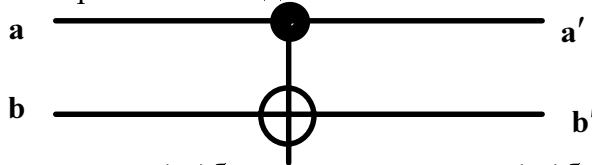


Лекция 2.

1. Сведения из термодинамики и статистической физики. Функция распределения. Теорема Лиувилля. Микроканоническое распределение. Первое начало термодинамики. Адиабатические процессы. Энтропия. Статистический вес. Формула Больцмана. Второе начало термодинамики. Обратимые и необратимые процессы.
2. Информационная энтропия Шеннона. Биты, наты, триты и проч. Связь энтропии и информации.

Эта часть относится к лекции 1. Ее лучше рассматривать в разделе V (“Концепция “перепутывания” (entanglement) квантовых состояний”).

ЛЭ CNOT изображается в виде:



Сохраняем значение (ку)бита a , в то время как (ку)бит b меняется по закону XOR:

a	b	a'	$b'=(a\oplus b)$
0	0	0	0
1	0	1	1
0	1	0	1
1	1	1	0

бит b (мишень = target) меняет свое состояние тогда и только тогда, когда состояние контрольного (control) бита a соответствует 1; при этом, состояние контрольного бита не меняется.

Логическая операция XOR (CNOT) иллюстрирует почему классические данные могут быть клонированы, а квантовые - нет. Заметим, что в общем случае под квантовыми данными мы будем понимать суперпозиции вида

$$|\Psi\rangle = \alpha|0\rangle + \beta|1\rangle, \quad (1)$$

где α и β - комплексные числа или амплитуды состояний, причем, $|\alpha|^2 + |\beta|^2 = 1$.

Согласно таблице истинности, если XOR применить к булевым данным, в которых второй бит находится в состоянии “0” (b), а первый - в состоянии “X” (a), то первый бит не изменяется, а второй становится его копией:

$$U_{XOR}(X, 0) = (X, X), \text{ где } X = \text{“0” или “1”}.$$

В квантовом случае, в качестве данных, обозначенных символом “X”, нужно рассматривать суперпозицию (1):

$$U_{XOR}|X, 0\rangle = |X, X\rangle.$$

Физически, данные можно закодировать, например, в поляризационном базисе $|V\rangle = 1$, $|H\rangle = 0$ (H, V) = (0, 1):

$$U_{XOR}|0, 0\rangle = |0, 0\rangle, \text{ и } U_{XOR}|1, 0\rangle = |1, 1\rangle.$$

Видно, что действительно имеет место копирование состояния. Теорема о запрете клонирования утверждает, что невозможно копирование *произвольного* квантового состояния. В рассмотренном примере копирование произошло, поскольку операция

производилась в собственном базисе ($|0\rangle$, $|1\rangle$), т.е. в частном случае квантового состояния.

Казалось бы, что операцию XOR можно использовать и для копирования суперпозиций двух булевых состояний, таких как $|45^0\rangle \rightarrow |V\rangle + |H\rangle$:

$$U_{XOR} |45^0, H\rangle = |45^0, 45^0\rangle.$$

Но это не так! Унитарность квантовой эволюции требует, чтобы суперпозиция входных состояний преобразовывалась в соответствующую суперпозицию выходных состояний:

$$U_{XOR} |45^0, H\rangle = \left(|45^0\rangle = 1/\sqrt{2} [|H\rangle + |V\rangle] \right) = 1/\sqrt{2} [|H, H\rangle + |V, V\rangle] \text{ или}$$

$$U_{XOR} |s, 0\rangle = \left(|s\rangle = 1/\sqrt{2} [|0\rangle + |1\rangle] \right) = 1/\sqrt{2} [|0, 0\rangle + |1, 1\rangle].$$

$$\Phi^+ = 1/\sqrt{2} [|0, 0\rangle + |1, 1\rangle] \quad (2)$$

Это т.н. перепутанное состояние (Φ^+), в котором каждый из двух выходных кубитов не имеет определенного значения (в данном случае - поляризации). Этот пример показывает, что логические операции, выполняемые над квантовыми объектами происходят по другим правилам, нежели в классических вычислительных процессах.

Возникает следующий вопрос: Вроде бы состояние в выходной моде **a** опять-таки можно представить в виде суперпозиции $|s\rangle = 1/\sqrt{2} [|0\rangle + |1\rangle]$, как и состояние в моде **b**.

Как показать, что это не так, т.е., что вообще нет смысла говорить о состояниях моды (бита) **a** и моды (бита) **b**?

Воспользуемся поляризационной аналогией, когда

$$|s\rangle \equiv |45^0\rangle = 1/\sqrt{2} [|H\rangle + |V\rangle] \quad (3).$$

Есть два пути. Путь 1 - длинный, но более последовательный. Надо посчитать средние значения параметров Стокса для обеих выходных мод. Средние берутся по волновой функции (2). Если все $\langle S_i \rangle$, кроме $\langle S_0 \rangle$ окажутся равными нулю - то это состояние неполяризованное, т.е. смешанное и суперпозиция (3) смысла не имеет. Работаем в представлении Гейзенберга, когда преобразуются операторы, а волновая функция - нет. Итак, находим $\langle S_i \rangle$ в моде **a**.

$$S_0 \equiv a_x^\dagger a_x + a_y^\dagger a_y, \text{ - общая интенсивность пучка a,}$$

$$S_1 \equiv a_x^\dagger a_x - a_y^\dagger a_y, \text{ - доля вертикальной поляризации,}$$

$$S_2 \equiv a_x^\dagger a_y + a_x a_y^\dagger \text{ - доля } +45^0\text{-ой поляризации,}$$

$$S_3 \equiv \frac{1}{i} \{ a_x^\dagger a_y - a_x a_y^\dagger \} \text{ - доля право-циркулярной поляризации.}$$

Волновая функция, по которой производится усреднение, берется в виде (2):

$$\Phi^+ = 1/\sqrt{2} [|H, H\rangle + |V, V\rangle] = 1/\sqrt{2} [a_x^\dagger b_x^\dagger + a_y^\dagger b_y^\dagger] |vac\rangle,$$

где операторы рождения и уничтожения в модах **a** и **b** действуют по правилам:

$$a |N\rangle = N^{1/2} |N-1\rangle,$$

$$a^\dagger |N\rangle = (N+1)^{1/2} |N+1\rangle.$$

Вычисления $\langle S_i \rangle$ сделать в разделе V (см.тетрадь). Там же рассчитать и вероятность регистрации совпадений или коррелятор вида

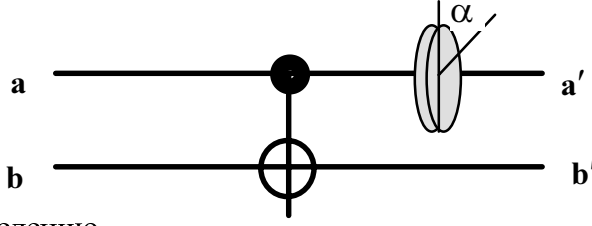
$$R_{ab} = \langle \Phi^+ | aba^\dagger b^\dagger | \Phi^+ \rangle$$

Путь II - более наглядный, но менее “честный”!

Найдем зависимость интенсивности света в моде **a** от угла поворота поляроида, помещенного в эту моду. Это стандартный квантово-оптический способ проверки состояния (2) - интенсивность не должна зависеть от поворота. В то же время, аналогичная зависимость числа совпадений имеет вид

$R_{совн} \sim \cos^2(\alpha \pm \beta)$. Впервые такие зависимости были получены Э.Фраем (1976) и А.Аспеком (1985) и часто интерпретируется как доказательство нелокальности квантовой механики.

Итак, экспериментальная ситуация изображена на рисунке:



По определению

$$R_a \sim \langle \Phi^+ | a^{\dagger} a | \Phi^+ \rangle,$$

где a^{\dagger} - оператор уничтожения в моде **a**. Известно, что преобразование операторов двух ортогонально поляризованных мод **x** и **y** при прохождении света через поляризатор, ориентированный под углом α имеет вид:

$$a' = a_x \sin \alpha + a_y \cos \alpha .$$

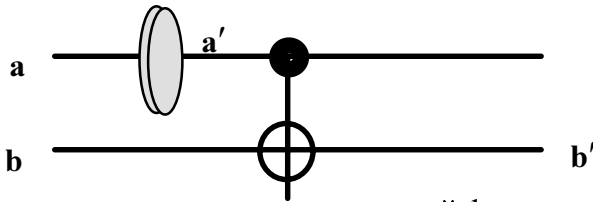
Тогда

$$\begin{aligned} R &\equiv \frac{1}{2} \left[\langle vac | (a_x b_x + a_y b_y) (a_x^{\dagger} \sin \alpha + a_y^{\dagger} \cos \alpha) (a_x \sin \alpha + a_y \cos \alpha) (a_x^{\dagger} b_x^{\dagger} + a_y^{\dagger} b_y^{\dagger}) | vac \rangle \right] = \\ &= \frac{1}{2} \left[\langle vac | (a_x b_x + a_y b_y) (a_x^{\dagger} a_x \sin^2 \alpha + a_x^{\dagger} a_y \sin \alpha \cos \alpha + a_y^{\dagger} a_x \sin \alpha \cos \alpha + a_x^{\dagger} a_y \cos^2 \alpha) (a_x^{\dagger} b_x^{\dagger} + a_y^{\dagger} b_y^{\dagger}) | vac \rangle \right] = \\ &= \frac{1}{2} \left[\langle vac | (a_x b_x + a_y b_y) (a_x^{\dagger} a_x a_x^{\dagger} b_x^{\dagger} \sin^2 \alpha + a_x^{\dagger} a_y a_y^{\dagger} b_y^{\dagger} \sin^2 \alpha + \right. \\ &+ a_x^{\dagger} a_y a_x^{\dagger} b_x^{\dagger} \sin \alpha \cos \alpha + a_x^{\dagger} a_y a_y^{\dagger} b_y^{\dagger} \sin \alpha \cos \alpha + \\ &+ a_y^{\dagger} a_x a_x^{\dagger} b_x^{\dagger} \sin \alpha \cos \alpha + a_y^{\dagger} a_x a_y^{\dagger} b_y^{\dagger} \sin \alpha \cos \alpha + \\ &+ a_y^{\dagger} a_y a_x^{\dagger} b_x^{\dagger} \cos^2 \alpha + a_y^{\dagger} a_y a_y^{\dagger} b_y^{\dagger} \cos^2 \alpha) | vac \rangle \left. \right] = (\text{только первое, четвертое, пятое и восьмое} \\ &\text{слагаемые отличны от нуля)} = \\ &= \frac{1}{2} \left[\langle vac | a_x b_x a_x^{\dagger} a_x a_x^{\dagger} b_x^{\dagger} \sin^2 \alpha + a_x b_x a_x^{\dagger} a_y a_y^{\dagger} b_y^{\dagger} \sin \alpha \cos \alpha + \right. \\ &+ a_x b_x a_y^{\dagger} a_x a_x^{\dagger} b_x^{\dagger} \sin \alpha \cos \alpha + a_x b_x a_y^{\dagger} a_y a_y^{\dagger} b_y^{\dagger} \cos^2 \alpha + \\ &+ a_y b_y a_x^{\dagger} a_x a_x^{\dagger} b_x^{\dagger} \sin^2 \alpha + a_y b_y a_x^{\dagger} a_y a_y^{\dagger} b_y^{\dagger} \sin \alpha \cos \alpha + \\ &+ a_y b_y a_y^{\dagger} a_x a_x^{\dagger} b_x^{\dagger} \sin \alpha \cos \alpha + a_y b_y a_y^{\dagger} a_y a_y^{\dagger} b_y^{\dagger} \cos^2 \alpha | vac \rangle \left. \right] = (\text{только первое и восьмое слагаемые} \\ &\text{отличны от нуля)} = \frac{1}{2} \left[\langle vac | \sin^2 \alpha + \cos^2 \alpha | vac \rangle \right] = \frac{1}{2} \left[(\sin^2 \alpha + \cos^2 \alpha) \langle vac | vac \rangle \right] = \frac{1}{2} - \text{не зависит} \\ &\text{от угла } \alpha! \end{aligned}$$

Физически это происходит потому, что волновая функция (2) не факторизуется и нет смысла говорить о состояниях в модах **a** и **b** по отдельности. Таким образом, нельзя утверждать, что мода **a** находится в суперпозиционном состоянии (3)!

Замечание. Прделанные вычисления (Путь II) вовсе не доказывают, что состояние в моде **a** неполяризованное. Например, при наличии в этой моде циркулярно-поляризованного света, результат получился бы таким же. Строгое доказательство - например, через параметры Стокса (в разделе V).

Заметим, что действуя таки же образом, можно доказать, что состояние в моде **a** до элемента CNOT - поляризованное.



Здесь усреднение нужно проводить по волновой функции исходного состояния (3). Результат получается таким:

$$R_a \sim \langle \Psi | a'^{\dagger} a' | \Psi \rangle = \frac{1}{2}(1 + \sin 2\alpha), \text{ т.е. максимум совпадений достигается при } \alpha = 45^{\circ}.$$

Информация и энтропия.

Не вводя пока “операционального” термина “информация” будем рассуждать, пользуясь “бытовым” языком. Т.е. информация - это некое знание об объекте.

За то, что понятия информация и энтропия тесно связаны, говорит следующий пример. Рассмотрим идеальный газ, находящийся в термодинамическом равновесии. Газ состоит из огромного количества молекул, которые двигаются в объеме V . Параметрами состояния являются давление, температура. Число состояний такой системы огромно. Энтропия газа при ТД равновесии максимальна и как следует из формулы Больцмана, определяется числом микросостояний системы. При этом мы ничего не знаем о том, какое конкретно состояние имеет система в данный момент времени у нас нет - информация минимальна. Допустим, что каким-то образом нам удалось с помощью очень быстрого прибора “подсмотреть состояние системы в данный момент времени. Значит мы получили о ней какую-то информацию. Можно даже представить, что мы сфотографировали не только координаты молекул, но и их скорости (например, сделав несколько фотографий одну за другой). При этом в каждые моменты времени, когда нам доступна информация о состоянии системы, энтропия стремится к нулю, т.к. система находится лишь в каком-то одном определенном состоянии из всего огромного их многообразия и это состояние сильно неравновесное. Этот пример показывает, что действительно информация и энтропия как-то связаны, причем уже вырисовывается характер связи: чем больше информация, тем меньше энтропия.

Сведения из термодинамики и статистической физики.

Физические величины, характеризующие макроскопические состояния тел (много молекул), называют термодинамическими (в том числе, энергия, объем). Существуют, однако, и величины, появляющиеся как результат действия чисто статистических

закономерностей и имеющие смысл в применении только к макроскопическим системам. Такова, например, энтропия и температура.

Классическая статистика

***Теорема Лиувилля.** Функция распределения постоянна вдоль фазовых траекторий подсистемы (речь идет о квазизамкнутых подсистемах, поэтому теорема справедлива только для не очень больших промежутков времени, в течение которых подсистема ведет себя как замкнутая).

$$\frac{d\rho}{dt} = \sum_i \left(\frac{\partial \rho}{\partial q_i} \dot{q}_i + \frac{\partial \rho}{\partial p_i} \dot{p}_i \right) = 0$$

Здесь - ρ - функция распределения или плотность вероятности. Она вводится через вероятность ω обнаружить подсистему в элементе фазового пространства $\Delta p \Delta q$ в данный момент времени: $d\omega = \rho(p_1, \dots, p_s, q_1, \dots, q_s) dp dq$, причем

$$\int \rho dp dq = 1. \quad (4)$$

Нахождение статистического распределения ρ для любой подсистемы и является основной задачей статистики. Если статистическое распределение известно, то можно вычислить вероятности различных значений любых физических величин, зависящих от состояний этой подсистемы (т.е. от значений координат и импульсов):

$$\bar{f} = \int f(p, q) \rho(p, q) dp dq.$$

***Микроканоническое распределение.**

Распределение ρ_{12} для совокупности двух подсистем (они полагаются замкнутыми, т.е. слабовазаимодействующими) равно $\rho_{12} = \rho_1 \rho_2$. Поэтому $\ln \rho_{12} = \ln \rho_1 + \ln \rho_2$ - логарифм функции распределения - величина *аддитивная*. Из теоремы Лиувилля следует, что функция распределения должна выражаться через такие комбинации переменных p и q , которые при движении подсистемы, как замкнутой, должны оставаться постоянными (такие величины называются интегралами движения). Значит сама функция распределения является интегралом движения. Более того, ее логарифм - тоже интеграл движения, причем *аддитивный*. Всего в механике существует семь интегралов движения - энергия, три компоненты импульса и три компоненты момента импульса - (для подсистемы a : $E_a(p, q)$, $P_a(p, q)$, $M_a(p, q)$). Единственная аддитивная комбинация этих величин есть

$\ln \rho_a = \alpha_a + \beta_a E_a(p, q) + \gamma P(p, q) + \delta M(p, q)$, причем коэффициенты β , γ , δ (и= семь штук)- должны оставаться одинаковыми для всех подсистем данной замкнутой системы, а α выбирается из условий нормировки (4).

Чтобы выполнялось условие нормировки (4), необходимо, чтобы функция $\rho(p, q)$ обращалась в точках E_0 , P_0 , M_0 в бесконечность. Более точная формулировка дает выражение

$$\rho = \text{const} \delta(E - E_0) \delta(P - P_0) \delta(M - M_0). \text{ - микроканоническое распределение.}$$

Наличие δ - функций обеспечивает обращение ρ в нуль для всех точек фазового пространства, в которых хотя бы одна из величин E , P , M не равна своему заданному (среднему) значению E_0 , P_0 , M_0 .

От шести интегралов P и M можно избавиться, заключив систему в твердый ящик, в котором она покоится.

Тогда

$$\rho = \text{const} \delta(E - E_0).$$

Физическая энтропия

Опять используем понятие идеального газа.

Пусть одноатомный идеальный газ с плотностью n и температурой T занимает объем V . Будем измерять температуру в энергетических единицах - не будет фигурировать постоянная Больцмана. Каждый атом газа имеет среднюю кинетическую энергию теплового движения, равную $3T/2$. Поэтому полная тепловая энергия газа равна

$$E = \frac{3}{2} TnV$$

Известно, что давление газа равно $p = nT$. Если газ может обмениваться теплом с внешней средой, то закон сохранения энергии газа выглядит так:

$$dE = -pdV + dQ. \quad (5)$$

Таким образом, изменение внутренней энергии газа может происходить как за счет совершаемой им работы, так и вследствие поступления некоторого количества тепла dQ извне. Это уравнение выражает первое начало термодинамики, т.е. закон сохранения энергии. При этом предполагается, что газ находится в равновесии, т.е. $p = const$ по всему объему.

Если же допустить, что газ находится и в состоянии ТД равновесия, $T = const$, то соотношение (5) можно рассматривать как элементарный процесс вариации параметров газа при их очень медленном изменении, когда ТД равновесие не нарушается. Именно для таких процессов и вводится понятие энтропии S с помощью соотношения

$$dS = \frac{dQ}{T} \quad (6)$$

Таким образом, утверждается, что у равновесного газа кроме внутренней энергии есть еще одна внутренняя характеристика, связанная с тепловым движением атомов. Согласно (5, 6) при постоянном объеме $dV = 0$, изменение энергии пропорционально изменению температуры, а в общем случае

$$dE = -pdV + TdS$$

Так как $p = nT = \frac{N}{V}T$, $E = \frac{3}{2}NT$, где $N = nV = const$ есть полное количество атомов газа, то последнее соотношение можно записать в виде

$$dS = N \left(\frac{3}{2} \frac{dT}{T} + \frac{dV}{V} \right).$$

После интегрирования получаем

$$S = N \left[\ln \left(VT^{3/2} \right) + const \right] \equiv NS.$$

Выражение в квадратных скобках представляет собой энтропию, приходящуюся на одну частицу.

Таким образом, если и температура и объем изменяются таким образом, что $VT^{3/2}$ остается постоянным, то и энтропия S не изменяется. Согласно (6) это означает, что газ не обменивается теплом с внешней средой, т.е. газ отделен от нее теплоизолирующими стенками. Такой процесс называется *адиабатическим*.

Поскольку

$$VT^{3/2} = const, \quad p = \frac{N}{V}T \rightarrow T = \frac{pV}{N} \rightarrow pV^\gamma = const.$$

где $\gamma = 5/3$ называется показателем адиабаты. Таким образом при адиабатическом процессе температура и давление изменяются с плотностью по закону

$$T = \text{const } n^{\gamma-1}, \quad p = \text{const } n^{\gamma}.$$

Формула Больцмана

Как следует из теоремы Лиувилля, функция распределения ρ имеет резкий максимум при $E = E_0$ (среднее значение) и отлична от нуля только в окрестности этой точки. Если ввести ширину ΔE кривой $\rho(E)$, определив ее как ширину прямоугольника, высота которого равна значению функции $\rho(E)$ в точке максимума, а площадь равна единице $\rho(E_0 = \bar{E})\Delta E = 1$. Можно перейти от интервала значений энергии к числу состояний $\Delta\Gamma$ с энергиями, принадлежащими ΔE (это, фактически, средняя флуктуация энергии системы). Тогда величина $\Delta\Gamma$ характеризует степень размазанности макроскопического состояния системы по ее микроскопическим состояниям. Другими словами, для классических систем $\Delta\Gamma$ - это размер той области фазового пространства, в которой данная подсистема проводит почти все время $\rho(E_0 = \bar{E})\Delta p\Delta q = 1$, $\Delta\Gamma \sim \Delta p\Delta q$. В квазиклассической теории устанавливается соответствие между объемом области фазового пространства и приходящимся на него числом квантовых состояний. А именно, на каждое квантовое состояние в фазовом пространстве приходится клетка с объемом $(2\pi\hbar)^s$, где s - число степеней свободы.

Величину $\Delta\Gamma$ называют статистическим весом макроскопического состояния, его можно записать в виде:

$$\Delta\Gamma = \frac{\Delta p\Delta q}{(2\pi\hbar)^s}.$$

Логарифм статистического веса называется энтропией:

$$S = \kappa \ln \Delta\Gamma.$$

где $\Delta\Gamma$ - статистический вес = число микросостояний, охватываемых рассматриваемым макросостоянием системы.

$$\frac{S}{\kappa} = \ln \Delta\Gamma \rightarrow e^{\frac{S}{\kappa}} = \Delta\Gamma \rightarrow \Delta\Gamma = e^{\frac{TS}{\kappa T}}.$$

В квантовой статистике показывается, что $\rho(E_0)\Delta\Gamma = 1$. Тогда

$$S = \ln \Delta\Gamma = -\ln \rho(E_0), \text{ где под } \rho \text{ понимается статистическая матрица (плотности).}$$

Ввиду линейности функции распределения по энергии $S = \ln \Delta\Gamma = -\ln \rho(E_0) = -\langle \ln \rho(E) \rangle$, где усреднение проводится по функции распределения ρ .

Поскольку число состояний $\Delta\Gamma$ во всяком случае не меньше единицы, то энтропия не может быть отрицательной. S определяет густоту уровней энергетического спектра макроскопической системы. Ввиду аддитивности энтропии можно сказать, что средние расстояния между уровнями макроскопического тела экспоненциально убывают с увеличением его размеров (т.е. числа частиц в нем). Наибольшее значение энтропии соответствует полному статистическому равновесию.

Характеризуя каждое макроскопическое состояние системы распределением энергии между различными подсистемами, можно сказать, что ряд последовательно проходимых системой состояний соответствует все более вероятному распределению энергии. Это возрастание вероятности велико в силу его экспоненциального характера e^S - в экспоненте стоит аддитивная величина - энтропия. Т.о. процессы, протекающие в неравновесной замкнутой системе, идут таким образом, что система непрерывно

переходит из состояний с меньшей энтропией в состояния с большей энтропией. В итоге энтропия достигает наибольшего возможного значения, соответствующего полному статистическому равновесию.

Таким образом, если замкнутая система в некоторый момент времени находится в неравновесном макроскопическом состоянии, то наиболее вероятным следствием в последующие моменты времени будет монотонное возрастание энтропии системы. Это - *второй закон термодинамики* (Р.Клаузиус, 1865г.). Его статистическое обоснование дано Л.Больцманом в 1870г. Другое определение:

если в некоторый момент времени энтропия замкнутой системы отлична от максимальной, то в последующие моменты энтропия не убывает. Она увеличивается или в предельном случае остается постоянной. Соответственно этим двум возможностям все происходящие с макроскопическими телами процессы принято делить на *необратимые* и *обратимые*. *Необратимые* - те процессы, которые сопровождаются увеличением энтропии всей замкнутой системы (процессы, которые бы являлись их повторениями в обратном порядке, не могут происходить, так как при этом энтропия должна была бы уменьшаться). Заметим, что уменьшение энтропии может быть вызвано флуктуациями. *Обратимыми* называются процессы, при которых энтропия замкнутой системы остается постоянной и которые, следовательно, могут проходить и в обратном направлении. Строго обратимый процесс представляет собой идеальный предельный случай.

При адиабатических процессах система не поглощает и не отдает тепло $\Delta Q = 0$.

Замечание: (существенное). Утверждение о том, что замкнутая система должна в течение достаточно длительного времени (большего, чем время релаксации) перейти в состояние равновесия относится лишь к системе, находящейся в стационарных внешних условиях. Пример - поведение доступной нашему наблюдению большой области Вселенной (свойства природы не имеют ничего общего со свойствами равновесной системы).

Информация.

Рассмотрим ленту, разбитую на ячейки - классический регистр. Если в каждой ячейке может быть помещен только один из двух символов, то говорят, что в ячейке содержится бит информации. Очевидно (см. лекцию 1), что в регистре, содержащем N ячеек содержится N бит информации и в нем можно записать 2^N сообщений. Итак, информационная энтропия измеряется в битах:

$$H_B \equiv N \equiv \log_2 Q_N. \quad (7)$$

Здесь $Q_N = 2^N$ - полное число различных сообщений. Из (7) ясно, что *информационная энтропия просто равна минимальному числу двоичных ячеек, с помощью которых можно записать некую информацию.*

Определение (7) можно переписать по-другому. Пусть у нас имеется множество Q_N различных сообщений. Найдем вероятность того, что необходимое нам сообщение совпадет со случайно выбранным из общего числа Q_N различных сообщений. Она равна, очевидно, $P_N = 1/Q_N$. Тогда определение (7) запишется как:

$$H_B \equiv N \equiv -\log_2 P_N. \quad (8)$$

Чем больше число ячеек N , тем меньше вероятность P_N и тем больше информационная энтропия H_B , содержащейся в данном конкретном сообщении.

Пример. Число букв алфавита равно 32 (без буквы ё). Число 32 есть пятая степень двойки $32 = 2^5$. Чтобы каждой букве сопоставить определенную комбинацию двоичных

чисел необходимо иметь 5 ячеек. Добавив к строчным буквам заглавные, мы удваиваем число символов, которые хотим закодировать - их станет $64 = 2^6$ - т.е. добавляется лишний бит информации $H_B = 6$. Здесь H_B - объем информации, приходящийся на одну букву (строчную или заглавную). Однако такой прямой подсчет информационной энтропии не совсем точен, поскольку в алфавите есть буквы, которые встречаются реже или чаще. Тем буквам, которые встречаются реже, можно отдать большее количество ячеек, а на часто встречающихся буквах - сэкономить и отдать им те состояния регистра, которые занимают меньшее количество ячеек. Точное определение информационной энтропии было дано Шенноном:

$$H \equiv -\sum_i p_i \ln p_i. \quad (9)$$

Формально вывод этого соотношения можно обосновать следующим образом.

Мы показали выше, что

$$S = \ln \Delta\Gamma = -\ln \rho(E_0) = -\langle \ln \rho(E) \rangle$$

из-за аддитивности логарифма функции распределения.

Пусть p - функция распределения какой-нибудь дискретной величины f_i (например, буквы "о" в этом тексте). Если с помощью функции p построить функцию распределения вероятностей различных значений величины $f = f_1, f_2, \dots, f_N$, то эта функция будет иметь максимум при $f = \bar{f}$, где $\bar{f} = \int f p d\Gamma$ и $\int p d\Gamma = 1$ (нормировка). Тогда $p(\bar{f})\Delta\Gamma = 1$ и $H \equiv -\ln p(\bar{f}) = -\langle \ln p(f_n) \rangle = -\sum_i p_i \ln p_i$.

Суммирование ведется по всем символам (буквам алфавита), а p_i означает вероятность появления символа с номером i . Как видно это выражение охватывает как часто используемые буквы, так и буквы, вероятность появления которых в данном сообщении мала.

Поскольку в выражении (9) используется натуральный логарифм, соответствующую единицу информации называют "нат".

Выражение (9) можно переписать в виде

$$H \equiv -\langle \ln p_i \rangle, \quad (10)$$

где скобки означают обычное классическое усреднение с помощью функции распределения p_i .

Замечание. В следующих лекциях будет показано, что для квантовых состояний

$$H_{quant} \equiv -\langle \ln p_i \rangle_\rho = -Sp \ln \rho.$$

где ρ - матрица плотности. Формально выражения (10) и (11) совпадают, однако есть и существенная разница. Классическое усреднение производится по ортогональным (собственным) состояниям системы, в то время как для квантового случая состояния могут быть и неортогональные (суперпозиции). Поэтому всегда $H_{quant} \leq H_{class}$!

В формулах (8) и (9) используются логарифмы при разных основаниях. В (8) - по основанию 2, а в (9) - по основанию e . Соответствующие этим формулам информационные энтропии можно легко выразить друг через друга. Воспользуемся соотношением, в котором M - произвольное число

$$M = 2^{\log_2 M} = e^{\ln M}.$$

Тогда, учитывая, что $H_{Bit} = \log_2 M$, а $H_{Nat} = \ln M$, получаем

$$e^{H_{Nat}} = 2^{H_{Bit}} \rightarrow H = H_{Nat} = \ln 2^{H_{Bit}} = H_{Bit} \ln 2.$$

Откуда

$$H_{Bit} = \frac{H_{Nat}}{\ln 2} = 1.44 H_{Nat} - \text{число бит почти в полтора раза больше числа нат!}$$

Рассуждая аналогично, можно получить соотношение между энтропиями, выраженными в тритах и битах:

$$M = 2^{\log_2 M} = 3^{\log_3 M} \rightarrow 2^{H_{Bit}} = 3^{H_{Trit}} \rightarrow H_{Bit} = \frac{H_{Trit}}{\log_3 2} = \frac{H_{Trit}}{0.631} = 1.58 H_{Trit}.$$

В компьютерной технике пользуются информацией по двоичному основанию (в битах). Для рассуждений в физике удобнее пользоваться информацией по Шеннону (в натах), которой можно характеризовать любую дискретную информацию. Всегда можно найти число соответствующих бит.

Связь энтропии и информации.

Рассмотрим идеальный газ, состоящий только из одной частицы (Кадамцев, динамика и информация). Это не абсурд. Если одна частица заключена в сосуде объемом V со стенками, находящимися при температуре T , то рано или поздно она придет в равновесие с этими стенками. В каждый момент времени она находится во вполне определенной точке пространства и с вполне определенной скоростью. Будем проводить все процессы настолько медленно, что частица успеет в среднем заполнить весь объем и многократно поменять величину и направление скорости при неупругих столкновениях со стенками сосуда. Таким образом, частица оказывает на стенки среднее давление, имеет температуру T и ее распределение по скоростям является максвелловским с температурой T . Эту систему из одной частицы можно адиабатически сжимать, можно менять ее температуру, давая ей возможность прийти в равновесие со стенками сосуда.

Среднее давление на стенку при $N = 1$, равно $p = T/V$, а средняя плотность $n = 1/V$. Рассмотрим случай изотермического процесса, когда $T = const$. Из первого начала при $T = const$ и $p = T/V$ получаем

$$\Delta Q = TdS = pdV = T \frac{dV}{V}, \text{ поскольку } dE = 0.$$

Отсюда находим, что изменение энтропии не зависит от температуры, так что

$$TdS = T \frac{dV}{V} \rightarrow S = \ln \frac{V}{V_0}.$$

Здесь введена постоянная интегрирования: “размер частицы” $\ll V_0 \ll V$ - чтобы не нарушалось приближение идеального газа.

Работа при изотермическом процессе

$$W = \int pdV = T \int \frac{dV}{V} = T \ln \frac{V_2}{V_1} = T(S_2 - S_1)$$

работа определяется разностью энтропий.

Пусть у нас имеются идеальные перегородки, которыми можно поделить сосуд на части без затраты энергии. Разделим наш сосуд на две равные части с объемом $V/2$ каждая. При этом частица будет находиться в одной из половин - но мы не знаем в какой. Допустим, что у нас есть прибор, который позволяет определить в какой из частей находится частица, например, прецизионные весы. Тогда из симметричного распределения вероятностей 50% на 50% нахождения в двух половинках мы получаем 100% вероятности для одной из половин - происходит “коллапс” распределения

вероятностей. Соответственно, новая энтропия $\Delta S = \ln \frac{V}{2V_0}$ окажется меньше исходной

энтропии на величину $\Delta S = S_1 - S_2 = \ln \frac{V}{V_0} - \ln \frac{V}{2V_0} = \ln V - \ln V_0 - \ln V + \ln 2 + \ln V_0 = \ln 2$.

За счет уменьшения энтропии можно совершить работу. Для этого достаточно двигать перегородку в сторону пустого объема вплоть до его исчезновения. Работа будет равна $W = T\Delta S = T \ln 2$. Если бы во внешнем мире ничего не менялось, то повторяя эти циклы, можно построить вечный двигатель второго рода. Но второй закон термодинамики запрещает получение работы только за счет тепла. Значит во внешнем мире должно что-то происходить. Что же это? Обнаружение частицы в одной из половин **меняет информацию о частице** - из двух возможных половинок указывается только одна, в которой находится частица. Это знание соответствует одному биту информации. Процесс измерения уменьшает энтропию частицы (перевод в неравновесное состояние) и ровно настолько же увеличивает информацию о системе (частице). Если совершать повторные деления пополам полученной ранее половинки, четвертушки, восьмушки и т.д., то энтропия будет последовательно уменьшаться, а информация - увеличиваться! Другими словами

$$S + I = \text{const.}$$

Чем больше известно о физической системе, тем меньше ее энтропия. Если о системе известно все - это значит, что мы перевели ее в сильнонеравновесное состояние, когда ее параметры максимально удалены от равновесных значений. Если в нашей модели частицу удастся поместить в элементарную ячейку объема V_0 , то при этом $S = 0$, а информация достигает своего максимального значения $I_{\max} = -\ln p_{\min} = \ln \frac{V}{V_0}$,

поскольку вероятность p_{\min} найти частицу в данной ячейке равна V_0/V . Если в последующие моменты времени частица начнет заполнять больший объем, то информация будет утрачиваться, а энтропия - расти. Подчеркнем, что за информацию нужно платить (по второму началу) увеличением энтропии S_e внешней системы, причем $\Delta S_e > I$. Действительно, если бы за один бит информации прибор (внешняя система) увеличивал свою энтропию на величину ΔS_e меньшую одного бита, то мы могли бы обратить тепловую машину. А именно, расширяя объем, занятый частицей, мы бы увеличивали ее энтропию на величину $\ln 2$, получая работу $T \ln 2$, а суммарная энтропия системы частица плюс прибор уменьшилась бы. Но это невозможно по второму началу.

Итак, информационная энтропия - это мера недостатка (или степень неопределенности) информации о действительном состоянии физической системы.

Информационная энтропия Шеннона:

$$H = \log_2 \Delta\Gamma = -\sum_n p_n \log_2 p_n, \text{ где } \sum_n p_n = 1. \text{ (это относится к двухуровневым системам,}$$

типа бит: "0" и "1". Если размерность равна n , то $H = \log_n \Delta\Gamma$. Так, для $n = 3$, $H = \log_3 \Delta\Gamma$, причем, $\Delta\Gamma = 3$.)

Количество информации I (или просто информация) о состоянии классической системы, получаемое в результате измерений внешним прибором, связанным с рассматриваемой системой некоторым каналом связи, определяется как разность информационной энтропии, соответствующей начальной неопределенности состояния системы H_0 , и информационной энтропии конечного состояния системы после измерения H . Таким образом,

$$I + H = H_0 = \text{const.}$$

В идеальном случае, когда отсутствуют шумы и помехи, создаваемые внешними источниками в канале связи, конечное распределение вероятностей после измерения сводится к одному определенному значению $p_n = 1$, т.е. $H = 0$, а максимальное значение полученной при измерении информации будет определяться: $I_{max} = H_0$. Таким образом, информационная энтропия Шеннона системы имеет смысл максимальной информации, заключенной в системе; она может быть определена в идеальных условиях измерения состояния системы в отсутствие шумов и помех, когда энтропия конечного состояния равна нулю:

$$H = 0.$$

Рассмотрим классический логический элемент, который может находиться в одном из двух равновероятных логических состояний “0” и “1”. Такой элемент вместе с окружающей средой - термостатом и генерируемым внешним теплоизолированным объектом сигналом единую неравновесную замкнутую систему. Переход элемента в одно из состояний, например, в состояние “0”, соответствует уменьшению стат. веса его состояния по сравнению с начальным состоянием в 2 раза (для трехуровневых систем - в 3 раза). Найдем уменьшение *информационной энтропии* Шеннона, которое соответствует увеличению количества информации об элементе на единицу, которая называется *битом*:

$$\Delta H = -\Delta I = \log_2 (\Delta\Gamma/2) - \log_2 \Delta\Gamma = \log_2 (2/2) - \log_2 2 = -\log_2 2 = -1\text{бит.}$$

Следовательно, информационная энтропия определяет число битов, которое требуется для кодирования информации в рассматриваемой системе или сообщении.