



METODOLOGIA COMPUTACIONAL PARA O ESTUDO DE GRUPOS
ESTELARES EM MOVIMENTO

Josina Oliveira do Nascimento

Tese de Doutorado apresentada ao Programa de Pós-graduação em Engenharia de Sistemas e Computação, COPPE, da Universidade Federal do Rio de Janeiro, como parte dos requisitos necessários à obtenção do título de Doutor em Engenharia de Sistemas e Computação.

Orientador: Valmir Carneiro Barbosa

Rio de Janeiro
Setembro de 2019

METODOLOGIA COMPUTACIONAL PARA O ESTUDO DE GRUPOS
ESTELARES EM MOVIMENTO

Josina Oliveira do Nascimento

TESE SUBMETIDA AO CORPO DOCENTE DO INSTITUTO ALBERTO LUIZ
COIMBRA DE PÓS-GRADUAÇÃO E PESQUISA DE ENGENHARIA (COPPE)
DA UNIVERSIDADE FEDERAL DO RIO DE JANEIRO COMO PARTE DOS
REQUISITOS NECESSÁRIOS PARA A OBTENÇÃO DO GRAU DE DOUTOR
EM CIÊNCIAS EM ENGENHARIA DE SISTEMAS E COMPUTAÇÃO.

Examinada por:

Prof. Valmir Carneiro Barbosa, Ph.D.

Prof. Felipe Maia Galvão França, Ph.D.

Profa. Priscila Machado Vieira Lima, Ph.D.

Prof. Vladimir Garrido Ortega, D.Sc.

Profa. Lúcia Maria de Assumpção Drummond, D.Sc.

RIO DE JANEIRO, RJ – BRASIL
SETEMBRO DE 2019

do Nascimento, Josina Oliveira

Metodologia Computacional para o estudo de Grupos Estelares em Movimento/Josina Oliveira do Nascimento. – Rio de Janeiro: UFRJ/COPPE, 2019.

XV, 83 p.: il.; 29, 7cm.

Orientador: Valmir Carneiro Barbosa

Tese (doutorado) – UFRJ/COPPE/Programa de Engenharia de Sistemas e Computação, 2019.

Referências Bibliográficas: p. 80 – 83.

1. Principal Component Analysis - PCA . 2. Workflows científicos. 3. stellar moving groups. 4. Beta Pictoris Moving Group - BPMG. 5. AB Doradus Moving Group. I. Barbosa, Valmir Carneiro. II. Universidade Federal do Rio de Janeiro, COPPE, Programa de Engenharia de Sistemas e Computação. III. Título.

Dedico à minha mãe como representante de todas as mães, avós e bisavós
que mantêm-se ativas cuidando de seus filhos, netos e bisnetos
Mesmo quando trabalham
Mesmo quando batalham
Mesmo quando são grandes no trabalho
Mas não deixam de cuidar
Unir, abraçar
Parar para brincar ou contar uma história
Deitar junto
Afagar, cuidar e cuidar
Ouvir, interferir
E saber calar
Dedico à minha mãe como representante de todas as mães, avós e bisavós
que mantêm-se ativas cuidando de seus filhos, netos e bisnetos
Mesmo quando parecem não estar aqui...
Por toda a eternidade

Agradecimentos

Astronomia e computação são minhas paixões profissionais. Juntar as duas em um trabalho de tese foi um sonho desde o tempo do mestrado. Realizar esse sonho sendo orientada pelo Valmir foi muito maravilhoso. Por uma razão desconhecida, o destino foi pródigo na criação de obstáculos desde o primeiro mês do doutorado. O resumo de todo esse tempo é que eu tentava estudar e calcular sobre as estrelas, enquanto perdia o chão. Só foi possível dar prosseguimento e finalizar esse trabalho porque meu orientador é um cientista incrível, um orientador extraordinário e um ser humano raro. Assim, todo agradecimento que eu lhe faça será sempre um infinitésimo.

Ao meu orientador, Professor Valmir Carneiro Barbosa, por sua orientação precisa, paciente, humana e persistente. Que os frutos desse trabalho sejam o seu melhor presente!

Ao Professor Vladimir Garrido Ortega pela amizade e pelas aulas diárias e entusiasmadas sobre galáxias, dinâmica e formação estelar.

À Professora Priscila Machado por sua inestimável ajuda e empenho na formatação do texto e cumprimento dos prazos finais.

Ao Professor Felipe França pelo conhecimento e força transmitidos.

Ao Jailson Alcaniz por ter me motivado a ingressar no doutorado e pela renovação dessa motivação ao longo de todo esse tempo.

Ao Claudio Bastos pela motivação, pela confiança em meu trabalho e pela liberação do que foi possível.

Aos amigos Alba Livia e Luciano Alberto pela amizade e pelo empenho na minha liberação dos processos de iluminação pública.

À Simone Daflon e Roberto Chan pela força e pela ajuda.

Ao Bruno Bazzanella por ter fornecido o fonte da implementação da abordagem de Ortega.

Ao Alexandre Henrique Porto pelo auxílio na implementação da PCA.

A todos que com seu trabalho e dedicação contribuíram com todos os dados e conhecimentos utilizados nesse trabalho. Em especial, agradeço à Profa. Maria de Fátima Saraiva e ao Prof. Kepler de Oliveira (UFRGS) que mantém o rico, preciso e atualizado sítio de Astronomia e Astrofísica e ao Prof. Carlos Alberto Torres (LNA) por seu árduo e dedicado trabalho na produção do catálogo SACY.

Aos funcionários do PESC por sua dedicação, presteza e cortesia.

Às colegas Andrea, Edijane, Gisela e Tamires pela ajuda sempre de boa vontade.

Ao Eduardo Matera por ter atendido a todos os meus pedidos de ajuda com a maior rapidez e perfeição.

Aos amigos Carol, Diego e Hugo pela força e auxílio em latex.

Ao amigo Eugênio Neto por ter se dedicado ao nosso projeto AstroEducadores, de forma a me deixar cuidar da finalização do doutorado (e também por me obrigar a almoçar).

Às amigas de vida e oração Irene, Claudia, Celina, Regina e Lorena.

Aos amados irmãos Jacira, Jussara, Josete e Dilson pela força e pela ajuda e especialmente à Janete por sua ajuda constante e por ser meu porto seguro.

Às minhas amadas filhas Carina, Diana, Marina e Ana Beatriz, aos meus amados filhos Ulisses e João Matheus e à minha amada sobrinha-filha Bruna, cada um ao seu jeito em seu tempo e lugar, por tudo que fizeram para ajudar e por todo empenho em fazer isso.

Aos meus amados netos Flor, Estrela, Sol, Tiê e Otto por seu encantamento e por saberem me esperar.

Ao meu amado Sérgio pelo companheirismo sempre.

A todos que me ajudaram, alguns anônimos, nos quais encontrei amparo, abraço e palavras de carinho, ânimo e coragem.

Resumo da Tese apresentada à COPPE/UFRJ como parte dos requisitos necessários para a obtenção do grau de Doutor em Ciências (D.Sc.)

METODOLOGIA COMPUTACIONAL PARA O ESTUDO DE GRUPOS ESTELARES EM MOVIMENTO

Josina Oliveira do Nascimento

Setembro/2019

Orientador: Valmir Carneiro Barbosa

Programa: Engenharia de Sistemas e Computação

O estudo de grupos estelares em movimento está proximamente ligado ao estudo da formação estelar e à determinação da idade de estrelas, bem como ao da origem e evolução de sistemas planetários extrassolares. A missão Hipparcos (1989 a 1993) realizou medidas astrométricas de alta qualidade e forneceu catálogos contendo posições, paralaxes e movimentos próprios de alta precisão. Com a adição de velocidades radiais obtidas por outros meios, isso tornou possível a simulação retrógrada da evolução de grupos contendo algumas dezenas de estrelas a fim de determinar se constituem ou não grupos em movimento, bem como seus locais de origem e suas idades. Esta tese apresenta uma nova metodologia computacional voltada a analisar milhares de estrelas concomitantemente com o objetivo de refinar o entendimento atual quanto à composição de grupos já conhecidos, acrescentando a eles novas estrelas ou retirando estrelas que neles se encontrem. O cerne da metodologia consiste em representar os milhares de estrelas sendo consideradas através das componentes principais dos dados que as descrevem (velocidades, magnitudes, tipos espectrais, etc.). Da análise dessa representação surgem estrelas candidatas a inclusão em determinado grupo ou a exclusão dele, as quais são então submetidas a sucessivas simulações retrógradas para confirmação. Como estudo de caso, a metodologia foi aplicada ao grupo estelar em movimento Beta Pictoris, o que resultou em propostas de alteração do entendimento atual sobre ele.

Abstract of Thesis presented to COPPE/UFRJ as a partial fulfillment of the requirements for the degree of Doctor of Science (D.Sc.)

COMPUTATIONAL METHODOLOGY FOR THE STUDY OF STELLAR MOVING GROUPS

Josina Oliveira do Nascimento

September/2019

Advisor: Valmir Carneiro Barbosa

Department: Systems Engineering and Computer Science

The study of stellar moving groups is closely tied to the study of stellar formation and the determination of stellar age, as well as to the study of the origin and evolution of extrasolar planetary systems. The Hipparcos mission (1989 to 1993) performed high-quality astrometric measurements and resulted in catalogs comprising high-precision stellar positions, parallaxes, and proper motions. With the addition of radial velocities obtained by other means, this made possible the retrograde simulation of the evolution of groups comprising a few tens of stars, aiming to establish whether or not they constitute moving groups as well as where and when they originated. This thesis presents a new computational methodology tailored to the concomitant analysis of thousands of stars, aiming to refine the field's current understanding of already known groups by enlarging them with the addition of new stars or by removing stars from them. The core of this methodology is the representation of such thousands of stars by the principal components of the data describing them (velocities, magnitudes, spectral types, etc.). Analyzing the resulting representation gives rise to candidate stars that are then subjected to a succession of retrograde simulations for confirmation. As a case study, the methodology was applied to the Beta Pictoris Moving Group, which resulted in proposals to change the current understanding of it.

Sumário

Lista de Figuras	xii
Lista de Tabelas	xiv
1 Introdução	1
2 Conceitos de Astronomia e os Grupos em Movimento	4
2.1 Movimento das estrelas	4
2.1.1 Os sistemas de referência e as posições das estrelas	4
2.1.2 A velocidade das estrelas em seu movimento	10
2.1.3 O potencial da Galáxia	13
2.2 Características Fotométricas e Espectroscópicas das Estrelas	13
2.2.1 Característica Fotométrica: magnitude	14
2.2.2 Característica espectroscópica: tipo espectral	15
2.3 Grupos Estelares em Movimento	16
2.3.1 Aglomerados, associações, grupos estelares em movimento	16
2.3.2 Breve histórico da detecção de grupos estelares em movimento	17
2.3.3 Estudos recentes de grupos em movimento	18
2.4 Abordagem dinâmica de Ortega <i>et al.</i>	19
3 Ciência de Dados	23
3.1 Ciência de Dados	23
3.1.1 Obtenção dos dados <u>O</u> SEMI	23
3.1.2 Limpeza dos dados <u>O</u> SEMI	24
3.1.3 Exploração dos dados <u>O</u> SEMI	24
3.1.4 Modelagem dos dados <u>O</u> SEMI	25
3.1.5 Interpretação dos dados <u>O</u> SEMI	25
3.2 Análise das Componentes Principais - PCA	25
3.2.1 Descrição da PCA	25
3.2.2 Passo a passo para calcular as componentes principais	26
3.2.3 Variâncias, <i>scores</i> e <i>loadings</i>	27
3.3 Regressão Linear	28

3.3.1	Passo a passo da Regressão Linear	28
3.3.2	Regressão Linear Simples e o método dos mínimos quadrados .	29
3.3.3	Coeficiente de Determinação	30
3.4	Conclusão Parcial	30
4	Metodologia	32
4.1	Obtenção dos dados (O <u>SE</u> MN)	32
4.1.1	Catálogos da Missão Hipparcos	33
4.1.2	Busca pelas velocidades radiais	34
4.1.3	Os grupos em movimento a serem estudados	36
4.2	Limpeza e tratamento dos dados (O <u>SE</u> MN)	37
4.2.1	Dados do catálogo Hipparcos	37
4.2.2	Mesclagem de Pulkovo com Sacy	38
4.2.3	Transformações de coordenadas	39
4.2.4	A base de dados completa	42
4.2.5	A construção dos dados dos grupos em movimento	42
4.3	Exploração dos dados (O <u>SE</u> MN)	43
4.4	Modelagem dos dados (O <u>SE</u> MN)	43
4.4.1	Experimentos com PCA	43
4.4.2	Regressão Linear Simples	45
4.5	Interpretação dos dados (O <u>SE</u> MN)	46
4.5.1	Seleção de estrelas	46
4.5.2	Evolução retrógrada com as estrelas selecionadas	47
4.5.3	Critérios para confirmação da estrela no grupo em movimento	47
4.6	Conclusões parciais	49
5	Resultados	52
5.1	Resultados dos experimentos com PCA	52
5.1.1	Experimentos com PCA	53
5.1.2	Discussão dos resultados com PCA	61
5.2	Resultados com Regressão Linear Simples	62
5.2.1	Ajustes com Regressão Linear Simples	62
5.2.2	Discussão dos resultados com regressão linear simples	67
5.2.3	Seleção de estrelas para a evolução retrógrada	67
5.3	Resultados da evolução retrógrada	68
5.4	Resultados dos testes com critérios	69
5.5	Modificações sugeridas ao grupo estelar em movimento Beta Pictoris segundo Ortega	75
6	Conclusão e Trabalhos Futuros	77

Lista de Figuras

2.1	Sistema de Referência Equatorial Celeste [1]	6
2.2	Sistemas de Referência Galáctico [2]	7
2.3	definição de paralaxe heliocêntrica (“p”) [2]	8
2.4	Definição de 1 parsec [3]	10
2.5	Movimento Próprio da Estrela Barnard [1]	11
2.6	Componentes da Velocidade de um corpo celeste visto da Terra (Adaptação)[4]	12
2.7	Componentes Estruturais da Galáxia. (Adaptação)[2]	13
2.8	Evolução retrógrada do grupo Beta Pictoris de Ortega <i>et al.</i>	21
4.1	<i>Workflow</i> da produção das bases de dados	37
4.2	Hipparcos: Erros médios dos cinco parâmetros astrométricos como função da magnitude [5]	38
4.3	O sistema Galáctico e o Sistema Equatorial Celeste [5]	41
4.4	Posições de todas as estrelas da base completa e dos grupos AB Doradus e Beta Pictoris	44
4.5	Componentes das velocidades de todas as estrelas da base completa e dos grupos AB Doradus e Beta Pictoris	45
4.6	Teste das Órbitas	48
4.7	<i>Workflow</i> da metodologia	51
5.1	Representação Gráfica PCA experimento 1	54
5.2	Representação Gráfica PCA experimento 2	55
5.3	Representação Gráfica PCA experimento 3	57
5.4	Representação Gráfica PCA experimento 4	58
5.5	Representação Gráfica PCA experimento 5	59
5.6	Representação Gráfica PCA experimento 6	60
5.7	Representação Gráfica PCA experimento 7	61
5.8	Ajuste da Reta para Beta Pictoris com u, v, w	63
5.9	Ajuste da Reta para Beta Pictoris com u, v, w, V_{mag}	63
5.10	Ajuste da Reta para Beta Pictoris com u, v, w, St	64

5.11	Ajuste da Reta para Beta Pictoris com u,v,w,V_{mag},St	64
5.12	Ajuste da Reta para AB Doradus com u,v,w	65
5.13	Ajuste da Reta para AB Doradus com u,v,w,V_{mag}	65
5.14	Ajuste da Reta para AB Doradus com u,v,w,St	66
5.15	Ajuste da Reta para AB Doradus com u,v,w,V_{mag},St	66
5.16	Seleção de Estrelas para Beta Pictoris	68
5.17	Local do nascimento de Beta Pictoris na época do nascimento	69
5.18	Órbitas do Grupo Beta Pictoris segundo Ortega	70
5.19	Órbitas do Grupo Beta Pictoris com 1 ^a modificação	71
5.20	Órbitas do Grupo Beta Pictoris modificado mais 6 estrelas sugeridas	72
5.21	Órbitas do Grupo Beta Pictoris modificado mais 5 estrelas sugeridas	72
5.22	Órbitas do Grupo Beta Pictoris modificado mais 4 estrelas sugeridas	73
5.23	Órbitas do Grupo Beta Pictoris modificado mais 3 estrelas sugeridas	74
5.24	Órbitas do Grupo Beta Pictoris com as 3 modificações confirmadas .	75

Lista de Tabelas

2.1	Estrelas do Grupo Beta Pictoris segundo Ortega <i>et.al.</i>	22
4.1	Estrelas do Grupo AB Doradus segundo Zuckerman e Bessel [6]	36
4.2	O tipo espectral expresso em números	39
4.3	A variáveis das estrelas na base de dados, suas representações e unidades	40
4.4	As características das estrelas na Base de Dados Completa	42
4.5	Estrelas da Base de Dados Beta Pictoris	43
4.6	Estrelas da Base de Dados AB Doradus	43
5.1	Características das estrelas usadas no experimento 1	53
5.2	Variâncias Explicadas para PCA experimento 1	54
5.3	Características das estrelas usadas no experimento 2	54
5.4	Variâncias Explicadas para PCA experimento 2	55
5.5	Características das estrelas usadas no experimento 3	56
5.6	Variâncias Explicadas para PCA experimento 3	56
5.7	Características das estrelas usadas no experimento 4	56
5.8	Variâncias Explicadas para PCA experimento 4	57
5.9	Características das estrelas usadas no experimento 5	58
5.10	Variâncias Explicadas para PCA experimento 5	58
5.11	Características das estrelas usadas no experimento 6	59
5.12	Variâncias Explicadas para PCA experimento 6	59
5.13	Características das estrelas usadas no experimento 7	60
5.14	Variâncias Explicadas para PCA experimento 7	60
5.15	Estrelas sugeridas para o Grupo Movente Beta Pictoris segundo Ortega	68
5.16	Confinamento máximo e idade dinâmica Beta Pictoris segundo Ortega	70
5.17	Confinamento e idade dinâmica Beta Pictoris com 1ª modificação . . .	71
5.18	Confinamento e idade dinâmica Beta Pictoris modificado mais 6 es- trelas sugeridas	71
5.19	Confinamento e idade dinâmica Beta Pictoris modificado mais 5 es- trelas sugeridas	72

5.20	Confinamento e idade dinâmica Beta Pictoris modificado mais 4 estrelas sugeridas	73
5.21	Confinamento e idade dinâmica Beta Pictoris modificado mais 3 estrelas sugeridas	74
5.22	Confinamento e idade dinâmica Beta Pictoris com as 3 modificações confirmadas	75

Capítulo 1

Introdução

Os seres vivos, e aqueles aparentemente inanimados vivem desde os tempos mais remotos de sua existência a mercê do dia e da noite, da escuridão e da claridade, que é a consequência do movimento da Terra em torno do seu eixo. Em uma escala de tempo um pouco maior, percebe-se algumas diferenças de acordo com o local em que se está no planeta Terra em relação a uma época que não se repete mais a cada dia, mas sim a cada ano: os dias e as noites possuem comprimentos diferentes, o trajeto que o Sol faz no céu, o local onde o Sol nasce e se põe e as estrelas que vemos no céu a noite também são diferentes. Mas, tudo se repete no ano seguinte. Tudo isso é consequência da inclinação do eixo da Terra e do movimento de translação da Terra em torno do Sol. Esses movimentos foram percebidos, estudados e utilizados pelas civilizações nos tempos mais remotos, com a tecnologia que lhes cabia: a percepção.

Outros movimentos ainda foram percebidos mesmo a olho nu, mas foi Galileo Galilei no século XVII que construiu os primeiros equipamentos para observar o céu, o tão incrível e fascinante céu que em todos os dias está disponível para nós e que traz em sua observação tanto aprendizado. A observação das imagens de hoje vistas de tantos locais, ângulos e datas diferentes, e em tantas profundidades poderia mesmo explicar a formação do Universo? Essa é uma das maiores, senão a maior questão que move a Astronomia de hoje.

Uma das inúmeras pesquisas da Astronomia moderna está voltada para a descoberta de planetas extrassolares, os chamados exoplanetas. Desde 1992 já foram descobertos, por diferentes métodos, mais de 4000 desses planetas [7] que orbitam outras estrelas que não são o nosso Sol. Um dos exemplos dessas estrelas é a Beta Pictoris, a segunda estrela mais brilhante da constelação Pictor, visível a olho nu no hemisfério sul. No ano de 2009 foi descoberto um planeta orbitando a estrela e posteriormente um disco protoplanetário em torno da estrela. Mas, muito antes disso já se sabia que essa estrela faz parte de um grupo em movimento, um grupo de estrelas que não estão próximas e nem estão ligadas gravitacionalmente, mas que

possuem movimentos semelhantes através da Galáxia e que estavam confinadas em uma região do espaço na época da origem do grupo. Os grupos em movimento são estudados há mais de 150 anos e mais recentemente, nos anos 90, com as missões espaciais astrométricas e de medidas de emissão de raios X, esse estudo tem sido muito ampliado. A detecção dos grupos em movimento jovens e na vizinhança do Sol está ligada à formação estelar e à formação de sistemas planetários.

Hoje são conhecidos alguns grupos estelares em movimento e alguns pesquisadores desenvolveram metodologias para previsão ou confirmação dos membros desses grupos. Em especial, Ortega *et al.* desenvolveram uma metodologia dinâmica em 3 dimensões, que confirma se as estrelas fazem parte do grupo, determina o local, o volume de confinamento das estrelas na origem do grupo, e a idade do grupo.[8]

A principal contribuição desse trabalho é a realização da análise de dezenas de milhares de estrelas concomitantemente com o objetivo de encontrar novos membros dos grupos em movimento já conhecidos. A metodologia desenvolvida é baseada na ciência de dados e é reprodutível a uma quantidade de dados de ordem de grandeza ainda maior, não somente no que se refere ao número de estrelas como também ao número de características das estrelas.

O cerne da metodologia consiste em representar os milhares de estrelas sendo consideradas através das componentes principais dos dados que as descrevem (velocidades, magnitudes, tipos espectrais, etc.). Da análise dessa representação surgem estrelas candidatas a inclusão em determinado grupo ou a exclusão dele, as quais são então submetidas a sucessivas simulações retrógradas para confirmação. Como estudo de caso, a metodologia foi aplicada ao grupo estelar em movimento Beta Pictoris, o que resultou em propostas de alteração do grupo.

O capítulo 2 é dedicado à descrição de todos os conceitos de Astronomia que estão envolvidos no assunto e principalmente dos grupos estelares em movimento, bem como de outros grupamentos estelares muito importantes para a compreensão da evolução estelar: os aglomerados abertos, aglomerados globulares e associações. É apresentado um breve histórico da detecção dos grupos em movimento e os trabalhos mais recentes envolvendo métodos e critérios para selecionar estrelas pertencentes aos grupos moventes. Também é feita a apresentação da metodologia dinâmica 3-D de Ortega e colaboradores baseada na evolução retrógrada a partir dos dados atuais completos de posição e velocidade, uma vez que o potencial Galáctico está inserido na implementação.

O capítulo 3 é dedicado à apresentação da ciência de dados e suas fases de obtenção, limpeza, exploração, modelagem e interpretação dos dados. São também descritas nesse capítulo a Análise das Componentes Principais e a Regressão Linear Simples, usadas na fase de modelagem.

No capítulo 4 é apresentada a metodologia, principal contribuição desse traba-

lho. É feita a descrição da montagem da base de dados, tendo em vista os catálogos Hipparcos disponíveis, que são produto da primeira missão astrométrica espacial, a Missão Hipparcos. Os catálogos Hipparcos fornecem os 5 parâmetros astrométricos necessários ao desenvolvimento desse trabalho, mas é necessário ainda um outro parâmetro para o estudo do movimento das estrelas: a velocidade radial, que é obtida por espectroscopia. Foi necessário um estudo cuidadoso para encontrar os catálogos mais apropriados para a montagem da base de dados completa, uma das contribuições desse trabalho. Na fase de modelagem da metodologia são realizadas experiências com Análise das Componentes Principais (PCA), variando as características das estrelas que são analisadas. É mostrado que no espaço das componentes principais as estrelas dos dois grupos estelares em movimento já conhecidos, AB Doradus e Beta Pictoris, apresentam configurações especiais. São selecionadas estrelas no entorno dessa configuração para o grupo Beta Pictoris e são, tais estrelas, submetidas à evolução retrógrada. Observa-se que há estrelas que estão muito próximas ao local de nascimento de Beta Pictoris, na época desse nascimento, além das estrelas conhecidas do grupo: são aquelas, então, as novas candidatas como membros do grupo. As estrelas do grupo e as novas candidatas são novamente submetidas à evolução retrógrada e a confirmação depende de alguns critérios, dentre os quais o critério das órbitas, desenvolvido nesse trabalho e que se constitui a terceira contribuição.

No capítulo 5 são apresentados os resultados onde se sugerem modificações ao grupo Beta Pictoris e no capítulo 6 a conclusão final e as expectativas de trabalhos futuros.

Capítulo 2

Conceitos de Astronomia e os Grupos em Movimento

Nesse capítulo são apresentadas as definições das características das estrelas que foram utilizadas nesse trabalho, suas unidades usuais e suas principais formas de medida. É apresentado também um breve histórico da detecção dos grupos estelares em movimento e sua definição, que está também relacionada com outros tipos de grupos de estrelas. Além disso, são citados os trabalhos recentes relacionados aos métodos e critérios para sugerir ou confirmar estrelas dos grupos moventes e, por fim, a apresentação da metodologia dinâmica 3-D de Ortega e colaboradores.

2.1 Movimento das estrelas

Observa-se que durante uma noite as estrelas mudam de posição em relação ao observador na Terra, mas esse movimento é resultante do movimento de rotação da Terra. As constelações que são vistas na Terra a noite vão mudando com o passar do ano, fenômeno que tem como origem o movimento de translação da Terra. Mas, as estrelas tem um movimento próprio? Sim, mas tal movimento das estrelas, por ser muito pequeno, não é perceptível a olho nu em uma noite e nem em um ano.

Para estudar o movimento é preciso sistema de referência, posição inicial, velocidade inicial. Se há um potencial, precisa-se da expressão do potencial. Essa seção apresenta as definições relacionadas a essas e outras características que foram usadas no presente trabalho.

2.1.1 Os sistemas de referência e as posições das estrelas

Para determinar a posição de uma estrela no céu é preciso definir um sistema de referência. Para isso é preciso escolher:

- um plano fundamental sobre o qual será medida uma coordenada angular

- um plano perpendicular ao primeiro sobre o qual será medida a outra coordenada angular
- a origem da contagem
- centro
- a forma de contagem das coordenadas angulares

Os sistemas celestes são como o sistema de coordenadas geográficas, cujo plano fundamental é o plano do equador sobre o qual se mede a longitude e o plano perpendicular é o meridiano local, sobre o qual se mede a latitude.

São mais utilizados em astronomia:

- sistema horizontal: sistema local, centrado no observador e cujo plano fundamental é o plano do horizonte do observador.
- sistema equatorial local: sistema local, cujo plano fundamental é a extensão celeste do equador terrestre
- sistema equatorial celeste: universal e geocêntrico, é o sistema de referência usado para expressar as posições das estrelas que são publicadas em catálogos e será descrito a seguir.
- sistema de coordenadas Galácticas: universal e tem como plano fundamental o plano mais próximo do disco da nossa Galáxia ¹ e será descrito a seguir.

O sistema de referência Equatorial Celeste

É o sistema de referência para expressar a posição dos astros como se vê de qualquer local da Terra. Utiliza como plano fundamental o equador celeste, extensão do equador terrestre. Assim como no sistema de coordenadas geográficas, os círculos máximos que passam pelos polos sul e norte são chamados de meridianos. As coordenadas equatoriais são a ascensão reta, representada pela letra grega (α) e a declinação representada por (δ). A ascensão reta é medida, em geral, em horas de 0h a 24h (mas pode igualmente ser contada em graus) sobre o equador celeste para leste com origem no ponto de Áries, também chamado de Ponto Vernal, que é o ponto do Equador Celeste ocupado pelo Sol quando ele passa do hemisfério sul celeste para o hemisfério norte celeste, definindo o equinócio de outono no hemisfério sul (e equinócio de primavera no hemisfério norte). A declinação é contada a partir

¹Por convenção, usa-se “G” (g maiúsculo) para referir-se à Via Láctea, que é a nossa Galáxia. Usa-se “G” quando se escreve, por exemplo, “a Galáxia” ou quando se escreve o adjetivo que se refere à Via Láctea, por exemplo “Galáctico”. Usa-se “galaxia” com “g” minúsculo para se referir às galáxias em geral

do equador celeste, sobre o meridiano da estrela, positiva para norte e negativa para o sul. A figura 2.1, obtida do livro de Maria de Fátima Saraiva e Kepler de Oliveira [1], mostra o sistema de referência Equatorial Celeste. O centro do sistema é a Terra e a eclíptica é o caminho que o Sol faz em um ano, sendo visto da Terra, inclinado de aproximadamente $23^{\circ}26'$ em relação ao Equador Celeste.

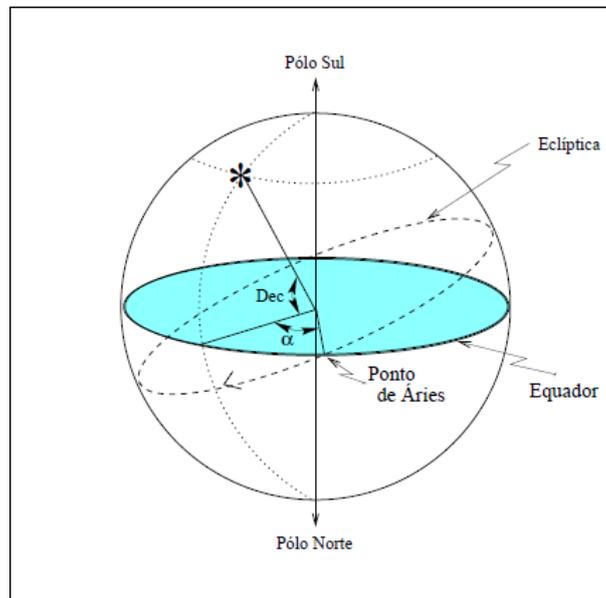


Figura 2.1: Sistema de Referência Equatorial Celeste [1]

O sistema de referência Galáctico

O sistema equatorial celeste, por ser geocêntrico, não é um bom sistema para observações, estudo e análises da estrutura e dinâmica da Galáxia. Para isso o sistema de coordenadas Galácticas é mais adequado e foi construído de forma a ter uma conexão fixa direta com a estrutura da Galáxia. O equador Galáctico é escolhido de forma a ser o círculo máximo que mais se aproxima do plano da Galáxia, contendo o centro da Galáxia e suas partes mais densas.[9] Esse plano é inclinado de cerca de $62,5^{\circ}$ em relação ao plano do equador celeste. Pode-se usar as coordenadas longitude (l) e latitude (b) ou as coordenadas cilíndricas R, ϕ e z . A figura 2.2, obtida do livro de Linda Sparker e Gallagher [2], mostra as duas representações do sistema de referência Galáctico: à esquerda são representadas as coordenadas longitude (l) e latitude (b) Galácticas com origem no Sol. À direita são representadas as coordenadas polares cilíndricas, R, ϕ e z com origem no centro Galáctico. A seta indica o sentido da rotação da Galáxia. As distâncias R, R_0 e z são medidas em parsecs, cuja definição é encontrada na próxima subseção.

As coordenadas longitude (l) e latitude (b) são as mais usadas. A latitude (b) é contada sobre o círculo máximo perpendicular ao equador Galáctico que passa pela

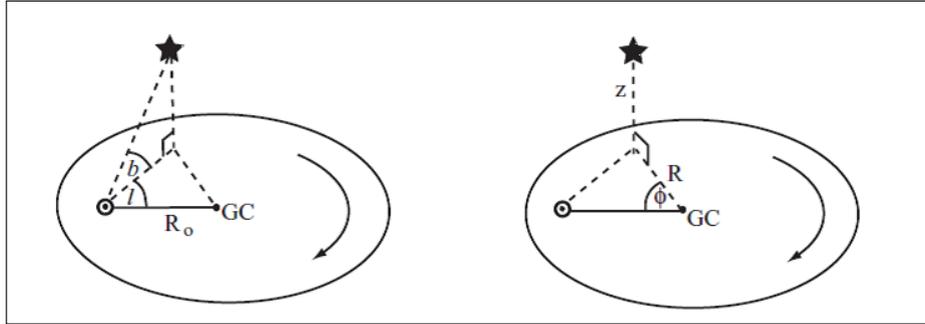


Figura 2.2: Sistemas de Referência Galáctico [2]

estrela a partir do equador Galáctico, como se fosse um meridiano Galáctico. A latitude é contada de 0 a 90°, positiva se a estrela está ao norte do plano Galáctico e negativa se está ao sul. A origem de contagem da longitude (l) está na interseção do plano Galáctico com a direção do centro da Galáxia. A longitude é contada sobre o equador Galáctico de 0 a 360° para leste. Como é difícil definir o centro da Galáxia, esse sistema passa frequentemente por revisões. Os valores que estão atualmente em vigor, são os seguintes, [3]:

Direção do centro da Galáxia em coordenadas Galácticas:

$$l = 0^\circ, b = 0^\circ \quad (2.1)$$

Direção do centro da Galáxia em coordenadas equatoriais celestes:

$$\alpha = 266,405^\circ, \delta = -28,936^\circ \quad (2.2)$$

Coordenadas equatoriais celestes do polo norte Galáctico:

$$\alpha = 192,85948^\circ, \delta = 27,12825^\circ \quad (2.3)$$

A distância das estrelas ao Sol, Paralaxe Heliocêntrica e o Parsec

As coordenadas celestes α e δ e as coordenadas Galácticas l e b , apresentadas nas seções anteriores, são coordenadas angulares e não incluem a distância.

A paralaxe é usada para medir a distância das estrelas e o parsec é a unidade dessa distância. A definição de parsec está ligada à definição de paralaxe, como veremos a seguir.

A ideia milenar para a medida de distância é baseada em simples geometria. O princípio fundamental está no fato de que há um deslocamento aparente na direção de um objeto devido à mudança de posição do observador. A esse deslocamento chamamos de paralaxe. [1]. Entendido isso, usa-se um triângulo retângulo conveniente

com componentes conhecidas, sendo a distância que se quer medir a componente desconhecida. Isso é usado, por exemplo, em topografia onde se usa o teodolito como instrumento para medir a paralaxe.

Para medir a distância das estrelas até nós, usamos a paralaxe heliocêntrica. Pela figura 2.3, obtida do livro de Linda Sparker e Gallagher [2], é possível observar que a mudança da posição da Terra em sua órbita em torno do Sol, provoca uma mudança angular na observação da estrela. Sendo assim, mede-se a direção da estrela em relação às estrelas de fundo quando a Terra está de um lado do Sol e faz-se a mesma medida seis meses depois, quando a Terra está do outro lado do Sol. Chama-se de paralaxe heliocêntrica a metade do desvio total na posição da estrela. Sendo conhecida a distância da Terra ao Sol, que é de aproximadamente 1 Unidade Astronômica “ua” em português e “au” em inglês) e obtendo-se medida a paralaxe heliocêntrica, calcula-se a distância da estrela ao Sol “d”.

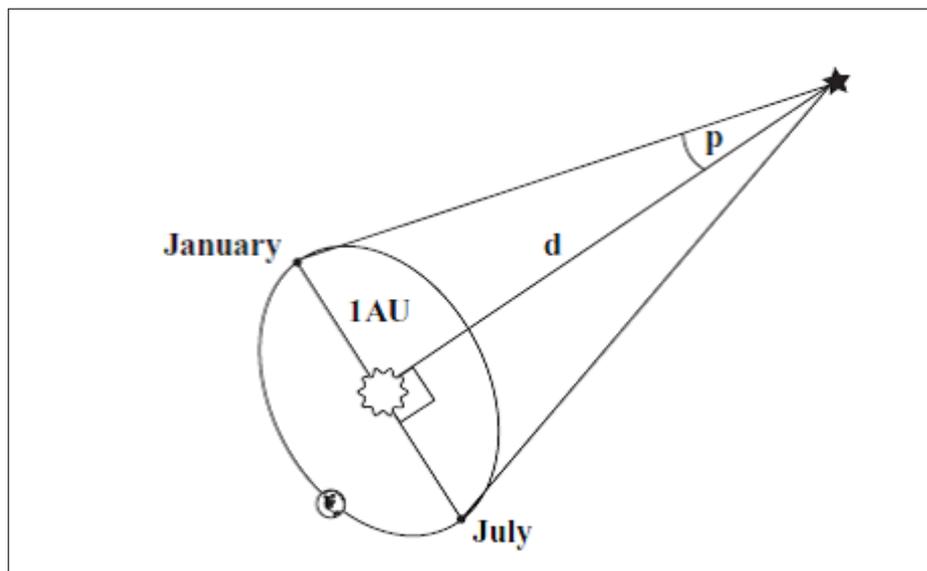


Figura 2.3: definição de paralaxe heliocêntrica (“p”) [2]

Sendo “r” o raio da órbita da Terra e “d” a distância da estrela ao Sol, então, pela figura 2.3 vê-se que

$$\frac{r}{d} = \tan(p) \quad (2.4)$$

Como a distância das estrelas ao Sol (“d”) é muito maior que o raio da órbita da Terra em torno do Sol (“r”), ou seja, como $d \gg r$, então “p” é um ângulo muito pequeno e

$$\frac{r}{d} \approx p \quad (2.5)$$

Isso porque, se tomarmos o ângulo “p” em radianos, a função tangente pode ser

representada por uma série de Taylor cujos termos são proporcionais às potências ímpares do ângulo. Para valores do ângulo (em radianos) menores do que 1 radianos, pode-se usar somente o primeiro termo, cuja potência é 1 e igual ao próprio ângulo “p”, porque esse primeiro termo é muito maior que a soma dos demais termos.[10]

Assim, fazendo $r = 1ua$ na equação 2.5, obtém-se para “p” em radianos

$$p \approx \frac{1ua}{d} \quad (2.6)$$

Essa paralaxe é chamada heliocêntrica porque está no Sol o ângulo reto do triângulo usado para a chamada triangulação. Em contraste, a paralaxe geocêntrica é aquela em que a base da triangulação é o raio da Terra, o que foi usado durante muito anos para medir a distância Terra-Lua e a distância dos planetas mais próximos.

Embora aqui denotada pela letra “p”, normalmente a letra grega π é usada para designar a paralaxe estelar.

E como medir as paralaxes? Francisco Catelli, Odilon Giovannini e Paula Hoffmann [10] publicaram um belo trabalho sobre as medidas de paralaxe, incluindo atividades práticas utilizando câmeras fotográficas e objetos terrestres. O teodolito como instrumento de medidas sobre a superfície da Terra, como já citado acima, é um exemplo.

Então, calcula-se a distância das estrelas em unidades astronômicas pela medida da paralaxe em radianos e a relação entre essas quantidades é:

$$d(ua) \approx \frac{1}{p(rad)} \quad (2.7)$$

Mas, a unidade astronômica não é uma boa unidade para distâncias estelares. Embora o ano-luz, que é a distância percorrida pela luz em um ano, seja uma boa unidade para expressar a distância das estrelas, o parsec, que será definido a seguir, é o mais utilizado.

O parsec (acrônimo de “par”alaxe de 1 “sec”, de “segundo” em inglês), abreviado por “pc”, é definido como a distância de uma estrela que tem a paralaxe de 1 segundo de arco, sendo a distância Terra-Sol de 1 unidade astronômica, ou seja, 1 parsec é a distância de um objeto tal que, um observador nesse objeto veria o raio da órbita da Terra com um tamanho angular de 1” (1 segundo de arco) [1], conforme representado na figura 2.4.

A definição da distância em parsec fica:

$$d(pc) = \frac{1}{p(")} \quad (2.8)$$

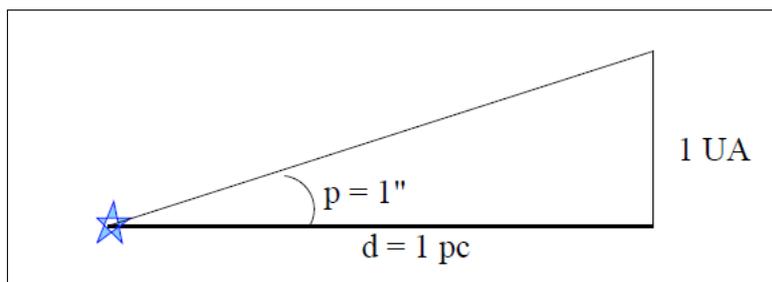


Figura 2.4: Definição de 1 parsec [3]

Como a expressão de 1 segundo de arco ($1''$) em relação ao radiano é:

$$1'' = \frac{\pi rad}{180 \times 3600} = 4,848 \times 10^{-6} \quad (2.9)$$

Assim, a equação 2.7 fica:

$$1pc = \frac{1ua}{4,848 \times 10^{-6}} = 206265ua \quad (2.10)$$

A paralaxe das estrelas é muito pequena e por isso difícil de medir com precisão. O Satélite Hipparcos realizou repetidas medidas de paralaxe para obter valores com precisão de milissegundo de arco, publicados nos catálogos Hipparcos, conforme será descrito no Capítulo 4. Usa-se a medida da paralaxe heliocêntrica para determinar medidas de distância até 1000pc (1 kpc). Para distâncias maiores que 1kpc e menores que 10kpc usa-se a paralaxe espectrocópica.[1]. A estrela mais próxima da Terra é a Próxima Centauri que está a uma distância de 1,32 pc e sua paralaxe é de $0,7687''$ [1].

2.1.2 A velocidade das estrelas em seu movimento

Em 1718 foi observado pela primeira vez que as estrelas tem um movimento intrínseco. Sir Edmund Halley observou para a estrela Arcturus uma mudança de posição de 1° em relação à posição obtida por Ptolomeu, no primeiro século da Era Cristã . Edmund Halley também observou uma mudança de $0,5^\circ$ na posição da estrela Sirius, que é a mais brilhante do céu. [1]. Uma mudança dessas representa muito. Para se ter uma ideia $0,5^\circ$, que é igual a $30'$ (30 minutos de arco), é a medida do diâmetro aparente do Sol ou da Lua, sendo diâmetro aparente o diâmetro do disco do Sol ou da Lua como se vê aqui da Terra.

Na figura 2.5 pode-se ver as medidas das posições da estrela de Barnard feitas pelo Satélite de Hipparcos, durante a sua missão. Observa-se que a variação na ascensão reta é de cerca de 0,1s (segundos de tempo) o que equivale a $1,5''$ em menos

de 2 anos. Em declinação a variação é de cerca de 15" (15 segundos de arco) em apenas 2 anos. Realmente, a estrela de Barnard tem um alto movimento próprio.

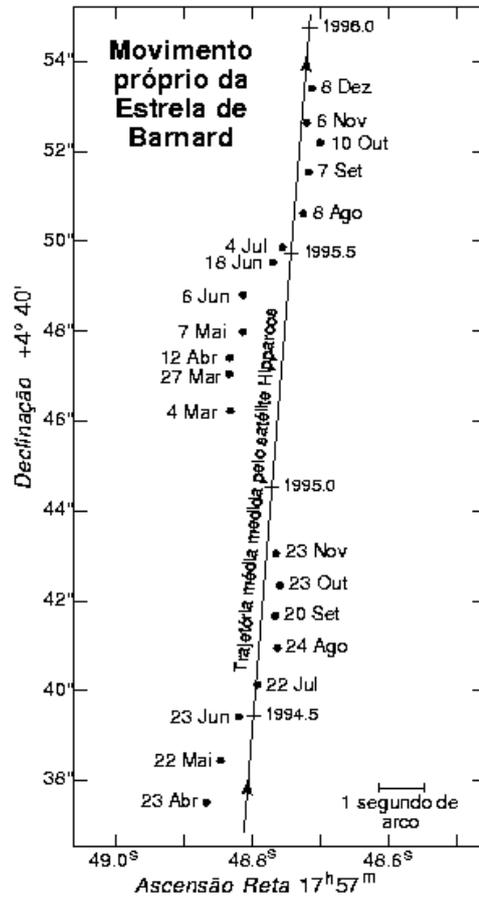


Figura 2.5: Movimento Próprio da Estrela Barnard [1]

O movimento próprio e a velocidade radial

Observou-se que as mudanças nas coordenadas ascensão reta (α) e declinação (δ) são devidas a movimentos na linha perpendicular à linha de visada e que movimentos paralelos à linha de visada vão mudar somente as distâncias das estrelas [4]. Isso está representado na figura 2.6: a velocidade na linha perpendicular à linha de visada é chamada de movimento próprio, o que está representados por v_p . A velocidade ao longo da linha de visada é chamada de velocidade radial, que está denotada v_r .

No sistema equatorial, como foi visto em 2.1.1, o movimento próprio é expresso para as coordenadas ascensão reta e declinação, obedecendo às mesmas regras de direção e sentido. Expressa-se normalmente:

- movimento próprio em ascensão reta: μ_α é expresso em segundos de tempo por ano (seg/ano ou sec/year em inglês)

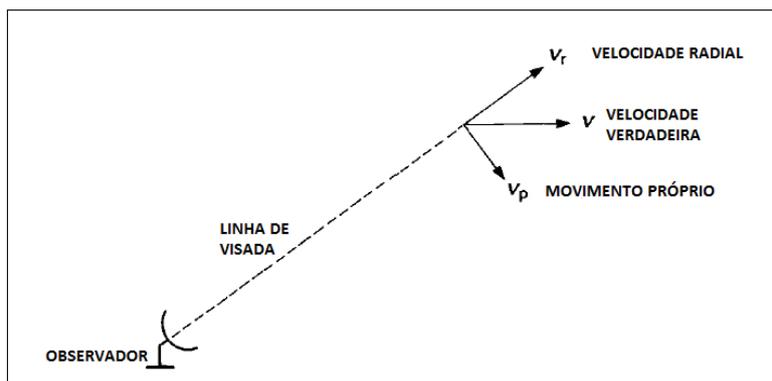


Figura 2.6: Componentes da Velocidade de um corpo celeste visto da Terra (Adaptação)[4]

- movimento próprio em declinação: μ_δ é expresso em segundos de arco por ano ("/ano ou "/year em inglês)

A velocidade radial é expressa em km/s e é medida pelo deslocamento Doppler. Em 1842 Christian Doppler (1803-1853) demonstrou que uma fonte que se distancia do observador tem todos os comprimentos de onda de seu espectro deslocados para o vermelho, isto é, o efeito Doppler desloca os comprimentos de onda para valores maiores (deslocamento para o vermelho ou *redshift*) quando a fonte se afasta e para valores menores, quando a fonte se aproxima (deslocamento para o azul ou *blueshift*. [1]

Assim:

- quando a fonte se afasta, há deslocamento para o vermelho (comprimentos de onda maiores) a diferença é positiva e a velocidade radial é positiva
- quando a fonte se aproxima, há deslocamento para o azul (comprimentos de onda menores) a diferença é negativa e a velocidade radial é negativa

As componentes das velocidades das estrelas no sistema Galáctico

Para estudar o movimento das estrelas do ponto de vista do plano da Galáxia é preciso fazer a transformação de coordenadas, conforme será explicitado no capítulo 4, do sistema equatorial celeste para o sistema de coordenadas Galácticas, pois os catálogos de estrelas em geral fornecem as coordenadas e velocidades referidas ao sistema equatorial celeste. Normalmente, no sistema Galáctico as velocidades são expressas nas componentes cartesianas, denominadas, "U", "V" e "W" (com letras maiúsculas). Essas componentes são expressas em km/s e se for usado o sistema da mão direita, U é positiva na direção do centro da Galáxia, V é positiva na direção

da rotação da Galáxia e W é positiva na direção do Polo Norte Galáctico. U, V e W são calculadas a partir dos valores do movimento próprio em ascensão reta (μ_α), do movimento próprio em declinação (μ_δ), da paralaxe (π) e da velocidade radial (V_r).

2.1.3 O potencial da Galáxia

O campo gravitacional geral da Galáxia é modelado como o campo total gerado pelas componentes estruturais da Galáxia que são: bojo, disco fino, disco espesso e halo, uma vez que as densidades de estrelas, gás e matéria escura diferem de acordo com a componente. A figura 2.7 mostra uma representação esquemática da Galáxia: o Sol está a cerca de 8500 pc do centro da Galáxia e as estrelas jovens e grupos em movimento, objeto desse trabalho, estão no disco fino.

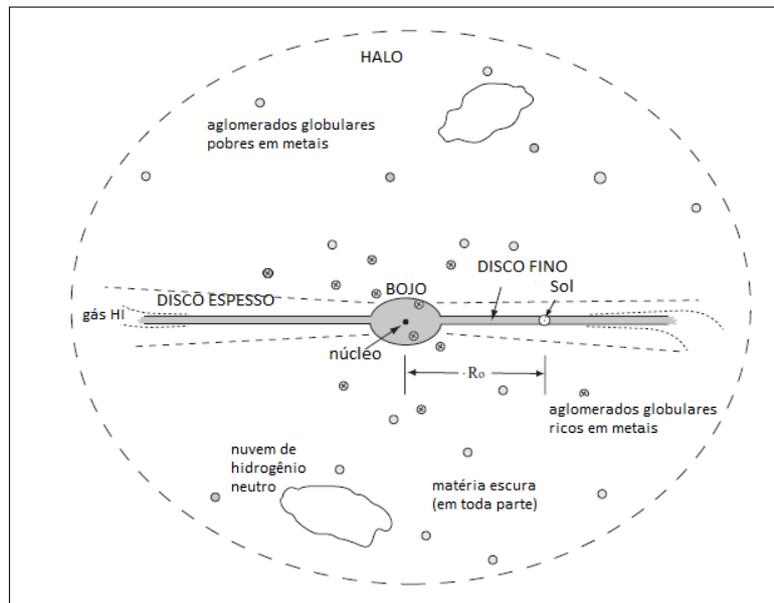


Figura 2.7: Componentes Estruturais da Galáxia. (Adaptação)[2]

O Potencial Galáctico de Hoogerwerf, de Brujine e de Zeeuw [11] é composto do potencial de Plummer para o bojo, do potencial de Miyamoto-Nagai para o disco e um potencial logarítmico para o halo. As expressões desses potenciais são detalhadas no livro Galactic Dynamics de James Binney e Scott Tremaine[3].

2.2 Características Fotométricas e Espectroscópicas das Estrelas

Em Astronomia, exceto a matéria escura, tudo é luz. Desde os tempos mais remotos, toda informação que nos chega dos astros chega através da luz que eles emitem ou

refletem. A fotometria é a medida da luz proveniente de um objeto e a espectroscopia é a análise do espectro de uma fonte, para obter informações a respeito de suas propriedades físicas. Nessa seção serão apresentados de resumidamente alguns conceitos relacionados a essas duas áreas, de forma a ser possível a compreensão de duas quantidades que serão usadas nesse trabalho: magnitude e o tipo espectral.

2.2.1 Característica Fotométrica: magnitude

Magnitude é a grandeza relacionada ao brilho da estrela. Foi Hiparco, considerado o maior astrônomo da era pré-cristã, nascido em de Nicéia (Grécia) no século II a.C., quem instituiu essa grandeza. Hiparco construiu um observatório, fez observações durante 20 anos e criou o primeiro catálogo de estrelas da humanidade. Ele criou também uma escala de magnitudes com 6 categorias e atribuiu o valor 1 à estrela mais brilhante e o valor 6 à estrela mais fraca visível a olho nu: as estrelas mais brilhantes seriam as vinte primeiras que aparecem depois do pôr do Sol. [1].

A medida que o tempo foi passando tanto estrelas mais brilhantes quanto estrelas mais fracas foram sendo observadas. Além disso, atribui-se também um valor de magnitude ao brilho de planetas e outros astros. Para manter a escala de Hiparco, foram introduzidos valores menores que 1, e até mesmo menores que zero, bem como valores muito maiores que 6.

Magnitude aparente

Chama-se magnitude aparente àquela relacionada ao brilho do astro como é visto aqui da Terra. A escala de magnitudes de Hiparco refere-se à magnitude aparente. Como foi explanado, essa escala baseia-se em uma comparação do brilho do astro, como é visto a olho nu aqui da Terra. Qual seria a relação da magnitude com a intensidade de energia recebida do astro? Quando os astrônomos perceberam que a escala de percepção do olho humano é logaritmica, mesmo assim, eles quiseram manter a escala de Hiparco, o que levou à seguinte equação, onde Δm_v é a diferença em magnitude como vista pelo olho humano e I_v é a intensidade de energia recebida com um instrumento que tem uma sensibilidade de cor semelhante ao olho humano[4]:

$$\Delta m_v = -2,5 \Delta \log I_v \quad (2.11)$$

Ou seja, a diferença do brilho entre 2 estrelas, que serão aqui chamadas de estrela “A” e estrela “B” fica:

$$m_v(A) - m_v(B) = -2,5[\log I_v(A) - \log I_v(B)] = -2,5 \log \frac{I_v(A)}{I_v(B)} \quad (2.12)$$

onde o sinal “menos” aparece porque ao maior brilho está relacionada um valor de magnitude menor.

Atualmente muitos detectores eletrônicos são usados para estudar a radiação electromagnética do espaço. Todo o espectro electromagnético, desde a radiação gama até as ondas de rádio são atualmente usadas para observações astronômicas. Chama-se magnitude visual, representada por m_v a magnitude aparente detectada nos comprimentos de onda da região visível do espectro (comprimentos de onda em torno de 5500 angstroms).[1]

Usa-se “m” (minúsculo) para denotar a magnitude aparente, em contraste ao uso de “M” (maiúsculo) para representar a magnitude absoluta.

Magnitude absoluta e o módulo de distância

A magnitude absoluta está relacionada ao brilho intrínseco do astro. Para podermos comparar os brilhos intrínsecos de duas estrelas usamos uma medida de brilho que independa da distância. Define-se como magnitude absoluta, a magnitude teórica que a estrela teria se estivesse a 10 parsecs da Terra:

$$M = -2,5 \log [I_v(10pc)] + constante \quad (2.13)$$

Usando-se a expressão para a intensidade da energia e propriedades de logaritmo, chega-se a

$$m - M = 5 \log d(pc) - 5 \quad (2.14)$$

onde “m” é a magnitude aparente, “M” a magnitude absoluta e “d” a distância do astro à Terra é expressa obrigatoriamente em parsecs.

Essa expressão dá origem à uma forma de cálculo de distância, chamada módulo de distância:

$$d(pc) = 10^{\frac{m-M+5}{5}} \quad (2.15)$$

onde, a distância “d” obtida é expressa obrigatoriamente em parsecs.

2.2.2 Característica espectrocópica: tipo espectral

O tipo espectral está relacionado com o tipo do espectro da estrela. Os tipos espectrais também estão relacionados às temperaturas efetivas das estrelas e são represen-

tados por uma letra e por uma sequência de números que são as classes espectrais, que variam de 0 a 9, sendo 0 a mais quente e 9 a mais fria. As letras "O", "B", "A", "F", "G", "K" e "M" são usadas para designar o tipo espectral sendo "O" a que equivale a maior temperatura e "M" à menor. Há ainda outros tipos "R", "S", "N", "C", "D" e "W" que estão relacionados a outras características e para os quais não há muitas mudanças de temperatura.

A designação completa dos tipos espectrais (letra+número) não são "fixos" e podem diferir um pouco de acordo com o autor e com as medidas mais recentemente realizadas.

2.3 Grupos Estelares em Movimento

2.3.1 Aglomerados, associações, grupos estelares em movimento

Os conceitos de aglomerados, associações e grupos estelares em movimento diferem um pouco de acordo com o autor e é frequente na literatura que logo no início o autor esclareça a definição que está usando ou a sua definição.

Sharpless [12] define aglomerado globular (*cluster*) como aglomerado com densidade de no mínimo uma ordem de grandeza maior do que a do campo e cujos membros possuem atração gravitacional mútua. Define ainda associação como um conjunto de estrelas com densidade consideravelmente menor que a do campo. Zuckerman and Song [13] e López-Santiago *et al.* [14], em seus estudos de estrelas jovens em movimento e próximas ao Sol, atribuem a grupos cinemáticos tanto associações, como aglomerados, quanto grupos em movimento uma vez que todos referem-se se a grupos de estrelas com movimento espacial semelhante.

Torres [15] prefere não usar o termo "grupo estelar em movimento" (*moving group*), mas somente "associação" e define associação como um grupo de estrelas aparentemente concentradas num pequeno volume do espaço com algumas características comuns como idade, composição química, distância e cinemática.

Ortega *et al.* [16] define "grupo estelar em movimento" como um conjunto de estrelas com cinemática atualmente semelhante e confinadas em uma região tridimensional na época da formação do grupo. Essa definição será adotada nesse trabalho. Também será considerado aqui que as estrelas de uma associação tem cinemática semelhante, mas a diferença está na densidade que para a associação é menor ou igual a 0,1 massas solares num volume de 1 parsec cúbico.

Os membros das associações e dos grupos estelares em movimento não possuem interação gravitacional entre si, mas encontram-se sob a ação do potencial da nossa Galáxia. De forma diferente, as estrelas dos aglomerados (*clusters*) sejam abertos

(*open clusters*) ou globulares possuem interação gravitacional entre si. Os aglomerados abertos são formados de dezenas a centenas de estrelas, mas os globulares possuem centenas de milhares de estrelas.

A ideia evolutiva que liga esses conceitos é que as estrelas nascem de nuvens moleculares e que vão formar posteriormente os aglomerados ou grupos de diferentes tipos. Com o decorrer do tempo, os aglomerados abertos e os grupos se dissociam e dão origem às associações e grupos em movimento.

O estudo dos grupos estelares, sejam aglomerados globulares ou abertos, associações ou grupos estelares em movimento é extremamente vasto e está intimamente ligado com a formação de estrelas e com a formação do universo. Em particular, o estudo dos grupos estelares em movimento apresenta uma grande importância relacionada com a determinação da idade das estrelas e no caso de grupos estelares em movimento jovens, com a origem e evolução recente de sistemas planetários.

2.3.2 Breve histórico da detecção de grupos estelares em movimento

A ideia dos grupos estelares em movimento já tem 150 anos. Proctor em seu trabalho datado de 1869 relata ter observado que em algumas partes do céu as estrelas tem uma tendência de fazer movimentos espaciais semelhantes. Ele acrescenta que essa tendência não é facilmente observável nos catálogos de movimentos próprios, pela forma como as estrelas são organizadas, entretanto, essa tendência ficaria evidente se as estrelas fossem mapeadas e seus movimentos próprios fossem representados com setas de comprimento e direção proporcionais, respectivamente, ao valor e direção daquela grandeza.[17].

Em sua série de artigos publicados ao longo de 3 décadas (1965 a 1994) Olin Eggen apresenta seu vasto estudo sobre grupos estelares em movimento. Em sua série de nove volumes denominada “Stars and Stellar Systems” de 1965, Eggen faz um histórico do estudo dos grupos localizados a poucas centenas de parsec do Sol. De acordo com Eggen, R.A.Proctor em sua publicação de 1869, notou não somente estrelas na vizinhança do aglomerado de estrelas Hyades movendo-se junto com a Galaxia, mas também cinco estrelas movendo-se junto com a o grupo Ursa Maior. Por cerca de um século, somente esses dois grupos eram conhecidos próximos da Terra; o primeiro foi associado com Hyades (45 pc da Terra e 600Myrs de idade)² e o segundo com a Ursa Maior (25 pc e 300Myr) [13]. Eggen criou um critério, conhecido na literatura como Critério de Eggen, que afirma que todas as estrelas que pertencem ao mesmo grupo estelar em movimento tem praticamente a mesma componente da velocidade V (componente da velocidade sobre o plano Galáctico na

²Myr significa milhões de anos

direção da rotação da Galáxia, se for positiva). Eggen também introduziu a ideia de Associação Local de estrelas, que se referia ao conjunto de vários grupos. [17].

Em seu artigo de revisão, Sharpless[12] também se referiu a Proctor como introdutor da ideia de feixe de estrelas em movimento na Galáxia. Sharpless discutiu os conceitos de associações e aglomerados estelares seguindo o trabalho de V.A. Ambarzumian na metade do século XX.[13].

Um avanço considerável sobre o reconhecimento de grupos em movimento ocorreu nos anos 90, quando se reconheceu que estrelas jovens são fortes emissores de raios X. Satélites de Raios X observaram aglomerados de estrelas jovens e principalmente o Satélite ROSAT, lançado em 1990, que mostrou que a emissão de raios X intensos persiste por um tempo da ordem de 100 Myr.

A missão Hipparcos da ESA (Agência Espacial Européia) foi a primeira missão astrométrica e um de seus objetivos foi a realização de medidas de alta qualidade de posição, de movimento próprio e de paralaxe de cerca de 120 mil estrelas. Os objetivos foram alcançados e os resultados dessa missão foram publicados nos catálogos de Hipparcos em 1997[5], Hipparcos 2 em 2002,[18] e Hipparcos Estendido (XHIP) em 2012 [19]. A publicação desses resultados tornou possível uma série de implementações e cálculos e, em vista disso membros de grupos em movimento e associações foram confirmados, ou sugeridos ou não confirmados.

Em 2013, a mesma ESA lançou o satélite Gaia com o objetivo de realizar medidas de posição de alta qualidade para cerca de 1,8 bilhões de estrelas e outros objetos celestes. Em 25 de abril de 2018 foi publicado o *Data Release 2* chamado de “Gaia DR2” com as posições, paralaxes, movimentos próprios e propriedades fotométricas. Nessa versão há um sub-catálogo com os valores de velocidades radiais obtidos a partir de medidas feitas pelo satélite, para cerca de 7 milhões de estrelas. Está previsto para ser publicado no segundo semestre de 2020 o chamado “Gaia EDR3 catalogue” com a revisão dos dados de posição, paralaxe e movimentos próprios [20], que vai alavancar nova série de trabalhos sobre o assunto. Já com os dados do “Gaia DR2” novos grupos foram identificados [21].

2.3.3 Estudos recentes de grupos em movimento

Recentemente passou-se a denominar por NYMGs (*Nearby Young Moving Groups*) os grupos em movimento jovens na vizinhança do Sol. Por jovens entendam-se aqueles com idade menor que 100 Myr e nas proximidades do Sol, aqueles localizados a menos de 100 pc do Sol.

Nos últimos 20 anos alguns NYMGs foram descobertos (da ordem de 10) e mais de dois mil estrelas desses grupos foram confirmadas pelos pesquisadores [22]. Cada grupo de autores utiliza diferentes métodos e critérios para confirmar os membros

de um grupo, sugerir outros e deixar de considerar outros. Assim, é possível que um autor considere uma estrela como membro de um grupo e outro autor considere, por outros critérios, que a mesma estrela não faz parte daquele mesmo grupo.

Alguns autores dedicaram-se ao estudo dos grupos mais próximos e mais jovens, como por exemplo Zuckerman e Song [23] e [13] que estudaram quatro grupos conhecidos com distâncias de até 60 pc do Sol e com idades entre 8 e 50Myr: Associação TW Hydrae, grupo em movimento Beta Pictoris, grupo em movimento AB Doradus e Associação Tucana/Horologium. Em seu trabalho mais recente, Zuckerman [21] confirma a existência da Associação Argus, com 40-50 Myr e contendo muitas estrelas a menos de 100 pc do Sol.

Outros trabalhos podem ser citados como Torres [15] e os mais recentes Móor [24], Elliot [25], Riedel [26], Evgenya Shkolnik [27] e Lee & Song [22], de agosto de 2019 que usou dois algoritmos de aprendizado de máquina não supervisionados (*K-means* e clusterização aglomerativa) para reconhecer grupos e seus membros. Nesses trabalhos são usados métodos e critérios relacionados às características das estrelas como idade e as propriedades fotométricas e espectroscópicas, além das posições e velocidades atuais das estrelas, ou seja, são utilizados critérios cinemáticos e não dinâmicos.

Na seção 2.4 a abordagem dinâmica de Ortega *et al.* será apresentada e serão citados os trabalhos de Ortega *et al.* usando os critérios dinâmicos.No presente trabalho foi utilizado o critério dinâmico.

2.4 Abordagem dinâmica de Ortega *et al.*

Ortega e colaboradores publicaram pela primeira vez em 2002 em seu “The Origin of the Beta Pictoris Moving Group” [8] a sua abordagem dinâmica para obter a idade de um grupo em movimento, especificamente naquele artigo, para o Beta Pictoris Moving Group (BPMG). Essa abordagem foi pioneira por trabalhar no espaço tridimensional de posições e demonstrou-se muito poderosa por determinar a idade dinâmica do grupo e simultaneamente o local de nascimento do grupo, partindo somente das posições e das velocidades atuais das estrelas e levando em conta o potencial da Galáxia. Esse método foi um progresso em relação aos anteriores apresentados que eram bidimensionais, no plano galáctico (X,Y). Um dos exemplos pode ser visto em Mamajek e Feigelson [28] que sem considerar o potencial galáctico, tentava encontrar as origens de algumas estrelas do BPMG assumindo a priori uma certa idade.

A metodologia de Ortega e colaboradores vem sendo utilizada com sucesso e determinou a origem de outros grupos em movimento como Eta Chamaeleontis e Epsilon Chamaeleontis [29], AB Doradus [30] e [31], além de embasar o estudo da

formação dos grupos estelares jovens na região da Associação OB Scorpio Centaurus [32]. Em seu artigo de revisão de 2004, Zuckerman e Song [13] citam a metodologia de Ortega *et al.* como calibradora das idades dos grupos em movimento.

O método consiste em calcular as posições das estrelas no espaço tridimensional no sistema de referência Galáctico (sendo XY o plano da Galáxia) em evolução retrógrada. Ou seja, os dados de entrada são as posições e as velocidades atuais das estrelas que são prováveis membros de um grupo estelar em movimento. Tendo em vista o potencial Galáctico são montadas as equações do movimento e por integração numérica são conhecidas as posições e velocidades de um tempo anterior. O método numérico de Runge-Kutta de 4ª ordem é empregado para integrar as equações de movimento e calcular as posições. Nessa implementação o tamanho do passo não é fixo como no trabalho de Hoogerwerf mas varia automaticamente com a precisão desejada. É utilizada a rotina “odeint” conforme descrito em Numerical Recipes in C [33].

Assim, partindo-se das posições e velocidades atuais, faz a evolução retrógrada sendo o passo determinado como se deseja, em unidades de milhões de anos (1 Myr) ou suas frações. A cada passo é calculado a posição média do grupo e o raio do grupo.

Se as estrelas realmente forem componentes de um grupo estelar em movimento, ficarão cada vez mais próximas entre si nessa evolução retrógrada. Considera-se que quando a distância média entre as estrelas for a menor possível, atingiu-se o primeiro mínimo, que é o local do nascimento do grupo. Assim determina-se a idade dinâmica do grupo e, simultaneamente o local do nascimento do grupo e o tamanho da região de confinamento. A figura 2.8 mostra a evolução retrógrada tridimensional com os 21 membros do grupo estelar em movimento Beta Pictoris, segundo Ortega *et al.*. As posições e velocidades foram calculadas a partir de compilação dos dados do Catálogo Hipparcos, 1997 com velocidades radiais obtidas de fontes diversas. Na tabela 2.1 são apresentadas as estrelas que compõem o Grupo Beta Pictoris segundo Ortega *et al.*: a grande maioria são estrelas do Catálogo Hipparcos e somente uma foi obtida do catálogo BD.³ A identificação das estrelas no Catálogo Hipparcos é feita por números compostos de 6 algarismos, precedidos por “Hip” ou “HIP”. Para sistemas compostos de mais de uma estrela é posposta uma letra para designar a componente do sistema.

Quando uma ou mais estrelas são adicionadas a um grupo já analisado, caso não seja um novo membro do grupo, observa-se que o comportamento dinâmico do grupo apresenta mudanças como, por exemplo, a relação do confinamento inicial com o maior confinamento fica muito diferente do valor já conhecido ou a idade do grupo fica muito diferente daquela determinada por outros métodos.

³Catálogo astrométrico compilado na Alemanha de 1859 a 1903

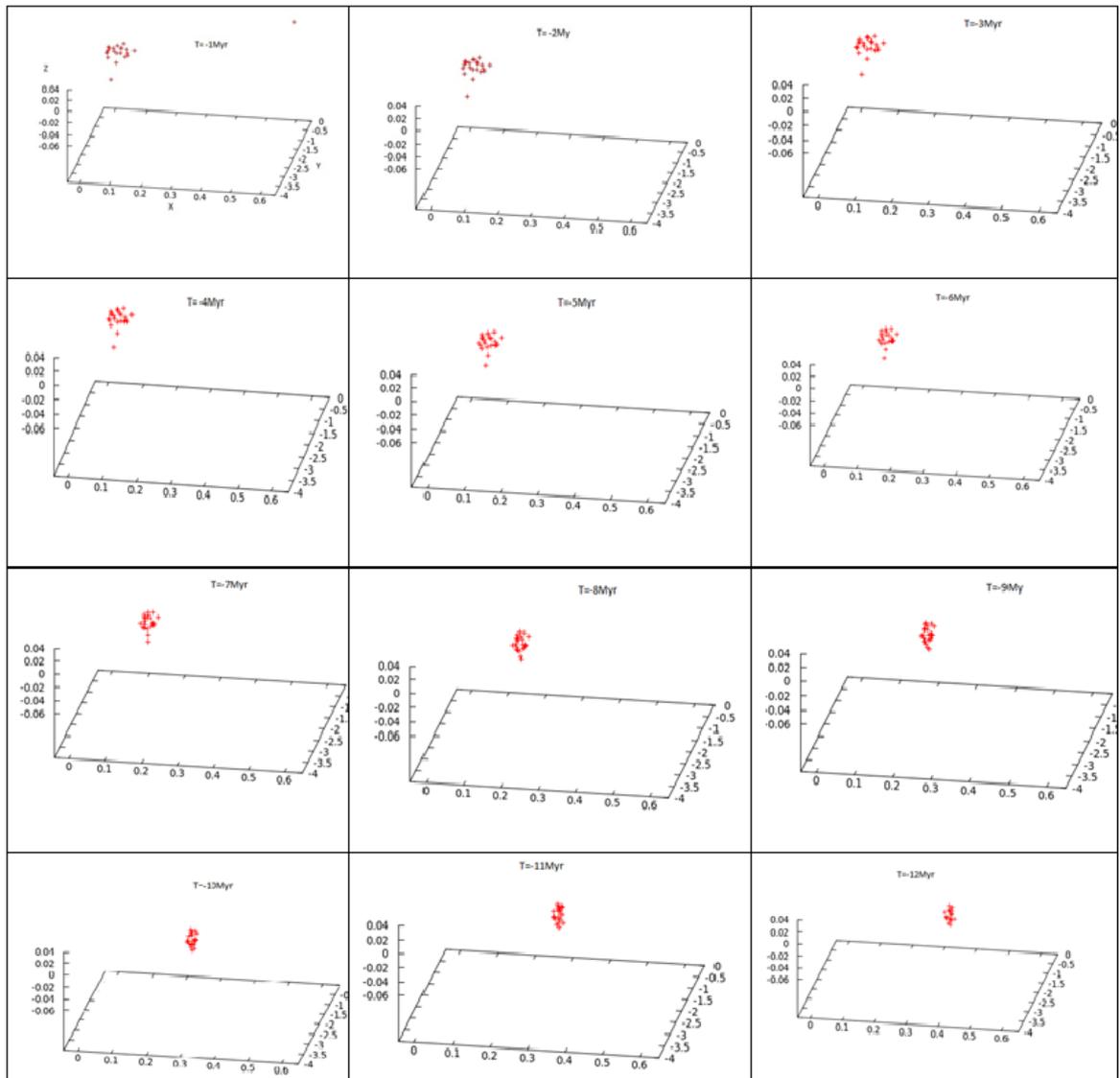


Figura 2.8: Evolução retrógrada do grupo Beta Pictoris de Ortega *et al.*

Tabela 2.1: Estrelas do Grupo Beta Pictoris segundo Ortega *et.al.*

Hip 560	Hip 103311	Hip 14361	Hip 105441
Hip 21547	Hip 10680	Hip 23200 B	Hip 102409
Hip 27321	Hip 11437 A	Hip 23309	Hip 88399
Hip 92024	Hip 84586	Hip 25486	Hip 102141
Hip 92680	Hip 12545	Hip 99273	Hip 76629
BD $-17^{\circ}6128$			

Ortega e col. [8] definem os grupos estelares em movimento como um conjunto de estrelas com cinemática atualmente similar e que estavam confinadas em uma região tridimensional na época da formação. A evolução retrógrada é realizada a partir das posições atuais e dos dados cinemáticos atuais completos sob a ação do campo gravitacional geral da nossa Galáxia. Não são consideradas as interações gravitacionais entre as estrelas pois, como foi visto na seção 2.3.1 os membros dos grupos em movimentos e das associações não possuem interação entre si.

Capítulo 3

Ciência de Dados

3.1 Ciência de Dados

A ciência de dados, em inglês *data science*, é uma ciência relativamente nova e há várias definições do que ela engloba. Por exemplo, Jeroen Janssens [34] em seu muito citado livro *Data Science at the Command Line* adotou a definição de Hilary Mason e Chris Wiggins [35], que define a Ciência de Dados de acordo com os seguintes cinco passos: obtenção, limpeza, exploração, modelagem e interpretação dos dados. Esses passos formam o modelo OSEM_N, cujo nome é um acrônimo criado com as letras dos nomes dos passos, em inglês, respectivamente: **O**btaining, **S**crubbing, **E**xploring, **M**odeling e **iN**terpreting. Para Jeffrey Stanton,[36] os passos são: coleta, preparação, análise, visualização, gerenciamento e preservação.

Qualquer que seja a definição dos passos, é certo que a ciência de dados é uma área interdisciplinar dedicada ao estudo e análise de dados, com o objetivo de extrair informação ou conhecimento.

Para esse trabalho adotou-se o modelo OSEM_N, cujos passos serão apresentados nas seguintes subseções. Observa-se que os cinco passos são executados em geral na ordem OSEM_N, mas na prática o fluxo pode ser diferente, por exemplo, executando múltiplos passos ao mesmo tempo ou retornando a um passo anterior, se for preciso.

3.1.1 Obtenção dos dados OSEMI

Os dados podem ser obtidos [34]

- de algum local pela internet ou de um servidor
- de uma base de dados ou aplicativo
- de uma outra base de dados ou arquivo
- por geração própria

A dificuldades nesse passo normalmente estão ligadas:

- ao formato do arquivo obtido que em muitas vezes precisa ser transformado e é preciso cuidado com a perda de dados nessa transformação.
- à descrição dos campos que pode não estar correta ou completamente documentada e que pode originar erros de atribuição.
- à necessidade de mesclar várias bases de dados para conseguir todos os dados que se deseja, que pode originar erro de perda, repetição ou atribuição incorreta dos dados

3.1.2 Limpeza dos dados OSEMI

Os dados obtidos podem conter colunas que não sejam de interesse, inconsistências, erros, caracteres estranhos. Então, é necessário o passo de limpeza dos dados com operações como: [34]

- filtragem de linhas: análise das linhas dos dados para verificar se pode ser passada como saída
- extração de colunas
- reordenação de colunas e/ou linhas
- troca de valores, por exemplo mudar uma string por um valor numérico
- extração de campos
- tratamento de valores ausentes
- conversão de dados de um formato para outro
- junção de colunas

3.1.3 Exploração dos dados OSEMI

O objetivo desse passo é explorar os dados usando ferramentas de visualização. As operações características desse passo são: [34]

- criação de estatísticas
- utilização de ferramentas de visualização

3.1.4 Modelagem dos dados OSEMI

Para explicar os dados ou obter informações, fazer previsões, etc. é preciso criar um modelo estatístico dos dados. As técnicas para criar um modelo incluem:[34]

- classificação
- regressão
- redução de dimensão

Nas seções 3.2 e 3.3 serão descritas, respectivamente, as técnicas de redução de dimensão e de regressão que foram usadas nesse trabalho.

3.1.5 Interpretação dos dados OSEMI

Finalmente, chega-se à interpretação. Os passos envolvem:[34]

- desenho da solução
- avaliação do significado dos resultados
- comunicação dos resultados

3.2 Análise das Componentes Principais - PCA

Análise das Componentes Principais ou *Principal Component Analysis* (PCA) é uma das ferramentas mais importantes da análise multivariada[37] e também uma das mais simples. Menos frequentemente é chamada de Transformada Discreta de Karhunen-Loève (KLT) ou ainda Transformada de Hotelling, em homenagem a Kari Karhunen, Michel Loève e Harold Hotelling.[38]

3.2.1 Descrição da PCA

Trata-se de um procedimento matemático que baseia-se numa transformação ortogonal com o objetivo de transformar um conjunto de variáveis discretas A - possivelmente correlacionadas entre si ou parcialmente correlacionadas entre si - em um outro conjunto de coeficientes discretos B que são linearmente decorrelacionadas entre si e são chamadas de componentes principais.

Esse método é utilizado para reduzir a dimensão do problema em componentes não correlacionadas, eliminar sobreposições e para escolha de formas mais representativas dos dados. [38]. A PCA é uma maneira de identificar a relação entre as

características extraídas dos dados. Essas componentes não correlacionadas (as componentes principais) são ortogonais entre si e são combinações lineares das variáveis originais. [39]

O número de componentes principais é igual ao número de variáveis originais, entretanto, as primeiras componentes concentram a maior parte da informação e normalmente são suficientes para representar todo o conjunto original, reduzindo, portanto, a dimensão do problema. A medida percentual do quanto uma componente principal representa as variáveis originais do problema é chamada de variância explicada.

Como as componentes principais são ortogonais entre si, elas formam uma base de um espaço vetorial, na qual a projeção das variáveis originais não contém nenhuma informação redundante.

Denomina-se por características, atributos ou *features* as variáveis do conjunto original.

O problema que o método da PCA resolve é o seguinte [39]: dado um conjunto de “n” características

$$\mathbf{X} = [X_1, \dots, X_n] \quad (3.1)$$

deve-se encontrar uma matriz \mathbf{T} de componentes principais, que faça com que o novo conjunto

$$\mathbf{Y} = [Y_1, \dots, Y_n] \quad (3.2)$$

seja não correlacionado, e

$$\mathbf{Y} = \mathbf{TX} \quad (3.3)$$

3.2.2 Passo a passo para calcular as componentes principais

O passo a passo para calcular as componentes principais é o seguinte:

1. obter os dados, ou seja, as \mathbf{M} amostras de vetores de dimensão \mathbf{N} (\mathbf{N} é o número de características)
2. calcular o vetor médio dos dados ou seja, calcular a média de cada uma das \mathbf{N} características pelas \mathbf{M} amostras
3. subtrair a média de todas as \mathbf{M} amostras
4. normalizar o valores, dividindo cada valor pelo desvio-padrão. Esse passo é muito importante quando as dimensões das características são diferentes entre si.

5. calcular a matriz de covariância usando todas as subtrações, fazendo a média do produto de cada subtração por ela mesma. Dimensão é $N \times N$
6. calcular os autovalores e os autovetores da matriz de covariância. A matriz de covariância é real e simétrica.
7. organizar a matriz da Transformada de Hotelling fazendo com que as linhas sejam formadas pelos autovetores da matriz de covariância em ordem decrescente dos seus autovalores associados

O autovetor com o maior autovalor corresponde à primeira componente principal do conjunto de dados usado. E assim, sucessivamente para a segunda componente principal e para a terceira, etc.[38] de forma que cada componente principal subsequente modele sempre a direção de maior variância no espaço multidimensional não modelado pelas componentes principais anteriores e sejam sempre ortogonais a todas elas. [37]

Nesse ponto faz-se a truncagem de acordo com o que se deseja, mantendo-se somente as primeiras componentes principais que já respondem por uma parcela significativa da informação total contida em toda a estrutura de dados. [37]

3.2.3 Variâncias, *scores* e *loadings*

Ao final dos 7 passos da seção anterior obteve-se a matriz da Transformada de Hotelling, sendo as linhas os autovetores em ordem decrescente de seus autovalores associados. O próximo passo após serem escolhidas as componentes principais (normalmente as duas ou três primeiras) é escrever a estrutura de dados inicial em função das componentes principais. Os *scores* fornecem a composição das componentes principais em relação às amostras e os *loadings* fornecem essa mesma composição em relação às variáveis. [38] Ou seja, os *scores* são as novas coordenadas das amostras, expressas nas “direções” das componentes principais. É no gráfico dos *scores* que se busca alguma relação entre as amostras, lembrando que a análise das componentes principais é uma ferramenta de reconhecimento de padrões.[37]

Uma interpretação geométrica para *score* e *loading* pode ser obtida fazendo-se um gráfico em duas dimensões para uma amostra, sendo os eixos coordenados duas variáveis. A direção de maior variabilidade será mostrada por uma reta, que será uma das componentes principais. Os *scores* são as projeções das amostras na direção das componentes principais e os *loadings* serão os ângulos entre cada componente principal e cada variável. [38]

As \mathbf{M} amostras iniciais que eram antes expressas através de \mathbf{N} características, podem agora ser expressas através de \mathbf{N} componentes principais. Como já foi visto acima, normalmente não se usa todas as \mathbf{N} componentes principais, mas somente

as primeiras desde que sejam suficientemente significativas para descrever toda a estrutura de dados. As variâncias explicadas e cumulativas expressam esses valores de representatividade das componentes principais, de forma percentual. A variância explicada expressa percentualmente quanto da variância está sendo explicada para cada componente principal, o que leva à soma de todas elas igual ter o valor de 100%. A variância cumulativa de cada componente principal é a soma das variâncias explicadas das componentes anteriores com a sua própria. Sendo assim, a primeira componente principal tem a variância cumulativa igual à sua variância explicada e a N ésima componente principal tem variância cumulativa igual a 100%.

3.3 Regressão Linear

A análise da regressão é uma análise estatística que visa verificar a existência de uma relação funcional entre uma variável dependente com uma ou mais variáveis independentes. O objetivo é obter uma equação que tenta expressar a variação da variável dependente com uma ou mais variáveis independentes.[40]

3.3.1 Passo a passo da Regressão Linear

1. fazer um gráfico, chamado de diagrama de dispersão, para visualizar como se comportam os valores da variável dependente, que será aqui chamada de “Y”, em relação à variável independente, que será aqui chamada de “X”.
2. observar o comportamento de Y em relação a X. Por exemplo, pode ser um comportamento linear, quadrático, cúbico, exponencial, logarítmico, ou outros.
3. escolher uma curva-modelo para representar o comportamento observado. O modelo que se vai escolher depende da resposta à seguinte pergunta: qual é a curva que melhor se adequa ao comportamento observado? Para escolher a curva deve-se levar em consideração as seguintes observações:
 - a curva-modelo escolhida deve estar de acordo com o comportamento observado e deve-se observar principalmente o grau do comportamento, escolhendo-se o mesmo grau para a curva-modelo
 - a curva-modelo deve conter apenas as variáveis que são relevantes para explicar o comportamento
4. escolher o método para obter a relação funcional

Um dos melhores e mais usados métodos para encontrar o modelo é o método dos mínimos quadrados.

3.3.2 Regressão Linear Simples e o método dos mínimos quadrados

Agora tem-se os pontos do comportamento observado e a curva-modelo. Observa-se que os pontos observados não pertencem todos à curva-modelo, ou possivelmente, nenhum deles.

O método dos mínimos quadrados baseia-se no seguinte objetivo: encontrar uma equação estimada para “Y” em função de “X” de forma que as distâncias entre os pontos do comportamento observado e os pontos da curva-modelo sejam, no total, as menores possíveis. Mais precisamente, a relação funcional entre “Y” e “X”, com o menor erro possível é obtida através da minimização da soma dos quadrados das distâncias entre os pontos do comportamento observado e os pontos da curva-modelo.

Será apresentado aqui o modelo linear simples, que considera somente uma variável independente, pois foi o utilizado no presente trabalho.

Deseja-se obter a relação funcional linear de “Y” em relação a “X”, ou seja, “Y” será escrito em função de “X” através da equação:[41]

$$Y_i = \beta_0 + \beta_1 X_i + \epsilon_i \quad (3.4)$$

onde:

Y_i e X_i são, respectivamente, a variável dependente e a variável independente de cada um dos “N” pontos “i” observados, ou seja, $i = 1, 2, 3, \dots, n$

β_0 é o ponto no qual a reta intercepta o eixo “Y”

β_1 é a inclinação da reta, ou coeficiente angular ou coeficiente de regressão

ϵ_i é o erro que está associado à distância entre o valor Y_i observado e o correspondente ponto na curva-modelo, para o mesmo i

A equação para o erro, fica: [40]

$$\epsilon_i = Y_i - \beta_0 - \beta_1 X_i \quad (3.5)$$

Elevando ambos os lados ao quadrado:

$$\epsilon_i^2 = (Y_i - \beta_0 - \beta_1 X_i)^2 \quad (3.6)$$

Colocando o somatório em ambos os lados:

$$\sum_{i=1}^N \epsilon_i^2 = \sum_{i=1}^N (Y_i - \beta_0 - \beta_1 X_i)^2 \quad (3.7)$$

Derivando-se em relação à β_1 e igualando-se a zero e, depois, derivando-se em relação à β_0 e igualando-se a zero, obtém-se

$$\beta_1' = \frac{\sum X_i Y_i - \frac{\sum X_i \sum Y_i}{N}}{\sum X_i^2 - \frac{(\sum X_i)^2}{N}} \quad (3.8)$$

$$\beta_0' = \sum Y_i - \beta_1' \sum X_i \quad (3.9)$$

onde todos os somatórios são de $i = 1$ a $i = N$

Assim, foi encontrada a relação funcional que se desejava determinar:

$$Y_i' = \beta_0' + \beta_1' X_i \quad (3.10)$$

3.3.3 Coeficiente de Determinação

A pergunta natural que se faz agora é: quanto essa relação funcional obtida “explica” o comportamento observado?

A resposta a essa pergunta está no cálculo do coeficiente de determinação.

Costuma-se denominar por R^2 (pronuncia-se “erre dois”), o valor que indica o quanto a curva-modelo representa os dados observados e calcula-se por

$$R^2 = 1 - \frac{SS_{resid}}{SS_{total}} \quad (3.11)$$

sendo que o coeficiente é tão melhor quanto mais próximo do valor 1 e onde:

SS_{resid} é a soma dos resíduos ao quadrado, ou seja, para cada ponto é calculado a diferença entre o “Y” observado e o “Y” da curva-modelo obtida. O valor é elevado ao quadrado, para que seja escalar e positivo.

$$SS_{resid} = \sum_{i=1}^N Y_i - Y_i' \quad (3.12)$$

SS_{total} é o quadrado da soma das diferenças da variável dependente em relação à sua média, ou seja, para cada ponto calcula-se $(Y - Y_{medio})^2$ e o final faz-se a soma.

$$SS_{total} = \sum_{i=1}^N (Y_i - Y_{medio})^2 \quad (3.13)$$

3.4 Conclusão Parcial

Foram apresentados no Capítulo 2 os conceitos de astronomia necessários à compreensão desse trabalho. No presente capítulo foi apresentada a Ciência de Dados em cinco passos, seguindo o modelo OSEMN, além das técnicas de modelagem dos dados que serão utilizadas nesse trabalho. No próximo capítulo será apresentada a

metodologia que foi utilizada para analisar concomitantemente milhares de estrelas e, em função dessa análise, sugerir novos membros ou a exclusão de outros, do já conhecido e bem estudado grupo em movimento Beta Pictoris.

Capítulo 4

Metodologia

O principal objetivo do presente trabalho é encontrar uma metodologia computacional para confirmar ou encontrar novos membros dos grupos estelares em movimento, analisando concomitantemente o maior número de estrelas possível e levando-se em conta a maior quantidade de características disponíveis das estrelas.

Neste capítulo serão apresentados todos os passos da metodologia desenvolvida, iniciando pelos passos para produção das bases de dados que serão utilizadas na metodologia computacional, que é uma das contribuições desse trabalho e cujo *workflow* é mostrado na figura 4.1.

Em seguida, na fase de exploração dos dados são apresentados os gráficos de dispersão nos espaços das velocidades e das posições. Na fase da modelagem são realizados experimentos usando a PCA e regressão linear, quando serão obtidas importantes conclusões a respeito da representação das estrelas no espaço das duas primeiras componentes principais, tendo sido feita a análise com um conjunto particular de características das estrelas. Essa percepção bem como o desenvolvimento da metodologia a partir da mesma, constitui-se a segunda e principal contribuição desse trabalho.

Na fase da interpretação dos dados são apresentados os passos da evolução retrógrada de estrelas selecionadas e os critérios para confirmação ou não de uma nova estrela como membro de um grupo em movimento conhecido. Um dos testes leva em conta o comportamento das órbitas de todas as estrelas do grupo e das estrelas selecionadas, sendo a terceira contribuição desse trabalho e aqui batizado de “teste das órbitas”.

4.1 Obtenção dos dados (OSEMN)

Foi visto na seção 2.1 que para estudar os grupos estelares em movimento são necessários os dados cinemáticos completos das estrelas, ou seja, posições e velocidades. É fundamental começar utilizando o melhor catálogo de posição e velocidades

disponível.

Nesta sessão serão apresentados os catálogos de estrelas originados da Missão Hipparcos, o catálogo escolhido para esse trabalho e os dados que foram obtidos de outros catálogos.

4.1.1 Catálogos da Missão Hipparcos

Catálogos de estrelas são escritos desde os tempos mais remotos. E, há até bem pouco tempo, todos eles eram construídos com base em observações realizadas aqui da Terra. Em muitas situações, observações das mesmas estrelas eram feitas de diferentes latitudes para que efeitos fossem eliminados na redução final. Nesse trabalho estamos interessados nas posições e nas velocidades das estrelas e para isso é fundamental usar um catálogo baseado em astrometria, que é o ramo da astronomia relacionado às medidas precisas de posição e movimento dos corpos celestes [1].

O primeiro catálogo de estrelas que se tem conhecimento foi compilado por Hiparco, astrônomo grego, baseado nas observações feitas por ele mesmo entre os anos 147 a.C e 127 a.C., no observatório de Rhodes (na Grécia) também construído por ele. O catálogo apresentava a posição e a magnitude de 850 estrelas, sendo que Hiparco criou uma escala de magnitude que, em sua essência, é usada até hoje. A magnitude específica de alguma maneira o brilho da estrela e a escala criada por Hiparco era dividida em 6 categorias, sendo 1 a mais brilhante e 6 a mais fraca visível a olho nu.

A Missão Astrométrica Hipparcos foi a primeira missão espacial a coletar dados astrométricos de alta precisão: cerca de 2 milissegundos de arco (mas) para a paralaxe, de 0,7 a 0,9 mas para a posição e 0,7 a 0,9 mas/ano para os movimentos próprios. A missão foi realizada pela ESA (Agência Espacial Européia) e o satélite Hipparcos operou de agosto de 1989 a março de 1993 colhendo dados em distâncias de até 150 pc da Terra.

Um dos produtos da Missão foi o catálogo Hipparcos, publicado em 1997, que possui uma grande quantidade de dados astrométricos e fotométricos para 118218 estrelas. Os catálogos originados da missão Hipparcos são os primeiros catálogos observacionais que contém a paralaxe das estrelas para todas as estrelas do programa de observação [5].

Embora se possa pensar que o nome da missão tenha sido uma homenagem ao astrônomo grego conhecido como o pioneiro da astronomia científica, Hiparco, o nome “ Hipparcos” é um acrônimo que significa *High Precision Parallax Collecting Satellite*. O ano de referência é $T_0 = 1991, 25$.

O catálogo Hipparcos 1997 foi obtido da base de dados Vizier, que reside na plataforma CDS *Centre de Données astronomiques de Strasbourg - Strasbourg as-*

tronomical Data Center da *Université de Strasbourg* [42], onde está designado por “I/239/hip_main”

Dez anos após a disponibilização dos dados de Hipparcos foi publicada uma nova redução dos dados astrométricos para estrelas mais brilhantes que magnitude 8 [18]. Esse novo catálogo é conhecido como Hipparcos2 (HIP 2) e apresenta as mesmas quantidades do Hipparcos anterior, que passou a ser chamado Hipparcos1 (HIP 1), mas para 117.955 estrelas. Embora mais recente, esse catálogo não tornou obsoleto o anterior, aqui referido como Hipparcos 1997. Pelo contrário, muitos trabalhos posteriores à essa publicação continuaram a usar o Hipparcos 1997. Além disso, não se compreendeu a nova redução realizada por Van Leeuwen e, por fim, o catálogo foi retirado por algum tempo da plataforma Vizier [42]. Por esses motivos no presente trabalho, foi utilizado o Catálogo Hipparcos,1997 [5].

Em 2012, foi disponibilizado o chamado XHIP, uma extensão do Hipparcos 2, [19] no qual foram acrescentados os valores das velocidades radiais para 46.392 estrelas do Hipparcos 2, obtidas de 47 publicações desde 1992. Segundo a publicação de Anderson e Francis, para sistemas estelares múltiplos os erros de paralaxe em HIP2 são maiores que os erros de paralaxe em HIP1 e para esses 1922 casos, os valores da paralaxe em XHIP são os de HIP1 e não os de HIP2. Como XHIP é baseado em HIP2, e tendo em vista os motivos apresentados acima, esse catálogo também não foi utilizado no presente trabalho.

4.1.2 Busca pelas velocidades radiais

Tendo obtido os cinco parâmetros astrométricos do Catálogo Hipparcos para cerca de 117 mil estrelas, é preciso agora buscar fontes para obtenção das velocidades radiais. Há algumas dezenas de catálogos baseados em observações e outros em revisão de literatura sendo que o número de estrelas apresentadas vai de poucas dezenas a até milhares de estrelas. Há catálogos que fornecem as velocidades radiais de estrelas que são de outros catálogos e que não estão em sua maioria no catálogo Hipparcos.

A medida das velocidades radiais está diretamente ligada às posições das linhas nos espectros das estrelas. Quando se tem uma espectroscopia de alta resolução é que se pode obter linhas bem determinadas no espectro e assim, medidas de alta precisão. O artigo de L.Lindegren e D.Dravins [43] intitulado *The fundamental definition of radial velocity* apresenta um estudo profundo das velocidades radiais, suas medidas, as dificuldades para a realização das medidas, o efeito relativístico e as limitações para as medidas espectroscópicas, dentre outros itens.

Após esse estudo e a análise de alguns dos catálogos disponíveis resolveu-se mesclar dois catálogos de velocidades radiais, Pulkovo e SACY, cujas descrições resu-

midas são apresentadas a seguir.

Catálogo de Pulkovo (PCRIV)

O chamado PCRIV, “*Pulkovo Compilation of Radial Velocities*” foi criado para o estudo da cinemática do braço espiral local da Galáxia, também conhecido como o Braço de Orion, que é o braço espiral menor que contém o Sistema Solar. O catálogo contém velocidades radiais para 35495 estrelas do catálogo Hipparcos de vários tipos espectrais e classes de luminosidade. A precisão média das velocidades radiais obtidas é de $0,7\text{km}^{-1}$. Para a construção desse catálogo foram utilizadas mais de 60000 velocidades radiais de 203 publicações, que foram cuidadosamente analisados. Com os valores já publicados foram calculadas médias e os erros sobre essas médias. [44]

No artigo sobre o catálogo [44], Gontcharov, do Observatório de Pulkovo, Russia, cita as dificuldades de se obter velocidades radiais absolutas para estudo da cinemática estelar, ou seja, velocidades radiais relativas ao baricentro do Sistema Solar. [44].

O catálogo Pulkovo foi obtido da base de dados Vizier, que reside na plataforma CDS *Centre de Données astronomiques de Strasbourg - Strasbourg astronomical Data Center* da *Université de Strasbourg* [42], onde está designado por “III/252/table8.dat”

Catálogo SACY

O catálogo de velocidades radiais SACY (*Search for Associations Containing Young Stars*) apresenta velocidades radiais obtidas por espectroscopia de alta resolução e dedicada justamente às estrelas jovens na vizinhança do Sol (termo conhecido como “*nearby young stars*”) do catálogo Hipparcos (I/239) e do catálogo Tycho, que é também produto da Missão Hipparcos. O catálogo possui 1626 estrelas, dentre as quais, 1511 tiveram pelo menos um espectro de alta-resolução obtido e analisado. Outras 115 foram incluídas para completar o conjunto.

Cerca de 70% das observações espectroscópicas foram obtidas com o espectrógrafo FEROS no telescópio de 1,5m do ESO (Observatório Europeu do Sul, *European Southern Observatory*) e as demais 30% foram coletadas pelo espectrógrafo coude acoplado ao telescópio de 1,60m do Observatório do Pico dos Dias (OPD) no Laboratório Nacional de Astrofísica, aqui no Brasi. [45]

O catálogo SACY também foi obtido da base de dados Vizier, [42], onde está designado por “J/A+A/460/695”

4.1.3 Os grupos em movimento a serem estudados

Dos grupos em movimento já conhecidos, foram escolhidos o Beta Pictoris e o AB Doradus que são os que possuem a maior parte das estrelas no Catálogo Hipparcos. Somente uma estrela de cada um dos referidos grupos não pertence ao Hipparcos. O Beta Pictoris é grupo mais jovem entre os grupos conhecidos com idade em torno de 20 myr e o AB Doradus um dos mais velhos com idade em torno dos 50 Myr. O AB Doradus é um dos mais próximos da Terra, com distância média de 20 pc e o Beta Pictoris tem distância média de 30 pc.

O grupo estelar em movimento Beta Pictoris

Na seção 2.3.2 foi visto que as estrelas dos grupos em movimento podem ser diferentes de acordo com o autor e seus critérios usados para definir o grupo e elencar seus membros. Na seção 2.4 foi apresentada a abordagem dinâmica de Ortega *et al.* para analisar os grupos em movimento fazendo a evolução retrógrada a partir dos dados de posição e de velocidade atuais e levando em conta o potencial galáctico. Isso é possível porque as estrelas dos grupos em movimento, como foi visto 2.3.1, não possuem interação gravitacional entre si, ou seja, estão sob ação somente do campo gravitacional geral da Galáxia.

Nesse trabalho, será usado o grupo em movimento Beta Pictoris segundo Ortega, cujas estrelas foram elencadas na tabela 2.1

O grupo estelar em movimento AB Doradus

O grupo estelar em movimento AB Doradus foi apresentado por Zuckerman e Bessel [6] com 36 estrelas, sendo 35 do Catálogo Hipparcos. É um dos grupos mais próximos da Terra, a uma distância média de 20 pc, e com maior idade, em torno de 50 Myr. No quadro 4.1 são apresentadas as identificações Hipparcos das estrelas que compõem o grupo, como apresentado por Zuckerman e Bessel, em 2004 [6].

Tabela 4.1: Estrelas do Grupo AB Doradus segundo Zuckerman e Bessel [6]

Hip 3589	Hip 16563 A	Hip 30314	Hip 106231
Hip 5191	Hip 16563 B	Hip 31711	Hip 110526
Hip 6276	Hip 17695	Hip 31878	Hip 113579
Hip 10272	Hip 18859	Hip 36349	Hip 113597
Hip 12635	Hip 19183	Hip 63742	Hip 114066
Hip 12638	Hip 25283	Hip 76768	Hip 114530
Hip 13027	Hip 25647	Hip 81084	Hip 115162
Hip 14807	Hip 26369	Hip 82688	Hip 118008
Hip 14809	Hip 26373	Hip 86346	

4.2 Limpeza e tratamento dos dados (OSEMN)

Nessa seção serão descritos todos os passos para limpeza dos dados, que incluíram: extração de colunas, extração de linhas, troca de campo *string* por valor numérico e tratamento de valores ausentes. Depois foi feito o tratamento dos dados através da mesclagem das bases Sacy e Pulkovo já limpas e das transformações de coordenadas, que resultou na produção das três bases de dados que serão usadas na metodologia computacional, e cujo *workflow* é mostrado na figura 4.1.

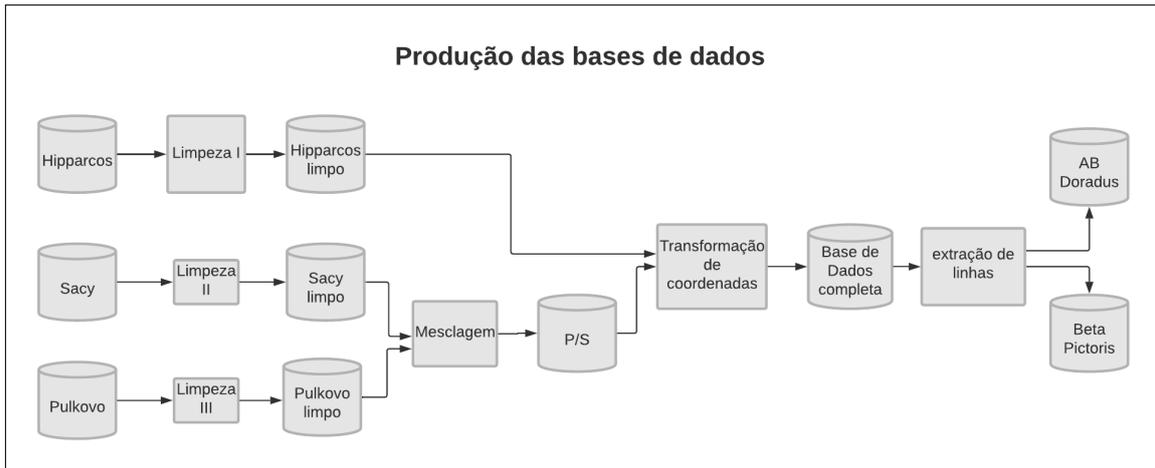


Figura 4.1: *Workflow* da produção das bases de dados

4.2.1 Dados do catálogo Hipparcos

Paralaxes negativas e de valor nulo no catálogo Hipparcos

Foram encontradas no Catálogo Hipparcos algumas estrelas com paralaxe (π) nula ou negativa, o que não tem sentido físico porque uma vez que a distância dessas estrelas em parsec é calculada por $\frac{1000}{\pi(mas)}$ a distância seria, respectivamente negativa ou indeterminada. Não foi encontrada a explicação para esses dados nas explicações sobre os dados do catálogo e nem nos artigos. Entendeu-se que esses valores estariam ligados à precisão, ou seja, para valores muito pequenos de paralaxe e com erros consideráveis o valor médio ficaria oscilando em torno do zero.

As estrelas com valores de paralaxe negativo ou nulo foram excluídas da base de dados desse trabalho. As estrelas que estão distantes até 100 pc do Sol, ou seja, com paralaxes maiores que 0,01" foram medidas com alta precisão pelo satélite da Missão Hipparcos, mas para estrelas com distâncias maiores que 100pc isso não é verdade. A precisão das medidas dos cinco parâmetros astrométricos: ascensão reta (α), declinação (δ), movimento próprio em ascensão reta (μ_α), movimento próprio

em declinação (μ_δ) e paralaxe (π) também variou com outros fatores, como, por exemplo, a magnitude conforme mostra a figura 4.2 [5].

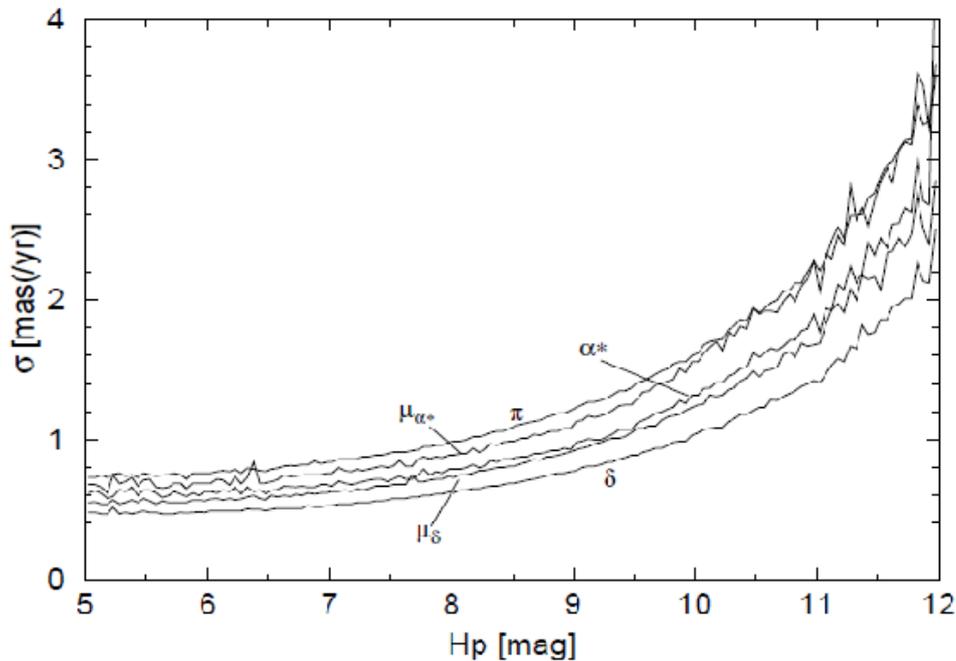


Figura 4.2: Hipparcos: Erros médios dos cinco parâmetros astrométricos como função da magnitude [5]

Construção da tabela do tipo espectral

No presente trabalho foi feita uma atribuição de valor numérico para o tipo espectral: a cada letra foi atribuído um número conforme a tabela 4.2 e a esse número foi somado o número da classe espectral. Por exemplo, para uma estrela “F8” o valor atribuído foi “48”. Além disso, há também as classes de luminosidade, que são expressas em algarismos romanos, mas que no presente trabalho não foram codificadas por ter sido entendido que trata-se de uma especificação fina e variável.

Na base de dados há 247 estrelas sem especificação do tipo espectral, ou seja, o campo está vazio.

4.2.2 Mesclagem de Pulkovo com Sacy

Para formar a base de dados do presente trabalho, foram utilizados os seguintes critérios para estrelas que possuem velocidade radial em ambos os catálogos, Pulkovo e SACY:

Tabela 4.2: O tipo espectral expresso em números

Tipo Espectral	Valor Numérico	Tipo Espectral	Valor Numérico
O	10	R	80
B	20	S	80
A	30	N	80
F	40	C	80
G	50	D	80
K	60	W	80
M	70		

1. se estrela de SACY está contida em Pulkovo e erro de SACY é menor que erro de Pulkovo, então usa velocidade radial de SACY, senão usa Pulkovo
2. se estrela de SACY está contida em Pulkovo e campo do erro está vazio, decisão manual, sendo observados outros parâmetros como a fonte do valor e número de observações envolvidas no cálculo do valor

A base mesclada foi nomeada de “P/S”. Embora o catálogo de Pulkovo tenha 35495 estrelas e o Catálogo SACY tenha 1626, a base de dados possui 34758 estrelas, uma vez que foram retiradas as estrelas com paralaxe nulas e negativas além do fato de muitas estrelas de SACY estarem também em Pulkovo.

4.2.3 Transformações de coordenadas

O segundo passo no tratamento dos dados é a implementação das funções de transformação. É preciso transformar as coordenadas de posição das estrelas que são publicadas nos catálogos referidas ao sistema equatorial celeste (ascensão reta e declinação) para o sistema de referência Galáctico (longitude Galáctica e latitude Galáctica). É também necessário transformar as quantidades obtidas dos catálogos: movimento próprio em ascensão reta, movimento próprio em declinação, velocidade radial e paralaxe nas componentes da velocidade da estrela no sistema Galáctico: U, V, W , sendo U e V sobre o plano Galáctico com U positiva na direção do centro da Galáxia, V positiva na direção da rotação da Galáxia e W positiva na direção do Polo Norte Galáctico (isso quando se usa a regra da mão direita).

Johnson e Soderblom iniciam seu artigo intitulado *Calculating Galactic Space Velocities and their Uncertainties, with an Application to the Ursa Major Group* [46] afirmando que “esse é um problema de trigonometria esférica e que deve ser direto, mas que erros são facilmente cometidos”. Assim, nesse artigo, as expressões de transformação são explicitadas, bem como as expressões das incertezas nas grandezas calculadas e é fornecido como exemplo os resultados para 16 estrelas do Grupo da Ursa Maior. Esse artigo, embora tenha sido publicado em 1987, é referência mundial para tal tipo de transformação e foi utilizado no presente trabalho.

Para facilitar a visualização, o quadro 4.3 apresenta as variáveis da base de dados, suas representações e unidades.

Tabela 4.3: A variáveis das estrelas na base de dados, suas representações e unidades

Representação	Descrição	Unidade
α	ascensão reta	horas ou radiano
δ	declinação	graus ou radiano
μ_α	movimento próprio	segundos de arco por ano
μ_δ	movimento próprio	segundos de arco por ano
π	paralaxe	segundos de arco
ρ	velocidade radial	km por segundo de tempo

O sistema de coordenadas Galáctico em relação ao Equatorial Celeste

Três ângulos definem o Sistema de Coordenadas Galáctico em relação ao Equatorial Celeste: os dois ângulos que definem a posição do Polo Norte Galáctico (PNG) no sistema equatorial celeste e o ângulo de posição do Polo Norte Galáctico em relação ao semi-círculo máximo que passa pelo PNG e pelo ponto zero da longitude Galáctica.

$$\alpha_{PNG} = 192,485948^\circ, \delta_{PNG} = 27,12825^\circ \quad (4.1)$$

$$\theta_0 = 122,93192^\circ \quad (4.2)$$

Os valores utilizados são os que foram publicados com o Catálogo Hipparcos que também definiu novos valores de constantes para o sistema de coordenadas celeste.[5]

Transformação das coordenadas de posição

O primeiro passo é obter a longitude e latitude Galácticas, a partir das coordenadas equatoriais celestes:

$$\begin{bmatrix} \cos b \cos l \\ \cos b \sin l \\ \sin b \end{bmatrix} = \mathbf{T} \begin{bmatrix} \cos \delta \cos \alpha \\ \cos \delta \sin \alpha \\ \sin \delta \end{bmatrix} \quad (4.3)$$

onde a matriz de transformação \mathbf{T} é dada por

$$\begin{bmatrix} +\cos\theta_0 & +\sin\theta_0 & 0 \\ +\sin\theta_0 & -\cos\theta_0 & 0 \\ 0 & 0 & +1 \end{bmatrix} \begin{bmatrix} -\sin\delta_{PNG} & 0 & +\cos\delta_{PNG} \\ 0 & -1 & 0 \\ +\cos\delta_{PNG} & 0 & -\sin\delta_{PNG} \end{bmatrix} \begin{bmatrix} +\cos\alpha_{PNG} & +\sin\alpha_{PNG} & 0 \\ +\sin\alpha_{PNG} & -\cos\alpha_{PNG} & 0 \\ 0 & 0 & +1 \end{bmatrix} \quad (4.4)$$

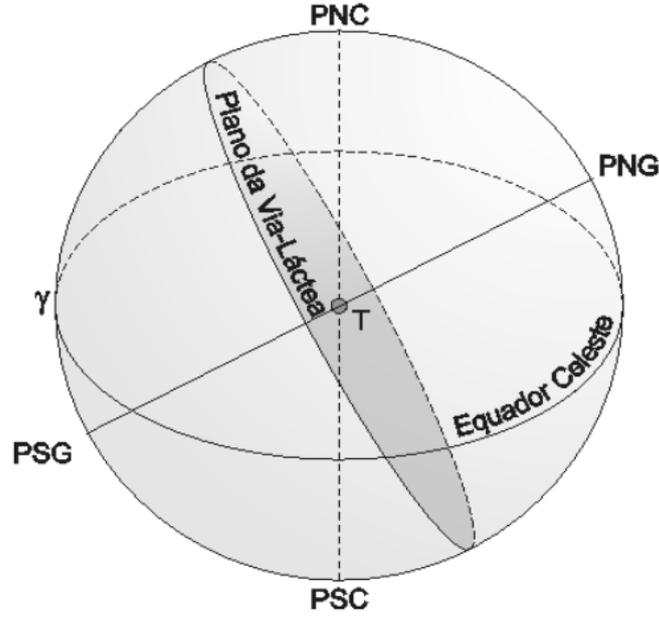


Figura 4.3: O sistema Galáctico e o Sistema Equatorial Celeste [5]

A matriz \mathbf{T} foi calculada e os valores conferidos com aqueles apresentados na publicação do catálogo Hipparcos [5]

Obtenção das componentes da velocidade da estrela no sistema Galáctico

Seja agora definida a matriz de coordenadas \mathbf{A}

$$\begin{bmatrix} +\cos\alpha\cos\delta & -\text{sen}\alpha & -\cos\alpha\text{sen}\delta \\ +\text{sen}\alpha\cos\delta & +\cos\alpha & -\text{sen}\alpha\text{sen}\delta \\ +\text{sen}\delta & 0 & +\cos\delta \end{bmatrix} = \begin{bmatrix} \cos\alpha & +\text{sen}\alpha & 0 \\ \text{sen}\alpha & -\cos\alpha & 0 \\ 0 & 0 & -1 \end{bmatrix} \begin{bmatrix} \cos\delta & 0 & -\text{sen}\delta \\ 0 & -1 & 0 \\ -\text{sen}\delta & 0 & -\cos\delta \end{bmatrix} \quad (4.5)$$

As componentes da velocidade da estrela no sistema de referência Galáctico, são obtidas por:

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} = \mathbf{B} \begin{bmatrix} \rho \\ k\mu_\alpha/\pi \\ k\mu_\delta/\pi \end{bmatrix} \quad (4.6)$$

onde

$$\mathbf{B} = \mathbf{T} \cdot \mathbf{A} \quad (4.7)$$

e

$$k = \frac{149597870,700km}{365,2422 \times 86400} \quad (4.8)$$

é o equivalente em km/s de 1 unidade astronômica por ano trópico, que tem 365,2422 dias e é o intervalo entre dois equinócios de primavera consecutivos, ou dois equinócios de outono consecutivos.

No presente trabalho, as implementações são feitas com os valores absolutos das coordenadas e suas velocidades, sem levar em conta os erros em l , b , U, V, W tendo em vista os erros nas variáveis da base de dados α , δ , μ_α , μ_δ , π e ρ .

Após a implementação de todas as equações foi realizado o teste, com resultado positivo, com as 16 estrelas do Grupo Ursa Maior, cujos parâmetros de velocidade estão apresentados no artigo de Johnson and Soderblom [46] e as coordenadas de posição que foram utilizadas são as do catálogo FK4 do equinócio de 1950.0, obtidas pelo Portal SIMBAD na plataforma CDS [42].

4.2.4 A base de dados completa

A base de dados completa é então formada das 34758 estrelas que possuem velocidade radial, dentre as quais 247 não possuem dados de tipo espectral. A cada estrela da referida base estão associados os dados conforme organizados na tabela 4.4.

Tabela 4.4: As características das estrelas na Base de Dados Completa

Característica	Tipo	Denominação
coordenadas equatoriais celestes	posição	α , δ
paralaxe (distância)	posição	π
velocidades equatoriais celestes	velocidade	μ_α , μ_δ
velocidade radial	velocidade	ρ
coordenadas Galácticas	posição	l , b
coordenadas Galácticas cartesianas	posição	X , Y , Z
velocidades Galáctias	velocidade	U, V, W
Magnitude	fotométrica	V_{mag}
Tipo Espectral	espectroscópica	$SpType$

4.2.5 A construção dos dados dos grupos em movimento

Montada a base de dados completa, foram dela extraídas as estrelas dos grupos Beta Pictoris e AB Doradus, para formar as duas bases de dados específicas de cada grupo em movimento, mas com as mesmas características da base de dados completa.

Nem todas as estrelas dos grupos em movimento Beta Pictoris e AB Doradus, possuem valores de velocidade radial nos catálogos Pulkovo e Sacy, e portanto na

base P/S, e por isso não estão na base de dados completa. As tabelas 4.5 e 4.6 apresentam as identificações das estrelas que estão nas bases, respectivamente, Beta Pictoris e AB Doradus.

Essa é a primeira contribuição desse trabalho: a produção da base de dados completa e a extração das bases de dados de cada grupo em movimento.

Tabela 4.5: Estrelas da Base de Dados Beta Pictoris

Hip 560	Hip 103311	Hip 14361	Hip 88395
Hip 21547	Hip 84586	Hip 92680	Hip 10680
Hip 27321	Hip 25486	Hip 95261	Hip 88726
Hip 92024	Hip 99273	Hip 95270	Hip 76629

Tabela 4.6: Estrelas da Base de Dados AB Doradus

Hip 3589	Hip 16563 A	Hip 25647	Hip 63742
Hip 10272	Hip 18859	Hip 26373	Hip 82688
Hip 12638	Hip 19183	Hip 30314	Hip 106231
Hip 13027	Hip 25283	Hip 31711	Hip 110526
Hip 113579	Hip 114530		

4.3 Exploração dos dados (OSEMN)

Os dados da base completa e dos grupos em movimento obtidos conforme descrito na seção 4.2 são explorados através de gráficos de dispersão.

O gráfico das componentes cartesianas das posições atuais no sistema Galáctico heliocêntrico, apresentado na figura 4.4, mostra como as estrelas dos grupos em movimento Beta Pictoris e AB Doradus estão muito distantes entre si.

No espaço das velocidades cartesianas no sistema Galáctico heliocêntrico, figura 4.5 observa-se em determinado ângulo o agrupamento distinto entre as estrelas de cada um dos grupos. Isso já era esperado, tendo em vista que os trabalhos correlatos desde Zuckerman *at al.*, [13] são em sua maioria relacionados ao espaço das componentes das velocidades naquele sistema de referência.

4.4 Modelagem dos dados (OSEMN)

4.4.1 Experimentos com PCA

Foram feitos várias experimentos usando-se a Análise das Componentes Principais (PCA) para todas as estrelas da base completa (34758 estrelas), considerando em cada um delas um conjunto diferente das características das estrelas. É intuitivo,

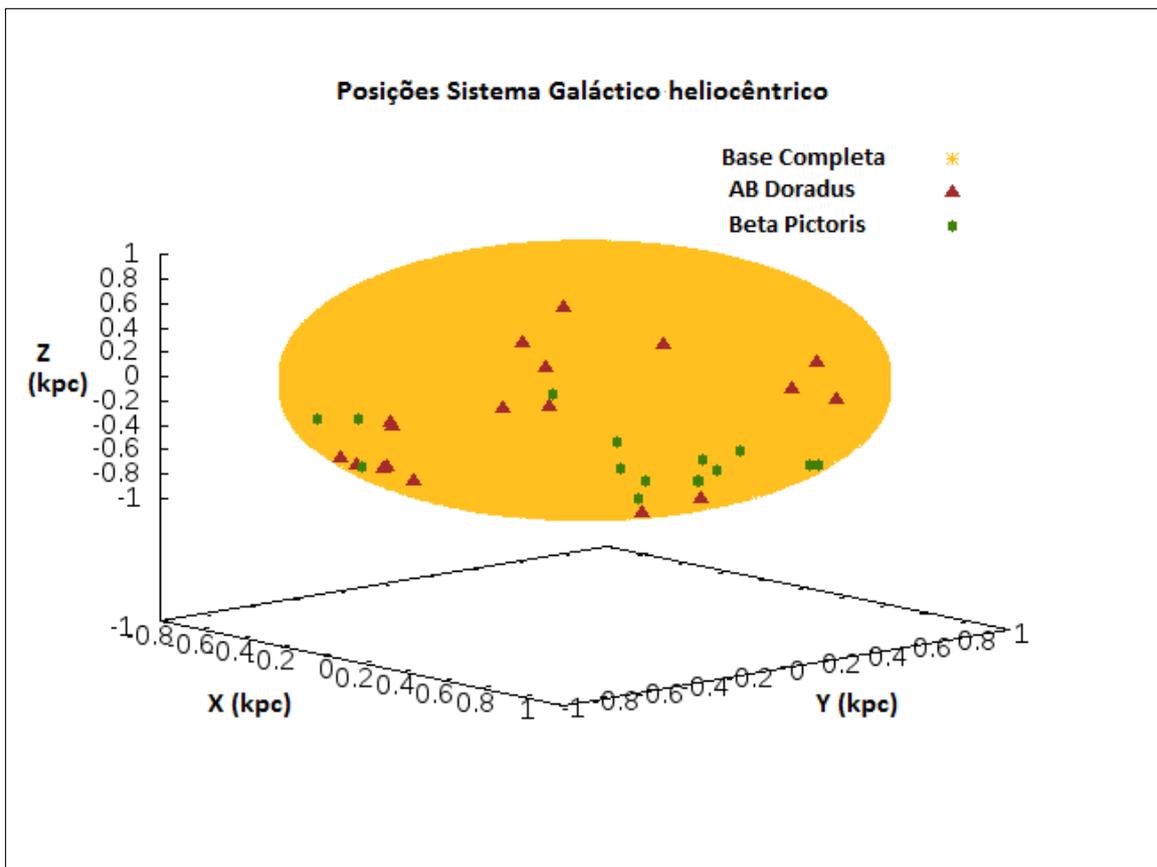


Figura 4.4: Posições de todas as estrelas da base completa e dos grupos AB Doradus e Beta Pictoris

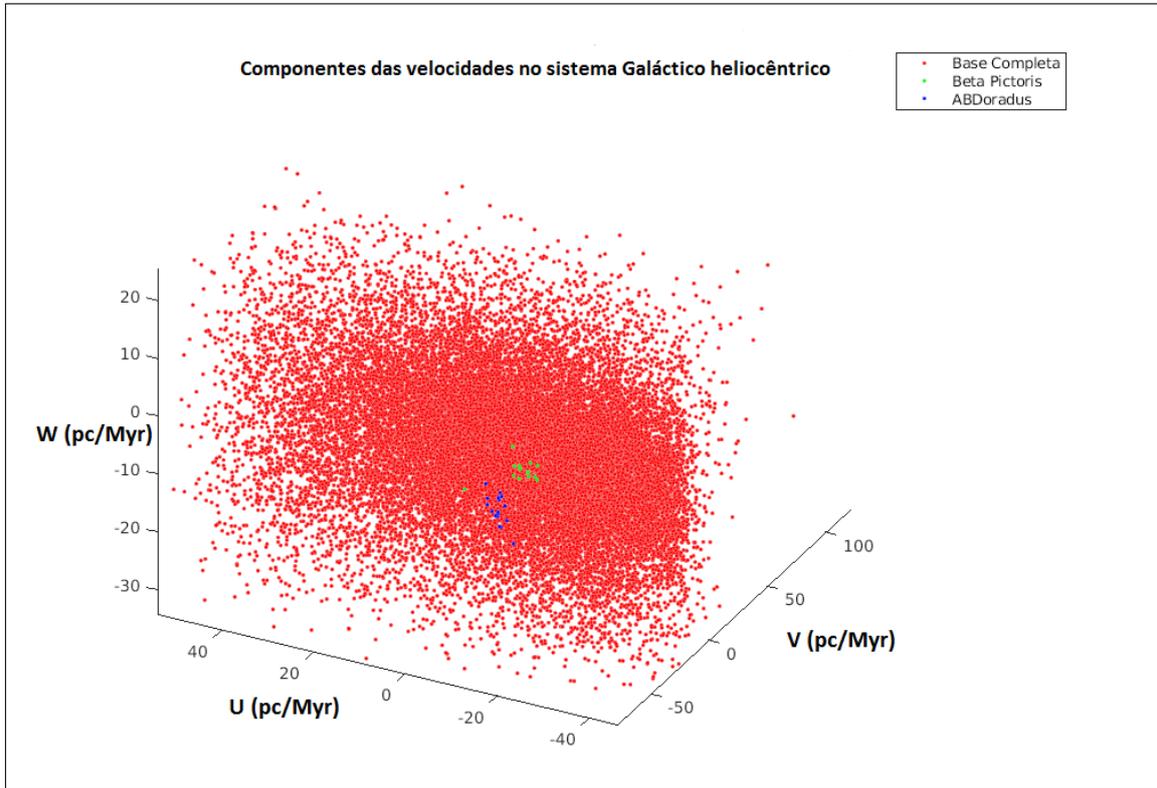


Figura 4.5: Componentes das velocidades de todas as estrelas da base completa e dos grupos AB Doradus e Beta Pictoris

pela definição dos grupos em movimento e pelos trabalhos relacionados, que as componentes das velocidades no sistema Galáctico sejam algumas dessas características, mas outras possibilidades foram também analisadas.

Somente nos casos em que uma das características escolhida é o tipo espectral é que foram retiradas da base completa as 247 estrelas que não possuem o campo tipo espectral.

Como foi visto na tabela 4.4 as dimensões das características são diferentes e, por causa disso, em todas os experimentos foi usado o passo da normalização. Os resultados dos experimentos serão apresentados na seção 5.1. Para cada conjunto de características foram apresentadas e analisadas as variâncias explicadas de cada componente principal, ou seja, o valor percentual do conjunto que é representado (“explicado”) por cada componente principal.

4.4.2 Regressão Linear Simples

Para algumas das experiências realizadas com a PCA, foi feita a regressão linear simples, no espaço das duas primeiras componentes principais.

Foi observado que além das redução de dimensão, o uso da PCA também favoreceu à observação de uma representação característica para os grupos em movimento:

quando a PCA é feita usando como características as componentes cartesianas da velocidade no sistema Galáctico helocêntrico “U”, “V” e “W” e, além disso, a magnitude Vmag “Vmag” e o tipo espectral “St”, as estrelas dos grupos em movimento, no espaço das duas primeiras componentes principais, estão muito bem ajustados a retas diferentes, com coeficiente de determinação muito próximo do valor “1”.

4.5 Interpretação dos dados (OSEMN)

As retas obtidas conforme citado seção 4.4.2, trazem consigo a ideia de que outros novos membros do mesmo grupo em movimento estejam em uma região próxima.

O cerne da metodologia está:

1. na análise das componentes principais usando as características “U”, “V”, “W”, “Vmag” e “SpType” de todas as estrelas das três bases,
 2. na representação dessas estrelas no espaço das duas primeiras componentes principais,
 3. na obtenção das retas de ajuste para os grupos em movimento Beta Pictoris e ABDoradus,
 4. na seleção das estrelas nas proximidades de cada reta, em busca de novos membros para o grupo representado pela reta,
 5. na busca de novos membros do grupo através da observação do local de nascimento na época do nascimento do grupo,
 6. nos testes para confirmar se os novos membros pertencem ao referido grupo
- e constitui-se a principal contribuição desse trabalho.

4.5.1 Seleção de estrelas

A seleção das estrelas, com objetivo de encontrar novos membros do grupo em movimento já conhecido, é feita no espaço das duas primeiras componentes principais, obtidas da PCA com as características “U”, “V”, “W”, “Vmag”, “St”.

Para seleção foi utilizado o seguinte critério:

- sejam d_i , as distâncias de cada uma das estrelas do grupo em movimento Beta Pictoris à reta obtida conforme descrito na seção 4.4.2
- seja D_{max} o valor máximo de todos os valores de d_i , ou seja a distância da estrela mais distante da reta

- Seleciona-se todas as estrelas que estão a uma distância da reta menor ou igual a $D_{max} + 0,05 \cdot D_{max}$

Somam-se às estrelas selecionadas as 247 estrelas sem especificação de tipo espectral que, por isso, não constaram da análise das componentes principais.

4.5.2 Evolução retrógrada com as estrelas selecionadas

Com as estrelas selecionadas poderia ser usado qualquer outro método conhecido para encontrar os novos membros sugeridos para o grupo em movimento estudado. Nesse trabalho, escolheu-se usar a abordagem dinâmica de Ortega *at al.*, que foi descrita na seção 2.4 e consiste na evolução retrógrada das estrelas, partindo-se das coordenadas e velocidades atuais e levando em conta o potencial gravitacional total da Galáxia.

Assim, foi feita inicialmente a evolução retrógrada somente das estrelas para o grupo escolhido para análise: o grupo em movimento Beta Pictoris. Para cada passo da evolução é calculado o confinamento do grupo conforme descrito em 2.4.

Foi observado o tempo e local para o qual o confinamento do grupo é o maior possível:

1. o tempo é a época do nascimento e fornece a idade dinâmica do grupo
2. o local é o local de nascimento do grupo e é definido pelo intervalo fechado entre:

- x_- e x_+ : sendo x_- a menor coordenada “X” e x_+ a maior coordenada “X”
- y_- e y_+ : sendo y_- a menor coordenada “Y” e y_+ a maior coordenada “Y”
- z_- e z_+ : sendo z_- a menor coordenada “Z” e z_+ a maior coordenada “Z”

Em segundo lugar foi feita a evolução retrógrada para todas as estrelas selecionadas e observado se existe alguma estrela diferente das estrelas de Beta Pictoris dentro do local de nascimento, na época do nascimento. Se existirem, essas estrelas serão os novos membros sugeridos do grupo.

Finalmente, os novos membros sugeridos são analisados através dos critérios apresentados na subseção 4.5.3 a fim de se tornarem ou não membros confirmados.

4.5.3 Critérios para confirmação da estrela no grupo em movimento

No primeiro passo da fase da evolução retrógrada, conforme descrito em 4.5.2 é feita a evolução retrógrada com as estrelas do grupo já conhecido, sem adicionar qualquer

outra.

Se esse grupo já foi estudado pela abordagem dinâmica, então ele já tem uma idade dinâmica conhecida e a idade dinâmica obtida nessa fase da metodologia não deve ser diferente.

Comportamento das órbitas

A figura 4.6 mostra o gráfico em três dimensões das coordenadas de posição no sistema Galáctico heliocêntrico de cada estrela de grupo em estudo, para cada passo da evolução retrógrada, 1 Myr, formando assim as órbitas de todas as estrelas ao longo do tempo, andando para trás a partir da data atual.

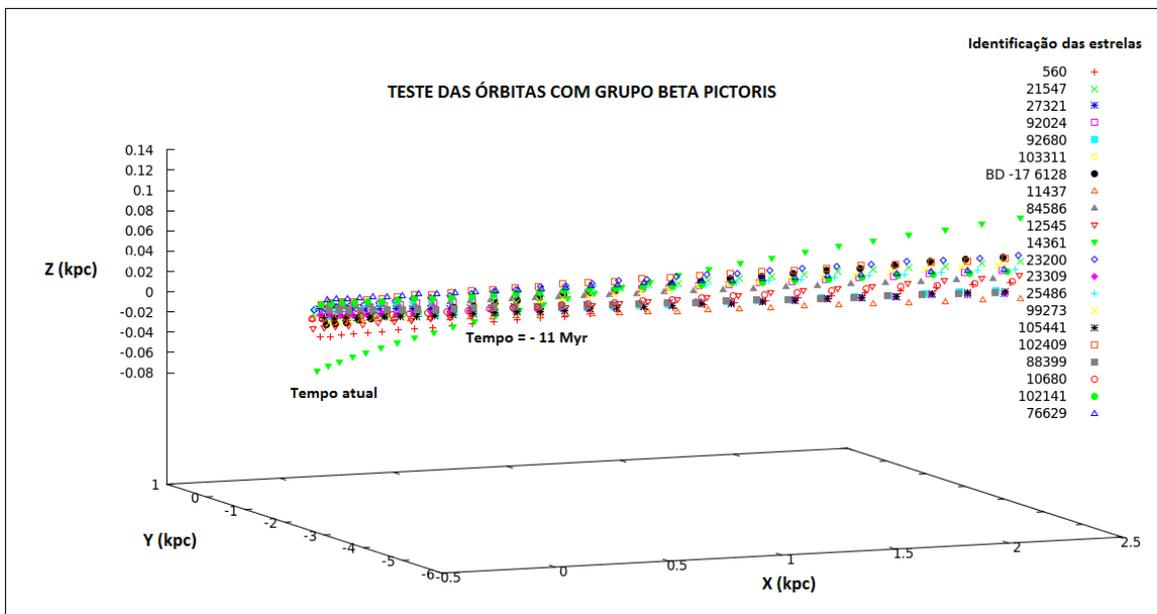


Figura 4.6: Teste das Órbitas

Observou-se que a maioria das estrelas do grupo fazem órbitas parecidas e que uma das estrelas nitidamente apenas cruza o grupo de órbitas na época do nascimento. Essa estrela está, portanto, no local de nascimento do grupo, mas não faz parte do grupo.

Esse critério foi desenvolvido nesse trabalho e constitui-se sua terceira contribuição, sendo nomeado de “teste das órbitas”. É o primeiro critério que será usado na metodologia para analisar os membros sugeridos: um membro sugerido não será confirmado se apresentar uma órbita que não seja coerente com as órbitas dos demais membros confirmados.

Confinamento do grupo em seu nascimento

A região de confinamento das estrelas do grupo é o local onde estão incluídas todas as estrelas do grupo para um determinado tempo da evolução retrógrada. Para essa região calcula-se o centro e as distâncias de todas as estrelas do grupo ao centro. O “raio do grupo” é o valor da média das distâncias das estrelas ao centro do grupo. Quanto mais próximas entre si estão as estrelas do grupo, menor é o raio do grupo e menor é a região de confinamento. Usou-se como medida desse confinamento a relação entre o raio do grupo para um determinado tempo e o maior raio do grupo em toda a evolução 4.9. O nascimento do grupo é determinado quando esse valor é o menor em toda a evolução retrógrada e, portanto, o confinamento é máximo.

$$\text{confinamento máximo} = \frac{\text{Raio mínimo}}{\text{Raio máximo}} \quad (4.9)$$

Quando uma estrela que é membro sugerido é acrescentada ao grupo e é feita nova evolução retrógrada, a estrela só poderá ser membro confirmado se a nova relação do confinamento não aumentar consideravelmente.

Idade Dinâmica do Grupo

A idade dinâmica do grupo é aquela que é determinada pela abordagem dinâmica de Ortega *at al.* e conforme descrito na seção 2.4 é o tempo da evolução retrógrada para o qual o confinamento é máximo.

Ao se fazer a evolução retrógrada incluindo a estrela que é membro sugerido junto com as estrelas já conhecidas e confirmadas do grupo, o valor da idade já determinado previamente não deve sofrer uma alteração considerável.

4.6 Conclusões parciais

Neste capítulo foi apresentada toda a metodologia computacional desenvolvida para estudar um grupo estelar em movimento já conhecido seja sugerindo, confirmando ou descartando alguma estrela. A metodologia é baseada na ciência de dados e está representada na junção das figuras 4.1 e 4.7.

A figura 4.1 está relacionada aos passos de obtenção, limpeza e tratamento dos dados com a finalidade de produzir as bases de dados que serão usadas na metodologia e que foram denominadas “base completa”, “Beta Pictoris” e “AB Doradus”, o que se constitui a primeira contribuição desse trabalho.

A figura 4.7 apresenta os passos de:

1. modelagem através de vários experimentos com a análise das componentes principais (PCA) e da escolha da PCA com um conjunto de características das

estrelas. Todas as estrelas da base completa são representadas no espaço das duas primeiras componentes principais e é nesse espaço que é feita a regressão linear com as estrelas das bases dos grupos em movimento.

2. análise através da seleção de algumas estrelas, da evolução retrógrada com essas estrelas e da busca de novos membros sugeridos para o grupo em movimento em estudo. Os membros sugeridos serão confirmados ou não utilizando-se o critério das órbitas, a terceira contribuição desse trabalho. Complementa-se esse critério, usando-se os testes da idade dinâmica e da razão de confinamento do grupo agora com os novos membros, que não podem ser muito diferentes dos valores obtidos inicialmente.

A metodologia computacional desenvolvida é a principal contribuição deste trabalho.

No capítulo 5 serão apresentados os resultados de cada uma dessas fases e o estudo do grupo em movimento Beta Pictoris segundo Ortega *at al.*

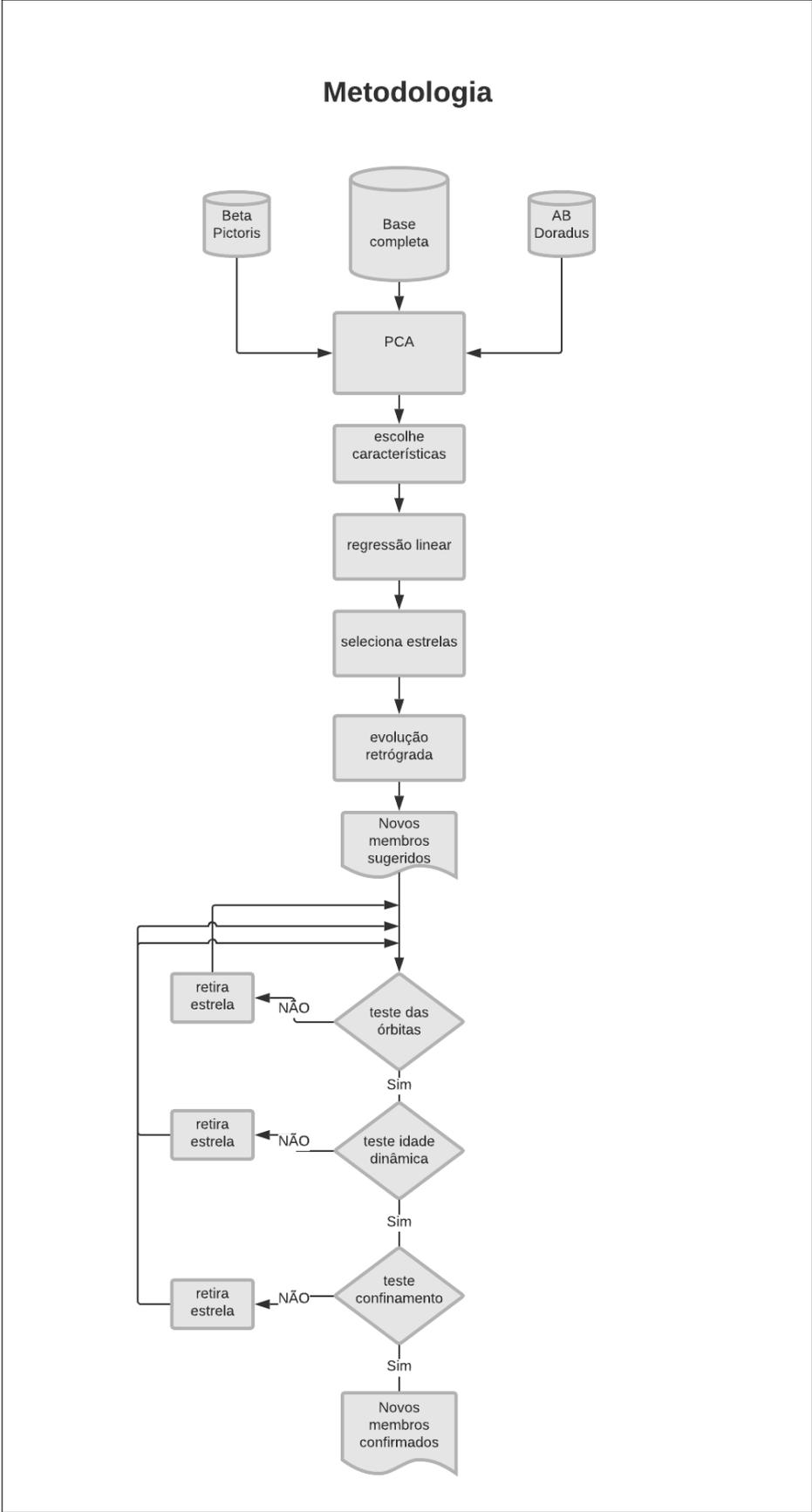


Figura 4.7: *Workflow* da metodologia

Capítulo 5

Resultados

Neste capítulo serão apresentados os resultados encontrados com a aplicação da metodologia apresentada no capítulo 4. Na 5.1 seção são mostrados os gráficos e tabelas referentes aos 7 experimentos mais significativos, dentre os que foram feitos com PCA, sendo que cada experimento utiliza um certo conjunto de características das estrelas. Na seção 5.2 são mostrados os resultados com Regressão Linear Simples para alguns dos experimentos com PCA. Da discussão dos resultados dessas duas seções uma representação para as bases de dados é escolhida: a representação das estrelas das bases de dados no espaço das duas primeiras componentes principais resultantes da PCA com as características componentes das velocidades Galácticas, magnitude e tipo espectral. Nesse espaço de representação são selecionadas estrelas para a evolução retrógrada e na seção 5.3 é observado que há outras estrelas no tempo e local de nascimento do grupo Beta Pictoris segundo Ortega descrito em 2.1. Essas estrelas são as novas estrelas sugeridas para o grupo, que serão ou não confirmadas de acordo com os testes de critérios realizados em 5.4.

5.1 Resultados dos experimentos com PCA

Foram realizados experimentos utilizando a Análise das Componentes Principais (PCA) com diversos conjuntos de combinações de características da base de dados completa e das bases de dados Beta Pictoris e AB Doradus. A implementação foi feita com funções do pacote Octave e utilizou-se a normalização, tendo em vista que as unidades das características são diferentes. Nessa seção são apresentados os resultados mais significativos de todos os experimentos realizados.

O teste da implementação foi feito com a base de dados fornecida no artigo de Lyra *et.al.* [37] e foi totalmente positiva a comparação com os resultados.

Na primeira subseção são apresentados os resultados e na segunda, as observações gerais e discussão dos resultados.

5.1.1 Experimentos com PCA

Para cada experimento são apresentados:

- a tabela com as características utilizadas das estrelas
- o gráfico dos escores de todas as estrelas da base completa no espaço das duas primeiras componentes principais. Em destaque são plotados os escores das bases Beta Pictoris 4.5 e AB Doradus 4.6. Esses escores são as projeções dos objetos originais no espaço das duas primeiras componentes principais, ou seja, são as coordenadas dos objetos originais nas duas primeiras componentes principais. [37]
- a tabela com as variâncias explicadas percentuais de cada componente principal, ou seja, o quanto cada componente principal explica o conjunto de dados

Experimento 1: PCA com as características posições e velocidades equatoriais, magnitude e tipo espectral

As características usadas no experimento 1 estão elencadas na tabela 5.1. Na figura 5.1 as estrelas da Base Completa e das bases Beta Pictoris e AB Doradus são plotadas no espaço das duas primeiras componentes principais. As variâncias explicadas percentuais de cada componente principal são apresentadas na tabela 5.2

Tabela 5.1: Características das estrelas usadas no experimento 1

Característica	Tipo	Denominação	Gráficos PCA
coordenadas equatoriais celestes	posição	α , δ	AR, DE
paralaxe (distância)	posição	π	Prlx
velocidades equatoriais celestes	velocidade	μ_α , μ_δ	pmRA, pmDE
velocidade radial	velocidade	ρ	RV
Magnitude	fotométrica	Vmag	Vmag
Tipo Espectral	espectroscópica	SpType	St

Experimento 2: PCA com as características posições cartesianas Galácticas, magnitude e tipo espectral

As características usadas no experimento 2 estão elencadas na tabela 5.3. Na figura 5.2 as estrelas da Base Completa e das bases Beta Pictoris e AB Doradus são plotadas no espaço das duas primeiras componentes principais. As variâncias explicadas percentuais de cada componente principal são apresentadas na tabela 5.4

Tabela 5.2: Variâncias Explicadas para PCA experimento 1

Número de Ordem da Componente Principal	Variância Explicada (%)
1 ^a	17
2 ^a	17
3 ^a	14
4 ^a	13
5 ^a	11
6 ^a	11
7 ^a	9
8 ^a	8

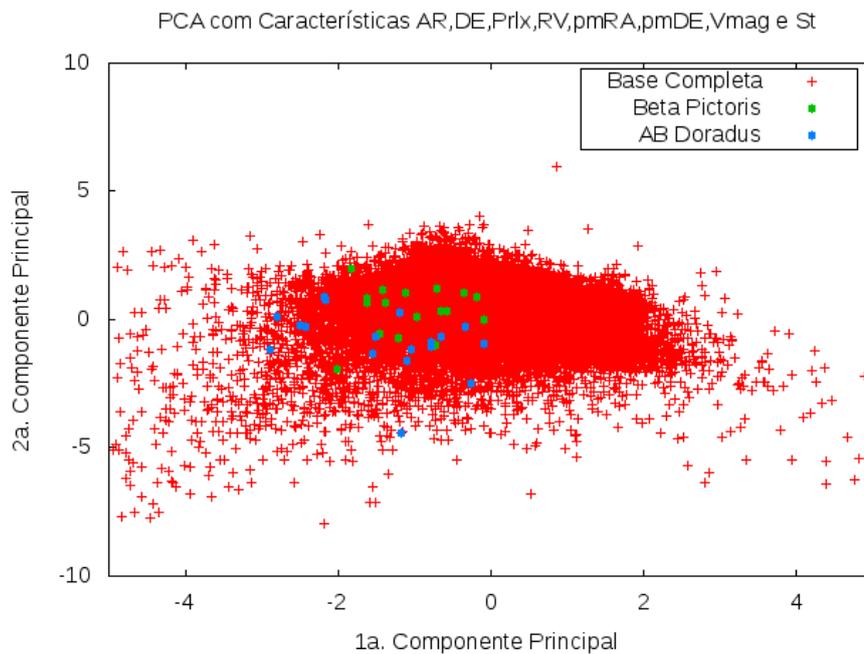


Figura 5.1: Representação Gráfica PCA experimento 1

Tabela 5.3: Características das estrelas usadas no experimento 2

Característica	Tipo	Denominação	Gráficos PCA
coordenadas Galácticas cartesianas	posição	X, Y, Z	x,y,z
Magnitude	fotométrica	Vmag	Vmag
Tipo Espectral	espectroscópica	SpType	St

Tabela 5.4: Variâncias Explicadas para PCA experimento 2

Número de Ordem da Componente Principal	Variância Explicada (%)
1 ^a	25
2 ^a	20
3 ^a	20
4 ^a	19
5 ^a	16

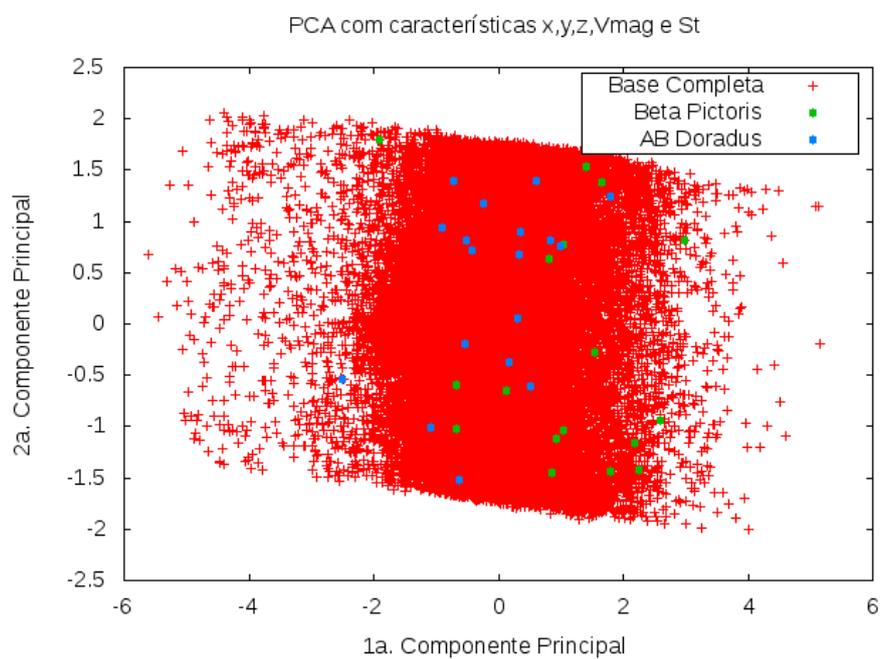


Figura 5.2: Representação Gráfica PCA experimento 2

Experimento 3: PCA com as características posições e velocidades cartesianas Galácticas, magnitude e tipo espectral

As características usadas no experimento 3 estão elencadas na tabela 5.5. Na figura 5.3 as estrelas da Base Completa e das bases Beta Pictoris e AB Doradus são plotadas no espaço das duas primeiras componentes principais. As variâncias explicadas percentuais de cada componente principal são apresentadas na tabela 5.6

Tabela 5.5: Características das estrelas usadas no experimento 3

Característica	Tipo	Denominação	Gráficos PCA
coordenadas Galácticas cartesianas	posição	X, Y, Z	x,y,z
velocidades Galáctias	velocidade	U,V,W	u,v,w
Magnitude	fotométrica	Vmag	Vmag
Tipo Espectral	espectroscópica	SpType	St

Tabela 5.6: Variâncias Explicadas para PCA experimento 3

Número de Ordem da Componente Principal	Variância Explicada (%)
1 ^a	16
2 ^a	15
3 ^a	13
4 ^a	13
5 ^a	12
6 ^a	12
7 ^a	10
8 ^a	9

Experimento 4: PCA com as características velocidades cartesianas Galácticas

As características usadas no experimento 4 estão elencadas na tabela 5.7. Na figura 5.4 as estrelas da Base Completa e das bases Beta Pictoris e AB Doradus são plotadas no espaço das duas primeiras componentes principais. As variâncias explicadas percentuais de cada componente principal são apresentadas na tabela 5.8

Tabela 5.7: Características das estrelas usadas no experimento 4

Característica	Tipo	Denominação	Gráficos PCA
velocidades Galáctias	velocidade	U,V,W	u,v,w

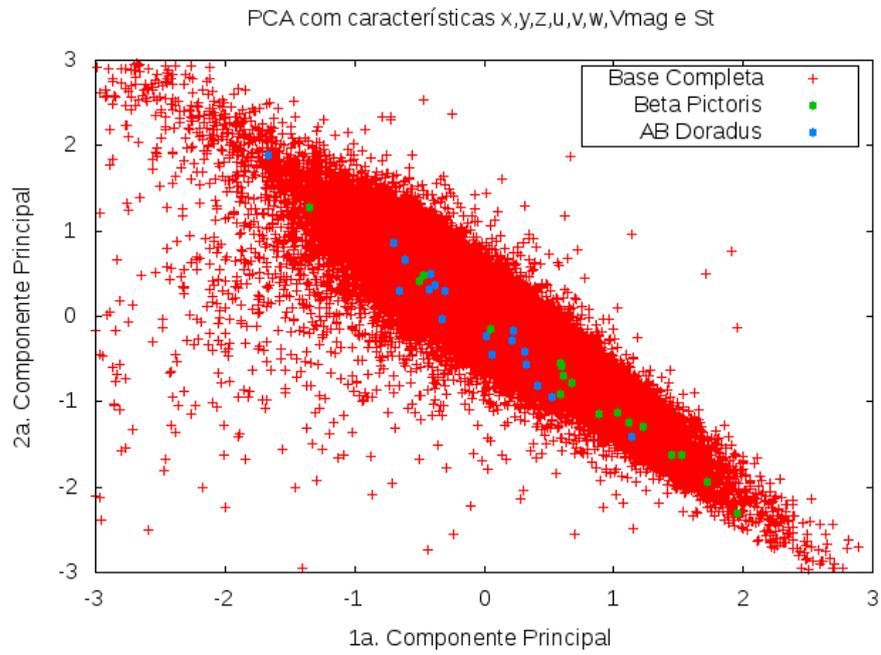


Figura 5.3: Representação Gráfica PCA experimento 3

Tabela 5.8: Variâncias Explicadas para PCA experimento 4

Número de Ordem da Componente Principal	Variância Explicada (%)
1 ^a	42
2 ^a	34
3 ^a	24

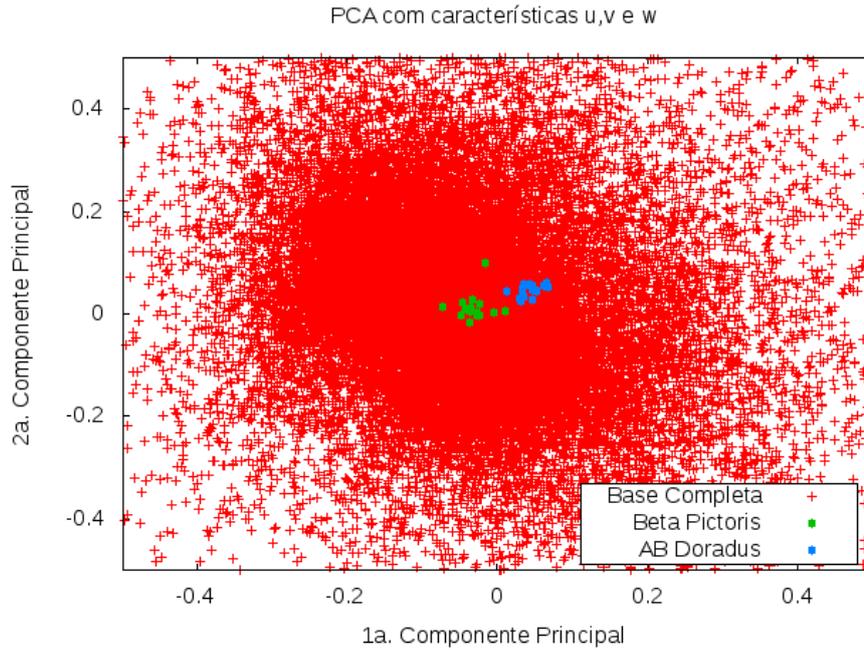


Figura 5.4: Representação Gráfica PCA experimento 4

Experimento 5: PCA com as características velocidades cartesianas Galácticas e magnitude

As características usadas no experimento 5 estão elencadas na tabela 5.9. Na figura 5.5 as estrelas da Base Completa e das bases Beta Pictoris e AB Doradus são plotadas no espaço das duas primeiras componentes principais. As variâncias explicadas percentuais de cada componente principal são apresentadas na tabela 5.10

Tabela 5.9: Características das estrelas usadas no experimento 5

Característica	Tipo	Denominação	Gráficos PCA
velocidades Galáctias	velocidade	U,V,W	u,v,w
Magnitude	fotométrica	Vmag	Vmag

Tabela 5.10: Variâncias Explicadas para PCA experimento 5

Número de Ordem da Componente Principal	Variância Explicada (%)
1 ^a	32
2 ^a	25
3 ^a	25
4 ^a	18

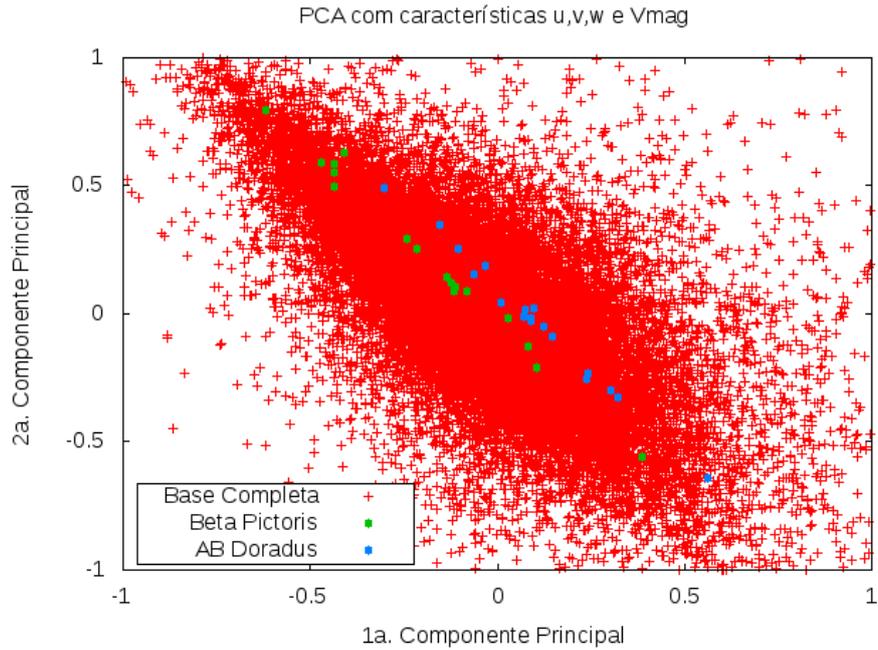


Figura 5.5: Representação Gráfica PCA experimento 5

Experimento 6: PCA com as características velocidades cartesianas Galácticas e tipo espectral

As características usadas no experimento 6 estão elencadas na tabela 5.11. Na figura 5.6 as estrelas da Base Completa e das bases Beta Pictoris e AB Doradus são plotadas no espaço das duas primeiras componentes principais. As variâncias explicadas percentuais de cada componente principal são apresentadas na tabela 5.12

Tabela 5.11: Características das estrelas usadas no experimento 6

Característica	Tipo	Denominação	Gráficos PCA
velocidades Galáctias	velocidade	U,V,W	u,v,w
Tipo Espectral	espectroscópica	SpType	St

Tabela 5.12: Variâncias Explicadas para PCA experimento 6

Número de Ordem da Componente Principal	Variância Explicada (%)
1 ^a	31
2 ^a	25
3 ^a	25
4 ^a	19

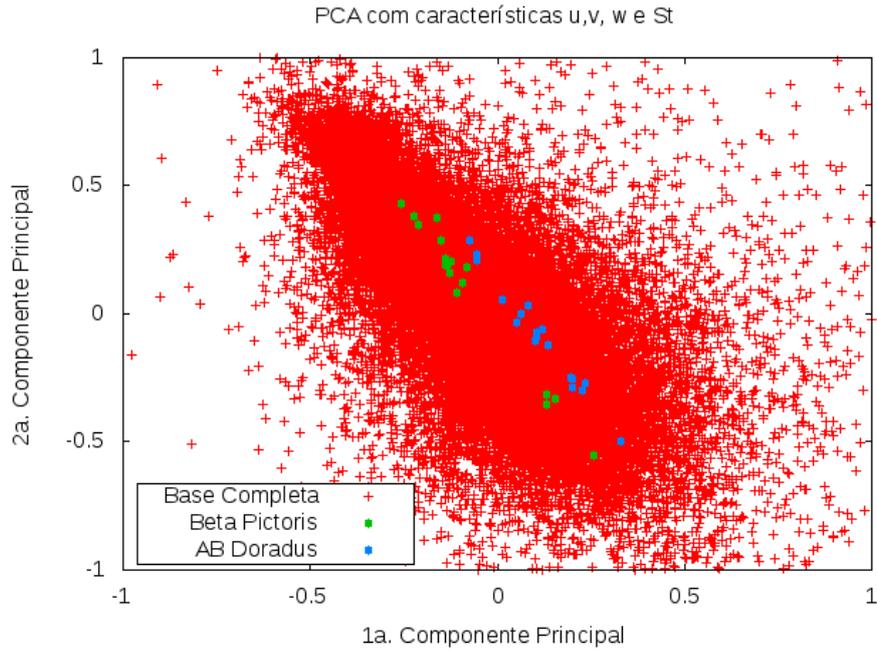


Figura 5.6: Representação Gráfica PCA experimento 6

Experimento 7: PCA com as características velocidades cartesianas Galácticas, magnitude e tipo espectral

As características usadas no experimento 7 estão elencadas na tabela 5.13. Na figura 5.7 as estrelas da Base Completa e das bases Beta Pictoris e AB Doradus são plotadas no espaço das duas primeiras componentes principais. As variâncias explicadas percentuais de cada componente principal são apresentadas na tabela 5.14

Tabela 5.13: Características das estrelas usadas no experimento 7

Característica	Tipo	Denominação	Gráficos PCA
velocidades Galáctias	velocidade	U,V,W	u,v,w
Magnitude	fotométrica	Vmag	Vmag
Tipo Espectral	espectroscópica	SpType	St

Tabela 5.14: Variâncias Explicadas para PCA experimento 7

Número de Ordem da Componente Principal	Variância Explicada (%)
1 ^a	26
2 ^a	23
3 ^a	20
4 ^a	16
5 ^a	15

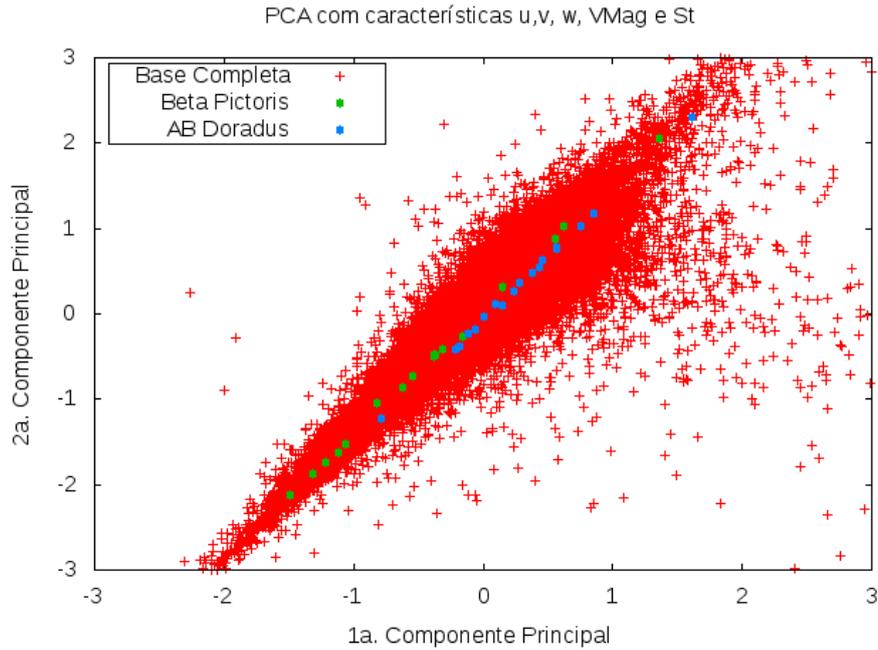


Figura 5.7: Representação Gráfica PCA experimento 7

5.1.2 Discussão dos resultados com PCA

Na seção 5.1.1 foram apresentados os gráficos e as tabelas com as variâncias explicadas correspondentes da PCA com os mais significativos conjuntos de características, embora vários outros tenham sido analisados. Para todos os conjuntos analisados observou-se o comportamento geral conforme descrito abaixo:

1. As tabelas com as variâncias explicadas mostram que:
 - as variâncias são distribuídas pelas componentes principais de maneira quase uniforme em todos os experimentos com características somente de posição
 - as variâncias são distribuídas pelas componentes principais de maneira quase uniforme em todos os experimentos com características de posição e de velocidade
 - as variâncias das primeiras componentes principais são mais representativas nos experimentos com características de velocidades e sem características de posição
 - as variâncias das primeiras componentes principais são mais significativas para as características do sistema Galáctico em relação às do equatorial
 - as variâncias das duas primeiras componentes são as mais significativas quando as características são as velocidades Galácticas

2. Os gráficos das bases de dados no espaço das duas primeiras componentes principais mostram que:

- os pontos de cada um dos dois grupos moventes tendem a agrupar-se quando dentre as características consideradas estão as velocidades Galácticas e não estão as posições
- os pontos de cada um dos dois grupos moventes tendem a agrupar-se cada um sobre uma reta quando as características são as velocidades Galácticas e pelo menos uma característica fotométrica ou espectroscópica

5.2 Resultados com Regressão Linear Simples

Como foi visto na seção 5.1.2 os membros dos grupos moventes tendem a agrupar-se sobre uma reta, diferente para cada grupo, quando são representados no espaço das duas primeiras componentes principais obtidas com PCA usando as componentes das velocidades Galácticas e pelo menos uma característica fotométrica ou espectroscópica.

Na subseção 5.2.1 serão mostrados os ajustes às retas, na 5.2.2 é feita a discussão sobre tais ajustes obtidos e na 5.2.3 é feita a escolha do ajuste que vai originar a seleção das estrelas para a evolução retrógrada.

5.2.1 Ajustes com Regressão Linear Simples

Nesta seção serão usados os resultados da seção 5.1 para os casos em que no espaço das duas primeiras componentes principais os elementos de cada um dos grupos moventes de estrelas, Beta Pictoris e AB Doradus, podem ser melhor ajustados a uma reta. Para ilustrar, também serão apresentados os casos em que as características consideradas são apenas as velocidades Galácticas. Nesse caso, observa-se que os elementos de cada um dos dois grupos tendem a agrupar-se no espaço das duas primeiras componentes, mas não estão sobre uma reta.

Os gráficos com os pontos de cada um dos grupos em movimento, a reta ajustada e o coeficiente de determinação do ajuste são mostrados nas figuras 5.8, 5.9, 5.10, 5.11, 5.12, 5.13, 5.14 e 5.15.

Algumas outras regressões podem ser feitas ainda no espaço de outros conjuntos de componentes principais, tendo em vista a semelhança das variâncias explicadas das primeiras componentes principais. Nesse ponto o fluxo será seguido e poder-se-á retornar para novos ajustes e conseqüentemente à seleção de novas estrelas.

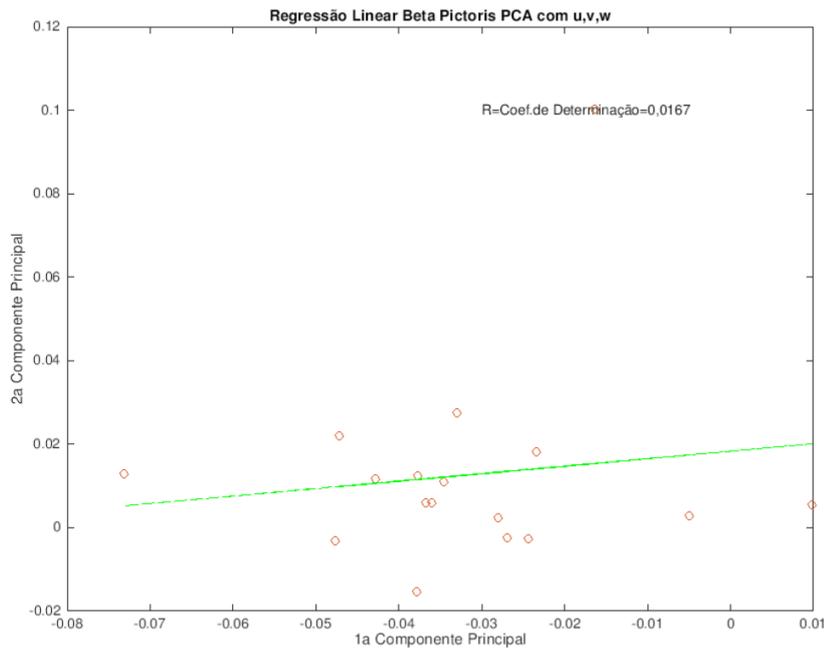


Figura 5.8: Ajuste da Retas para Beta Pictoris com u,v,w

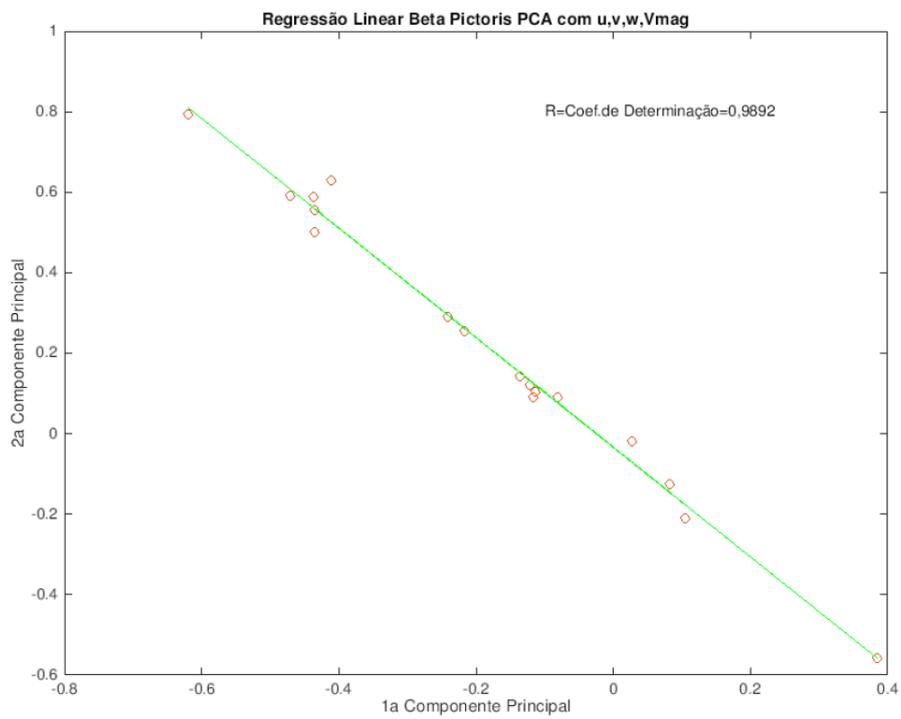


Figura 5.9: Ajuste da Retas para Beta Pictoris com u,v,w,Vmag

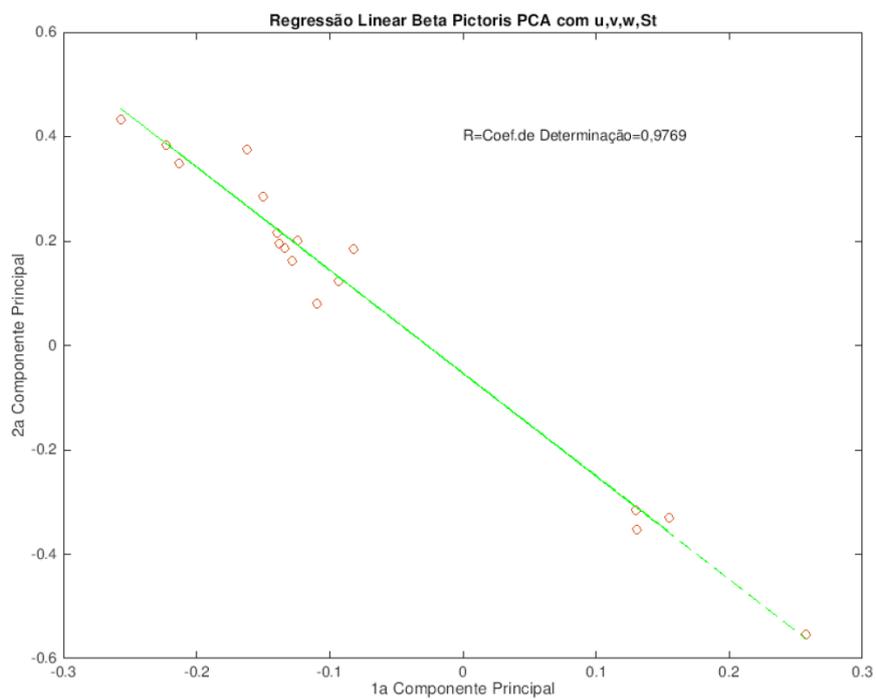


Figura 5.10: Ajuste da Reta para Beta Pictoris com u,v,w,St

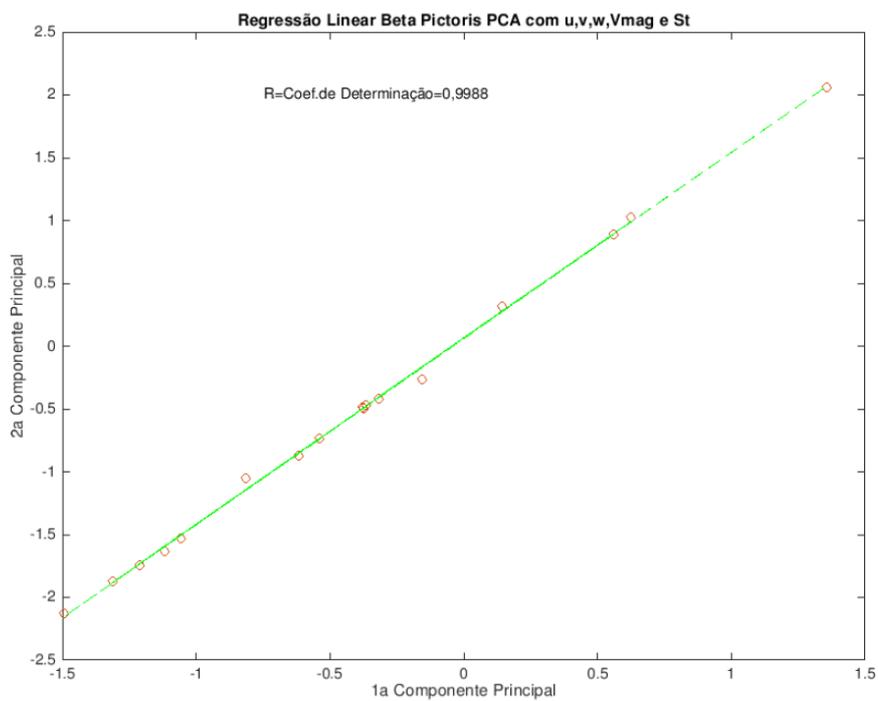


Figura 5.11: Ajuste da Reta para Beta Pictoris com u,v,w,Vmag,St

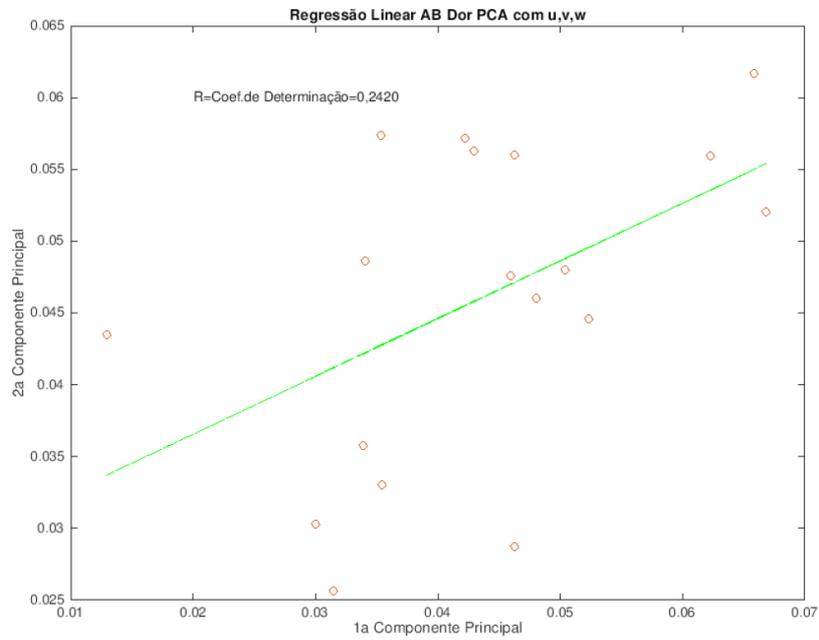


Figura 5.12: Ajuste da Reta para AB Doradus com u,v,w

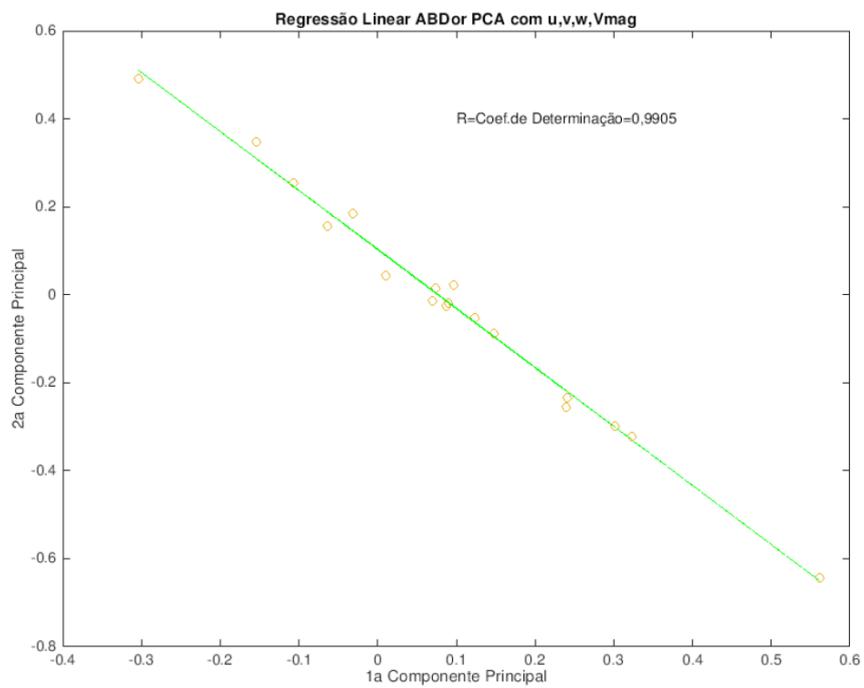


Figura 5.13: Ajuste da Reta para AB Doradus com u,v,w,Vmag

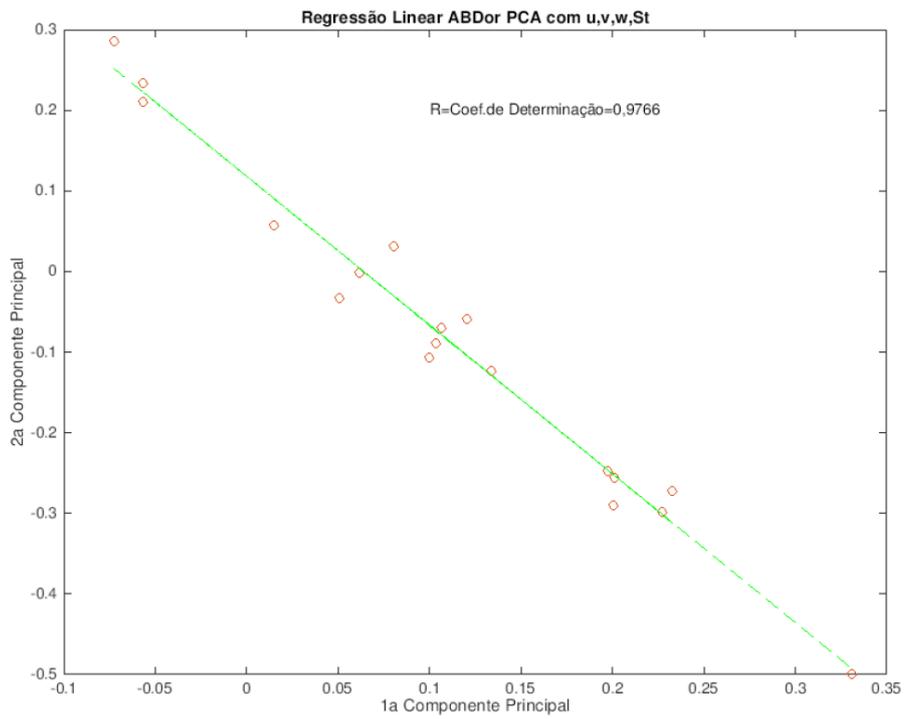


Figura 5.14: Ajuste da Reta para AB Doradus com u,v,w,St

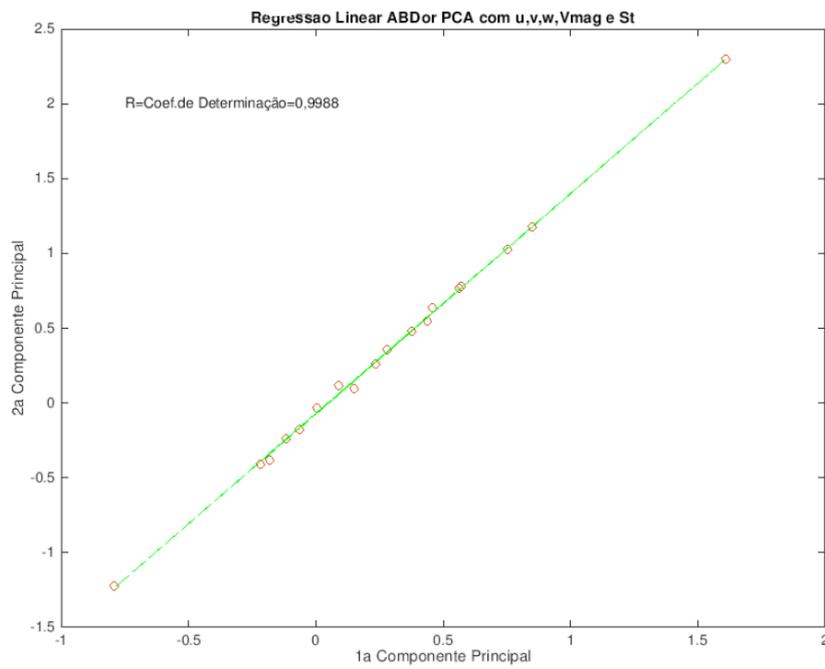


Figura 5.15: Ajuste da Reta para AB Doradus com u,v,w,Vmag,St

5.2.2 Discussão dos resultados com regressão linear simples

Foram analisados os gráficos das figuras 5.8, 5.9, 5.10, 5.11, 5.12, 5.13, 5.14 e 5.15.

Observou-se que as estrelas dos dois grupos moventes estudados, Beta Pictoris e AB Doradus, estão praticamente sobre uma reta - cada um em uma reta diferente - no espaço da 1ª e da 2ª componentes principais quando a PCA é realizada com as seguintes características: as três componentes cartesianas Galácticas da velocidade, a magnitude e o tipo espectral. Para os demais conjuntos analisados, são feitos também bons ajustes de retas (exceto para somente u,v,w), entretanto, é bem interessante o fato do coeficiente de determinação ser muito próximo de 1 e exatamente igual para os dois grupos moventes.

Como a análise das componentes principais é uma técnica de reconhecimento de padrões, é muito significativo que as estrelas dos grupos moventes tendam a agrupar-se sobre uma reta, quando são representadas no espaço das duas primeiras componentes principais, sendo essas componentes obtidas da PCA com as características: componentes cartesianas Galácticas da velocidade, magnitude e tipo espectral.

Assim, é provável que também estejam sobre a mesma reta ou muito próximas da reta outras estrelas que também pertençam ao grupo movente. A seleção de tais estrelas é apresentada na seção 5.2.3.

5.2.3 Seleção de estrelas para a evolução retrógrada

Tendo por base a discussão dos resultados com a Regressão Linear Simples será feita a seleção das estrelas para a evolução retrógrada baseada na reta da figura 5.11. Como foi especificado na seção 5.2.1, pode-se voltar nesse ponto e fazer a seleção tendo em vista outras retas obtidas da representação no espaço de outros pares de componentes principais.

Foi feita a seleção das estrelas usando-se o seguinte critério:

sejam d_i , as distâncias de cada estrela de Beta Pictoris à reta da fig 5.11

seja D_{max} o valor máximo entre todos os valores de d_i

Seleciona-se todas as estrelas que estão a uma distância da reta menor ou igual a $D_{max} + 0,05 \cdot D_{max}$

A seleção está representada em rosa na figura 5.16. O número de estrelas selecionadas foi 11899. Somam-se a essas as 247 estrelas sem especificação de tipo espectral que, por isso, não constaram da análise das componentes principais.

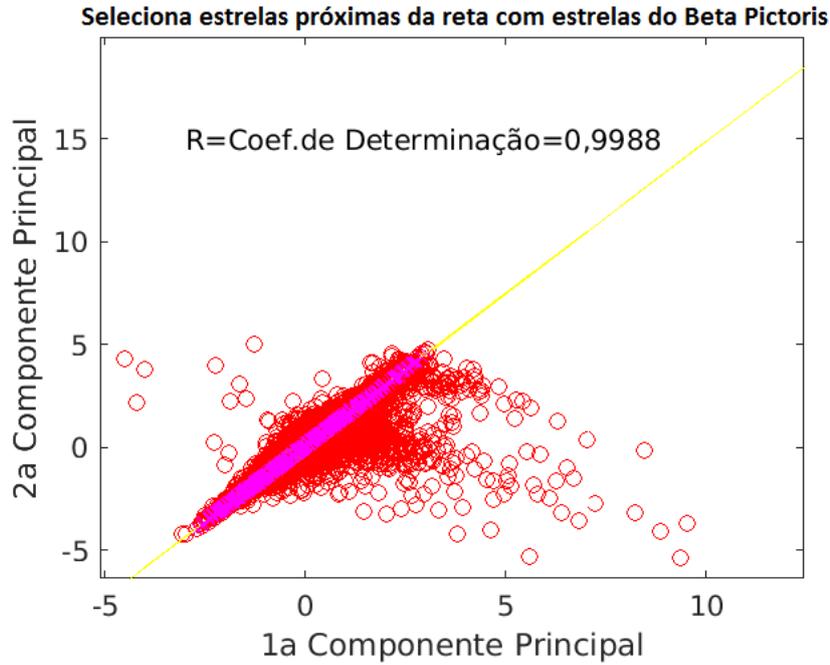


Figura 5.16: Seleção de Estrelas para Beta Pictoris

5.3 Resultados da evolução retrógrada

Nesse passo é realizada a evolução retrógrada com todas as estrelas selecionadas. Para os tempos próximos à época de nascimento do Grupo Beta Pictoris foram feitas imagens do local de nascimento de Beta Pictoris para cada passo da evolução. Algumas estrelas já conhecidas do Grupo Beta Pictoris começam a aparecer, mas surgem também outras estrelas. O passo da evolução é 0,04 Myr.

A figura 5.17 mostra as estrelas do grupo movente Beta Pictoris, segundo Ortega 2.1, que estão em seu local de nascimento no tempo do nascimento, conforme já calculado anteriormente e igual a $-11,22$ Myr. A figura mostra também que há outras seis estrelas. Essas outras estrelas são as novas candidatas para fazer parte do grupo. As identificações pelo número do catálogo Hipparcos são os seguintes: 27305, 3014, 19121, 29154, 34043 e 24394.

Tabela 5.15: Estrelas sugeridas para o Grupo Movente Beta Pictoris segundo Ortega

Identificação Hipparcos
27305
3014
19121
29154
34043
24394

Na seção 5.4 são apresentados os resultados dos testes usando os critérios apre-

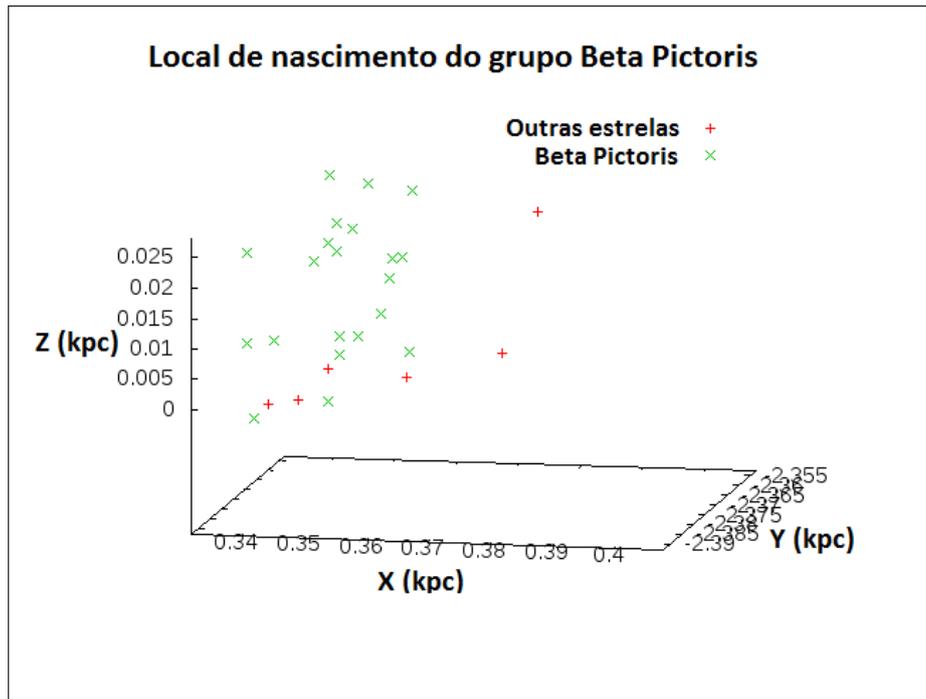


Figura 5.17: Local do nascimento de Beta Pictoris na época do nascimento

sentados na subseção 4.5.3 que irão indicar as modificações ao grupo Beta Pictoris segundo Ortega 2.1.

5.4 Resultados dos testes com critérios

Para a realização dos testes usando os critérios apresentados na subseção 4.5.3 faz-se nova evolução retrógrada, mas agora somente com os membros do grupo movente Beta Pictoris segundo Ortega 2.1 e com as estrelas sugeridas na seção 5.3.

Os testes são realizados juntos no mesmo passo porque durante a evolução retrógrada são calculados, para cada passo da evolução:

- as posições cartesianas Galácticas, que vão permitir a análise do comportamento das órbitas e, assim, a realização do “teste das órbitas”
- o “raio do grupo” conforme descrito na seção 4.5.3, que vai permitir calcular ao final:
 - o confinamento máximo, conforme descrito pela equação 4.9 e a realização do teste do confinamento
 - a idade dinâmica do grupo, igual ao tempo para o confinamento máximo e, assim, a realização do teste da idade dinâmica

Para cada teste realizado serão apresentados:

- a representação gráfica das órbitas das estrelas
- o valor do confinamento máximo
- a idade dinâmica do grupo

1. Teste com Grupo Beta Pictoris segundo Ortega

A evolução retrógrada com o grupo Beta Pictoris segundo Ortega apresenta os resultados conforme a tabela 5.16 e as órbitas conforme a figura 5.18. Observa-se que a estrela de identificação Hipparcos 14361 faz uma órbita que cruza o grupo de órbitas no tempo do nascimento, mas que não a caracteriza como membro do grupo.

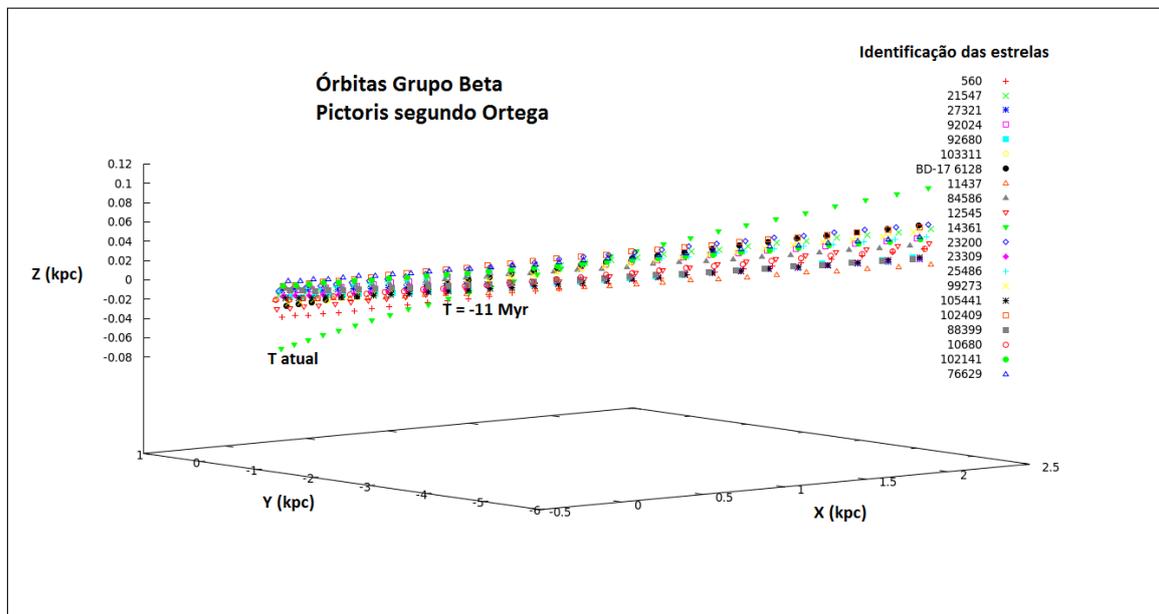


Figura 5.18: Órbitas do Grupo Beta Pictoris segundo Ortega

Tabela 5.16: Confinamento máximo e idade dinâmica Beta Pictoris segundo Ortega

confinamento máximo	idade dinâmica (Myr)
0,316	11 Myr

2. Órbitas do Grupo Beta Pictoris com 1ª modificação

A 1ª modificação é, portanto, a retirada da estrela Hipparcos 14361 do grupo Beta Pictoris. Os novos valores para confinamento máximo e idade do grupo são mostrados na tabela 5.17 e a representação gráfica das órbitas do grupo passa a ser a que é mostrada na figura 5.19.

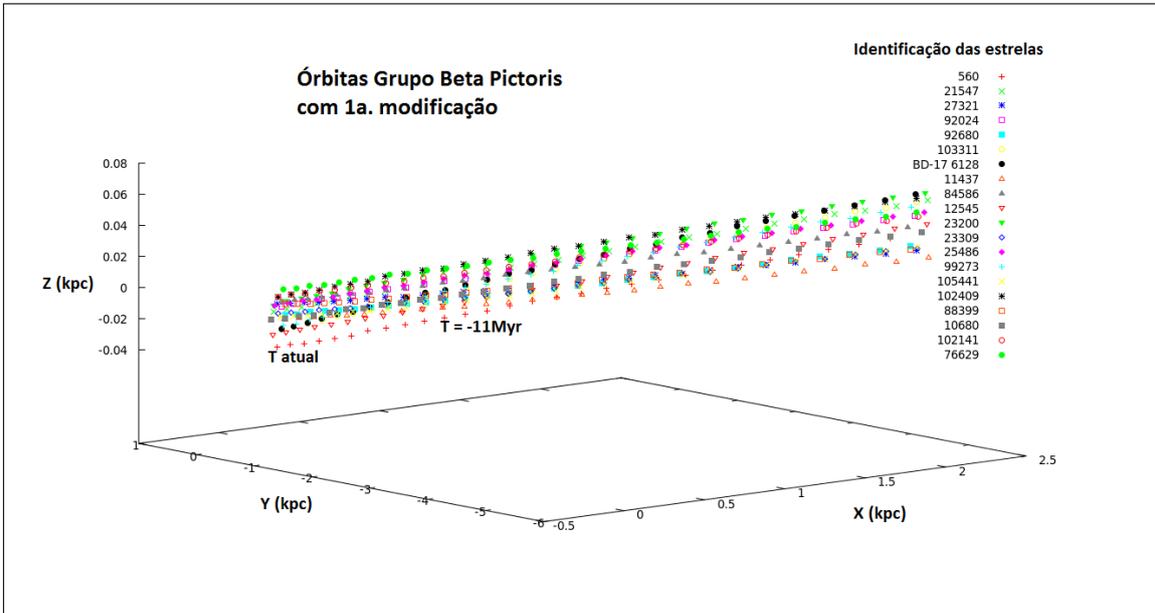


Figura 5.19: Órbitas do Grupo Beta Pictoris com 1ª modificação

Tabela 5.17: Confinamento e idade dinâmica Beta Pictoris com 1ª modificação

confinamento máximo	idade dinâmica (Myr)
0,350	11 Myr

3. Teste com Grupo Beta Pictoris modificado mais 6 estrelas sugeridas

Ao grupo Beta Pictoris com a 1ª modificação são acrescentadas as 6 estrelas sugeridas conforme a tabela 5.15. A figura 5.20 mostra as órbitas e a tabela 5.18 os valores do confinamento máximo e da idade dinâmica. A estrela de identificação Hipparcos 34043 é a indicada para ser retirada.

Tabela 5.18: Confinamento e idade dinâmica Beta Pictoris modificado mais 6 estrelas sugeridas

confinamento máximo	idade dinâmica (Myr)
0,202	10 Myr

4. Teste com Grupo Beta Pictoris modificado mais 5 estrelas sugeridas

Nova evolução retrógrada é realizada com o grupo Beta Pictoris com a 1ª modificação e são acrescentadas as 5 estrelas da tabela 5.15 exceto a de identificação Hipparcos 34043. A figura 5.21 mostra as órbitas e a tabela 5.19 os valores do confinamento máximo e da idade dinâmica. A estrela de identificação Hipparcos 3014 é a indicada para ser retirada.

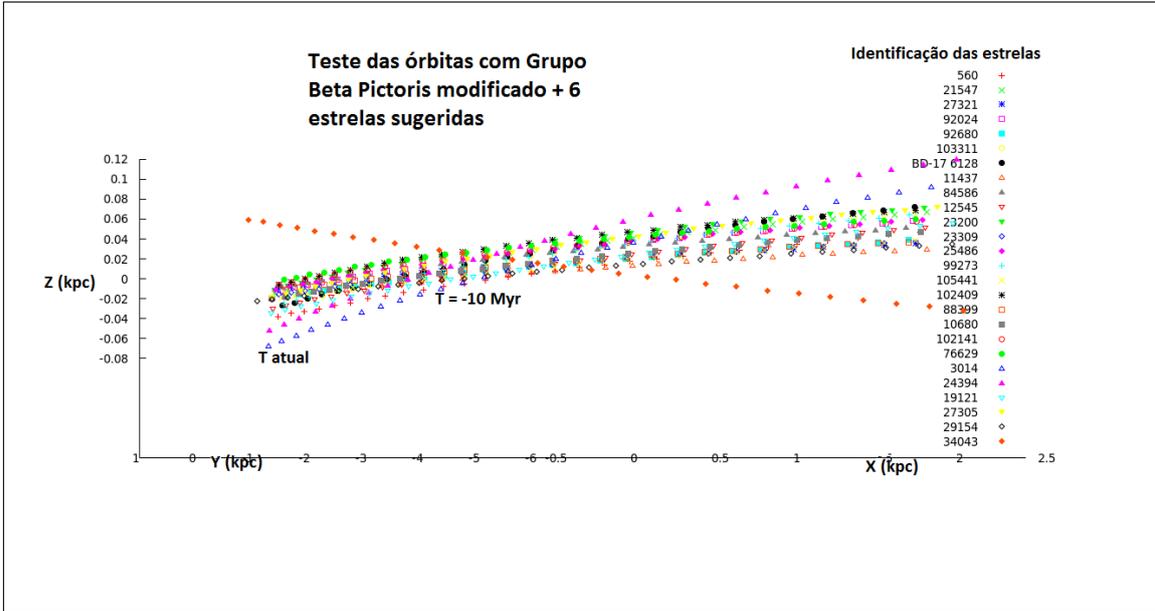


Figura 5.20: Órbitas do Grupo Beta Pictoris modificado mais 6 estrelas sugeridas

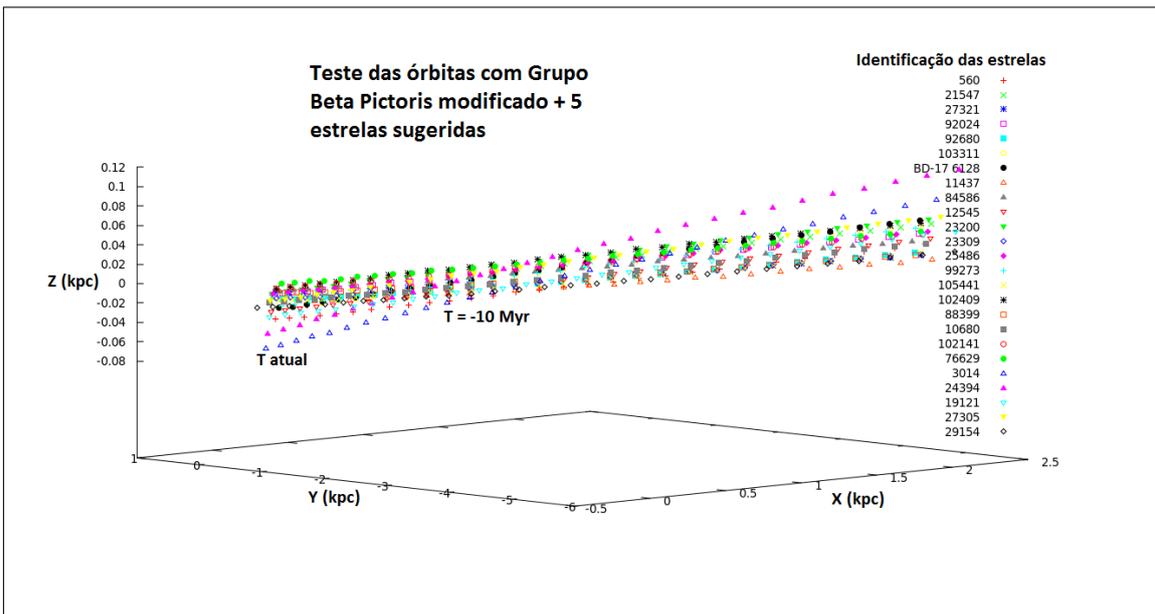


Figura 5.21: Órbitas do Grupo Beta Pictoris modificado mais 5 estrelas sugeridas

Tabela 5.19: Confinamento e idade dinâmica Beta Pictoris modificado mais 5 estrelas sugeridas

confinamento máximo	idade dinâmica (Myr)
0,226	10 Myr

5. Teste com Grupo Beta Pictoris modificado mais 4 estrelas sugeridas

Nova evolução retrógrada é realizada com o grupo Beta Pictoris com a 1ª modificação e são acrescentadas as 4 estrelas da tabela 5.15 exceto as de identificação Hipparcos 34043 e 3014. A figura 5.22 mostra as órbitas e a tabela 5.20 os valores do confinamento máximo e da idade dinâmica. A estrela de identificação Hipparcos 24394 é a indicada para ser retirada.

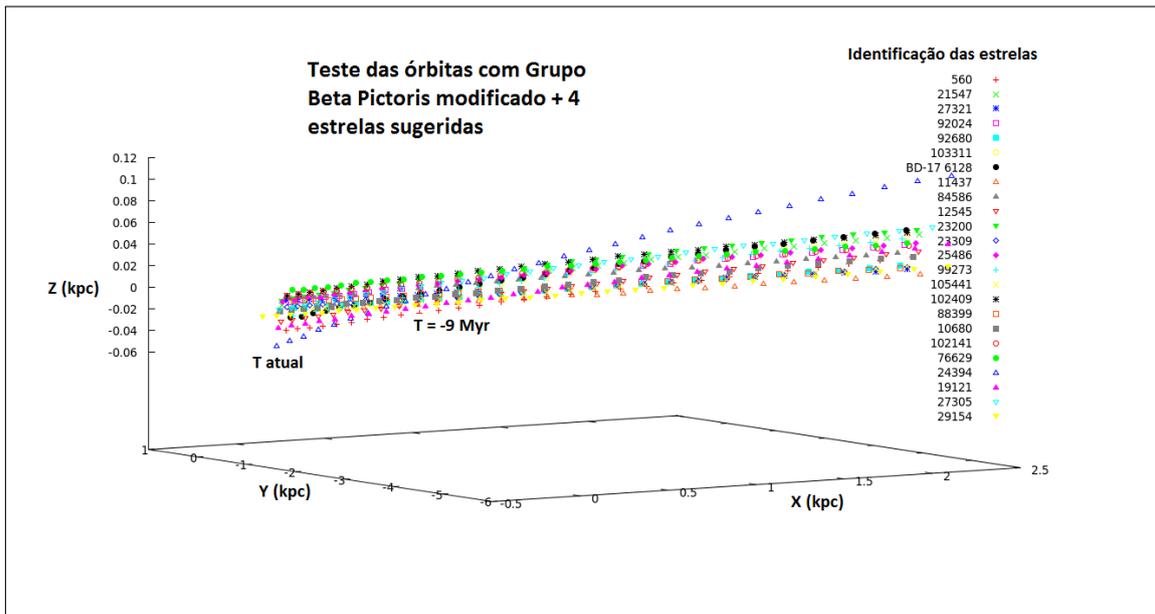


Figura 5.22: Órbitas do Grupo Beta Pictoris modificado mais 4 estrelas sugeridas

Tabela 5.20: Confinamento e idade dinâmica Beta Pictoris modificado mais 4 estrelas sugeridas

confinamento máximo	idade dinâmica (Myr)
0,221	10 Myr

6. Teste com Grupo Beta Pictoris modificado mais 3 estrelas sugeridas

Nova evolução retrógrada é realizada com o grupo Beta Pictoris com a 1ª modificação e são acrescentadas as 3 estrelas da tabela 5.15 exceto as de identificação Hipparcos 34043, 3014 e 24394. A figura 5.23 mostra as órbitas e a tabela 5.21 os valores do confinamento máximo e da idade dinâmica. A estrela de identificação Hipparcos 19121 é a indicada para ser retirada.

7. Órbitas do Grupo Beta Pictoris com as 3 modificações

Com a nova evolução retrógrada que é realizada com grupo Beta Pictoris com a 1ª modificação e mais as estrelas de identificação Hipparcos 27305 e 29154

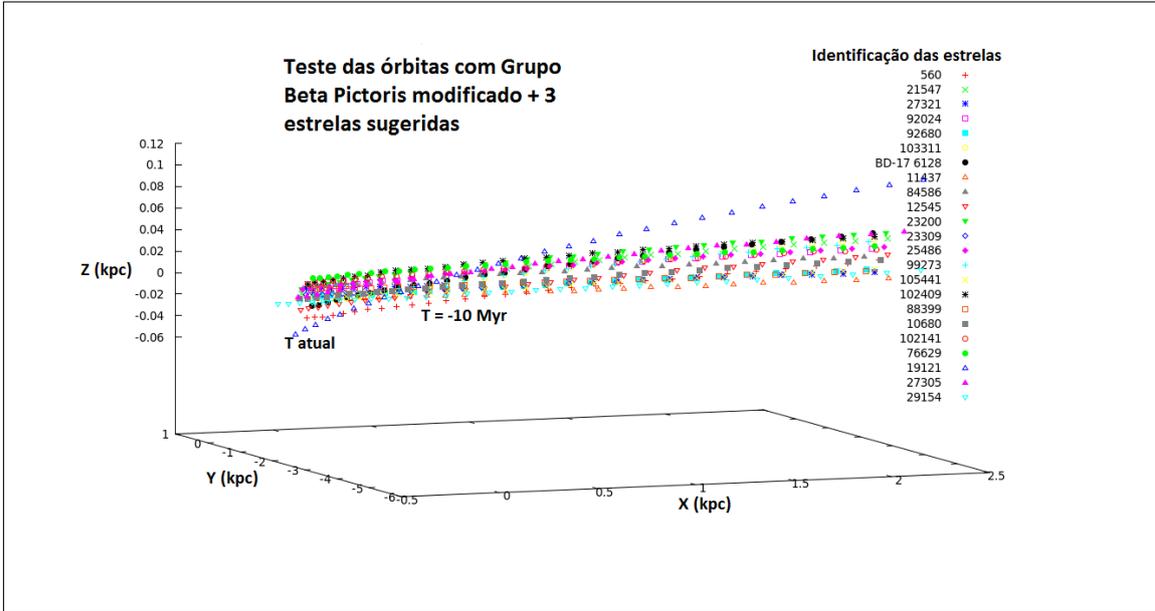


Figura 5.23: Órbitas do Grupo Beta Pictoris modificado mais 3 estrelas sugeridas

Tabela 5.21: Confinamento e idade dinâmica Beta Pictoris modificado mais 3 estrelas sugeridas

confinamento máximo	idade dinâmica (Myr)
0,224	10 Myr

observa-se pela figura 5.24 que o teste das órbitas é positivo. Pelos valores mostrados na tabela 5.22 observa-se que a idade dinâmica é a mesma do Grupo Beta Pictoris segundo Ortega, conforme tabela 5.16 e que o confinamento máximo é ainda menor, o que é positivo também.

Assim, inserção das estrelas de identificação Hipparcos 27305 e 29154 marcam a 2ª e a 3ª modificações sugeridas ao grupo Beta Pictoris segundo Ortega.

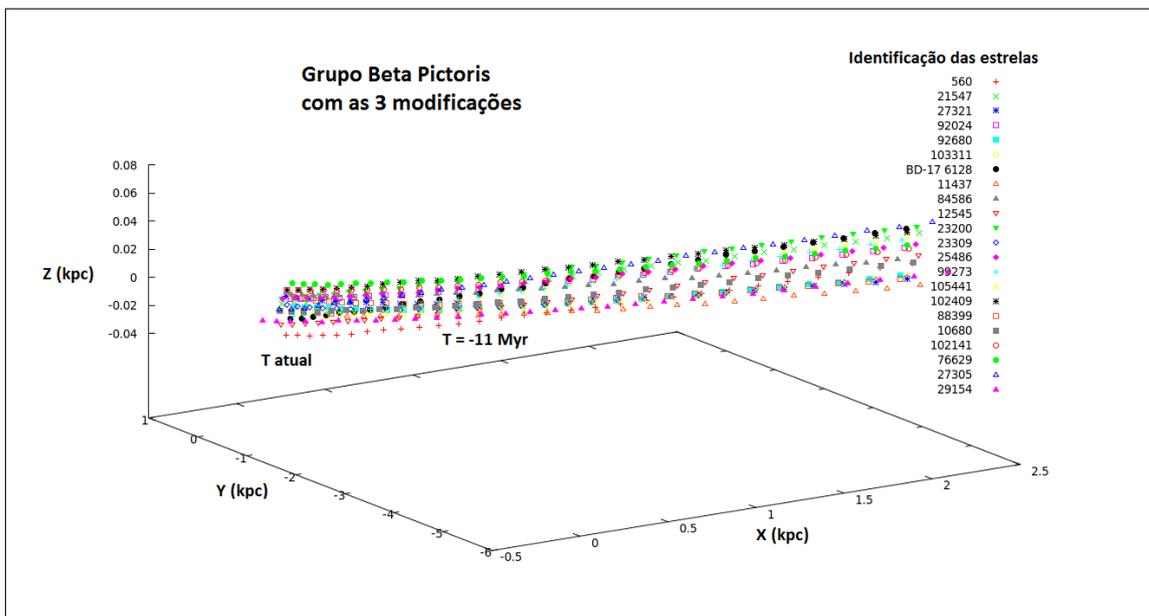


Figura 5.24: Órbitas do Grupo Beta Pictoris com as 3 modificações confirmadas

Tabela 5.22: Confinamento e idade dinâmica Beta Pictoris com as 3 modificações confirmadas

confinamento máximo	idade dinâmica (Myr)
0,227	11 Myr

5.5 Modificações sugeridas ao grupo estelar em movimento Beta Pictoris segundo Ortega

Como foi visto na seção 5.4 a metodologia indicou 3 modificações ao grupo em movimento Beta Pictoris segundo Ortega:

1. a retirada da estrela de identificação Hipparcos 14361
2. a inclusão da estrela de identificação Hipparcos 27305
3. a inclusão da estrela de identificação Hipparcos 29154

O grupo Beta Pictoris foi estudado por vários autores e, como foi visto na seção 2.3.3, os métodos e critérios diferem de autor para autor, ocasionando diferença dos membros dos grupos de acordo com o autor. A estrela 14361 não faz parte do grupo Beta Pictoris segundo outro autor.

As estrelas 27305 e 29154 são estrelas de tipo espectral, respectivamente, F8V e A2IV e não foram detectadas como membros do grupo movente Beta Pictoris por nenhum dos autores estudados.

Capítulo 6

Conclusão e Trabalhos Futuros

Neste trabalho foi apresentada uma nova metodologia computacional para analisar concomitantemente dezenas de milhares de estrelas com o objetivo de encontrar novos membros dos grupos em movimento já conhecidos. A metodologia apresentada é a principal contribuição deste trabalho e está baseada na ciência de dados sendo reprodutível para uma quantidade de dados de ordem de grandeza ainda maior, tanto no que se refere ao número de estrelas quanto ao número de características de cada estrela, sendo a principal contribuição deste trabalho. Seu *workflow* é apresentado na figura 4.7.

Seguindo as fases de obtenção e limpeza da ciência de dados, foi montada a base de dados completa e dela extraídas as estrelas dos grupos em movimento Beta Pictoris segundo Ortega, conforme a tabela 4.5 , e AB Doradus segundo Zuckerman conforme tabela 4.6. A produção das bases de dados constitui-se a primeira contribuição deste trabalho e o *workflow* dessa fase é mostrado na figura 4.1.

Os dados astrométricos das estrelas foram obtidos de um dos catálogos resultantes da 1^a missão espacial astrométrica: a Missão Hipparcos, cujo satélite recolheu dados entre os anos 1990 e 1993. O catálogo utilizado foi o publicado em 1997 e chamado Hipparcos 1997. Tal catálogo fornece as posições de cerca de 118 mil estrelas com precisão de milissegundos de arco. Além das posições fornece também os movimentos próprios e paralaxe das estrelas com precisão de milissegundos de arco, precisão essa nunca anteriormente conseguida. Para estudar o movimento das estrelas, como parte da metodologia apresentada, faz-se necessário, entretanto, um dado que não foi observado na missão Hipparcos e que não está, portanto, no catálogo utilizado: a velocidade radial. Após analisar os catálogos de velocidade radial disponíveis decidiu-se obter as velocidades radiais dos Catálogos de Pulkovo e SACY, conseguindo-se montar uma base de dados com cerca de 35 mil estrelas que foram, então, acopladas aos dados dos catálogos Hipparcos 1997.

Na fase de exploração dos dados, foram analisados os gráficos de dispersão das três bases nos espaços das componentes cartesianas das velocidades Galácticas e

das componentes cartesianas das posições Galácticas, que são usados em alguns dos trabalhos correlatos mais recentes conforme descrito na seção 2.3.3.

Na fase de modelagem foram realizados vários experimentos usando-se a Análise das Componentes Principais (PCA) e Regressão Linear Simples para diferentes conjuntos de características das estrelas. Nessa fase percebeu-se que as estrelas dos grupos moventes são muito bem ajustadas a uma reta, sendo uma reta diferente para cada grupo, quando são representadas no espaço das duas primeiras componentes principais da PCA realizada com as componentes das velocidades Galácticas, com a magnitude e com o tipo espectral. Essas retas trazem consigo a ideia que outros novos membros daqueles grupos estejam em uma região muito próxima.

Com base naquela ideia foi realizada, na fase de interpretação dos dados, a seleção de estrelas que estão próximas à reta obtida para o Grupo Beta Pictoris. Usando-se a abordagem dinâmica de Ortega *at al.*, descrita na seção 2.4, foi feita a evolução retrógrada com as estrelas selecionadas. Foi observado o local do nascimento no tempo do nascimento do grupo Beta Pictoris segundo Ortega, conforme tabela 2.1 e verificado que há seis estrelas além das estrelas já conhecidas do grupo que são, então, as novas candidatas do grupo Beta Pictoris. Essas estrelas são submetidas aos testes com critérios para serem confirmadas como membros do grupo. O primeiro teste consiste na análise das órbitas de todas as estrelas do grupo com as novas candidatas, que foi desenvolvido neste trabalho e constitui-se em sua terceira contribuição. Os outros dois testes analisam a nova idade dinâmica do grupo e o novo valor do confinamento máximo, que não devem ser muito diferentes dos valores já conhecidos.

Assim, aplicando-se a metodologia ao grupo movente Beta Pictoris, segundo Ortega, foram indicadas 3 modificações: a exclusão da estrela de identificação Hipparcos 14361 e a inclusão das estrelas de identificação Hipparcos 27305 e 29154.

Essa metodologia foi aplicada nesse trabalho para o grupo conhecido Beta Pictoris segundo Ortega. Para ser aplicada a outro grupo de estrelas já conhecido e sendo utilizada a mesma base de dados, é necessário somente que se especifique as estrelas que compõem o grupo conhecido e seguir o mesmo *workflow*. Para se compor uma nova base de dados é preciso refazer os passos de obtenção, limpeza e tratamento de dados descrito no *workflow* de produção das bases de dados como é mostrado na figura 4.1.

Trabalhos Futuros

A sequência natural desse trabalho consiste em primeiro lugar na aplicação da metodologia para o mesmo Beta Pictoris, mas fazendo-se alterações de parâmetros e em segundo lugar para outros grupos conhecidos, iniciando-se pelo grupo AB Doradus, usando a mesma base de dados completa.

Como foi visto na seção 2.3.3 outros autores levando em consideração outros catálogos, métodos e critérios encontraram novos membros para o grupo Beta Pictoris. Assim, o próximo passo será montar novas bases de dados com os catálogos usados por esses autores, como por exemplo o 2 MASS, e aplicar a metodologia.

Recentemente, em outubro de 2018, foram publicados os resultados da missão Gaia, com o nome DR2 *data release 2*. O DR2 publicou dados astrométricos, fotométricos e espectroscópicos para cerca de 1 bilhão de corpos celestes. Foram publicados para 7 milhões de estrelas os dados cinemáticos completos, tendo sido as velocidades radiais medidas pela própria missão. Espera-se para o segundo semestre de 2020 o terceiro *release* com os dados revistos, já denominado “Gaia EDR3 catalogue”. Pretende-se utilizar a metodologia aqui desenvolvida montando uma nova base de dados com os dados de Gaia DR2 e posteriormente com aqueles do Gaia EDR3.

Referências Bibliográficas

- [1] KEPLER, S. O., SARAIVA, M. F. O. *Astronomia e Astrofísica*. 3 ed. São Paulo, Editora Livraria da Física, 2013. ISBN: 978-85-7861-187-3.
- [2] L.SPAPKE, J.GALLAGHER. *Galaxies in the Universe: An Introduction*. 2 ed. Cambridge, Cambridge University Press, 2007.
- [3] BINNEY, J., TREMAINE, S. *Galactic Dynamics*. 2 ed. New Jersey, Princeton University Press, 2008.
- [4] BOHM-VITENSE, E. *Introduction to Stellar Astrophysics*. 1 ed. Cambridge, Cambridge University Press, 1989. ISBN: 0521348692.
- [5] PERRYMAN, M., LINDEGREN, L., K., J., et al. “The Hipparcos Catalog”, *Astronomy and Astrophysics*,323, L49-L52, 1997.
- [6] ZUCKERMAN, B., SONG, I., BESSEL, M. “The AB Doradus Moving Group”, *The Astrophysical Journal*, 613:L65-L68, 2004.
- [7] NASA. “exoplanet exploration”. 2019. Disponível em: <<https://exoplanets.nasa.gov/>>. Acesso em 18/11/2019.
- [8] ORTEGA, V., DE LA REZA, R., JILINSKY, E., et al. “The origin of the Beta Pictoris Moving Group”, *The Astrophysical Journal*,575:L75-L78, 2002.
- [9] MIHALAS, D., BINNEY, J. *Galactic Astronomy*. 2 ed. San Francisco, W.H.Freeman and Company, 1981.
- [10] CATELLI, F., GIOVANNINI, O., HOFFMANN, P. “Um problema didático: como determinar ângulos de paralaxe trigonométrica”, *Revista Brasileira de Ensino de Física*, vol.40, n1,e1307, 2018.
- [11] HOOGERWERF, R., DE BRUIJNE J.H.J., DE ZEEUW, P. “On the Origin of the O and B-type stars with high velocities”, *Astronomy and Astrophysics*, 365, 49-77, 2001.

- [12] SHARPLESS, S. “Distribution of Associations, Emission Regions, Galactic Clusters and Supergiants”. In: *Galactic structure. Edited by Adriaan Blaauw and Maarten Schmidt. Published by the University of Chicago Press, Chicago, ILL USA, p.131*, 1965.
- [13] ZUCKERMAN, B., SONG, I. “Young Stars Near the Sun”, *Annu.Rev.Astronn.Astrophys.*, 42:685-721, 2004.
- [14] LOPEZ-SANTIAGO, J. AND MONTES, D., CRESPO-CHACON, I., FERNANDEZ-FIGUEROA, J. “The Nearest Young Moving Groups”, *The Astrophysical Journal*, 643:1160-1165, 2006 June 1, 2006.
- [15] TORRES, C., QUAST, G.R., M., C.H.F., et al. “Young Nearby Loose Association Handbook of Star Forming Regions Vol.II, Astronomical Society of the Pacific, Bo Reipurth, ed.(2008)”. In: *Handbook of Star Forming Regions Vol.II, Astronomical Society of the Pacific, Bo Reipurth*, 2008.
- [16] ORTEGA, V., DE LA REZA, R., JILINSKI, E. “Dynamical 3-D Evolution of Nearby Young Moving Groups”, *RevMexAA(Serie de Conferencias)*, 26:91-91, 2006.
- [17] SKULJAN, J. “A Study of Moving Groups of Stars in our Galaxy”. 1999.
- [18] LEEUWEN, V. “The Hipparcos Catalog”, *Astronomy and Astrophysics*, 500, 505-506, 2009.
- [19] ANDERSON, E., FRANCIS, C. “XHIP: An Extended Hipparcos Compilation”, *Astronomy Letters*, 2012, Vol.38, No.5, 331-346, 2012.
- [20] AGENCY, E. S. “Gaia Data Release Scenario”. 11 2019. Disponivel em: <<https://www.cosmos.esa.int/web/gaia/release>>. Acesso em 18/11/2019.
- [21] ZUCKERMAN, B. “The Nearby, Young, Argus Association: Membership, Age and Dusty Debris Disks”, *ASTrophysical Journal*, 870:27 (8pp), 2019.
- [22] LEE, J., SONG, I. “Evaluation of nearby young moving groups based on unsupervised machine learning”, *MNRAS* 489, 2189-2194, 2019.
- [23] ZUCKERMAN, B., SONG, I., BESSEL, M., et al. “The Beta Pictoris Moving Group”, *The Astronomical Journal*, 562:L87-L90, 2001.
- [24] MÓOR, A., SZABÓ, G., L.L.KISS, et al. “Unveiling new members in five nearby young moving groups”, *MNRAS*, 435, 1376-1388, 2013.

- [25] ELLIOTT, P., BAYO, A., MELO, C. H. F., et al. “Search for associations containing young stars (SACY) VII. New stellar and substellar candidate members in the young associations”, *Astronomy and Astrophysics*, 590, A13, 28, 2016.
- [26] R.RIEDEL, A., BLENT, S. C., LAMBRIDES, E. L., et al. “LACEwing: A New Moving Group Analysis Code”, *The Astronomical Journal*, 153:95 (36pp), 2017.
- [27] SHKOLNIK, E. L., ALLERS, K. N., KRAUS, A. L., et al. “All-sky Co-moving Recovery of Nearby Young Members (ACRONYM). II. The Beta Pictoris Moving Group”, *The Astronomical Journal*, 154:69 (23pp), 2017.
- [28] MAMAJEK, E., FEIGELSON, E. “Young Stars Near Earth: Progress and Prospects”. In: *ASP Conf.Ser.244,104*, 2001.
- [29] JILINSKI, E., ORTEGA, V., DE LA REZA, R. “On The Origin of the Very Young Groups eta and epsilon Chamaeleontis”, *The Astrophysical Journal*, 619:945-947, 2005 February 1, 2005.
- [30] ORTEGA, V., JILINSKI, E., DE LA REZA, R., et al. “On the Common Origin of the AB Doradus Moving Group and the Pleiades Cluster”, *MNRS*, 377:441-445, 2007.
- [31] JILINSKI, E., ORTEGA, V., DE LA REZA, R., et al. “Dynamical Evolution and Spectral Characteristics of the Stellar Group Mamajek 2”, *The Astrophysical Journal*, 691:212-218, 2009 January 20, 2009.
- [32] ORTEGA, V., JILINSKI, E., DE LA REZA, R., et al. “A Formation Scenario of Young Stellar Groups in the Region of the Scorpio Centaurus OB Association”, *The Astrophysical Journal*, 137:3922-3930, 2009.
- [33] PRESS, W., TEUKOLSKY, S., VETTERLING, W., et al. *Numerical Recipes in C*. 2 ed. New York, Cambridge University Press, 1992.
- [34] JANSSENS, J. *Data Science at the Command Line*. 1 ed. Sebastopol, CA, O’Reilly Media, Inc., 2015. ISBN: 978-1-491-94785-2.
- [35] MASON, H., WIGGIN, C. “A taxonomy of Data Science”. 2010. Disponivel em: <<http://www.dataists.com/2010/09/a-taxonomy-of-data-science/>>. Acesso em 18/11/2019.
- [36] STATON, J. M., SALTZ, J. S. *An Introduction to Data Science*. 3 ed. Londres, SAGE Publication, Inc., 2018.

- [37] LYRA, W., DA SILVA, E., ARAÚJO, M. T. “Classificação Periódica: um Exemplo Didático para Ensinar Análise de Componentes Principais”, *Química Nova*, Vol 33, No 7, 1594-1597, 2010.
- [38] VASCONCELOS, S. “Análise das Componentes Principais (PCA)”. Disponível em: <<http://www2.ic.uff.br/~aconci/PCA-ACP.pdf>>. Acesso em 18/11/2019.
- [39] FERRARI, L. I. “Uma Metodologia para Extração de Informação sobre o Sistema Imunológico”. 2008.
- [40] PETERNELLI, L. A. “Regressão Linear e Correlação”. 2019. Disponível em: <<http://www.dpi.ufv.br/~peterneli/inf162.www.16032004/materiais/CAPITULO9.pdf>>. Acesso em 18/11/2019.
- [41] MATLAB. “Linear Regression”. 2019. Disponível em: <https://www.mathworks.com/help/matlab/data_analysis/linear-regression.html>. Acesso em 18/11/2019.
- [42] DE STRASBOURG, U. “Centre de Données astronomiques de Strasbourg - Strasbourg astronomical Data Center”. 2019. Disponível em: <<https://cds.u-strasbg.fr/>>. Acesso em 18/11/2019.
- [43] LINDEGREN, L., DRAVINS, D. “The fundamental definition of ”radial velocity””, *Astronomy and Astrophysics* 401, 1185-1201, 2003.
- [44] GONTCHAROV, G. “Pulkovo Compilation of Radial Velocities for 35495 Hipparcos Stars in a Common System”, *Astronomy Letter*, Vol.31, No 11, pp 759-771, 2006.
- [45] TORRES, C., QUAST, G., DA SILVA, L. T. “Search for Association Containing Young Stars”, *Astronomy and Astrophysics*, 460,695, 2006.
- [46] JOHNSON, D., SODERBLOM, D. “Calculating Galactic Space Velocities and Their Uncertainties, with an Application to the Ursa Major Group”, *The Astronomical Journal*, vol 93, numer 4, 1987.