

Molekulare Bioinformatik

Wintersemester 2011/2012

Martinetz

Institut für Neuro- und Bioinformatik
Universität zu Luebeck

10.1.2012





Statistische Physik

Moleküle sind ständig in Bewegung und verändern Ort und Form (mikroskopischer Zustand).

Energieinhalt E spielt eine Rolle.

Temperatur T spielt eine Rolle.

Auch Druck p spielt häufig eine Rolle.

Die Parameter E , T , p sind makroskopische Parameter.

Statistische Physik schafft Verbindung zwischen mikroskopischen Zuständen und makroskopischen Parametern.



Beispiel Gas

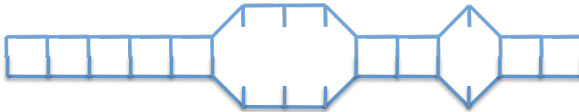
Der mikroskopische Zustand eines Gases ist durch Ort und Impuls jedes der $N \approx 10^{23}$ Gasteilchen beschrieben.

Die makroskopischen Meßgrößen sind der Druck p und die Temperatur T .

Was kann man bei Messung von p und T über den mikroskopischen Zustand des Gases sagen?

Die Konzepte sind sehr universell und auf beliebige Systeme mit vielen Freiheitsgraden übertragbar.

Beispiel Doppelhelix



Uns interessiert, an wievielen Stellen die DNA-Helix aufgetrennt ist.

Dies hängt statistisch ab von der Umgebungstemperatur.

Bei niedriger Temperatur T treten nur wenige aufgetrennte Basenpaare auf, bei hohen Temperaturen kann sich die Helix vollständig auftrennen.

Wie können wir zu Aussagen über den mikroskopischen Zustand der Helix kommen?



Beispiel Doppelhelix

Kodierung des mikroskopischen Zustandes: $s_i \in \{0, 1\}$ beschreibt, ob das i -te Basenpaar zusammen ($s_i = 0$) oder aufgetrennt ist ($s_i = 1$).

Der Zustand der Helix ist dann beschrieben durch den Vektor $\vec{s} = (s_1, s_2, \dots, s_N)$, mit N als die Gesamtzahl der Basenpaare.

Die entscheidende Größe eines physikalischen Systems ist dessen Energie E .

Für die Auftrennung eines Basenpaares ist die Energie ε notwendig.

Sind n Paare aufgetrennt, so besitzt das System die Energie

$$E = n \cdot \varepsilon.$$



Energie des Systems gegeben

Frage: Wenn wir nur den aktuellen Energiegehalt E des Systems kennen (System sei isoliert und damit $E = \text{const.}$), welche Aussagen können wir über \vec{s} treffen?

Dann sind

$$n = \frac{E}{\varepsilon}$$

Basenpaare aufgetrennt.

Aber welche mit welcher Wahrscheinlichkeit?

Was können wir über die Wahrscheinlichkeit $P(\vec{s})$ sagen, die Helix im Zustand \vec{s} anzutreffen?



Maximierung der Unkenntnis

Wir wissen nur, daß

$$E = \sum_{i=1}^N \varepsilon \cdot s_i = \text{const.}$$

Ansatz: Da wir sonst nichts über den Zustand des Systems wissen, setzen wir dasjenige $P(\vec{s})$ an, welches unsere Unkenntnis über das System unter der Randbedingung

$$\sum_{i=1}^N \varepsilon \cdot s_i = E = \text{const.}$$

maximiert. Wir nehmen also maximale Entropie an.



Maximierung der Entropie

Unsere Unkenntnis ist gegeben durch die Entropie

$$H = - \sum_{\text{alle } \vec{s}} P(\vec{s}) \log P(\vec{s})$$

Wir suchen also diejenige Verteilungsfunktion $P(\vec{s})$, die H maximiert und für die $\sum_{\text{alle } \vec{s}} P(\vec{s}) = 1$ (Normierung), gleichzeitig aber auch $\sum_{i=1}^N \varepsilon \cdot s_i = E$ gilt.

$P(\vec{s})$ ist also überall Null, wo $\sum_{i=1}^N \varepsilon s_i = E$ nicht gilt.



Shannon Entropie

Aufgabe: Suche $P(\vec{s})$ welches

$$H = - \sum_{\text{alle } \vec{s}} P(\vec{s}) \log P(\vec{s})$$

maximiert, unter den Nebenbedingungen

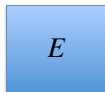
$$\sum_{\text{alle } \vec{s}} P(\vec{s}) = 1 \quad \text{und} \quad \sum_{i=1}^N \varepsilon \cdot s_i = E$$

Mit Lagrange-Multiplikatoren erhalten wir $P(\vec{s}) = \text{const.}$. Also gilt

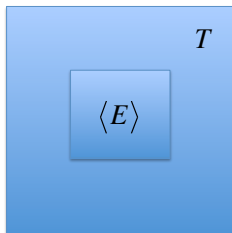
$$P(\vec{s}) = 0 \quad \text{falls} \quad \sum_{i=1}^N \varepsilon \cdot s_i \neq E$$

$$= \left(\frac{N!}{n!(N-n)!} \right)^{-1} \quad \text{falls} \quad \sum_{i=1}^N \varepsilon \cdot s_i = E.$$

System im Wärmebad



Bislang war das System isoliert. Dann E konstant. Die Zustände, die das System einnimmt, nennt man mikrokanonische Gesamtheit.



Für uns realistischer ist System im Wärmebad mit Temperatur T . Der Energieinhalt des Systems fluktuiert dann um einen Mittelwert $\langle E \rangle$. Die Zustände, die das System einnimmt, nennt man kanonische Gesamtheit.



System im Wärmebad

Die Energie des Gesamtsystems (System plus Wärmebad) ist konstant, der Energieinhalt des Systems schwankt um einen Mittelwert $\langle E \rangle$.

Auch hier stellen wir die Frage, mit welcher Wahrscheinlichkeit $P(\vec{s})$ das System den Zustand \vec{s} annimmt.

Nebenbedingung ist diesmal, dass $\langle E \rangle$ konstant ist.

Auch hier bestimmen wir $P(\vec{s})$ wieder durch Maximierung unserer Unkenntnis, also der Entropie.



Maximale Entropie im Wärmebad

Aufgabe: Wir suchen dasjenige $P(\vec{s})$, welches

$$H = - \sum_{\text{alle } \vec{s}} P(\vec{s}) \log P(\vec{s})$$

maximiert, diesmal unter den Nebenbedingungen

$$\sum_{\text{alle } \vec{s}} P(\vec{s}) = 1 \quad \text{und} \quad \sum_{\text{alle } \vec{s}} P(\vec{s}) E(\vec{s}) = \langle E \rangle$$

$E(\vec{s})$ bezeichnet die Energie des Systems im Zustand \vec{s} .



Maximale Entropie im Wärmebad

Ansatz mit Lagrangemultiplikatoren:

$$L = - \sum_{\text{alle } \vec{s}} P(\vec{s}) \ln P(\vec{s}) + \alpha \left(1 - \sum_{\text{alle } \vec{s}} P(\vec{s}) \right) + \beta \left(\langle E \rangle - \sum_{\text{alle } \vec{s}} P(\vec{s}) E(\vec{s}) \right)$$

Es gibt (bei der DNA) 2^N verschiedene \vec{s} . Wir nehmen jetzt ein \vec{s}^* heraus und leiten L nach $P(\vec{s}^*)$ ab:

$$\frac{\partial L}{\partial P(\vec{s}^*)} = -\ln P(\vec{s}^*) - 1 - \alpha - \beta E(\vec{s}^*)$$

Dies wird Null für

$$P(\vec{s}^*) = \underbrace{e^{-1-\alpha}}_{\text{Normierungskonstante } Z^{-1}} \cdot e^{-\beta E(\vec{s}^*)}$$



Boltzmann-Verteilung

Z wird so gewählt, dass die Summe über alle Wahrscheinlichkeiten 1 ergibt:

$$\sum_{\text{alle } \vec{s}} P(\vec{s}) = Z^{-1} \sum_{\text{alle } \vec{s}} e^{-\beta E(\vec{s})} = 1$$

also

$$Z = \sum_{\text{alle } \vec{s}} e^{-\beta E(\vec{s})}$$

Z wird auch als Zustandssumme bezeichnet.

Für die Wahrscheinlichkeiten $P(\vec{s})$ erhalten wir dann die berühmte **Boltzmann-Verteilung**

$$P(\vec{s}) = \frac{e^{-\beta E(\vec{s})}}{Z}$$

Boltzmann-Verteilung

Den Lagrange-Parameter β kann man der Temperatur eines Systems zuordnen.

Es gilt $\beta = 1/k_B T$ mit k_B als der Boltzmann-Konstante und T als der absoluten Temperatur (Kelvin).



Boltzmann-Verteilung

$$P(\vec{s}) = \frac{e^{-\frac{E(\vec{s})}{k_B T}}}{Z}$$

Boltzmann-Konstante

$$k_B = 1,38 \cdot 10^{-23} \text{ J/K}$$