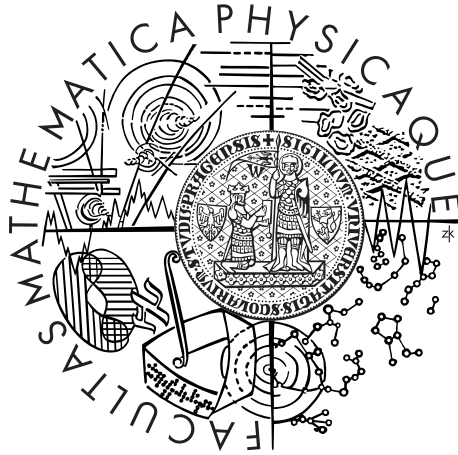


Univerzita Karlova v Praze  
Matematicko-fyzikální fakulta

## DIPLOMOVÁ PRÁCE



Václav Pavlík

## Modelování Velké mlhoviny v Orionu

Astronomický ústav Univerzity Karlovy

Vedoucí diplomové práce: doc. RNDr. Ladislav Šubr, Ph.D.

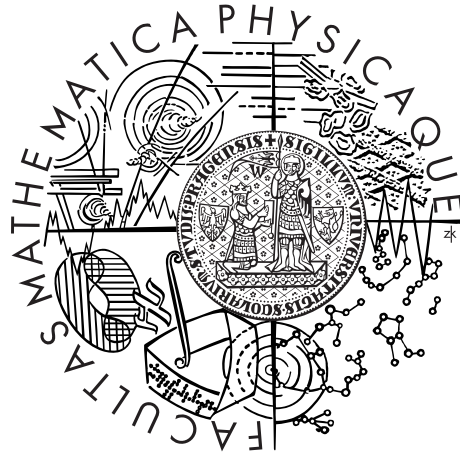
Studijní program: Fyzika

Studijní obor: Astronomie a astrofyzika

Praha 2014

Charles University in Prague  
Faculty of Mathematics and Physics

## MASTER THESIS



Václav Pavlík

# Modelling the Orion Nebula Cluster

Astronomical Institute of the Charles University

Supervisor of the master thesis: doc. RNDr. Ladislav Šubr, Ph.D.

Study programme: Physics

Specialization: Astronomy and Astrophysics

Prague 2014

Na prvním místě děkuji své rodině za poskytnuté zázemí a svému vedoucímu doc. Šubrovi za trpělivé vedení a za spoustu hodin strávených na konzultacích. Dále děkuji pánům Marku Wengerovi, João Alvesovi a Hervé Bouymu za jejich přispění při analýze dat. V neposlední řadě pak děkuji své přítelkyni Heleně, která pro mě měla za jakékoliv situace pochopení, a kamarádovi Honzovi za věcné připomínky a všechnu kávu, kterou jsem při psaní vypil na jeho účet.

First of all I thank to my family for the support provided and to my supervisor doc. Šubr for his patient leading and for many hours spent on consultations. I also thank Mark Wenger, João Alves and Hervé Bouy for the contributions in the data analysis. Last but not least, I thank to my girlfriend Helena, who had always understandings of me, and my friend Jan for his constructive notes and many cups of coffee that I drank up on his account while I was writing this thesis.

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně a výhradně s použitím citovaných pramenů, literatury a dalších odborných zdrojů.

Beru na vědomí, že se na moji práci vztahují práva a povinnosti vyplývající ze zákona č. 121/2000 Sb., autorského zákona v platném znění, zejména skutečnost, že Univerzita Karlova v Praze má právo na uzavření licenční smlouvy o užití této práce jako školního díla podle §60 odst. 1 autorského zákona.

I declare that I carried out this master thesis independently, and only with the cited sources, literature and other professional sources.

I understand that my work relates to the rights and obligations under the Act No. 121/2000 Coll., the Copyright Act, as amended, in particular the fact that the Charles University in Prague has the right to conclude a license agreement on the use of this thesis as a school work pursuant to Section 60 paragraph 1 of the Copyright Act.

V Praze dne 10. dubna 2014

.....

**Název práce:** Modelování Velké mlhoviny v Orionu

**Autor:** Václav Pavlík

**Katedra:** Astronomický ústav Univerzity Karlovy

**Vedoucí diplomové práce:** doc. RNDr. Ladislav Šubr, Ph.D. (Astronomický ústav Univerzity Karlovy)

**Abstrakt:** Otevřené hvězdokupy jsou často diskutovány z hlediska jejich vývoje a struktury. V této práci jsme se zaměřili na studium typického představitele těchto objektů — Velké mlhoviny v Orionu (M 42) — na základě observačních dat, včetně jejich porovnání s  $N$ -částicovými modely od Šubr et al. (2012). Tyto numerické modely byly inspirovány nedávno navrženým vývojovým schématem, podle něhož hvězdokupy vznikají z velmi hustých počátečních podmínek. Z analýzy rentgenových zdrojů se ukázalo jako pravděpodobné, že M 42 je rotačně symetrická ve vnitřních oblastech do 0.7 pc. Další analýzou zahrnující též observační data z optického a infračerveného oboru jsme došli k závěru, že M 42 má na větších poloměrech (do 2 pc) protáhlý tvar od severovýchodu na jihozápad. Porovnali jsme rovněž radiální profily různých hmotnostních skupin hvězd. Objevili jsme náznak inverzní hmotové segregace mezi hvězdami s hmotností v rozsahu od 1 do  $5 M_{\odot}$  a hvězdami s hmotností menší než  $0.5 M_{\odot}$ , a to v oblasti od 0.5 pc do 1.5 pc. Tento výsledek je v nesouladu s teoretickými očekáváními (i s výsledky numerických modelů) a zaslouží si další studium.

**Klíčová slova:** metody:  $N$ -částicové simulace, metody: analýza dat, hvězdokupy: konkrétní (M 42), hvězdy: dynamika

**Title:** Modelling the Orion Nebula Cluster

**Author:** Václav Pavlík

**Department:** Astronomical Institute of the Charles University

**Supervisor:** doc. RNDr. Ladislav Šubr, Ph.D. (Astronomical Institute of the Charles University)

**Abstract:** Young star clusters are widely discussed from the point of view of their evolution and structure. In this work we focused our attention on studying a typical representative of these objects — the Orion Nebula Cluster (ONC, M 42) — based on the observational data, including their confrontation with  $N$ -body models from Šubr et al. (2012). These numerical models were inspired by the recently proposed evolutionary scenario, according to which the star clusters begin their evolution from very dense initial conditions. From the analysis of the X-ray sources we revealed that the ONC is likely to be rotationally symmetric in the inner area ( $\lesssim 0.7$  pc). Further analysis including also optical and IR observational data led us to the conclusion that the ONC is elongated from the North-East to the South-West on large scales (up to 2 pc). We also compared radial profiles of different mass groups of stars and we discovered a possibly inverse mass segregation between stars with masses in the interval  $(1 ; 5) M_{\odot}$  and the stars less massive than  $0.5 M_{\odot}$  in the range from 0.5 pc to 1.5 pc. This result does not correspond to the theoretical expectations (neither to the results from the numerical models) and it deserves further investigation.

**Keywords:** methods:  $N$ -body simulations, methods: data analysis, star clusters: individual (ONC), stars: dynamics

# Contents

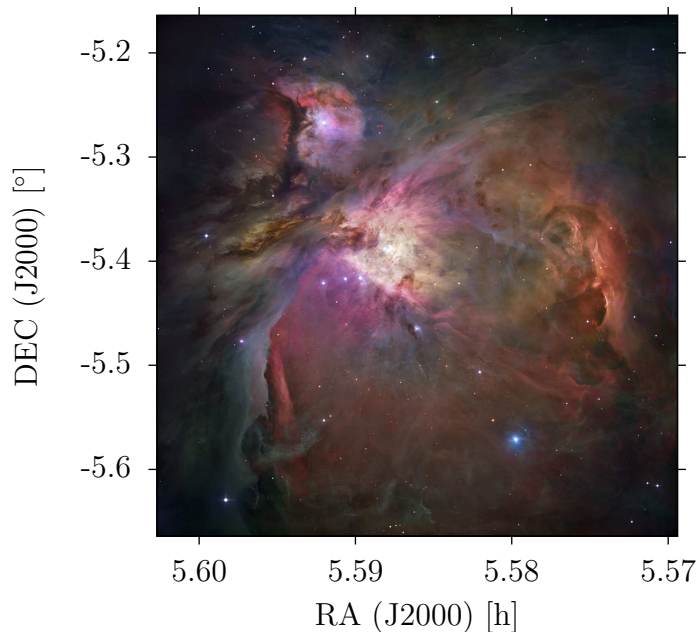
|          |  |           |
|----------|--|-----------|
| <b>1</b> | <b>Introduction</b>                                      | <b>2</b>  |
| <b>2</b> | <b>Theoretical expectations</b>                          | <b>4</b>  |
| 2.1      | Analytical estimates . . . . .                           | 4         |
| 2.2      | Numerical models . . . . .                               | 5         |
| 2.2.1    | Temporal evolution . . . . .                             | 7         |
| 2.2.2    | Mass segregation . . . . .                               | 8         |
| <b>3</b> | <b>Observational characteristics</b>                     | <b>10</b> |
| 3.1      | Properties of the cluster . . . . .                      | 10        |
| 3.2      | Datasets . . . . .                                       | 12        |
| 3.3      | Mass distribution . . . . .                              | 15        |
| <b>4</b> | <b>Structure of the cluster</b>                          | <b>17</b> |
| 4.1      | Centre of mass / Numerical centre . . . . .              | 17        |
| 4.2      | Density centre . . . . .                                 | 20        |
| 4.3      | New coordinates . . . . .                                | 20        |
| 4.4      | Angular profiles . . . . .                               | 21        |
| 4.4.1    | Centre of symmetry . . . . .                             | 23        |
| 4.4.2    | Large scale shape . . . . .                              | 23        |
| 4.5      | Radial profiles and mass segregation . . . . .           | 25        |
| 4.5.1    | Possible detection of inverse mass segregation . . . . . | 25        |
| 4.5.2    | Comparison with the numerical models . . . . .           | 26        |
| <b>5</b> | <b>Conclusions</b>                                       | <b>29</b> |
|          | <b>Bibliography</b>                                      | <b>31</b> |
|          | <b>Appendix</b>  | <b>34</b> |
| A        | Kolmogorov–Smirnov test . . . . .                        | 34        |
| A.1      | One sample testing . . . . .                             | 34        |
| A.2      | Two sample testing . . . . .                             | 37        |
| B        | Centre of the cluster . . . . .                          | 40        |
| B.1      | Centre of mass / Numerical centre . . . . .              | 40        |
| B.2      | Density centre . . . . .                                 | 42        |

# 1. Introduction

Star clusters are numbered among the objects understanding of whose is essential for many fields in astrophysics. On the one hand they are important for the evolution of galaxies (sometimes being referred to as the building blocks of galaxies). On the other hand, it is now widely accepted that the vast majority of stars, if not all of the stars, were born in numerous groups. Those groups that we observe nowadays in star forming regions are called the young star clusters.

A typical representative of this kind of objects is the Orion Nebula Cluster (also known as M 42, NGC 1976, or the ONC, see Figure 1.1). It is a part of a large star forming region in our Galaxy, the Orion Molecular Cloud (OMC-1), and it can be identified even with the naked eye as the middle “star” of the Orion’s sword asterism. Its estimated age is less than 2.5 million years (Tobin et al., 2009), characteristic radius is approximately 2 pc and it contains about 3 000 stars with total mass of nearly  $10^3 M_{\odot}$  (Robberto et al., 2013). Due to its relative proximity to the Earth ( $\sim 400$  pc) it is one of the most observed and theoretically studied star clusters. Despite all of the work, that was invested into the study of this object (with more than 1 000 scientific papers mentioning it), many uncertainties still remain.

In the most of astronomical observations (especially the ones concerning the star clusters), there is not enough of the human lifetime to see even a little fraction of the object’s evolution. We are provided only with a snapshot in time. This is what leads to the biggest question that we are facing when making numerical simulations of the star clusters: “What is the evolutionary state of the star cluster that we observe?” Considering the ONC, the raw estimate from



**Figure 1.1:** Image of the Orion Nebula in visible light placed in RA/DEC coordinates. ©NASA/ESA (Hubble’s sharpest view of the Orion Nebula, <http://www.spacetelescope.org/images/heic0601a/>, September 2013)



the observationally determined characteristics give us an answer, that the time needed for this star cluster to pass through a substantial dynamical evolution is about one order of magnitude longer than its age. Therefore, the ONC is not yet dynamically evolved. On the other hand, the recent theoretical studies (e. g. Baumgardt & Kroupa, 2007; Marks & Kroupa, 2012; Šubr et al., 2012) are emphasising the fact, that dynamics of the young star cluster is too complicated to assume that the current evolutionary state is representative for its whole lifetime. In order to create stars, it is essential for the young star cluster to contain a lot of interstellar matter. The observations of the ONC, however, indicate a lack of the residual gas. Hence, the above mentioned authors deduced that the initial state of young star cluster has been much more compact than what we observe now and that the residual gas was expelled at some time during its evolution. Consequently, the initially deep potential well was shallowed, which permitted the less bounded stars to escape, and the star cluster expanded. This kind of the initial conditions means bigger star density and significantly faster evolution of the whole system. Therefore, the observed star cluster might be dynamically relaxed.

A relevant answer to the question about the past evolution of the ONC cannot be given by the observations nor by the numerical models separately. Only a careful comparison of these two approaches, with considering various initial conditions and evolutionary scenarios, can support either one theory or the other. In this work, we attempt to confront a set of numerical models of the ONC (presented in the work of Šubr et al., 2012) with the publicly available observational data.

First, in Chapter 2 we are showing our analytical estimates on the star cluster evolution and introduce the properties of our numerical models. Then in Chapter 3, we are presenting the measured parameters of the Orion Nebula Cluster, such as its distance from the Earth or extinction, and our collection of data. In Chapter 4 we are introducing a set of Kolmogorov–Smirnov tests to study symmetries, radial profiles and level of mass segregation of the ONC. Later in Chapter 4 we are comparing the numerical models with the observational data. Finally in Chapter 5 we are discussing our results and we are putting several suggestions on the future work. In the Appendix, we comment the programs written for the purpose of this work.

## 2. Theoretical expectations

The dynamical evolution of the star cluster is generally a complicated process. The equations which describe it lead to stochastic solutions. Nevertheless, we can use the analytical estimates to reveal some of its characteristics.

### 2.1 Analytical estimates

The observations of the Orion Nebula Cluster make us believe in the possibility, that we are looking at the young star cluster at the beginning of its life (Hillenbrand, 1997). We can estimate its characteristic time of evolution, the relaxation time (Binney & Tremaine, 1994) from

$$t_{\text{relax}} \approx \frac{\sigma^3}{8\pi G^2 \langle \rho \rangle \langle M_{\star} \rangle \ln N_{\star}}, \quad (2.1)$$

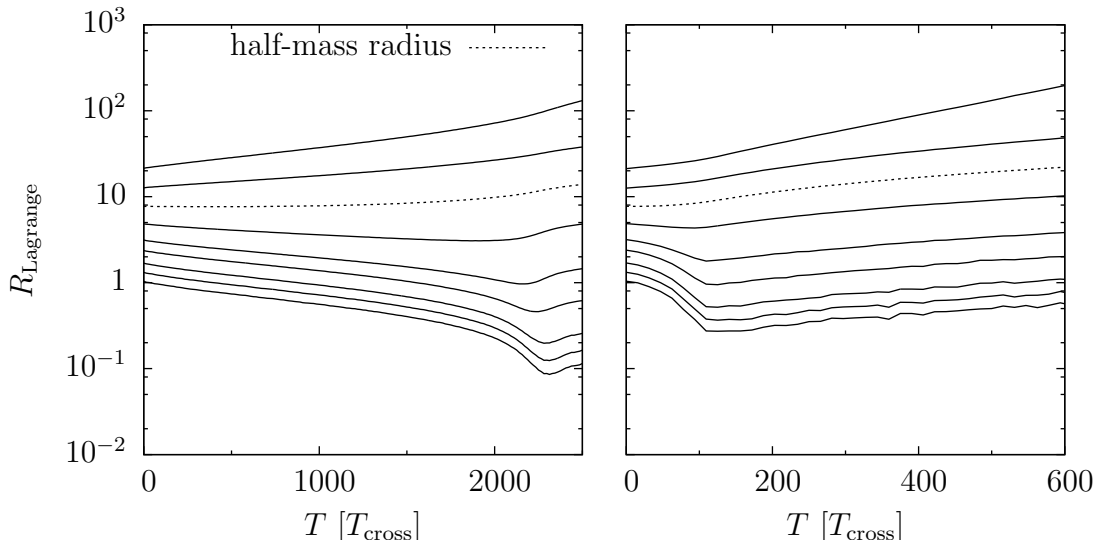
where  $\sigma$  is the mean velocity dispersion,  $G$  is the gravitational constant,  $\langle \rho \rangle$  is the mean star density,  $\langle M_{\star} \rangle$  is the mean mass per star and  $N_{\star}$  is the number of stars. Inserting values relevant for the ONC, we obtain  $t_{\text{relax}} \simeq 9.5$  Myr. We can see that this result is well beyond the age of the ONC ( $\lesssim 2.5$  Myr), but there are some facts which have not been considered in this calculation.

First, the equation (2.1) gives good estimates only for star clusters with the equal-mass stars. But the process which drives the evolution of a star cluster — the two-body relaxation (Binney & Tremaine, 1994; Heggie & Hut, 2003) — is much more effective, when there is the dispersion of masses. As the example, see Figure 2.1, where we plot the evolution of Lagrange radii of two star clusters. One of them has equal-mass stars and its evolution is slower, while the other one has the range of masses, which makes its evolution faster. We can get more precise results, if we consider the above mentioned dynamical process. Therefore, we define the dynamical time

$$t_{\text{dyn}} \approx \frac{\langle M_{\star} \rangle}{M_{\star}} t_{\text{relax}}, \quad (2.2)$$

where  $M_{\star}$  is the mass of a star we estimate the dynamical time for. Plugging in the values for the more massive stars from the ONC, we obtain  $t_{\text{dyn}} \simeq 1$  Myr, which is comparable to its estimated age.

Another correction can be obtained by following the recently discussed evolutionary scenario (e. g. Baumgardt & Kroupa, 2007; Marks & Kroupa, 2012; Šubr et al., 2012) which is supported by the observed lack of the residual gas in the ONC (Wilson et al., 1997). These works point out that initially the ONC must have had considerably more interstellar matter inside. At some time during its evolution, the radiation pressure of massive stars led to the gas expulsion. The consequence to this is that the initially deep potential well was reduced and the less bounded stars escaped. So the ONC expanded, maybe even by a factor of 3 to 5 (Marks & Kroupa, 2012). Such result also means, that this star cluster must have been denser by a factor of  $\gtrsim 10$  in the past (Baumgardt & Kroupa, 2007). Similarly the velocity dispersion has been higher as well. The initial



**Figure 2.1:** Lagrange radii (0.5 %, 1 %, 2 %, 5 %, 10 %, 25 %, 50 %, 75 % and 90 %) for two numerical models of two star clusters, both of them have the total mass of  $10\,000 M_{\odot}$ . The model on the left side has 10 000 stars with  $1 M_{\odot}$  and the model on the right side has 8 000 stars with  $1 M_{\odot}$  and 200 stars with  $10 M_{\odot}$ . The time is in both cases measured in the crossing times. The core collapse occurred later for the model with equally massive stars (left) than for the other model (right). This evolutionary scenario is caused by the dispersion of masses and cannot be explained by lower number of stars in the second model.

size, density and velocity dispersion will be reducing the characteristic time of evolution, see equation (2.1), which supports the assumption that the ONC is dynamically relaxed.

## 2.2 Numerical models

The dynamics of the young star clusters is, however, very complicated and we cannot deduce it just from the analytical estimates. This is a reason for making numerical simulations. In this work, we studied four different numerical models of the Orion Nebula Cluster, all of which are  $N$ -body models. The stars are represented by point masses and the interactions between them are calculated from the Newton's law of universal gravitation. The acceleration of the  $i$ -th star, caused by the  $j$ -th star, can be written as

$$\vec{a}_i = \frac{d^2 \vec{x}_i}{dT^2} = -\frac{GM_{\star,j}}{|\vec{x}_i - \vec{x}_j|^3} (\vec{x}_i - \vec{x}_j), \quad (2.3)$$

where  $\vec{x}$  is the position vector,  $T$  is the time,  $G$  stands for the gravitational constant and  $M_{\star,j}$  is the mass of the  $j$ -th star. Deriving from this equation, the input parameters of every star in the numerical model has to be its position vector, velocity and mass. The numerical integrator, which was used is NBODY6 (Aarseth, 2003).

One of these models (`modA`) was presented by Šubr et al. (2012) and the other three are from L. Šubr (private communication, 2014). Initially these models

| id   | radial profile     | IMS | $R_{\text{hm},0}$<br>[pc] | $T_{\text{ex}}$<br>[Myr] | $f_{\text{gas}}$ | runs |
|------|--------------------|-----|---------------------------|--------------------------|------------------|------|
| modA | Subr et al. (2008) | yes | 0.11                      | 0.50                     | 50 %             | 100  |
| modB | Plummer (1911)     | no  | 0.10                      | 0.70                     | 50 %             | 30   |
| modC | Plummer (1911)     | no  | 0.15                      | 0.50                     | 50 %             | 30   |
| modD | Plummer (1911)     | no  | 0.10                      | 0.70                     | 33 %             | 40   |

**Table 2.1:** Parameters of the numerical models of the ONC used in this work. IMS means the initial mass segregation,  $R_{\text{hm},0}$  is the initial half-mass radius (the radius comprising exactly one half of the total mass of stars),  $T_{\text{ex}}$  is the time, when the gas expulsion started and  $f_{\text{gas}}$  is the initial mass fraction of the gas. In the last column is the number of runs of each model.

are constructed as very compact (by a factor of  $\sim 7$  smaller than the ONC, see Table 2.1), spherically symmetric and isolated. The latter characteristic means, that the Galactic potential and the tidal effects are neglected. All these models of the ONC have total mass of  $2\,700\,M_{\odot}$ . The individual mass of each star ( $M_{\star}$ ) is defined with respect to the standard Kroupa’s initial mass function (Kroupa, 2001)

$$\xi(M_{\star}) = \begin{cases} M_{\star}^{-0.3}, & \text{if } M_{\star} < 0.08 \\ M_{\star}^{-1.3}, & \text{if } 0.08 \leq M_{\star} < 0.5 \\ M_{\star}^{-2.3}, & \text{if } M_{\star} \geq 0.5 \end{cases} \quad (2.4)$$

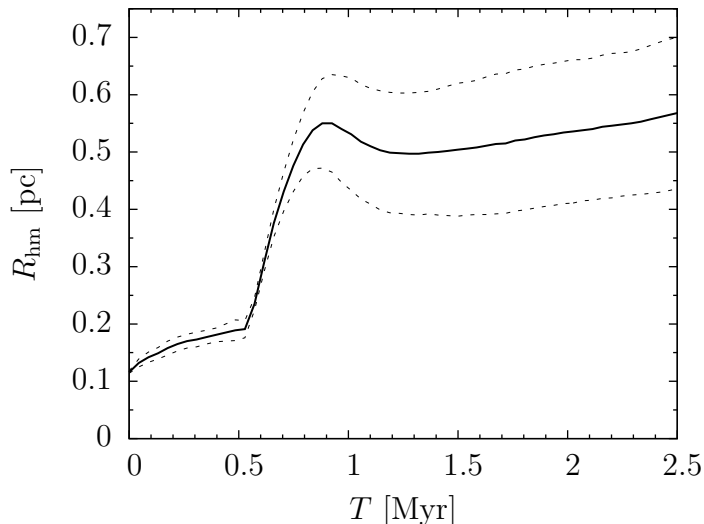
with the upper limit of  $M_{\star}$  at  $80\,M_{\odot}$  and the bottom cutoff at  $0.1\,M_{\odot}$ . The function  $\xi(M_{\star})dM_{\star}$  gives the number of stars in the interval  $(M_{\star}; M_{\star} + dM_{\star})$ . The stellar evolution in these models is present in a form of star collisions and merges only. Although globally stars behave like point masses, during close encounters they are treated as spheres of previously defined radius. When two stars overlap too much, they merge in one star.

The gas was modelled as well. The correct procedure would be to incorporate hydrodynamical equations and integrate them. But such models are so complicated and use so much computational time, that it is impossible to use them to study the whole evolution of the star cluster. Therefore, we have a simplification in our models — the gas is represented by point masses with low mass. It forms a substantial fraction of the total mass of the star cluster (see Table 2.1 for  $f_{\text{gas}}$ ). At the defined time  $T_{\text{ex}}$  during the simulations (see Table 2.1), the gas has been expelled and the star cluster begun to expand. In Figure 2.2 we show evolution of the half-mass radius of the model modA. We can clearly see that shortly after  $0.5\,\text{Myr}$  the star cluster started the expansion. The hump at  $0.9\,\text{Myr}$  is the sign of resettling of the virial equilibrium. After that, the model does not evolve so much, which is in agreement with our analytical estimates.

In general, the models of self-gravitating systems are scalable. But we have to keep the constraints, which come from the equations of motion. For example, if we want to multiply length by a factor  $R_{\text{scale}}$ , the factor  $T_{\text{scale}}$ , which multiplies time, is given by

$$T_{\text{scale}} \propto R_{\text{scale}}^{3/2}, \quad (2.5)$$

as derives from the equation (2.3). If we do not want an unrealistic evolution rate, such as 3 times the speed, a good limitation is  $R_{\text{scale}} \lesssim 2$ . These models,



**Figure 2.2:** Evolution of the half-mass radius of model `modA` (solid line) and the standard deviation  $\pm\sigma$  (dashed lines). Graph taken from Šubr et al. (2012).

however, have a set length scale. The motivation to define it was to set the stellar radii in order to study the collisions between the stars. If we shrunk the model, the collisions would be more frequent and vice versa. We also know, that there were only few collisions during the whole star cluster’s evolution ( $\lesssim 10$ ). So we can assume, that the global characteristics of the star cluster will not change dramatically (except the wrong number of collisions), when we free the length scale and rescale the model to our needs.

### 2.2.1 Temporal evolution

To be consistent with the observations, we will from now on look only at the planar projection of the models. We also define a new variable, the half-count radius, which is the analogy to the half-mass radius, except that it stands for the radius that comprises one half of the total number of stars, not mass. Note, that the projected value of the half-count radius is smaller in comparison to the three-dimensional half-count radius. When projecting, we do not take just the stars in the central sphere, but we include all stars in the cylinder in our line of sight, which means that we take into account more stars. Therefore, we come to the wanted result quicker, hence the radius is smaller.

We do not know the exact age of the ONC, only its upper limit (2.5 Myr). We also showed that the numerical models can be rescaled, which will modify the time scale by the relation (2.5). Because of these two reasons, it is sufficient to analyse the models only in chosen time frames, which roughly represent the model evolution. We chose  $T = 0$  Myr (at the beginning),  $T \simeq 1.2$  Myr (after the gas expulsion and resettling of the virial equilibrium) and  $T \simeq 2.5$  Myr (at the maximum estimated age of the ONC and also at the end of integration). From the computed values of the half-count radii in Table 2.2, we can see that all models expanded by the factor of  $> 3$  after the phase of the gas expulsion. Between 1.2 and 2.5 Myr these models did not evolve as much, as has been predicted by our analytical estimates.

| id   | $R_{\text{hc}} (0 \text{ Myr})$ [pc] | $R_{\text{hc}} (1.2 \text{ Myr})$ [pc] | $R_{\text{hc}} (2.5 \text{ Myr})$ [pc] |
|------|--------------------------------------|--|--|
| modA | 0.13                                 | 0.58                                   | 0.57                                   |
| modB | 0.07                                 | 0.32                                   | 0.40                                   |
| modC | 0.11                                 | 0.37                                   | 0.48                                   |
| modD | 0.08                                 | 0.42                                   | 0.45                                   |

**Table 2.2:** Projected half-count radii of the models in three analysed time frames.

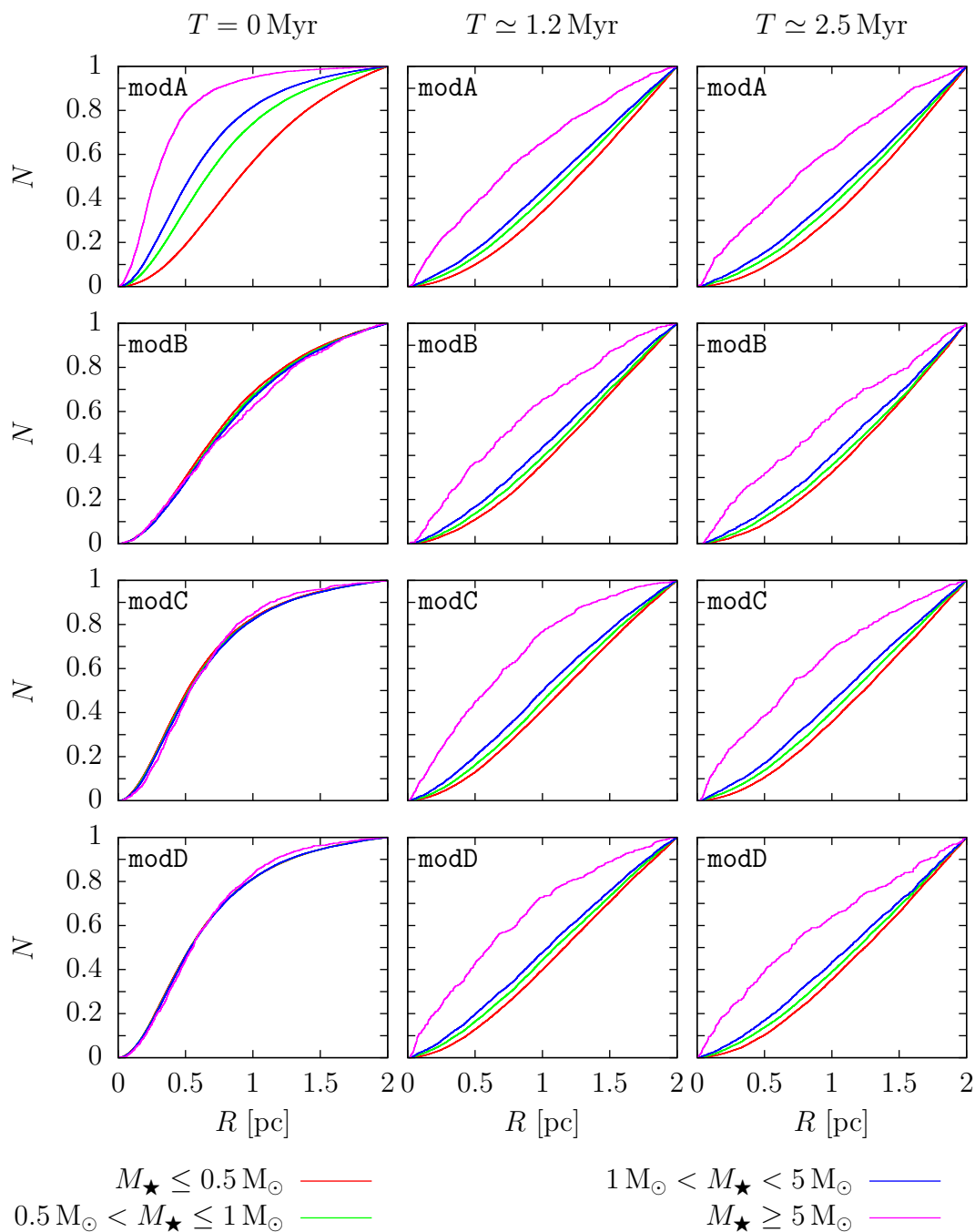
## 2.2.2 Mass segregation

One of the consequences of the two-body relaxation is the mass segregation — different distribution of more massive stars in comparison to the less massive stars. In the case of the star clusters, this usually means that more massive stars are concentrated towards the centre, while the less massive ones are spread in the outer regions. Due to the negative heat capacity of the self-gravitating system, mass segregation will lead to the process called core-collapse (Binney & Tremaine, 1994).

To test the numerical models for the mass segregation, we divided all stars into four groups of masses within the intervals

$$\begin{aligned}
 M_{\star} &\leq 0.5 M_{\odot}, \\
 0.5 M_{\odot} &< M_{\star} \leq 1 M_{\odot}, \\
 1 M_{\odot} &< M_{\star} < 5 M_{\odot} \\
 \text{and } M_{\star} &\geq 5 M_{\odot}
 \end{aligned}$$

and compared their cumulative distribution functions by the one-dimensional Kolmogorov–Smirnov test (KS test) from Press et al. (2007). This test analyses the maximal difference of two cumulative distribution functions and returns the value  $Q_{\text{KS}}$  (between 0 and 1), which is the probability of this difference to be a coincidence. We set the boundary value of  $Q_{\text{KS}}$  at 0.05 — when result of the KS test is below this value, we reject the null hypothesis. As a centre of the model of the star cluster (a point from which we started to build these cumulative distribution functions), we took the centre calculated by NBODY6 via the algorithm of Casertano & Hut (1985). Our tests revealed that all models with no initial mass segregation (IMS, see Table 2.1) evolved already after 1.2 Myr into the mass segregated clusters between all mass groups. They preserved in that state up to 2.5 Myr, as shown in Figure 2.3. The probability ( $Q_{\text{KS}}$ ) that the stars of any mass group are distributed with respect to the same distribution function is  $\lesssim 10^{-5}$  at these two times. The initially mass segregated model (modA) kept the mass segregation during its evolution also with  $Q_{\text{KS}} \lesssim 10^{-5}$ . This result was expected, because theoretically, denser star clusters tend to evolve faster, as we showed by the analytical estimates.



**Figure 2.3:** Evolution of the cumulative distribution functions of different mass groups ( $M_{\star} \leq 0.5 M_{\odot}$ ,  $0.5 M_{\odot} < M_{\star} \leq 1 M_{\odot}$ ,  $1 M_{\odot} < M_{\star} < 5 M_{\odot}$  and  $M_{\star} \geq 5 M_{\odot}$ ) of four numerical models of the ONC (see Table 2.1).

# 3. Observational characteristics

One of the issues we are facing, is the ongoing discussion about the physical parameters of the Orion Nebula Cluster. Different methods of measurements and specific assumptions lead to different results of the age of this star cluster or its distance from the Earth. Another issue is in the inconsistency between measurable parameters from the observations and the parameters from the numerical models. For example the masses of individual stars, which are the essential input parameters of the numerical models, are not easy to determine from the observations. In general, masses of the stars cannot be observed, therefore, we have to use some additional methods, e. g. the models of the stellar evolution, to estimate them from the observable parameters, such as luminosities. Note, that this approach might not work on all kinds of the stars. Below we present the physical properties of the ONC and the observational data, that we used in this thesis.

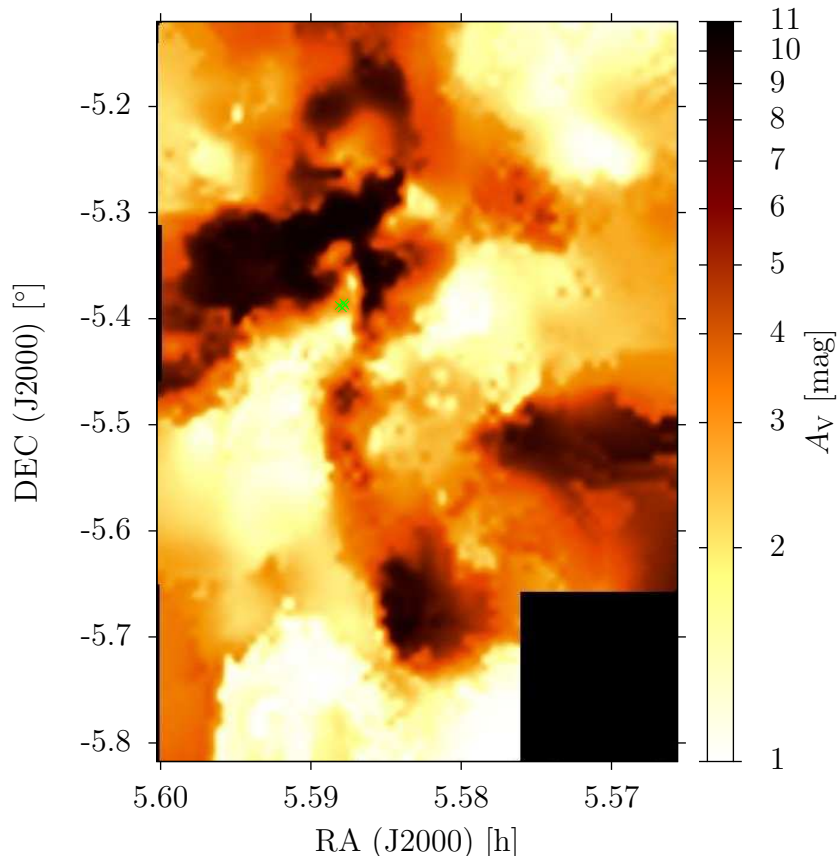
## 3.1 Properties of the cluster

Due to the character of the astronomical observations, we are able to determine precisely only the angular sizes of objects. To get the absolute size (e. g. in parsecs), which is natural for the numerical models, it is crucial to know the distance from the Earth to the studied object. In Chapter 1, we mentioned that distance to the ONC is about 400 pc, but this is just the rounded value and certainly is not definite. In older works, we can find the measurements of the distance modulus to the star forming regions, using the UBV photometry (Walker, 1969) or the maser radiation (Genzel et al., 1981), which led to value 8.37 for the ONC, or equivalently 472 pc, where precision of the latter method is only about  $\pm 20\%$ . The number 470 pc was thought of as canonical and sometimes used without any reference or further discussion (e. g. as in Hillenbrand & Hartmann, 1998). Newer methods measured the ONC to be closer with more precise results. For example Jeffries (2007) got two results from studying the pre-main sequence stars — first is  $(440 \pm 34)$  pc and the second result, when omitting stars with the accretion disks, is  $(392 \pm 32)$  pc. In the same year Menten et al. (2007) obtained the value of  $(414 \pm 7)$  pc from measuring the trigonometric parallaxes of the stars from the ONC. Another recent study (O’Dell & Henney, 2008) places this star cluster at  $(436 \pm 20)$  pc. In our work, we adopt the value 414 pc, because it is close to the average of the previously discussed distances and it also has the smallest uncertainty. For this distance, the conversion between the angular and the proper separation reads

$$1' = 0.1204 \text{ pc} . \tag{3.1}$$

The similar problem comes out, when estimating the age of the star cluster. We have evidence for stars within the ONC which are more than 5 Myr old (Huff & Stahler, 2006). Some works even mention stars, that are up to 12 Myr old (Warren & Hesser, 1977, based on the measurements of Balmer’s  $H\beta$ ). On the other hand, the youngest stars are less than 1 Myr old. From this spread of ages, we can see that the ONC did not form in one burst (Palla et al., 2005), as we assume in the numerical models. The question is, when the forming star cluster

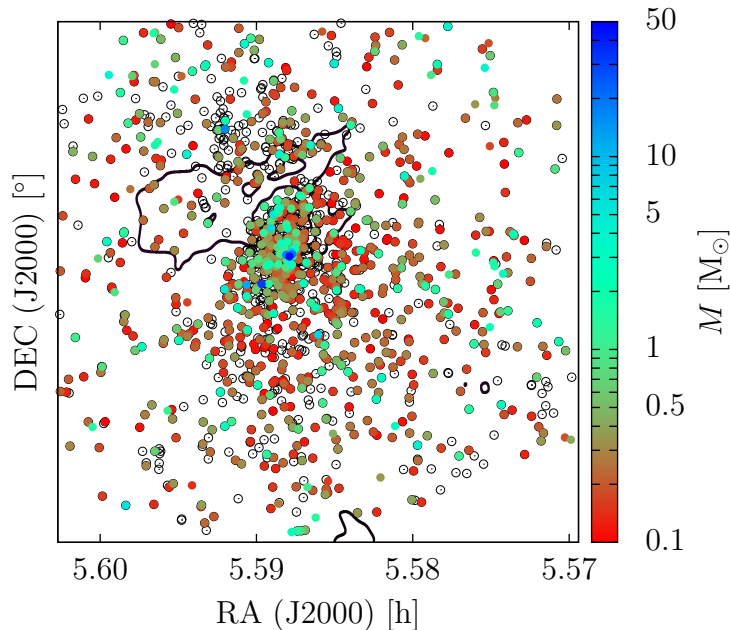




**Figure 3.1:** Extinction map of the Orion Nebula with the value of  $A_V$  defined by the horizontal colour bar. Data from Scandariato et al. (2011). Green points are the positions of the Trapezium stars.

becomes really cluster-like, i. e. when we consider that its “age is set to zero”. In this work and the numerical models, that we used, we assume that the age of the ONC is estimated from the time of the most active star formation, i. e. 1 to 2 Myr ago and does not exceed 2.5 Myr (Hillenbrand, 1997; Palla & Stahler, 1999).

The structure of the ONC is complicated as well. The stars are surrounded and also partially covered by the interstellar medium (see Figure 1.1), which is part of the Orion Molecular Cloud (Hillenbrand, 1997; Hillenbrand & Hartmann, 1998). This resolves in varying extinction in different wavelengths (Scandariato et al., 2011, see the extinction map in Figure 3.1). Even though the optical photometry and spectroscopy can be used in certain areas of the ONC, in some regions, where the gas is too dense (black colour in the extinction map), only X-ray observations can give us reliable results, because for the X-ray light, the interstellar medium is likely more transparent (Flaccomio et al., 2003a,b).



**Figure 3.2:** Positions of the ONC members from Hillenbrand (1997). Colour represents the masses of stars (the range is on the right), the blank circles shows stars with unknown mass. On the background is the isocontour for  $A_V = 5$  mag.

| data source                | spectral region           | number of stars |
|----------------------------|---------------------------|-----------------|
| Hillenbrand (1997)         | optical                   | 1449            |
| Flaccomio et al. (2003a,b) | X-ray                     | 742             |
|                            | optical                   | 696             |
| Da Rio et al. (2009)       | near-infrared and optical | 1228            |
| Da Rio et al. (2012)       | near-infrared             | 1583            |
| Simbad: NAME ONC           | IR, optical and X-ray     | 2156            |

**Table 3.1:** Data used in this work. The tables with detailed informations about each source are in the electronic version.

## 3.2 Datasets

Here is the list of the existing data, that we used in this work, which were taken from publicly accessible database at VizieR<sup>1</sup>.

### Hillenbrand (1997)<sup>2</sup>

This dataset is a merge of existing optical observations up to 1994 and the new data acquired by the author of that paper. It gives us large sample of stars (almost 1500) in the Orion Nebula region. We also adopt masses, which are determined in the catalogue for 929 stars from their luminosities and temperatures, using Hertzsprung–Russell diagram and evolutionary models of D’Antona & Mazzitelli (1994), Swenson et al. (1994) and Ezer & Cameron (1967).

<sup>1</sup><http://vizier.u-strasbg.fr/viz-bin/VizieR>

<sup>2</sup><http://vizier.u-strasbg.fr/viz-bin/VizieR?-source=J/AJ/113/1733>, march 2014

Following the membership probability discussion in Hillenbrand (1997), from the total number of 1 578 stars, we excluded those, which were characterised as the proper motion non-members (Jones & Walker, 1988, total of 127 stars). We also excluded all the new stars found by the author (their identification number is from 3 000 to 3 999, from 5 000 to 5 999 or from 7 000 to 7 999), which are farther from  $\theta_1$  Ori C than  $18.29'$  (or equivalently farther than 2.5 pc at 470 pc distance).

The complete dataset is in Table 3.2. For the informational purpose we also plot all these stars in Figure 3.2 with the colour range representing their masses, or the blank circles, where the mass is unknown.

### Flaccomio et al. (2003a,b)<sup>3</sup>

The authors of these two consecutive papers identified 742 objects active in the X-ray and 696 in the optical wavelengths. All stars from the optical sample have been pronounced as likely ONC members — they either have membership probability greater than 50% or yet unknown. For most of the X-ray sources, the optical or infrared counterparts have been found. The rest of the stars was, based on the extinction and other parameters, “associated” with the possible stars which are deeply embedded in the molecular cloud and not yet observed. We take the whole X-ray and optical sample in the following calculations. Masses are present for each star in the optical sample. All sources that we used from these two papers are listed in Table 3.3 and 3.4.

### Da Rio et al. (2012)<sup>4</sup>

For measurements in the near-infrared wavelengths we used data from Da Rio et al. (2012), see Table 3.6 for the complete dataset. These observations were focused on brown dwarfs and pre-main sequence stars in the Orion Nebula Cluster. To estimate their masses, two evolutionary models were used (Baraffe et al., 1998; D’Antona & Mazzitelli, 1998). Vast majority of these sources have  $M_\star < 1 M_\odot$ . From all of the objects in this dataset, we distinguish between probable members of the ONC and the foreground/background contamination. This differentiation is based on the reddening of sources done by Alves & Bouy (2012).

### Online database

In addition to the data sources listed above, we also used the online database service Simbad<sup>5</sup>, which is populated by data, identifiers and references taken from large set of journals. The advantage of this website is not only, that it gathers data from various observations of the individual objects, but that it also connects them together. One of these datasets is called NAME ONC and is the collection of the sources, that are part of the Orion Nebula Cluster — the selection of the objects is done by evaluating their properties (M. Wenger, personal communication, January 29, 2014). Some of the large catalogues are cross-identified, which leads to the possibility of making one more complete dataset without doubling any

<sup>3</sup><http://vizier.u-strasbg.fr/viz-bin/VizieR?source=J/ApJ/582/382>, march 2014

<sup>4</sup><http://vizier.u-strasbg.fr/viz-bin/VizieR?source=J/ApJ/748/14>, march 2014

<sup>5</sup><http://simbad.u-strasbg.fr/simbad>

data. In NAME ONC, there are incorporated data from Hillenbrand (1997) with identifier [H97b], Da Rio et al. (2009) with the identifier [DRS2009], the X-ray sources from Flaccomio et al. (2003a,b) with identifier [FDM2003] X and several others. Together it forms the dataset with the total number of 2 156 stars. From the list in Table 3.1, there are missing informations about Da Rio et al. (2012) measurements — request of completing the NAME ONC was transmitted to the CDS Helpdesk (personal communication, January 30, 2014).

As for the later analysis and comparisons with the numerical models, we need coordinates and masses of the stars. Simbad makes possible to display multiple identifiers from different catalogues for each star, which was useful to retrieve necessary informations about stars listed in the NAME ONC. The coordinates are present for every star, while the masses are estimated only in Hillenbrand (1997) and Da Rio et al. (2009) datasets, which means 1 023 stars. See Table 3.5 for the complete set of the stars.

| RA J2000<br>[h m s] | DEC J2000<br>[° ' '' ] | [H97b] | prob<br>[%] | $R_{\text{proj}}$<br>[m] | $M_{\star}$<br>[ $M_{\odot}$ ] |
|---------------------|------------------------|--------|-------------|--------------------------|--------------------------------|
| 05 34 14.87         | −05 15 48.96           | 3024   |             | 17.14                    | 0.34                           |
| 05 34 31.71         | −05 35 20.32           | 21     | 95          | 16.37                    |                                |
| ...                 | ...                    | ...    | ...         | ...                      | ...                            |

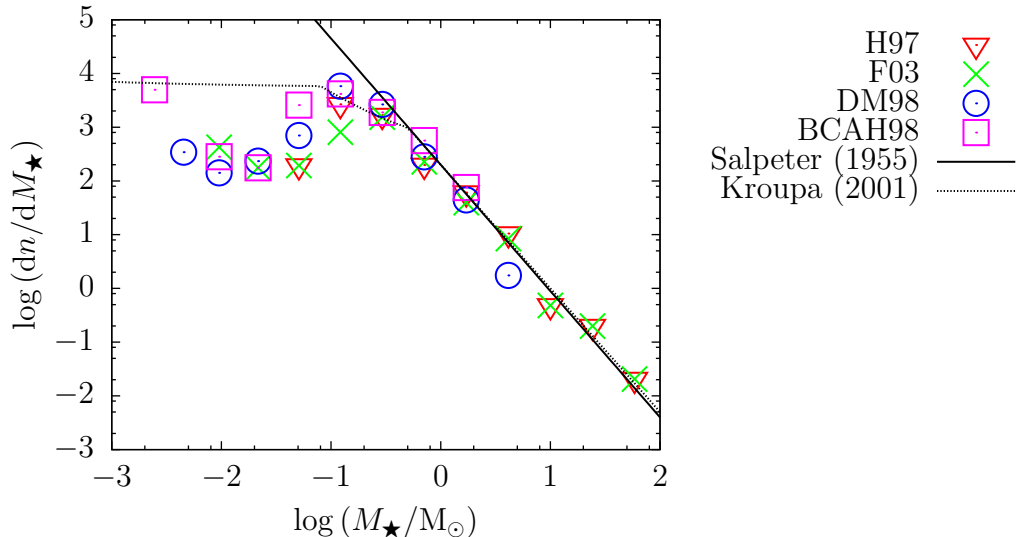
**Table 3.2:** The header and few lines of the dataset from Hillenbrand (1997). First two columns are the right ascension and declination, [H97b] is the identifier of each star, in the next column is the value of proper motion membership probability,  $R_{\text{proj}}$  stands for the projected radius from  $\theta_1$  Ori C in arc minutes and  $M_{\star}$  is the mass of each star in  $M_{\odot}$ . The electronic version is in the file Table\_H97.

| RA J2000<br>[h m s] | DEC J2000<br>[° ' '' ] | [FDM2003] X |
|---------------------|------------------------|-------------|
| 05 34 14.46         | −05 28 16.62           | 1           |
| 05 34 18.23         | −05 33 28.53           | 2           |
| ...                 | ...                    | ...         |

**Table 3.3:** The header and few lines of the X-ray dataset from Flaccomio et al. (2003a,b). First two columns are the right ascension and declination and [FDM2003] X is the identifier of each star. The electronic version is in the file Table\_F03\_X.

| RA J2000<br>[h m s] | DEC J2000<br>[° ' '' ] | [FDM2003] Opt | $M_{\star}$<br>[ $M_{\odot}$ ] |
|---------------------|------------------------|---------------|--------------------------------|
| 05 34 12.81         | −05 28 48.28           | 1             | 0.72                           |
| 05 34 14.39         | −05 28 16.77           | 2             | 1.74                           |
| ...                 | ...                    | ...           | ...                            |

**Table 3.4:** The header and few lines of the optical dataset from Flaccomio et al. (2003a,b). First two columns are the right ascension and declination, [FDM2003] Opt is the identifier of each star, in the next column is the mass of each star in  $M_{\odot}$ . The electronic version is in the file Table\_F03\_Opt.



**Figure 3.3:** Mass histograms of stars from Hillenbrand (1997) — H97, Flaccomio et al. (2003a,b) — F03, Da Rio et al. (2012) with masses from D’Antona & Mazzitelli (1998) — DM98, and Da Rio et al. (2012) with masses from Baraffe et al. (1998) — BCAH98. Dashed and solid lines are representing the initial mass functions from Kroupa (2001) and Salpeter (1955) respectively.

### 3.3 Mass distribution

As it is mentioned above, some of the stars from our observational datasets have the mass unknown. Nevertheless, we assume that all heavy stars ( $M_{\star} \gtrsim 5 M_{\odot}$ ) and the majority of the moderate stars ( $1 M_{\odot} \lesssim M_{\star} \lesssim 5 M_{\odot}$ ) have their masses estimated (Hillenbrand, 1997). In Figure 3.3 we present the histogram of the star masses and the mass functions from Kroupa (2001), also shown in the equation (2.4), and from Salpeter (1955), described by

$$\xi(M_{\star}) = M_{\star}^{-2.3}. \quad (3.2)$$

Both of these theoretical functions are compatible with our datasets at higher masses, but neither of them fits the low mass stars well, most probably because of the observational limit. In this graph we can also see the dependency on the model of the stellar evolution which is being used — we have clearly different counts of the stars from Da Rio et al. (2012) in various mass intervals, depending on whether the masses are estimated by Baraffe et al. (1998) or by D’Antona & Mazzitelli (1998).

| RA J2000<br>[h m s] | DEC J2000<br>[° ' "] | [DRS2009] | [H97b] | $M_{\star}^{\text{H97}}$<br>[ $M_{\odot}$ ] | $M_{\star}^{\text{S00}}$<br>[ $M_{\odot}$ ] | $M_{\star}^{\text{P99}}$<br>[ $M_{\odot}$ ] |
|---------------------|----------------------|-----------|--------|---|---|---|
| 05 34 30.25         | −05 11 48.20         | 344       | 10     | 0.59  | 1.02  | 1.005                                       |
| 05 34 48.82         | −05 33 33.06         | 899       | 100    | 0.26  | 0.343                                       | 0.32  |
| ...                 | ...                  | ...       | ...    | ...   | ...   | ...   |

**Table 3.5:** The header and few lines of the dataset from NAME ONC. First two columns are the right ascension and declination. The Simbad identifier of Da Rio et al. (2009) is in the column [DRS2009], the identifier of Hillenbrand (1997) is in [H97b]. The masses from Hillenbrand (1997) are denoted  $M_{\star}^{\text{H97}}$ , the masses determined by Siess et al. (2000) are in  $M_{\star}^{\text{S00}}$  and the masses from Palla & Stahler (1999) in the column  $M_{\star}^{\text{P99}}$ . The electronic version is in the file Table\_NAME\_ONC.

| RA J2000<br>[h m s] | DEC J2000<br>[° ' "] | id  | $M_{\star}^{\text{BCAH98}}$<br>[ $M_{\odot}$ ] | $\sigma_{M_{\star}^{\text{BCAH98}}}$<br>[ $M_{\odot}$ ] | $M_{\star}^{\text{DM98}}$<br>[ $M_{\odot}$ ] | $\sigma_{M_{\star}^{\text{DM98}}}$<br>[ $M_{\odot}$ ] | $A_{\text{V}}$<br>[mag] | $\sigma_{A_{\text{V}}}$<br>[mag] |
|---------------------|----------------------|-----|--|---|--|---|-------------------------|----------------------------------|
| 05 34 15.10         | −05 23 00.0          | 1   | 0.243  | 0.053   | 0.255  | 0.038   | 0                       | 0.30                             |
| 05 34 17.26         | −05 22 36.7          | 2   | 0.481  | 0.028   | 0.447  | 0.034   | 0.77                    | 0.238                            |
| ...                 | ...                  | ... | ...  | ...   | ...  | ...   | ...                     | ...                              |

**Table 3.6:** The header and few lines of the dataset from Da Rio et al. (2012). First two columns are the right ascension and declination, in the next column is the identifier of each star (not integrated in the Simbad database),  $M_{\star}^{\text{BCAH98}}$  is the mass determined by Baraffe et al. (1998),  $M_{\star}^{\text{DM98}}$  stands for the mass calculated by D’Antona & Mazzitelli (1998) and  $A_{\text{V}}$  is the extinction in the V band. The value  $\sigma$  is the uncertainty. The electronic version is in the file Table\_D12.

## 4. Structure of the cluster

In order to analyse the structure of the Orion Nebula Cluster from the point of view of the symmetries and the radial profiles, we have to find its centre. There is, however, no unambiguous definition of such a point. We can use several methods that will find some centres, but there is no direct answer to the question, which of these methods is correct. For example the centre of mass might be a good solution, but in the case of the star cluster, one star far away from the rest can shift it outside the area occupied by the majority of stars (where we would expect it to be). Here we introduce two methods that we applied on the data presented in Chapter 3. We also discuss their advantages and limitations.

### 4.1 Centre of mass / Numerical centre

The coordinates of the centre of mass (CoM) are given by

$$\vec{x}_{\text{CoM}} = \frac{\sum_i^{N_\star} \vec{x}_i M_{\star,i}}{\sum_i^{N_\star} M_{\star,i}}, \quad (4.1)$$

where  $\vec{x}_i$  is the position vector of the  $i$ -th star,  $M_{\star,i}$  is its mass and  $N_\star$  is the total number of the stars. Because this method is applicable only on those stars which have determined masses, we defined its simplification — the numerical centre (NC). The only difference between these two approaches is that the latter works with counts of stars, instead of their masses. The equation which describes it, is defined as

$$\vec{x}_{\text{NC}} = \frac{\sum_i^{N_\star} \vec{x}_i}{N_\star}. \quad (4.2)$$

Both of these variations, equation (4.1) and (4.2), have to be used carefully because of their dependence on the layout of stars. If we take the whole star cluster, we might get a not reliable result, as is mentioned above. To eliminate such error as much as possible, we looked for the centre of the ONC only in the restricted areas close to the central region only. For this purpose we used the iterative process (source code is included in the Appendix B), which can determine the centre inside a given radius or for a fraction of stars, e. g. the half-count.

1. The program estimates the position of the centre for all the stars,
2. from that point it either
  - (a) sets the limit radius to the value we input and then a new number of stars inside this area ( $N_\star$ ) is counted,
  - (b) or calculates the limit radius for the fraction of stars we input (this gives the number  $N_\star$  automatically),
3. then it uses either

- (a) the equation (4.1), if we want to calculate the centre of mass,
  - (b) or the equation(4.2) for the numerical centre,
4. from which it computes the position of the new centre for the stars within the wanted area.
  5. Now it resets  $N_{\star}$  to the total number of stars in the dataset and loop back to the step 2, until the results from two consecutive loops are the same.

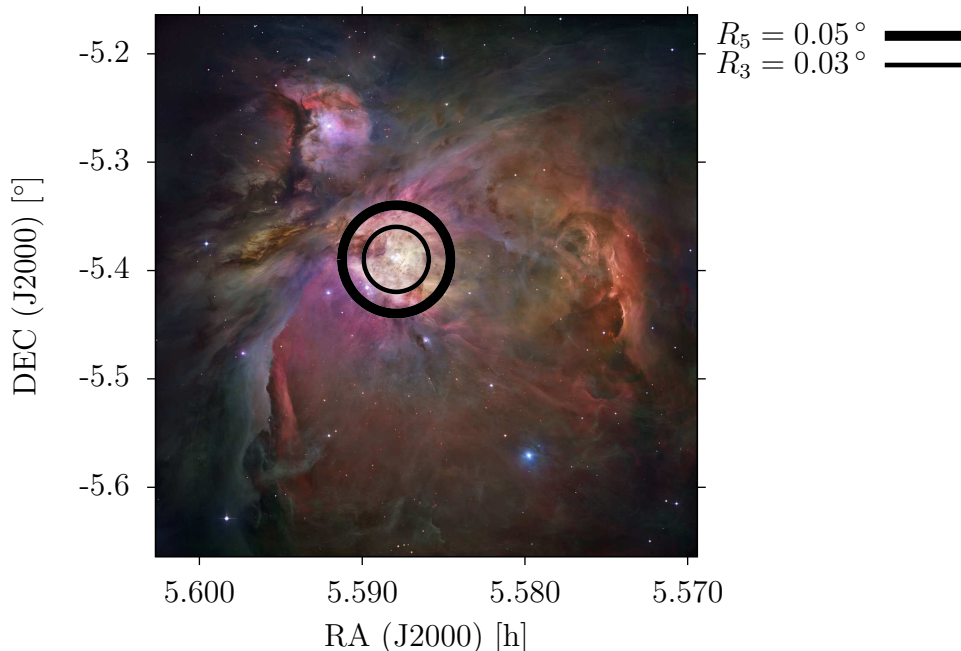
This iterative process can give us up to four different centres for one input dataset just by combining the methods in the step 3 with the type of centre from the step 4. Other results are dependent on the input parameters. We can see in the upper panel of Figure 4.2 that the centre for 90 % of the stars is slightly shifted upwards and to the right from the centre for 50 % because of the lack of stars on the left. The assumption that all observations we have are uniformly distributed through the ONC led us to the conclusion, that into further searching we can include only so many stars, that their surrounding radius does not intercept the area with high extinction ( $A_V \gtrsim 5$  mag). That is why we defined two limit radii (see Figure 4.1), one at

$$R_5 = 0.05^\circ = 3', \quad (4.3)$$

which roughly comprises one half of the X-ray sample or one third of the stars from Hillenbrand (1997). We also defined one limit radius even more centred at

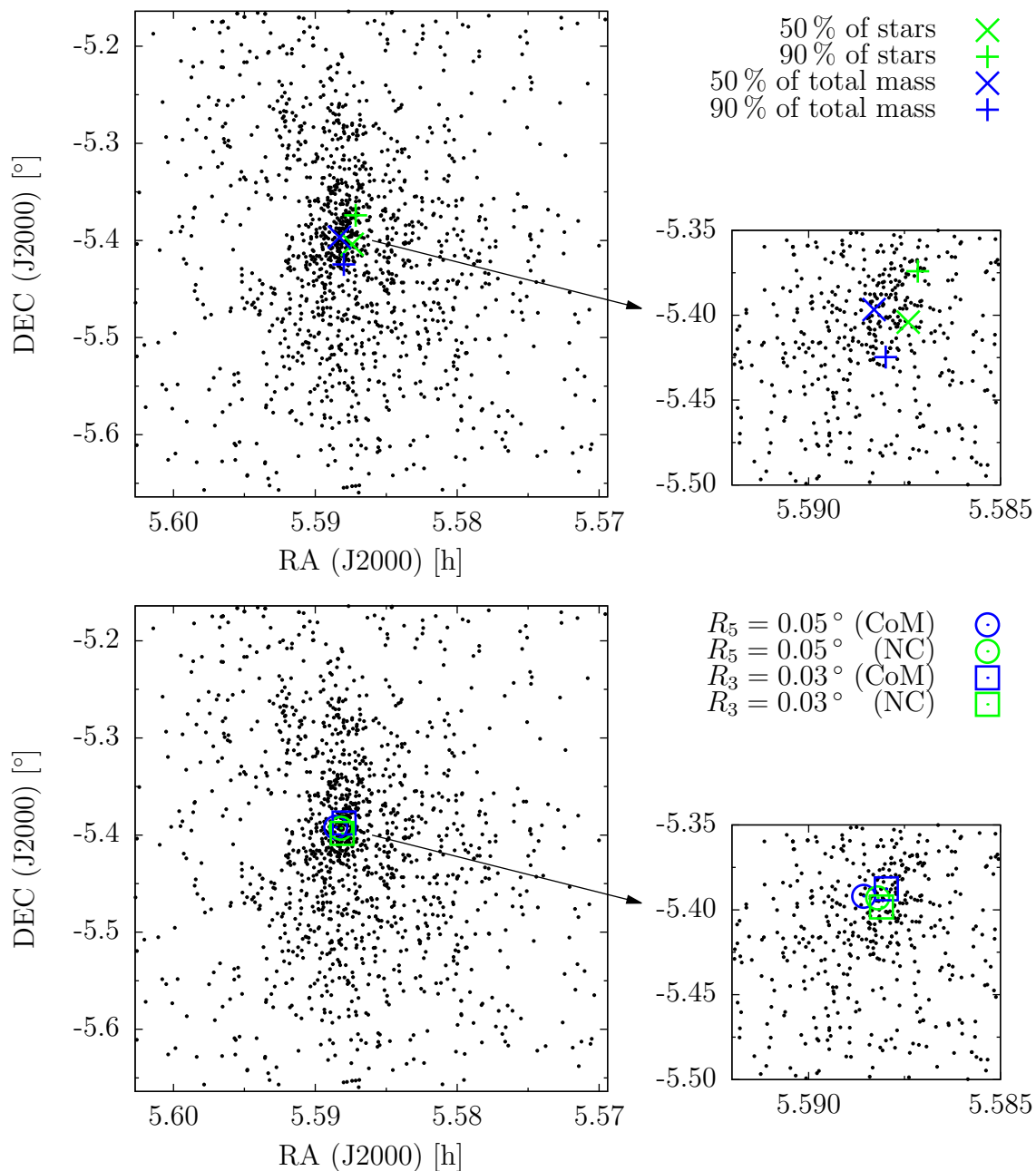
$$R_3 = 0.03^\circ = 1.8'. \quad (4.4)$$

The results for these two radii are plotted in the bottom panel of Figure 4.2. We can see that they are less spread than the centres for 50 % and 90 %.



**Figure 4.1:** Circles with limit radii defined in equations (4.3) and (4.4). The bigger circle ( $R_5$ ) is set to barely touch the area with the extensive extinction ( $A_V \gtrsim 5$  mag).

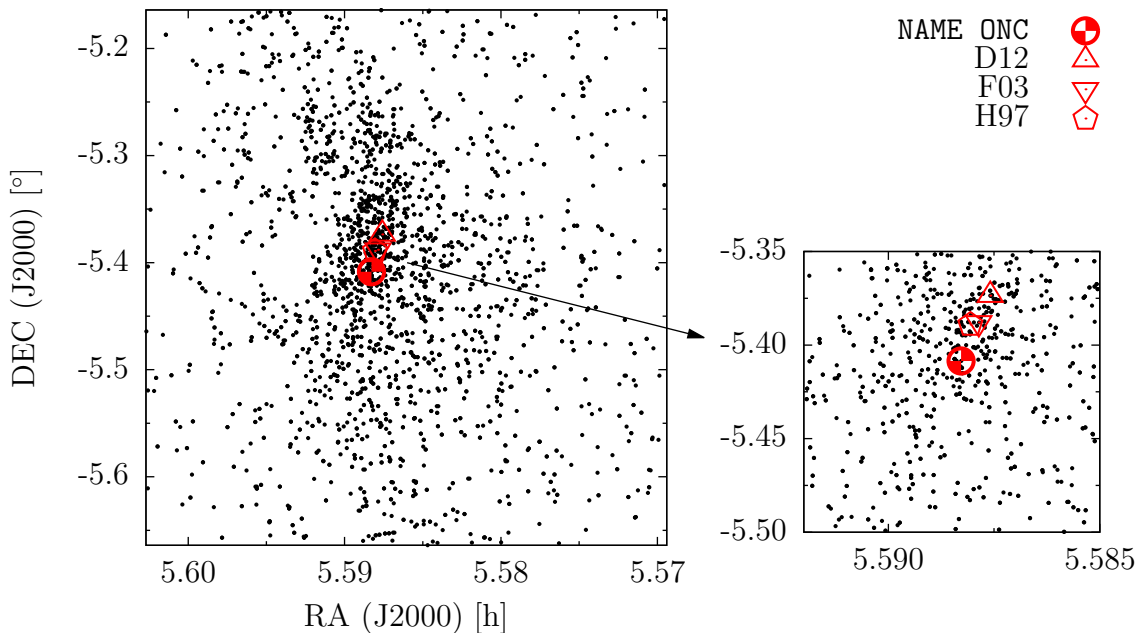




**Figure 4.2:** The centre of mass (blue) and the numerical centre (green) of the Orion Nebula Cluster represented by the dataset NAME ONC (see Chapter 3).

**Top panel:** Plus signs are used for 50% and crosses for 90% of the stars (or the total mass, depending on the method used).

**Bottom panel:** Circles are centres for stars up to  $R_5$  and squares have limit at  $R_3$ , defined in the equations (4.3) and (4.4).



**Figure 4.3:** Density centres for the Simbad sample NAME ONC (half-filled circle), Da Rio et al. (2012, triangle) — D12, X-ray sources of Flaccomio et al. (2003a,b, up side down triangle) — F03, Hillenbrand (1997, pentagon) — H97. On the background, there are all the stars from NAME ONC and Da Rio et al. (2012).

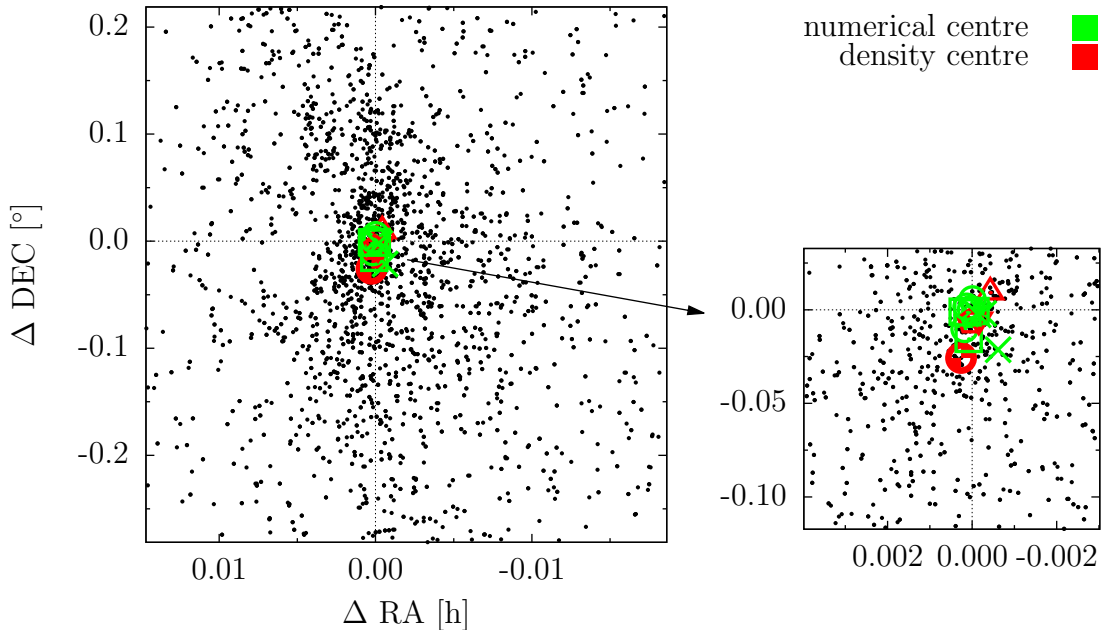
## 4.2 Density centre

Due to the core-collapse, the star clusters form during their evolution a very compact core and wide outer regions with low density of stars (Binney & Tremaine, 1994). This motivated Casertano & Hut (1985) to define the density centre. It is done by using an algorithm, which finds the  $n$ -th closest star to each star in the particular dataset. The centre is then the star, where this calculated distance is minimal (for the numerical implementation see Appendix B). The positions of these centres are shown in Figure 4.3 for the datasets that we analysed.

A considerable advantage of this method is its independence on masses, the limit radius and partly on the observational biases. The only thing, that is considered, is the star density. On the other hand, the result might be misled by some local fluctuation of density. To overcome this problem, we set the number of the closest stars to  $n = 10$ . This was enough for our datasets, because we obtained the same results when setting  $n$  between 10 and 20, while setting it between 5 and 10 gave us different centres.

## 4.3 New coordinates

We can see from our results that all the methods of finding the centre of the ONC led us to the area of the Trapezium association — a place considered to be the core of this star cluster (Hillenbrand & Hartmann, 1998). It was discussed, that the centre of mass is not a reliable solution because we do not have information about mass of every star in each dataset. We can see as well that the centre of mass is shifted from the numerical centre and the density centre (Figure 4.2). For



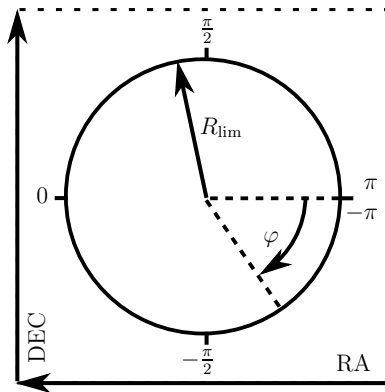
**Figure 4.4:** New coordinate system inside the Orion Nebula Cluster. Its origin is defined as an average from the centres (red is for density centres, green represents numerical centres), that we found using methods described above.

these reasons, we excluded results from this method from further analysis. As we do not see any clear preference between the other approaches (the numerical and the density centre), we took an average from all solutions, that we obtained, as origin of the new coordinate system (see Figure 4.4). We transformed all of our datasets into these newly defined coordinates and all of the following analyses are made with respect to them. The exact position of our new centre of the Orion Nebula Cluster is

$$\begin{aligned} \text{RA}_{\text{new}} &= 83^{\circ} 49' 13.08'' \\ \text{DEC}_{\text{new}} &= -5^{\text{h}} 22^{\text{m}} 57.90^{\text{s}}. \end{aligned} \quad (4.5)$$

## 4.4 Angular profiles

Using the observational data described in Chapter 3 we studied the angular profiles of the ONC up to the defined limit radii ( $R_{\text{lim}}$ ). To create the angular cumulative distribution functions, we ran continuously from  $-\pi$  to  $\pi$  (see Figure 4.5 for our definition of angular limits) through all the stars from the given dataset that are inside a disk of radius  $R_{\text{lim}}$ , and added 1 every time we spotted a star. We tested the rotational symmetry with the KS test at a predetermined confidence level  $Q_{\text{KS}} = 0.05$ . The analysis of the X-ray dataset from Flaccomio et al. (2003a,b) revealed, that the angular profiles are compatible with the homogeneous distribution function up to  $R_{\text{lim}} = 0.7$  pc (see Figure 4.6 and Table 4.1). We also got an agreement with the rotational symmetry from the optical sample (Hillenbrand, 1997) but only up to  $R_{\text{lim}} = 0.5$  pc. Tests of the IR dataset from Da Rio et al. (2012) gave lower confidence level than 0.05, but we can assume



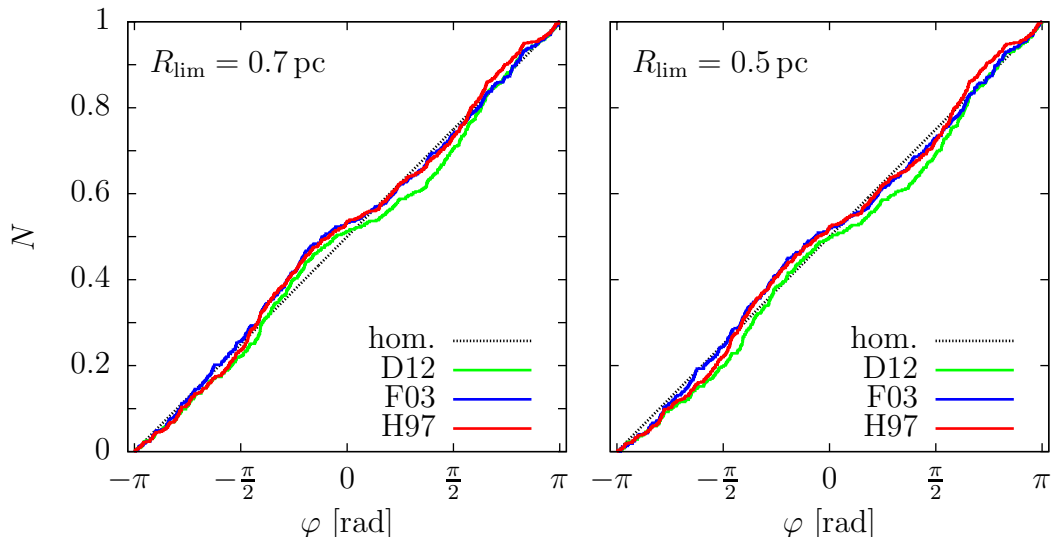
**Figure 4.5:** Definition of the angle  $\varphi$  and the limit radius  $R_{\text{lim}}$  used in the construction of the angular profiles, displayed in the coordinate system.

| data                       | $R_{\text{lim}}$ |        |             |             |             |
|----------------------------|------------------|--------|-------------|-------------|-------------|
|                            | 0.5 pc           | 0.7 pc | 1.0 pc      | 1.5 pc      | 2.0 pc      |
| Flaccomio et al. (2003a,b) | 0.34             | 0.06   | $< 10^{-3}$ | $< 10^{-5}$ | $< 10^{-5}$ |
| Hillenbrand (1997)         | 0.26             | 0.04   | $< 10^{-4}$ | $< 10^{-5}$ | $< 10^{-6}$ |
| Da Rio et al. (2012)       | 0.03             | 0.02   | $< 10^{-3}$ | $< 10^{-3}$ | $< 10^{-7}$ |

**Table 4.1:** The results of Kolmogorov–Smirnov test ( $Q_{\text{KS}}$ ) for the X-ray dataset (Flaccomio et al., 2003a,b), the optical dataset (Hillenbrand, 1997) and the infrared dataset (Da Rio et al., 2012) at the given limit radius  $R_{\text{lim}}$ .

that it is a consequence of some observational bias. The objects measured by Da Rio et al. (2012) are mostly brown dwarfs and pre-main sequence stars, which means light objects, as is presented in Figure 3.3, therefore, harder to detect. From all the results in Table 4.1, we can see that the closer we were to the centre of the ONC, the higher value of  $Q_{\text{KS}}$  we obtained. Deducing from the results above and assuming that there is no physical mechanism, which would distribute sources active in different wavelengths differently, and that our results are influenced only by the extinction, we can extend the rotational symmetry up to the area limited by  $R_{\text{lim}} = 0.7$  pc, where the X-ray sources are showing statistically significant agreement.

Our results derived from the analysis of the optical data correspond to the results of Hillenbrand & Hartmann (1998), who fitted the spherically symmetric models of the star clusters to the optical and IR data and came to a result that the central regions of the ONC are in agreement with spherical symmetry. Further, they stated that the ONC is clearly not spherically symmetric on larger radii than 0.5 pc. Taking into account their conversion between angular and proper distance, 0.5 pc is in our length scale 0.44 pc. However, considering more recent X-ray data, which are also better seen through the nebula (Flaccomio et al., 2003a,b), we found the rotational symmetry to the larger radii. The probability that we observe the ONC from a special direction (e. g. looking on a base of a cylinder) is negligible. Therefore, we assume that the rotational symmetry of the central area of the ONC implies the spherical symmetry in that area too.



**Figure 4.6:** The cumulative distribution functions of Da Rio et al. (2012) — D12, Flaccomio et al. (2003a,b) — F03, Hillenbrand (1997) — H97, and the homogeneous distribution in angle (dashed line). Left panel corresponds to the area with the limit radius 0.7 pc and the right panel is for 0.5 pc.

#### 4.4.1 Centre of symmetry

The results from the previous testing motivated us to formulate an alternative approach to the question of the rotational (perhaps spherical) symmetry. If we assume that the observational data are rotationally symmetric, there must exist a point from which the rotational symmetry originates. We expect from such point to reach the maximal level of confidence when compared to the rotationally symmetric profile. Therefore, we covered the central region of the ONC by a grid and in each of its node we created the angular profile of the star number density. We evaluated the confidence level by the KS test. Because of the computational difficulty, we present only the result for the X-ray dataset (in Figure 4.7). We can see that the other centres of the X-ray dataset are very close to the maximal result of the KS test ( $Q_{KS} \simeq 0.86$ ). This supports the other methods of finding the centre of the ONC and also the conclusion we made on the rotational symmetry.

#### 4.4.2 Large scale shape

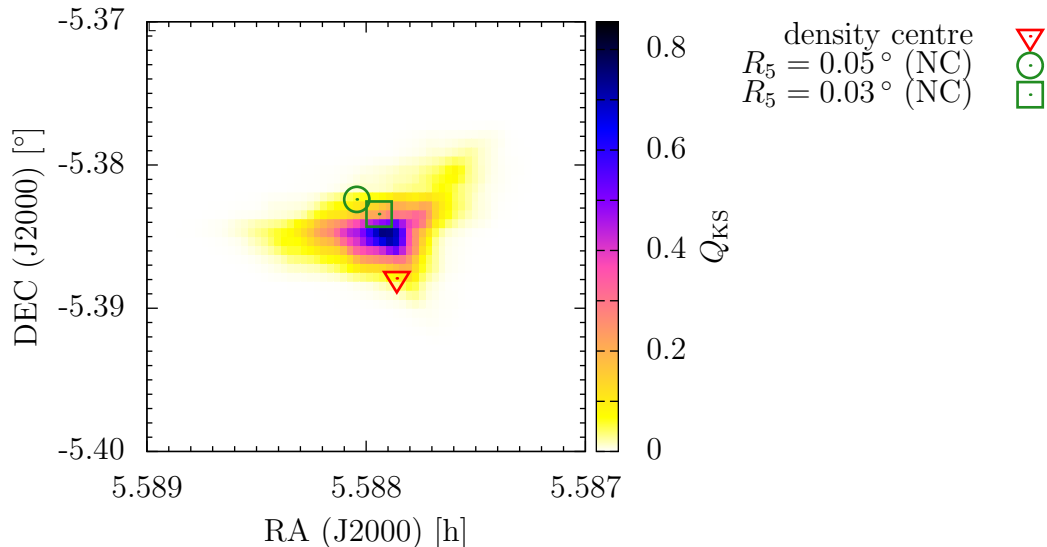
We can see that there is varying extinction in the ONC (Figure 3.1). This fact motivated us to divide this star cluster into the angular sections with roughly homogeneous distribution of stars and use them to analyse the overall shape of the ONC (up to the radius of 2 pc). As the first attempt we cut the ONC into four quadrants

$$Q0: \varphi \in \left(-\pi; -\frac{\pi}{2}\right),$$

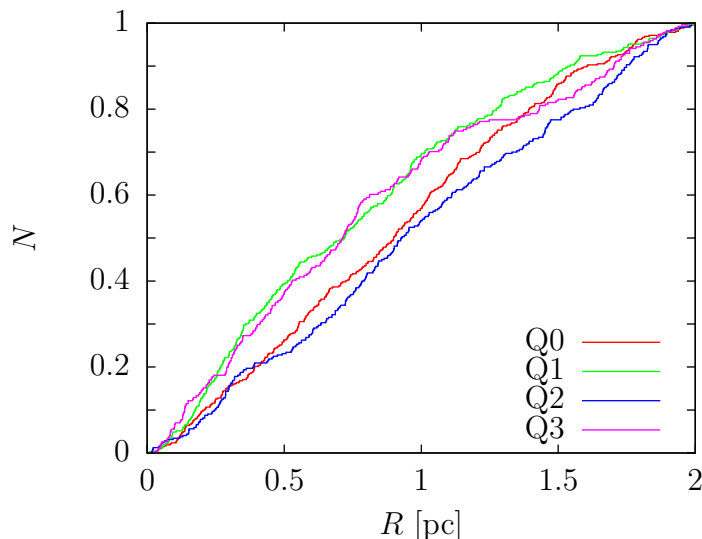
$$Q1: \varphi \in \left(-\frac{\pi}{2}; 0\right),$$

$$Q2: \varphi \in \left(0; \frac{\pi}{2}\right),$$

$$Q3: \varphi \in \left(\frac{\pi}{2}; \pi\right).$$

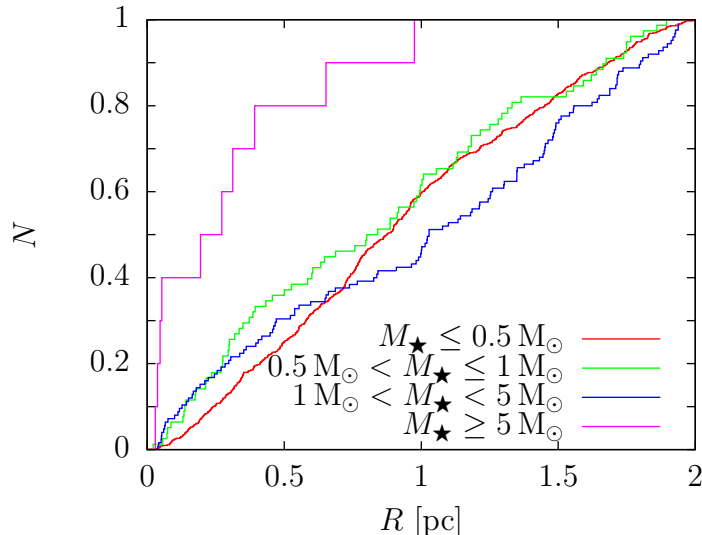


**Figure 4.7:** Visualisation of the results of the KS test on the X-ray data (Flaccomio et al., 2003a,b) in the central region of the ONC. Limit radius, for which the KS test was evaluated is  $R_5 = 0.05^\circ$ . We also plot the density centre (red up side down triangle), numerical centre for  $R_5$  (green circle) and for  $R_3$  (green square).



**Figure 4.8:** The normalised cumulative distribution functions of the count of stars from different quadrants. Data are taken from NAME ONC.

We then used the KS test to compare the cumulative distribution functions of the star radial number density of these quadrants with respect to each other. We found out that the opposite quadrants, i. e. the pair Q0 with Q2 and the pair Q1 with Q3, showed statistically more similar radial profiles than the adjacent quadrants, e. g. Q0 and Q1. Although our division into the quadrants was rough, we can still see, that this results disproves the rotational symmetry of the ONC at large scales. This stays even for the X-ray data on large radii. Instead, our results indicate (with no direct proof) that the whole cluster is rather centrally symmetric (e. g. like an ellipsoid). From the Figure 4.8, where we plot the radial profiles of these quadrants, we can see that the ONC is likely to be elongated from the South-West (Q1) to the North-East (Q3). This result is in agreement



**Figure 4.9:** The radial profiles of different mass groups of the ONC. The curves are made of the stars with known masses from NAME ONC dataset.

with Hillenbrand & Hartmann (1998), who derived to the vertically elongated elliptical shape of the ONC by analysing the distribution of the optical and the infrared sources within the nebula.

## 4.5 Radial profiles and mass segregation

The next step in the analysis of the ONC was to evaluate the level of the mass segregation within this star cluster. We set the limit radius for the whole star cluster to 2 pc. Accordingly to the model analysis (Chapter 2.2.2), we divided the stars (with known masses) from the NAME ONC dataset into four mass groups ( $M_{\star} \leq 0.5 M_{\odot}$ ,  $0.5 M_{\odot} < M_{\star} \leq 1 M_{\odot}$ ,  $1 M_{\odot} < M_{\star} < 5 M_{\odot}$  and  $M_{\star} \geq 5 M_{\odot}$ ) and compared their radial profiles by the KS test. From the cumulative distribution functions in Figure 4.9, we can see that stars with  $M_{\star} \geq 5 M_{\odot}$  are more concentrated to the centre of the ONC and thus mass segregated from the other mass groups. We can base this hypothesis on low compatibility ( $Q_{\text{KS}} < 10^{-3}$ ) of this group with the less massive stars. We do not observe a statistically significant mass segregation between the groups of less massive stars alone for  $R \in [0 \text{ pc}; 2 \text{ pc}]$ .

### 4.5.1 Possible detection of inverse mass segregation

From all tests we realized, only the comparison of the stars with  $1 M_{\odot} < M_{\star} < 5 M_{\odot}$  and the stars with  $M_{\star} \leq 0.5 M_{\odot}$  in the region of  $R \in [0; 2] \text{ pc}$  gave us the result almost equal to the predetermined level of significance of the KS test. This motivated us to focus on the radial profiles of these two groups. From the graphs in Figure 4.9, we can see that the cumulative distribution function of the more massive stars from these two groups is clearly under the cumulative distribution function of the less massive stars at  $R \gtrsim 0.75 \text{ pc}$ . This indicates, that the group

| annulus<br>[pc] | $M_{\star}$ [ $M_{\odot}$ ] |          |        |             |
|-----------------|-----------------------------|----------|--------|-------------|
|                 | less than 0.5               | 0.5 to 1 | 1 to 5 | more than 5 |
| 0 to 0.5        | 182                         | 28       | 38     | 8           |
| 0.5 to 1.0      | 251                         | 20       | 19     | 2           |
| 1.0 to 1.5      | 167                         | 16       | 38     | 0           |
| 1.5 to 2.0      | 124                         | 14       | 30     | 0           |

**Table 4.2:** The counts of stars in different annuli for specific mass groups.

of stars with masses in  $1 M_{\odot} < M_{\star} < 5 M_{\odot}$  might be inversely mass segregated from the group of stars with  $M_{\star} \leq 0.5 M_{\odot}$ .

We can support the above hypothesis in two ways. The first way is by comparing the star count ratios ( $f_{\star}$ ) in the annulus with  $R \in [0.5; 1]$  pc and in the annulus defined by  $R \in [1; 1.5]$  pc.

$$f_{\star}^{M_{\star} \leq 0.5 M_{\odot}} = \frac{N_{\star}^{M_{\star} \leq 0.5 M_{\odot}} (R \in [0.5; 1] \text{ pc})}{N_{\star}^{M_{\star} \leq 0.5 M_{\odot}} (R \in [1; 1.5] \text{ pc})} = \frac{3}{2} \quad (4.6)$$

$$f_{\star}^{1 M_{\odot} < M_{\star} < 5 M_{\odot}} = \frac{N_{\star}^{1 M_{\odot} < M_{\star} < 5 M_{\odot}} (R \in [0.5; 1] \text{ pc})}{N_{\star}^{1 M_{\odot} < M_{\star} < 5 M_{\odot}} (R \in [1; 1.5] \text{ pc})} = \frac{1}{2}, \quad (4.7)$$

From the theory, we expect the ratio of low mass stars, i. e. the equation (4.6), to be lower than the ratio of more massive stars, i. e. equation (4.7). Our results are opposite to that (for the total counts of stars in all mass groups, see Table 4.2). The other way is to determine the statistical weight of this result. We created the cumulative distribution functions of these two mass groups in the annulus with  $R \in [0.5; 1.5]$  pc and compared them by the KS test. We obtained the result, that the probability of these two functions being from the same distribution is around  $2 \times 10^{-4}$ .

The previous argumentation supported our conjecture of the inverse mass segregation. This result goes opposite to the predictions, where the mass segregation is a strong effect, which drives the star cluster evolution (Binney & Tremaine, 1994; Heggie & Hut, 2003). We must, however, be careful when interpreting the results. There are some limitations to the KS test — we can influence the results by selecting the border values, e. g. the radii, or by a low number of stars. There is also no way how to decide between the inverse and classic mass segregation solely from the KS test. We must also analyse the shape of the cumulative distribution functions or evaluate the ratios of stars, as we did.

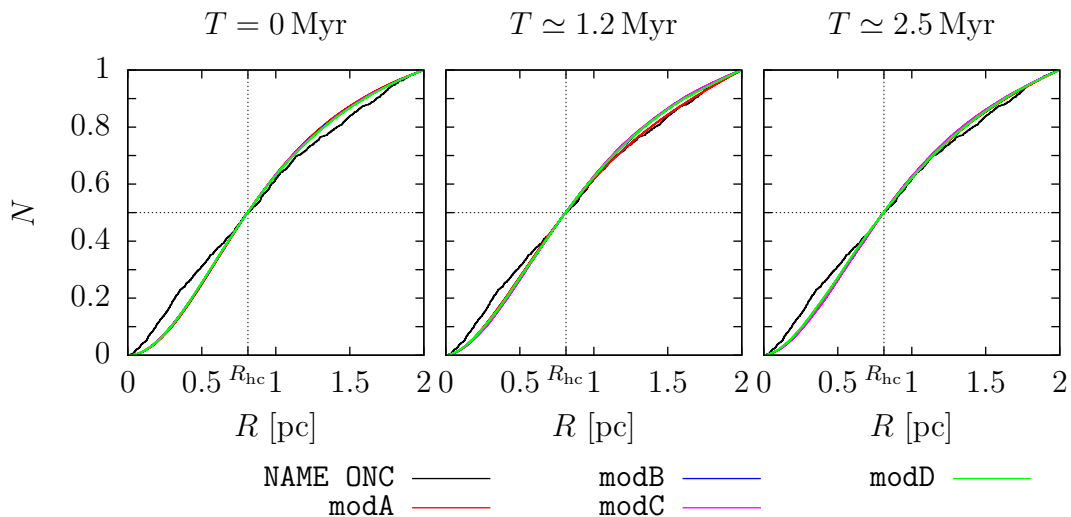
## 4.5.2 Comparison with the numerical models

The following analysis is oriented to compare the models, that we presented in Chapter 2, to the observational data (Chapter 3). To do this comparison properly, we have to scale the models to the same size. We used the half-count radii for the size check. The half-count radius of the ONC is  $R_{\text{hc}} = 0.81$  pc for stars up to 2 pc. Note, that this scaling cannot be done by taking the values of the half-count radii for models from Table 2.2 and multiply all distances in the models of the star cluster by the factor derived from the size of the ONC. The reason was in



| id   | $S$ (0 Myr) | $S$ (1.2 Myr) | $S$ (2.5 Myr) |
|------|-------------|---------------|---------------|
| modA | 7.13        | 2.06          | 1.66          |
| modB | 12.91       | 3.65          | 2.52          |
| modC | 8.31        | 2.81          | 1.99          |
| modD | 12.41       | 3.33          | 2.30          |

**Table 4.3:** The values of the scaling factor  $S$  of the numerical models presented in Chapter 2. The limit radius of the ONC and of the rescaled models is 2 pc.



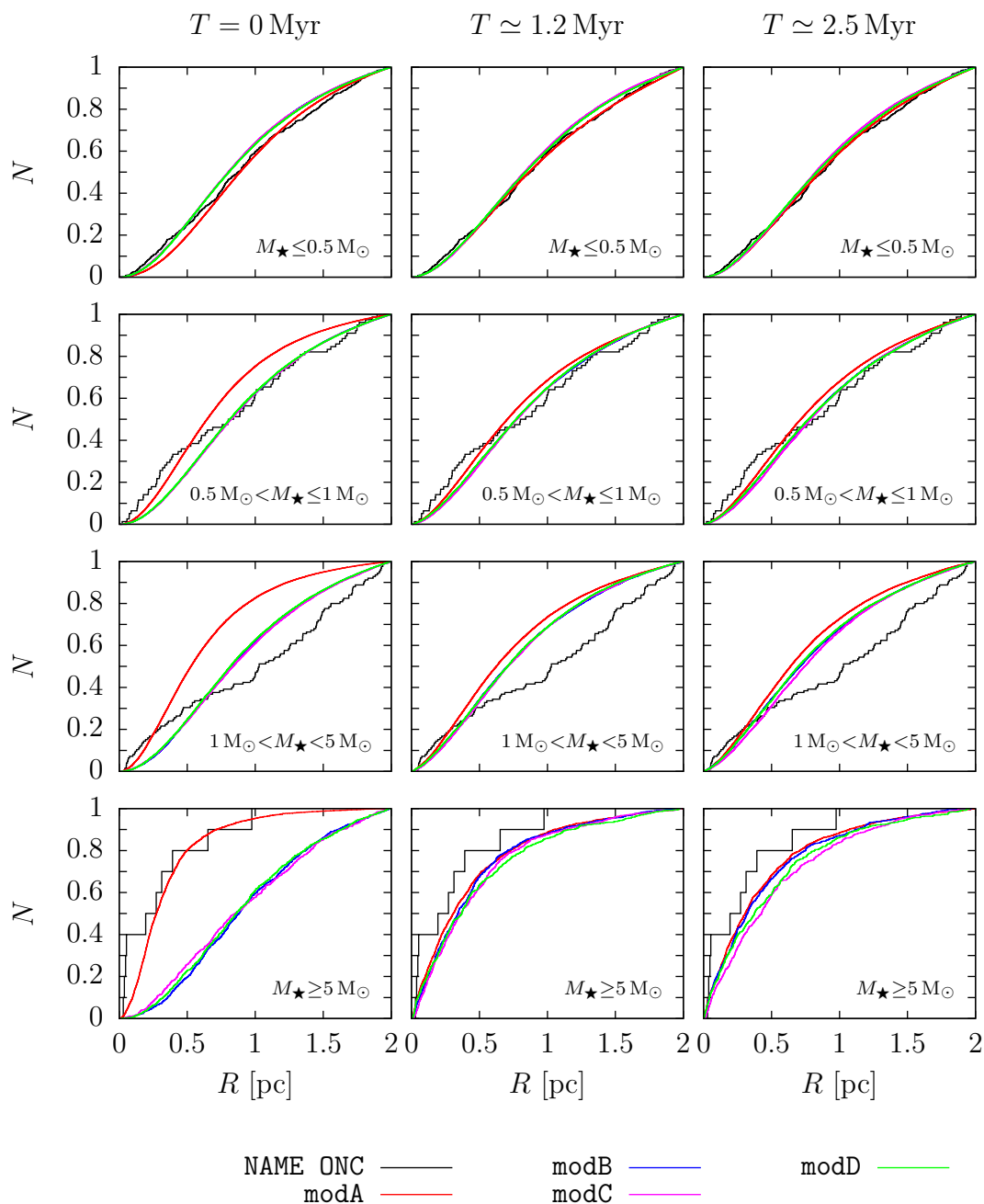
**Figure 4.10:** The cumulative distribution functions of the models and the ONC, fitted to the half-count radius of the ONC,  $R_{\text{hc}} = 0.81$  pc.

the normalisation of the KS test, which we used. Although all half-count radii were originally counted from stars within the same limit radius ( $R_{\text{lim}} = 2$  pc), by rescaling the model, its limit radius increases as well. Therefore, the half-count radii would not be comparable any more. We had to use an iterative process, which determined the limit radius of the models of the star cluster with the following constraint — after rescaling the whole model to the size where the half-count radii of the model and the ONC fit, the new modelled cluster radius must be equal to 2 pc. We are interested only in the time frames  $T \simeq 1.2$  Myr and  $T \simeq 2.5$  Myr. The initial cluster is not adequate for the comparison, because its size has been intentionally defined several times more compact than the ONC. From the values of the scaling factor  $S$  of all models (shown in Table 4.3), we can see that **modA**, which has been published as the best representation of the ONC (Šubr et al., 2012), is in fact the best fitting one. Other models are struggling in fitting by less than a factor of 2, which is crucial for having realistic time evolution of the model (see Chapter 2).

Next we created the cumulative distribution functions of the star radial number density from all stars of the models and all stars in the **NAME ONC** dataset and compared them among themselves. Although the evolution of these functions was roughly the same (as seen in Figure 4.10), we were not able to find any model among the four tested, which would be also statistically comparable.

The possible reason for this is in the inverse mass segregation between the mass group with  $1 M_{\odot} < M_{\star} < 5 M_{\odot}$  and the less massive stars. When analysing

the radial profiles of the same mass groups from the model and the observational data (plotted in Figure 4.11), we came to an agreement in all of the tests, except for the mass group of  $1 M_{\odot} < M_{\star} < 5 M_{\odot}$ , as it has been expected from the previous results.



**Figure 4.11:** Cumulative distribution functions of the models and the ONC for different mass groups in three time frames. We can see the initial mass segregation of the model modA and also the compatibility of measured data with the numerical models, except for the mass group of  $1 M_{\odot} < M_{\star} < 5 M_{\odot}$ .

## 5. Conclusions

In this work, we focused on studying the structure of the Orion Nebula Cluster and comparing the results of numerical simulations with the observational data. First, we analysed the numerical models of this star cluster. The level of the mass segregation has been tested on the evolved models by the Kolmogorov–Smirnov test, using four different mass groups ( $\leq 0.5 M_{\odot}$ ,  $0.5 M_{\odot}$  to  $1 M_{\odot}$ ,  $1 M_{\odot}$  to  $5 M_{\odot}$  and  $\geq 5 M_{\odot}$ ). We revealed, that all four mass groups have different cumulative distribution functions at high level of statistical significance. Also more massive stars were more concentrated than the less massive stars. We can summarize, that all models of this star cluster were mass segregated after their evolution, including those without the initial mass segregation.

When analysing the observational data we had to overcome several issues caused by the observational incompleteness. One of them is varying extinction in the nebula, from  $A_V \simeq 1$  mag up to  $A_V \simeq 10$  mag in certain regions, due to which we cannot see all stars properly. Because the interstellar medium is more transparent in the X-ray than in the visible light, we collected data from the multi-wavelength observations, i. e. IR, optical and X-ray, and tested different characteristics of the ONC on these datasets separately.

We found out that stars from the ONC are distributed with respect to the rotational symmetry (and perhaps spherical symmetry) in the central region. The limit radius of this area is for the optical sources at 0.5 pc, which corresponds to the results from Hillenbrand & Hartmann (1998), who stated that the central regions of the ONC are in agreement with spherical symmetry. Their results further showed that the ONC is not spherically symmetric above 0.44 pc. However, when analysing more recent X-ray sources, we revealed that the area, where the rotational symmetry is statistically significant, has the limit radius of 0.7 pc — almost the half-count radius of the ONC. Under assumption that stars visible in different wavelengths are not distributed differently, we concluded that the ONC is rotationally (spherically) symmetric in the central region of maximum radius 0.7 pc. We also studied the structure of the ONC on larger scales (up to 2 pc). Our test showed that this star cluster is likely to be elongated from the North-East to the South-West, which is also in agreement with the above cited paper.

Next, we focused on the question of the mass segregation in the ONC. We found out, that the stars more massive than  $5 M_{\odot}$  are mass segregated from the less massive stars, as in the models of the star cluster. But the mass segregation among the less massive stars alone is not as evident as in the numerical models. Focussing on the radial profiles, we discovered the potential inverse mass segregation of stars with masses between 1 and  $5 M_{\odot}$  and stars less massive than  $0.5 M_{\odot}$  in the annulus from 0.5 to 1.5 pc. Such result is unexpected, because it was not predicted by the analytical view on the evolution of the star clusters nor by its implementation into the  $N$ -body simulations that we used. Among the works that studied the mass segregation in the ONC (e. g. Allison et al., 2009; Hillenbrand & Hartmann, 1998), none discussed the possibility that this star cluster is partly inversely mass segregated.

We believe that the inverse mass segregation may be caused by an uneven star formation in the outer regions of the ONC, perhaps due to the gas expulsion. To support this conjecture, we will have to consider in the numerical models the varying star formation in space and time. We can improve in that manner the existing  $N$ -body models that we used in this thesis. One possible way to do it is to concentrate on the local density of gas particles. We then set a condition according to which some gas particles will transform into one star particle, if their local density reaches a predetermined value.

# Bibliography

- Aarseth, S. J. (2003). *Gravitational N-Body Simulations*. Cambridge University Press, Cambridge, UK.
- Allison, R. J., Goodwin, S. P., Parker, R. J., Portegies Zwart, S. F., de Grijs, R., and Kouwenhoven, M. B. N. (2009). *Using the minimum spanning tree to trace mass segregation*. MNRAS, **395**:1449–1454.
- Alves, J. and Bouy, H. (2012). *Orion revisited I.: The massive cluster in front of the Orion nebula cluster*. A&A, **547**:97.
- Baraffe, I., Chabrier, G., Allard, F., and Hauschildt, P. H. (1998). *Evolutionary models for solar metallicity low-mass stars: mass-magnitude relationships and color-magnitude diagrams*. A&A, **337**:403–412.
- Baumgardt, H. and Kroupa, P. (2007). *A comprehensive set of simulations studying the influence of gas expulsion on star cluster evolution*. MNRAS, **380**:1589–1598.
- Binney, J. and Tremaine, S. (1994). *Galactic Dynamics*. Princeton University Press, Princeton, New Jersey.
- Casertano, S. and Hut, P. (1985). *Core radius and density measurements in N-body experiments Connections with theoretical and observational definitions*. ApJ, **298**:80–94.
- Da Rio, N., Robberto, M., Hillenbrand, L. A., Henning, T., and Stassun, K. G. (2012). *The Initial Mass Function of the Orion Nebula Cluster across the H-burning Limit*. ApJ, **748**:14.
- Da Rio, N., Robberto, M., Soderblom, D. R., Panagia, N., Hillenbrand, L. A., Palla, F., and Stassun, K. (2009). *A Multi-color Optical Survey of the Orion Nebula Cluster I.: The Catalog*. ApJS, **183**:261–277.
- D’Antona, F. and Mazzitelli, I. (1994). *New pre-main-sequence tracks for M less than or equal to 2.5 solar mass as tests of opacities and convection model*. ApJS, **90**:467–500.
- D’Antona, F. and Mazzitelli, I. (1998). *A Role for Superadiabatic Convection in Low Mass Structures?* Brown Dwarfs and Extrasolar Planets, Astronomical Society of the Pacific Conference Series, **134**:442.
- Ezer, D. and Cameron, A. G. W. (1967). *Early and main sequence evolution of stars in the range 0.5 to 100 solar masses*. Canadian Journal of Physics, **45**:3429.
- Flaccomio, E., Damiani, F., Micela, G., Sciortino, S., Harnden, Jr., F. R., Murray, S. S., and Wolk, S. J. (2003a). *Chandra X-Ray Observation of the Orion Nebula Cluster. I. Detection, Identification, and Determination of X-Ray Luminosities*. ApJ, **582**:382–397.

- Flaccomio, E., Damiani, F., Micela, G., Sciortino, S., Harnden, Jr., F. R., Murray, S. S., and Wolk, S. J. (2003b). *Chandra X-Ray Observation of the Orion Nebula Cluster. II. Relationship between X-Ray Activity Indicators and Stellar Parameters*. ApJ, **582**:398–409.
- Genzel, R., Reid, M. J., Moran, J. M., and Downes, D. (1981). *Proper motions and distances of H<sub>2</sub>O maser sources. I - The outflow in Orion-KL*. ApJ, **244**:884–902.
- Heggie, D. and Hut, P. (2003). *The Gravitational Million-Body Problem: A Multidisciplinary Approach to Star Cluster Dynamics*. Cambridge University Press, Cambridge, UK.
- Hillenbrand, L. A. (1997). *On the Stellar Population and Star-Forming History of the Orion Nebula Cluster*. AJ, **113**:1733–1768.
- Hillenbrand, L. A. and Hartmann, L. W. (1998). *A Preliminary Study of the Orion Nebula Cluster Structure and Dynamics*. ApJ, **492**:540.
- Huff, E. M. and Stahler, S. W. (2006). *Star Formation in Space and Time: The Orion Nebula Cluster*. ApJ, **644**:355–363.
- Jeffries, R. D. (2007). *The distance to the Orion Nebula cluster*. MNRAS, **376**:1109–1119.
- Jones, B. F. and Walker, M. F. (1988). *Proper motions and variabilities of stars near the Orion Nebula*. AJ, **95**:1755–1782.
- Kroupa, P. (2001). *On the variation of the initial mass function*. MNRAS, **322**:231–246.
- Marks, M. and Kroupa, P. (2012). *Inverse dynamical population synthesis. Constraining the initial conditions of young stellar clusters by studying their binary populations*. A&A, **543**:A8.
- Menten, K. M., Reid, M. J., Forbrich, J., and Brunthaler, A. (2007). *The distance to the Orion Nebula*. A&A, **474**:515–520.
- O’Dell, C. R. and Henney, W. J. (2008). *High Spatial Velocity Features in the Orion Nebula*. AJ, **136**:1566–1586.
- Palla, F., Randich, S., Flaccomio, E., and Pallavicini, R. (2005). *Age Spreads in Star-forming Regions: The Lithium Test in the Orion Nebula Cluster*. ApJ, **626**:L49–L52.
- Palla, F. and Stahler, S. W. (1999). *Star Formation in the Orion Nebula Cluster*. ApJ, **525**:772–783.
- Plummer, H. C. (1911). *On the problem of distribution in globular star clusters*. MNRAS, **71**:460–470.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (third edition, 2007). *Numerical Recipes — The Art of Scientific Computing*. Cambridge University Press, Cambridge, New York.

- Robberto, M., Soderblom, D. R., Bergeron, E., Kozhurina-Platais, V., Makidon, R. B., McCullough, P. R., McMaster, M., Panagia, N., Reid, I. N., Levay, Z., Frattare, L., Da Rio, N., Andersen, M., O'Dell, C. R., Stassun, K. G., Simon, M., Feigelson, E. D., Stauffer, J. R., Meyer, M., Reggiani, M., Krist, J., Manara, C. F., Romaniello, M., Hillenbrand, L. A., Ricci, L., Palla, F., Najita, J. R., Ananna, T. T., Scandariato, G., and Smith, K. (2013). *The Hubble Space Telescope Treasury Program on the Orion Nebula Cluster*. *ApJS*, **207**:10.
- Salpeter, E. E. (1955). *The Luminosity Function and Stellar Evolution*. *ApJ*, **121**:161.
- Scandariato, G., Robberto, M., Pagano, I., and Hillenbrand, L. A. (2011). *The extinction map of the OMC-1 molecular cloud behind the Orion nebula*. *A&A*, **533**:A38.
- Siess, L., Dufour, E., and Forestini, M. (2000). *An internet server for pre-main sequence tracks of low- and intermediate-mass stars*. *A&A*, **358**:593–599.
- Šubr, L., Kroupa, P., and Baumgardt, H. (2008). *A new method to create initially mass segregated star clusters in virial equilibrium*. *MNRAS*, **385**:1673–1680.
- Šubr, L., Kroupa, P., and Baumgardt, H. (2012). *Catch Me If You Can: Is There a “Runaway-mass” Black Hole in the Orion Nebula Cluster?* *ApJ*, **757**:37.
- Swenson, F. J., Faulkner, J., Rogers, F. J., and Iglesias, C. A. (1994). *The Hyades lithium problem revisited*. *ApJ*, **425**:286–302.
- Tobin, J. J., Hartmann, L., Fűrész, G., Mateo, M., and Megeath, S. T. (2009). *Kinematics of the Orion Nebula Cluster: Velocity Substructure and Spectroscopic Binaries*. *ApJ*, **697**:1103–1118.
- Walker, M. F. (1969). *Studies of extremely young clusters. V. Stars in the vicinity of the Orion nebula*. *ApJ*, **155**:447.
- Warren, Jr., W. H. and Hesser, J. E. (1977). *A photometric study of the Orion OB 1 association. 2: Photometric analysis*. *ApJS*, **34**:207–231.
- Wilson, T. L., Filges, L., Codella, C., Reich, W., and Reich, P. (1997). *Kinematics and electron temperatures in the core of OrionA*. *A&A*, **327**:1177–1184.

# Appendix

## A Kolmogorov–Smirnov test

For analysing the Orion Nebula Cluster, e. g. finding symmetries and comparing radial profiles of the observational data and the numerical model, we used the statistical one-dimensional Kolmogorov–Smirnov test (KS test), which gives the confidence level ( $Q_{\text{KS}}$ ) of the tested hypothesis. In this appendix, we want to illustrate the use of KS test in this work and connect the method to a program written in programming language `awk` (inspired by Press et al., 2007).

There is only one requirement for the input data — that they are sorted in ascending order by the value, we want to test. Sorting data can be done by external tools, such as `sort` (in Linux), so it is not part of any source code listed below.

The confidence level  $Q_{\text{KS}}$  can be calculated from its complement  $P_{\text{KS}}$  defined as

$$P_{\text{KS}}(z) = \frac{\sqrt{2\pi}}{z} \sum_{j=1}^{\infty} \exp\left(-\frac{(2j-1)^2\pi^2}{8z^2}\right) = 1 - Q_{\text{KS}}(z), \quad (\text{A.1})$$

where the argument  $z$  derives from the maximal difference between two normalised cumulative distribution functions ( $D_{\text{max}}$ ), and the effective number of data points ( $N_{\text{eff}}$ ) as

$$z = D_{\text{max}} \left( \sqrt{N_{\text{eff}}} + 0.12 + \frac{0.11}{\sqrt{N_{\text{eff}}}} \right) \quad (\text{A.2})$$

(see Press et al., 2007, and references in there).

### A.1 One sample testing

We must initialize several values. The total number of entries `n` for later normalisation, two variables which represent the cumulative distribution functions. One for the sum of data points (`data`) and second for the sum of theoretical points (`dist`). Then we define the current difference between the data and the distribution, called `D`, the maximal value `Dmax` of this difference, the argument `z` for the KS statistics, variable `e` to store the value of the exponential function and of course `Q` and `P`, which are the results.

```
BEGIN{
  n = 0;
  data = 0.; dist = 0.;
  D = 0.; Dmax = 0.;
  z = 0.;
  e = 0.;
  Q = 0.; P = 0.;
  limit = 0.;
}
```



In our case, we often set a limit radius around the data, in which we tested the null hypothesis (see the parameter `limit`). If its value is equal to zero (as in this example), the border is unset and we test the whole input datafile. Radius was stored in the fourth column of the input file, which is why we compare `limit` to `$4`.

The main program will now read the input file line by line and store necessary data into the array `var[]` — for us, it was the information about angle, which was in the third column, hence `$3`. Useful trick is written on the first line, which tells the program to skip empty lines and lines beginning by `#` in the input datafile (e. g. for the possibility of having comments in there).

```
NF && $1!~/^#/{
  if (limit)
  {
    if ($4 <= limit)
    {
      var[n] = $3;
      n++;
    }
  }
  else
  {
    var[n] = $3;
    n++;
  }
}
```

Before proceeding to the KS test, two additional functions are needed. By definition, we shall put them either at the beginning or at the end of the source code, but the logical steps require to place them here. First one is the absolute value.

```
function abs(x){
  return ((x < 0.0) ? -x : x);
}
```

The second one — also the most important one — is evaluating the distribution function for a homogeneous disk at any given angle `x` in radians inside the interval  $[-\pi; \pi)$ .

```
function f(x){
```

Beginning by the definition of  $\pi$

```
  PI = atan2(0, -1);
```

recalculating the angle for the interval  $[0; 2\pi)$ , because it is easier to work with positive numbers.

```
  x += PI;
```

And finally return how full the disk should be in theory.

```

return (x / (2 * PI));
}

```

Now, for the data, that we gathered in the main part of our program, we want to calculate the absolute difference from the corresponding theoretical value.

```

END{
  for (i = 0; i < n; i++)
  {
    dist = f(var[i]);

```

Note, that the difference for one theoretical value `dist` has to be calculated before ...

```

    D = abs(dist - data);
    Dmax = (Dmax >= D ? Dmax : D);

    data = i / n;

```

... and after increasing the cumulative function `data` by one step.

```

    D = abs(dist - data);
    Dmax = (Dmax >= D ? Dmax : D);
  }

```

We now use equation (A.2), where the effective number of data points  $N_{\text{eff}}$  is equal to the actual number of data points  $n$ . Then we define  $\pi$  and finally set the summation upper limit  $N$  to “almost” infinity — the series (A.1) converges quickly, which means that bigger number than 200 is not needed. The exponential function in `awk` does not support very small numbers and returns zeros instead. This is why there is the command to break the sum, when this occurs, mainly due to the reduction of the computational time.

```

z = (sqrt(n) + 0.12 + (0.11 / sqrt(n))) * Dmax

PI = atan2(0, -1);

N = 200;
for (i = 1; i <= N; i++)
{
  e = exp(0. - (((2. * i - 1.)**2) * PI * PI) / (8. * z * z));
  if (e > 0) P += sqrt(2. * PI) / z * e;
  else break;
}

```

Finally we calculate the complement of  $P_{\text{KS}}$ .

```

Q = 1. - P;

```

The only thing left is to print the results.

```

print "D_max = " Dmax ", n_tot = " n;
print "z = " z ", P = " P ", Q = " Q;
}

```

At this point, we can compare the value of  $Q_{\text{KS}}$ , to the predetermined significance level, that we wanted to achieve.

## A.2 Two sample testing

Instead of comparing one dataset to a theoretical distribution, we can compare two different datasets together. Here lies one major advantage of the KS test — we do not need to modify the data in any way, e. g. same binning. But the requirement, which was mentioned before, that data have to be sorted in ascending order of the value, we wish to test, remains.

At the beginning, we need few definitions. The number of stars in both datasets (`n1`, `n2`), the value of the cumulative distribution functions (`data1`, `data2`) and two variables for later use (`d1`, `d2`). Other declarations are the same as in the one sample testing.

```
BEGIN{
  n1 = 0; n2 = 0;
  data1 = 0.; data2 = 0.;
  d1 = 0.; d2 = 0.;
  D = 0.; Dmax = 0.;
  z = 0.;
  e = 0.;
  Q = 0.; P = 0.;
  limit = 0.05;
}
```

In this case we set a value to `limit` to compare both datasets only up to a certain radius, here 0.05, and instead of the angle we compare radii (in the fourth column, \$4). Arrays for storing the value to compare are called `var1[]` and `var2[]`. Now we proceed to read the first file. That is controlled by the condition that the number of records read from the current file (`FNR`) has to be equal to the total number of records read so far (`NR`).

```
(FNR == NR) && NF && $1!~/^#/{
  if (limit)
  {
    if ($4 <= limit)
    {
      var1[n1] = $4;
      n1++;
    }
  }
  else
  {
    var1[n1] = $4;
    n1++;
  }
}
```

At this point, we can read the second file. Look at the condition, that total number of records is not equal to the number of records in the current file.

```

(FNR != NR) && NF && $1!~/^#/{
  if (limit)
  {
    if ($4 <= limit)
    {
      var2[n2] = $4;
      n2++;
    }
  }
  else
  {
    var2[n2] = $4;
    n2++;
  }
}

```

Now we define two indices `i1` and `i2`, which will be used to go through our arrays.

```

END{
  i1 = 0;
  i2 = 0;

  while ((i1 < n1) && (i2 < n2))
  {

```

This statement says, that for the purpose of finding the biggest difference between two datasets, there is no need to continue after one of them achieves its maximal value — that function will stay constant, while the other will grow, so the difference will only decrease. Now we save both radii in `d1` and `d2` respectively and work with the lower one.

```

    if ((d1 = var1[i1]) <= (d2 = var2[i2]))
    {
      do
      {
        i1++;
        data1 = i1 / n1;

```

We increase the value of its distribution function until the radius is the same (or we still have data points).

```

      } while ((i1 < n1) && (d1 == var1[i1]));
    }
    if (d1 >= d2)
    {

```

If `d1` was not smaller than `d2`, this part would have been executed.

```

      do
      {

```

```

    i2++;
    data2 = i2 / n2;
} while ((i2 < n2) && (d2 == var2[i2]));
}

```

At the end, we calculate the actual difference and compare it to the maximum. Then we loop the previously mentioned steps.

```

    D = abs(data1 - data2);
    Dmax = (Dmax > D ? Dmax : D);
}

```

There is one difference in calculating the effective number of points. We use the following equation

$$N_{\text{eff}} = \frac{N_1 N_2}{N_1 + N_2}, \quad (\text{A.3})$$

where  $N_1$  and  $N_2$  correspond to  $n1$  and  $n2$  respectively. Finally we use equations (A.2) and (A.1) and print out the results in order to compare them later.

```

n = (n1 * n2) / (n1 + n2);

z = (sqrt(n) + 0.12 + (0.11 / sqrt(n))) * Dmax;

PI = atan2(0, -1);

N = 200;
for (i = 1; i <= N; i++)
{
    e = exp(0. - (((2. * i - 1.)**2) * PI * PI) / (8. * z * z));
    if (e > 0) P += sqrt(2. * PI) / z * e;
    else break;
}
Q = 1. - P;

print "# n1_tot = " n1 ", n2_tot = " n2;
print "# D_max = " Dmax ", n_eq = " n;
print "# z = " z ", P = " P ", Q = " Q;
}

```

## B Centre of the cluster

In this appendix we show the source code, that was used to calculate different types of centres of the ONC. In order to give the following explanation a cleaner look, all the bits, where input or output files are managed are intentionally excluded (the complete source code is attached on CD).

### B.1 Centre of mass / Numerical centre

The following function, which computes the position of the centre of mass or the numerical centre (see Chapter 4), needs some parameters. First is the number of lines (entries or stars) in the input data file (`no1`). As it is described in the main text, we searched for a cluster centre either in a given limit radius (`limR`) or in the area, where is the specified fraction of stars (`per`, number between 0 and 100). The number of iterations is `iter` — there is no need for setting this number very high, because the iterative process converges quickly (in few steps).

We also define some local variables, such as the summation index `i`, right ascension and declination of the centre (`c_ra` and `c_dec` respectively), total mass `M_tot` of the cluster and the number `arr`, which will later determine the last star, that is included in the calculation. For the first estimate, we use every star in the input dataset, which is why it is set to `no1`.

```
int mass_centre(int no1, float per, int iter, float limR)
{
    int i;
    float c_ra, c_dec, M_tot;
    int arr = no1;
```

Before we proceed to the iteration itself, we must present one last unknown — the array `stars[no1][3]`. The first dimension, `stars[i][ ]`, serves for counting stars and the second one has been filled by their properties, namely the right ascension (`stars[i][0]`), declination (`stars[i][1]`) and mass, or the number 1, if we are calculating the numerical centre (`stars[i][2]`).

```
while (iter >= 0)
{
    c_ra = 0.;
    c_dec = 0.;
    M_tot = 0.;
    for (i = 0; i < arr; i++)
    {
        c_ra += stars[i][0] * stars[i][2];
        c_dec += stars[i][1] * stars[i][2];
        M_tot += stars[i][2];
    }
    c_ra /= M_tot;
    c_dec /= M_tot;
```

At this place we can print out the results from each step of the iteration, which is not shown here. The following lines will tell us the number of stars (`arr`),

for which we wanted to calculate the centre initially (depending on the method used). Both of the following functions will be explained afterwards.

```

    if (limR > 0.) arr = limit_R(limR, nol, c_ra, c_dec);
    else arr = rl(per, nol, c_ra, c_dec);

```

The number of steps left is decreased by 1.

```

    iter--;
}

```

At the end of this function, we shall return some integer, for example `nol`.

```

return nol;
}

```

We presented two functions to deduce the number of stars, for whose we calculated the centre of mass (or numerical centre). Here they are explained.

The function used to deduce the number of stars inside the limit radius needs the number `limR`, then the total number of stars (`nol`) and position of the current centre (`c_ra, c_dec`). All stars have to be sorted in ascending order of radius from the current centre, which can be done by various methods — for example a bubble sort (here called `sort_r()`).

```

int limit_R(float limR, int nol, float c_ra, float c_dec)
{
    sort_r(nol, c_ra, c_dec);

    int i;

    for (i = 0; i < nol; i++)
    {

```

Running through all the stars, it finds the first, which is on bigger radius than `limR`. Then it breaks this loop, because it is not necessary to continue, and returns its index — the cycle in `mass_centre()` runs only for `i` smaller than `arr`, which is why, we returned the index of the first star, which is not inside the radius.

```

        if (limR < sqrt((stars[i][0] - c_ra) * (stars[i][0] - c_ra) +
            (stars[i][1] - c_dec) * (stars[i][1] - c_dec)))
        {
            break;
        };
    }

    return i;
}

```

The other method is based on calculating the Lagrange radius (radius containing given percentage of the total mass). We parse the fraction, the number of stars and coordinates of the current centre. Data must be sorted in ascending order of radius as well as in the previous function. Then we define mass of the whole sample `M_tot`, mass `M_lagr` comprised inside the examined area (or equally the number of stars, if the mass is set to 1) and index `i`.

---

```

int rl(float per, int nol, float c_ra, float c_dec)
{
    sort_r(nol, c_ra, c_dec);

    float M_tot = 0.;
    float M_lagr = 0.;
    int i;

    for (i = 0; i < nol; i++) M_tot += stars[i][2];

```

Note that this value of `M_tot` is different from the `M_tot` used in `mass_centre()`. Earlier it stood for a mass given by the first `arr` stars, whereas now it means the mass of the whole dataset. We proceed to the computation of the Lagrange radius. The next few lines are determining whether the mass of the `i`-th star plus the mass cumulated so far is still less or equal to the wanted percentage of the total mass. If not, the loop breaks and the current index is returned to be stored in `arr`.

```

    for (i = 0; i < nol; i++)
    {
        if ((M_lagr + stars[i][2]) <= (per / 100. * M_tot))
        {
            M_lagr += stars[i][2];
        }
        else break;
    }

    return i;
}

```

## B.2 Density centre

Another approach to find the centre, discussed in Chapter 4, is via the star density. Similarly to the centre of mass, one of the input parameters for this function is the total number of stars in the dataset (`nol`). No other parameter, except the number of the closest stars `n`, is needed for this function to work. Before we start explaining the code, several declarations have to be done. All the integers are defined in order to count elements in the arrays, then we need an array `star_min[]`, which will store the radii of `n` closest stars. Variable `min` will be used to save the smallest distance to the `n`-th star, hence the reason to initialize it at a very high value. Finally, the coordinates of the centre (`c_ra` and `c_dec`), differences in right ascension and declination and one dummy variable to store whatever is needed.

```

int density_centre(int n, int nol)
{
    int i, j, k, l, dn, count;
    float *star_min, min = INT_MAX;
    float c_ra, c_dec;

```



```
float dra, ddec, dummy;
```

```
star_min = (float*)malloc(n * sizeof(float));
```

For every single star in the dataset (index  $i$ ), the algorithm will run through all other stars (index  $j$ ),

```
for (i = 0; i < nol; i++)
{
    count = 0;
    for (j = 0; j < nol; j++)
    {
        if (j != i)
        {
```

calculate the distance from the  $i$ -th to the  $j$ -th star, not including itself,

```
dra = stars[i][0] - stars[j][0];
ddec = stars[i][1] - stars[j][1];
dummy = sqrt(dra * dra + ddec * ddec);
```

and put the first  $n$  stars (theirs distances from the  $i$ -th one) into the array `star_min[]` in ascending order, which is why we set `count` to zero earlier.

```
if (count < n)
{
    star_min[count] = dummy;
    count++;
    for (k = 0; k < count; k++)
    {
        for (l = k + 1; l < count; l++)
        {
            if (star_min[k] > star_min[l])
            {
                dummy = star_min[k];
                star_min[k] = star_min[l];
                star_min[l] = dummy;
            }
        }
    }
}
```

For the remaining stars, it tests if they are closer to the  $i$ -th star than the others, previously saved.

```
else
{
    for (dn = 0; dn < n; dn++) if (dummy < star_min[dn])
    {
```

When this occurs, it will put that star into `star_min[]` at the correct position — to maintain the ascending order of the calculated distance — and move the others inside the array one position up, while discarding the current outermost star.

```

        for (k = n - 1; k > dn; k--)
        {
            star_min[k] = star_min[k - 1];
        }
        star_min[dn] = dummy;
        break;
    }
}
};
}

```

At this point, we shall have the closest `n` stars to the `i`-th one saved in `star_min[]`. Now we will compare the biggest distance in this array (the last record) to the current minimal distance stored in `min`. If it is smaller, we rewrite `min`, if not, nothing happens. In either case, we proceed to the next star in the dataset and do this procedure again. Note, that it is not strictly necessary to initialize `min` to a very high value at the beginning, because the following condition will be true in the first step either way.

```

    if ((star_min[count - 1] < min) || (i == 0))
    {
        c_ra = stars[i][0];
        c_dec = stars[i][1];
        min = star_min[count - 1];
    }
}

```

Before ending this function, we must free the memory of `star_min[]` and return an integer value.

```

    free(star_min);

    return nol;
}

```