Теорема Цермело, аксиома выбора и лемма Цорна.

Определение 1 *Линейно упорядоченное множество $(X, \prec)$ называется вполне упорядоченным множеством (мы будем использовать сокращение ВУМ), если любое его подмножество $A \subset X$ содержит минимальный элемент (т.е. $\exists$ такой $a_0 \in A$, что $\forall a \in A$ $a_0 \prec a$).*

Простейшим примером ВУМ является множество натуральных чисел $\mathbb{N}$. Изоморфное (как ЧУМ) ему подмножество интервала $(0, 1)$ выглядит, например, так: $\{\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots \frac{n}{n+1}, \dots\}$. Теперь нетрудно привести более сложный пример: $M = \{m + \frac{n}{n+1}, \quad m, n \in \mathbb{N}\}$. (Нарисуйте!) Дальше этот пример можно усложнить, переведя, например, множество M в интервал $(0, 1)$ монотонной функцией $f(x) = 1 - \frac{1}{x}$ и рассмотрев множество, состоящее из всевозможных сумм $x + y$, где $x \in f(M)$, а $y \in \mathbb{N}$. Далее это построение можно многократно повторять. Таким образом, мы имеем много разнообразных примеров счетных вполне упорядоченных множеств. А вот построить такой же наглядный пример несчетного вполне упорядоченного множества невозможно. Более того, оказалось, что утверждение, что на любом множестве можно ввести структуру ВУМ (его обычно называют теоремой Цермело), не зависит от остальных аксиом теории множеств. Обсуждению этого утверждения и двух, равносильных ему (аксиомы выбора и леммы Цорна) как раз и посвящена эта лекция.

Отметим несколько простых свойств вполне упорядоченных множеств.

Во-первых любое ВУМ X имеет минимальный элемент, это следует из определения 1, если в качестве взять A само X. А вот наибольшего элемента в ВУМ, как мы видели на примерах, может и не быть.

Во-вторых, всякий элемент $z \in X$ имеет непосредственного последователя, то есть такой элемент $a \in X$, что $z \prec a$ и из $z \prec x \prec a$ следует, что либо $x = z$ либо $x = a$. Для доказательства достаточно взять в определении 1 в качестве A множество $\{x \in X,$ таких что $z \prec x$ и $z \neq x\}$. А вот непосредственного предшественника, как мы видели на примерах, может и не быть.

В дальнейшем мы для обозначения того, что $z \prec x$ и $z \neq x$ будем использовать обозначение $z \precneqq x$.

Очевидно, что любое подмножество в ВУМ само также является ВУМ.

Еще одно важное определение.

Определение 2 *Дано ВУМ X. Подмножество $Y \subset X$ называется начальным отрезком X, если из $y \in Y$ и $z \prec y$ следует $z \in Y$.*

Нетрудно видеть, что любой (отличный от X) начальный отрезок Y можно описать как

$$Y = S_a = \{x \in X, \text{ таких что } x \precneqq a\}$$

для некоторого $a \in X$ — в качестве a нужно взять минимальный элемент множества $X \setminus Y$. Из этого сразу следует, что из любых двух начальных отрезков один содержится внутри другого. Отметим еще, что $\emptyset$ также является начальным отрезком.

Оказывается, ВУМ является очень "жестким" объектом.

Теорема 1 *У ВУМ нет автоморфизмов, кроме тождественного, и, более того, ВУМ не изоморфно никакому своему собственному начальному отрезку.*

Напомним, что изоморфизмом двух упорядоченных множеств A и B называется биекция $f : A \to B$, согласованная с имеющимися на множествах A и B отношениями порядка, т.е. такая, что $a < a'$ равносильно $f(a) < f(a')$.

Для доказательства теоремы рассмотрим некоторый начальный отрезок $Y \subset X$ и изоморфизм $f : X \to Y$. Для доказательства теоремы нам надо убедиться в том, что это возможно только если $Y = X$ и $f = \mathrm{Id}_X$.

Рассмотрим множество

$$Z = \{x \in X \text{ таких что } f(x) = x\}.$$

Если мы докажем, что множество Z совпадает со всем X то наша теорема будет доказана, так как из этого будет следовать, что начальный отрезок Y, являющийся образом отображения f, также совпадает со всем X, и отображение f тождественно.

Предположим противное, то есть что Z не совпадает со всем X, тогда множество $X \setminus Z$ непусто и обладает минимальным элементом a.

Что можно сказать об $f(a)$? Во-первых, $f(a) \neq a$, поскольку $a \notin Z$. Во-вторых, если бы было $f(a) \prec a$, то это значило бы, что $f(a) \in Z$, но тогда $f(f(a)) = f(a)$, откуда в силу биективности f следовало бы $f(a) = a$, что противоречит тому, что $a \notin Z$. Следовательно, $a \precneqq f(a)$. Поскольку $f(a) \in Y$ и Y является начальным отрезком, это означает, что $a \in Y$ и потому можно рассмотреть его прообраз $f^{-1}(a) = c$. Соображения, аналогичные предыдущим, показывают, что $c = a$ или $c \prec a$ невозможно, поэтому $a \precneqq c$. Тем самым мы получили противоречие, поскольку из монотонности биекции f тогда должно следовать, что $f(a) \precneqq f(c) = a$, а мы только что получили противоположное неравенство. Теорема доказана.

Оказывается, вполне упорядоченные множества всегда сравнимы по мощности.

Теорема 2 *Если X и X' — два ВУМ, то одно из них изоморфно начальному отрезку другого.*

Для доказательства заметим, что в X можно найти такой начальный отрезок Y, что он изоморфен некоторому начальному отрезку $Y' \subset X'$, и пусть $f_Y : Y \to Y'$ соответствующий изоморфизм. (Например, в качестве Y можно взять начальный отрезок, состоящий из одного элемента — минимального элемента X, тогда Y' будет состоять из минимального элемента X'.) Назовем такие начальные отрезки хорошими. Обозначим через W объединение всех хороших начальных отрезков в X, а через W' — объединение изоморфных им начальных отрезков в X'. Заметим, что объединение начальных отрезков само является начальным отрезком. Действительно, если $y \in W$, то y является элементом какого-то начального отрезка Y, поэтому для любого $z \in X$ из $z \prec y$ следует $z \in Y$, но Y является подмножеством W, поэтому и $z \in W$.

Осталось заметить, что все изоморфизмы хороших отрезков можно склеить в один изомофизм $F : W \to W'$. Определим требуемое отображение F следующим образом. Пусть $x \in W$, тогда x является элементом какого-то хорошего начального отрезка Y, поэтому имеется изоморфизм $f_Y : Y \to Y'$. Положим тогда по определению $F(x) = f_Y(x)$.

Нам необходимо сначала доказать, что отображение F определено корректно, т.е. что если бы мы выбрали другой хороший начальный отрезок Z, содержащий x, то значения $f_Y(x)$ и $f_Z(x)$ совпадут. Из двух начальных отрезков Y и Z один обязательно содержится в другом, будем

для определенности считать, что $Y \subset Z$. Нетрудно понять, что тогда и $Y' \subset Z'$. Если бы это было не так, то имелось бы противоположное включение, $Z' \subset Y'$ и тогда бы $f_Y^{-1}(Z')$ был бы начальным отрезком в Y, изоморфным Z (изоморфизм $f_Y^{-1} \circ f_Z : Z \to f_Y^{-1}(Z')$), что противоречит теореме 1. Итак, имеем $Y \subset Z$, $Y' \subset Z'$, и два изоморфизма $f_Y : Y \to Y'$ и $f_Z : Z \to Z'$. Тогда $f_Z^{-1}(Y')$ является начальным отрезком в Z, изоморфным Y (изоморфизм $f_Z^{-1} \circ f_Y : Y \to f_Z^{-1}(Y')$), поэтому по той же теореме 1 эти отрезки должны совпадать, а изоморфизм между ними, который при их совпадении становится автоморфизмом, должен быть тождественным. Итак, $Y = f_Z^{-1}(Y')$, и $f_Z^{-1} \circ f_Y$ является тождественным на Y, то есть $\forall y \in Y \ f_Z^{-1}(f_Y(y)) = y$. Это означает, что $f_Z(y) = f_Y(y))$, что и доказывает корректность определения $F : W \to W'$. Осталось проверить, что это отображение является изоморфизмом, но это простое упражнение мы оставляем слушателям.

Итак, мы показали, что объединение W всех хороших начальных отрезков само является хорошим начальным отрезком, так что имеется изоморфизм $F : W \to W'$. Предположим, что $W \neq X$ и $W' \neq X'$, сейчас мы легко приведем это предположение к противоречию, что и завершит доказательство теоремы. Действительно, если оба наших предположения верны, то множества $X \setminus W$ и $X' \setminus W'$ непусты и потому в них имеются минимальные элементы, обозначим их t и t'. Но тогда легко видеть, что множество $\overline{W} = W \cup \{t\}$ также является хорошим начальным отрезком, поскольку оно очевидно изоморфно $\overline{W'} = W' \cup \{t'\}$: изоморфизм $\overline{F} : \overline{W} \to \overline{W'}$ совпадает с F на W и переводит t в t'. Тем самым мы нашли хороший начальный отрезок, не содержащийся в W, что противоречит определению W как объединения всех хороших отрезков. Следовательно, либо $W = X$, либо $W' = X'$, и наша терема доказана.

Перейдем теперь к аксиоме выбора, утверждающей, что для любого семейства непустых множеств можно образовать новое множество, содержащее по одному элементу из каждого множества семейства. Конечно, эта формулировка не является еще математической, и использованные в ней слова требуют уточнения.

Во-первых, что такое "семейство непустых множеств"? Это правило, которое каждому элементу некоторого множества B (которое часто называется *базой семейства*) сопоставляет некоторое множество. Поскольку мы договорились, что все рассматриваемые нами множества являются подмножествами некоторого множества Ω, сказанное означа-

ет, что у нас имеется некоторое отображение $F : B \to 2^\Omega \setminus \{\emptyset\}$. Теперь уже понятно, что "выбрать по одному элементу из каждого множества" означает задать функцию $f : B \to \Omega$, такую что $\forall b \in B \ f(b) \in F(b)$. Такую функцию часто называют *функцией выбора*. Так мы получаем первую формулировку аксиомы выбора.

A1. *Для любого отображения $F : B \to 2^\Omega \setminus \{\emptyset\}$ существует функция выбора $f : B \to \Omega$, такая что $\forall b \in B$ верно $f(b) \in F(b)$.*

Как мы сейчас увидим, достаточно потребовать наличия "универсальной" функции выбора, выбирающей по одному элементу из любого непустого подмножества Ω.

A2. *Существует "универсальная" функция выбора $\Phi : 2^\Omega \setminus \{\emptyset\} \to \Omega$, такая что для любого непустого подмножества $A \subset \Omega$ верно $\Phi(A) \in A$.*

Наконец, укажем еще одну удобную форма аксиомы выбора.

A3. *Для любой сюръекции $p : X \to Y$ существует правое обратное отображение $q : Y \to X$, т.е. такое, что $p \circ q = \mathrm{Id}_Y$.*

Теорема 3 *Формулировки аксиомы выбора* **A1, A2 и A3** *равносильны.*

Проще всего вывести A1 из A2: функция выбора f получается сразу как композиция F и "универсальной" функции выбора Φ: $f = \Phi \circ F$.

Чтобы вывести A3 из A1 достаточно заметить, что любое сюръективное отображение $p : X \to Y$ задает семейство подмножеств в X с базой Y: отображение $F : Y \to 2^X \setminus \{\emptyset\}$ определяется как $F(y) = p^{-1}(y)$. (Множество $p^{-1}(y) \subset X$ — прообраз элемента $y \in Y$, непустой в силу сюръективности p.) Тогда в качестве q можно взять как раз функцию выбора $f : Y \to X$, поскольку условие $f(y) \in F(y) = p^{-1}(y)$ как раз означает, что $p(f(y)) = y$, т.е. что $p \circ f = \mathrm{Id}_Y$.

Осталось вывести A2 из A3. Итак, нам дано только множество Ω, и для применения A3 нам нужно где-то взять сюръективное отображение. Его нам доставит следующая очень часто используемая конструкция. Рассмотрим декартово произведение $\Omega \times (2^\Omega \setminus \{\emptyset\})$ и в нем график соответствия принадлежности элемента подмножеству, т.е. подмножество $R \subset \Omega \times (2^\Omega \setminus \{\emptyset\})$, состоящее из таких пар (a, A), что $a \in A$:

$$R = \{(a, A), \text{ таких что } a \in A\}.$$

У прямого произведения $\Omega \times (2^\Omega \setminus \{\emptyset\})$ имеются две проекции на сомножители: $\pi_1 : \Omega \times (2^\Omega \setminus \{\emptyset\}) \to \Omega$ и $\pi_2 : \Omega \times (2^\Omega \setminus \{\emptyset\}) \to 2^\Omega \setminus \{\emptyset\}$, сопоставляющие каждой паре ее, соответственно, первую и вторую компоненту:

$\pi_1(a, A) = a$ и $\pi_2(a, A) = A$. Обозначим через p_1 и p_2 ограничения этих проекций на R. Очевидно, отображение p_2 является сюръекцией, это и есть искомая сюръекция $p_2 : R \to 2^\Omega \setminus \{\emptyset\}$. Согласно A3 у p_2 существует правое обратное отображение $q : 2^\Omega \setminus \{\emptyset\} \to R$, сопоставляющее непустому подмножеству $A \subset \Omega$ такую пару $q(A) = (\alpha, A)$, что $\alpha \in A$. Тогда универсальной функцией выбора будет композиция $p_1 \circ q : 2^\Omega \setminus \{\emptyset\} \to \Omega$.

Теорема доказана.

Лемма Цорна формулируется в терминах частично упорядоченных множеств. Напомним, что *цепью* в частично упорядоченном множестве X называется любое его линейно упорядоченное подмножество. (В частности, $\emptyset$ также является цепью.) Далее, элемент $a \in X$ называется *верхней гранью* цепи $Y \subset X$, если $y < a \ \forall y \in Y$. Обратите внимание, что верхняя грань цепи не обязательно является ее элементом! И, наконец, *максимальным элементом* частично упорядоченного множества X называется такой элемент $b \in X$, что из $b < x$ следует $b = x$. Обратите внимание, что в нашем определении не требуется, чтобы максимальный элемент был больше всех элементов частично упорядоченного множества X; требуется лишь, чтобы он был больше тех, с которыми он сравним.

Лемма Цорна. *Если любая цепь частично упорядоченного множества X обладает верхней гранью, то в X имеется максимальный элемент.*

На первый взгляд, из трех равносильных утверждений лемма Цорна имеет наиболее сложную формулировку, но, как показывает опыт, именно она наиболее удобна для применения в различных математических рассуждениях. Примеры такого применения мы оставим до следующей лекции, а пока что сформулируем и докажем основную теорему.

Теорема 4 *Теорема Цермело, аксиома выбора и лемма Цорна равносильны.*

Проще всего вывести аксиому выбора в формулировке A2 из теоремы Цермело: если множество Ω вполне упорядочено, то мы можем сопоставить любому подмножеству $A \subset \Omega$ его минимальный элемент — это и будет универсальная функция выбора.

Выведем теперь лемму Цорна из аксиомы выбора. Предположим, что в частично упорядоченном множестве X любая цепь обладает верхней гранью, а максимального элемента в X, тем не менее, нет. Мы хотим

привести это предположение к противоречию. Стратегия получения противоречия будет примерно следующая: отсутствие максимального элемента означает, что у любого элемента есть строго больший его, и, многократно пользуясь этим обстоятельством, мы построим "слишком длинную" цепь, у которой не будет верхней грани.

Чтобы реализовать этот план, заметим, что если $x \in X$ не является максимальным элементом, то $\exists y \in X$ такой, что $x < y$ и $y \neq x$. (Мы для такого случая будем использовать обозначение $x \lneqq y$.) следовательно, для любого $x \in X$ множество $H_x = \{y \in X \text{ таких что } x \lneqq y\}$ непусто, поэтому получается отображение $F : X \to 2^X \setminus \{\emptyset\}$, заданное тем, что $F(x) = H_x$. Следовательно, из аксиомы выбора следует, что имеется функция выбора $f : X \to X$, сопоставляющая каждому элементу $x \in X$ какой-то строго больший его элемент: $x \lneqq f(x)$.

Теперь из каждого элемента $x \in X$ мы можем сконструировать цепь $\{x, f(x), f(f(x)), \ldots\}$. У этой цепи, по предположению, есть верхняя грань, мы можем к этой грани применить функцию f и многократно повторять подобную процедуру, но, к сожалению, слова "многократно повторять" не имеют пока точного значения, и не очень ясно, как на этом пути может получиться противоречие. (Вспомните, как мы конструировали все более и более сложные счетные ВУМы в начале этой лекции.) Вместо этого воспользуемся аксиомой выбора еще раз, теперь для множества всех цепей множества X. Точнее, рассмотрим множество

$$C = \{Y \in 2^X, \text{ таких что } Y \text{ является цепью } \}$$

По предположению, любая цепь Y обладает верхней гранью $b \in X$, при этом b может быть или не быть элементом Y. Но элемент $f(b)$ уже точно не будет элементом Y, и при этом, конечно, он по транзитивности будет верхней гранью для Y. Следовательно, для любой цепи Y множество

$$H_Y = \{x \in X \setminus Y, \text{ таких что } x \text{ является верхней гранью } Y\}$$

непусто, и мы получаем новое отображение $G : C \to 2^X \setminus \{\emptyset\}$, заданное тем, что $G(C) = H_C$. Применяя к нему аксиому выбора, мы получим функцию выбора $g : C \to X$, сопоставляющую каждой цепи Y ее верхнюю грань $g(Y) \notin Y$.

Теперь мы готовы к построению "слишком длинной цепи"; нефомально говоря, мы хотим устроить ее таким образом, чтобы она в каждом месте продолжалась именно с помощью функции g. Чтобы фор-

мализовать это пожелание, дадим следующее определение: цепь Y называется хорошей, если она является вполне упорядоченным множеством и любой ее элемент y получается из ее начального отрезка

$$S_y = \{z \in Y, \text{ таких что } z \lneqq y\}$$

применением функции g: $y = g(S_y)$.

Нетрудно придумать примеры хороших цепей; такими цепями будут, например, $\{g(\emptyset)\}$, $\{g(\emptyset), g(\{g(\emptyset)\})\}$, $\{g(\emptyset), g(\{g(\emptyset)\}), g(\{g(\emptyset), g(\{g(\emptyset)\})\})\}$ и т.д.

Заметим, что у всех хороших цепей начало будет общим, действительно, если y_0 это минимальный элемент Y, то $S_{y_0} = \emptyset$, поэтому $y_0 = g(\emptyset)$. Из этого сразу легко вывести, что из любых двух хороших цепей Y и Y' одна всегда будет начальным отрезком другой. Действительно, если множества $Y \setminus Y'$ и $Y' \setminus Y$ оба непусты, то в них есть минимальные элементы $v \in Y \setminus Y'$ и $v' \in Y' \setminus Y$, причем $v \neq v'$. Но, с другой стороны, их начальные отрезки $S_v = \{z \in Y, \text{ таких что } z \lneqq v\}$ и $S_{v'} = \{z' \in Y', \text{ таких что } z' \lneqq yv'\}$ оба должны содержаться в $Y \cap Y'$ и потому совпадать, т.е. $S_v = S_{v'}$, но тогда должны совпасть и $v = g(S_v) = g(S_{v'}) = v'$, а мы только что видели, что они разные.

Теперь в качестве "слишком длинной цепи" можно взять объединение W всех хороших цепей. Нетрудно проверить, что W тогда тоже будет хорошей цепью. Но тогда $W \cup \{g((W))\}$ также будет хорошей цепью, но не будет подмножеством W — это и есть искомое противоречие.

Нам остался последний шаг: вывести теорему Цермело из леммы Цорна. Итак, дано некоторое множество X, мы хотим, пользуясь леммой Цорна, построить на нем отношения порядка, превращающего X во ВУМ. Схема применения леммы Цорна, как правило, выглядит так: если нам надо доказать существование чего-то (в данном случае отношения порядка, превращающего X в ВУМ), то нам достаточно придумать частично упорядоченное множество, в котором это искомое что-то очевидно было бы максимальным элементом, а потом проверить, что придуманное нами частично упорядоченное множество удовлетворяет условию леммы Цорна.

В нашем случае элементами частично упорядоченного множества W будут всевозможные пары $(A, <_A)$, где A — это некоторое подмножество в X, а $<_A$ это некоторое отношение частичного порядка на множестве A, превращающее A во ВУМ. Множество W очевидно, непусто; оно

содержит, например, все одноэлементные подмножества X. Отношение порядка $\prec$ на W мы определим следующим образом: $(A, <_A) \prec (B, <_B)$, если $A \subset B$ и ограничение отношения $<_B$ на множество A совпадает с $<_A$ (т.е. если $a, a' \in A$, то $a <_A a'$ верно тогда и только тогда, когда верно $a <_B a'$) и при этом ВУМ A является начальным отрезком ВУМа B.

Что представляет собой цепь V ЧУМа W? Элементами V являются вполне упорядоченные множества, причем из любых двух одно из них является начальным отрезком другого. Как и в прошлом доказательстве, рассмотрим их объединение

$$E = \bigcup_{(A,<_A)\in V} A.$$

На E естественно можно ввести отношение порядка $<_E$: любые два элемента $x, y \in E$ лежат в каком-нибудь A из V, и тогда мы по определению положим, что $x <_E y$ тогда и только тогда, когда $x <_A y$ — это, очевидно, не зависит от выбора A. Легко видеть, что E тогда также является ВУМом, причем каждое A из V является его начальным отрезком, поэтому $(E, <_E)$ является верхней гранью цепи V. Тем самым мы видим, что условие леммы Цорна выполнено, поэтому W содержит некоторый максимальный элемент $(M, <_M)$. Может ли подмножество M не совпадать со всем X? Если $M \neq X$, то существует $x \in X$, $x \notin M$. Тогда очень легко определить отношение порядка на множестве $N = M \cup \{x\}$ таким образом, чтобы N было ВУМом, а M его начальным отрезком. Для этого отношение порядка $<_N$ надо, конечно, определить так: при $m, m' \in M$ $m <_N m'$ тогда и только тогда, когда $m <_M m'$, и $m <_N x$ $\forall m \in M$. Тогда $(M, <_M) \prec (N, <_N)$, что противоречит предположению о максимальности $(M, <_M)$. Полученное противоречие доказывает, что $M = X$, и тем самым мы превратили наще множество X в ВУМ, что и требовалось.

Теорема о равносильности трех утверждений доказана.