

Galaxy Generation

Emile Hansmaennel

Theodor Fliedner Gymnasium
emile.hansmaennel@gmail.com



Zusammenfassung

Das Ziel meines Projektes war es, Galaxien drei-dimensional zu visualisieren. Dazu verwendete ich das sogenannte "Navarro-Frenk-White" Profil als Dichtefunktion in Kombination mit der Random-Sampling Methode zum Generieren von Koordinaten.

Die Generierung der Koordinaten hat jedoch ein Problem: Es ist extrem rechenaufwendig. Um dieses Problem zu lösen, habe ich mehrere Methoden zur Optimierung angewendet, darunter die Nutzung von Lookup-Tabellen und die Verwendung von mehreren Computer-Rechenkernen.

Im Verlauf des Projektes habe ich viel Neues dazugelernt, darunter die Handhabung mit großen Datenmengen und die Organisation eines größeren Projektes, jedoch auch mit vielen tiefer gehenden Funktionen und Bibliotheken in der Programmiersprache Python umzugehen.

An das Praktikum im Astronomischen Recheninstitut in Heidelberg gelangte ich, indem ich mein Jugend Forscht Projekt aus dem letztem Jahr (Satellite Computation), auf der Social-Media Plattform Reddit hochgeladen habe, wonach mich der Doktorand Tim Tugendhat anschrieb, ob ich Interesse an einem Praktikum hätte, das ich im Sommer 2017 absolviert habe.

Hauptziele

1. Generieren von elliptischen Galaxien mithilfe des Navarro-Frenk-White-Profiles in Verbindung mit der Random-Sampling Methode
2. Verbesserung des Generierungsprozesses mithilfe verschiedener Methoden
 - Lookup-Tabellen
 - Unterteilung der Galaxie in verschiedene Zellen
 - Effizientere Speichernutzung
3. Generierung von Dunkle-Materie-Halos und Spiralgalaxien
4. Nutzung von konkurrierenden neuronalen Netzen zum unbeaufsichtigten Generieren von Galaxien und deren Anpassung an "echte" Galaxien.

Nutzen

- Visualisierung zur Verbesserung der Vorstellung von Galaxien
- Möglichkeit, sich eine Galaxie aus allen Perspektiven anzusehen
- Modell einer Galaxie, welches genutzt werden kann, um Vorhersagen für die echte Galaxie zu treffen

Methoden

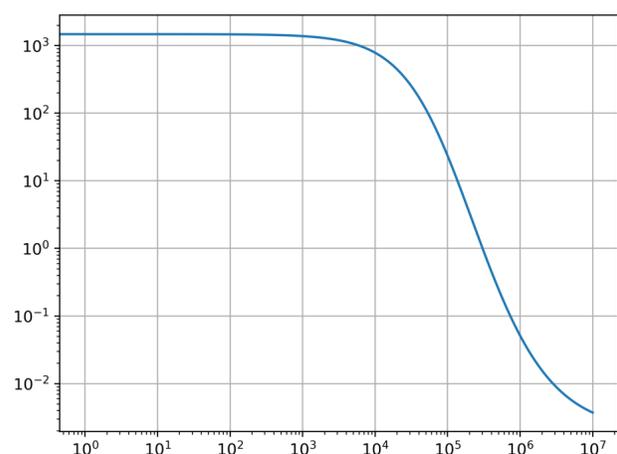
- Random Sampling des NFW-Profiles
- Generierung von Lookup-Tabellen
- Code Vectorization
- Parallelization
- Kräfteberechnung mithilfe der universellen Gravitationsformel

Random-Sampling des NFW Profils

Beim Random-Sampling wird eine zufällige Koordinate generiert und dessen Abstand zum Mittelpunkt der Galaxie r berechnet. Diesen Wert r setzt man in das NFW-Profil ein, um eine Wahrscheinlichkeit $\phi = \rho(r)$ zu berechnen, mit der der Stern generiert wird. Es wird ein zufälliger Wert s im Intervall $[\rho_{min}; \rho_{max}]$ generiert. Gilt $s < r$, dann werden die Koordinaten des Sterns beibehalten, ansonsten werden neue Koordinaten generiert und der Prozess wird wiederholt.

$$\rho_{NFW}(r) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \cdot \exp\left(\frac{-\phi(r)}{\sigma^2}\right) \quad (1)$$

$$\phi(r) = \frac{4\pi \cdot G \cdot f_0 \cdot R_s^3}{r} \cdot \ln\left(1 + \frac{r}{R_s}\right) \quad (2)$$



Der entsprechende Funktionsgraph zum NFW-Profil

Lookup Tabellen

Um den Rechenaufwand der beim Berechnen eines Wertes aus dem NFW-Profil entsteht zu mindern, wird die Funktion $\rho(r)$ im Vorhinein berechnet und zum Weiterverwenden gespeichert:

2e8.csv (~ 500 MB)

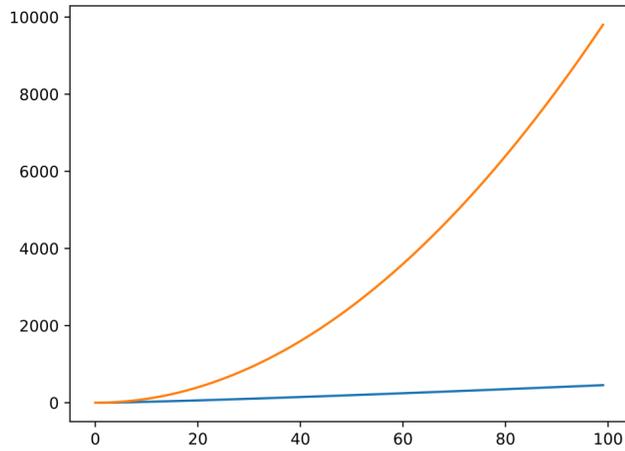
```
0, 1477.1586582000994
1, 1477.0588424006478
2, 1476.9590343243835
3, 1476.8592346184294
4, 1476.7594429495975
...
19999995, 0.0028544345590963767
19999996, 0.0028544345175450904
19999997, 0.0028544344759938085
19999998, 0.002854434434442531
19999999, 0.002854434392891257
```

Möchte man zu einem Wert r aus der Funktion die Wahrscheinlichkeit $\rho(r)$ erhalten, muss man nur noch aus der Tabelle ablesen. Hier entsteht jedoch ein Problem: Alle Prozesse müssen aus dieser Tabelle die Werte auslesen, weshalb sie für jeden Prozess einmal in den Arbeitsspeicher geladen werden müssen. Dies lässt sich jedoch durch effizientes Parallelisieren umgehen.

Rechenaufwand (n-Körper Problem)

$$O(n) = n^2 \Rightarrow O(n) = n \cdot \log(n) \quad (3)$$

Um die Kräfte die zwischen allen Sternen in einer Galaxie wirken zu berechnen werden n^2 Rechenschritte benötigt (n entspricht der Anzahl der) Sterne.



Orange n^2 , Blau $n \log(n)$

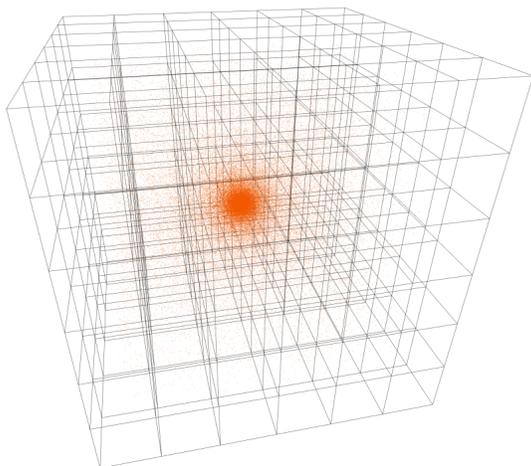
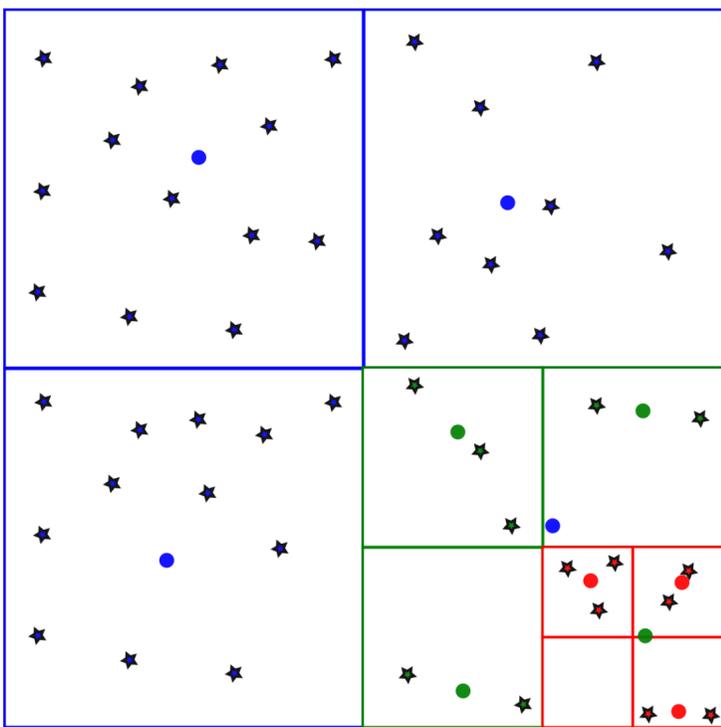
Kräfte die auf einen Stern wirken

Um die Gesamtkraft die auf einen Stern wirkt zu berechnen, habe ich das Newton'sche Gravitationsgesetz in der folgenden Form verwendet:

$$\vec{F}_1 = G \cdot m_1 \cdot \sum_{i=2}^n m_i \cdot \frac{\vec{r}_i - \vec{r}_1}{|\vec{r}_i - \vec{r}_1|^3} \quad (4)$$

Unterteilung der Galaxie in Zellen

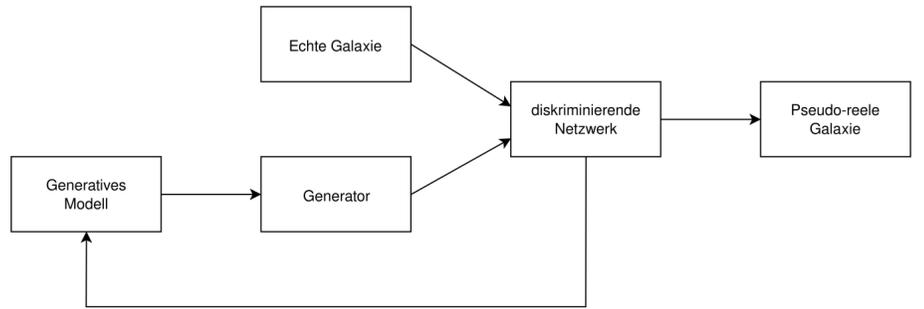
Um den Rechenaufwand, der bei der Berechnung der wirkenden Kräfte entsteht, zu minimieren, wird die Galaxie in verschiedene Zellen unterteilt:



Genutzte Rechner

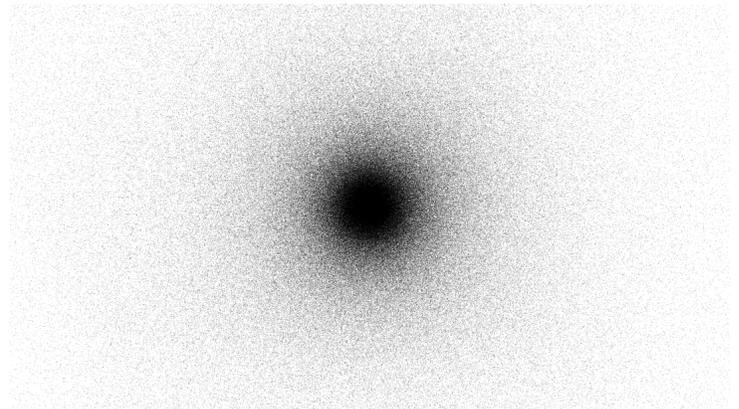
Bezeichnung	Kerne	Threads	Taktrate
Acer Laptop	2	4	3.1 GHz
Heidelberg Cluster	44	88	~ 2.5 GHz
Flammkuchenblech	6	24	2.5 GHz

Generative Adversarial Neural Networks



Aufbau eines GAN Netzwerkes zum Generieren von Galaxien

Ergebnisse



Diese generierte Galaxie bestehend aus 680.000 Sternen

Zukunft

- OpenMP
- OpenACC statt CUDA oder OpenCL
- Code Vectorization
- Shared Memory
 - Distributed Memory
- Hilbert Spiral
- Nutzung des L1, L2 und L3 caches

Danksagungen

Hier möchte ich mich bei ...

... Tim Tugendhat

... Konstantin Bosbach

... Jörg Thar

... Martin Dessauer

bedanken, denn ohne sie wäre die Durchführung des Projektes nicht möglich gewesen.