$	$R_{\odot} = 8$ kpc ; $h_R = 0.28 R_{\odot}$ $z_{\odot} = 14.5$ pc ; disque tronqué à 10.5 kpc

Tab. 7.2: Comparaison des caractéristiques du gauchissement et de l'évasement du disque de la Galaxie, relevées dans la littérature. La deuxième colonne indique le domaine d'observation, le type de données et éventuellement la zone sondée. Les deuxième et troisième colonnes résument les résultats obtenus pour le gauchissement et l'évasement, respectivement. Les variables utilisées sont celles intervenant dans les équations 6.14, 6.15 et 6.16 (voir p. 131). Dans la mesure du possible, j'ai estimé les amplitudes du gauchissement et de l'évasement en termes de γ_{warp} et γ_{flare} . La dernière colonne donne quelques indications supplémentaires (troncature du disque, par exemple). La détermination des distances dépend pour les études en H I d'un modèle de rotation, et j'ai donc signalé les R_{\odot} et V_{\odot} utilisés, le cas échéant.

Strip	Slot	α°	δ
7367	2858	134.2	I
8787	641	106.8	E
7368	3004	162.3	I
7238	652	108.6	E
8779	642	107.0	E
7752	2917	145.5	I

Tab. 7.3: Liste des strips DENIS sélectionnés, ascension droite centrale et zone de déclinaison (*I* : intermédiaire ; *E* : équatoriale).

Restent les troisième et quatrième quadrants galactiques, où le gauchissement se traduit par une inclinaison du disque à des latitudes négatives. La faible amplitude du gauchissement impose de rechercher ses effets à basse latitude. J'ai donc sélectionné plusieurs strips DENIS traversant le plan galactique à différentes longitudes pour contraindre les paramètres du gauchissement dans le modèle. Le tableau 7.3 donne la liste de ces strips.

La figure 7.2 montre la distribution en coordonnées galactiques des zones et des strips retenus (Tab. 7.1 et 7.3).

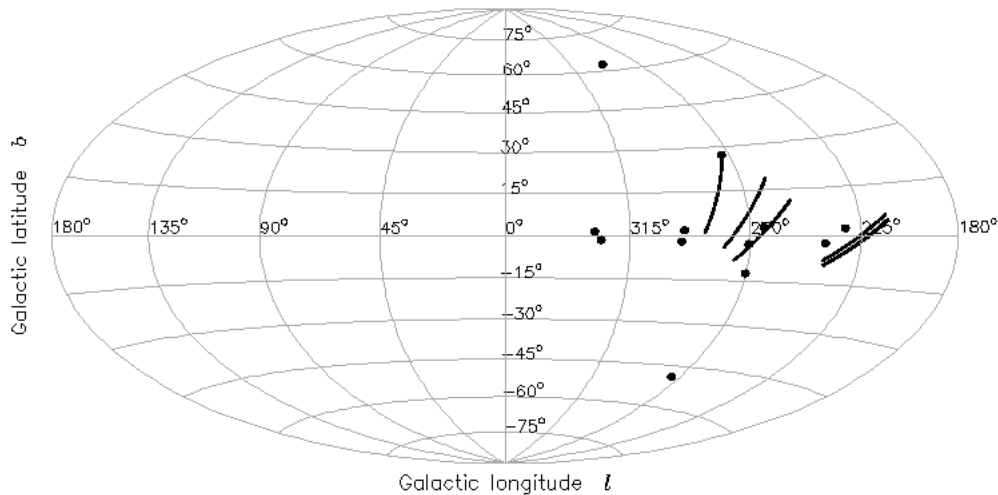


Fig. 7.2: Emplacement des zones et strips sélectionnés.

7.2 Premiers ajustements

L'utilisation des comptages d'étoiles pour l'étude de la structure galactique paraît intuitive : les modèles de Galaxie permettent de faire des prédictions de comptages

d'étoiles (comptages différentiels en magnitude, en couleurs, ...) qui sont directement comparables aux observations.

Cette approche a été largement appliquée avec succès (voir Bahcall (1986) pour une revue, mais aussi Robin & Crézé (1986b); Bienaymé et al. (1987); Wainscoat et al. (1992); Cohen (1994); Haywood et al. (1997b); Ruphy et al. (1997); Hammersley et al. (1999); Robin et al. (2000), ...). Dans la majorité des cas, cependant, les auteurs se restreignent aux hautes latitudes galactiques afin de minimiser les effets de l'extinction interstellaire sur les comptages.

7.2.1 Observables et estimateurs

Nous disposons avec les données DENIS de trois bandes photométriques (I , J , K_s), mais un objet n'est pas forcément détecté simultanément dans toutes les bandes (à cause de sa couleur et/ou des limites de détection). Pour que les catalogues simulés soient directement comparables aux comptages, nous devons les affecter des mêmes effets de sélection¹.

Limites de complétude

A partir de comptages différentiels en magnitudes, j'ai estimé les limites de complétude des 12 zones du tableau 7.1, dans chacune des trois bandes, afin de pouvoir effectuer une sélection identique sur les catalogues simulés par le modèle (voir Fig. 7.3)².

Il est toujours délicat de déterminer exactement la limite de complétude d'un échantillon. Dans le cas extrême d'un espace rempli d'une densité uniforme d'étoiles de même magnitude absolue, le logarithme du nombre d'étoiles détectées en fonction de la magnitude apparente augmenterait linéairement. Il suffirait alors de détecter à partir de quelle magnitude les comptages observés dévient de cette loi pour trouver la limite de complétude de l'échantillon.

En pratique, on traverse des zones de densité variable contenant des populations d'étoiles diverses. On ne connaît donc pas la véritable forme de la distribution du nombre d'objets en fonction de la magnitude. J'ai considéré, comme sur la figure 7.3 que la limite de complétude des échantillons se situait au point culminant de la distribution, juste avant le décrochement des comptages.

Erreurs observationnelles

Un point critique pour les comparaisons modèle–données concerne les erreurs observationnelles. Lors des simulations que nous effectuons avec le modèle de Besançon, les magnitudes des étoiles dans les différentes bandes sont calculées en utilisant des

¹Bien que le modèle utilise la bande K (et non K_s), les différences entre les deux bandes ne dépassent jamais 0.05 mag (Loup, communication privée), et ne jouent pas sur nos résultats.

²Les zones considérées ici se situant majoritairement à basse latitude galactique, la contamination attendue des comptages stellaires par des détectations de sources extragalactiques reste négligeable, et n'a pas été prise en compte dans la suite

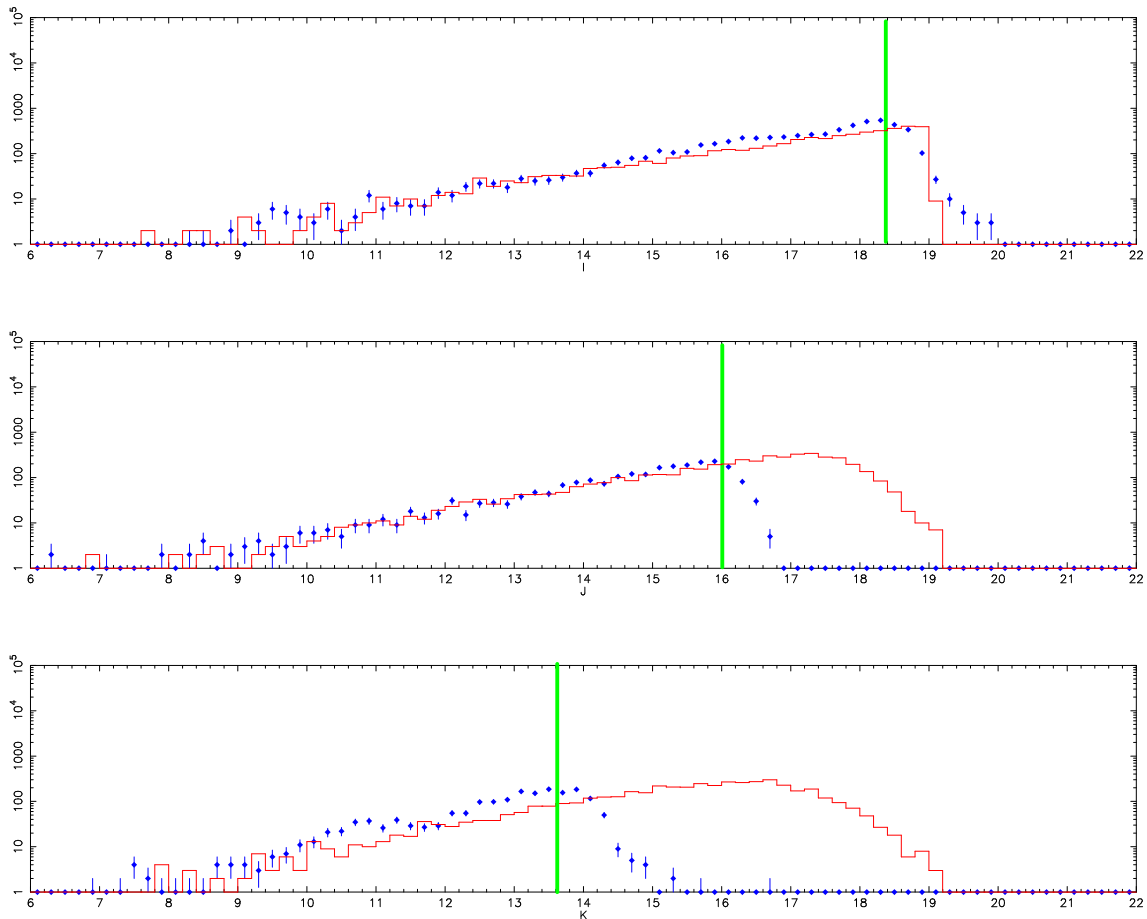


Fig. 7.3: Détermination de la limite de complétude dans les trois bandes pour la zone 01A. Les symboles correspondent aux comptages différentiels (échelle logarithmique!) des données DENIS (avec un pas de 0.2 mag) pour les objets détectés en I (haut), J (milieu) et K_s (bas). Le trait continu correspond à un fichier simulé (avant ajustement). Les limites adoptées sont indiquées par les barres verticales.

librairies de modèles stellaires (Lejeune et al., 1997; Lejeune et al., 1998). Nous avons vu p. 128 la simulation des erreurs de mesure dans les catalogues de pseudo-étoiles.

Le choix des paramètres A , B et C de l'équation 6.13 peut avoir de grandes répercussions sur les résultats. Si l'on effectue une simulation avec une loi d'erreur sous-estimée, la dispersion dans le diagramme couleur-couleur du catalogue simulé, par exemple, sera bien moindre que celle présente dans les données.

En toute rigueur, chaque strip DENIS possède sa propre loi d'erreur pour chacune des bandes. Ceci est valable même pour deux strips adjacents, généralement observés au cours de nuits différentes.

Dans notre cas, nous avons pris soin de ne pas inclure dans nos échantillons de strips présentant une loi d'erreur anormale, et des paramètres moyens ont pu être déterminés pour les lois d'erreur. Les formules retenues sont :

$$\begin{cases} \sigma_I &= 0.03 + \exp(0.5 \cdot I - 12.0) \\ \sigma_J &= 0.06 + \exp(0.7 \cdot J - 14.0) \\ \sigma_{K_s} &= 0.07 + \exp(0.8 \cdot K_s - 13.5) \end{cases} \quad (7.1)$$

La dispersion entre les différents strips ne dépasse pas 50% de σ .

On prend soin de tronquer les lois d'erreur à 0.5 mag pour ne pas introduire de dispersion trop importante (les lois d'erreur du PDAC sont, elles aussi, tronquées).

Estimateurs

Pour les comparaisons modèle-données, j'ai utilisé 7 observables :

- les comptages différentiels en magnitude dans chacune des 3 bandes ;
- les histogrammes en couleur $[I - J]$ et $[J - K_s]$;
- le diagramme couleur-magnitude $[K_s, J - K_s]$;
- le diagramme couleur-couleur.

Pour chaque observable, la qualité des ajustements est mesurée par trois estimateurs :

- un χ^2 classique ;
- la déviation σ_{χ^2} , obtenue avec l'approximation de Wilson-Hilferty pour un χ^2 à ν degrés de liberté (si ν n'est pas trop petit) :

$$\sigma_{\chi^2} = \left[\left(\frac{\chi^2}{\nu} \right)^{1/3} + \frac{2}{9\nu} - 1 \right] \left(\frac{9\nu}{2} \right)^{1/2} ; \quad (7.2)$$

- une vraisemblance comme celle définie dans l'annexe C de Bienaymé et al. (1987).

Le χ^2 simple n'est pas un très bon estimateur dans notre cas, en particulier dans le cas des histogrammes de complétude : la très grande différence de peuplement entre les intervalles correspondant aux faibles et aux grandes magnitudes introduit un biais, et, de plus, les intervalles ne sont pas vraiment indépendants puisqu'ils sont corrélés par les incertitudes de mesure.

7.2.2 Paramètres de densité

Comme nous l'avons vu au chapitre 6, un grand nombre de paramètres interviennent dans le contrôle des lois de densité dans le modèle de Besançon. Il n'était pas question de re-déterminer l'ensemble de ces paramètres, et j'ai donc adopté, pour certains, des valeurs déterminées de façon fiable par des études antérieures.

R_{\odot} et z_{\odot}

J'ai conservé pour le Soleil une distance galactocentrique $R_{\odot} = 8.5$ kpc, valeur recommandée par l'UAI.

On sait depuis longtemps que le Soleil ne se situe pas exactement dans le plan galactique ($z = 0$), mais légèrement au-dessus. Les valeurs de z_{\odot} vont de 4 à 40 pc selon les auteurs (voir Cohen (1995) pour une revue), mais les résultats les plus récents penchent pour une valeur de 15 pc environ (Cohen, 1995; Drimmel & Spergel, 2001). C'est cette valeur que nous adopterons dans la suite.

Il faut noter que le choix de z_{\odot} n'est pas dénué de conséquence pour une étude du gauchissement du disque. Porcel et al. (1997) ont montré qu'une élévation de 15 pc produisait un décalage de 0.2° sur les profils en latitude, pour toutes les longitudes galactiques.

Normalisations

Les proportions d'étoiles de magnitude absolue $M_v < 8$ retenues sont, par rapport au disque mince, de 5.6% pour le disque épais (Robin et al., 1996)³, et 0.07% pour le halo. On adopte un aplatissement de 0.7 pour le sphéroïde (Robin et al. (2000) trouvent $c = 0.76$).

Nous avons seulement deux régions à haute latitude, ce qui ne nous permet pas de contraindre les paramètres du halo. Par contre, les comptages à plus basse latitude excluent une fraction de disque épais supérieure à 7%, avec des échelles de longueur $h_R = 2800$ pc et de hauteur $h_z = 760$ pc pour cette composante.

Gauchissement, évaselement

On peut remarquer dans le tableau 7.1 qu'on a systématiquement, pour des longitudes comparables (entre 230° et 330°), des densités de sources plus grandes dans les zones situées à $b < 0$ que dans les zones où $b > 0$. C'est bien le genre d'effet attendu si le disque stellaire est gauchi vers les latitudes négatives.

On considérera par la suite que le Soleil se trouve sur la ligne de nœuds, et donc $\Theta_{\max} = 90^{\circ}$. Dans un premier temps, on adopte pour les amplitudes γ et rayons de départ R du gauchissement et de l'évasement, des paramètres comparables à ceux trouvés dans la littérature (Gyuk et al., 1999) : $\gamma_{\text{warp}} = 0.18$, $\gamma_{\text{flare}} = 0.54 \cdot 10^{-3}$, $R_{\text{warp}} = 10$ kpc, $R_{\text{flare}} = 12$ kpc. Ces paramètres seront ajustés par la suite.

Disque mince

Deux paramètres concernant le disque mince nous intéressent particulièrement : l'échelle de longueur h_R et le rayon de coupure du disque R_{cut} .

Les valeurs proposées pour l'échelle de longueur exponentielle du disque mince sont très variées : de 1 à 6 kpc (Kent et al., 1991). Cependant, les estimations récentes

³Une nouvelle densité de $7.1 \cdot 10^{-4} M_{\odot} \cdot \text{pc}^{-3}$, soit 6.2% du disque mince, vient d'être déterminée pour le disque épais par Reylé & Robin (2001).

semblent converger vers une valeur plus faible : 2.5 kpc pour Robin et al. (1992) et Fux & Martinet (1994), 2.3 kpc pour Ruphy et al. (1996) et Drimmel & Spergel (2001), 2.1 kpc pour Porcel et al. (1998).

Le disque stellaire apparaît moins étendu que le disque gazeux, mais la distance galactocentrique à laquelle se situe la troncature reste mal définie.

Comme on peut le voir dans le tableau 7.2, les auteurs situent en général R_{cut} entre 12 et 14 kpc. Ruphy et al. (1996) trouvaient $R_{\text{cut}} = 15 \pm 2$ kpc avec les premières données DENIS.

On peut cependant remarquer que ces valeurs faibles sont toujours obtenues sans tenir compte de l'effet d'un évasement du disque. Gyuk et al. (1999), qui font intervenir un évasement à partir de R_{\odot} , utilisent $R_{\text{cut}} = 25$ kpc. Seul Drimmel & Spergel (2001) dans un travail récent, trouve un disque tronqué très près ($R_{\text{cut}} = 10.5$ kpc, avec $R_{\odot} = 8$ kpc) avec un évasement.

Avec des paramètres standard pour le gauchissement et l'évasement, un ajustement simultané des diverses observables présentées p. 144 sur l'ensemble des zones nous donne une échelle de longueur $h_R = 2.5$ kpc et une coupure du disque $R_{\text{cut}} = 14$ kpc.

Les diagrammes de complétude, seuls, ne sont pas très contraignants, car ils contiennent une information très dégénérée. Les histogrammes en couleur ($I - J$ et $J - K$) semblent être les observables les plus sensibles, que la qualité de l'ajustement soit mesurée par un χ^2 ou par maximum de vraisemblance, les deux estimateurs donnant le même couple (h_R, R_{cut}).

Une étude sur un plus grand nombre de régions, et surtout une meilleure estimation de l'extinction au cas par cas à ces basses latitudes permettra d'affiner les déterminations de h_R et R_{cut} , et d'estimer les incertitudes sur ces paramètres.

En effet, lors des ajustements, cinq des zones (02A, 02B, et dans une moindre mesure 07A, 07B, et 09C) présentaient systématiquement de très mauvaises valeurs des estimateurs, quels que soient les paramètres utilisés. Visiblement, l'extinction dans ces régions était supérieure à celle simulée par le modèle.

Je décris dans la section suivante une méthode permettant de déterminer le profil de l'extinction sur la ligne de visée. Avec les valeurs de l'extinction déduites, les ajustements sur ces zones sont du même ordre que pour les autres zones.

7.3 Effets de l'extinction

Nous avons vu p. 127 que l'extinction interstellaire est modélisée par un disque très mince ($h_z = 140$ pc, $h_R = 5$ kpc), avec une normalisation au voisinage solaire $A_v = 0.7$ mag/kpc.

Ce modèle est largement valide à haute latitude, mais peut s'avérer trop simplificateur dans le plan, où des structures non axisymétriques (comme les bras spiraux) ou des surdensités locales sont présentes. La distribution de la poussière n'est pas homogène, et donc un modèle d'extinction diffuse ne peut donner qu'une valeur moyenne de l'extinction.

On peut certes ajouter des nuages absorbants sur la ligne de visée pour reproduire les effets de ces structures, mais cette démarche reste très empirique.

J'ai tenté d'automatiser la reconstruction du profil d'extinction le long de la ligne de visée. Puisqu'on dispose, dans les catalogues générés par le modèle, de la distance de chaque étoile, on peut essayer de placer de façon itérative des « grumeaux » d'extinction à différentes distances, et observer l'effet produit sur diverses observables.

J'utilise les valeurs de Bessell & Brett (1988) : $A_I = 0.600A_v$, $A_J = 0.288A_v$, $A_K = 0.109A_v$.

Le principe est simple : on se donne une observable (un histogramme en couleur par exemple), et on ajoute progressivement, dans le catalogue issu d'une simulation du modèle, des grumeaux absorbants $\delta A_v(r)$ jusqu'à ce que le modèle modifié ajuste correctement les observations.

Chaque élément absorbant $\delta A_v(r)$ agit sur toutes les étoiles du catalogue situées au-delà de r . Par exemple, si une étoile située à une distance r' possède une magnitude K_0 avant modification de l'extinction, celle-ci devient après ajout de plusieurs éléments :

$$K = K_0 + \sum_{r < r'} 0.109 \cdot \delta A_v(r).$$

Des éléments $\delta A_v(r)$ de taille variable sont placés itérativement à différentes distances jusqu'à ce qu'il y ait convergence d'un estimateur.

Validation

Afin de valider la méthode, je suis parti d'un catalogue simulé, avec une loi d'extinction standard. J'ai ensuite produit un fichier de données à partir de ce catalogue, en faisant subir une absorption de $A_v = 1.3$ mag à toutes les étoiles situées à plus de 2.23 kpc. Cet ajout simule un nuage ponctuel sur la ligne de visée.

La figure 7.4 montre la distribution $A_v(r)$ reconstruite en tentant d'ajuster le modèle original sur le fichier de données de test, par la méthode itérative, en minimisant σ_{χ^2} (défini par l'équation 7.2).

Seule la distribution reconstruite à partir de la densité dans le plan $[I-J], [J-K_s]$ est représentée sur la figure, mais l'utilisation des distributions en $[I-J]$ ou en $[J-K_s]$ donne sensiblement le même résultat, dans ce cas simplifié.

On voit que même si l'amplitude de l'élément absorbant diffère légèrement de celui introduit, la méthode permet bien de retrouver l'emplacement du nuage. Des simulations avec plusieurs nuages à diverses distances donnent également des résultats encourageants. La méthode permet bien de retrouver la forme des zones d'absorption, à condition toutefois que les nuages ne soient pas situés trop loin (il faut évidemment qu'il reste un nombre d'étoiles détectées derrière le nuage suffisant pour produire un effet mesurable dans la distribution totale sur laquelle se fait l'ajustement).

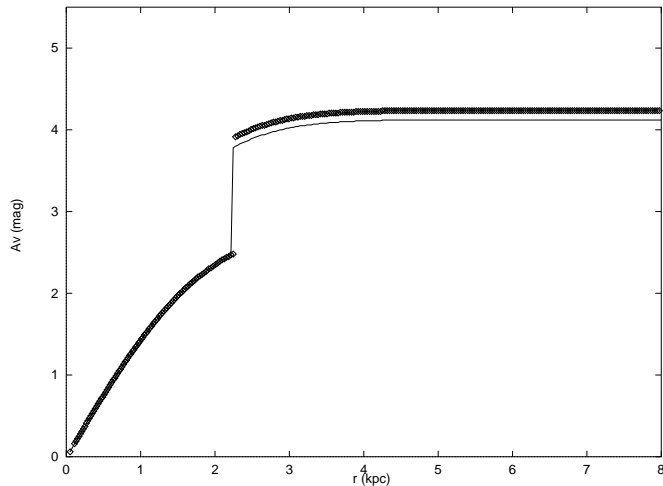


Fig. 7.4: Reconstruction du profil d'extinction dans le cas simple où un seul nuage a été ajouté sur la ligne de visée. Les symboles représentent le fichier test fabriqué, et le trait plein la reconstruction à partir du diagramme $[I - J], [J - K_s]$.

Application pratique

J'ai ensuite appliqué cette méthode à l'une des zones (07B) souffrant vraisemblablement d'une extinction assez forte. J'ai utilisé successivement trois observables pour effectuer l'ajustement : les histogrammes en $[I - J]$, $[J - K_s]$, et un histogramme à deux dimensions dans le plan $[I - J], [J - K_s]$.

La figure 7.5 présente, d'une part, l'histogramme en $[J - K_s]$ avant et après reconstruction de l'extinction, et d'autre part le profil $A_v(r)$ reconstruit à partir des trois observables.

Le profil reconstruit diffère selon que l'on utilise l'une ou l'autre des observables. Cela tient en partie à la méthode d'ajustement employée et à la sensibilité différente à l'extinction dans les 3 bandes.

C'est la bande I qui est la plus sensible à l'extinction. Sur une zone fortement éteinte, on réduira fortement le χ^2 en plaçant les premiers grumeaux relativement près, ce qui peut expliquer que le profil reconstruit à partir de $[I - J]$ montre une augmentation de $A(r)$ assez rapide. A l'inverse, $[J - K_s]$ ne devient proportionnellement sensible que dans les zones fortement éteintes, et le pic d'extinction se situe plus loin pour cette observable. On a une situation intermédiaire pour l'ajustement sur $[I - J], [J - K_s]$.

On note une brusque augmentation (4 magnitudes visibles) de l'extinction à une distance héliocentrique de 8 à 10 kpc. Or, on trouve, dans cette direction ($l=264^\circ$), le bras spiral de Persée à une distance d'environ 9.3 kpc (Vallee, 1995; Drimmel & Spergel, 2001). La brusque montée de A_v est donc très certainement associée à la présence de matériel associé au bras spiral à cet endroit.

L'utilisation de cette méthode dans des régions adjacentes (à condition que le modèle décrive correctement les populations, bien sûr) pourrait permettre d'obtenir une cartographie tridimensionnelle de l'extinction.

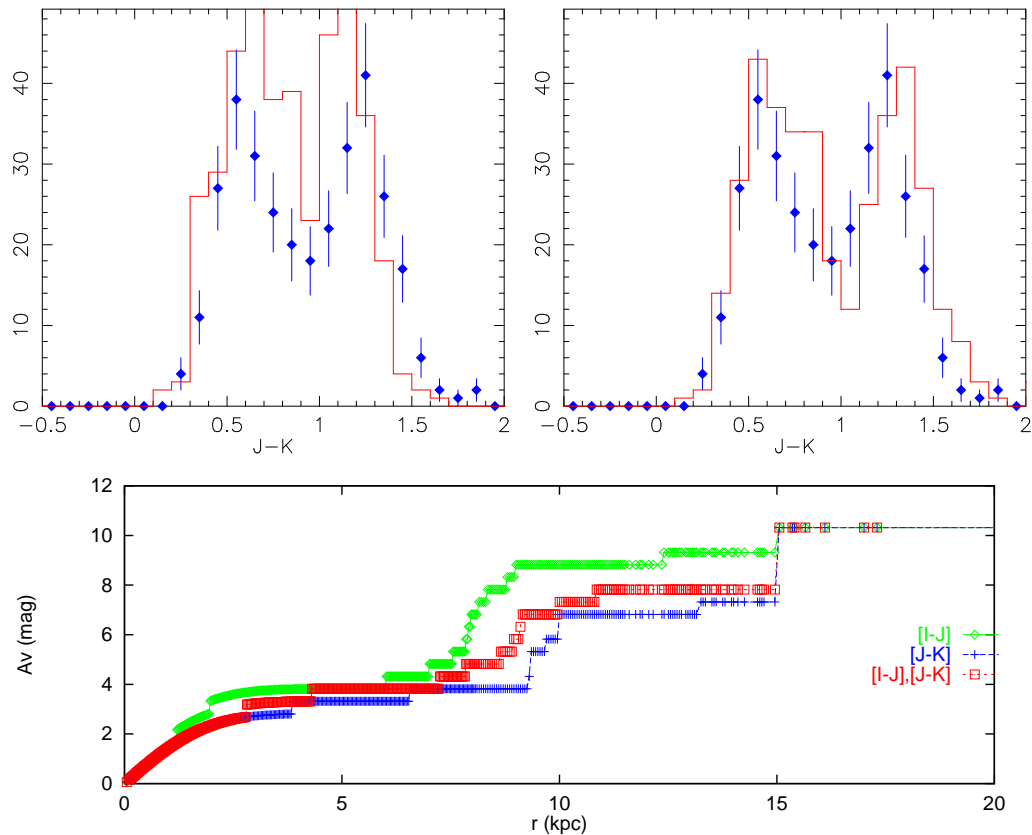


Fig. 7.5: Détermination du profil d'extinction cumulée sur la ligne de visée $A_v(r)$ pour la zone 07B. Les histogrammes du haut représentent les comptages observés (ronds) et simulés (trait plein) avant (gauche) et après (droite) modification de l'extinction, pour les sources avec $12 < K_s < 13$. Le graphique du bas présente le profil reconstruit pour trois observables.

7.4 Gauchissement et évaselement⁴

La figure 7.6 illustre les effets de l'introduction d'un gauchissement et d'un évaselement du disque sur un profil en latitude galactique, traversant le plan en $b = 266.1^\circ$.

Dans les trois graphiques, les carrés représentent les mêmes comptages DENIS (on ne garde que les sources détectées au moins en K_s , sous la limite de complétude, afin de réduire la sensibilité à l'extinction).

Les traits pleins correspondent à différentes modélisations. Le graphique du haut correspond à un disque normal. La distribution correspondante est centrée sur $b = 0^\circ$, ce qui n'est pas le cas des données pour lesquelles le maximum se situe à environ -2.5° .

Dans le deuxième graphique, on introduit un gauchissement du plan. Le pic de la distribution modélisée se trouve décalé vers $b < 0^\circ$.

⁴Voir également le compte-rendu d'une présentation orale sur ce sujet en annexe, p. 210. Un article plus détaillé est en préparation (Derrière et al., 2001).

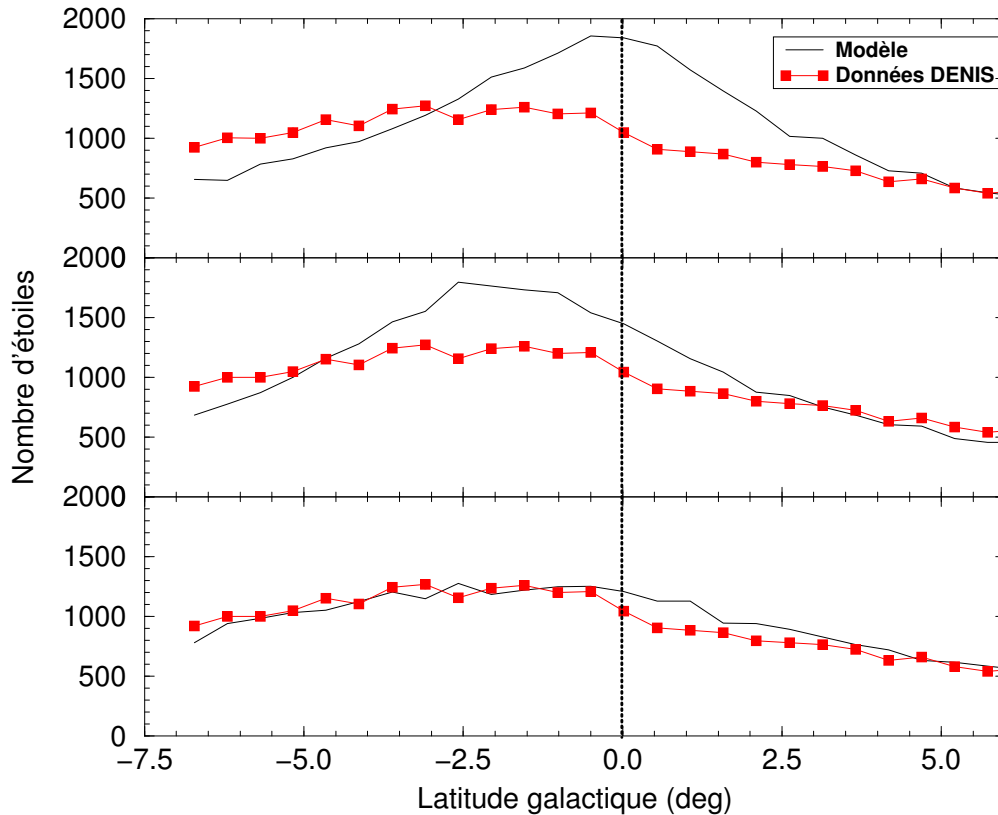


Fig. 7.6: Effet du gauchissement et de l'évasement sur un profil en latitude galactique traversant le plan ($b = 0^\circ$) en $l = 266.1^\circ$. Les données sont groupées par pas de 0.8° . On a respectivement, de haut en bas, un disque normal (sans gauchissement ni évasement), un disque gauchi, et enfin un disque gauchi et évasé.

Enfin, dans le dernier cas, on ajoute un évasement du disque, ce qui a pour effet d'étaler la distribution simulée, qui est alors en bien meilleur accord avec les données.

On peut voir sur la figure 7.7 un exemple d'ajustement des valeurs optimales de R_{warp} et R_{flare} , pour deux strips traversant le plan galactique. Pour chaque couple de valeur, on mesure la qualité de l'ajustement du profil en latitude (où seuls les objets détectés au moins en K_s en deçà de la limite de complétude sont retenus dans les deux cas).

Pour les strips situés autour de 270° de longitude, on obtient un ajustement optimal pour $R_{\text{warp}} = 8400 \pm 200$ pc, soit une position très proche de la valeur retenue pour R_{\odot} (8500 pc).

L'évasement dans cette direction commence à $R_{\text{flare}} = 8800$ pc. Cependant, vers $l = 225^\circ$, le meilleur ajustement est obtenu pour $R_{\text{flare}} = 10500$ pc. Le gauchissement est mal contraint dans cette direction, en raison de sa faible amplitude.

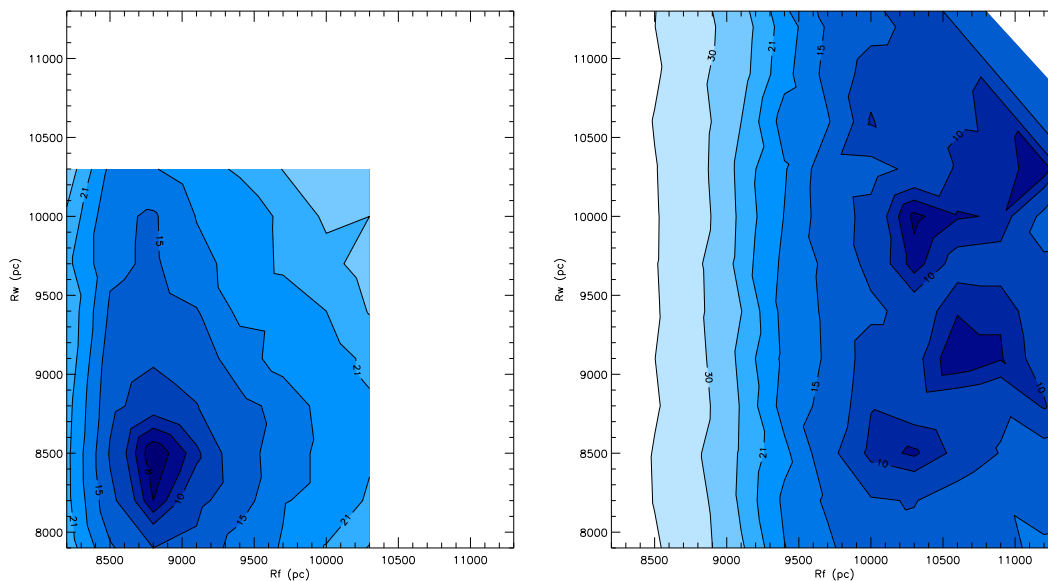


Fig. 7.7: Isocontours (σ_{χ^2}) de l'ajustement du profil en latitude pour les strips 7367 (à gauche, ce strip traverse $b = 0^\circ$ en $l = 266.1^\circ$) et 8787 (à droite, ce strip traverse le plan en $l = 225^\circ$).

La valeur trouvée ici pour R_{warp} est en accord avec de nombreux auteurs qui trouvent également que le Soleil occupe une place privilégiée, sur la ligne de nœuds, et à la frontière du disque gauchi (cf. Tab. 7.2). Elle diffère cependant singulièrement de valeurs récentes qui font commencer le gauchissement bien avant R_\odot ($R_{\text{warp}} = 4.3$ kpc pour Freudreich (1998), ou 6.5 kpc pour Drimmel et al. (2000)). La comparaison n'est toutefois pas évidente, car ces auteurs utilisent, pour décrire l'amplitude radiale du gauchissement, des polynômes plutôt que des droites : l'amplitude z_c reste faible sur de grandes distances avant d'augmenter plus rapidement qu'un modèle linéaire.

Pour l'évasement, les différences observées dans R_{flare} pourraient aussi s'expliquer par une échelle de longueur inhomogène dans le disque, ou des déviations de la loi exponentielle.

Avec les échelles de longueur et les paramètres du gauchissement et de l'évasement trouvés ici, il est exclu qu'une fraction significative de la profondeur optique τ_{LMC} des événements de microlentille gravitationnelle observés vers le LMC soient dus à des objets du disque déplacés sur la ligne de visée du Nuage. Ceci rejoint la conclusion de Gyuk et al. (1999) : il faut, pour augmenter sensiblement τ_{LMC} , à la fois une fraction énorme de disque épais (25%), et des échelles de longueur très élevées (4.5 kpc), ce qui est incompatible avec les comptages.

Conclusions et perspectives

J'AI eu la chance d'effectuer ma thèse dans un contexte unique, avec l'arrivée simultanée à maturité de deux grands projets de cartographie du ciel dans l'infrarouge proche : DENIS et 2MASS.

Les premières données publiques de ces deux projets initiés au début des années 1990 ont été diffusées pendant ma thèse. Ces très grands catalogues contiennent des dizaines de paramètres pour des dizaines de millions de sources.

La première partie du travail que j'ai effectué portait sur la mise au point de systèmes de gestion des énormes volumes de données impliqués. J'ai ainsi été amené à développer des méthodes de codage en binaire dédiées à ces deux projets. L'optimisation du codage permet de gagner en compression un facteur 1.5 à 2 par rapport à des algorithmes de type LZ77 (utilisé par `gzip`), mais surtout de conserver une indexation en position des catalogues, permettant un accès rapide.

L'accès aux catalogues préliminaires DENIS et 2MASS, au fur et à mesure de leurs publications, a ainsi pu être intégré aux services du CDS, comme `VizieR` ou `ALADIN`. Lors d'une requête sur `VizieR`, le temps d'accès moyen par source est inférieur à 20 μ s.

Les algorithmes développés pourront être appliqués pour le codage et l'accès en ligne aux catalogues complets, qui devraient être disponibles au cours de l'année 2002.

Un effort particulier a été accompli dans le cadre du projet DENIS. La validation des données, tant dans le cadre privé du consortium que pour les données publiques (Epchtein et al., 1999), a représenté une étape essentielle. Ainsi, chaque strip est associé à un ensemble de valeurs et de graphiques permettant de contrôler la qualité des données. Le tout étant intégralement consultable en ligne, de même que diverses interfaces de requêtes dédiées.

La comparaison des données réduites par les deux chaînes de traitement de données DENIS (LDAC et PDAC) m'a permis de définir des critères de filtrage des données du PDAC, et d'opter pour l'utilisation exclusive de ces données par la suite.

Le codage des données du PDAC est également prévu prochainement au CDS. Les algorithmes sont d'ores et déjà développés et devraient permettre une diffusion de données au cours de l'année 2001.

Bénéficier en avance de publication des données infrarouges du projet DENIS

était une opportunité que nous avons saisie pour contraindre le modèle de structure galactique de Besançon.

Nous avons introduit quelques nouveautés dans celui-ci, afin de pouvoir rendre compte du gauchissement et de l'évasement du disque galactique.

Des ajustements sur quelques régions sélectionnées nous donnent pour le disque mince une échelle de longueur $h_R = 2.5$ kpc et une troncature à $R_{\text{cut}} = 14$ kpc.

J'ai expérimenté une méthode de reconstruction de l'extinction le long d'une ligne de visée par ajustement itératif d'un catalogue synthétique du modèle de Besançon sur diverses observables. Il serait intéressant d'étendre cette méthode à des zones plus étendues afin de voir si une reconstruction tridimensionnelle de la distribution de matériau absorbant à grande échelle est possible.

Pour toutes les longitudes entre 230 et 330° , on observe à très basse latitude dans les données DENIS, une asymétrie avec une densité de sources systématiquement plus élevée pour $b < 0$ que pour $b > 0$. Nous attribuons cet effet au gauchissement du disque galactique dans les troisième et quatrième quadrants.

En gardant des amplitudes similaires à celles observées dans le gaz pour le gauchissement, on trouve que celui-ci commence très près de la position du Soleil ($R_{\text{warp}} = 8.4$ kpc). Il semble qu'il y ait une dépendance de R_{flare} en longitude.

L'observation d'une contrepartie stellaire du disque gauchi indique que ce phénomène ne résulte pas de l'action d'un champ magnétique extragalactique sur le disque.

Un important travail d'ajustement des paramètres reste à faire, et il sera facilité par l'avancement des réductions. L'utilisation des données 2MASS devrait aussi permettre de contraindre les caractéristiques du gauchissement et de l'évasement, y compris dans l'hémisphère nord (même si la présence du bras spiral Orion-Cygnus risque de rendre cette tâche plus ardue).

Une fois les deux catalogues complets (et accessibles en ligne!), les astronomes disposeront d'un support inestimable pour l'étude de la structure galactique.

Bibliographie

- Alard C., 2000, *Image subtraction using a space-varying kernel*, *A&AS*, **144**, 363
- Alard C. et Lupton R. H., 1998, *A Method for Optimal Image Subtraction*, *ApJ*, **503**, 325
- Arp H. C., 1959, *Stars in the direction of the galactic center*, *AJ*, **64**, 33
- Azcárate I. N., Cersosimo J. C., et Colomb F. R., 1990, *Evidence for the Presence of a Warping of the Ionizing Gas Layer Derived from H166 α Emission Observations in the Outer Galaxy*, *Revista Mexicana de Astronomia y Astrofisica*, **20**, 75
- Baade W., 1944, *The Resolution of Messier 32, NGC 205, and the Central Region of the Andromeda Nebula*, *ApJ*, **100**, 137
- Bahcall J. N., 1986, *Star counts and galactic structure*, *ARA&A*, **24**, 577
- Bahcall J. N. et Soneira R. M., 1980, *The universe at faint magnitudes. I - Models for the galaxy and the predicted star counts*, *ApJS*, **44**, 73
- Bartlett J. et Egret D., 1998, *Cross wavelength comparison of images and catalogs, dans New Horizons from Multi-Wavelength Sky Surveys (Baltimore - Maryland, 26-30 August, 1996)*, Vol. 179 de *IAU Symposium*, p. 437
- Battaner E., Florido E., et Sanchez-Saavedra M. L., 1990, *Intergalactic magnetic field and galactic WARPS*, *A&A*, **236**, 1
- Battaner E. et Jimenez-Vicente J., 1998, *A simple model of magnetically induced WARPS*, *A&A*, **332**, 809
- Bekki K. et Chiba M., 2000, *Formation of the Galactic Stellar Halo : Origin of the Metallicity–Eccentricity Relation*, *ApJ*, **534**, L89
- Bell E. F. et de Jong R. S., 2000, *The stellar populations of spiral galaxies*, *MNRAS*, **312**, 497
- Bergbusch P. A. et Vandenberg D. A., 1992, *Oxygen-enhanced models for globular cluster stars. II - Isochrones and luminosity functions*, *ApJS*, **81**, 163
- Bergeron P., Wesemael F., et Beauchamp A., 1995, *Photometric Calibration of Hydrogen- and Helium-Rich White Dwarf Models*, *PASP*, **107**, 1047
- Bertin E. et Arnouts S., 1996, *SExtractor : Software for source extraction*, *A&AS*, **117**, 393
- Bessell M. S. et Brett J. M., 1988, *JHKLM photometry - Standard systems, passbands, and intrinsic colors*, *PASP*, **100**, 1134

- Bienaymé O., 1999, *Dynamical Models of the Thick Disc* dans *Galaxy Evolution : Connecting the Distant Universe with the Local Fossil Record*, M. Spite (ed.), p. 225
- Bienaymé O., Robin A. C., et Crézé M., 1987, *The mass density in our Galaxy*, *A&A*, **180**, 94
- Binney J., 1992, *WARPS*, *ARA&A*, **30**, 51
- Binney J. et Merrifield M., 1998, *Galactic astronomy*, Vol. QB857 .B522 de *Princeton series in astrophysics*, Princeton University Press, Princeton, NJ
- Bonnarel F., Fernique P., Bienaymé O., Egret D., Genova F., Louys M., Ochsenbein F., Wenger M., et Bartlett J. G., 2000, *The ALADIN interactive sky atlas. A reference tool for identification of astronomical sources*, *A&AS*, **143**, 33
- Borsenberger J., 1997, *DENIS Data Processing at PDAC* dans *Second DENIS Euroconference, The Impact of Large Scale Near-IR Sky Surveys, Tenerife (Spain)*, F. Garzón, N. Epchtein, A. Omont, W. Burton, et P. Persi. (eds.), Vol. 210 de *ASSL*, p. 181, Kluwer
- Bosma A., 1991, *Warped and Flaring HI Disks*, dans *Warped Disks and Inclined Rings around Galaxies*, p. 181
- Briggs F. H., 1990, *Rules of behavior for galactic WARPS*, *ApJ*, **352**, 15
- Bruzual A. G. et Charlot S., 1993, *Spectral evolution of stellar populations using isochrone synthesis*, *ApJ*, **405**, 538
- Burgasser A. J., Kirkpatrick J. D., Brown M. E., Reid I. N., Gizis J. E., Dahn C. C., Monet D. G., Beichman C. A., Liebert J., Cutri R. M., et Skrutskie M. F., 1999, *Discovery of Four Field Methane (T-Type) Dwarfs with the Two Micron All-Sky Survey*, *ApJ*, **522**, L65
- Burke B. F., 1957, *Systematic distortion of the outer regions of the Galaxy.*, *AJ*, **62**, 90
- Burton W. B., 1988, *The structure of our Galaxy derived from observations of neutral hydrogen*, p. 295, *Galactic and Extragalactic Radio Astronomy*, Springer-Verlag
- Burton W. B. et de Lintel Hekkert P., 1986, *The shape of the outer-Galaxy HI layer. I - Atlas of volume densities in cuts through a composite, galactocentric data cube*, *A&AS*, **65**, 427
- Buser R. et Kurucz R. L., 1992, *A library of theoretical stellar flux spectra. I - Synthetic UBVRI photometry and the metallicity scale for F- to K-type stars*, *A&A*, **264**, 557
- Caldwell J. A. R. et Ostriker J. P., 1981, *The mass distribution within our Galaxy - A three component model*, *ApJ*, **251**, 61
- Cambrésy L., Copet E., Epchtein N., de Batz B., Borsenberger J., Fouqué P., Kimeswenger S., et Tiphène D., 1998, *New Young Stellar Object candidates in the Chamaeleon I molecular cloud discovered by DENIS*, *A&A*, **338**, 977

- Cambrésy L., Epchtein N., Copet E., de Batz B., Kimeswenger S., Le Bertre T., Rouan D., et Tiphène D., 1997, *Extinction map of Chamaeleon I molecular cloud with DENIS star counts*, *A&A*, **324**, L5
- Camm G. L., 1950, *Self-gravitating star systems*, *MNRAS*, **110**, 305
- Cannon A. et Pickering E., 1918-1924, *The Henry Draper Catalogue*, *Ann. Astron. Obs. Harvard College*, p. 91
- Cannon A. J., 1925-1936, *The Henry Draper extension*, *Annals of Harvard College Observatory*, **100**, 17
- Carney B. W. et Seitzer P., 1993, *Optical detection of the Galaxy's southern stellar warp and outer disk*, *AJ*, **105**, 2127
- Carpenter E. F., 1931, *The Distribution of Color in Two Extra-Galactic Nebulae*, *PASP*, **43**, 294
- Carraro G., 2001, *The age of the galactic disk* dans *The Chemical Evolution of the Milky Way : Stars vs. Clusters*, F. Matteucci (ed.), Vol. 253 de *ASSL*, Kluwer, in press, astro-ph/9911382
- Chiba M. et Beers T. C., 2000, *Kinematics of Metal-poor Stars in the Galaxy. III. Formation of the Stellar Halo and Thick Disk as Revealed from a Large Sample of Non-kinematically Selected Stars*, *AJ*, **119**, 2843
- Cioni M. R., Habing H. J., et Israel F. P., 2000a, *The morphology of the Magellanic Clouds revealed by stars of different age : results from the DENIS survey*, *A&A*, **358**, L9
- Cioni M. R., Loup C., Habing H. J., Fouqué P., Bertin E., Deul E., Egret D., Alard C., de Batz B., Borsenberger J., Denefeld M., Epchtein N., Forveille T., Garzón F., Hron J., Kimeswenger S., Lacombe F., Le Bertre T., Mamon G. A., Omont A., Paturel G., Persi P., Robin A., Rouan D., Simon G., Tiphène D., Vauglin I., et Wagner S., 2000b, *The DENIS Point Source Catalogue towards the Magellanic Clouds*, *A&AS*, **144**, 235
- Cioni M. R., van der Marel R. P., Loup C., et Habing H. J., 2000c, *The tip of the red giant branch and distance of the Magellanic Clouds : results from the DENIS survey*, *A&A*, **359**, 601
- Cohen M., 1994, *Powerful model for the point source sky : Far-ultraviolet and enhanced midinfrared performance*, *AJ*, **107**, 582
- Cohen M., 1995, *The displacement of the sun from the galactic plane using IRAS and FAUST source counts*, *ApJ*, **444**, 874
- Copet E., 1996, *Etude de la fonction de luminosité des étoiles de faible masse à l'aide du relevé infrarouge DENIS*, *Thèse*, Université Pierre et Marie Curie, Paris
- Copet E., Epchtein N., et Rouan D., 2001, *The DENIS instrument*, *A&AS*, en préparation
- Crézé M., 1979, dans *Comptes Rendus de l'école de Goutelas : Classification stellaire*, Ed. D. Ballereau

- Cr ez e M., Chereul E., Bienaym e O., et Pichon C., 1998, *The distribution of nearby stars in phase space mapped by Hipparcos. I. The potential well and local dynamical mass*, *A&A*, **329**, 920
- Cutri R. M., 1997, *The 2MASS Point Source Processing Pipeline : Design and Prototype Results* dans *The Impact of Large Scale Near-IR Sky Surveys, April 1996, Tenerife (Spain)*, F. Garz on, N. Epchtein, A. Omont, W. Burton, et P. Persi. (eds.), Vol. 210 de *ASSL*, p. 187, Kluwer
- Dahn C. C., Liebert J., et Harrington R. S., 1986, *LHS 292 and the luminosity function of the nearby M dwarfs*, *AJ*, **91**, 621
- de Jager C. et Nieuwenhuijzen H., 1987, *A new determination of the statistical relations between stellar spectral and luminosity classes and stellar effective temperature and luminosity*, *A&A*, **177**, 217
- de Jong R. S., 1996, *Near-infrared and optical broadband surface photometry of 86 face-on disk dominated galaxies. IV. Using color profiles to study stellar and dust content of galaxies*, *A&A*, **313**, 377
- Dekel A. et Shlosman I., 1983, *Galactic WARPS in Tilted Halos*, dans *IAU Symp. 100 : Internal Kinematics and Dynamics of Galaxies*, Vol. 100, p. 187
- Delfosse X., 1997, *Naines brunes et  toiles de tr s faible masse*, Th se, Universit  Joseph Fourier, Grenoble
- Delfosse X., Forveille T., Tinney C. G., et Epchtein N., 1998, *The DENIS Sky Survey : First Results on Very Low Mass Stars and Substellar Objects*, dans *ASP Conf. Ser. 134 : Brown Dwarfs and Extrasolar Planets*, p. 67
- Delfosse X., Tinney C. G., Forveille T., Epchtein N., Bertin E., Borsenberger J., Copet E., de Batz B., Fouque P., Kimeswenger S., Le Bertre T., Lacombe F., Rouan D., et Tiphene. D., 1997, *Field brown dwarfs found by DENIS*, *A&A*, **327**, L25
- Derriere S., Ochsenbein F., et Egret D., 2000, *On-line Access to Very Large Catalogues* dans *Astronomical Data Analysis Software and Systems IX*, N. Manset, C. Veillet, et D. Crabtree (eds.), Vol. 216 de *ASP Conf. Ser.*, p. 235, San Francisco : ASP
- Diplas A. et Savage B. D., 1991, *Neutral hydrogen in the outer Galaxy*, *ApJ*, **377**, 126
- Djorgovski S. et Meylan G., 1994, *The galactic globular cluster system*, *AJ*, **108**, 1292
- Djorgovski S. et Sosin C., 1989, *The warp of the Galactic stellar disk detected in IRAS source counts*, *ApJ*, **341**, L13
- Dorman B., 1992, *Oxygen-enhanced models for globular cluster stars. III - Horizontal-branch sequences*, *ApJS*, **81**, 221
- Dreyer J. L. E., 1888, *A New General Catalogue of Nebulae and Clusters of Stars, being the Catalogue of the late Sir John F.W. Herschel, Bart., revised, corrected, and enlarged.*, *MmRAS*, **49**, 1

- Dreyer J. L. E., 1895, *Index Catalogue of Nebulae found in the years 1888 to 1894, with Notes and Corrections to the New General Catalogue*, *MmRAS*, **51**, 185
- Dreyer J. L. E., 1908, *Index Catalogue of Nebulae and Clusters of Stars, containing objects found in the years 1895 to 1907; with Notes and Corrections to the New General Catalogue and to the Index Catalogue for 1888-94*, *MmRAS*, **59**, 105
- Drimmel R., Smart R. L., et Lattanzi M. G., 2000, *The Galactic warp in OB stars from Hipparcos*, *A&A*, **354**, 67
- Drimmel R. et Spergel D. N., 2001, *Three Dimensional Structure of the Milky Way Disk : the distribution of stars and dust beyond $0.35R_{\odot}$* , *ApJ*, **556**, 181
- Dwek E., Arendt R. G., Hauser M. G., Kelsall T., Lisse C. M., Moseley S. H., Silverberg R. F., Sodroski T. J., et Weiland J. L., 1995, *Morphology, near-infrared luminosity, and mass of the Galactic bulge from COBE DIRBE observations*, *ApJ*, **445**, 716
- Eaton N., Adams D. J., et Giles A. B., 1984, *The 2.2 microns stellar distribution in the galactic plane*, *MNRAS*, **208**, 241
- Eggen O. J., Lynden-Bell D., et Sandage A. R., 1962, *Evidence from the motions of old stars that the Galaxy collapsed*, *ApJ*, **136**, 748
- Egret D., Bartlett J. G., Bonnarel F., Cochery J.-M., Derriere S., Fernique P., Genova F., Louys M., Ochsenbein F., Ortiz P., et Wenger M., 1998a, *CDS Services in Support of Survey Activity : Data Access, Identification and Validation*, dans *Wide Field Surveys in Cosmology, 14th IAP meeting held May 26-30, 1998, Paris.*, p. 394, Editions Frontières
- Egret D., Derriere S., Genova F., Bartlett J. G., Bonnarel F., Cochery J.-M., Fernique P., Louys M., Ochsenbein F., Ortiz P., et Wenger M., 1999, *Mining survey data : CDS services for catalogue access, identification and validation* dans *Astrophysics with Infrared Surveys : A prelude to SIRTf*, M. D. Bica, C. A. Beichman, R. M. Cutri, et B. F. Madore (eds.), Vol. 177 de *ASP Conf. Ser.*, p. 438
- Egret D., Didelon P., McLean B. J., Russell J. L., et Turon C., 1992, *The Tycho Input Catalogue - Cross-matching the Guide Star Catalog with the Hipparcos INCA Data Base*, *A&A*, **258**, 217
- Egret D., Fernique P., et Genova F., 1998b, *Prototype of a Discovery Tool for Querying Heterogeneous Services*, dans *Astronomical Data Analysis Software and Systems VII*, Vol. 145 de *ASP Conf. Ser.*, p. 416
- Egret D., Hanisch R. J., et Murtagh F., 2000, *Search and discovery tools for astronomical on-line resources and services*, *A&AS*, **143**, 137
- Einasto J., 1979, *Galactic mass modeling*, *IAU Symposia*, **84**, 451
- Epchtein N., 1997, *The Deep Near Infrared Southern Sky Survey (denis) : Progress Status and Scientific Achievements* dans *Second DENIS Euroconference, The Impact of Large Scale Near-IR Sky Surveys, April 1996, Tenerife (Spain)*, F. Garzón, N. Epchtein, A. Omont, W. Burton, et P. Persi. (eds.), Vol. 210 de *ASSL*, p. 15, Kluwer

- Epchtein N., Deul E., Derriere S., Borsenberger J., Egret D., Simon G., Alard C., Balázs L. G., de Batz B., Cioni M. ., Copet E., Dennefeld M., Forveille T., Fouqué P., Garzón F., Habing H. J., Holl A., Hron J., Kimeswenger S., Lacombe F., Le Bertre T., Loup C., Mamon G. A., Omont A., Paturel G., Persi P., Robin A. C., Rouan D., Tiphène D., Vauglin I., et Wagner S. J., 1999, *A preliminary database of DENIS point sources*, *A&A*, **349**, 236
- ESA, 1997, *The Hipparcos and Tycho catalogues*, ESA SP-1200
- Evans N. W., Gyuk G., Turner M. S., et Binney J., 1998, *The LMC Microlensing Events : Evidence for a Warped and Flaring Milky Way Disk ?*, *ApJ*, **501**, L45
- Feissel M. et Mignard F., 1998, *The adoption of ICRS on 1 January 1998 : meaning and consequences*, *A&A*, **331**, L33
- Fernique P., Ochsenbein F., et Wenger M., 1998, *CDS GLU, a Tool for Managing Heterogeneous Distributed Web Services*, dans *Astronomical Data Analysis Software and Systems VII*, Vol. 145 de *ASP Conf. Ser.*, p. 466
- Flower P. J., 1975, *Bolometric Corrections for Late-type Giants and Supergiants*, *A&A*, **41**, 391
- Flower P. J., 1977, *Transformations from Theoretical H-R Diagrams to C-M Diagrams*, *A&A*, **54**, 31
- Fouqué P., Chevallier L., Cohen M., Galliano E., Loup C., Alard C., de Batz B., Bertin E., Borsenberger J., Cioni M. R., Copet E., Dennefeld M., Derriere S., Deul E., Duc P., Egret D., Epchtein N., Forveille T., Garzón F., Habing H. J., Hron J., Kimeswenger S., Lacombe F., Le Bertre T., Mamon G. A., Omont A., Paturel G., Pau S., Persi P., Robin A. C., Rouan D., Schultheis M., Simon G., Tiphène D., Vauglin I., et Wagner S. J., 2000, *An absolute calibration of DENIS (deep near infrared southern sky survey)*, *A&AS*, **141**, 313
- Freudenreich H. T., 1998, *A COBE Model of the Galactic Bar and Disk*, *ApJ*, **492**, 495
- Freudenreich H. T., Berriman G. B., Dwek E., Hauser M. G., Kelsall T., Moseley S. H., Silverberg R. F., Sodroski T. J., Toller G. N., et Weiland J. L., 1994, *DIRBE evidence for a warp in the interstellar dust layer and stellar disk of the galaxy*, *ApJ*, **429**, L69
- Fricke W., Schwan H., Lederle T., Bastian U., Bien R., Burkhardt G., Du Mont B., Hering R., Jährling R., Jahreiß H., Röser S., Schwerdtfeger H., et Walter H. G., 1988, *Fifth fundamental catalogue (FK5). Part 1 : The basic fundamental stars*, *Veroeffentlichungen des Astronomischen Rechen-Instituts Heidelberg*, **32**, 1
- Frogel J. A., 1988, *The Galactic nuclear bulge and the stellar content of spheroidal systems*, *ARA&A*, **26**, 51
- Fux R. et Martinet L., 1994, *The disc scale length of the Galaxy*, *A&A*, **287**, L21
- García-Ruiz I., Kuijken K., Dubinski J., et Balcells M., 2000, *The Warp of the Galaxy and the Orientation of the LMC Orbit*, *soumis à MNRAS*, astro-ph-0002057

- Garzón F., Hammersley P. L., Mahoney T., Calbet X., Selby M. J., et Hepburn I. D., 1993, *A Two-Micron Galactic Survey*, *MNRAS*, **264**, 773
- Genova F., Bartlett J. G., Bienaymé O., Bonnarel F., Dubois P., Egret D., Fernique P., Jasniewicz G., Lesteven S., Monier R., Ochsenbein F., et Wenger M., 1996, *CDS as an astronomical information hub*, *Vistas in Astronomy*, **40**, 429
- Genova F., Bartlett J. G., Bonnarel F., Dubois P., Egret D., Fernique P., Jasniewicz G., Lesteven S., Ochsenbein F., et Wenger M., 1998, *The CDS Information Hub*, *ASP Conf. Ser. 145 : Astronomical Data Analysis Software and Systems VII*, **7**, 470
- Genova F., Egret D., Bienaymé O., Bonnarel F., Dubois P., Fernique P., Jasniewicz G., Lesteven S., Monier R., Ochsenbein F., et Wenger M., 2000, *The CDS information hub. On-line services and links at the Centre de Données astronomiques de Strasbourg*, *A&AS*, **143**, 1
- Genzel R., Pichon C., Eckart A., Gerhard O. E., et Ott T., 2000, *Stellar dynamics in the Galactic Centre : proper motions and anisotropy*, *MNRAS*, **317**, 348
- Gilmore G. et Reid N., 1983, *New light on faint stars. III - Galactic structure towards the South Pole and the Galactic thick disc*, *MNRAS*, **202**, 1025
- Gilmore G., Wyse R. F. G., et Kuijken K., 1989, *Kinematics, chemistry, and structure of the Galaxy*, *ARA&A*, **27**, 555
- Gómez A. E., Grenier S., Udry S., Haywood M., Meillon L., Sabas V., Sellier A., et Morin D., 1997, *Kinematics of Disk Stars in the Solar Neighbourhood*, *ESA SP-402 : Hipparcos - Venice '97*, **402**, 621
- Gould A., 1999, *Galactic Structure From Infrared Surveys (invited talk)*, dans *ASP Conf. Ser. 177 : Astrophysics with Infrared Surveys : A Prelude to SIRTF*, p. 234
- Grabelsky D. A., Cohen R. S., Bronfman L., Thaddeus P., et May J., 1987, *Molecular clouds in the Carina arm - Large-scale properties of molecular gas and comparison with H I*, *ApJ*, **315**, 122
- Gredel R. (ed.), 1996, *The galactic center*, Vol. 102 de *ASP Conf. Series*
- Groenewegen M. A. T., 1997, *From I, J, K to m_{Bol} in AGB Stars* dans *Second DENIS Euroconference, The Impact of Large Scale Near-IR Sky Surveys, April 1996, Tenerife (Spain)*, F. Garzón, N. Epchtein, A. Omont, W. Burton, et P. Persi. (eds.), Vol. 210 de *ASSL*, p. 165, Kluwer
- Guillout P., Schmitt J. H. M. M., Egret D., Voges W., Motch C., et Sterzik M. F., 1999, *The stellar content of soft X-ray surveys. II. Cross-correlation of the ROSAT All-Sky Survey with the Tycho and Hipparcos catalogs*, *A&A*, **351**, 1003
- Gyuk G., Flynn C., et Evans N. W., 1999, *Star Counts and the Warped and Flaring Milky Way Disk*, *ApJ*, **521**, 190
- Hammersley P. L., Cohen M., Garzón F., Mahoney T., et López-Corredoira M., 1999, *Structure in the first quadrant of the Galaxy : an analysis of TMGS star counts using the SKY model*, *MNRAS*, **308**, 333

- Hammersley P. L., Garzón F., Mahoney T. J., López-Corredoira M., et Torres M. A. P., 2000, *Detection of the old stellar component of the major Galactic bar*, *MNRAS*, **317**, L45
- Hansen B. M. S. et Phinney E. S., 1998, *Stellar forensics. I - Cooling curves*, *MNRAS*, **294**, 557
- Harris W. E., 1976, *Spatial structure of the globular cluster system and the distance to the galactic center*, *AJ*, **81**, 1095
- Hawkins M. R. S., 1984, *A study of the Galactic halo from a complete sample of RR Lyrae variables to $B = 21$* , *MNRAS*, **206**, 433
- Haywood M., 1994a, *L'évolution du disque de la Voie Lactée*, Thèse, Paris, Université Paris-VI
- Haywood M., 1994b, *The model of stellar evolution of the galactic disc : The IMF at low masses*, *A&A*, **282**, 444
- Haywood M., Robin A. C., et Crézé M., 1997a, *The evolution of the Milky Way disc. I. Vertical structure and local constraints*, *A&A*, **320**, 428
- Haywood M., Robin A. C., et Crézé M., 1997b, *The evolution of the Milky Way disc. II. Constraints from star counts at the galactic poles*, *A&A*, **320**, 440
- Heck A., 1997, *Electronic Yellow-Page Services : The Star*s Family as an Example of Diversified Publishing*, dans *Electronic Publishing for Physics and Astronomy*, Vol. 224 de *ASSL*, p. 211
- Helfer H. L., Wallerstein G., et Greenstein J. L., 1959, *Abundances in Some Population II K Giants.*, *ApJ*, **129**, 700
- Henderson A. P., Jackson P. D., et Kerr F. J., 1982, *The distribution of neutral atomic hydrogen in our galaxy beyond the solar circle*, *ApJ*, **263**, 116
- Høg E., Fabricius C., Makarov V. V., Urban S., Corbin T., Wycoff G., Bastian U., Schwekendiek P., et Wicenec A., 2000, *The Tycho-2 catalogue of the 2.5 million brightest stars*, *A&A*, **355**, L27
- Høg E., Kuzmin A., Bastian U., Fabricius C., Kuimov K., Lindegren L., Makarov V. V., et Roeser S., 1998, *The TYCHO Reference Catalogue*, *A&A*, **335**, L65
- Hubble E. P., 1925, *NGC 6822, a remote stellar system.*, *ApJ*, **62**, 409
- Hubble E. P., 1926, *A spiral nebula as a stellar system : Messier 33.*, *ApJ*, **63**, 236
- Hubble E. P., 1929, *A spiral nebula as a stellar system, Messier 31.*, *ApJ*, **69**, 103
- Hunter C. et Toomre A., 1969, *Dynamics of the Bending of the Galaxy*, *ApJ*, **155**, 747
- Hunter D. A., Light R. M., Holtzman J. A., Lynds R., O'Neil E. J., et Grillmair C. J., 1997, *The Intermediate-Mass Stellar Population of the Large Magellanic Cloud Cluster NGC 1818 and the Universality of the Stellar Initial Mass Function*, *ApJ*, **478**, 124
- Ibata R. A., Gilmore G., et Irwin M. J., 1995, *Sagittarius : the nearest dwarf galaxy*, *MNRAS*, **277**, 781

- Ibata R. A. et Razoumov A. O., 1998, *Archer of the Galactic disk? The effect on the outer HI disk of the Milky Way of collisional encounters with the Sagittarius dwarf galaxy*, *A&A*, **336**, 130
- Ibata R. A., Richer H. B., Gilliland R. L., et Scott D., 1999, *Faint, Moving Objects in the Hubble Deep Field : Components of the Dark Halo ?*, *ApJ*, **524**, L95
- Isern J., Garcia-Berro E., Hernanz M., Mochkovitch R., et Torres S., 1998, *The Halo White Dwarf Population*, *ApJ*, **503**, 239
- Jackson R., Wells D., Adorf H. M., Egret D., Heck A., Koekemoer A., et Murtagh F., 1994, *AstroWeb. A database of links to astronomy resources. (Announcement of a database)*, *A&AS*, **108**, 235
- Jahreiß H. et Wielen R., 1997, *The impact of HIPPARCOS on the Catalogue of Nearby Stars. The stellar luminosity function and local kinematics*, *ESA Symp. 402 : Hipparcos - Venice '97*, **402**, 675
- Jarrett T. H., 2000, *Near-Infrared Galaxy Morphology Atlas*, *PASP*, **112**, 1008
- Jarrett T. H., Chester T., Cutri R., Schneider S., Rosenberg J., Huchra J. P., et Mader J., 2000, *2MASS Extended Sources in the Zone of Avoidance*, *AJ*, **120**, 298
- Johnston K. J. et de Vegt C., 1999, *Reference Frames in Astronomy*, *ARA&A*, **37**, 97
- Kawara K., Kozasa T., Sato S., Kobayashi Y., Okuda H., et Jugaku J., 1982, *Near Infrared Source Counts in the Galactic Plane*, *PASJ*, **34**, 389
- Kent S. M., Dame T. M., et Fazio G., 1991, *Galactic structure from the Spacelab infrared telescope. II - Luminosity models of the Milky Way*, *ApJ*, **378**, 131
- Kerr F. J., 1957, *A Magellanic effect on the galaxy.*, *AJ*, **62**, 93
- Kirkpatrick J. D., Reid I. N., Liebert J., Cutri R. M., Nelson B., Beichman C. A., Dahn C. C., Monet D. G., Gizis J. E., et Skrutskie M. F., 1999, *Dwarfs Cooler than "M" : The Definition of Spectral Type "L" Using Discoveries from the 2 Micron All-Sky Survey (2MASS)*, *ApJ*, **519**, 802
- Kron R. G., 1980, *Photometry of a complete sample of faint galaxies*, *ApJS*, **43**, 305
- Kroupa P. et Tout C. A., 1997, *The theoretical mass-magnitude relation of low mass stars and its metallicity dependence*, *MNRAS*, **287**, 402
- Kuijken K., 1991, *Galactic disk WARPS*, *ApJ*, **376**, 467
- Larson R. B., 1976, *Models for the formation of disc galaxies*, *MNRAS*, **176**, 31
- Lasker B. M., Sturch C. R., McLean B. J., Russell J. L., Jenkner H., et Shara M. M., 1990, *The Guide Star Catalog. I - Astronomical foundations and image processing*, *AJ*, **99**, 2019
- Lejeune T., Cuisinier F., et Buser R., 1997, *Standard stellar library for evolutionary synthesis. I. Calibration of theoretical spectra*, *A&AS*, **125**, 229
- Lejeune T., Cuisinier F., et Buser R., 1998, *A standard stellar library for evolutionary synthesis. II. The M dwarf extension*, *A&AS*, **130**, 65

- Liebert J., Reid I. N., Kirkpatrick J. D., Burrows A., Gizis J. E., Burgasser A., Dahn C. C., Monet D. G., Cutri R., Beichman C., et Skrutskie M., 2000, *2MASS Brown Dwarfs and a First Estimate of the Substellar Mass Function* dans *From Extrasolar Planets to Cosmology : The VLT Opening Symposium*, J. Bergeron et A. Renzini (eds.), ESO Symposium, Antofagasta, Chili, 1-4 Mars 1999, p. 505, Springer-Verlag
- Lindblad B., 1925, *Star-streaming and the structure of the stellar system*, *Arkiv för Matematik, Astronomi och Fysik*, 19A(21)
- Lindblad B., 1926a, *Cosmogonic consequences of a theory of the stellar system*, *Arkiv för Matematik, Astronomi och Fysik*, 19A(35)
- Lindblad B., 1926b, *On the dynamics of the system of globular clusters*, *Arkiv för Matematik, Astronomi och Fysik*, 19A(27)
- Lindblad B., 1926c, *Star-streaming and the structure of the stellar system (Second paper)*, *Arkiv för Matematik, Astronomi och Fysik*, 19B(7)
- Lindblad B., 1927, *On the state of motion in the galactic system*, *MNRAS*, **87**, 553
- Lindblad B., 1930, *The velocity ellipsoid, galactic rotation, and the dimensions of the stellar system*, *MNRAS*, **90**, 503
- Lonsdale C., Conrow T., Evans T., Fullmer L., Moshir M., Chester T., Yentis D., MacGillivray H., Wolstencroft R., et Egret D., 1998, *The OPTID database : deep optical identifications to the IRAS faint source survey*, dans *New Horizons from Multi-Wavelength Sky Surveys, (Baltimore - Maryland, 26-30 August, 1996)*, Vol. 179 de *IAU Symposium*, p. 450
- Lortet M., Borde S., et Ochsenbein F., 1994, *Second Reference Dictionary of the Nomenclature of Celestial Objects*, *A&AS*, **107**, 193
- Ma C., Arias E. F., Eubanks T. M., Fey A. L., Gontier A. ., Jacobs C. S., Sovers O. J., Archinal B. A., et Charlot P., 1998, *The International Celestial Reference Frame as Realized by Very Long Baseline Interferometry*, *AJ*, **116**, 516
- Majewski S. R., 1993, *Galactic structure surveys and the evolution of the Milky Way*, *ARA&A*, **31**, 575
- Majewski S. R., 1999, *The (Sub) Structure of the Halo*, dans *Galaxy Evolution : Connecting the Distant Universe with the Local Fossil Record*, p. 115
- Malhotra S., 1995, *The Vertical Distribution and Kinematics of HI and Mass Models of the Galactic Disk*, *ApJ*, **448**, 138
- Mamon G. A., Tricottet M., Bonin W., et Banchet V., 1997, *Galaxies and Cosmology with DENIS* dans *Extragalactic Astronomy in the Infrared*, G. Mamon, T. Thuan, et J. Trân Than Vân (eds.), p. 369, Frontières
- May J., Alvarez H., et Bronfman L., 1997, *Physical properties of molecular clouds in the southern outer Galaxy*, *A&A*, **327**, 325
- Mayor M., 1974, *Kinematics and Age of Stars*, *A&A*, **32**, 321

- Merrifield M. R., 1992, *The rotation curve of the Milky Way to $2.5 R_{\odot}$ from the thickness of the H I layer*, *AJ*, **103**, 1552
- Misiriotis A., Kylafis N. D., Papamastorakis J., et Xilouris E. M., 2000, *Is the exponential distribution a good approximation of dusty galactic disks?*, *A&A*, **353**, 117
- Miyamoto M., Yoshizawa M., et Suzuki S., 1988, *An optical warp of the Galaxy*, *A&A*, **194**, 107
- Monet D., Bird A., Canzian B., Dahn C., Guetter H., Harris H., Henden A., Levine S., Luginbuhl C., Monet A., Rhodes A., Riepe B., Sell S., Stone R., Vrba F., et Walker R., 1998, *USNO-A V2.0, A Catalog of Astrometric Standards*, Technical report, U.S. Naval Observatory, Washington DC
- Monet D., Bird A., Canzian B., Harris H., Reid N., Rhodes A., Sell S., Ables H., Dahn C., Guetter H., Henden A., Leggett S., Levison H., Luginbuhl C., Martini J., Monet A., Pier J., Riepe B., Stone R., Vrba F., et Walker R., 1997, *USNO-A V1.0, A Catalog of Astrometric Standards*, Technical report, U.S. Naval Observatory, Washington DC
- Nelson B., Cutri R., Kirkpatrick J. D., Beichman C., et Lonsdale C., 1999, *New Populations of IR-Selected AGN : Initial Results From 2MASS*, dans *Astrophysics with Infrared Surveys : A Prelude to SIRTf*, Vol. 177 de *ASP conference*, p. 150
- Neugebauer G. et Leighton R. B., 1969, *Two-micron sky survey. A preliminary catalogue*, NASA SP, Washington : NASA, 1969
- Ochsenbein F., 1998, *The VizieR System for Accessing Astronomical Data* dans *Astronomical Data Analysis Software and Systems VII*, R. Albrecht, R. N. Hook, et H. A. Bushouse (eds.), Vol. 145 de *ASP Conf. Ser.*, p. 387, San Francisco : ASP
- Ochsenbein F. et Lequeux J., 1995, *The A&A Tables and Abstracts : an Example of Collaboration Between Data Centres and Editors*, *Vistas in Astronomy*, **39**, 227
- O'Connell D. J. K. (ed.), 1958, *Ricerche Astronomiche, Specola Vaticana, Proceedings of a Conference at Vatican Observatory, Castel Gandolfo, May 20-28, 1957*, Amsterdam : North-Holland, and New York : Interscience
- Oort J. H., 1927a, *Investigations concerning the rotational motion of the galactic system together with new determinations of secular parallaxes, precession and motion of the equinox (Errata : 4 94)*, *Bull. Astron. Inst. Netherlands*, **4**, 79
- Oort J. H., 1927b, *Observational evidence confirming Lindblad's hypothesis of a rotation of the galactic system*, *Bull. Astron. Inst. Netherlands*, **3**, 275
- Oort J. H., 1928, *Dynamics of the galactic system in the vicinity of the Sun*, *Bull. Astron. Inst. Netherlands*, **4**, 269
- Oort J. H., Kerr F. J., et Westerhout G., 1958, *The galactic system as a spiral nebula (Council Note)*, *MNRAS*, **118**, 379
- Pelat D., 1996, *Bruits et signaux*, Cours

- Perryman M. A. C., 1979, *Optical identifications of radio sources in the 5C 6 survey*, *MNRAS*, **187**, 223
- Pfenniger D., Combes F., et Martinet L., 1994, *Is dark matter in spiral galaxies cold gas? I. Observational constraints and dynamical clues about galaxy evolution*, *A&A*, **285**, 79
- Porcel C., Battaner E., et Jimenez-Vicente J., 1997, *Geometric differences between the gaseous and stellar WARPS in the Milky Way*, *A&A*, **322**, 103
- Porcel C., Garzón F., Jimenez-Vicente J., et Battaner E., 1998, *The radial scale length of the Milky Way*, *A&A*, **330**, 136
- Portinari L. et Chiosi C., 2000, *On radial gas flows, the Galactic Bar and chemical evolution in the Galactic Disc*, *A&A*, **355**, 929
- Price S. D., 1988, *The infrared sky - A survey of surveys*, *PASP*, **100**, 171
- Price S. D. et Walker R. G., 1976, *The AFGL four color infrared sky survey : Catalog of observations at 4.2, 11.0, 19.8, and 27.4 micrometer*, Environmental Research Papers, Hanscom AFB, Mass. : Air Force Geophysics Laboratory, Optical Physics Division, 1976
- Reed B. C., 1996, *Evidence for a Warped Distribution of OB Stars in the Southern Milky Way*, *AJ*, **111**, 804
- Reylé C. et Robin A. C., 2001, *Early galaxy evolution from deep wide field star counts. II. First estimate of the thick disc mass function*, *A&A*, **373**, 886
- Rich R. M., 1988, *Spectroscopy and abundances of 88 K giants in Baade's Window*, *AJ*, **95**, 828
- Robin A. et Crézé M., 1986a, *Stellar populations in the Milky Way - A synthetic model*, *A&A*, **157**, 71
- Robin A. et Crézé M., 1986b, *Stellar populations in the Milky Way - Comparisons of a synthetic model with star counts in nine fields*, *A&AS*, **64**, 53
- Robin A. C., Crézé M., et Mohan V., 1992, *The edge of the Galactic disk*, *ApJ*, **400**, L25
- Robin A. C., Haywood M., Crézé M., Ojha D. K., et Bienaymé O., 1996, *The thick disc of the Galaxy : sequel of a merging event*, *A&A*, **305**, 125
- Robin A. C., Reylé C., et Crézé M., 2000, *Early galaxy evolution from deep wide field star counts. I. The spheroid density law and mass function*, *A&A*, **359**, 103
- Rocca-Volmerange B., Lequeux J., et Maucherat-Joubert M., 1981, *Studies of the Magellanic Clouds. III - Colours, gas and past star formation rate*, *A&A*, **104**, 177
- Rolleston W. R. J., Smartt S. J., Dufton P. L., et Ryans R. S. I., 2000, *The Galactic metallicity gradient*, *A&A*, **363**, 537
- Romano D., Matteucci F., Salucci P., et Chiappini C., 2000, *The Mass Surface Density in the Local Disk and the Chemical Evolution of the Galaxy*, *ApJ*, **539**, 235

- Rood R. T., Carretta E., Paltrinieri B., Ferraro F. R., Fusi Pecci F., Dorman B., Chieffi A., Straniero O., et Buonanno R., 1999, *The Luminosity Function of M3*, *ApJ*, **523**, 752
- Ruphy S., Epchtein N., Cohen M., Copet E., de Batz B., Borsenberger J., Fouqué P., Kimeswenger S., Lacombe F., Le Bertre T., Rouan D., et Tiphène D., 1997, *Stellar populations and inhomogeneities of the galactic plane from DENIS star counts*, *A&A*, **326**, 597
- Ruphy S., Robin A. C., Epchtein N., Copet E., Bertin E., Fouqué P., et Guglielmo F., 1996, *New determination of the disc scale length and the radial cutoff in the anticenter with DENIS data*, *A&A*, **313**, L21
- Russell J. L., Lasker B. M., McLean B. J., Sturch C. R., et Jenkner H., 1990, *The Guide Star Catalog. II - Photometric and astrometric models and solutions*, *AJ*, **99**, 2059
- Salpeter E. E., 1955, *The Luminosity Function and Stellar Evolution*, *ApJ*, **121**, 161
- Sanchez-Saavedra M. L., Battaner E., et Florido E., 1990, *Frequency of Warped Spiral Galaxies at Visible Wavelengths*, *MNRAS*, **246**, 458
- Sanchez-Salcedo F. J., 1999, *Massive dark clusters and the heating problem of galactic discs*, *MNRAS*, **303**, 755
- Sandage A., 1986, *The population concept, globular clusters, subdwarfs, ages, and the collapse of the Galaxy*, *ARA&A*, **24**, 421
- Sandage A. et Fouts G., 1987, *New subdwarfs. VI - Kinematics of 1125 high-proper-motion stars and the collapse of the Galaxy*, *AJ*, **93**, 74
- Schaller G., Schaerer D., Meynet G., et Maeder A., 1992, *New grids of stellar models from 0.8 to 120 solar masses at $Z = 0.020$ and $Z = 0.001$* , *A&AS*, **96**, 269
- Schmidt M., 1956, *A model of the distribution of mass in the Galactic System*, *Bull. Astron. Inst. Netherlands*, **13**, 15
- Schmidt M., 1959, *The Rate of Star Formation*, *ApJ*, **129**, 243
- Schmidt-Kaler T., 1975, *The spiral structure of our Galaxy - A review of current studies*, *Vistas in Astronomy*, **19**, 69
- Schmitz M., Helou G., Dubois P., Lague C., Madore B., Corwin H. G., et Lesteven S., 1995, *A Uniform Bibliographic Code*, *Vistas in Astronomy*, **39**, 272
- Schultheis M., Ganesh S., Glass I. S., Omont A., Ortiz R., Simon G., van Loon J. T., Alard C., Blommaert J. A. D. L., Borsenberger J., Fouqué P., et Habing H. J., 2000, *DENIS and ISOGAL properties of variable star candidates in the Galactic Bulge*, *A&A*, **362**, 215
- Schultheis M., Ganesh S., Simon G., Omont A., Alard C., Borsenberger J., Copet E., Epchtein N., Fouqué P., et Habing H., 1999, *Interstellar extinction towards the inner Galactic Bulge*, *A&A*, **349**, L69

- Schwarzkopf U. et Dettmar R., 2000, *The influence of interactions and minor mergers on the structure of galactic disks. II. Results and interpretations*, *A&A*, **361**, 451
- Seares F. H., 1916, *Color Photographs of Nebulae*, *PASP*, **28**, 123
- Searle L. et Zinn R., 1978, *Compositions of halo clusters and the formation of the galactic halo*, *ApJ*, **225**, 357
- Sion E. M. et Liebert J., 1977, *The space motions and luminosity function of white dwarfs*, *ApJ*, **213**, 468
- Skrutskie M. F., Schneider S. E., Stiening R., Strom S. E., Weinberg M. D., Beichman C., Chester T., Cutri R., Lonsdale C., Elias J., Elston R., Capps R., Carpenter J., Huchra J., Liebert J., Monet D., Price S., et Seitzer P., 1997, *The Two Micron All Sky Survey (2MASS) : Overview and Status* dans *The Impact of Large Scale Near-IR Sky Surveys, April 1996, Tenerife (Spain)*, F. Garzón, N. Epchtein, A. Omont, W. Burton, et P. Persi. (eds.), Vol. 210 de *ASSL*, p. 25, Kluwer
- Sparke L. S. et Casertano S., 1988, *A model for persistent galactic WARPS*, *MNRAS*, **234**, 873
- Spitzer L. J. et Schwarzschild M., 1953, *The Possible Influence of Interstellar Clouds on Stellar Velocities. II.*, *ApJ*, **118**, 106
- Toomre A., 1983, *Theories of WARPS*, dans *IAU Symp. 100 : Internal Kinematics and Dynamics of Galaxies*, Vol. 100, p. 177
- Trumpler R. J., 1930, *Absorption of Light in the Galactic System*, *PASP*, **42**, 214
- Twarog B. A., 1980, *The chemical evolution of the solar neighborhood. II - The age-metallicity relation and the history of star formation in the galactic disk*, *ApJ*, **242**, 242
- Urban S. E., Corbin T. E., et Wycoff G. L., 1998, *The ACT Reference Catalog*, *AJ*, **115**, 2161
- Vallee J. P., 1995, *The Milky Way's Spiral Arms Traced by Magnetic Fields, Dust, Gas, and Stars*, *ApJ*, **454**, 119
- Vauglin I., Paturel G., Borsenberger J., Fouqué P., Epchtein N., Kimeswenger S., Tiphène D., Lanoix P., et Courtois H., 1999, *First DENIS I-band extragalactic catalog*, *A&AS*, **135**, 133
- Velazquez H. et White S. D. M., 1999, *Sinking satellites and the heating of galaxy discs*, *MNRAS*, **304**, 254
- Wainscoat R. J., Cohen M., Volk K., Walker H. J., et Schwartz D. E., 1992, *A model of the 8-25 micron point source infrared sky*, *ApJS*, **83**, 111
- Wenger M., Ochsenbein F., Egret D., Dubois P., Bonnarel F., Borde S., Genova F., Jasniewicz G., Laloë S., Lesteven S., et Monier R., 2000, *The SIMBAD astronomical database. The CDS reference database for astronomical objects*, *A&AS*, **143**, 9

- Wetterer C. J. et McGraw J. T., 1996, *RR Lyrae Variable Star Distribution in the Galactic Halo*, *AJ*, **112**, 1046
- Wielen R., Jahreiß H., et Krüger R., 1983, *The Determination of the Luminosity Function of Nearby Stars*, dans *IAU Colloq. 76 : Nearby Stars and the Stellar Luminosity Function*, p. 163
- Wouterloot J. G. A., Brand J., Burton W. B., et Kwee K. K., 1990, *IRAS sources beyond the solar circle. II - Distribution in the Galactic warp*, *A&A*, **230**, 21
- Wyse R. F. G., Gilmore G., et Franx M., 1997, *Galactic Bulges*, *ARA&A*, **35**, 637

Annexe A

Notations

A_V	Absorption dans la bande V
b	Latitude galactique
$F(M_V, \log T_{\text{eff}}, \text{Age})$	Fonction de luminosité à 2 dimensions
g	Gravité superficielle (on utilise $\log g$)
K_R, K_z	Dérivées du potentiel galactique
l	Longitude galactique
M_V	Magnitude absolue dans la bande V
m_V	Magnitude apparente dans la bande V
ω	Angle solide
$\phi(r, z)$	Potentiel galactique
$\varphi(m)$	Fonction de masse initiale (IMF)
$\Phi(M_V)$	Fonction de luminosité
$\Psi(t)$	Taux de formation d'étoiles (SFR)
$\psi(R, Z, \text{Age})$	Répartition en métallicité : $\sum_i \psi(R, Z_i, \text{Age}) = 1$
R	Distance galactocentrique
r	Distance héliocentrique
$\rho(R, \Theta, z)$	Fonction de densité spatiale
$\sigma_R, \sigma_\Theta, \sigma_z$	Dispersion des vitesses
T_{eff}	Température effective d'une étoile
Θ	Angle en coordonnées cylindriques galactocentriques
V	Vitesse du centroïde
Z	Métallicité
z	Distance au dessus du plan $b = 0$
z_\odot	Hauteur du soleil au dessus du plan $b = 0$

Tab. A.1: Description des symboles utilisés pour les fonctions et variables dans l'étude de la structure galactique.

Annexe B

Identification croisée

CETTE annexe présente une approche simplifiée du problème de l'identification croisée des sources dans les catalogues astronomiques. Après quelques rappels statistiques, et la présentation du problème, on traite le cas de la recherche du plus proche voisin, avec une application sur des simulations numériques et sur de vraies données.

B.1 Rappels de Statistique ¹

B.1.1 Flux de Poisson 2D

Un flux à deux dimensions (2D) est formé d'événements répartis au hasard sur le plan \mathbb{R}^2 .

Un flux de Poisson 2D est un flux ordinaire, homogène, uniforme et sans action à distance. Si on nomme λ la densité moyenne d'événements par unité de surface, la probabilité d'avoir n objets dans une surface ΔS est donnée par :

$$p_n(\Delta S) = \frac{(\lambda \Delta S)^n}{n!} \exp(-\lambda \Delta S) \quad (\text{B.1})$$

La fonction de répartition de la distance au plus proche voisin (ppv) s'écrit, avec $\Delta S = \pi d^2$:

$$F_1(d) = 1 - \exp(-\lambda \pi d^2) \quad (\text{B.2})$$

On en déduit la densité de probabilité de la distance au ppv (voir figure B.1) :

$$\Phi(d, \lambda) = 2\pi \lambda d \exp(-\pi \lambda d^2) \quad (\text{B.3})$$

¹Voir le cours de D. Pelat (1996) pour les démonstrations. pelat@daec.obspm.fr

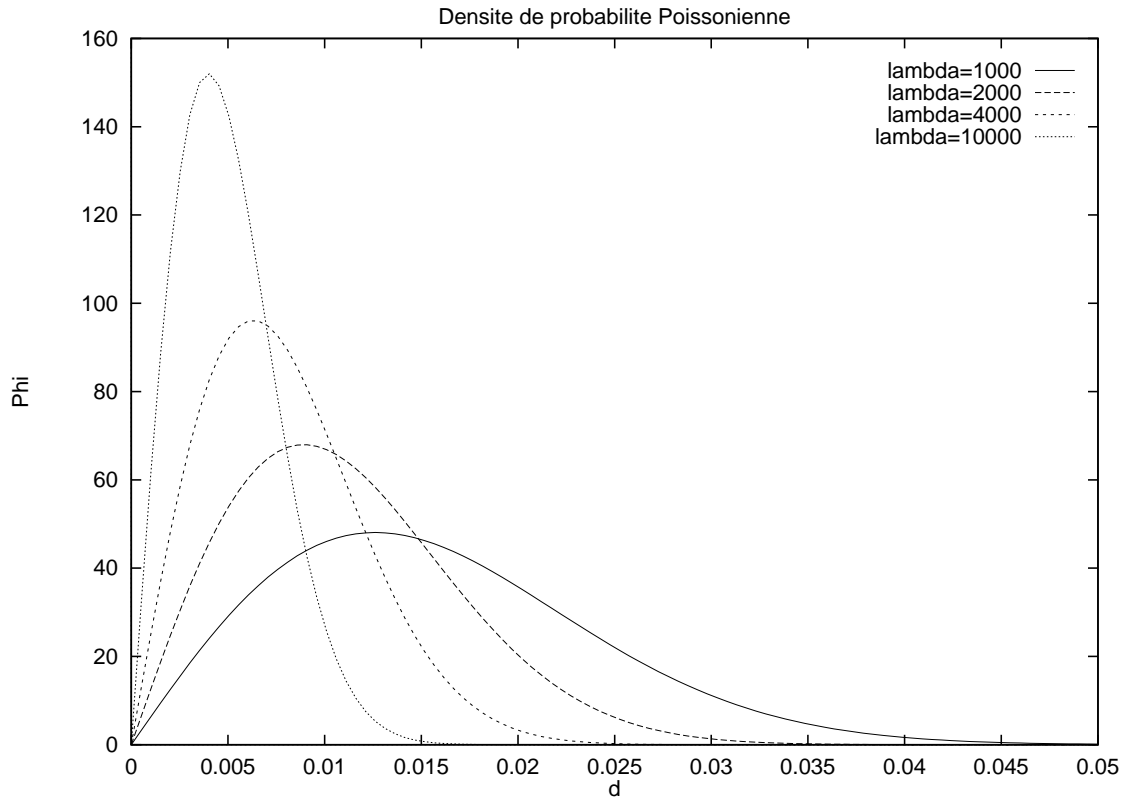


Fig. B.1: Densité de probabilité de la distance au ppv, $\Phi(d, \lambda)$ pour différentes densités λ , dans le cas d'un flux de Poisson 2D.

B.1.2 Flux Normal 2D

La loi normale 2D s'écrit dans le cas général :

$$N_{2D}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2(1-\rho^2)^{\frac{1}{2}}} \exp \left\{ -\frac{1}{2(1-\rho^2)} \right. \quad (\text{B.4})$$

$$\left. \left(\frac{(x_1 - \mu_1)^2}{\sigma_1^2} - \frac{2\rho(x_1 - \mu_1)(x_2 - \mu_2)}{\sigma_1\sigma_2} + \frac{(x_2 - \mu_2)^2}{\sigma_2^2} \right) \right\} \quad (\text{B.5})$$

Dans le cas où $\rho = 0$, $\mu_1 = \mu_2 = 0$ et $\sigma_1 = \sigma_2 = \sigma$, elle se réduit à :

$$N_{2D}(d) = \frac{1}{2\pi\sigma^2} \exp \left(-\frac{d^2}{2\sigma^2} \right) \quad (\text{B.6})$$

avec $d^2 = x_1^2 + x_2^2$. La fonction de densité associée est :

$$f_G = 2\pi d \cdot N_{2D}(d) = \frac{d}{\sigma^2} \exp \left(-\frac{d^2}{2\sigma^2} \right) \quad (\text{B.7})$$

Et la fonction de répartition :

$$F_G = 1 - \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (\text{B.8})$$

On voit qu'on retrouve la même forme de loi que pour le flux de Poisson 2D avec $\frac{1}{\sigma^2} = 2\pi\lambda$.

B.2 Principe de l'identification croisée

B.2.1 Introduction au problème

Soient deux catalogues 1 et 2, sur une même région du ciel de surface S : on a N_1 objets dans le premier catalogue et N_2 objets dans le deuxième.

L'identification croisée a pour but d'associer les objets communs aux deux catalogues.

Au mieux, l'identification croisée permet de récupérer N objets, N étant le nombre d'objets communs aux deux catalogues. En effet, certains objets d'un catalogue peuvent ne pas avoir de contrepartie dans l'autre (à cause, par exemple, de l'utilisation de détecteurs de sensibilités différentes, de longueurs d'onde différentes...).

Restent donc dans les catalogues respectivement $\bar{N}_1 = N_1 - N$ et $\bar{N}_2 = N_2 - N$ objets qui ne doivent pas être associés, mais qui peuvent perturber l'identification des autres.

B.2.2 Difficultés rencontrées

Un critère essentiel pour effectuer l'identification croisée de deux objets est leur **proximité spatiale**. L'approche la plus simplifiée consiste à rechercher le ppv d'un catalogue à l'autre.

En pratique, l'identification croisée de deux catalogues n'est pas aussi simple. La présence d'objets n'ayant pas de contrepartie dans l'autre catalogue est une première difficulté, mais il y en a d'autres :

- **Erreurs sur les positions** : les positions des sources dans les deux catalogues ne sont pas connues avec une précision infinie. Chaque objet se voit associé, en général, une ellipse dont la taille des axes est proportionnelle à l'erreur sur la position. Quand on recherche la contrepartie d'un objet dans un autre catalogue, la recherche doit être effectuée dans une zone dont l'étendue dépend des erreurs sur les positions dans les deux catalogues ;
- **Mouvements propres** : les étoiles de la Galaxie ont un certain mouvement propre. Si les deux catalogues correspondent à des époques d'observation différentes, un même objet ne sera pas situé au même endroit dans les deux, ce qui va compliquer l'identification (Egret et al., 1992) ;

- **Vraisemblance** : dans le cas où les deux catalogues ont des densités d'objets très différentes (par exemple, $N_1 \ll N_2$), on peut trouver, pour un objet du catalogue le moins peuplé, plusieurs partenaires dans l'autre catalogue. Il en est de même si les erreurs sur les positions dans un catalogue sont très supérieures aux erreurs dans l'autre. Il peut être judicieux, dans ces cas, de proposer pour chaque objet du premier catalogue plusieurs candidats dans le deuxième, avec une estimation de la probabilité que chaque candidat soit le bon : la **vraisemblance de l'identification croisée** (Lonsdale et al., 1998) ;
- **Distance généralisée** : si l'on dispose de paramètres autres que la position des objets (magnitude, couleur, taille, ...), on peut envisager d'utiliser une distance généralisée, tenant compte par exemple de la différence de magnitude entre deux objets. Ce genre de critère doit cependant être employé avec précautions, car il peut conduire à rejeter des associations d'objets aux caractéristiques singulières (couleurs extrêmes, très grands mouvements propres), qui sont souvent les plus intéressants .

Les cas les plus délicats correspondent à des identifications croisées entre des domaines de longueur d'onde très différents (rayons X et optique (Guillout et al., 1999), radio et optique (Perryman, 1979)). C'est en effet surtout dans ces cas que l'on rencontre des erreurs sur les positions et des densités d'objets très différentes. La proportion d'objets détectés à la fois dans les deux catalogues est aussi plus faible.

Un cas limite d'identification croisée consiste à trouver, dans un catalogue donné, la contrepartie d'un seul objet.

Production d'un catalogue d'objets associés

La production d'un catalogue d'objets communs lors d'un processus d'identification croisée impose de prendre des décisions sur les seuils d'acceptation (rayon de recherche du ppv par exemple). Il est intéressant, dans ce cas, d'estimer la qualité du catalogue produit (pertinence des identifications, taux de contamination par de fausses associations).

B.3 Le cas du plus proche voisin

Je parlerai essentiellement, dans la suite de cette annexe, de l'identification croisée d'un point de vue statistique, pour de grands ensembles de sources, comme c'est le cas avec les catalogues optiques ou infrarouges proches (Bartlett & Egret, 1998).

Dans cette section, j'étudie le cas d'une identification croisée où le seul critère retenu est la distance entre les objets.

B.3.1 Positions exactes

Partons du cas idéal où les erreurs sur les positions sont nulles : on a N_1 objets dans le premier catalogue, N_2 dans le deuxième, dont N objets communs aux deux.

Soit d la distance entre un objet du premier catalogue et son ppv dans le deuxième. Si on étudie la répartition du nombre d'objets ayant leur ppv à la distance d , on trouvera :

- Les N objets identiques en $d = 0$;
- Les $(N_1 - N)$ objets restants associés aléatoirement (car ils n'ont pas de partenaire réel!) avec un autre objet du deuxième catalogue. Supposons que la distribution des objets du deuxième catalogue suive une loi de Poisson $2D^2$ avec une densité $\lambda = N_2/S$, S étant la surface de la zone de comparaison. La densité de probabilité associée se réduit donc à la loi $\Phi(d, \lambda)$ définie à l'équation B.3, et la distribution pour les $(N_1 - N)$ objets sans partenaire réel est donc donnée par

$$\varphi(N_1, N_2, N, d) = (N_1 - N) \cdot \Phi(d, \lambda) \quad (\text{B.9})$$

(la dépendance en N_2 de φ vient de λ).

Un point important est que les $N_1 - N$ objets qui n'ont pas de partenaire dans le deuxième catalogue donnent **toujours** une contribution de cette forme, dépendant principalement de N_2/S , et ce tant que l'hypothèse de la distribution Poissonienne est valide.

La distribution totale du nombre d'objets en fonction de la distance au ppv peut s'écrire

$$D(d, N_1, N_2, N) = N \cdot \delta(d - 0) + \varphi(N_1, N_2, N, d) \quad (\text{B.10})$$

où $\delta(d - 0)$ est une distribution de Dirac centrée sur 0.

B.3.2 Incertitude sur les positions

Introduisons maintenant une incertitude sur les positions, dans le premier catalogue³. Par rapport au cas où les positions sont exactes, nous constatons que :

- le Dirac qu'on obtenait en zéro contenant les N objets va s'élargir, à cause de la distance que l'erreur en position va introduire entre la position de chaque objet dans les deux catalogues ;
- on aura toujours une partie des objets du premier catalogue correctement associés à leur vrai partenaire, mais dans certains cas, si l'erreur sur la position est assez grande, *un objet peut perdre son vrai partenaire* et se retrouver associé à un autre.

²Cette hypothèse n'est valable que si la densité est homogène sur la zone de comparaison, et qu'il n'y a aucune corrélation entre les positions des différentes sources. On revient dans la section B.6 sur les limites de cette approximation.

³On verra dans la section B.6 que ce cas se généralise simplement à des erreurs sur les positions dans les deux catalogues.

La distribution totale du nombre d'objets en fonction de la distance au ppv est donc désormais la somme de 3 composantes :

- $\alpha(d)$: Objets correctement associés à leur vrai partenaire ;
- $\beta(d)$: Objets ayant un vrai partenaire, mais mal associés ;
- $\varphi(d)$: Objets sans vrai partenaire, associés avec un partenaire aléatoire.

B.3.3 Formalisme

Supposons que les positions dans le deuxième catalogue soient exactes et que l'erreur sur les positions dans le premier catalogue suive une loi normale à deux dimensions (loi d'erreur gaussienne).

Pour l'un des N objets du premier catalogue ayant un vrai partenaire dans le deuxième, on aura deux cas de figure :

- composante $\alpha(d)$: il est associé à son (vrai) partenaire gaussien (B.11) et donc il n'a pas de (faux) partenaire Poissonien pour une distance inférieure (B.12) ;
ou bien
- composante $\beta(d)$: il est associé à un partenaire Poissonien (il a donc perdu son vrai partenaire, B.13) et donc il n'a pas rencontré son partenaire gaussien pour une distance inférieure (B.14).

Si on exprime les différentes lois (voir section B.1) :

- Fonction de densité normale 2D :

$$f_G = \frac{d}{\sigma^2} \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (\text{B.11})$$

- Probabilité de ne pas trouver un élément Poissonien à une distance inférieure à d :

$$1 - F_1(d) = \exp(-\lambda\pi d^2) \quad (\text{B.12})$$

- Fonction de densité Poissonienne :

$$\Phi(d, \lambda) = 2\pi\lambda d \exp(-\pi\lambda d^2) \quad (\text{B.13})$$

- Probabilité de ne pas trouver un élément gaussien à une distance inférieure à d :

$$1 - F_G = \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (\text{B.14})$$

La composante α s'exprime comme le produit des expressions B.11 et B.12 :

$$\alpha = N \cdot \frac{d}{\sigma^2} \exp\left(-\frac{d^2}{2\sigma^2}\right) \cdot \exp(-\lambda\pi d^2) \quad (\text{B.15})$$

et la composante β comme le produit des expressions B.13 et B.14 :

$$\beta = N \cdot 2\pi\lambda d \exp(-\pi\lambda d^2) \cdot \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (\text{B.16})$$

On vérifie bien la normalisation :

$$\int_{d=0}^{\infty} (\alpha(d) + \beta(d)) dd = N \cdot (1 + 2\pi\lambda\sigma^2) \int_{d=0}^{\infty} \frac{d}{\sigma^2} \exp\left\{-\left(\frac{1}{2\sigma^2} + \pi\lambda\right) d^2\right\} dd = N \quad (\text{B.17})$$

On peut réécrire les composantes α et β en utilisant l'équation B.3 :

$$\alpha = \frac{N}{2\pi\lambda\sigma^2} \Phi(d, \lambda) \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (\text{B.18})$$

$$\beta = N\Phi(d, \lambda) \exp\left(-\frac{d^2}{2\sigma^2}\right) \quad (\text{B.19})$$

Pour les sources sans vrai partenaire, on a toujours la relation B.9 :

$$\varphi = (N_1 - N) \cdot \Phi(d, \lambda)$$

La distribution totale observée de la distance au ppv pour les objets du premier catalogue est la somme de ces 3 fonctions, et son intégrale vaut évidemment N_1 :

$$D(d, N_1, \lambda, \sigma, N) = \alpha + \beta + \varphi$$

Ce qu'on peut écrire :

$$D(d, N_1, \lambda, \sigma, N) = N_1\Phi(d, \lambda) + N\Phi(d, \lambda) \left[\frac{1 + 2\pi\lambda\sigma^2}{2\pi\lambda\sigma^2} \exp\left(-\frac{d^2}{2\sigma^2}\right) - 1 \right] \quad (\text{B.20})$$

On connaît N_1 et λ (via N_2 et S). L'ajustement des deux paramètres σ et N sur la distribution observée permet de trouver le nombre d'objets ayant un vrai partenaire dans le deuxième catalogue : N .

B.3.4 Vraisemblance des identifications

On peut estimer pour chaque valeur de d , une vraisemblance que l'association soit correcte :

$$L(d) = \frac{\alpha}{\alpha + \beta + \varphi}. \quad (\text{B.21})$$

Ce qui s'exprime en simplifiant :

$$L(d) = \frac{1}{1 + 2\pi\lambda\sigma^2 \left[1 + \left(\frac{N_1}{N} - 1\right) \exp\left(+\frac{d^2}{2\sigma^2}\right) \right]}. \quad (\text{B.22})$$

Admettons qu'on décide d'une distance seuil d_s pour la création d'un catalogue de sources identifiées : on accepte une association pour $d < d_s$ et on la rejette sinon.

Le nombre d'objets N_s ainsi sélectionnés n'a aucune raison d'être égal à N . On le trouve en intégrant la formule B.20 :

$$N_s = \int_0^{d_s} D(d)dd. \quad (\text{B.23})$$

Ce nombre N_s contient un ensemble N_e d'erreurs, donné par

$$N_e = \int_0^{d_s} [\beta(d) + \varphi(d)] dd = N_s - \int_0^{d_s} \alpha(d)dd \quad (\text{B.24})$$

et on sait qu'on a manqué N_m objets qui avaient vraiment un partenaire :

$$N_m = \int_{d_s}^{\infty} \alpha(d)dd. \quad (\text{B.25})$$

B.4 Simulation numérique

Pour tester le formalisme, on utilise dans un premier temps des fichiers simulés. On crée deux catalogues identiques (objets uniformément répartis), puis on ajoute dans chaque catalogue un nombre arbitraire d'objets supplémentaires (différents dans les deux). La distribution des objets dans les deux catalogues suit une loi de Poisson 2D.

On ajoute ensuite, dans l'un des deux catalogues, une erreur gaussienne sur les positions.

Dans l'exemple suivant, nous avons créé deux catalogues contenant respectivement $N_1 = 10000$ et $N_2 = 14000$ objets, dont $N = 8000$ objets identiques, distribués uniformément sur une surface $S = 1.693^{\circ 2}$. La densité d'objets dans le deuxième catalogue est donc $\lambda \approx 8270^{\circ 2^{-1}}$.

Le bruit gaussien introduit sur chaque objet du premier catalogue, correspond à $\sigma = 5 \cdot 10^{-4}$.

On recherche ensuite, pour chaque objet du premier catalogue, le ppv dans le deuxième. L'histogramme de la distribution ainsi obtenue est reporté sur la figure B.2 (intervalles de largeur 10^{-4} °).

On a également superposé sur la figure la distribution totale $N(d, N_1, \lambda, \sigma, N)$ théorique et les 3 composantes (α, β, φ) .

On voit que les distributions théoriques suivent très fidèlement la répartition issue des catalogues simulés.

La vraisemblance (formule B.22) correspondant à cette situation simulée est reportée sur la figure B.3.

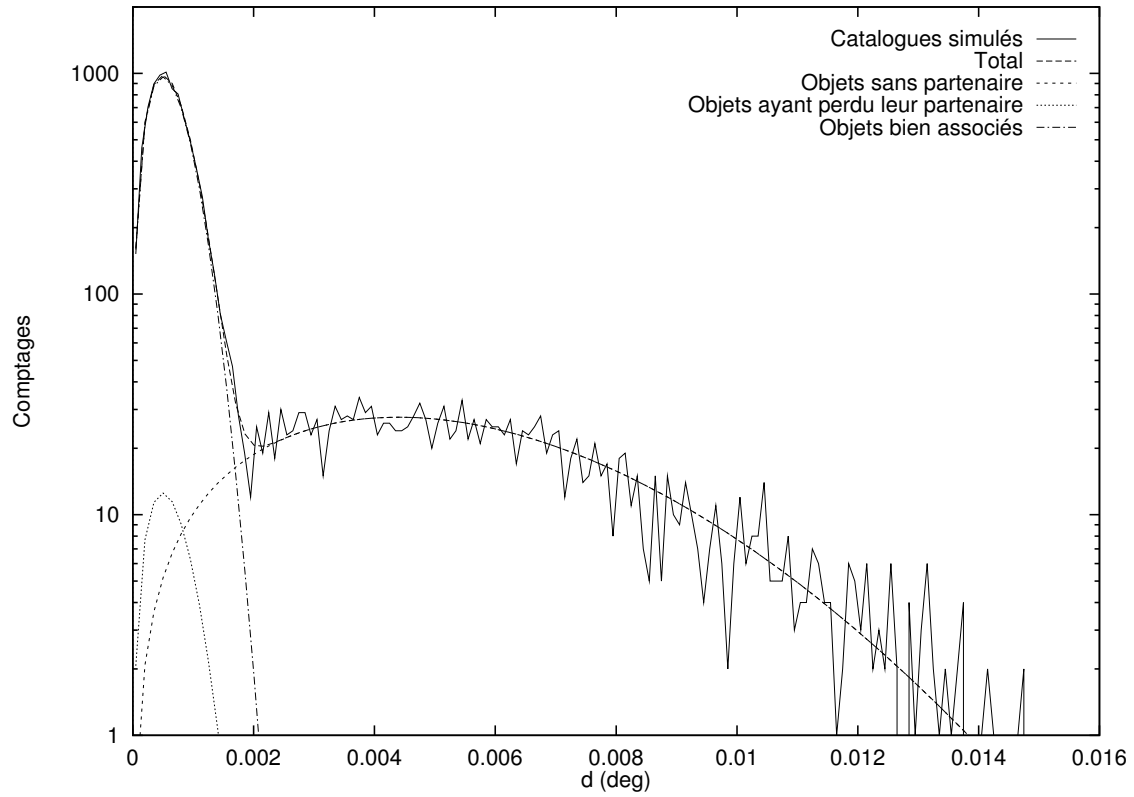


Fig. B.2: Distribution de la distance aux plus proches voisins pour des catalogues simulés. On a superposé les distributions théoriques $D(d)$, $\varphi(d)$, $\beta(d)$, $\alpha(d)$.

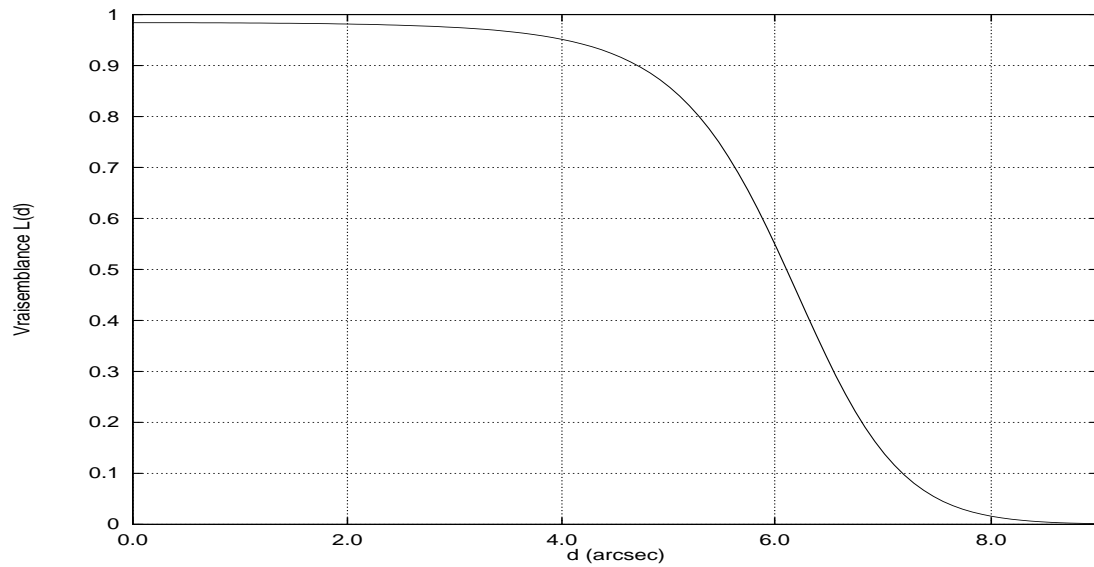


Fig. B.3: Vraisemblance de l'identification croisée en fonction de la distance d au ppv, pour un catalogue simulé.

On voit que dans ce cas, avec une erreur imposée sur les positions de $1.8''$ ($\sigma = 5 \cdot 10^{-4}$), tous les objets pour lesquels on trouve le ppv à moins de $4''$ ont plus de 95% de chances d'être bien associées. Par contre, il y aura moins de 2% de chance qu'une association soit correcte si la distance au ppv est supérieure à $8''$.

B.5 Données réelles

Nous avons testé notre formalisme sur des données réelles, en étudiant l'identification croisée sur une petite portion des catalogues DENIS et USNO A1.0.

Prenons, par exemple, la zone suivante :

$$8.75^\circ < \alpha < 8.95^\circ, \quad -35^\circ < \delta < -32^\circ$$

La surface correspondante fait $S = 0.5^{\circ 2}$

On prend d'un côté les données DENIS du strip 4818 (traité au LDAC) comprises dans cette zone. On obtient une première liste de $N_1 = 4230$ objets.

D'un autre côté, on extrait du catalogue USNO A1.0 toutes les sources dans la même zone : $N_2 = 1809$ objets.

On cherche directement, à partir des positions seules, le ppv de chaque objet DENIS dans la liste d'objets USNO A1.0.

La figure B.4 donne la distribution observée pour la distance aux ppv. On a superposé les courbes théoriques correspondant aux valeurs $N = 1020$ et $\sigma = 1.4 \cdot 10^{-4}$ qui donnent un bon ajustement des données.

On a un excellent accord entre les distributions calculées et observées. Ceci confirme que l'erreur sur les positions entre les sources DENIS et les sources USNO est de l'ordre de $0.5''$ (valeur de σ pour le meilleur ajustement). Mais surtout, cela nous permet d'estimer de façon fiable le nombre d'objets détectés à la fois dans DENIS et dans l'USNO : $N = 1020$. Cette valeur est à comparer avec les nombres d'objets dans les deux listes de départ : 4230 (DENIS), et 1809 (USNO).

La vraisemblance correspondant à cette comparaison, déduite des valeurs estimées pour N et σ , est reportée sur la figure B.5. On voit que pour une séparation de plus de $2''$ des deux sources, la probabilité de faire une bonne identification est quasiment nulle, avec le seul critère de proximité considéré ici.

On a également reporté sur la figure B.5, en fonction du seuil d adopté comme coupure sur la distance au ppv pour la production d'un catalogue de sortie, le nombre total d'objets retenus pour l'identification croisée N_s (équation B.23, ici normalisée à N), et le nombre d'erreurs commises N_e (équation B.24, également normalisée).

B.6 Vers une approche plus générale

- Nous avons fait l'hypothèse que les positions étaient sans erreur dans le premier catalogue, et que seules les positions dans le second souffraient d'une

incertitude gaussienne. La situation se généralise en fait simplement au cas où il y a des erreurs sur les positions dans les deux catalogues : le σ devient simplement la combinaison des incertitudes dans les deux catalogues (situation triviale dans le cas gaussien).

- Nous avons ici limité la recherche au plus proche voisin des sources d'un catalogue dans l'autre. Mais à aucun moment nous n'avons imposé l'unicité de l'appariement. Une même source du deuxième catalogue peut être le plus proche voisin de plusieurs sources du premier catalogue. Si l'on veut que chaque source ne puisse être associée qu'une seule fois, il faut adopter une démarche itérative, en rejetant progressivement les associations si on en trouve de meilleures pour les mêmes objets (Egret et al., 1992).
- De façon implicite, nous avons supposé que la recherche du plus proche voisin se fait dans un cercle. Si on utilise une boîte carrée, par exemple, la forme des lois sera tronquée aux grandes distances (comme on peut le voir sur la Fig. 1 de Guillout et al. (1999)).
- Sur les diagrammes de validation de l'astrométrie dont il est question dans la section 3.3.1, c'est une recherche de tous les voisins dans un rayon donné qui est effectuée. La distribution de la distance au n^e plus proche voisin est donnée, dans le cas d'un flux de Poisson 2D, par la relation :

$$f_{D_n}(d) = 2\pi\lambda d \frac{(\lambda\pi d^2)^{n-1}}{(n-1)!} e^{-\lambda\pi d^2} . \quad (\text{B.26})$$

La somme des distributions des distances à tous les plus proches voisins vaut :

$$\sum_{n=1}^{\infty} f_{D_n}(d) = 2\pi\lambda d \cdot e^{-\lambda\pi d^2} \sum_{n=1}^{\infty} \frac{(\lambda\pi d^2)^{n-1}}{(n-1)!} \quad (\text{B.27})$$

$$= 2\pi\lambda d \quad (\text{B.28})$$

La densité surfacique dans les diagrammes comme ceux de la figure 3.2 est donc constante pour la composante Poissonienne (on divise par $2\pi d$ la distribution linéaire).

Sur un diagramme comme celui de la figure 3.3-a, on voit cependant clairement qu'il y a une zone où la densité du bruit de fond est plus faible. On peut l'expliquer car la distribution des sources dans un plan (α, δ) n'est pas strictement identique à un flux de Poisson 2D : lors du processus d'extraction des sources, il y a un rayon autour de la position de chaque source dans lequel on ne peut pas détecter un autre objet. Ceci explique la sous-densité relative du bruit de fond à proximité immédiate des objets situés au même endroit dans les deux catalogues.

Ceci prouve que l'approximation de la distribution des objets sans partenaire par un flux de Poisson 2D est exagérée dans le formalisme que nous avons développé ici. Toutefois, les résultats restent satisfaisants, comme on l'a vu dans l'application entre DENIS et l'USNO.

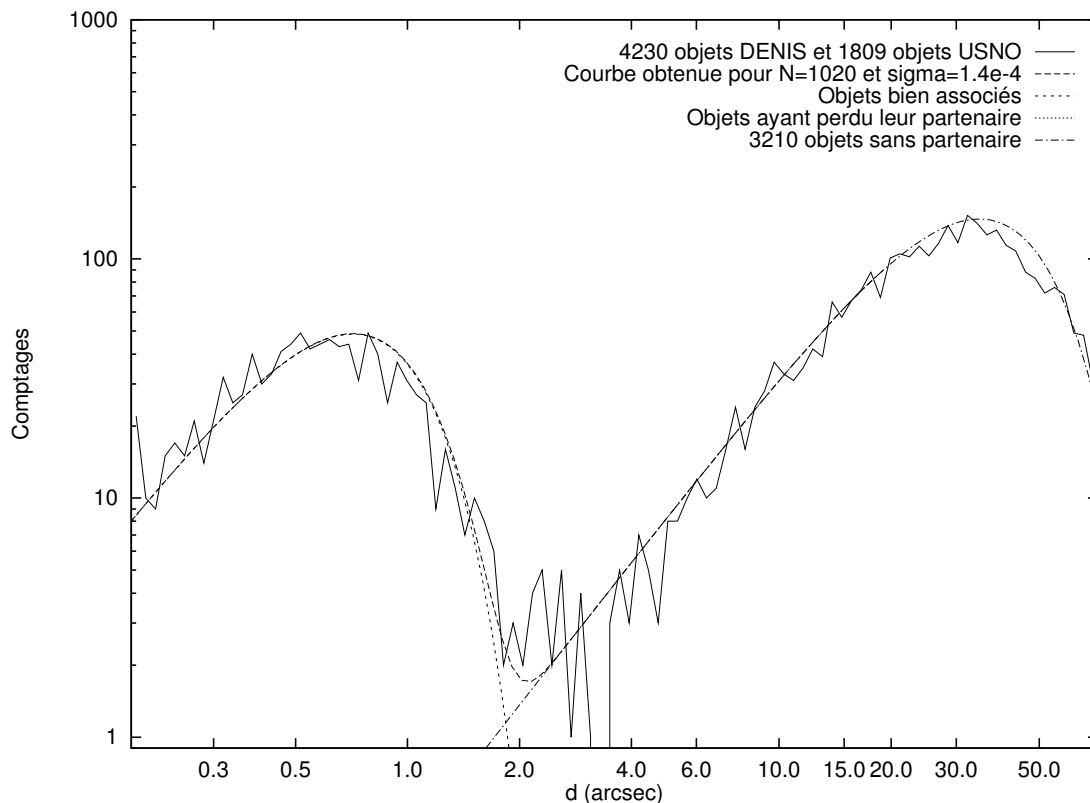


Fig. B.4: Distribution de la distance aux ppv dans le cas de l'identification croisée de sources DENIS et USNO (échelle log-log). La ligne irrégulière en trait plein correspond à l'histogramme observé (avec un pas variable, logarithmique) des valeurs de d . On a superposé les différentes lois présentées au paragraphe B.3.3. Dans la partie droite de la courbe (grandes valeurs de d), c'est la fonction $\varphi(d)$ qui domine, c'est à dire les objets DENIS n'ayant pas de partenaire, et associés aléatoirement avec des objets USNO (loi de type Poisson). A gauche, on a les vraies associations (fonction $\alpha(d)$), avec des ppv à petite distance. La probabilité qu'un objet ait perdu son partenaire est ici trop faible et n'apparaît pas ($\beta(d) \approx 0$). On voit que la distribution globale théorique $D(d)$ (en tirets), pour les paramètres $N = 1020$ et $\sigma = 0.5''$ est en très bon accord avec la distribution observée.

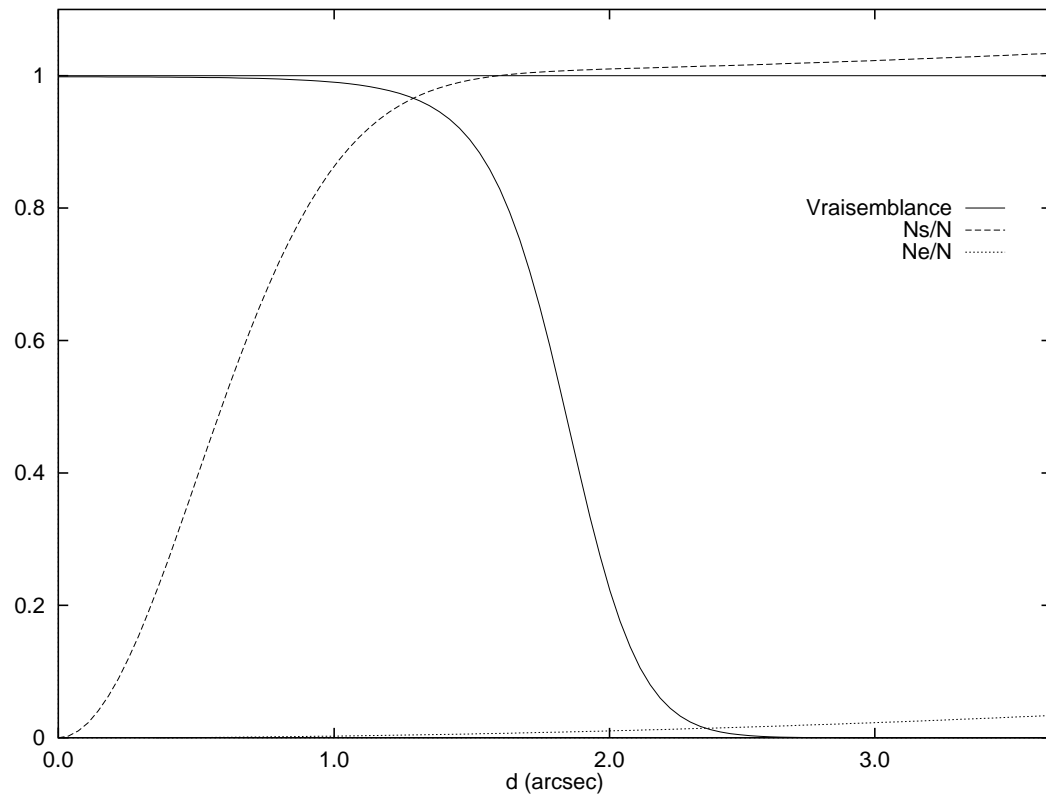


Fig. B.5: Vraisemblance de l'identification croisée des catalogues DENIS et USNO (trait plein). Le nombre total de sources retenues N_s/N et le nombre d'erreurs commises N_e/N sont également représentés en fonction du seuil de sélection d . N_s devient supérieur à N avant que la vraisemblance ne chute sensiblement à cause des fausses associations accumulées.

Annexe C

Publications

C.1 Journaux à comité de lecture

- Page 188 : Epchtein N., Deul E., Derriere S. et al., 1999, **A preliminary database of DENIS point sources**, *A&A*, **349**, 236
- Page 195 : Fouqué P. et al., 2000, **An absolute calibration of DENIS (deep near infrared southern sky survey)**, *A&AS*, **141**, 313

C.2 Autres publications

- Page 200 : Egret D. et al., 1998, **CDS Services in Support of Survey Activity : Data Access, Identification and Validation**, dans *Wide Field Surveys in Cosmology, 14th IAP meeting held May 26-30, 1998, Paris*, Editions Frontières, p. 394 ;
- Page 202 : Egret D., Derriere S. et al., 1999, **Mining survey data : CDS services for catalogue access, identification and validation**, dans *Astrophysics with Infrared Surveys : A prelude to SIRTf*, Vol. **177** de *ASP Conf. Ser.*, 438 ;
- Page 206 : Derriere S., Ochsenbein F., et Egret D., 2000 **On-line Access to Very Large Catalogues** dans *Astronomical Data Analysis Software and Systems IX*, N. Manset, C. Veillet, et D. Crabtree (eds.), Vol. **216** de *ASP Conf. Ser.*, p. 235 ;
- Page 210 : Derriere S. et Robin A. C., 2001, **Near-infrared surveys and the shape of the Galactic disc** dans *The New Era of Wide Field Astronomy*, R. Clowes, A. Adamson, et G. Bromage (eds.), Vol. **232** de *ASP Conf. Ser.*, sous presse.

A preliminary database of DENIS point sources*

N. Epchtein¹, E. Deul², S. Derriere³, J. Borsenberger⁴, D. Egret³, G. Simon⁵, C. Alard⁵, L.G. Balázs¹⁸, B. de Batz⁵, M.-R. Cioni², E. Copet⁶, M. Dennefeld⁴, T. Forveille⁹, P. Fouqué^{6,10}, F. Garzón¹¹, H.J. Habing², A. Holl¹⁸, J. Hron¹², S. Kimeswenger¹³, F. Lacombe⁶, T. Le Bertre⁷, C. Loup⁴, G.A. Mamon^{4,8}, A. Omont⁴, G. Paturel¹⁴, P. Persi¹⁵, A.C. Robin¹⁶, D. Rouan⁶, D. Tiphène⁶, I. Vauglin¹⁴, and S.J. Wagner¹⁷

- ¹ Observatoire de la Côte d'Azur, Département Fresnel, F-06304 Nice Cedex, France
² Leiden Observatory, University of Leiden, P.O. Box 9513, 2300 RA Leiden, The Netherlands
³ CDS, Observatoire astronomique de Strasbourg, UMR 7550, 11 rue de l'Université, F-67000 Strasbourg, France
⁴ IAP, Institut d'Astrophysique de Paris, 98 bis Bd. Arago, F-75014 Paris, France
⁵ DASGAL, Observatoire de Paris, 61 avenue de l'Observatoire, F-75014 Paris, France
⁶ DESPA, Observatoire de Paris, 5 place J. Janssen, F-92195 Meudon Cedex, France
⁷ DEMIRM, Observatoire de Paris, 5 place J. Janssen, F-92195 Meudon Cedex, France
⁸ DAEC, Observatoire de Paris, 5 place J. Janssen, F-92195 Meudon Cedex, France
⁹ Observatoire de Grenoble, 414 rue de la Piscine, Domaine Universitaire de Saint Martin d'Hères, F-38041 Grenoble, France
¹⁰ European Southern Observatory, ESO, La Silla, Chile
¹¹ Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain
¹² Institut für Astronomie der Universität Wien, Türkenschanzstrasse 17, A-1180 Wien, Austria
¹³ Institut für Astronomie, Innsbruck University, A-6020 Innsbruck, Austria
¹⁴ CRAL, Observatoire de Lyon, F-69561 Saint-Genis Laval Cedex, France
¹⁵ Istituto di Astrofisica Spaziale, CNR, Area di Ricerca Roma-Tor-Vergata, Italy
¹⁶ Observatoire de Besançon, B.P. 1615, F-25010 Besançon Cedex, France
¹⁷ Landessternwarte Heidelberg, Königstuhl, D-69117 Heidelberg, Germany
¹⁸ Konkoly Observatory, Box 67, H-1525 Budapest, Hungary

Received 17 June 1999 / Accepted 21 June 1999

Abstract. This paper announces the release at CDS of a substantial set of point sources detected by the DENIS project. DENIS is the first astronomical survey of the Southern sky in two near-infrared bands (J at $1.25 \mu\text{m}$, and K_s at $2.15 \mu\text{m}$) and one optical band (Gunn- i at $0.82 \mu\text{m}$), conducted by a European consortium, using the 1m telescope at ESO, La Silla (Chile).

The first data release, described here, consists of a preliminary set of about 17 million extracted point sources, corresponding to 102 strips (2% of the Southern sky), and resulting from observations performed in 1996.

Data are available through a World-Wide Web server at the CDS (Strasbourg astronomical Data Center):
<http://cdsweb.u-strasbg.fr/denis.html>.

Key words: infrared: stars – Galaxy: stellar content – stars: low-mass, brown dwarfs – astronomical data bases: miscellaneous – surveys

Send offprint requests to: S. Derriere

* Based upon observations collected at the European Southern Observatory, La Silla, Chile

Correspondence to: denis@simbad.u-strasbg.fr

1. Introduction

The main purposes of this paper are to announce the first public release of DENIS data, through the World-Wide Web server of the CDS, and to present the first steps in the validation process of these data, concerning their astrometric and photometric reliability.

The Deep Near-Infrared Survey of the Southern sky (DENIS) is a project to survey the southern sky in three wavelength bands (Gunn- i , $0.82 \mu\text{m}$; J , $1.25 \mu\text{m}$; and K_s , $2.15 \mu\text{m}$) simultaneously, with limiting magnitudes 18.5, 16.5 and 14.0, respectively (for a detailed description, see Epchtein 1998).

The project is managed by a European consortium, led by N. Epchtein (Observatoire de la Côte d'Azur), involving fifteen institutes in eight different countries. The observations are performed with the ESO 1m telescope at La Silla (Chile). The survey is carried out by observing strips of 30° in declination and $12'$ in right ascension with an overlap of $2'$ between consecutive strips. The observations started at the end of 1995 and will be completed by the end of 2000.

The data are reduced in two consecutive steps: the first is performed as a joint effort of the Institut d'Astrophysique de Paris (IAP) and Observatoire de Paris; the second is performed at the Leiden Observatory. The position of a general extracted

point source is provided with an accuracy better than $1''$ and its magnitude to better than 0.1 mag.

The *Centre de Données Astronomiques de Strasbourg* (CDS) is implementing the final databases and provides access to a preliminary set of processed and calibrated data, through a WWW server.

The scientific exploitation of the already existing data is carried out by several working groups, allowing a continuous monitoring of the quality of the survey data.

2. The data

2.1. The survey observations

The survey observations with all 3 cameras started in routine mode in December 1995 and were slowed down by technical and meteorological problems in 1997 and 1998. The survey is currently (June 1999) progressing at nominal speed.

Special efforts have been devoted to cover priority regions of high scientific interest, such as the Magellanic Clouds, the Galactic bulge, and giant molecular clouds.

The observing technique is the following: the sky, south of declination $+2^\circ$, has been divided into 5112 regions named *slots*. A slot is $12'$ wide in right ascension, and spans $30'$ in declination. Two adjacent slots have $2'$ overlaps in right ascension or declination. A DENIS *strip* corresponds to the observation of a slot. A given slot may be observed several times (for example if the photometric quality has been detected as insufficient by the data reduction software). Thus several strip numbers can refer to the same slot. A strip is made of 180 individual images ($12' \times 12'$), with $2'$ overlaps between consecutive images.

2.2. Point source data reduction

All the images received from La Silla are first processed at the Paris Data Analysis Center (PDAC), where the images are checked and calibrated (Borsenberger 1997); afterwards, the point source extraction (Bertin & Arnouts 1996) and the astrometric and photometric calibrations (Deul et al. 1995) are carried out at the Leiden Data Analysis Center (LDAC). Each strip is processed independently. Analysis of extended sources, carried out at PDAC, is not considered in the present article, which deals only with point sources.

2.3. Data selection

The preliminary database of DENIS point sources released at CDS provides, for 102 *strips* (as of June 1999), the three-colour information resulting from the reduction pipeline.

The available strips were selected according to their astrometric and photometric quality, in order to form a useful sample of the current DENIS point source catalogue. It is to be noted, however, that this is not a subset of the final catalogue of point sources, because some improvements in the reduction pipeline are still currently being implemented, and this will imply a new reduction of already observed data. These improvements are not,

however, of such a nature that they could change the statistical interpretation of the currently released data sets.

Within a strip, association has been done between bands (I , J , K_s), and overlapping images (with the notable exception of bad quality flagged sources which are not matched and appear duplicated). The positional coincidence is determined by examining the elliptical shape parametrization of each entry. When a source is present several times in a given band, the resulting merged entry in the catalogue carries the simple flux average for the magnitudes, and the weighted average for the coordinates. Eventual overlaps between adjacent ($12' \times 30'$) strips have not been matched in this preliminary release. These elements should be carefully taken into account when working on star counts derived from the released data.

It should also be noted that in the final DENIS catalogue, overlaps between adjacent strips will lead to a better accuracy in the photometry (see Sect. 3.2).

3. Data validation

One of us (S. Derriere) has performed a series of tests in order to assess the astrometric and photometric quality of the released point source data.

3.1. Astrometric validation

The catalogues of the strips in the first DENIS release contain a single entry for objects cross-identified in the same colour (frame overlap) or among different colours. About 10% of the objects are identified simultaneously in the three colour channels.

DENIS astrometry is performed in two steps: first, pairing of extracted objects is done between channels (I , J , K_s), in overlaps between consecutive frames, and with sources from an astrometric reference catalogue (ARC hereafter); then, the actual astrometric calibration based on this pairing information is performed, and a sky position is derived for each source.

Comparison of objects lying in overlapping strips indicates an internal accuracy of $0.5''$.

In the released DENIS data, the Guide Star Catalog (GSC, Lasker et al. 1990), which contains about 15 million stars, is used as the ARC. The astrometry of the final DENIS point source catalogue will be based on the Tycho catalogue (ESA 1997), through the use of the USNO A2.0 catalogue (Monet et al. 1997), itself calibrated on Tycho. Since an independent offset is indeed determined for each DENIS frame, the ARC must have a large average density (at least of the order of 1 star per DENIS frame).

In order to check for the accuracy of the DENIS astrometry, comparisons between positions in the DENIS and USNO A1.0 (also calibrated on the GSC) catalogues have been performed. The USNO A1.0 was preferred in order to have, for each DENIS frame, a sufficient number of objects likely to be detected in both catalogues. We selected for each DENIS frame all the DENIS and USNO A1.0 sources detected on that sky region, and for *each* single USNO A1.0 source we searched for all the DENIS sources located in its vicinity. We over-plotted on a single chart

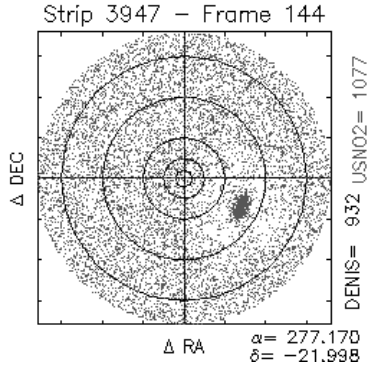


Fig. 1. Chart showing the systematic shifts between USNO A1.0 and DENIS positions for a given frame (example of bad quality astrometry leading to the rejection of the strip). Circles show the 1, 2.5, 5, 10 and 15'' ranges. The shift illustrated here was in the range 5 to 10 arcsec.

(as shown on Fig. 1) for each USNO A1.0 source, the relative position of all its DENIS neighbours closer than a given radius. Provided that some objects are detected in both catalogues, and that there is no significant shift between the two catalogues, there will be a cluster of points around relative position (0,0). In case there is a systematic shift, this cluster will be located around the relative position corresponding to the shift.

This operation was systematically applied to each DENIS frame, in order to reject strips with anomalous astrometric solution.

3.2. Photometric validation

The photometric calibration of each night is derived from observations of standard star fields; a fixed extinction coefficient is used to determine the zero point (hereafter ZP) for each strip (see details in Fouqué et al. 1999).

Although each standard star is observed 8 times on different parts of the chip, some of these measurements might later be rejected in the phase of source extraction and object characterization. On average this reduces the number of measurements from 8 to 6. In addition, only standard stars fainter than $I=10.5$ mag, $J=8.0$ mag, and $K=6.5$ mag are used for the photometric calibration.

A series of graphics (including colour-colour diagrams and differential star counts histograms) are routinely provided for each DENIS strip, together with the data. These graphs allow the user to check the contents and quality of individual strips. For example, colour-colour diagrams clearly show the effect of interstellar reddening, while star counts provide an estimate of the completeness limit for the corresponding strip.

Only strips observed during good photometric nights and with good zero points were released. Care was also exercised to

Table 1. Summary of catalogue record contents for DENIS point sources

Column	Contents
DENIS Id.	Identifier JHHMMSS.s±DDMMSS
Strip	DENIS strip number
RA, DEC	J2000 Right Ascension and Declination, and associated error
<i>For each I, J, K_s channel</i>	
Mag_Aper 7''	Magnitude derived from the flux collected in a fixed 7'' aperture centered on the source, and associated error
Mag_Aper 15''	Same as above with 15'' aperture, and associated error
Mag_Auto	Automatic (Kron type) scaling elliptical aperture magnitude
a, b, e, θ	Parameters of the object ellipse
Stellarity	1 for a star; 0 for an extended source; see Bertin & Arnouts (1996)
Flags	Extraction and Image Flags; Image Artifact marker; Merged object indicator
Field	Field Position reference

remove strips with heavily parasited frames. A small number of artifacts may, however, remain among the point sources: some of them, appearing as vertical triplets in J or K_s , are micro-scanning features due to poor pixels on the NICMOS cameras. They will, as much as possible, be removed from the final DENIS catalogue.

Fig. 2 shows the histogram of I , J , and K_s magnitudes for the 17 million point sources belonging to the first data release.

In the final DENIS catalogue, overlaps between adjacent strips will lead to a better accuracy in the photometry. Global photometry will be performed when a sufficient number of adjacent overlapping strips are reduced, by deriving variations in the originally derived nightly based ZPs. Comparison of objects lying in overlapping strips indicates an accuracy of 0.05 to 0.1 mag for the photometric zero-points.

4. Data organization

4.1. Catalogue contents

Table 1 gives an overview of the contents of an individual record in the preliminary database of DENIS point sources.

4.2. Data access

The released data are being made available through the CDS Web server at the following address:
<http://cdsweb.u-strasbg.fr/denis.html>

This server provides access to DENIS documentation, including scientific reports, list of publications, etc., and to the publicly released data sets. The database implementation benefits from the CDS environment (Genova et al. 1998): SIMBAD name resolver, VIZIER query engine (Ochsenbein 1998), ALADIN cross-identification atlas, etc.

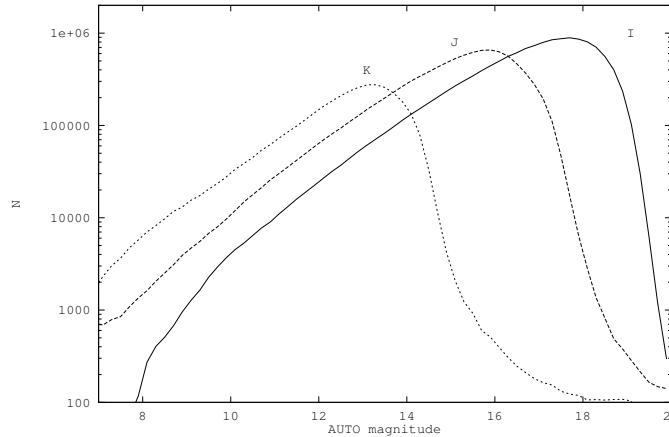


Fig. 2. Raw differential star counts for *I*, *J* and *K_s* (0.2 mag bins) cumulated for the 102 released strips (17.5 million sources).

4.3. Query modes

An interactive sky map provides a graphical access to individual strips, while an auxiliary database provides information about night and strip qualities. Strips can be searched directly by position or *strip* or *slot* number.

In addition an optimized query mechanism provides access to individual data records (including source number, position, *I*, *J*, *K_s* magnitudes, and estimated errors, and additional flags related to the source extraction) for the extracted point sources. The first release of the preliminary database includes about 17 million entries in 102 strips.

Queries can be made for all DENIS sources around a given position (center and radius, where the center is to be specified either as J2000 coordinates, or by the name of a central object to be resolved by SIMBAD). Selection of objects in the colour-colour diagrams of specified strips is also possible.

Finally, the DENIS point source database is also available through the VIZIER catalogue service, and through the ALADIN interactive sky atlas, where additional query modes and functionalities are available.

4.4. Future evolutions

The DENIS point source data server at CDS will continue to evolve in the future, with newly reduced data being gradually incorporated.

Changes in the data reduction pipeline will be posted in the corresponding information pages.

New features for an improved support of the query mechanism and source selection will also be made available as they are being developed. A mirror copy of the publicly released 2MASS *J*, *H*, *K_s* data (see e.g. Skrutskie 1998) is planned to become available, in parallel, at CDS, for easier cross-comparison.

In complement to the CDS services, observational DENIS data for galaxies are also being made available through the Lyon-Meudon Extragalactic Database (LEDA; see Sect. 5.6, below). DENIS data of Asymptotic Giant Branch (AGB) stars recognized as associated to sources included in the ASTRID database are also gradually integrated into this base.

5. Highlights of DENIS scientific results

5.1. Faint stars in the solar neighbourhood: red and brown dwarfs

The large range of DENIS wavelengths (from the *I*-band to the *K*-band is a factor 2.5 in wavelength) makes the survey data well suited for the detection of red dwarfs of the latest spectral types (M7 or later) about which little is known. The analysis of 220 square degrees of high galactic data led to the identification of 25 dwarfs later than M7V, doubling the known inventory of these very cool objects (Delfosse et al. 1999), and 3 brown dwarf candidates. One of these three is DENIS-P J1228-1547 (Delfosse et al. 1997); Keck spectroscopy showed the presence of a strong lithium absorption line at 680.7 nm (Martin et al. 1997; Tinney et al. 1997); in such fully convective objects lithium nuclear burning never took place and thus there has been no hydrogen burning either. In addition, this object is the first identified binary brown dwarf (Martin et al. 1999). Interestingly, one of the other two candidates DENIS-P J0205-1159 has no lithium (Martin, private communication) in spite of a significantly lower effective temperature.

Very few of these enigmatic objects are known. The extreme M dwarfs and brown dwarfs contained in the DENIS database constitute potentially the first statistically significant sample of these objects, sufficiently large to determine their local density (i.e. the disc luminosity function) and to form the basis for the calibration of the observable properties of brown dwarfs (flux

calibrated spectra and distances) into more fundamental physical quantities (effective temperature, luminosity, mass).

5.2. The interstellar medium

Detailed maps of the extinction in nearby molecular clouds (in Chameleon, Ophiuchus, Orion and Serpens) can be obtained via star counts in the DENIS data. An example is the study by Cambr sy et al. (1997) of the nearby Cham 1 cloud, that has produced an extinction map of much higher angular resolution than before.

The DENIS survey observations make possible to obtain similar maps of other well-known and nearby clouds, e.g. those in Ophiuchus and Serpens, but also in small globules. When these extinction maps are combined with maps in the CO millimeter line emission, with 60 and 100 μm maps obtained by IRAS, the distribution of gas, dust and of the temperature will be known much better than so far.

Counting stars in the whole DENIS data base will eventually provide the first deep maps of interstellar extinction throughout the Galaxy.

5.3. Young stars in dark clouds

The exploration of DENIS data of dense molecular clouds will lead to the detection of a large number of stars in the process of formation: young stellar objects (YSOs), protostars, etc. Copet (1996) has studied DENIS data of the Orion molecular cloud and discovered many new faint YSOs which are likely to be T Tau stars.

From an analysis of similar data, Cambr sy et al. (1998) detected YSOs that probably form a new population of low-luminosity T Tau stars: a full survey of the clouds in Chameleon (an area of several tens of square degrees) and in other similar regions will improve our knowledge of the low-end of the luminosity function.

5.4. The distribution of stars in the disc and in the bulge of our Galaxy

At visual wavelength the search for stars in the inner galactic disc is limited to distances smaller than one or two kiloparsecs, while at 2 μm red giants and supergiants are detected up to galactic center distances: the extinction in the infrared is much smaller ($A_K \approx A_V/10$). The DENIS large-scale digital survey in the infrared, with its homogeneous magnitude limits in its photometric bands – with the exception of a modest variation (0.6 mag) in the early data in the K_s -band – provides an ideal support for new studies of the distribution of dust, red giants and red supergiants in our Galaxy. The newly detected supergiants will give the distribution of the star formation rate and the red giants the distribution and evolution of older stellar populations in the galactic plane.

Detailed new information are expected to be obtained about the inner spiral arms, the ring, the bar and the bulge of the Galaxy. The COBE/DIRBE data have been important in find-

ing such large scale structures; the much higher angular resolution of DENIS (a few arcseconds instead of a few degrees with COBE) will give more detailed information.

The DENIS data will be used to interpret the data on the inner Galaxy at 7 and 15 μm obtained through the ISOGAL program (see P rault et al. 1996). Several other surveys of the inner Galaxy exist already, and we expect that the additional information from the DENIS and ISOGAL survey on the stellar content will be important, perhaps essential, to obtain a comprehensive model for the inner Galaxy.

The existing survey data have already significantly constrained models for the large scale structure of the Galaxy (Ruphy et al. 1996, 1997). The scale length in the galactic plane has been determined as well as the outer edge of the stellar disc, the position of the Sun relative to the galactic plane and the shape and density of some spiral arms and of the 3 kpc ring.

There is a strong interest in DENIS survey data on the bulge and the bar of our Galaxy: Schultheis (1998) analysed the DENIS measurements of $\sim 30,000$ sources in the Palomar-Groningen field #3 (PG3). Among these there are 36 previously known AGB variables; carbon stars are missing, confirming earlier studies.

Planetary nebulae (PNe) are an interesting category of objects as well. Already 250 of the 750 known PNe in the southern sky have been detected by the DENIS survey (Kimeswenger et al. 1998). The survey will provide better surface photometry at its three wavelengths than previously known. It will also discover faint, red, background stars that often contaminate photometry at visual wavelengths but that also may be used as distance indicators. A further goal concerns the few thousand PNe in the inner Galaxy that must exist, but that have escaped detection so far because of interstellar extinction. Kimeswenger et al. expect that DENIS will find many new of those.

Even in medium or high latitude fields, where interstellar extinction at visual wavelengths is not a problem, the DENIS survey data will identify red giants that can be used to trace their distribution in the halo, before optical surveys (such as the Sloan) produce more complete results.

5.5. Stars in the Magellanic Clouds

With a distance modulus of $(m - M) = 18.6$ for the Large Magellanic Cloud and $(m - M) = 19.1$ for the SMC, DENIS records data of all stars with $(V - I) > 3$ that have a luminosity higher than about $2500 L_\odot$ or $M_{\text{bol}} < -4.0$ and thus all AGB stars except the few that are very red. There are large differences in luminosities and element abundances between populations of AGB stars in different surroundings, e.g. the SMC, the LMC, the outer disc of our Galaxy and our bulge. Most AGB stars in the SMC are carbon-rich whereas in the galactic bulge the AGB stars are all of M-type (“oxygen-rich”). The reasons for these variations are not well understood at present and the DENIS samples may provide the data for a breakthrough.

Knowledge so obtained will be of great importance in the study of AGB stars in other Local Group dwarf galaxies, within

the grasp of large ground-based telescopes such as the ESO VLT.

A specific catalogue of DENIS point sources in both Magellanic Clouds is planned to become publicly available around mid 1999 (see Cioni, Loup & Habing 1999).

5.6. Galaxies and cosmology

Near infrared surveys such as DENIS offer substantial advantages for constructing samples of galaxies: the low extinction allows for a fair view of external galaxies, virtually unaffected by their interstellar media, and for samples extending to low galactic latitudes. Moreover, the near infrared domain provides galaxy samples that are more mass-weighted and less affected by recent star formation than galaxy samples in other wavebands.

Vauglin et al. (1999) have produced the first DENIS *I*-band extragalactic catalog, by extracting galaxies from highly compressed DENIS images, in a homogeneous manner. The *I*-band data are the best suited to separate stars from galaxies and to determine the parameters of extended objects. The first catalog, corresponding to DENIS observations prior to June 1997, contains 20,260 galaxies, of which roughly 2/3 are new, and is complete to $I \simeq 14.5$. The catalogue is available from CDS. The comparison of this catalogue with the sample of Mathewson et al. (1992) and Mathewson & Ford (1996) shows that the uncertainty in DENIS *I*-magnitude is about 0.18 mag at $I \simeq 14$.

An updated version of the catalogue, containing 41,000 galaxies, and covering the observations before February 1999, is in preparation.

A special project is the search for galaxies in the Zone of Avoidance in the direction of the "Great Attractor" (Kraan-Korteweg et al. 1998). A systematic search is being made at Lyon for galaxies at latitudes smaller than 15° . Routine analysis of the DENIS data base started in March 1997. More than 1500 extended objects have already been listed; some of these are not galaxies but newly discovered globular clusters or planetary nebulae.

It is possible to go as faint as $I = 16.5$ and extract galaxies with very high ($> 95\%$) completeness and reliability, as shown by Mamon et al. (1998). The galaxy counts in *I* and *J* follow the Euclidean 0.6 slope up to the completeness limit, with no lack of bright objects, in contrast to counts performed in the optical (Heydon-Dumbleton et al. 1989; Maddox et al. 1990), and in agreement with recent optical counts by Bertin & Dennefeld (1997). A complete and reliable catalogue with 900,000 galaxies in the *I*-band, 500,000 at *J* and 50,000 at K_s , is expected. The homogeneity of the extraction is essential for statistical cosmological studies of galaxies such as the measurement of the angular two-point correlation function of galaxies, the derivation of the primordial density fluctuation spectrum, the building of catalogs of groups and clusters, and the study of the variation of galaxy colours with environment. So far, such studies have been based on photographic photometry, which suffers from non-linearities and large systematics in photometric calibration.

The galaxies discovered in the DENIS survey are expected to be the input for future spectroscopic surveys of the southern sky (see e.g. Theureau et al. 1997; Mamon 1998).

6. Conclusion

The preliminary database of DENIS point sources (about 17 million sources for 2% of the Southern sky) can be used for assessing the interest of DENIS data for large scale scientific studies in a variety of fields.

Acknowledgements. The DENIS project was partly funded by the European Commission through *Science and Human Capital and Mobility* grants. It is also supported by INSU-CNRS, and Ministère de l'Éducation Nationale, de la Recherche et de la Technologie, in France; by the Land of Baden-Württemberg in Germany; by DGICYT in Spain; by CNR in Italy; by FWF and BMWF in Austria; by FAPESP, in Brazil; by OTKA grants F-4239 and F-013990 in Hungary; and by the ESO C&EE grant A-04-046. CDS acknowledges the support of INSU-CNRS, the Centre National d'Études Spatiales (CNES), and Université Louis Pasteur. François Ochsenbein has been very helpful in the final implementation of the database at CDS.

References

- Bertin E., Arnouts S., 1996, A&AS 117, 393
- Bertin E., Dennefeld M., 1997, A&A 317, 43
- Borsenberger J., 1997, In: The Impact of Near-Infrared Sky Surveys. Proc. 2nd Euroconf. on Near-Infrared Surveys, Kluwer ASSL 210, 181
- Cambrésy L., Epchtein N., Copet E., et al., 1997, A&A 324, L5
- Cambrésy L., Copet E., Epchtein N., et al., 1998, A&A 338, 977
- Cioni M.-R., Loup C., Habing H., 1999, In: Whitelock P., Cannon R. (eds.) IAU Symposium 192, The Stellar Content of Local Group Galaxies. Cape Town, South Africa, 7-11 September 1998, to be published by ASP, E2
- Copet E., 1996, Thèse de doctorat de l'Université Paris 6
- Delfosse X., Tinney C.G., Forveille T., et al., 1997, A&A 327, L25
- Delfosse X., Tinney C.G., Forveille T., et al., 1999, A&A 135, 41
- Deul E., Holl A., Guglielmo F., et al., 1995, In: Proc. of 1st Euro-conference on near-infrared sky surveys. San Miniato, Mem. Soc. Astron. Ital. 66, 549
- Epchtein N., 1998, The impact of near-infrared surveys on galactic and extragalactic astronomy. Proc. 3rd. Euroconf., Kluwer ASSL, 230
- ESA, 1997, The Hipparcos and Tycho Catalogues. ESA SP-1200
- Fouqué P., Chevallier L., Cohen M., et al., 1999, A&AS, in press
- Genova F., Bartlett J.G., Bonnarel F., et al., 1998, In: Astronomical Data Analysis Software and Systems VII, ASP Conf. Ser. 145, p. 470
- Heydon-Dumbleton N.H., Collins C.A., MacGillivray H.T., 1989, MNRAS 238, 379
- Kimeswenger S., Kerber F., Roth M., et al., 1998, A&A 332, 300
- Kraan-Korteweg R.C., Schröder A., Mamon G.A., Ruphy S., 1998, In: The Impact of Near-Infrared Sky Surveys on Galactic and Extragalactic Astronomy. Proc. 3rd Euroconf. on Near-Infrared Surveys, Kluwer ASSL 230, 209
- Lasker B., Sturch C.H., McLean B.J., et al., 1990, AJ 99, 2019
- Maddox S.J., Sutherland W.J., Efsthathiou G., Loveday J., Peterson B.A., 1990, MNRAS 247, 1P

- Mamon G.A., 1998, In: Colombi S., Mellier Y., Raban B. (eds.) 14th IAP meeting, Wide Field Surveys in Cosmology. Editions Frontières, Paris, p. 323 (astro-ph/9809376)
- Mamon G.A., Borsenberger J., Tricottet M., Banchet V., 1998, In: The impact of near-infrared surveys on galactic and extragalactic astronomy. Proc. 3rd. Euroconf., Kluwer ASSL 230, 177 (astro-ph/9712169)
- Martin E., Basri G., Delfosse X., Forveille T., 1997, A&A 327, L29
- Martin E.L., Brandner W., Basri G., 1999, Sci 283, 1718
- Mathewson D.S., Ford V.L., Buchhorn M., 1992, ApJS 81, 413
- Mathewson D.S., Ford V.L., 1996, ApJS 107, 97
- Monet D., Canzian B., Harris H., et al., 1997, The PMM USNO A1.0 Catalogue. US Naval Observatory Flagstaff Station
- Ochsenbein F., 1998, In: Astronomical Data Analysis Software and Systems VII, ASP Conf. Ser. 145, p. 387
- Pérault M., Omont A., Simon G., et al., 1996, A&A 315, L165
- Ruphy S., Robin A.C., Epchtein N., et al., 1996, A&A 313, L21
- Ruphy S., Epchtein N., Cohen M., et al., 1997, A&A 326, 597
- Schultheis M., 1998, Ph.D. Thesis, Vienna astronomical Observatory
- Skrutskie M., 1998, In: The impact of near-infrared surveys on galactic and extragalactic astronomy. Proc. 3rd. Euroconf., Kluwer ASSL 230, 11
- Theureau G., Patrel G., Vauglin I., 1997, In: Mamon G.A., Thuan T.X., Tran Van Thanh J. (eds.) XVIIth Moriond Astrophysics Mtg, Extragalactic Astronomy in the Infrared. Editions Frontières, Paris, p. 393
- Tinney C.G., Delfosse X., Forveille T., 1997, ApJ 490, L95
- Vauglin I., Patrel G., Borsenberger J., et al., 1999, A&AS 135, 133

An absolute calibration of DENIS (deep near infrared southern sky survey)

P. Fouqué^{1,2}, L. Chevallier², M. Cohen³, E. Galliano², C. Loup⁴, C. Alard¹⁵, B. de Batz⁵, E. Bertin⁴, J. Borsenberger⁴, M.R. Cioni⁶, E. Copet¹, M. Dennefeld⁴, S. Derriere⁷, E. Deul⁶, P.-A. Duc⁸, D. Egret⁷, N. Epchtein⁹, T. Forveille¹⁰, F. Garzón¹⁶, H.J. Habing⁶, J. Hron¹¹, S. Kimeswenger¹², F. Lacombe¹, T. Le Bertre¹³, G.A. Mamon^{4,14}, A. Omont⁴, G. Patureau¹⁷, S. Pau¹, P. Persi¹⁸, A.C. Robin¹⁹, D. Rouan¹, M. Schultheis⁴, G. Simon¹⁵, D. Tiphène¹, I. Vauglin¹⁷, and S.J. Wagner²⁰

¹ DESPA, Observatoire de Paris, 5 place J. Janssen, F-92195 Meudon Cedex, France

² European Southern Observatory, Casilla 19001, Santiago 19, Chile

³ Radio Astronomy Laboratory, 601 Campbell Hall, University of California, Berkeley, CA 94720, U.S.A.

⁴ Institut d'Astrophysique de Paris, 98 bis Bd. Arago, F-75014 Paris, France

⁵ DASGAL, Observatoire de Paris, 5 place J. Janssen, F-92195 Meudon Cedex, France

⁶ Leiden Observatory, University of Leiden, P.O. Box 9513, 2300 RA Leiden, The Netherlands

⁷ CDS, Observatoire Astronomique de Strasbourg, UMR 7550, 11 rue de l'Université, F-67000 Strasbourg, France

⁸ CEA, DSM, DAPNIA, Centre d'Études de Saclay, F-91191 Gif-sur-Yvette Cedex, France

⁹ Observatoire de la Côte d'Azur, Département Fresnel, F-06304 Nice Cedex, France

¹⁰ Observatoire de Grenoble, 414 rue de la Piscine, Domaine Universitaire de Saint Martin d'Hères, F-38041 Grenoble, France

¹¹ Institut für Astronomie der Universität Wien, Türkenschanzstrasse 17, A-1180 Wien, Austria

¹² Institut für Astronomie, Innsbruck University, A-6020 Innsbruck, Austria

¹³ DEMIRM, Observatoire de Paris, 61 Av. de l'Observatoire, F-75014 Paris, France

¹⁴ DAEC, Observatoire de Paris, 5 place J. Janssen, F-92195 Meudon Cedex, France

¹⁵ DASGAL, Observatoire de Paris, 61 Av. de l'Observatoire, F-75014 Paris, France

¹⁶ Instituto de Astrofísica de Canarias, E-38200 La Laguna, Tenerife, Spain

¹⁷ CRAL, Observatoire de Lyon, F-69561 Saint-Genis Laval Cedex, France

¹⁸ Istituto di Astrofisica Spaziale, CNR, C.P. 67, I-00044 Frascati, Italy

¹⁹ Observatoire de Besançon, BP. 1615, F-25010 Besançon Cedex, France

²⁰ Landessternwarte Heidelberg, Königstuhl, D-69117 Heidelberg, Germany

e-mail: pfouque@eso.org

Received June 24; accepted October 20, 1999

Abstract. An absolute calibration of the DENIS photometric system is presented. It includes the determination of the overall transmission profiles in the 3 bands, namely i , J and K_s , combining contributions from atmosphere, telescope mirrors, instrument lenses and dichroics, filters, and detectors. From these normalized profiles, isophotal and effective wavelengths are computed, using the same synthetic Vega spectrum as that used to support the absolute calibration of many other ground-based and spaceborne photometric systems. Flux densities at zero magnitude are derived and integrated to give in-band fluxes, which are used to compute theoretical zero-points and compare them to observed ones, yielding estimates of the overall throughput of the whole system.

Key words: surveys — infrared: general — instrumentation: miscellaneous

1. Introduction

The main goal of the DENIS survey (Deep Near-Infrared Southern Sky Survey, see Epchtein et al. (1994) for a complete introduction to DENIS) is to bridge the gap between the optical surveys on Schmidt plates and the far-infrared IRAS survey. Many aspects of astrophysics will benefit from such a survey, particularly studies of cool stars and heavily obscured regions.

Each night, roughly 1 million stars are detected by at least one of the 3 cameras of the DENIS instrument.

Send offprint requests to: P. Fouqué

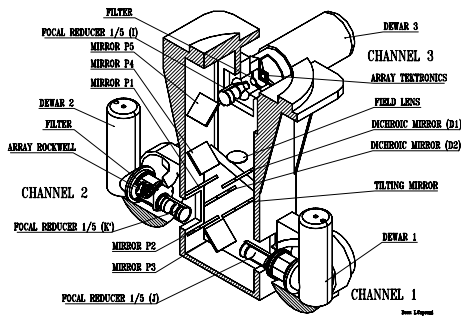


Fig. 1. Optics of the DENIS instrument

Photometric calibration is derived by observation of standard star fields. In order to compare our magnitude system to published ones, we need a precise definition of our photometric bands and an absolute calibration of the DENIS photometric system.

In Sect. 2, we describe the DENIS instrument and show the response curve for the complete system in the three bands. In Sect. 3, we estimate the conversion factors (ADUs to electrons) from the typical characteristics of DENIS images. Absolute calibration, based upon a synthetic Vega spectrum, is performed in Sect. 4, and observed and theoretical zero-points are compared.

2. Instrument characteristics

The DENIS instrument has been described in detail by Copet et al. (1999). A sketch of its main optical components is displayed in Fig. 1. It is located at the Cassegrain focus of the ESO 1 m telescope at La Silla Observatory (Chile). After reflection from the two telescope mirrors, the light beam goes through a field lens at the telescope focus, covered by a protective blade, both of CaF_2 and uncoated. Then a dichroic splits the i beam in reflection from the J/K_s beam in transmission.

The i beam has two more reflections from coated mirrors before entering the objective (3 CaF_2 and 2 silica coated lenses), then goes through the Gunn i filter, a shutter, the cryostat entrance window (BK7) and arrives at the Tektronix 1 K CCD detector, cooled to 180 K.

The J/K_s beam is reflected off a microscanning mirror (uncoated Al), then J is reflected by a second dichroic and a coated mirror before entering the J objective (3 CaF_2 and 2 fused silica coated lenses), then the cryostat entrance window (coated fused silica), the filter and finally the NICMOS-3 detector, both cooled to 80 K.

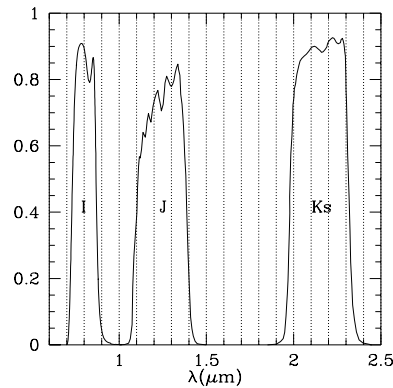


Fig. 2. The 3 DENIS filter response curves

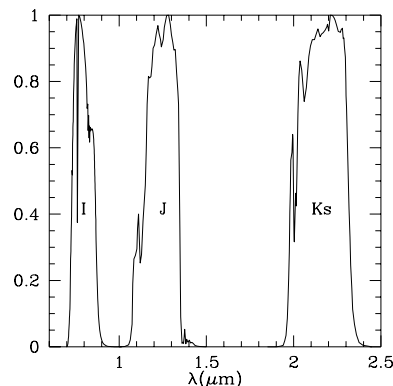


Fig. 3. The complete DENIS system normalized response curves (atmosphere, optics, filters, detectors)

The K_s beam is transmitted by the second dichroic, then reflects off two more coated mirrors, passes through the K_s objective (4 ZnS-Cleartran and 1 fused silica coated lenses), the cryostat entrance window (coated fused silica), the filter and the other NICMOS-3 detector.

We have tried to obtain response curves for all these elements. When this was not possible, a reasonable estimate of the transmission was adopted. Filter response curves as provided by Barr Associates (U.S.A.) for J and K_s , at the nominal detector temperature (77 K), and by MTO-France for the Gunn i filter (at ambient temperature) are displayed in Fig. 2. The full system response curves (atmosphere, optics, filter and detector) are shown in Fig. 3.

Table 1. Conversion factors (expected and measured, in e^-/ADU) in the three bands

Gain	<i>i</i>	<i>J</i>	<i>K_s</i>
Expected	3.08	18.4	58.6
Measured	3.04	11.4	56.2
Std dev	0.01	0.1	0.1

3. Characteristics of DENIS images and conversion factors

The first characteristics generally measured on astronomical images are the “sky emission” level (which includes telescope and instrument background in the *K_s* band), and the noise on sky images and darks.

Variations in sky level are observed in all three bands. In *i*, they are related to the distance to, and phase of, the Moon. In *J*, they are due to variations in the hydroxyl radical’s emission intensity (OH^- Meinel bands), connected to the passage of density and temperature perturbations through the upper atmosphere (Ramsay et al. 1992). In *K_s* they come from the temperature variations. Also note that, in crowded fields, the background value is set by the confusion level (background of faint undistinguishable stars).

From these variations, the conversion factor between ADU (analog-to-digital units) and electrons can be deduced. Table 1 compares the result of these gain measurements with the expected values, calculated from the electronics characteristics of the chips, the preamplifiers gains, and the analog-to-digital conversions. Table 2 gives for each band the median value of the read-out noise in electrons, the median and faintest values of the sky level, in mag arcsec^{-2} , and the median and minimal values of the sky image noise in electrons, calculated from more than 2000 images taken during the last year of observations (April 1998 to April 1999), and adopting zero-points of 23.5, 21.3, and 19.2 in *i*, *J*, and *K_s*, respectively (see Sect. 4 and Table 5). The large values of sky and sky noise in *K_s* come from the thermal background of the instrument, which does not include cool stop optics. The “best” value of the sky level in *J* is suspect.

Note that some care must be taken in applying the conversion factors and zero-points to the whole DENIS survey: first, *J* and *K_s* conversion factors seem to be slightly variable (11.4 to 15.3 in *J*, 51.6 to 57.0 in *K_s*), and second, changes in the instrument have altered these values: the pre-amplifier boards of *J* and *K_s* cameras were changed in June 1996, and a spare *J* camera has been in use from April 3, 1998 to May 9, 1999. Old conversion factors valid until June 1996 were 12.47 in *J* and 39.3 in *K_s* (Chevallier 1996). Zero-point variations will be analyzed in a future paper.

Table 2. Read-out noise, sky level, and sky image noise in the three bands

Band	RON e^-	SKY		SKY noise	
		median mag arcsec^{-2}	best mag arcsec^{-2}	median e^-	min e^-
<i>i</i>	7.3	19.2	20.2	15	11
<i>J</i>	21	16.6	18.1?	37	27
<i>K_s</i>	40	11.2	12.0	311	222

Table 3. Mean filter wavelengths (pure filter, and filter + atmosphere + detector, hereafter “fad”), effective and isophotal wavelengths for Vega

Wavelength (μm)	<i>i</i>	<i>J</i>	<i>K_s</i>
λ_{fil}	0.802	1.248	2.152
λ_{fad}	0.795	1.235	2.160
λ_{eff}	0.788	1.221	2.144
λ_{iso}	0.791	1.228	2.145

Table 4. Flux densities of a zero magnitude star for the three DENIS bands

Band	λ_{iso} μm	F_{λ} $\text{W}/\text{m}^2/\mu\text{m}$	F_{ν} Jy
<i>i</i>	0.791	$1.20 \cdot 10^{-8}$	2499
<i>J</i>	1.228	$3.17 \cdot 10^{-9}$	1595
<i>K_s</i>	2.145	$4.34 \cdot 10^{-10}$	665

4. Absolute calibration

The next interesting characteristics to establish for absolute photometric work is the flux of a zero magnitude star in the three DENIS bands. Many infrared systems and several ways to calibrate them exist. We have decided to use the calibration scheme described by Cohen et al. (1992): they start from a model of Vega from Kurucz (1991), taking into account its lower than solar metallicity, and normalize it to $F_{5556} = 3.44 \cdot 10^{-8} \text{ W m}^{-2} \mu\text{m}^{-1}$ from Hayes (1985). Additionally, we adopt $V = 0.03 \text{ mag}$, $V - I = 0$, so $I = 0.03 \text{ mag}$, but $JHKLM = 0.00 \text{ mag}$ for this star. For a more detailed discussion of Vega magnitudes and colours, see Bessell et al. (1998).

We must now determine the isophotal wavelength of each filter, taking into account the filter response curve, the atmospheric transmission, the detector radiance response and the Vega spectrum. Isophotal wavelengths are preferred over effective wavelengths, because the latter vary with input source spectrum much more than do the former (see Golay 1974, for details and definitions). Results are given in Table 3.

Using the isophotal wavelengths of our bands, and the Vega spectrum, we can compute the flux densities for a zero magnitude star. Table 4 gives the results in wavelength and frequency units.

Table 5. Theoretical and observed zero-points. Derived overall transmission and its components (τ , ρ and QE correspond to transmission, reflection and quantum efficiency, respectively)

Parameter	i	J	K_s
Theoretical zero-point	25.10	23.29	20.43
Observed zero-point ($7''$)	23.4	21.1	19.1
Corrected zero-point	23.5	21.3	19.2
Measured overall transmission	0.24	0.16	0.32
$\tau_{\text{atmosphere}}$	0.955	0.912	0.912
τ_{mirrors}	0.868 ²	0.964 ³	0.978 ³
$\tau_{\text{blade, field lens}}$	0.938 ²	0.939 ²	0.940 ²
$\tau_{\text{dichroics}}$		0.784	0.810 \times 0.870
$\rho_{\text{dichroics}}$	0.993	0.970	
$\tau_{\text{coated mirrors}}$	0.993 ²	0.970	0.97 ²
$\tau_{\text{objective}}$	0.979 ³ \times 0.982 ²	0.979 ³ \times 0.985 ²	0.70 \times 0.982
$\tau_{\text{cryostat window}}$	0.94	0.985	0.982
τ_{filter}	0.909	0.846	0.926
QE_{detector}	0.65	0.8	0.8
Resulting overall transmission	0.31	0.32	0.25

From these flux densities, we can estimate how many ADUs would be measured if the atmosphere, telescope and instrument totally transmitted the photons from this zero magnitude star. Comparing to the actually observed zero-point will give the overall transmission of the system. We first integrate the product of the Vega spectrum (shifted by 0.03 mag in i) by the transmission of the full system (fad), over the wavelength domain of our filters (λ_0 and λ_1 correspond to the first and last wavelengths where filter transmission reaches 0), to obtain the measured flux of a zero magnitude star:

$$F_t = \int_{\lambda_0}^{\lambda_1} S(\lambda) \frac{F_\lambda(\text{Vega})}{h\nu} d\lambda. \quad (1)$$

The theoretical zero-point is given by:

$$ZP_{\text{th}} = 2.5 \log(F_t \times A t / G), \quad (2)$$

where A is the unobscured telescope collecting area (0.68 m²), t is the effective integration time (8.998 s in i , 8.809 s in J and K_s), and G is the conversion factor. Table 5 gives the results.

Zero-points are measured during calibration nights, where only photometric standards are observed, and routinely during survey nights, to follow possible instrumental variations and make a rough estimate of the extinction coefficients. They are measured from aperture magnitudes inside a 7 arcsec diameter circle around the standard star. To make a valid comparison with the theoretical zero-points, a first correction is necessary to include flux falling outside this aperture. This has been estimated to amount to 0.1 mag in all three bands from a comparison of observations through a 15 arcsec diameter aperture.

A linear fit of the observed magnitudes vs. airmass (assuming that Bouguer's 1729, law is valid) gives the extinction coefficient as the slope and the zero-point as the

intercept. However, it is well known that extrapolation to zero airmass leads to a systematic error in the near-infrared (the Forbes 1842, effect), which has been quantified for the J and K bands by Manduca & Bell (1979). For the La Silla typical water vapour contents (1 to 10 mm of precipitable water), the error is about 0.10 mag in J and 0.02 mag in K , and should be similar in K_s . Therefore, we also add this J correction to the observed zero-point, while we neglect the K correction, given the uncertainty in measured zero-points. The corrected zero-points (for infinite aperture and non-linear variation with airmass) are given in Table 5, and the derived overall transmissions follow.

To interpret the measured overall transmission, we have tried to estimate the contribution of each component of the system, namely atmosphere, aluminium reflections (telescope mirrors, and microscanning mirror for J and K_s), thin blade protecting the field lens, field lens itself, dichroics, coated mirrors, objectives, cryostat entrance windows, filters, and detector quantum efficiency (converted to radiance response). For details about each value, see Galliano (1999). Table 5 gives all these estimates, and their final product. The agreement with the measured overall transmission is satisfying, and shows a good performance of the instrument, with overall throughput of 20 to 30% in all three bands.

Acknowledgements. Thanks to Maria Eugenia Gómez for finding the original reference of Bouguer's law.

The DENIS project is supported by the *SCIENCE and the Human Capital and Mobility* plans of the European Commission under grants CT920791 and CT940627, by the French Institut National des Sciences de l'Univers, the Ministère de l'Éducation Nationale and the Centre National de la Recherche Scientifique, in Germany by the State of Baden-Württemberg, in Spain by the DGICYT, in Italy by the Consiglio Nazionale delle Ricerche, by the Austrian Fonds zur Förderung der wissenschaftlichen Forschung und

Bundesministerium für Wissenschaft und Forschung, in Brazil by the Foundation for the development of Scientific Research of the State of São Paulo (FAPESP), and by the Hungarian OTKA grants F-4239 and F-013990, and the ESO C & EE grant A-04-046.

References

- Bessell M.S., Castelli F., Plez B., 1998, *A&A* 333, 231 (and erratum in *A&A* 337, 321)
- Bouguer P., 1729, *Essai d'optique sur la gradation de la lumière*, C. Jombert, Paris, reprinted in: *Les maîtres de la pensée scientifique*, Paris, 1921
- Chevallier L., 1996, *Performances et qualité d'images du survey DENIS*, DEA (master's) thesis. University of Strasbourg
- Cohen M., Walker R.G., Barlow M.J., Deacon J.R., 1992, *AJ* 104, 1650
- Copet E., Epchtein N., Rouan D., et al., 1999, *A&AS* (in preparation)
- Epchtein N., de Batz B., Copet E., et al., 1994, in: *Proceedings of a Conference held at Les Houches: "Science with astronomical near-infrared sky surveys"*, Epchtein N., Omont A., Burton W.B., Persi P. (eds.), "A Deep Near-Infrared Survey of the Southern Sky", september 1993, p. 3, Les Houches, France, reprinted from *Astrophysics and Space Science* 217, 3
- Forbes J.D., 1842, *Phil. Trans.* 132, 225
- Galliano E., 1999, *Une calibration absolue du système photométrique du survey DENIS*, DEA (master's) thesis. University of Grenoble
- Golay M., 1974, in: *Introduction to astronomical photometry*. Reidel, Dordrecht, p. 39
- Hayes D.S., 1985, in: *Calibration of Fundamental Stellar Quantities*, Proc. IAU Symposium No. 111, Hayes D.S., Pasinetti L.E. & Davis Philip A.G. (eds.). Reidel, Dordrecht, p. 225
- Kurucz R.L., 1991, quoted in Cohen et al. (1992) (as private communication)
- Manduca A., Bell R.A., 1979, *PASP* 91, 848
- Ramsay S.K., Mountain C.M., Geballe T.R., 1992, *MNRAS* 259, 751

CDS Services in support of survey activity: Data access, identification and validation.

D. Egret, J. G. Bartlett, F. Bonnarel, J.-M. Cochery, S. Derriere, P. Fernique,
F. Genova, M. Louys, F. Ochsenbein, P. Ortiz, M. Wenger

¹*CDS, Observatoire astronomique de Strasbourg, UMR CNRS 7550,
11 rue de l'Université, 67000 Strasbourg, France*

Abstract. CDS develops services providing integrated access to information related to astronomical objects: basic data, catalogues and tables, bibliographical references, digitized sky images, etc. New projects are under way for the management of very large catalogues and surveys, and for providing data mining tools for future surveys.

1 Introduction

Access to the results of sky surveys is critical for many studies: search for new or rare categories of objects, statistical studies over complete samples of objects, analysis of spatial distribution, reference sky for new observations, etc.

As a service to the astronomical community, the Strasbourg astronomical data centre (CDS; see e.g. [3]) develops specific tools for managing and making accessible published data from sky surveys, deep fields, or very large catalogues at any wavelength; and for helping cross-identification, and validation of data from new surveys.

2 Current services at CDS in support of survey activity

Among all the services available at CDS¹, those described below are more specifically designed for approaching survey data.

2.1 Vizier

The Vizier and Astronomer's Bazaar services [4] provide access to the most complete astronomical library of published catalogues and data tables, including key catalogues and surveys, such as, for example, Hipparcos and Tycho catalogues, IRAS catalogs, VLA FIRST survey, USNO A1.0 catalog, etc. Those services allow respectively sampling and full retrieval of the catalogues over the Internet, through World-Wide Web or ftp protocols.

Specifically, original dedicated management techniques have been developed

¹Internet address: <http://cdsweb.u-strasbg.fr>

for fast coordinate access to very large catalogues (of size > 20 million objects) such as the Hubble Guide Star Catalog (GSC) or USNO A1.0.

2.2 SIMBAD

The SIMBAD database is the reference database for identification and bibliography of astronomical objects [2]. It contains millions of cross-identifications of sources for objects referenced in the literature.

2.3 Aladin

The ALADIN interactive sky atlas provides access to reference images from the large Schmidt surveys (DSS I and later DSS II, completed with high resolution MAMA scans) as well as a set of tools to help source diagnostic and identification, including the capacity of overlaying information from VizieR or Simbad [1].

3 New projects at CDS

Two new projects are also currently developed at CDS:

- an object-oriented management system for very large catalogues, storing several very large catalogues together, with fast multi-criterion access; current studies at CDS make use of a commercial system: Objectivity. As part of this project, strategies for multi-dimensional indexing and sky partitioning are also studied;
- an automatic cross-referencing engine for catalogues —a collaboration between ESO and CDS, in the framework of the VLT scientific environment. This tool will support the user for cross-correlations between “any” kind of data tables: ESO catalogues, VizieR catalogues, surveys, or tables, user table, etc.

These projects, which will be interfaced in the future, will allow CDS to provide convenient access and data mining tools for large survey catalogues such as DENIS, 2MASS, Megaprime/TERAPIX, etc.

References

- [1] Bartlett, J. G., Egret, D. 1997, in *IAU Symp. 179*, Kluwer Acad. Publ. p. 437
- [2] Egret, D., Wenger, M., Dubois, P. 1991, in *Databases & On-line Data in Astronomy*, Kluwer Acad. Publ., p. 79
- [3] Genova F. et al. 1998, in *Astronomical Data Analysis Software and Systems VII*, ASP Conf. Ser. 145, p. 466
- [4] Ochsenbein F. 1998, in *Astronomical Data Analysis Software and Systems VII*, ASP Conf. Ser. 145, p. 387

Mining survey data: CDS services for catalogue access, identification and validation.

D. Egret, S. Derriere, F. Genova, J. Bartlett, F. Bonnarel, P. Fernique, M. Louys, F. Ochsenbein, P. Ortiz and M. Wenger

CDS, Observatoire astronomique de Strasbourg, UMR CNRS 7550, 11 rue de l'Université, 67000 Strasbourg, France

Abstract.

CDS already provides the astronomical community with services like SIMBAD, VizieR and ALADIN. New projects are under way for mining survey data, and managing very large catalogues, like the DENIS point source catalogue.

1. Introduction

Mining sky survey data is becoming critical for many studies at infrared and other wavelengths: search for new or rare categories of objects, statistical studies over complete samples of objects, analysis of spatial distribution, references for new kinds of observations, etc.

The Strasbourg astronomical data center (CDS) has developed a series of tools for:

- managing and making accessible published data extracted from sky surveys, deep fields, or more generally, very large catalogues at any wavelength;
- helping cross-identification and validation of data from new surveys.

The main CDS services¹ are briefly reviewed in section 2, while section 3 describes the new projects currently developed for the management of very large catalogues and cross-referencing. Section 4 is devoted to the role that CDS will play in the DENIS project.

2. The key services at CDS

2.1. SIMBAD

SIMBAD is the reference database for the identification and bibliography of astronomical objects (Egret *et al.* 1991).

It contains about 2,000,000 objects, 5,000,000 object names, and more than 100,000 bibliographical references containing 2,600,000 citations of objects (as

¹<http://cdsweb.u-strasbg.fr/>

of July 1998), making it an invaluable tool for cross-identification of objects referenced in the literature.

2.2. Vizier

The Vizier and Astronomer's Bazaar services (Ochsenbein 1998) provide access to a wide library of data tables, including key catalogues and surveys (such as: Hipparcos, Tycho, IRAS, VLA FIRST, etc.), and allow respectively sampling and full retrieval of the catalogues over the network (1815 catalogues available in July 1998).

2.3. ALADIN

The ALADIN interactive sky atlas provides access to reference images from the major Schmidt surveys (DSS I and later DSS II, and high resolution MAMA scans) as well as a set of tools to help source diagnostic and identification.

The user can also visualize his/her own digitized images, and overlay catalogue (or user) data on the images, making it a very useful tool for multi-wavelength cross-identifications.

X-Windows and JAVA interfaces to ALADIN are currently developed.

2.4. Interoperability and performance

The different CDS services are completely inter-connected (Genova *et al.* 1998), and one can navigate easily between them (for example from ALADIN to Vizier, and then from Vizier to SIMBAD).

The GLU system (Fernique *et al.* 1998), also developed at CDS, helps in maintaining connexion between those services (and also with external services).

In the frame of the Vizier service, dedicated management techniques have been developed for fast coordinate access to very large catalogues (of size > 20 Million objects) such as GSC or USNO A1.0. This approach gives very good response times, for example, in USNO A1.0 (488,006,860 sources):

- 0.09 seconds for a search radius of 2.5 arcmin (14,000 records);
- 0.67 seconds for a search radius of 30 arcmin (200,000 records).

Although no indexing is made for magnitudes and colors, a query based on color criteria can be done on the complete catalogue in about 35 minutes, thanks to heavy (but lossless) data compression.

3. New projects at CDS

Two new projects are currently under way at CDS:

- a very large catalogue management system, fully object-oriented, storing several very large catalogues together, with fast multi-criterion access. Current studies at CDS make use of a commercial DBMS: Objectivity.

As a part of this project, new methods for multi-dimensional indexing and sky partitioning are also studied;

- an automatic cross-referencing engine for catalogues (collaboration between ESO and CDS in the framework of the VLT scientific environment). This tool will allow the user to perform cross-correlation between “any” kind of data (ESO catalogue, VizierR catalogue, user table, ...).

Fig. 1 shows the access page of a prototype for this facility.

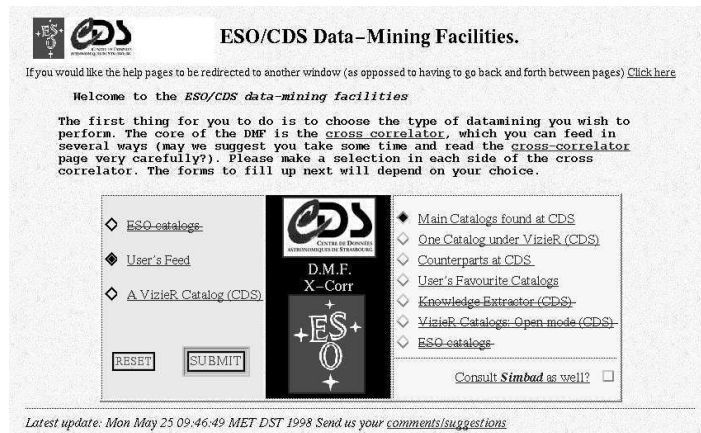


Figure 1. Prototype of the cross-referencing engine for catalogues: access page.

These projects, which will be interfaced in the future, will allow CDS to provide convenient access and data mining tools for large survey catalogues such as DENIS, MegaPrime/Terapix, 2MASS, etc.

4. DENIS data access at CDS

Observations and on-line checks for the DENIS project (Epchtein 1998) are made at ESO, La Silla. The PDAC (IAP, Paris) performs flat-fielding and quality control on the images, and the LDAC (Leiden) produces source extraction and calibration. A systematic catalogue of extragalactic sources (REDCAT) is also produced by the LEDA group in Lyon. The final release of the DENIS catalogues will be done through the CDS. Access to the images will be through PDAC.

Before DENIS data becoming public, CDS is implementing the software tools needed for managing DENIS catalogues. As a first step, a display of data reduction progress at LDAC, via a queriable map giving information and strip access is provided (figure 2).

The next step will be to integrate a dedicated management technique for providing direct on-line access to the DENIS point sources, pointers to the image

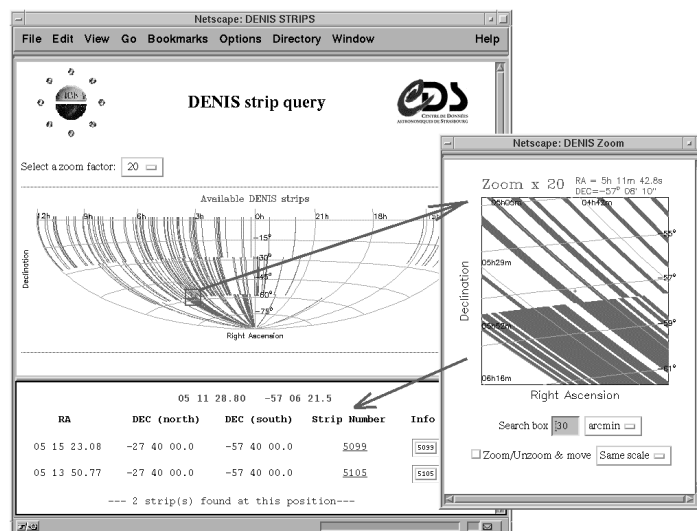


Figure 2. Prototype of the interface for DENIS data access: strip selection from an interactive map.

archive, and facilities for validation and cross-identification through the other CDS services.

References

- Egret, D., Wenger, M., & Dubois, P. 1991, in *Databases and on-line data in astronomy*, Albrecht, M. A. & Egret, D., eds., Kluwer Acad. Publ., pp.79-88.
- Epchtein, N. 1998 (this conference)
- Fernique, P., Ochsenbein, F., & Wenger, M. 1998, *Astronomical Data Analysis Software and Systems VII*, A.S.P. Conference Series, Vol. 145, R. Albrecht, R.N. Hook & H.A. Bushouse, eds., p.466
- Genova, F., Bartlett, J. G., Bonnarel, F., Dubois, P., Egret, D., Fernique, P., Jasniewicz, G., Lesteven, S., Ochsenbein, F., & Wenger, M. 1998, *Astronomical Data Analysis Software and Systems VII*, A.S.P. Conference Series, Vol. 145, R. Albrecht, R.N. Hook & H.A. Bushouse, eds., p.470
- Ochsenbein, F. 1998, *Astronomical Data Analysis Software and Systems VII*, A.S.P. Conference Series, Vol. 145, R. Albrecht, R.N. Hook & H.A. Bushouse, eds., p.387

On-line access to very large catalogues

S. Derriere, F. Ochsenbein, D. Egret

CDS, Observatoire Astronomique de Strasbourg, UMR 7550, 11 rue de l'Université, 67000 Strasbourg, France

Abstract.

Dedicated tools have been developed at CDS in order to face the challenge of fast on-line access to very large astronomical catalogues. Powerful compression methods, keeping the direct access to the data on the basis of their celestial position, allow very fast queries; their usage for CDS services such as VizieR or Aladin is presented.

1. Introduction

Astronomy is entering the era of very large catalogues. Digitizations of photographic plates (USNO, GSC II) or current digital sky surveys (DENIS, 2MASS, Sloan DSS) lead to 10^8 — 10^9 object catalogues. And the number of parameters for each object (positions, magnitudes, errors, flags, etc.) is also considerably increasing.

Gigabyte is the basic size unit for the resulting catalogues, and it becomes impossible for each astronomer to copy locally the full amount of data. Queries over the net, extracting smaller samples with pertaining information, are thus a requisite for such catalogues.

CDS¹ already provides fast on-line access to some very large catalogues which are fully integrated into other CDS services.

2. Compression and Query Mechanism

2.1. Binary Compression

When running queries on databases of several hundred million objects, response time becomes critical. A fast access to the database requires that i/o be reduced as far as possible, and thus to have a powerful compression tool to downsize the amount of data.

We use, for each catalogue, a dedicated and optimized compression method to store the data into binary files:

- all parameters are first converted into integer numbers, and all unused bits are removed. The records are then grouped by their position on the sky, meaning that contiguous records have quite similar values for α and δ ;

¹<http://cdsweb.u-strasbg.fr/>

- a “reference” position is then stored as a “header” of a group of records, and individual positions are stored as offsets (i.e. very small integers) from the “reference” position;
- additional flags can be added if they help in compressing the data. For example, if a catalogue contains measurements in several photometric bands which are not necessarily available for each object, we can add a single bit, telling if there are data in the band, to every record. That way, we can avoid coding empty values.

E.g., simply by storing only position offsets, the 6 Gbytes of the USNO A1.0 catalogue (Monet et al. 1997) were compressed down to 3.4 Gbytes (Ochsenbein 1998).

2.2. Query Processing

Most of the queries include a selection based on celestial positions, and thus the compressed binary database is indexed on position. Additional filters (on colors or flags) may be applied to the selected sample if necessary.

When a query to such a large catalogue is submitted to VizieR², the CDS database of catalogues and tables (Ochsenbein et al. 2000), a dedicated program handles the request, extracts and decodes the matching data, and sends its output back to VizieR.

An overview of the system is shown in Fig. 1.

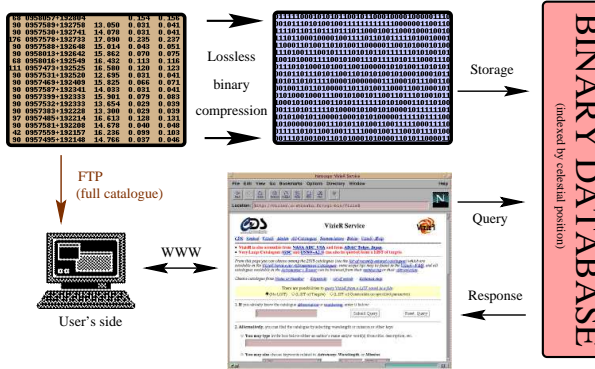


Figure 1. Construction of the database and queries processing.

Table 1 gathers some figures for queries performed on some large catalogues using a Sparc-20 station (72 MHz). On this machine, the complete USNO A1.0 is read in about 40 minutes.

²<http://vizier.u-strasbg.fr/>

Table 1. Example of performances achieved.

Catalogue	Record length (bytes)	Reading time (per object)
USNO A1.0	7	5 μ s
DENIS PSC	42	40 μ s
2MASS PSC	60	138 μ s

3. Integration in CDS Services

Thanks to the specific tools that have been developed, very large catalogues can be fully integrated in CDS services such as VizieR or Aladin.

For example, one can select in VizieR all the USNO sources around a given target, with additional constraints on the magnitudes. In Aladin³, the interactive digitized sky atlas at CDS (Bonnarel et al. 2000), very large catalogues can be overlaid on digitized sky images just like any other catalogue, as illustrated in Fig. 2.

The currently (November 1999) available large catalogues include:

- the Guide Star Catalog (Lasker et al. 1990, Russel et al. 1990, Jenkner et al. 1990) (25,258,765 entries), in 3 different versions (GSC 1.1, GSC 1.2, GSC-ACT);
- the USNO catalogues (Monet et al. 1997) versions 1.0 and 2.0 (488,006,860 and 526,280,881 sources, respectively);
- the first DENIS release (17,471,214 point sources);
- the 2MASS (Skrutskie et al. 1997) Spring 1999 release (20,204,378 point sources).

For the DENIS project (Epchtein et al. 1999), beyond the access to the point source catalogue, the CDS provides other on-line services⁴, like access to a database of observation strips (quality, number of sources, etc.).

Note that the DENIS and 2MASS catalogues are growing as new data become public.

4. Conclusions

Thanks to dedicated tools, large-size catalogues can now be handled as easily as the more “classical” ones. This will allow further integration of upcoming very large catalogues into CDS services.

³<http://aladin.u-strasbg.fr/>

⁴<http://cdsweb.u-strasbg.fr/denis.html>

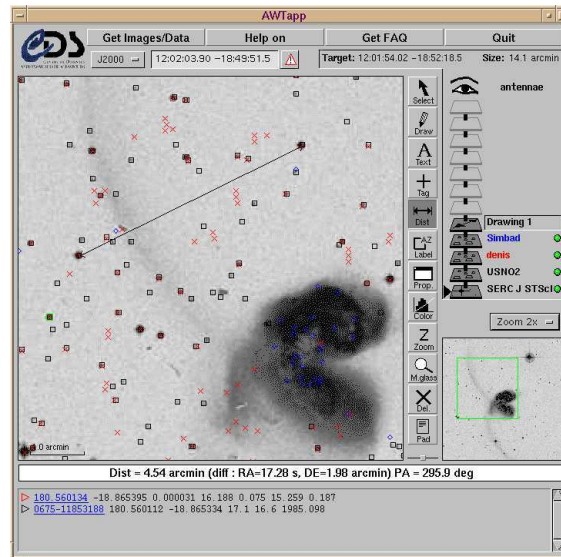


Figure 2. Projection of DENIS (crosses) and USNO A2.0 (squares) sources on an DSS-I blue plate image using Aladin (the field displayed is a 7 arcmin box in the Antennae galaxies).

References

- Bonnarel, F., Fernique, P., Bienaymé, O. et al., 2000, A&A, in preparation
 Epchtein, N., Deul, E., Derriere, S. et al., 1999, A&A, 349, 236
 Jenkner, H., Lasker, B.M., Sturch, C.R. et al., 1990, AJ, 99, 2082
 Lasker, B.M., Sturch, C.R., McLean, B.J. et al., 1990, AJ, 99, 2019
 Monet, D., Bird, A., Canzian, B. et al., 1997, USNO-A1.0, (U.S. Naval Observatory, Washington DC)
 Ochsenbein, F., 1998, in ASP Conf. Ser., Vol. 145, Astronomical Data Analysis Software and Systems VII, ed. R. Albrecht, R. N. Hook, & H. A. Bushouse (San Francisco: ASP), 387
 Ochsenbein, F., Bauer, P., Genova, F. et al., 2000, A&A, in preparation
 Russell, J.L., Lasker, B.M., McLean, B.J. et al., 1990, AJ, 99, 2059
 Skrutskie M.F., Schneider S.E., Stiening R. et al., 1997, Proc. Workshop "The Impact of Large Scale Near-IR Sky Surveys", 25

****TITLE****
*ASP Conference Series, Vol. **VOLUME**, **PUBLICATION YEAR***
****EDITORS****

Near-infrared surveys and the shape of the Galactic Disc

S. Derriere

*CDS, Observatoire de Strasbourg, 11, rue de l'Université, Strasbourg,
F-67000, France*

A.C. Robin

*Observatoire de Besançon, CNRS-ERA6091, BP 1615, F-25010,
Besançon cedex, France*

Abstract.

We present results from the comparison of selected fields observed in the near-infrared by DENIS with predictions from a modified version of the Besançon synthetic model, taking into account a possibly warped and flaring galactic disc. We find evidence for a warp and flare in the stellar disc. The warp starts approximately at the solar circle, whereas the galactocentric radius for the beginning of the flare depends on galactic longitude.

1. Introduction

The near-infrared surveys currently under completion (DENIS and 2MASS) will provide 3-colors observations of thousands of square degrees with much fainter magnitudes than achieved before. The lower extinction at those wavelengths makes these surveys well suited for studying galactic structure. In this paper we focus on the study of a warp and flare of the Milky Way disc using DENIS data.

Warps are quite a common feature in external galaxies (Bosma 1991). Since radio observations, it has been known that the Milky Way itself hosts a warped and flaring disc. Maps of the gas distribution were obtained with 21 cm observations (Burton & de Lintell Hekker 1986; Burton 1988). The warp and flare in the Galaxy have also been detected in the interstellar dust (Freudenreich et al. 1994), in molecular clouds (May et al. 1997), in IRAS sources (Djorgovski & Sosin 1989), or in the distribution of O-B stars (Miyamoto et al. 1988; Reed 1996; Drimmel et al. 2000).

In order to study the warp in the stellar disc, we use a modified version of the Besançon stellar population synthesis model (Robin & Crézé 1986; Bienaymé et al. 1987; Haywood et al. 1997) and compare model predictions with DENIS data in selected regions. This allows to improve our knowledge of the warped and flaring disc.

2

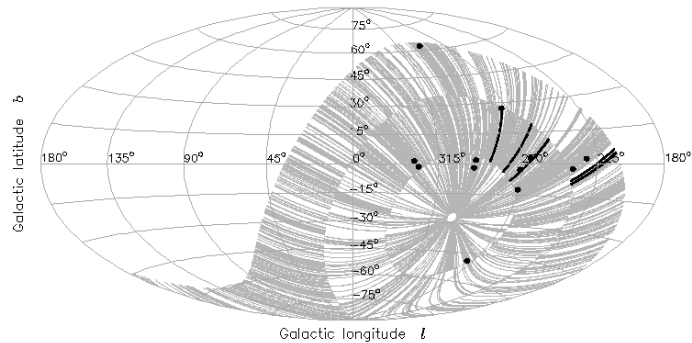
Derriere, S. & Robin, A.C.

Figure 1. Location of the selected regions of the DENIS survey. Circles stand for small regions, and lines are DENIS selected strips. Pre-reduced strips appear as a grey background.

2. Data selection

The DENIS survey (Epchtein 1997) is conducted on a 1 m telescope at La Silla. Observations are made simultaneously in three bands: I , J and K_S (Fouqué et al. 2000). The limiting magnitudes are respectively 18.5, 16.5 and 14.0, with photometric accuracy better than 0.1 mag.

We here use DENIS data point sources recently extracted at the Paris Data Analysis Center (Borsenberger, 1997) with photometric values derived from a PSF fit procedure.

According to radio observations (Burton 1988), the sun lies within 10° of the warp's line of nodes in the galactic plane. This implies that the maximum displacement of the disc below the plane in the south will occur between the 3rd and 4th quadrants.

In order to study this effect, we have selected two kinds of data sets: small regions symmetrically distributed on both sides of the plane, and DENIS strips ($30^\circ \times 12'$ regions) crossing the plane at various longitudes, as plotted on Fig. 1. The small regions will be used to constrain global parameters, before using the strips for constraining the warp and the flare.

3. Warping and flaring the galactic disc

The selected data are compared with a model of stellar population synthesis of the Galaxy, in which we include parameters describing a possibly warped and flaring disc.

Assuming stellar evolution theory, a scheme of galactic evolution and dynamical constraints, the Besançon model (Robin & Crézé 1986; Bienaymé et al. 1987; Haywood et al. 1997), allow to simulate catalogues of stars which can be directly compared with observations.

To be able to reproduce the data in the selected regions, we have to model the warp of both the thin and thick discs with a tilted ring model, such as described in (Porcel et al. 1997). For galactocentric radii $R > R_{\text{warp}}$, the mid-disc is shifted perpendicular to the plane by a value

$$z_{\text{warp}} = z_c \cos(\theta - \theta_{\text{max}}). \quad (1)$$

The warp has its maximum amplitude z_c towards θ_{max} , and this amplitude increases linearly as $z_c = \gamma_{\text{warp}}(R - R_{\text{warp}})$.

As in Gyuk et al. (1999), we model the flaring of the discs by increasing the scale heights by a factor k_{flare} , beyond a galactocentric radius R_{flare} , with

$$k_{\text{flare}} = 1 + \gamma_{\text{flare}}(R - R_{\text{flare}}). \quad (2)$$

We must ensure that the surface density remains constant in this process. One must compute a normalization factor f , when increasing the scale height from h_z to $k_{\text{flare}}h_z$. For the thick disc, for example, the spatial density is a truncated exponential:

$$\rho = \rho_0 \exp\left(-\frac{R}{h_R}\right) \begin{cases} 1 - \frac{z^2}{2h_z h_c + h_c^2} & \text{for } |z| < h_c \\ \frac{\exp(h_c/h_z)}{1+h_c/2h_z} \exp\left(-\frac{|z|}{h_z}\right) & \text{for } |z| > h_c. \end{cases} \quad (3)$$

With a flaring disc, ρ has to be corrected by a factor:

$$f = \frac{(2h_z + h_c)(3k_{\text{flare}}h_z h_c + 3k_{\text{flare}}^2 h_z^2 + h_c^2)}{(2k_{\text{flare}}h_z + h_c)(3h_z h_c + 3h_z^2 + h_c^2)}. \quad (4)$$

We assume in the following that the sun exactly lies on the line of nodes ($\theta_{\text{max}} = 90^\circ$).

4. Dealing with extinction

The stellar warp can be detected only at low galactic latitudes, where the extinction plays a very important role. The use of near-infrared data benefits from the lower absorption at those wavelengths (Bessel & Brett (1988) give $A_J/A_V = 0.60$, $A_J/A_V = 0.29$, and $A_{K_s}/A_V = 0.11$), but extinction cannot be neglected.

The density law of the absorbing material in the Besançon model corresponds to a thin disc, with an Einasto (1979) density law of scale length 140 pc and scale length 5000 pc.

This model is valid for high galactic latitudes. For some of the small selected regions at low latitudes, there are some strong discrepancies between the data and simulations with this simple absorption model, suggesting a strongest absorption (see Fig. 2). In such regions the modeled extinction is modified by adding clouds at different positions along the line of sight. The number of

4

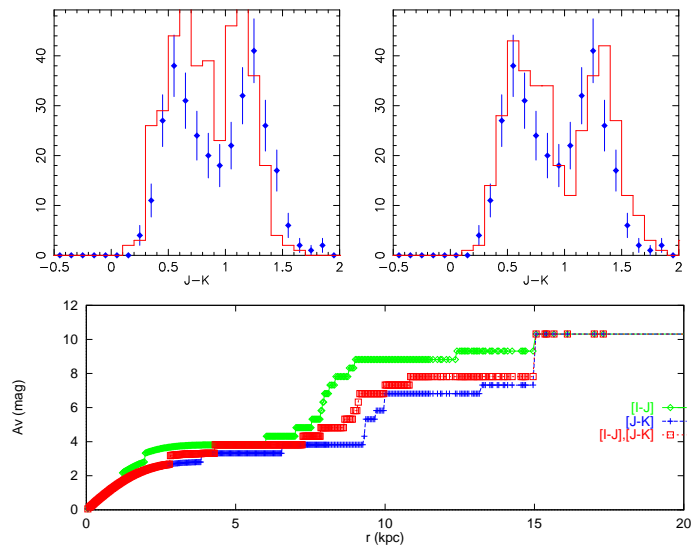
Derriere, S. & Robin, A.C.

Figure 2. Effect of the extinction towards $l = 264^\circ, b = +2.7^\circ$. The two upper graphs show DENIS [J-K] histogram for sources with $12 < K < 13$ (symbols), and the model (solid line) before (left) and after (right) adjusting the extinction. The bottom graph displays the absorption distribution fit for 3 different observables.

clouds, the internal extinction of each cloud and its distance are determined by fitting simulated colour histograms to the observed ones. We progressively add, in the simulated catalogue, clouds of variable visual absorption at different distances along the line of sight, in order to minimize the χ^2 between the observed histogram and the modelled one. Using clouds of decreasing absorption, this iterative procedure converges, and a new distribution of the absorbing material is found.

In Fig. 2, we apply this method to a field located at $l = 264^\circ$ and $b = +2.7^\circ$. The lower figure represents the reconstructed absorption: there is a steep increase (nearly 4 visual magnitudes) at distances of 7–9 kpc. This position coincides with the Perseus arm (Vallée 1995).

5. Results on galactic structure

Assuming that the stellar warp and flare follow the same shape as the gas, we make comparisons on small regions to put constraints on structure parameters

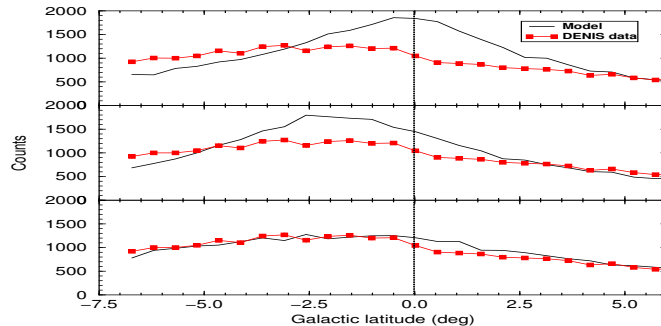


Figure 3. Latitude profile, crossing $b = 0^\circ$ at $l = 266.1^\circ$ (star counts in 0.8° bins in declination, in the DENIS strip, compared to those of the model). Top panel: flat disc. Middle panel: warped disc. Bottom panel: warped and flaring disc.

(with a fitted absorption when necessary). We find for the thin disc a scale length $h_R = 2.5$ kpc, and a radial cutoff $R_{\text{cut}} = 14$ kpc (Ruphy et al. (1996) found $h_R = 2.3$ kpc and $R_{\text{cut}} = 15$ kpc with the first DENIS data). The local normalization of the thick disc is assumed to be about 6% (Robin et al., 1996).

For all small regions symmetrically distributed around $b = 0^\circ$, the density of sources is higher at negative latitudes. We have plotted on Fig. 3 the effect of a warp and flare (identical to the gas layer) on a given latitude profile. Introducing a warp shifts the peak of the star counts towards negative latitudes, and the flaring spreads the overall profile.

We attempt to fit latitude profiles (star counts for bright objects detected in the K band) in the selected strips, adjusting the value of the galactocentric radii where the warp and flare start. It appears impossible to constrain R_{warp} towards $l = 225^\circ$, as the amplitude is very small. The strips around $l = 270^\circ$ favour a warp starting at $R_{\text{warp}} = (8400 \pm 200)$ pc, very close to the distance adopted for the Sun ($R_\odot = 8500$ pc).

The starting position for the flare has been found to depend on galactic longitude. It starts further away towards the anticenter : $R_{\text{flare}} = 8800$ pc at 270° and $R_{\text{flare}} = 10500$ pc at 225° .

We assumed here a unique scale length for the discs. Any dependency of the scale length on longitude, or deviations from the exponential distribution might contribute to this effect.

6. Conclusions

Basic star counts from the DENIS survey and latitude profiles show a strong evidence for a warp and flare in the stellar disc. From selected fields of the DENIS

survey, the stellar warp is found to start at $R_{\text{warp}} = (8400 \pm 200)$ pc, while the flare might depend on longitude (R_{flare} ranging from 8800 to 10500 pc).

Further work on an extended set of zones should allow to tighten the constraints on the structural parameters of the stellar warp and flare.

References

- Bessell, M. S. & Brett, J. M. 1988, *PASP*, 100, 1134
 Bienaymé, O., Robin, A. C., & Crézé, M. 1987, *A&A*, 180, 94
 Borsenberger, J. 1997, *The Impact of Large Scale Near-IR Sky Surveys*, *ASSL 230*, Kluwer ed., p. 181
 Bosma, A. 1991, in *Warped Disks and Inclined Rings around Galaxies*, Knudsen ed., p. 181
 Burton, W. B. 1988, in *Galactic and Extragalactic Radio Astronomy*, Springer-Verlag, p. 295
 Burton, W. B. & te Lintel Hekkert, P. 1986, *A&AS*, 65, 427
 Djorgovski, S. & Sosin, C. 1989, *ApJ*, 341, L13
 Drimmel, R., Smart, R. L., & Lattanzi, M. G. 2000, *A&A*, 354, 67
 Einasto, J. 1979, *IAU Symposia*, 84, 451
 Epchtein, N. 1997, *The Impact of Large Scale Near-IR Sky Surveys*, p. 15, Springer-Verlag
 Fouqué, P., Chevallier, L., Cohen, M., Galliano, E., Loup, C., Alard, C., de Batz, B., Bertin, E., Borsenberger, J., Cioni, M. R., Copet, E., Dennefeld, M., Derriere, S., Deul, E., Duc, P., Egret, D., Epchtein, N., Forveille, T., Garzón, F., Habing, H. J., Hron, J., Kimeswenger, S., Lacombe, F., Le Bertre, T., Mamon, G. A., Omont, A., Paturel, G., Pau, S., Persi, P., Robin, A. C., Rouan, D., Schultheis, M., Simon, G., Tiphène, D., Vauglin, I., & Wagner, S. J. 2000, *A&AS*, 141, 313
 Freudenreich, H. T., Berriman, G. B., Dwek, E., Hauser, M. G., Kelsall, T., Moseley, S. H., Silverberg, R. F., Sodroski, T. J., Toller, G. N., & Weiland, J. L. 1994, *ApJ*, 429, L69
 Gyuk, G., Flynn, C., & Evans, N. W. 1999, *ApJ*, 521, 190
 Haywood, M., Robin, A. C., & Crézé, M. 1997, *A&A*, 320, 428
 May, J., Alvarez, H., & Bronfman, L. 1997, *A&A*, 327, 325
 Miyamoto, M., Yoshizawa, M., & Suzuki, S. 1988, *A&A*, 194, 107
 Porcel, C., Battaner, E., & Jimenez-Vicente, J. 1997, *A&A*, 322, 103
 Reed, B. C. 1996, *AJ*, 111, 804
 Robin, A. C. & Crézé, M. 1986, *A&A*, 157, 71
 Robin, A. C. Haywood, M., Crézé, M., Ojha, D.K., & Bienaymé, O. 1986, *A&A*, 157, 71
 Ruphy, S., Robin, A. C., Epchtein, N., Copet, E., Bertin, E., Fouque, P. & Guglielmo, F. 1996, *A&A*, 313, L21
 Vallee, J. P. 1995, *ApJ*, 454, 119

Résumé

La première partie de cette thèse concerne la gestion des très grands catalogues astronomiques. Deux grands projets de cartographie du ciel dans l'infrarouge proche, DENIS et 2MASS, étaient en cours pendant ce travail de thèse. Je présente les méthodes de compression, sans perte d'information et préservant une indexation sur la base des positions célestes, développées pour les catalogues de sources ponctuelles de ces deux relevés.

Le codage optimisé permet un gain de compression d'un facteur 1.5 à 2 par rapport à des algorithmes de type LZ77. Les premières versions des catalogues DENIS et 2MASS sont pleinement intégrées aux services du CDS (VizieR, Aladin), et donc interrogeables par le Web, avec des temps d'accès moyens inférieurs à 20 μ s par source.

Dans le cadre du projet DENIS, le travail de validation des données a permis, dans un premier temps, la première diffusion publique pour ce relevé (<http://cdsweb.u-strasbg.fr/denis-public/>). La comparaison des données réduites par les deux chaînes de traitement DENIS (LDAC et PDAC) a ensuite permis de définir des critères de filtrage et d'utiliser les données PDAC pour l'étude de la structure Galactique.

La deuxième partie de la thèse s'appuie sur une version modifiée du modèle de synthèse de populations stellaires de Besançon. Les lois de densité du disque ont été modifiées pour pouvoir simuler un disque stellaire gauchi et/ou évasé.

Les comparaisons avec les données PDAC indiquent une échelle de longueur $h_R = 2.5$ kpc pour le disque, avec une troncature à $R = 14$ kpc. Je discute une méthode de reconstruction tridimensionnelle de l'extinction par ajustement du modèle sur les données DENIS. Le gauchissement du disque galactique est observé entre $l = 230$ et 330° , et se traduit par une inclinaison du disque stellaire vers $b < 0$, comme pour le gaz. Le gauchissement commence près de la position du Soleil ($R = 8.4$ kpc), et l'origine de l'évasement montre une dépendance en longitude.

Abstract

The first part of this thesis concerns the management of very large astronomical catalogues. Two large projects aiming at covering the sky at near-infrared wavelength were running during this thesis : DENIS and 2MASS. I present the dedicated lossless compression methods, preserving an indexation based on celestial positions, that have been developed for the point source catalogues of these two surveys.

Optimized coding leads to compression factors 1.5 to 2 times better than those achieved by LZ77-type algorithms. The first versions of the DENIS and 2MASS catalogues are fully integrated into the CDS services (VizieR, Aladin), and can be queried via the Web, with average response times below 20 μ s per source.

Within the frame of the DENIS project, data validation has allowed the first public release for this survey (<http://cdsweb.u-strasbg.fr/denis-public/>). Comparison of data reduced by the two DENIS pipelines (LDAC and PDAC) led to the definition of filtering criteria for selecting PDAC data suitable for Galactic structure studies.

The second part of this thesis is based on the use of a modified version of the Besançon synthetic model of stellar population synthesis. The disc density laws have been changed to allow simulation of a warped and/or flaring stellar disc.

Comparisons with PDAC data yield a 2.5 kpc scale length for the thin disc, with radial cutoff at 14 kpc. I present a method for reconstructing the 3D structure of extinction by fitting the model on DENIS data. The warp is observed between $l = 230$ and 330° , bending the disc towards $b < 0$, as in the gas. The warp is found to start close to the Sun ($R = 8.4$ kpc), and the radius for the flare depends on the galactic longitude.