

# 统计物理的数学模型及动理学理论

Ziyang ZHU

School of Mathematical Sciences, Capital Normal University

<https://ziyangzhu.github.io/home/>

October 22, 2025

When we talk about artificial intelligence, we often mean machine learning using artificial neural networks. This technology was originally inspired by the structure of the brain. In an artificial neural network, the brain's neurons are represented by nodes that have different values. In 1983–1985, Geoffrey Hinton used tools from statistical physics to create the Boltzmann machine, which can learn to recognise characteristic elements in a set of data. The invention became significant, for example, for classifying and creating images.

我们谈论的人工智能通常指使用人工神经网络的机器学习。这项技术最初的灵感来自大脑的结构。在人工神经网络中，大脑的神经元由具有不同值的节点表示。1983-1985 年，Geoffrey Hinton 使用统计物理学的工具创建了玻尔兹曼机，该机器可以学习识别一组数据中的特征元素。例如，该发明在图像分类和创建方面具有重要意义。

——2024 年诺贝尔物理学奖

# 目录

统计物理的数学模型及系综理论

Ziyang ZHU

School of Mathematical Sciences, Capital Normal University

<https://ziyangzhu.github.io/home/>

December 25, 2023

Ludwig Boltzmann, who spent much of his life studying statistical mechanics, died in 1906, by his own hand. Paul Ehrenfest, carrying on the work, died similarly in 1933. Now it is our turn to study statistical mechanics.

—Dr. David Louis Goodstein (1939- )




图: Dr. Chris Ferrie 的著作《宝宝的统计物理学》, 中国科大出版

- 1 前情提要：平衡态统计与系综
- 2 非平衡态统计：Boltzmann 的动理学方程
- 3 热力学第二定律与“熵”定理

# 1、什么是平衡态？

Boltzmann 最初的想法

# 1、什么是平衡态？

## Boltzmann 最初的想法

考虑固定粒子个数  $N$  以及总能量  $E$  的全同粒子作成的系统, 假设每个粒子的“能级”为  $\{\epsilon_i : i \geq 1\}$ . 方程组

$$\sum_{i=1}^{\infty} n_i = N; \quad \sum_{i=1}^{\infty} \epsilon_i n_i = E$$

的解  $\mathbf{n} = (n_i)$  的重数为  $\Omega(\mathbf{n}) = \frac{N!}{\prod_i n_i!}$ . 其中, 有一个解 (如果存在的话)

$$\mathbf{n}_B = \left( e^{-(\alpha + \beta \epsilon_i)} \right)$$

的重数最大, 且在  $N \rightarrow \infty$  时远远超出其余解的重数 (甚至是总和). 我们称这个解为 **Boltzmann 分布**.

- Boltzmann 分布的物理意义是: 系统以能级  $\epsilon_i$  上有  $e^{-(\alpha+\beta\epsilon_i)}$  个粒子这个宏观状态出现的概率最高 (假设各微观状态满足等概率原理).

- Boltzmann 分布的物理意义是: 系统以能级  $\epsilon_i$  上有  $e^{-(\alpha+\beta\epsilon_i)}$  个粒子这个宏观状态出现的概率最高 (假设各微观状态满足等概率原理).
- 理论上将  $n_B$  代入原方程组可以解出  $\alpha$  和  $\beta$ .

- Boltzmann 分布的物理意义是: 系统以能级  $\epsilon_i$  上有  $e^{-(\alpha+\beta\epsilon_i)}$  个粒子这个宏观状态出现的概率最高 (假设各微观状态满足等概率原理).
- 理论上将  $\mathbf{n}_B$  代入原方程组可以解出  $\alpha$  和  $\beta$ .
- 通常称  $Z(\beta) := \sum_i e^{-\beta\epsilon_i}$  为配分函数 (partition function).



- Boltzmann 分布的物理意义是: 系统以能级  $\epsilon_i$  上有  $e^{-(\alpha+\beta\epsilon_i)}$  个粒子这个宏观状态出现的概率最高 (假设各微观状态满足等概率原理).
- 理论上将  $\mathbf{n}_B$  代入原方程组可以解出  $\alpha$  和  $\beta$ .
- 通常称  $Z(\beta) := \sum_i e^{-\beta\epsilon_i}$  为配分函数 (partition function).
- $\alpha = \log \frac{Z}{N}$ ;  $\frac{dZ}{d\beta} = -Ee^\alpha$ ;  $\log \Omega(\mathbf{n}_B) \sim N \log Z + \beta E$  ( $N \rightarrow \infty$ ).

- Boltzmann 分布的物理意义是: 系统以能级  $\epsilon_i$  上有  $e^{-(\alpha+\beta\epsilon_i)}$  个粒子这个宏观状态出现的概率最高 (假设各微观状态满足等概率原理).
- 理论上将  $\mathbf{n}_B$  代入原方程组可以解出  $\alpha$  和  $\beta$ .
- 通常称  $Z(\beta) := \sum_i e^{-\beta\epsilon_i}$  为配分函数 (partition function).
- $\alpha = \log \frac{Z}{N}$ ;  $\frac{dZ}{d\beta} = -Ee^\alpha$ ;  $\log \Omega(\mathbf{n}_B) \sim N \log Z + \beta E$  ( $N \rightarrow \infty$ ).
- $\log \Omega(\mathbf{n}_B) \sim \log \sum_{\mathbf{n}} \Omega(\mathbf{n})$  的物理意义是熵 (entropy);  $\beta = \frac{\partial \text{熵}}{\partial E}$  的物理意义是  $\frac{1}{\text{温度}}$ .

- Boltzmann 分布的物理意义是: 系统以能级  $\epsilon_i$  上有  $e^{-(\alpha+\beta\epsilon_i)}$  个粒子这个宏观状态出现的概率最高 (假设各微观状态满足等概率原理).
- 理论上将  $\mathbf{n}_B$  代入原方程组可以解出  $\alpha$  和  $\beta$ .
- 通常称  $Z(\beta) := \sum_i e^{-\beta\epsilon_i}$  为配分函数 (partition function).
- $\alpha = \log \frac{Z}{N}$ ;  $\frac{dZ}{d\beta} = -Ee^\alpha$ ;  $\log \Omega(\mathbf{n}_B) \sim N \log Z + \beta E$  ( $N \rightarrow \infty$ ).
- $\log \Omega(\mathbf{n}_B) \sim \log \sum_{\mathbf{n}} \Omega(\mathbf{n})$  的物理意义是熵 (entropy);  $\beta = \frac{\partial \text{熵}}{\partial E}$  的物理意义是  $\frac{1}{\text{温度}}$ .
- 平衡态指熵最大时对应的分布. 根据定义 Boltzmann 分布就是平衡态的分布.

- Boltzmann 分布的物理意义是: 系统以能级  $\epsilon_i$  上有  $e^{-(\alpha+\beta\epsilon_i)}$  个粒子这个宏观状态出现的概率最高 (假设各微观状态满足等概率原理).
- 理论上将  $\mathbf{n}_B$  代入原方程组可以解出  $\alpha$  和  $\beta$ .
- 通常称  $Z(\beta) := \sum_i e^{-\beta\epsilon_i}$  为配分函数 (partition function).
- $\alpha = \log \frac{Z}{N}$ ;  $\frac{dZ}{d\beta} = -Ee^\alpha$ ;  $\log \Omega(\mathbf{n}_B) \sim N \log Z + \beta E$  ( $N \rightarrow \infty$ ).
- $\log \Omega(\mathbf{n}_B) \sim \log \sum_{\mathbf{n}} \Omega(\mathbf{n})$  的物理意义是熵 (entropy);  $\beta = \frac{\partial \text{熵}}{\partial E}$  的物理意义是  $\frac{1}{\text{温度}}$ .
- 平衡态指熵最大时对应的分布. 根据定义 Boltzmann 分布就是平衡态的分布.
- Boltzmann 最初的想法并不严谨, 而系综 (ensemble) 的诞生弥补了平衡态统计物理的一切.

## Gibbs 的想法——系综

考虑限制在区域  $\Lambda \subseteq \mathbb{R}^d$  内的  $N$  个自由度为  $d$  的粒子的相空间

$$\Gamma = (\Lambda \times \mathbb{R}^d)^N$$

并配备一个 (不带势能的) Hamilton 量  $H: \Gamma \rightarrow \mathbb{R}, x = (q, p) \mapsto \sum_{i=1}^N p_i^2$ .  
定义其对应的正则系综 (canonical ensemble) 为相空间上的概率

$$\rho(x) = \frac{e^{-\beta H(x)}}{\int_{\Gamma} e^{-\beta H(x)} dx},$$

其中  $dx$  为典范辛形式诱导的体积形式.

## Gibbs 的想法——系综

考虑限制在区域  $\Lambda \subseteq \mathbb{R}^d$  内的  $N$  个自由度为  $d$  的粒子的相空间

$$\Gamma = (\Lambda \times \mathbb{R}^d)^N$$

并配备一个 (不带势能的) Hamilton 量  $H: \Gamma \rightarrow \mathbb{R}, x = (q, p) \mapsto \sum_{i=1}^N p_i^2$ .  
定义其对应的正则系综 (canonical ensemble) 为相空间上的概率

$$\rho(x) = \frac{e^{-\beta H(x)}}{\int_{\Gamma} e^{-\beta H(x)} dx},$$

其中  $dx$  为典范辛形式诱导的体积形式. 如果固定某个能量  $E$ , 就获得了微正则系综 (microcanonical ensemble, NEV): 它是相空间中体积有限的闭子集  $H^{-1}(E)$  上由正则系综诱导的常值概率 (由 Liouville 定理得到等概率原理: 微正则系综中系统的各微观状态出现的概率相等).

- 相比于 Boltzmann 最初的想法, 系综的优势在于不需要再使用微观状态数  $\Omega(\mathbf{n})$ , 它被相空间完全记录.

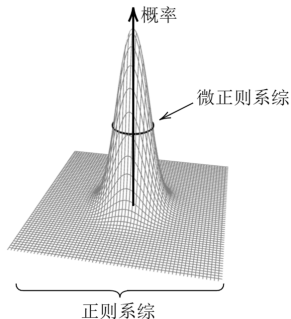
- 相比于 Boltzmann 最初的想法, 系综的优势在于不需要再使用微观状态数  $\Omega(\mathbf{n})$ , 它被相空间完全记录.
- 通常称  $Z(\beta) := \int_{\Gamma} e^{-\beta H(x)} dx$  为**配分函数**, 这里  $\beta$  的物理意义是  $\frac{1}{\text{温度}}$ . 利用配分函数可以直接获得其它热力学函数.

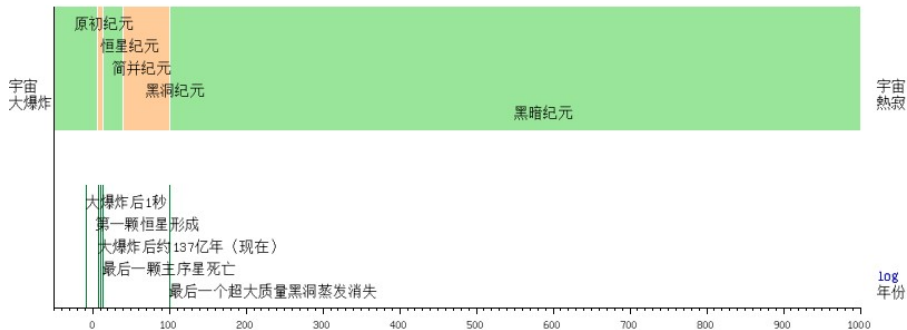


- 相比于 Boltzmann 最初的想法, 系综的优势在于不需要再使用微观状态数  $\Omega(\mathbf{n})$ , 它被相空间完全记录.
- 通常称  $Z(\beta) := \int_{\Gamma} e^{-\beta H(\mathbf{x})} d\mathbf{x}$  为**配分函数**, 这里  $\beta$  的物理意义是  $\frac{1}{\text{温度}}$ . 利用配分函数可以直接获得其它热力学函数.
- 微正则系综熵就是  $H^{-1}(E)$  的体积, 它是 Boltzmann 熵的系综版本.

- 相比于 Boltzmann 最初的想法, 系综的优势在于不需要再使用微观状态数  $\Omega(\mathbf{n})$ , 它被相空间完全记录.
- 通常称  $Z(\beta) := \int_{\Gamma} e^{-\beta H(\mathbf{x})} d\mathbf{x}$  为**配分函数**, 这里  $\beta$  的物理意义是  $\frac{1}{\text{温度}}$ . 利用配分函数可以直接获得其它热力学函数.
- 微正则系综熵就是  $H^{-1}(E)$  的体积, 它是 Boltzmann 熵的系综版本.
- 正则系综的定义沿用了 Boltzmann 分布, 所以它致力于描述平衡态, 也就是熵最大时的分布.

- 相比于 Boltzmann 最初的想法, 系综的优势在于不需要再使用微观状态数  $\Omega(\mathbf{n})$ , 它被相空间完全记录.
- 通常称  $Z(\beta) := \int_{\Gamma} e^{-\beta H(\mathbf{x})} d\mathbf{x}$  为配分函数, 这里  $\beta$  的物理意义是  $\frac{1}{\text{温度}}$ . 利用配分函数可以直接获得其它热力学函数.
- 微正则系综熵就是  $H^{-1}(E)$  的体积, 它是 Boltzmann 熵的系综版本.
- 正则系综的定义沿用了 Boltzmann 分布, 所以它致力于描述平衡态, 也就是熵最大时的分布.





图：热寂假说——处于热平衡时的宇宙熵最大，各向同性达到极致

## 2、什么是非平衡态？

一言以蔽之：熵最大是平衡态，故熵还没有达到最大就是非平衡态。

## 2、什么是非平衡态？

一言以蔽之：熵最大是平衡态，故熵还没有达到最大就是非平衡态。

### 问题

熵是什么？

我们无法从平衡态入手定义熵！

## 2、什么是非平衡态？

一言以蔽之：熵最大是平衡态，故熵还没有达到最大就是非平衡态。

### 问题

熵是什么？

我们无法从平衡态入手定义熵！

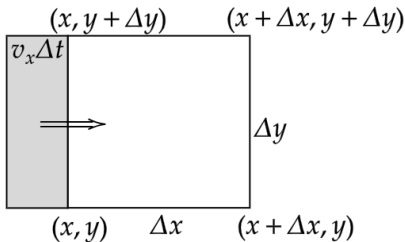
(与系综完全不同的模型) 不妨考虑二维粒子的相空间  $\Gamma = \mathbb{R}^2 \times \mathbb{R}^2$ . 既然要描述熵变的过程，我们引入一个时间轴  $\mathbb{R}$ . 因此相空间上的演化就是一族分布函数

$$f(r, v, t) : \Gamma \times \mathbb{R} \rightarrow \mathbb{R},$$

其中  $r = (x, y)$  表示位置， $v = (v_x, v_y)$  表示动量. 它表示在  $t$  时刻处，在位置  $r$ ，动量  $v$  处的粒子密度.

# Boltzmann 方程的推导

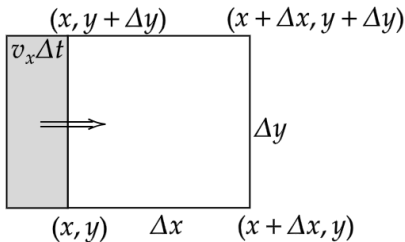
固定相空间中  $r = (x, y)$ ,  $v = (v_x, v_y)$  处的微元如下. 那么微元内的粒子数为  $f(r, v, t)\Delta x\Delta y\Delta v_x\Delta v_y$ .





# Boltzmann 方程的推导

固定相空间中  $r = (x, y)$ ,  $v = (v_x, v_y)$  处的微元如下. 那么微元内的粒子数为  $f(r, v, t)\Delta x\Delta y\Delta v_x\Delta v_y$ .



当时间  $t \rightarrow t + \Delta t$  时 (假设  $\Delta t \ll \Delta x, \Delta y$ ), 那么从微元左面进入的粒子数为 (此处  $f$  在  $y + \Delta y$  方向用一阶近似)

$$L.in = \int_{v_x > 0} f(x, y, v_x, v_y, t) (v_x \Delta t \Delta y) dv_x dv_y.$$

# Boltzmann 方程的推导

从微元左面进入的粒子数为

$$L.in = \int_{v_x > 0} f(x, y, v_x, v_y, t) (v_x \Delta t \Delta y) dv_x dv_y.$$

# Boltzmann 方程的推导

从微元左面进入的粒子数为

$$L.in = \int_{v_x > 0} f(x, y, v_x, v_y, t) (v_x \Delta t \Delta y) dv_x dv_y.$$

同理, 从微元

- 左面离开粒子数  $L.out = \int_{v_x < 0} f(x, y, v_x, v_y, t) (-v_x \Delta t \Delta y) dv_x dv_y.$
- 右面进入粒子数  $R.in = \int_{v_x < 0} f(x + \Delta x, y, v_x, v_y, t) (-v_x \Delta t \Delta y) dv_x dv_y.$
- 右面离开粒子数  $R.out = \int_{v_x > 0} f(x + \Delta x, y, v_x, v_y, t) (v_x \Delta t \Delta y) dv_x dv_y.$

# Boltzmann 方程的推导

从微元左面进入的粒子数为

$$L.in = \int_{v_x > 0} f(x, y, v_x, v_y, t) (v_x \Delta t \Delta y) dv_x dv_y.$$

同理, 从微元

- 左面离开粒子数  $L.out = \int_{v_x < 0} f(x, y, v_x, v_y, t) (-v_x \Delta t \Delta y) dv_x dv_y.$
- 右面进入粒子数  $R.in = \int_{v_x < 0} f(x + \Delta x, y, v_x, v_y, t) (-v_x \Delta t \Delta y) dv_x dv_y.$
- 右面离开粒子数  $R.out = \int_{v_x > 0} f(x + \Delta x, y, v_x, v_y, t) (v_x \Delta t \Delta y) dv_x dv_y.$

因此水平方向的总通量为

$$\Delta H_{or} = L.in + R.in - L.out - R.out.$$

# Boltzmann 方程的推导

$$\begin{aligned}\Delta H_{or} &= \int_{v_x > 0} \left( f(x, y, v_x, v_y, t) - f(x + \Delta x, y, v_x, v_y, t) \right) (v_x \Delta t \Delta y) dv_x dv_y \\ &\quad + \int_{v_x < 0} \left( f(x + \Delta x, y, v_x, v_y, t) - f(x, y, v_x, v_y, t) \right) (-v_x \Delta t \Delta y) dv_x dv_y \\ &\sim \int_{v_x > 0} -\frac{\partial f}{\partial x} \Delta x (v_x \Delta t \Delta y) dv_x dv_y + \int_{v_x < 0} \frac{\partial f}{\partial x} \Delta x (-v_x \Delta t \Delta y) dv_x dv_y \\ &\sim - \int_{v_x \in \mathbb{R}} \frac{\partial f}{\partial x} v_x \Delta x \Delta y \Delta t dv_x dv_y \quad (\text{if } \Delta x \rightarrow 0).\end{aligned}$$

# Boltzmann 方程的推导

$$\begin{aligned}\Delta H_{or} &= \int_{v_x > 0} \left( f(x, y, v_x, v_y, t) - f(x + \Delta x, y, v_x, v_y, t) \right) (v_x \Delta t \Delta y) dv_x dv_y \\ &\quad + \int_{v_x < 0} \left( f(x + \Delta x, y, v_x, v_y, t) - f(x, y, v_x, v_y, t) \right) (-v_x \Delta t \Delta y) dv_x dv_y \\ &\sim \int_{v_x > 0} -\frac{\partial f}{\partial x} \Delta x (v_x \Delta t \Delta y) dv_x dv_y + \int_{v_x < 0} \frac{\partial f}{\partial x} \Delta x (-v_x \Delta t \Delta y) dv_x dv_y \\ &\sim - \int_{v_x \in \mathbb{R}} \frac{\partial f}{\partial x} v_x \Delta x \Delta y \Delta t dv_x dv_y \quad (\text{if } \Delta x \rightarrow 0).\end{aligned}$$

同理, 竖直方向的总通量为

$$\Delta V_{er} \sim - \int_{v_y \in \mathbb{R}} \frac{\partial f}{\partial y} v_y \Delta x \Delta y \Delta t dv_x dv_y \quad (\text{if } \Delta y \rightarrow 0).$$

# Boltzmann 方程的推导

当  $\Delta x \rightarrow 0, \Delta y \rightarrow 0$  (所以  $\Delta t \rightarrow 0$ ) 时, 由粒子数守恒可知密度函数  $f$  (的一部分) 关于时间的变化率  $(\frac{\partial f}{\partial t})_{\text{流动部分}}$  满足

$$\begin{aligned}\lim(\Delta Hor + \Delta Ver) &= - \int_{v_x, v_y \in \mathbb{R}} \left( \frac{\partial f}{\partial x} v_x + \frac{\partial f}{\partial y} v_y \right) \Delta x \Delta y \Delta t dv_x dv_y \\ &= \int_{v_x, v_y \in \mathbb{R}} \left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} \Delta x \Delta y \Delta t dv_x dv_y.\end{aligned}$$

# Boltzmann 方程的推导

当  $\Delta x \rightarrow 0, \Delta y \rightarrow 0$  (所以  $\Delta t \rightarrow 0$ ) 时, 由粒子数守恒可知密度函数  $f$  (的一部分) 关于时间的变化率  $\left(\frac{\partial f}{\partial t}\right)_{\text{流动部分}}$  满足

$$\begin{aligned}\lim(\Delta Hor + \Delta Ver) &= - \int_{v_x, v_y \in \mathbb{R}} \left( \frac{\partial f}{\partial x} v_x + \frac{\partial f}{\partial y} v_y \right) \Delta x \Delta y \Delta t dv_x dv_y \\ &= \int_{v_x, v_y \in \mathbb{R}} \left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} \Delta x \Delta y \Delta t dv_x dv_y.\end{aligned}$$

所以我们让

$$\left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} + (\mathbf{v} \cdot \nabla) f = 0,$$

这就是 **Boltzmann-Vlasov 方程** (不带场的部分).



## $f$ 的另一部分: 碰撞项

粒子不光有流动, 有时还需考虑相互作用 (Interaction). 所以完整的 Boltzmann 方程会带有一个碰撞项:

### Boltzmann 方程

$$\frac{\partial f}{\partial t} = \left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} + \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}}$$

## $f$ 的另一部分: 碰撞项

粒子不光有流动, 有时还需考虑相互作用 (Interaction). 所以完整的 Boltzmann 方程会带有一个碰撞项:

### Boltzmann 方程

$$\frac{\partial f}{\partial t} = \left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} + \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}}$$

其中:

- $\left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} = -(\mathbf{v} \cdot \nabla) f$  (不含场的项).

## $f$ 的另一部分: 碰撞项

粒子不光有流动, 有时还需考虑相互作用 (Interaction). 所以完整的 Boltzmann 方程会带有一个碰撞项:

### Boltzmann 方程

$$\frac{\partial f}{\partial t} = \left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} + \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}}$$

其中:

- $\left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} = -(\mathbf{v} \cdot \nabla) f$  (不含场的项).
- $\left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}}$  基本上是一事一议——
  - ▶ **经典款:** 粒子混沌假设 + 刚体小球碰撞模型 (分子动理论).

## $f$ 的另一部分: 碰撞项

粒子不光有流动, 有时还需考虑相互作用 (Interaction). 所以完整的 Boltzmann 方程会带有一个碰撞项:

### Boltzmann 方程

$$\frac{\partial f}{\partial t} = \left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} + \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}}$$

其中:

- $\left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} = -(\mathbf{v} \cdot \nabla) f$  (不含场的项).
- $\left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}}$  基本上是一事一议——
  - ▶ **经典款:** 粒子混沌假设 + 刚体小球碰撞模型 (分子动理论).
  - ▶ **光学款:** 散射模型 (X-ray 成像原理).

## $f$ 的另一部分: 碰撞项

粒子不光有流动, 有时还需考虑相互作用 (Interaction). 所以完整的 Boltzmann 方程会带有一个碰撞项:

### Boltzmann 方程

$$\frac{\partial f}{\partial t} = \left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} + \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}}$$

其中:

- $\left( \frac{\partial f}{\partial t} \right)_{\text{流动部分}} = -(\mathbf{v} \cdot \nabla) f$  (不含场的项).
- $\left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}}$  基本上是一事一议——
  - ▶ **经典款:** 粒子混沌假设 + 刚体小球碰撞模型 (分子动理论).
  - ▶ **光学款:** 散射模型 (X-ray 成像原理).
  - ▶ **流体力学, 大气科学, 人工智能 (神经网络), 宇宙学 (天体力学).....**

# 经典碰撞项

我们有:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{碰撞部分}} = \int \rho(u, v, u', v') \left( f(r, u', t) f(r, v', t) - f(r, u, t) f(r, v, t) \right) du dv du' dv',$$

其中  $r, u, u', v, v'$  均是向量; 而

$$\rho(u, v, u', v') f(r, u, t) f(r, v, t) du dv du' dv'$$

表示单位体积单位时间的总碰撞数, 它依赖不同的 (刚体小球) 碰撞模型.

# 经典碰撞项

我们有:

$$\left(\frac{\partial f}{\partial t}\right)_{\text{碰撞部分}} = \int \rho(u, v, u', v') \left( f(r, u', t) f(r, v', t) - f(r, u, t) f(r, v, t) \right) du du' dv',$$

其中  $r, u, u', v, v'$  均是向量; 而

$$\rho(u, v, u', v') f(r, u, t) f(r, v, t) du dv du' dv'$$

表示单位体积单位时间的总碰撞数, 它依赖不同的 (刚体小球) 碰撞模型.

## 警告

这里建立  $\rho$  用到了所谓的**粒子混沌假设**, 它认为所有的粒子碰撞都是独立事件, 但该过程会有信息丢失. 这个假设会导致 (经典碰撞项的) Boltzmann 方程不能时间反演, 也就是有一个时间箭头: “熵” 增.

# 经典碰撞项

没有人比他们更懂小球的碰撞:



图: 斯诺克球员 Mark Selby (左), Judd Trump (右). 2025 年 10 月 5 日于徐州



# 光学碰撞项: 成像原理

由于光速是常数  $c$ , 所以可以稍稍修改相空间以及 Boltzmann 方程:

$$f(x, y, \theta, t) : (\mathbb{R}^2 \times S^1) \times \mathbb{R} \rightarrow \mathbb{R};$$

# 光学碰撞项: 成像原理

由于光速是常数  $c$ , 所以可以稍稍修改相空间以及 Boltzmann 方程:

$$f(x, y, \theta, t) : (\mathbb{R}^2 \times S^1) \times \mathbb{R} \rightarrow \mathbb{R};$$

和

$$\frac{\partial f}{\partial t} + c \left( \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta \right) = -pf + (1-p) \int_0^{2\pi} s(x, y, \theta - \tau) f(x, y, \tau, t) d\tau.$$

# 光学碰撞项: 成像原理

由于光速是常数  $c$ , 所以可以稍稍修改相空间以及 Boltzmann 方程:

$$f(x, y, \theta, t) : (\mathbb{R}^2 \times S^1) \times \mathbb{R} \rightarrow \mathbb{R};$$

和

$$\frac{\partial f}{\partial t} + c \left( \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta \right) = -pf + (1-p) \int_0^{2\pi} s(x, y, \theta - \tau) f(x, y, \tau, t) d\tau.$$

其中,  $p = p(x, y)$  表示位置  $(x, y)$  处对粒子的吸收率,  $s(x, y, \cdot)$  表示在  $(x, y)$  处将某方向的粒子散射到其它方向的概率.

# 光学碰撞项: 成像原理

由于光速是常数  $c$ , 所以可以稍稍修改相空间以及 Boltzmann 方程:

$$f(x, y, \theta, t) : (\mathbb{R}^2 \times S^1) \times \mathbb{R} \rightarrow \mathbb{R};$$

和

$$\frac{\partial f}{\partial t} + c \left( \frac{\partial f}{\partial x} \cos \theta + \frac{\partial f}{\partial y} \sin \theta \right) = -pf + (1-p) \int_0^{2\pi} s(x, y, \theta - \tau) f(x, y, \tau, t) d\tau.$$

其中,  $p = p(x, y)$  表示位置  $(x, y)$  处对粒子的吸收率,  $s(x, y, \cdot)$  表示在  $(x, y)$  处将某方向的粒子散射到其它方向的概率.

该模型的物理意义是: 初始时刻输入一个分布 (点射线源), 这些粒子在运动过程中被不均匀的空间吸收、散射, 最终输出  $t$  时刻粒子在空间中的分布 (位置分布  $\times$  方向分布).

# 数值模拟

平面内存在一个“回”字形区域, 粒子从左上角均匀射出并以不同的碰撞模型在其中流动. 左图的碰撞项为 0; 中图的碰撞项为光学散射模型; 右图为一个随意模型.

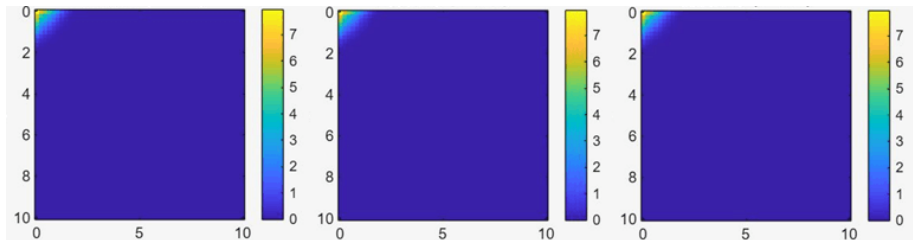


图: ● ○ ○ ○ ○ ○ ○ ○

# 数值模拟

平面内存在一个“回”字形区域, 粒子从左上角均匀射出并以不同的碰撞模型在其中流动. 左图的碰撞项为 0; 中图的碰撞项为光学散射模型; 右图为一个随意模型.

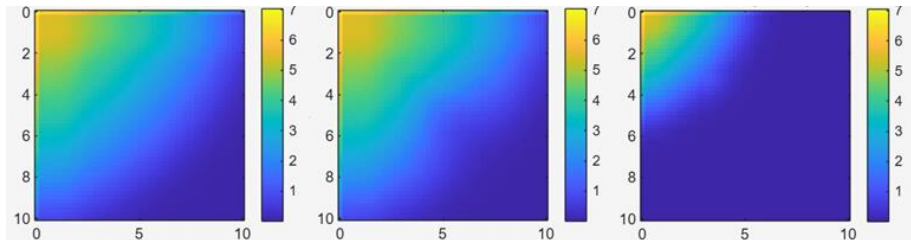


图: ● ● ○ ○ ○ ○ ○

# 数值模拟

平面内存在一个“回”字形区域, 粒子从左上角均匀射出并以不同的碰撞模型在其中流动. 左图的碰撞项为 0; 中图的碰撞项为光学散射模型; 右图为一个随意模型.

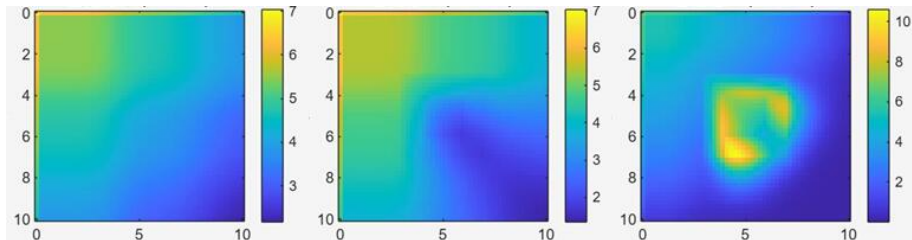


图: ● ● ● ○ ○ ○ ○

# 数值模拟

平面内存在一个“回”字形区域, 粒子从左上角均匀射出并以不同的碰撞模型在其中流动. 左图的碰撞项为 0; 中图的碰撞项为光学散射模型; 右图为一个随意模型.

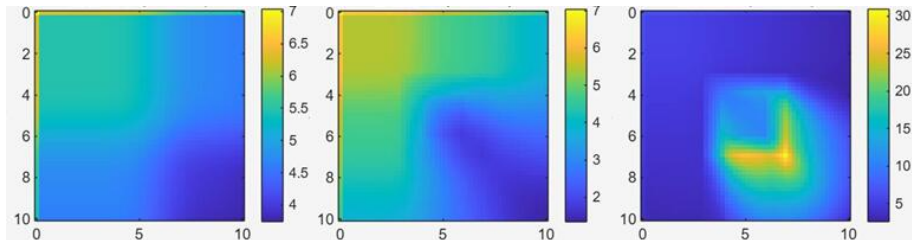


图: ● ● ● ● ○ ○ ○



# 数值模拟

平面内存在一个“回”字形区域, 粒子从左上角均匀射出并以不同的碰撞模型在其中流动. 左图的碰撞项为 0; 中图的碰撞项为光学散射模型; 右图为一个随意模型.

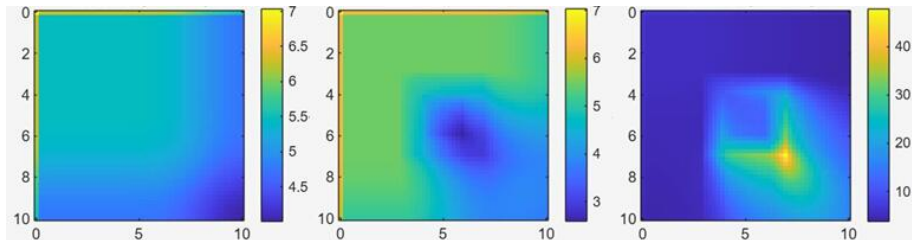


图: ● ● ● ● ● ○ ○

# 数值模拟

平面内存在一个“回”字形区域, 粒子从左上角均匀射出并以不同的碰撞模型在其中流动. 左图的碰撞项为 0; 中图的碰撞项为光学散射模型; 右图为一个随意模型.

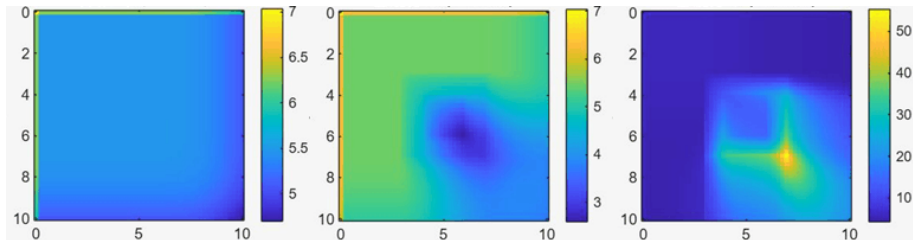


图: ● ● ● ● ● ● ○

# 数值模拟

平面内存在一个“回”字形区域, 粒子从左上角均匀射出并以不同的碰撞模型在其中流动. 左图的碰撞项为 0; 中图的碰撞项为光学散射模型; 右图为一个随意模型.

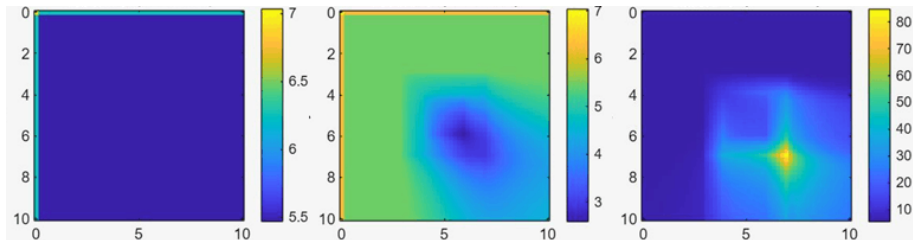


图: ● ● ● ● ● ● ● ●

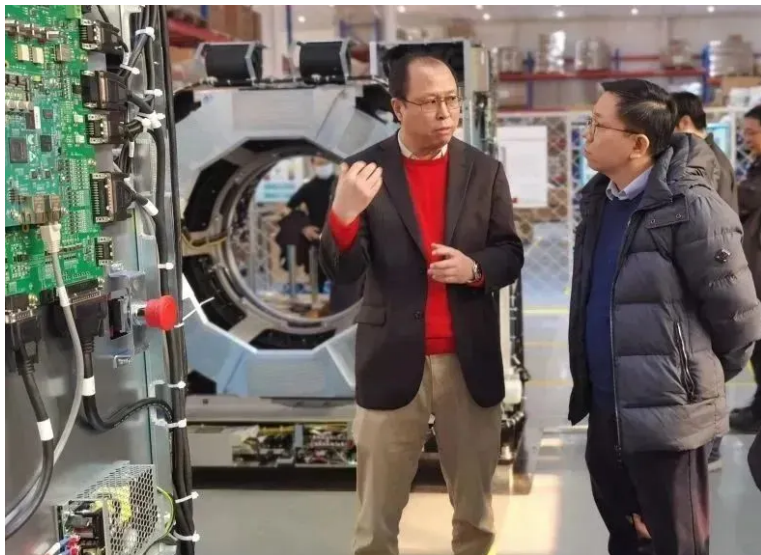


图: 学校代表参观高精尖成像实验室

# 动力 (Dynamic) V.S. 动理 (Kinetic)

## 行业黑话

都是在描述运动系统, Dynamic 和 Kinetic 到底有啥区别?

# 动力 (Dynamic) V.S. 动理 (Kinetic)

## 行业黑话

都是在描述运动系统, Dynamic 和 Kinetic 到底有啥区别?

众说纷纭. 具体历史已不可考, 且很多人在混淆术语. 有说法认为:

- “Kinetic” 比较古老 (可查的是出现在 17 世纪一帮哲学家的骂战), 它似乎专指 “与速度直接相关的动力学”.
- “Dynamic” 相对时髦 (成型于 19 世纪 Poincaré 对天体力学的研究), 用来描述 “随时间演化的因果律” 等更宽泛的概念.

# 动力 (Dynamic) V.S. 动理 (Kinetic)

## 行业黑话

都是在描述运动系统, Dynamic 和 Kinetic 到底有啥区别?

众说纷纭. 具体历史已不可考, 且很多人在混淆术语. 有说法认为:

- “Kinetic” 比较古老 (可查的是出现在 17 世纪一帮哲学家的骂战), 它似乎专指 “与速度直接相关的动力学”.
- “Dynamic” 相对时髦 (成型于 19 世纪 Poincaré 对天体力学的研究), 用来描述 “随时间演化的因果律” 等更宽泛的概念.

不过混淆术语貌似不是什么稀奇事, 就好比我不明白:

Number Theory (数论) V.S. Arithmetic (算术? 算数?) 有什么区别?

### 3、熵定理 (H-定理)

#### 问题

给定初始分布, 当  $t \rightarrow \infty$  时, 系统会趋于平衡态吗?



### 3、熵定理 (H-定理)

#### 问题

给定初始分布, 当  $t \rightarrow \infty$  时, 系统会趋于平衡态吗?

什么是平衡态? 按之前的说法: 熵最大, 系统随时间趋于稳定. 后者很好说,  $\frac{\partial f}{\partial t} = 0$  (只有这个条件的话只能称为稳态).

### 3、熵定理 (H-定理)

#### 问题

给定初始分布, 当  $t \rightarrow \infty$  时, 系统会趋于平衡态吗?

什么是平衡态? 按之前的说法: 熵最大, 系统随时间趋于稳定. 后者很好说,  $\frac{\partial f}{\partial t} = 0$  (只有这个条件的话只能称为稳态).

#### 问题

什么是熵?

### 3、熵定理 (H-定理)

#### 问题

给定初始分布, 当  $t \rightarrow \infty$  时, 系统会趋于平衡态吗?

什么是平衡态? 按之前的说法: 熵最大, 系统随时间趋于稳定. 后者很好说,  $\frac{\partial f}{\partial t} = 0$  (只有这个条件的话只能称为稳态).

#### 问题

什么是熵?

对于  $f = f(r, v, t) > 0$ , 定义

$$H(t) := - \int f \log f dr dv.$$

### 3、熵定理 (H-定理)

#### 问题

给定初始分布, 当  $t \rightarrow \infty$  时, 系统会趋于平衡态吗?

什么是平衡态? 按之前的说法: 熵最大, 系统随时间趋于稳定. 后者很好说,  $\frac{\partial f}{\partial t} = 0$  (只有这个条件的话只能称为稳态).

#### 问题

什么是熵?

对于  $f = f(r, v, t) > 0$ , 定义

$$H(t) := - \int f \log f dr dv.$$

这个定义来自于最原始的熵形式  $\log W = -W \cdot \frac{1}{W} \log \frac{1}{W} = -\sum_i p_i \log p_i$ , 其中  $W = \sum_n \Omega(n)$  且根据等概率原理  $p_i = \frac{1}{W}$ .

### 3、熵定理 (H-定理)

#### 定理

假设  $f > 0$  且在  $|r| \rightarrow \infty$  时足够小, 那么

$$\frac{dH}{dt} = \int \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}} (1 + \log f) dr dv.$$

### 3、熵定理 (H-定理)

#### 定理

假设  $f > 0$  且在  $|r| \rightarrow \infty$  时足够小, 那么

$$\frac{dH}{dt} = \int \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}} (1 + \log f) dr dv.$$

Proof. 直接计算

$$\frac{dH}{dt} = \int \frac{\partial f}{\partial t} (1 + \log f) dr dv = \int \left( -v \cdot \nabla f + \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}} \right) (1 + \log f) dr dv.$$

### 3、熵定理 (H-定理)

#### 定理

假设  $f > 0$  且在  $|r| \rightarrow \infty$  时足够小, 那么

$$\frac{dH}{dt} = \int \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}} (1 + \log f) dr dv.$$

Proof. 直接计算

$$\frac{dH}{dt} = \int \frac{\partial f}{\partial t} (1 + \log f) dr dv = \int \left( -v \cdot \nabla f + \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}} \right) (1 + \log f) dr dv.$$

而  $\int v \cdot \nabla f dr dv = \int v \cdot \left( \int \nabla f dr \right) dv = 0$  (Gauss 定理), 同理

$$\int (v \cdot \nabla f) \log f dr dv = \int v \cdot \left( \int \nabla (f \log f - f) dr \right) dv = 0.$$

### 3、熵定理 (H-定理)

我们希望  $\frac{dH}{dt} \leq 0$  (对应熵增). 至少在某些模型中确实如此:

例 (分子动理论, 热力学第二定律)



### 3、熵定理 (H-定理)

我们希望  $\frac{dH}{dt} \leq 0$  (对应熵增). 至少在某些模型中确实如此:

例 (分子动理论, 热力学第二定律)

在经典碰撞模型中, 由参与碰撞的粒子数守恒得到

$$\int \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}} dv = 0.$$

另一方面, 可以证明 (略, 我们并没有给出  $\rho$  的具体构造)

$$\int \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}} \log f dv \leq 0.$$

所以  $\frac{dH}{dt} \leq 0$ .

### 3、熵定理 (H-定理)

我们希望  $\frac{dH}{dt} \leq 0$  (对应熵增). 至少在某些模型中确实如此:

例 (分子动理论, 热力学第二定律)

在经典碰撞模型中, 由参与碰撞的粒子数守恒得到

$$\int \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}} dv = 0.$$

另一方面, 可以证明 (略, 我们并没有给出  $\rho$  的具体构造)

$$\int \left( \frac{\partial f}{\partial t} \right)_{\text{碰撞部分}} \log f dv \leq 0.$$

所以  $\frac{dH}{dt} \leq 0$ . 特别的,  $\frac{dH}{dt} = 0$  当且仅当微观上来看碰撞与其逆碰撞平衡, 我们称其为**细致平衡**. 此时, 粒子在空间中均匀分布, 速度服从 Gauss (Boltzmann) 分布——这就是该模型的平衡态. 它对应一个正则系综.

但是

平衡态  $\implies \frac{\partial f}{\partial t} = 0$  (系统稳定) &  $\frac{dH}{dt} = 0$  (熵最大).

# 但是

平衡态  $\implies \frac{\partial f}{\partial t} = 0$  (系统稳定) &  $\frac{dH}{dt} = 0$  (熵最大).

- $\nLeftarrow$ . 右边的信息严重不足. 物理学家的“补丁”: 速度局部服从 Boltzmann (Maxwell, Gauss, 正态) 分布, 空间服从均匀分布等.

但是

平衡态  $\implies \frac{\partial f}{\partial t} = 0$  (系统稳定) &  $\frac{dH}{dt} = 0$  (熵最大).

- $\nRightarrow$ . 右边的信息严重不足. 物理学家的“补丁”: 速度局部服从 Boltzmann (Maxwell, Gauss, 正态) 分布, 空间服从均匀分布等.
- 并非所有的系统都会熵增; 并非所有系统都能达到熵最大的状态. 例如平凡碰撞项的熵守恒, 也有一些特殊的碰撞项会导致熵减.

但是

平衡态  $\implies \frac{\partial f}{\partial t} = 0$  (系统稳定) &  $\frac{dH}{dt} = 0$  (熵最大).

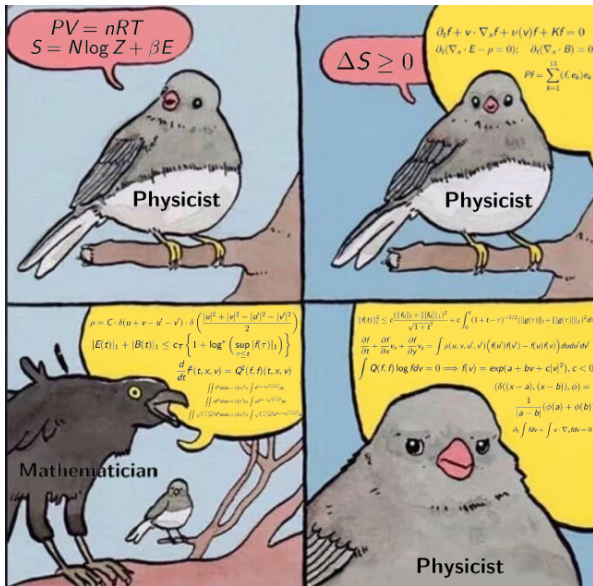
- $\nRightarrow$ . 右边的信息严重不足. 物理学家的“补丁”: 速度局部服从 Boltzmann (Maxwell, Gauss, 正态) 分布, 空间服从均匀分布等.
- 并非所有的系统都会熵增; 并非所有系统都能达到熵最大的状态. 例如平凡碰撞项的熵守恒, 也有一些特殊的碰撞项会导致熵减.
- 这个熵并不是“系统一切混乱程度的度量”. 它只在粒子混沌假设下建立了细致平衡的方向 (假设碰撞项平凡, 则  $H$  不随系统的演化而改变).

但是

平衡态  $\implies \frac{\partial f}{\partial t} = 0$  (系统稳定) &  $\frac{dH}{dt} = 0$  (熵最大).

- $\nRightarrow$ . 右边的信息严重不足. 物理学家的“补丁”: 速度局部服从 Boltzmann (Maxwell, Gauss, 正态) 分布, 空间服从均匀分布等.
- 并非所有的系统都会熵增; 并非所有系统都能达到熵最大的状态. 例如平凡碰撞项的熵守恒, 也有一些特殊的碰撞项会导致熵减.
- 这个熵并不是“系统一切混乱程度的度量”. 它只在粒子混沌假设下建立了细致平衡的方向 (假设碰撞项平凡, 则  $H$  不随系统的演化而改变).

还是那句话,  $H$  真的是我们希望的熵吗?



图：我知道你还有很多问题，但我们是搞数学的



*Thank You!*