Конспект

Мы предлагаем регуляризацию Моро-Йосиды для
максимальных монотонных операторов типа (D) в нерефлексивных банаховых пространствах. Она обобщает классическую регуляризацию Моро-Йосиды, а также расширение этой регуляризации **Брези-Крандалл-Пази** на строго выпуклые (рефлексивные) банаховы пространства со строго выпуклыми сопряженными. Наши основные результаты получены с использованием недавних результатов авторов о выпуклых представлениях максимальных монотонных операторов в нерефлексивных банаховых пространствах.

---

*Ключевое:

авторы **используют выпуклые представления операторов**, то есть аппроксимации через выпуклые функции, чтобы получить свои результаты.

Идея двойственной задачи

Когда мы решаем **оптимизационную задачу**, мы хотим **минимизировать** или **максимизировать** какую-то функцию, например:
$min_{x \in X} f(x) + g(x)$
Это называется **прямая задача** (или primal).

Вместо того, чтобы решать задачу напрямую, мы можем перейти в **сопряжённое пространство** X^* (пространство линейных функционалов на X) и построить **двойственную задачу**, в которой участвуют **сопряжённые функции** f^* и g^*:
$\max_{x^* \in X^*} -f^*(-x^*) - g^*(x^*)$
Это **двойственная задача** (или dual)

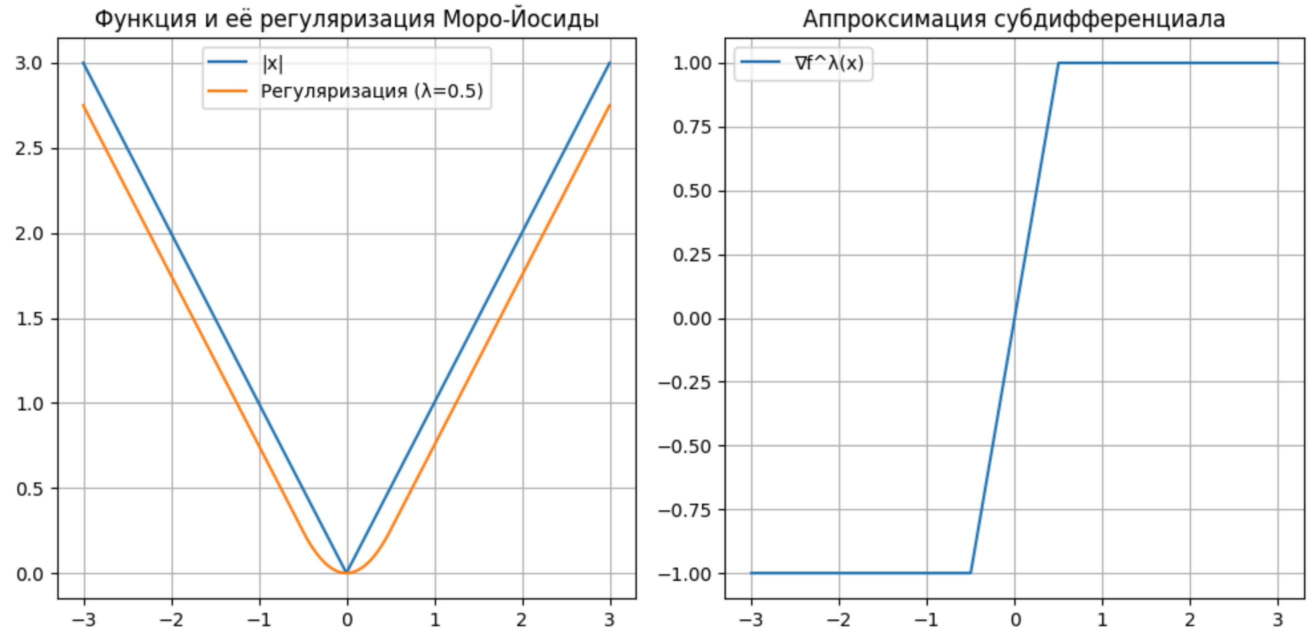
Аппроксимация субдифференциала

Так как ∂f — **не функция**, а **множество**, с ним тяжело работать. Поэтому мы хотим найти **гладкую приближающую функцию**, у которой есть производная и которая близка к поведению ∂f.

Вот тут и появляется **регуляризация Моро-Йосиды**:
- Она превращает «рваную» функцию типа $|x|$ в гладкую.
- А её производная становится **аппроксимацией субдифференциала**.

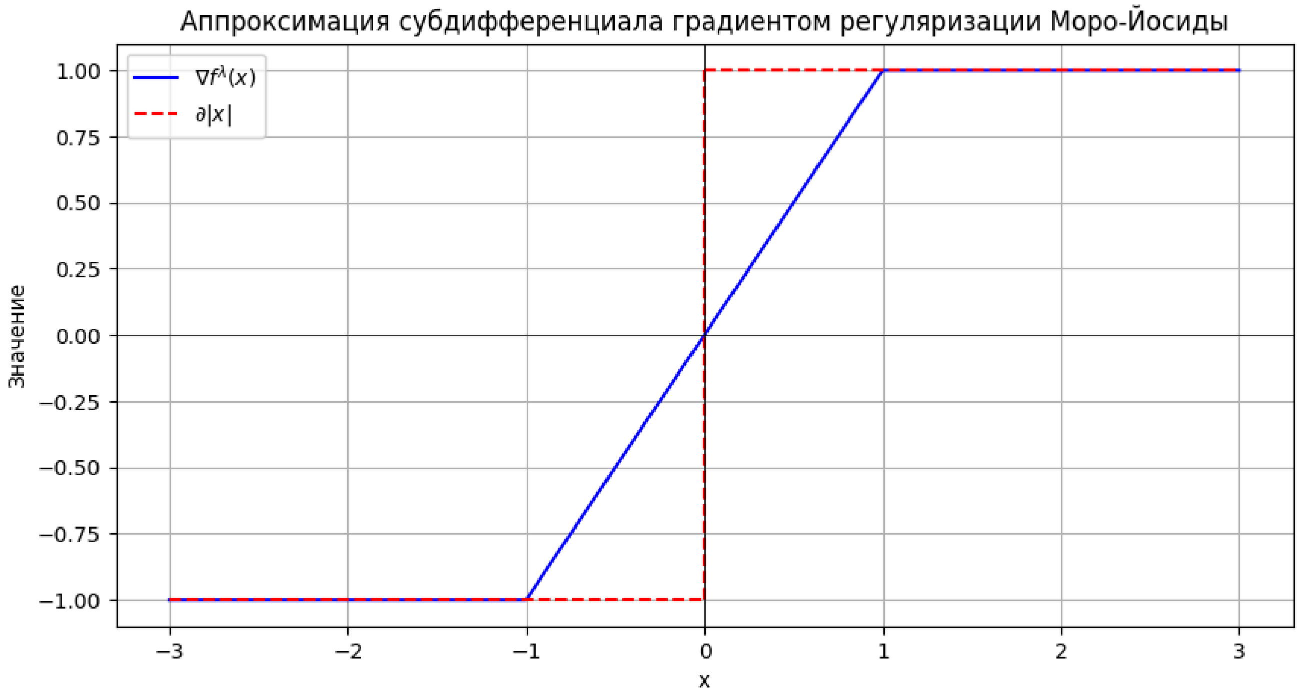
То есть:

$\nabla f\lambda(x) \approx \partial f(x)$ - это аппроксимация субдифференциала

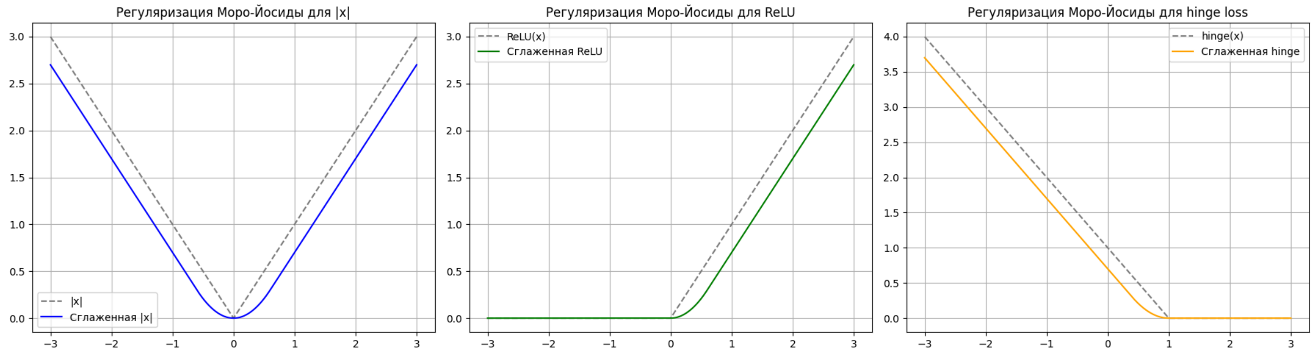


∇f^{λ}

субдифференциал — это **множество всех касательных прямых, которые лежат ниже графика функции**. Обозначается $\partial f(x)$



Регуляризация для других негладких функций:



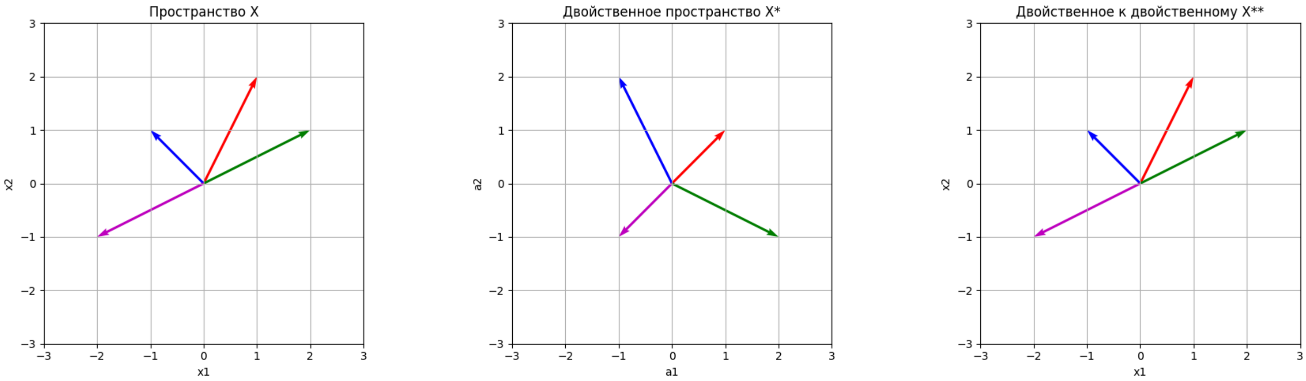
- Регуляризация **заменяет** оригинальную функцию f на её **сглаженное приближение**.
- Чем меньше λ, тем ближе $f^\lambda(x)$ к оригинальной функции (и наоборот — при увеличении λ сглаживание сильнее).

Рефлексивные банаховы пространства

Банахово пространство

Банахово пространство — это нормированное векторное пространство, полное по метрике, порождённой нормой (полное нормированное линейное пространство).
*Банахово пространство — это просто **пространство с нормой**, в котором можно **измерять длину и расстояния**, и в котором **все "почти-сходящиеся" последовательности действительно сходятся.**
Полное означает, что если у нас есть последовательность точек, которая всё ближе и ближе друг к другу (Коши-последовательность), то она обязательно **сойдётся к какой-то точке в этом пространстве**.
Двойственное (сопряженное) пространство — это просто **множество линейных функций**, которые "измеряют" вектора. Эти функции называют **функционалами**.
- Вектор $x = (x1, x2)$
- Функционал $f(x) = 2x_1 + 3x_2$ — это **линейный функционал**
- Пространство всех таких функционалов и есть **сопряжённое пространство**.
Рефлексивное означает, что если возять:
1. Пространство X
2. Потом его двойственное $X^{\star}$
3. Потом двойственное от двойственного $X^{\star\star}$
То в итоге **вернёмся обратно** к исходному пространству X. То есть, пространство **равно своему двойственному двойственному пространству**.
Крч Если знаешь как f ∈ $X^{\star}$ взаимодействует с x, то можно восстановить сам вектор.

Пример рефлективности:



1 Введение

X — **вещественное банахово пространство**.
Пример: $\mathbb{R}^n, L^p([a,b]), \ell^p$, и др.

X^* — **топологическое сопряжённое пространство** (или **дуальное пространство**).
Оно состоит из **всех непрерывных линейных функционалов** на X.
То есть, каждый элемент $x^* \in X^*$ — это функция:
$x^* : X \to \mathbb{R}, \quad x \mapsto x^*(x)$

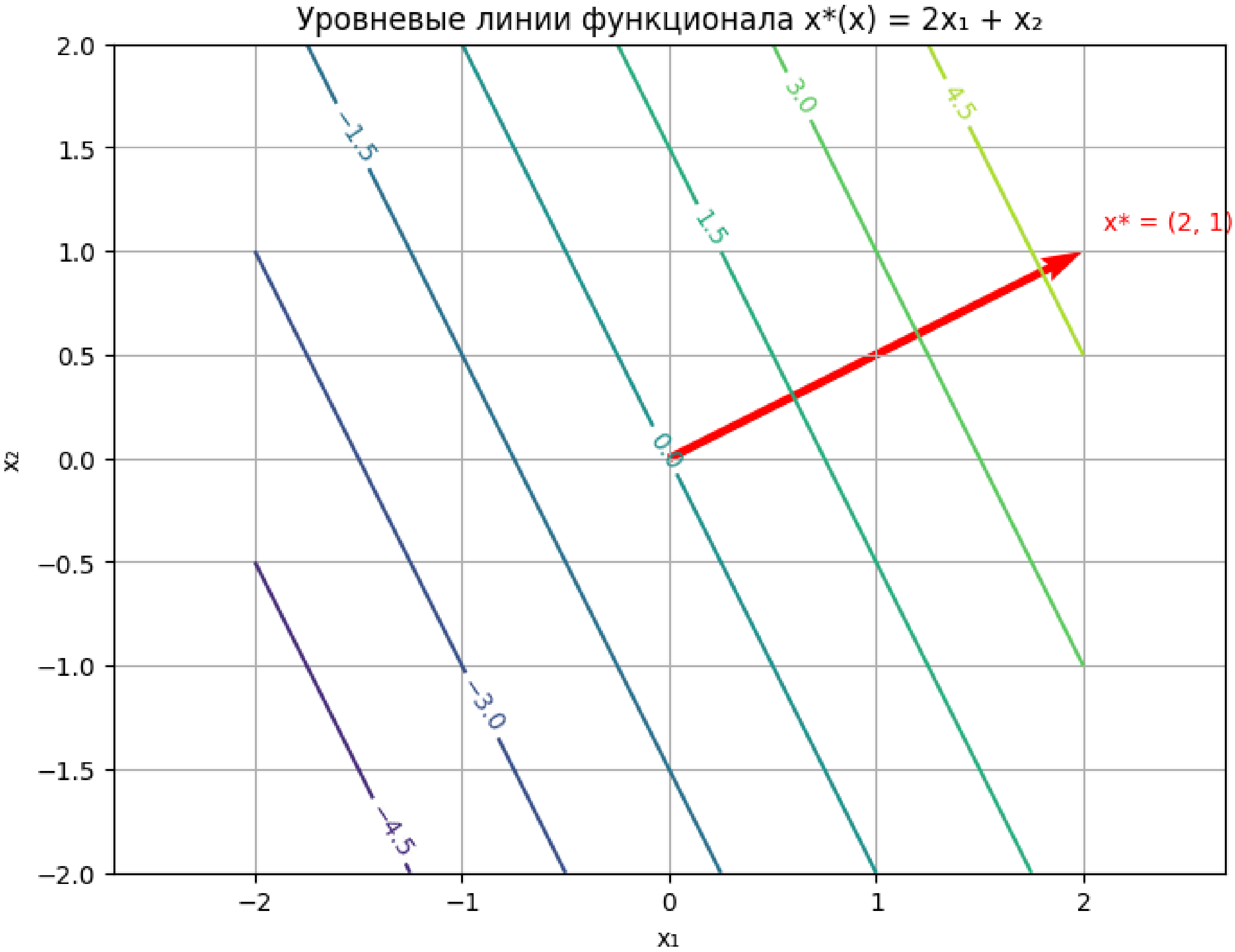
Пусть:

- $x = (x_1, x_2) \in \mathbb{R}^2$
- $x^* = (2, 1)$, тогда:

$$x^*(x) = 2x_1 + x_2$$

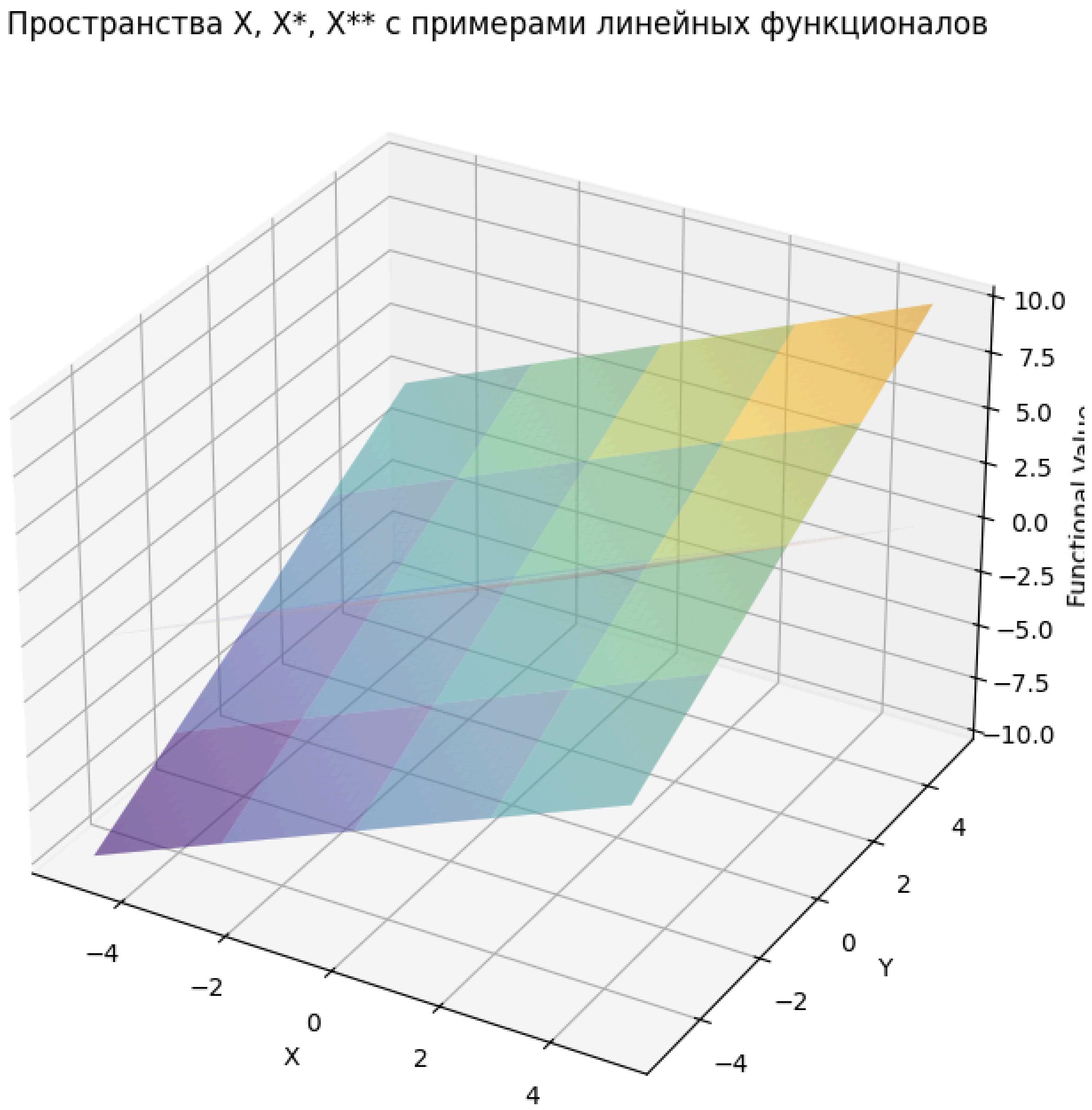
Это — линейная функция, значение которой зависит от положения точки x. Нарисуем **уровневые линии** $x^*(x) = c$, то есть:

$$2x_1 + x_2 = c \Rightarrow x_2 = c - 2x_1$$



X^{**} — **второе сопряжённое** пространство (или **бисопряжённое**).
Это множество **всех непрерывных линейных функционалов на $X^{\star}$.
То есть каждый $x^{\star\star} \in X^{\star\star}$ — это функция, которая действует на $x^{\star} \in X^{\star}$.

Плоскости на графике это линейные функционалы X^* (а X^* действует на X, а X будет соответствовать x который можно оценить с помощью линейного функционала. (X отображает вещ числа) sum up: это действия линейных функционалов на функционалах).



Двойственное скалярное произведение

Оно обозначается как:

$$\langle x, x^* \rangle = x^*(x), \quad \langle x^*, x^{**} \rangle = x^{**}(x^*), \quad x \in X, x^* \in X^*, x^{**} \in X^{**}.$$

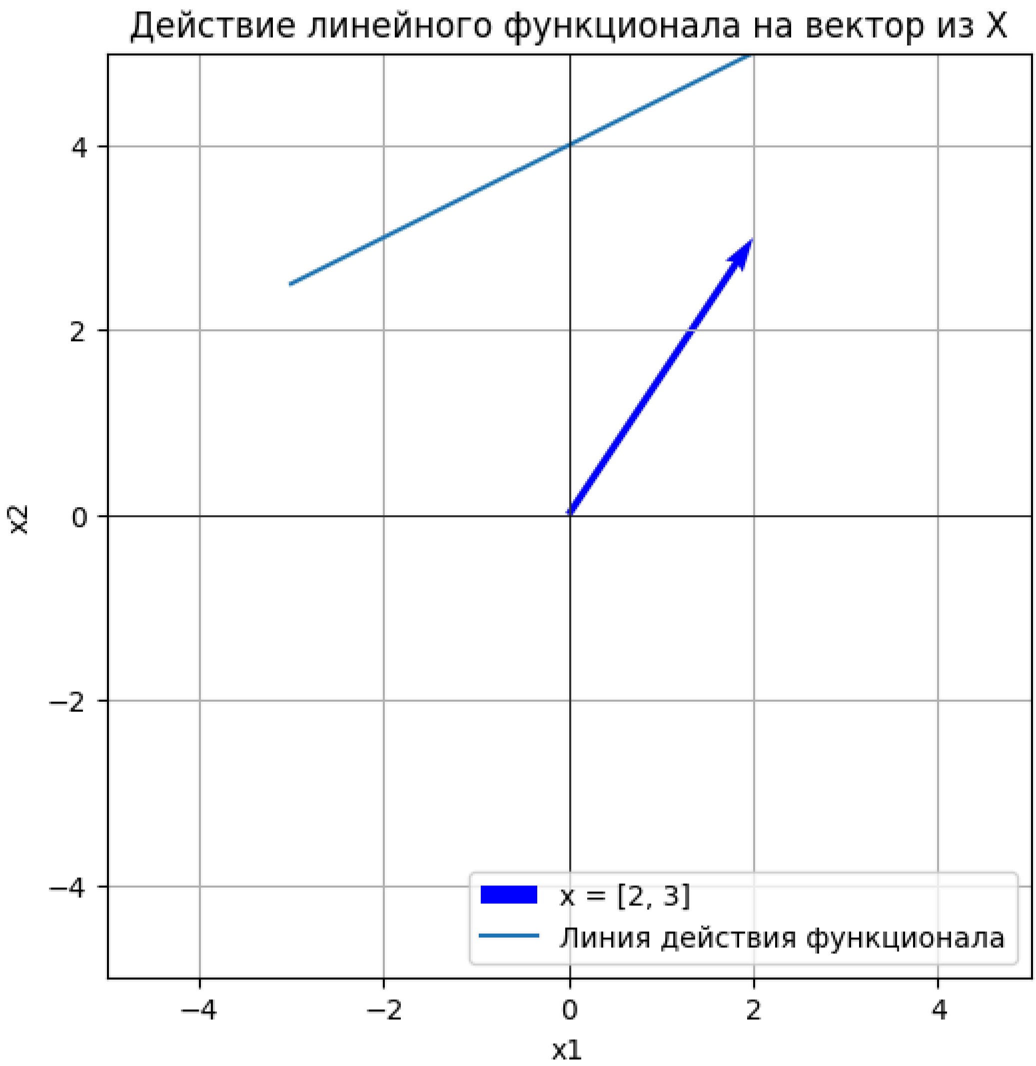
Это не обычное скалярное произведение, как в Евклидовых пространствах, а **действие линейного функционала на элемент**.

Почему это не обычное скалярное произведение?

Обычное скалярное произведение в евклидовых пространствах вычисляется по компонентам векторов, и оно всегда дает вещественное число. В случае двойственного скалярного произведения мы **не работаем с компонентами векторов**, а применяем **линейные функционалы** (которые могут быть функциями, а не просто

числами) к элементам из других пространств.
График:

- Синий вектор x=[2,3] — это элемент пространства $X = \mathbb{R}^2$.
- Линия — это действия линейного функционала $f(x) = x1 + 2x_2$, то есть отображение функционала на элементы пространства.
  Таким образом, двойственное скалярное произведение $\langle x, x^* \rangle = x^*(x)$ можно рассматривать как **применение линейного функционала** x^* к элементу x. В этом контексте x^* может быть любым линейным функционалом, который действует на x, и это дает вещественное число, которое отражает взаимодействие между элементами этих пространств.
  Крч x_1+2x_2 это линейный функционал, а [2,3] исходный вектор. (ну и $2*1+2*3=8$ то есть функционал принимает значение 8)



Оператор $T : X \rightrightarrows X^*$

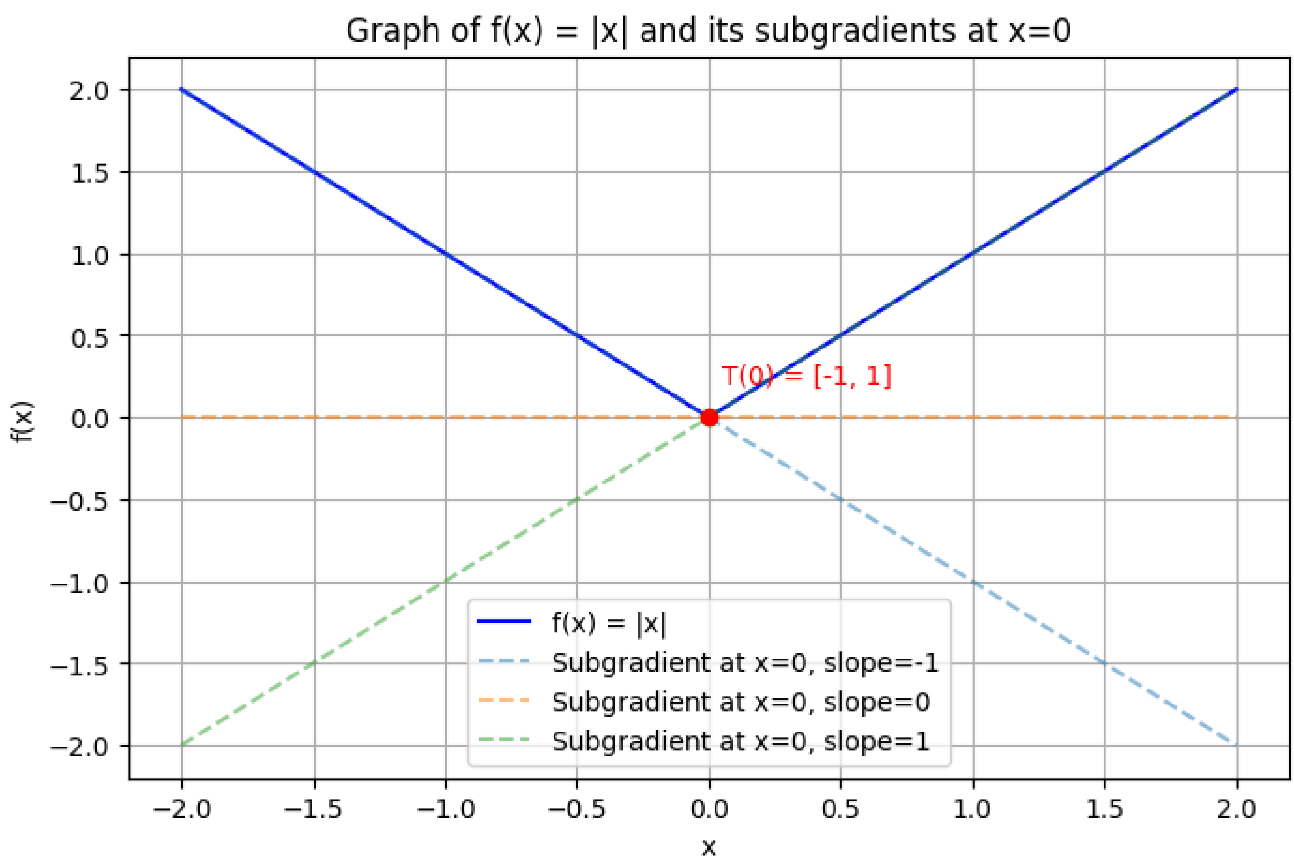
— это отношение на X в X^*:

$$T \subset X \times X^*,$$

— это **подмножество** декартова произведения $X \times X^*$. Такое множество можно воспринимать как **отношение** между $x \in X$ и $x^* \in X^*$.

Если $(x, x*) \in T$, то мы говорим: $x* \in T(x)$
Это обобщение обычного оператора, где одному x соответствует **не один**, а **множество** значений x^*.
(также и градиент-это один вектор, а подградиент это множество векторов)

график для выпуклой функции f(x)=|x|
касательные линии это подградиенты (все возможные для x=0)



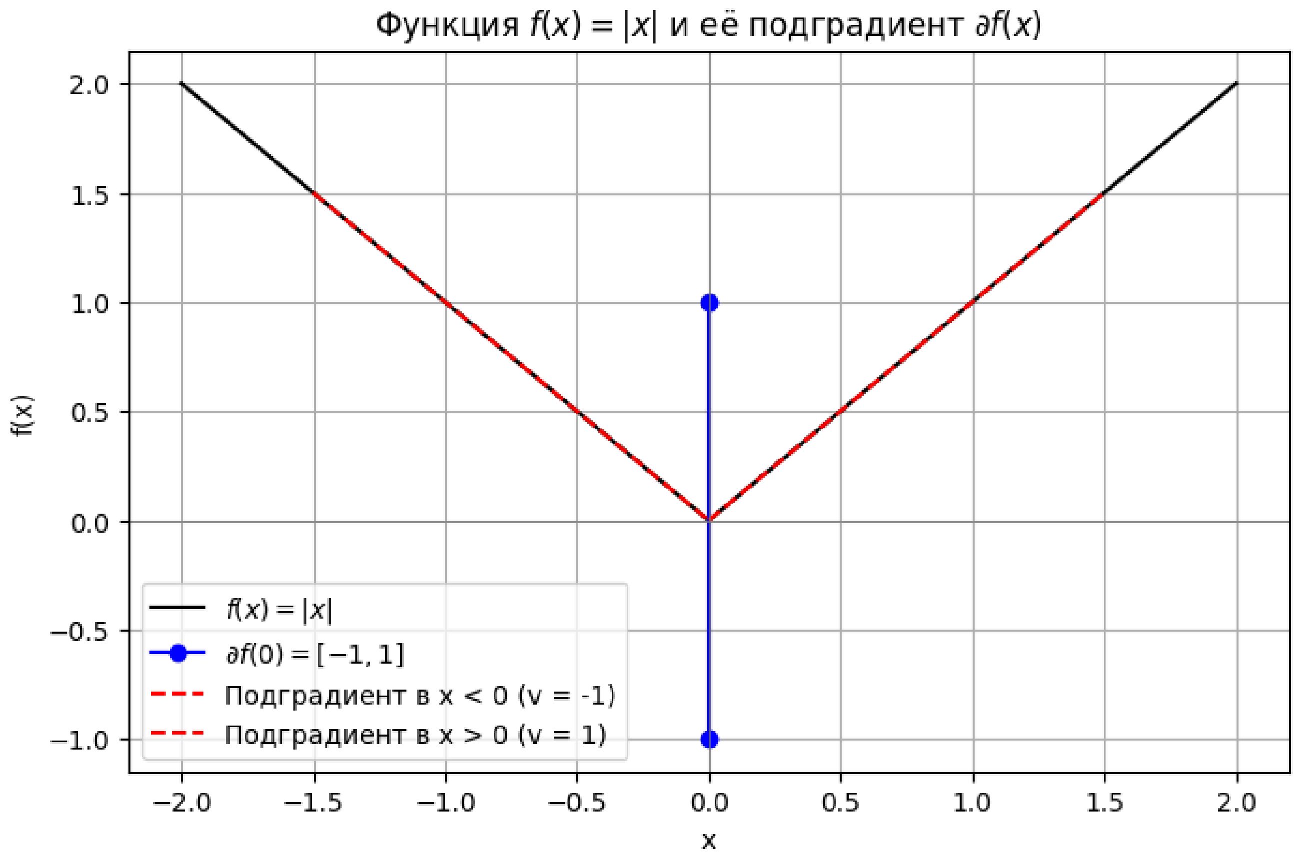
Область определения

Обозначается:

$$\mathrm{Dom}(T) = \{x \in X \mid T(x) \neq \emptyset\}.$$

$Dom(T)$ — это **область определения оператора** T, то есть множество всех таких элементов $x \in X$, для которых оператор T имеет хотя бы одно значение $x* \in X^*$, то есть $T(x) \neq \emptyset$.

В данном случае Dom это вся пряма |x|



Монотонный оператор

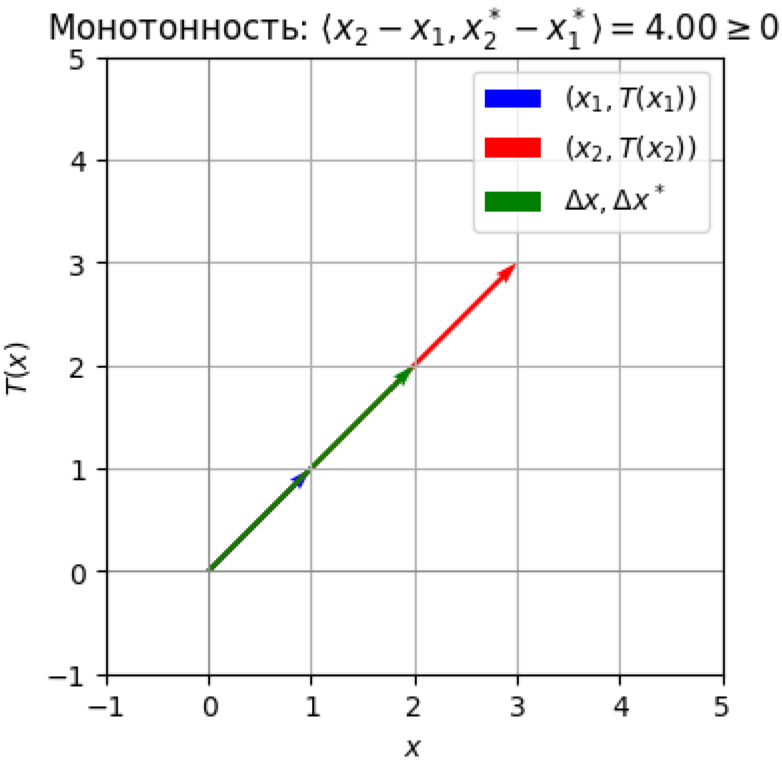
** $\rightrightarrows$ означает, что T может быть мультизадачным, то есть одному $x \in X$ может соответствовать множество $x^{\star} \in X^{\star}$

Пусть $T : X \rightrightarrows X^*$ — оператор, который отображает элементы из банахова пространства X в его сопряжённое пространство X^*.
Оператор $T : X \rightrightarrows X^*$ называется **монотонным**, если:

$$\langle x - y, x^* - y^* \rangle \geq 0, \quad \forall (x, x^*), (y, y^*) \in T.$$

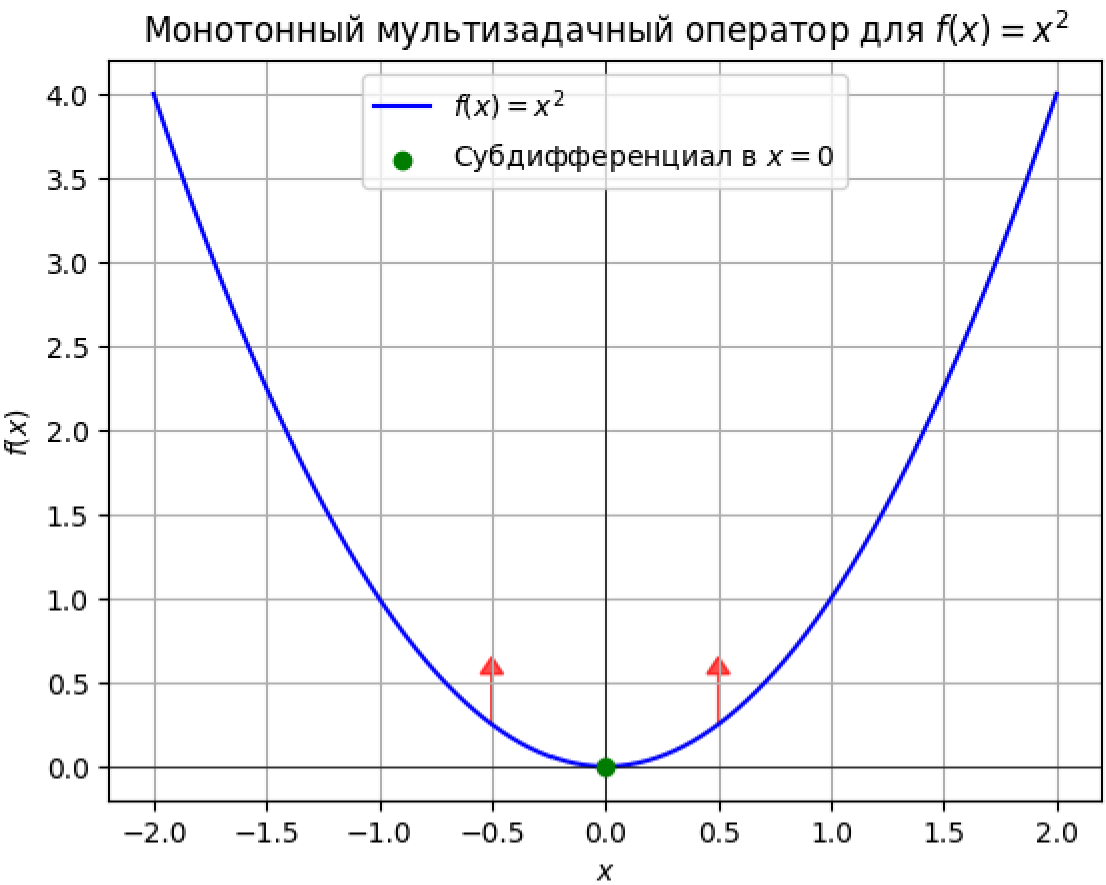
= Монотонность требует, чтобы скалярное произведение разности между любыми двумя разными точками было неотрицательно.

!! Монотонность - положительное движение в определенном направлении



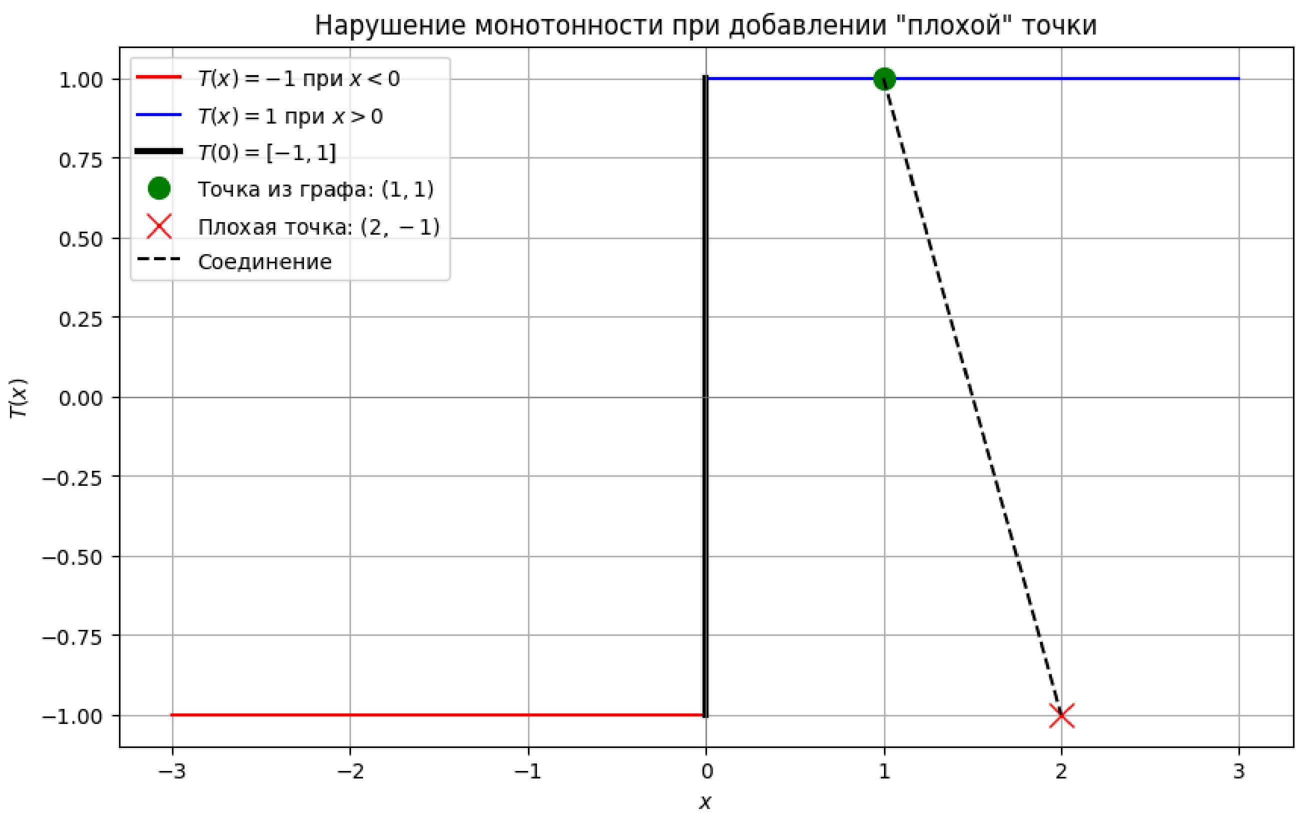
!!! А для мультизадачных операторов это означает, что любые два значения, взятые в любых двух точках согласованы по направлению

График:
для x=0 соответствует интервал [0,0] это мультизадачный оператор



Максимальный монотонный оператор

Оператор T называется **максимальным монотонным**, если он монотонный **и нельзя расширить его граф** (множество пар (x, x^*)), не потеряв монотонность. То есть он "максимально полный" среди всех монотонных операторов.
*Можно представить, что граф монотонного оператора — это "множество точек, расположенных так, что все углы между соединяющими их отрезками — острые или прямые".
Тогда максимальность означает, что **больше таких точек добавить нельзя, не сделав угол тупым**.
График: (2-1, -1-1) = -2 < 0 => условие монотонности нарушено



Контекст: H — вещественное гильбертово(см. начало) пространство

это просто определение регуляризации Моро-Йошида:

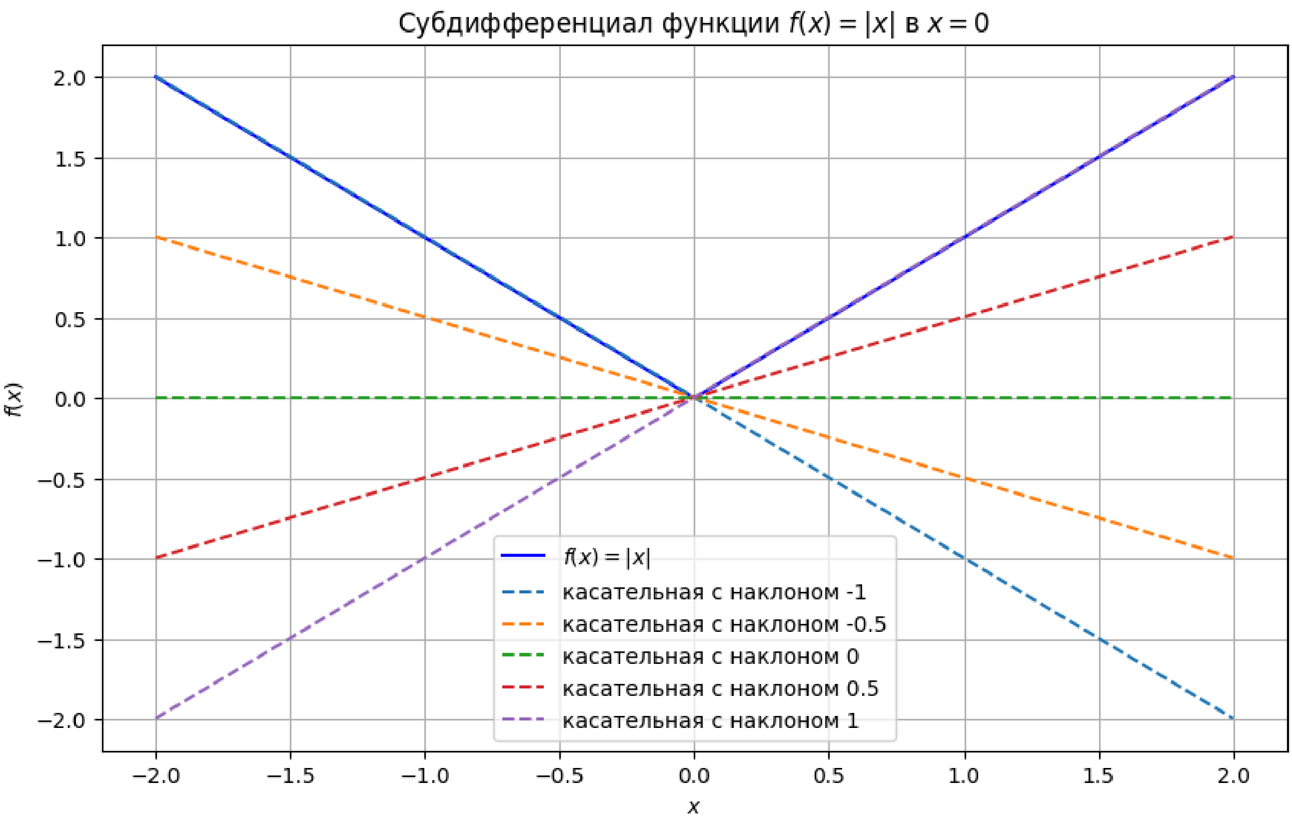
Пусть $T : H \rightrightarrows H$ — максимальный монотонный оператор. Регуляризация Моро-Йосиды — это способ **приблизить "жесткий" оператор T гладким оператором T_λ, который удобнее анализировать и применять в численных методах.
$T_\lambda = \frac{I - R_\lambda}{\lambda}, \quad R_\lambda = (\lambda T + I)^{-1}.$
Здесь:

- I — тождественный оператор (identity).

- $\lambda > 0$ — параметр регуляризации.
- R_λ — **резольвентный оператор** (он хорошо определён даже когда T не является функцией, а отношением).
- T_λ — **регуляризация Моро-Йосиды**, сглаженный и Липшицев оператор, приближающий T.

Субдифференциал функции

Субдифференциал функции $f : X \to \mathbb{R} \cup \{\infty\}$ — это точечно-множественный оператор
$\partial f : X \rightrightarrows X^*$:
$\partial f(x) = \{x^* \in X^* \mid f(y) \geq f(x) + \langle y - x, x^* \rangle, \forall y \in X\}$.
Это значит, что $x^{\star}$ — это такой гиперплоскостной косатель, который не пересекает график функции f снизу. в начале конспекта упоминается.
График: Эти касательные лежат под графиком функции |x| — это и есть геометрическая интерпретация субдифференциала.
*



тождественный оператор в уравнении заменяется на отображение двойственности:
$J = \partial \frac{1}{2} \| \cdot \|^2$,

которое является точечным. Регуляризация Брези-Крандалл-Пази* обладает некоторыми свойствами классической регуляризации Моро-Йосиды и в гильбертовых пространствах совпадает с классической регуляризацией Моро-Йосиды.

*Регуляризация Брези-Крандалл-Пази: Это более **общая версия** регуляризации Моро-Йосиды, предложенная для **рефлексивных** банаховых пространств.

Однако, чтобы применить её, **часто перенормируют** пространство, т.е. заменяют исходную норму на **строго выпуклую** и **гладкую** (дифференцируемую).

---

2 Основные результаты и обозначения

Мы используем обозначение $\overline{\mathbb{R}}$ для расширенных вещественных чисел:
$\overline{\mathbb{R}} = \{-\infty\} \cup \mathbb{R} \cup \{\infty\}$.
Это просто означает, что мы расширяем множество действительных чисел, добавляя к нему бесконечности.

Выпуклая функция

Функция $f : X \to \overline{\mathbb{R}}$ называется **выпуклой**, если $f > -\infty$ и существует точка $\hat{x} \in X$, для которой $f(\hat{x}) < \infty$.

Сопряжённая функция Фенхеля–Лежандра для $f : X \to \overline{\mathbb{R}}$ — это
$f^* : X^* \to \overline{\mathbb{R}}$, определённая как:
$f^*(x^*) = \sup_{x \in X} \langle x, x^* \rangle - f(x)$.$

Это функциональный аналог **преобразования Лапласа**, но в выпуклом анализе. Она показывает, как функция "реагирует" на линейные функционалы x^*.
Пример: $f(x) = \frac{1}{2}x^2$
$f^*(x^*) = \sup_{x \in X} \left(x * x^x - \frac{1}{2}x^2\right)$.
$\frac{d}{dx}\left(x * x^* - \frac{1}{2}x^2\right) = x^* - x$
Приравниваем к 0 для max
$x * -x = 0 \implies x = x^*$

Максимум будет достигаться при $x = x^*$, поэтому:
$f^*(x^*) = xx^* - \frac{1}{2}x^2 = x^*x^* - \frac{1}{2}(x^*)^2 = \frac{1}{2}(x^*)^2$

Бисопряженная функция

Если f^* — сопряжённая функция, то $f^{**} := (f^*)^*$ — **бисопряжённая**.

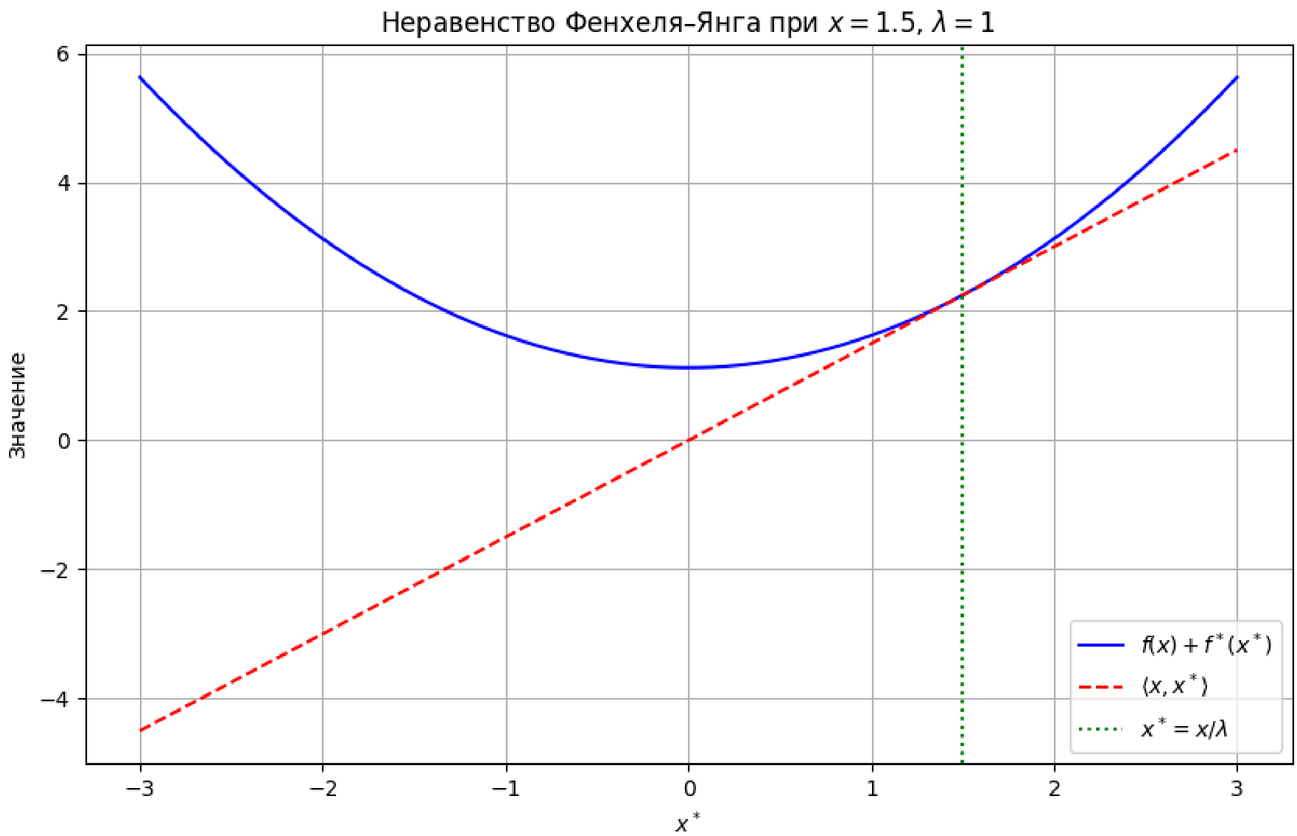
Теорема Фенхеля–Моро:

Если f выпуклая, собственная и полунепрерывная снизу, то

$f^{**}(x) = f(x), \quad \forall x \in X.$
То есть функция "восстанавливается" через двойное сопряжение. (см начало)

Неравенство Фенхеля–Янга

для всех $x \in X$, $x^* \in X^*$:
$\langle x^*, x \rangle \leq f(x) + f^*(x^*)$
$f(x) + f^*(x^*) = \langle x, x^* \rangle \quad \text{тогда и только тогда, когда} \quad x^* \in \partial f(x).$
— субдифференциал функции в точке x



В частном случае $f(x) = \|x\|^2/(2\lambda)$, где $\lambda > 0$,
мы получаем характеризуй отображения двойственности
$J : X \rightrightarrows X^*$.

$$\frac{1}{2\lambda}\|x\|^2 + \frac{\lambda}{2}\|x^*\|^2 \geq \langle x, x^* \rangle,$$

Это **уравнение действительно описывает точку касания (или пересечения) графиков** в **неравенстве Фенхеля–Янга**, когда оно **переходит в равенство**:
$\frac{1}{2\lambda}\|x\|^2 + \frac{\lambda}{2}\|x^*\|^2 = \langle x, x^* \rangle \quad \text{тогда и только тогда, когда} \quad x^* \in \lambda^{-1}J(x).$

Теорема 2.1. (см начало про прямую и двойственную задачу)

Пусть $f, g : X \to \overline{\mathbb{R}}$ — собственные выпуклые и полунепрерывные снизу функции.
Если существует $\hat{x} \in X$ такое, что f (или g) непрерывна в $\hat{x}$,
и $f(\hat{x}) < \infty$, $g(\hat{x}) < \infty$, то:
$\inf_{x \in X} f(x) + g(x) = \max_{x^* \in X^*} -f^*(-x^*) - g^*(x^*)$.
Здесь:

- $f, g : X \to R$ - выпуклые функции в линейном пространстве X
- f^*, g^* — **сопряженные функции** (по Фенхелю):
  $f^*(x^*) = \sup_{x \in X} (\langle x, x^* \rangle - f(x))$
- X^* — **сопряженное пространство** к X (например, если $X = \mathbb{R}^n$, то $X^* \cong R^n$)
- inf - инфимум (наименьшая нижняя грань)
  То есть **оптимизационная задача в прямом пространстве (левая часть)** может

быть эквивалентно выражена как **двойственная задача в сопряжённом пространстве (правая часть)**. Если выполнены технические условия (выпуклость, нижняя полунепрерывность, непрерывность хотя бы одной из функций в какой-то точке), то **двойственность сильная**, и инфимум совпадает с максимумом.

🧠 Пример: $f(x) = \frac{1}{2}x^2$, $g(x) = |x - 1|$

Это простой случай в одномерном пространстве $X = \mathbb{R}$.

---

1. Прямая задача:

$$\inf_{x \in \mathbb{R}} f(x) + g(x) = \inf_{x \in \mathbb{R}} \left(\frac{1}{2}x^2 + |x - 1| \right)$$

---

2. Сопряжённые функции:

Для $f(x) = \frac{1}{2}x^2$:

$$f^*(x^*) = \sup_{x \in \mathbb{R}} \left(xx^* - \frac{1}{2}x^2 \right) = \frac{1}{2}(x^*)^2$$

Для $g(x) = |x - 1|$, сопряжённая функция:

$$g^*(x^*) = x^* \cdot 1, \quad \text{если } |x^*| \leq 1, \text{ иначе } \infty$$

---

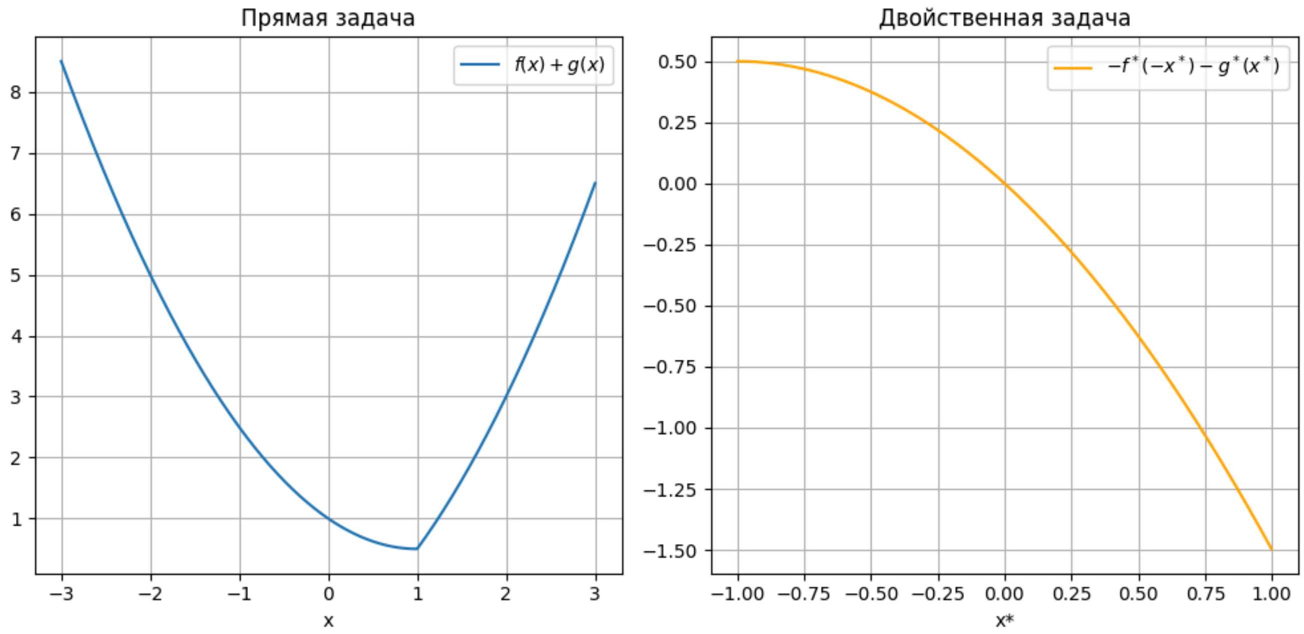
3. Двойственная задача:

$$\max_{x^* \in [-1,1]} -\frac{1}{2}(-x^*)^2 - x^* = \max_{x^* \in [-1,1]} -\frac{1}{2}(x^*)^2 - x^*$$

Мы можем перейти от **задачи минимизации суммы функций** к **двойственной задаче максимизации**, где используются сопряжённые функции.
Прямая задача:
$f(x) = \frac{1}{2}x^2$
$g(x) = |x - 1|$
f(x)+g(x)
Двойственная задача:

```
x_star = np.linspace(-1.2, 1.2, 500) #ось сопряжённого пространства
dual = -0.5 * (x_star**2) - x_star # значение двойственной задачи
dual[np.abs(x_star) > 1] = np.nan  # вне допустимой области сопряженной
функции (inf)
```

$-f^*(-x^*) - g^*(x^*) = -\frac{1}{2}(x^*)^2 - x^*$



Фицпатрик доказал, что каждому максимальному
$T : X \rightrightarrows X^*$ соответствует семейство $\mathcal{F}_T$
Семейство $\mathcal{F}_T$: выпуклые представления **монотонных операторов (см выше)**

Функция Фицпатрика (Позволяет перейти от оператора к выпуклой функции, с которой легче работать)

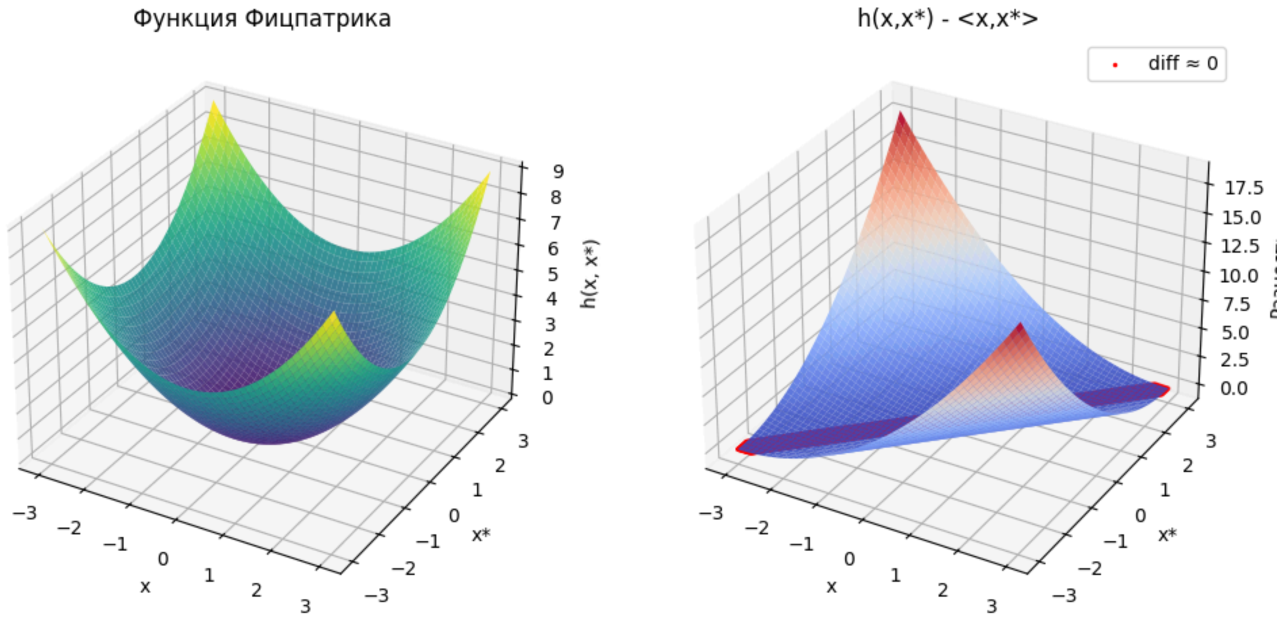
Описывает поведение оператора через выпуклую геометрию.

$$\mathcal{F}_T = \left\{ h : X \times X^* \to \overline{\mathbb{R}} \;\middle|\; \begin{array}{l} \text{для любого оператора T существует семейство функций h, которые:} \\ h \text{ — выпуклая и полунепрерывная снизу} \\ h(x, x^*) \geq \langle x, x^* \rangle, \quad \forall (x, x^*) \in X \times X^* \\ \text{Совпадают со сколярным произведением на графике оператора:} \\ (x, x^*) \in T \Rightarrow h(x, x^*) = \langle x, x^* \rangle \end{array} \right\}.$$

```
x = np.linspace(-3, 3, 300)
x_star = np.linspace(-3, 3, 300)
X, X_star = np.meshgrid(x, x_star)
# Скалярное произведение
scalar_prod = X * X_star
# Функция Фицпатрика
fitzpatrick = 0.5 * X**2 + 0.5 * X_star**2
```

```
# Разность между h и скалярным произведением
diff = fitzpatrick - scalar_prod
```

Красная линия второго графика - место где проходит график оператора



Если взять $X = \mathbb{R}$, то $X^* = \mathbb{R}$(в линейных пространствах с евклидовой метрикой $X \cong X^*$). Тогда h — функция **двух вещественных переменных**:

$h : \mathbb{R}^2 \to \mathbb{R}, h(x, x*)$

Чтобы изобразить такую функцию, мы **по стандарту математики** отображаем:

- Ось x — первая переменная,
- Ось x^* — вторая переменная,
- Ось z (вверх) — значение функции $h(x, x^*)$

Фицпатрик также дал явную формулу для наименьшего элемента $\mathcal{F}_T$ и доказал, что для любой $h \in \mathcal{F}_T$:

$$(x, x^*) \in T \iff h(x, x^*) = \langle x, x^* \rangle.$$

В связи с вышеуказанным уравнением, отныне мы будем называть любую $(h \in \mathcal{F}_T)$ **выпуклым представлением** T.

Он также дал **явную формулу** (ЭТОГО НЕТ В СТАТЬЕ) для наименьшего $h \in \mathcal{F}_T$, которая выглядит так:
$hT(x, x*) = \sup_{(y,y^*)\in T} \{\langle x, y^* \rangle + \langle y, x^* \rangle - \langle y, y^* \rangle\}$

Теорема 2.2. (Проверяет, задаёт ли функция h максимально монотонный оператор)

Каждая функция $h \in \mathcal{F}_T$ обладает свойствами:

- **выпуклость и нижняя полунепрерывность**;
- **мажорирование скалярного произведения**:
  $h(x, x^*) \geq \langle x, x^* \rangle$;
- **равенство скалярному произведению на графике оператора**:
  если $(x, x^*) \in T$, то $h(x, x^*) = \langle x, x^* \rangle$.

Пусть X — вещественное банахово пространство. Если $h : X \times X^* \to \overline{\mathbb{R}}$ — выпуклая функция и:
$h(x, x^*) \geq \langle x, x^* \rangle, \quad \forall (x, x^*) \in X \times X^*,$
$h^*(x^*, x^{**}) \geq \langle x^*, x^{**} \rangle, \quad \forall (x^*, x^{**}) \in X^* \times X^{**},$
где h^* — **выпуклое сопряжение** (сопряжённая функция в смысле Фенхеля).
то оператор $T : X \rightrightarrows X^*$, определённый как:
$T = \{(x, x^*) \in X \times X^* \mid h^*(x^*, x) = \langle x, x^* \rangle\}$, является **максимальным монотонным**, и функция $g(x, x^*) := h^*(x^*, x)$ — **выпуклое представление** T.
Кроме того, если h полунепрерывна снизу, то сама h - выпуклое представление T, а:

$T = \{(x, x^*) \in X \times X^* \mid h(x, x^*) = \langle x, x^* \rangle\}.$

```
   h(x, x*) ≥ ⟨x, x*⟩                  ⇨          «всё выше диагонали»
              |
              |
              v
    h*(x*, x) ≥ ⟨x, x*⟩                  ⇨          «двойственность сохраняет
выпуклость»
              |
              |
              v
     h*(x*, x) = ⟨x, x*⟩   ⇔   (x, x*) ∈ T      ⇨   определение T как max.
monotone
```

Следствие 2.3
Следствие 2. 3 уточняет, что **замыкание функции снизу** сохраняет её выпуклость и возможность быть выпуклым представлением оператора T. Это результат использования важного свойства сопряжённой функции, которая инвариантна относительно замыкания снизу.

С помощью этого следствия можно утверждать, что если мы работаем с функциями, которые имеют разрывы, то можно использовать их замыкание для построения выпуклых представлений монотонных операторов, не нарушая важные математические свойства, такие как выпуклость и непрерывность.

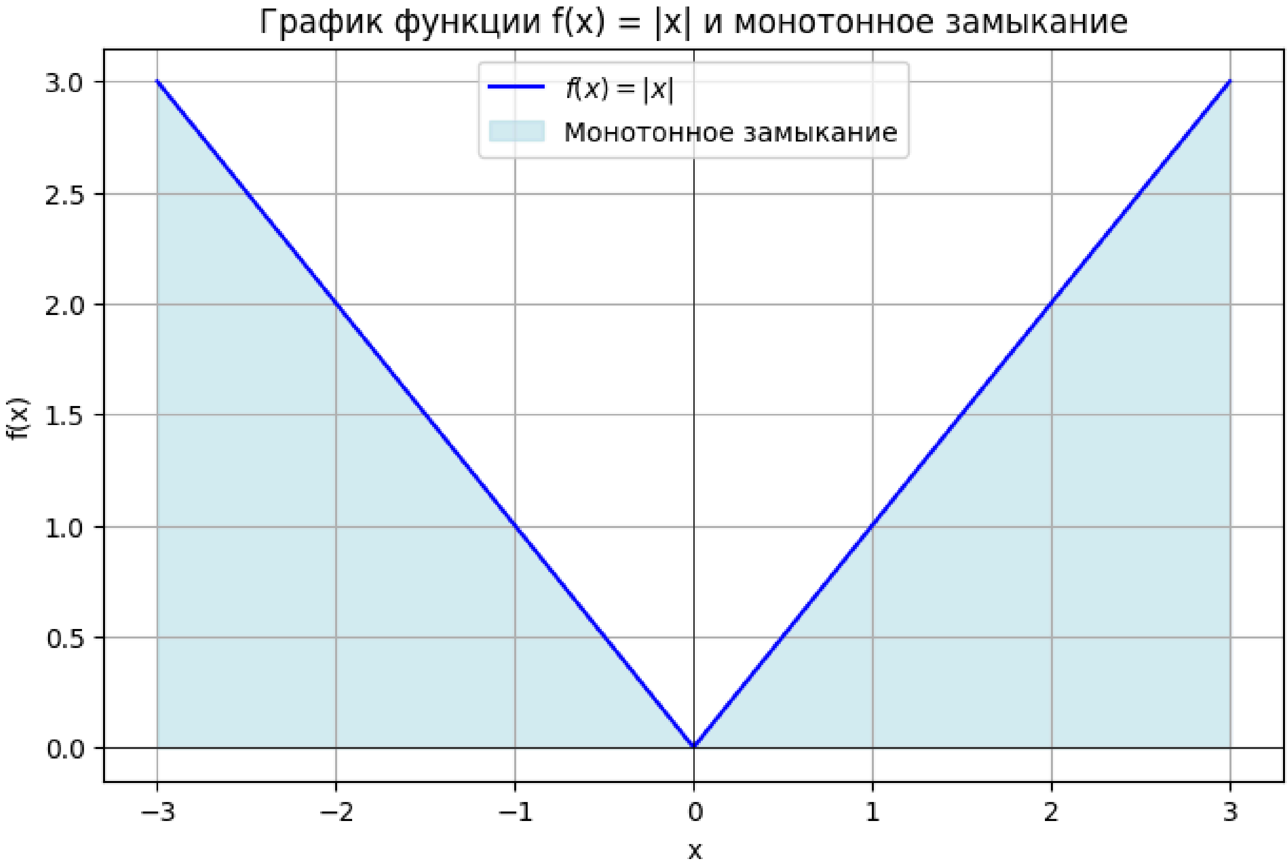
3 Операторы типа (D) Госсеза

Госсез определил класс операторов типа (D), чтобы расширить некоторые свойства максимальных монотонных операторов в рефлексивных банаховых пространствах на нерефлексивные пространства.

Оператор T называется **оператором типа (D)**, если для любого элемента, который лежит в его **монотонном замыкании**, выполняется следующее условие:
Элемент из монотонного замыкания $\widetilde{T}$ является пределом в слабой *−$*$ и сильной топологиях на пространстве $X^{\star} \times X^{\star\star}$.

Монотонное замыкание Госсеза [4] оператора $(T : X \rightrightarrows X^*)$ — это оператор $(\widetilde{T} : X^{**} \rightrightarrows X^*)$:
*Монотонное замыкание $\widetilde{T}$ оператора T — это оператор, который расширяет T, определяя более общее отношение между элементами пространств $X^{\star\star}$ и $X^{\star}$.
Пространство $X^{\star\star}$ — это **двойственное пространство** для $X^{\star}$, то есть пространство всех линейных функционалов на $X^{\star}$..

$\widetilde{T} = \{(x^{**}, x^*) \in X^{**} \times X^* \mid \langle x^{**} - y, x^* - y^* \rangle \geq 0, \quad \forall (y, y^*) \in T\}.$



Теорема 3.1

Теорема утверждает, что если T — максимальный монотонный оператор типа (D), то его расширение $\tilde{T}$ является уникальным максимальным монотонным расширением оператора T на пространство $X^{**} \times X^*$.

- $\tilde{T}$ — это расширение оператора T на более широкий контекст, то есть на пространство двойственного пространства X^{**} и пространство X^*. Теорема говорит о том, что такое расширение существует и оно единственно.

Теорема 3.2

Если f — собственная, выпуклая и полунепрерывная снизу функция, то её **субдифференциал** ∂f является оператором типа (D). При этом существует связь между субдифференциалом функции и её двойственным оператором:
$\widetilde{\partial f} = (\partial f^*)^{-1}$.

- Если функция f удовлетворяет условиям выпуклости и полунепрерывности снизу, то её субдифференциал является оператором типа (D), что означает, что его можно рассматривать как максимальный монотонный оператор, который можно расширить на более высокие пространства.
- Также указывается, что наилучшее расширение субдифференциала f в контексте максимальных монотонных операторов можно выразить через обратное преобразование двойственного оператора.

Теорема 3.3

Эта теорема утверждает эквивалентность нескольких условий для оператора T, который является максимальным монотонным оператором. Среди эквивалентных условий выделяются следующие:

1. Оператор T является оператором типа (D).
2. Для любой функции $h \in \mathcal{F}_T$ выполняется неравенство:
$h^*(x^*, x^{**}) \geq \langle x^{**}, x^* \rangle, \quad \forall (x^*, x^{**}) \in X^* \times X^{**}$,
3. Существует функция $h \in \mathcal{F}_T$, такая что:
$h^*(x^*, x^{**}) \geq \langle x^{**}, x^* \rangle, \quad \forall (x^*, x^{**}) \in X^* \times X^{**}$.

Кроме того, если выполнено любое из этих условий и $h \in \mathcal{F}_T$, то:

$h^* \in \mathcal{F}_{\widetilde{T}^{-1}}$.
*в теореме определяет расширение оператора T на более высокое пространство $X^{\star\star} \times X^{\star}$.

---

4 Регуляризация Моро-Йосиды в общих банаховых пространствах

Если X — строго выпуклое, гладкое и рефлексивное банахово пространство, то отображение двойственности:
$J = \partial \frac{1}{2} \| \cdot \|^2$

,где $\|\cdot\|$-норма пространства X
является биекцией между X и X^*.
Если пространство X обладает указанными свойствами, то отображение J (субдифференциал функции $\frac{1}{2}\|\cdot\|^2$) становится **биекцией** между пространством X и его двойственным пространством X^*∗. Это значит, что каждый элемент в X можно однозначно сопоставить элементу в X^* через операцию субдифференциала, и наоборот.

В этом контексте для $T : X \rightrightarrows X^*$, максимального монотонного оператора, обобщение Брези-Крандалл-Пази резольвента и регуляризации Моро-Йосиды оператора T с параметром $\lambda > 0$ определяются соответственно как:
Резольвента(resolver) оператора T, которая обозначается как R_λ. $R_\lambda : X \to X$,
Регуляризация оператора T, которая обозначается как T_λ. $\quad T_\lambda : X \to X^*$,
где для каждого $x \in X$, $\lambda > 0$, $(z, x^*) \in X \times X^*$ — единственное решение:
$\lambda x^* + J(z - x) = 0, \quad x^* \in T(z),$
и таким образом:
$R_\lambda(x) := z, \quad T_\lambda(x) := x^*$. (8)
Это значит, что для каждого $x \in X$, регуляризация оператора T приводит к элементам $z \in X$ и $x^* \in X^*$
Первая попытка обобщить версию регуляризации Брези-Крандалл-Пази (уравнение 8) заключалась бы в замене первого из этих уравнений на:
$\lambda x^* + J(z - x) \ni 0, \quad x^* \in T(z).$
В нерефлексивном банаховом пространстве J не является сюръективным.
Следовательно, если $T \equiv x^*$ и $x^* \notin J(X)$, то вышеуказанное включение не имело бы решения для любого x, и мы получили бы пустой T_λ. Чтобы обойти эту проблему, мы будем работать с монотонными замыканиями Госсеза T и J, то есть $\tilde{T}$ и $\tilde{J}$. Отметим, что поскольку J является субдифференциалом, в виду Теоремы 3.2, J — оператор типа (D), и:
$\widetilde{J} = (J_{X^*})^{-1}$,
где J_{X^*} обозначает отображение двойственности X^*.

Определение 4.1. Пусть X — вещественное банахово пространство, а $T : X \rightrightarrows X^*$ — максимальный монотонный оператор типа (D). Регуляризация Моро-Йосиды и резольвента T с параметром регуляризации $\lambda > 0$ определяются соответственно как $T_\lambda : X \rightrightarrows X^*$ и $R_\lambda : X \rightrightarrows X^{**}$:

$$T_\lambda = \left\{ (x, x^*) \in X \times X^* \,\middle|\, \begin{array}{l} \exists z^{**} \in X^{**} \text{ такое, что} \\ \lambda x^* + \widetilde{J}(z^{**} - x) \ni 0, \quad x^* \in \widetilde{T}(z^{**}) \end{array} \right\},$$

Здесь $\widetilde{J}$ — это оператор, который обычно связан с нормой в пространстве, а $\widetilde{T}(z^{**})$ — это некоторое расширение оператора T.
Резольвента:

$$R_\lambda = \left\{ (x, z^{**}) \in X \times X^{**} \,\middle|\, \begin{array}{l} \exists x^* \in X^* \text{ такое, что} \\ \lambda x^* + \widetilde{J}(z^{**} - x) \ni 0, \quad x^* \in \widetilde{T}(z^{**}) \end{array} \right\}.$$

В строго выпуклом, гладком и рефлексивном банаховом пространстве вышеуказанное определение тривиально приводит к версии регуляризации Брези-Крандалл-Пази. Основными инструментами для доказательства максимальной монотонности T_λ будут выпуклые представления Фицпатрика T, Теоремы 2.2 и 3.3.
Для $h : X \times X^* \to \overline{\mathbb{R}}$ и $\lambda > 0$ определим:

$$h_\lambda(x, x^*) = \inf_{z \in X} \left\{ h(z, x^*) + \frac{1}{2\lambda} \|z - x\|^2 \right\} + \frac{\lambda}{2} \|x^*\|^2.$$

Лемма 4.2. Пусть $h : X \times X^* \to \overline{\mathbb{R}}$ — выпуклая, собственная и полунепрерывная снизу функция. Тогда:

$$h_\lambda^*(x^*, x^{**}) = \min_{z^{**} \in X^{**}} \left\{ h^*(x^*, z^{**}) + \frac{1}{2\lambda} \|z^{**} - x^{**}\|^2 \right\} + \frac{\lambda}{2} \|x^*\|^2,$$

```
(13)
```

и h_λ^*, h_λ — выпуклые и собственные функции, причем h_λ^* полунепрерывна снизу. Кроме того, если X рефлексивно, то h_λ также полунепрерывна снизу, и инфимум в уравнении (12) является минимумом.
Доказательство. Прямой расчет дает:

$$h_\lambda^*(x^*, x^{**}) = \sup_{(y, y^*, z)} \langle y, x^* \rangle + \langle y^*, x^{**} \rangle - h(z, y^*) - \frac{1}{2\lambda} \|z - y\|^2 - \frac{\lambda}{2} \|y^*\|^2.$$

Определим $g(z, y^*) = \frac{\lambda}{2} \|y^*\|^2$. Тогда супремум в последнем члене вышеуказанного уравнения:

$$(h + g)^*(x^*, x^{**}).$$

Поскольку g выпукла и непрерывна, а h выпукла, собственна и полунепрерывна снизу, вышеуказанное сопряженное является (точной) инфи-конволюцией сопряженных h и g. Следовательно:

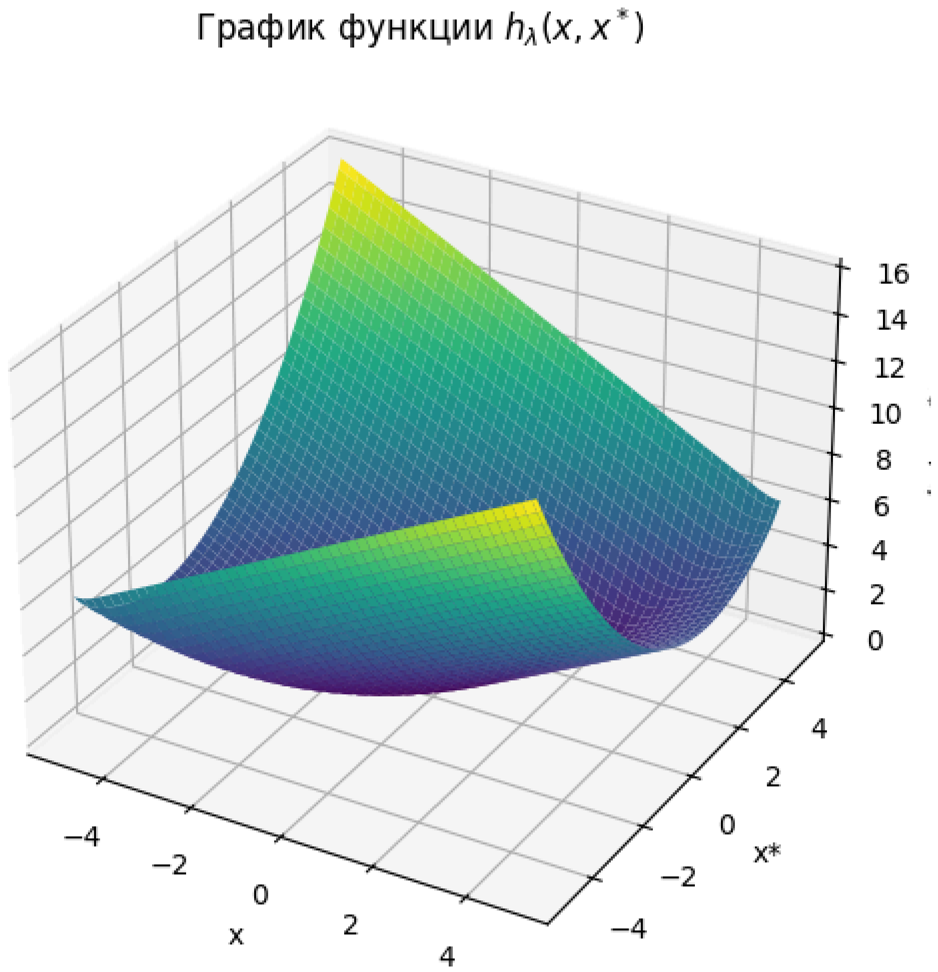
$$h_\lambda^*(x^*, x^{**}) = \frac{\lambda}{2} \|x^*\|^2 + \inf_{(u^*, u^{**})} h^*(x^* - u^*, x^{**} - u^{**}) + g^*(u^*, u^{**}),$$

где $g^*(u^*, u^{**}) = \frac{1}{2\lambda} \|u^{**}\|^2 + \delta_0(u^*)$. Используя подстановку $z^{**} = x^{**} - u^{**}$, получаем уравнение (13), что доказывает первую часть леммы.
Легко проверить, что h_λ выпукла. Поскольку h конечна в некоторой точке, h_λ собственна. Функция h_λ^* выпукла и полунепрерывна снизу, так как она сопряженная к h_λ. Используя то, что h_λ собственна, заключаем, что $h_\lambda^* > -\infty$. Кроме того, из уравнения (13) следует, что $h_\lambda^* < \infty$ в некоторой точке. Следовательно, h_λ^* собственна.
Теперь предположим, что X рефлексивно. Для доказательства того, что h_λ полунепрерывна снизу и что инфимум в её определении является минимумом, применим первую часть леммы к функции $\tilde{h}(x, x^*) = h^*(x^*, x)$, чтобы получить $(\tilde{h}_\lambda)^* = h_\lambda$.

Теорема 4.3. Пусть $T : X \rightrightarrows X^*$ — максимальный монотонный оператор типа (D). Для любого $\lambda > 0$ оператор T_λ является максимальным монотонным оператором типа (D). Кроме того, если h — выпуклое представление T, и h_λ определено как в уравнении (12):

$$h_\lambda(x, x^*) = \inf_{z \in X} \left\{ h(z, x^*) + \frac{1}{2\lambda} \|z - x\|^2 \right\} + \frac{\lambda}{2} \|x^*\|^2,$$

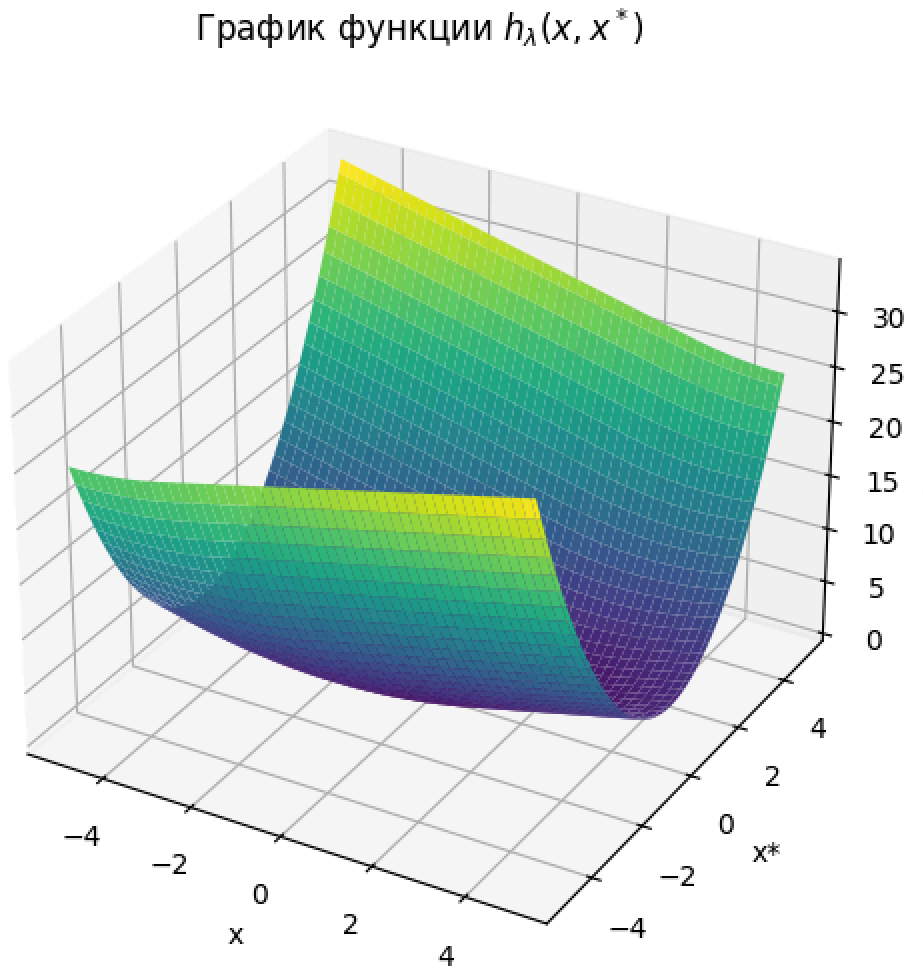
$h(z, x*) = |z - x^*|$



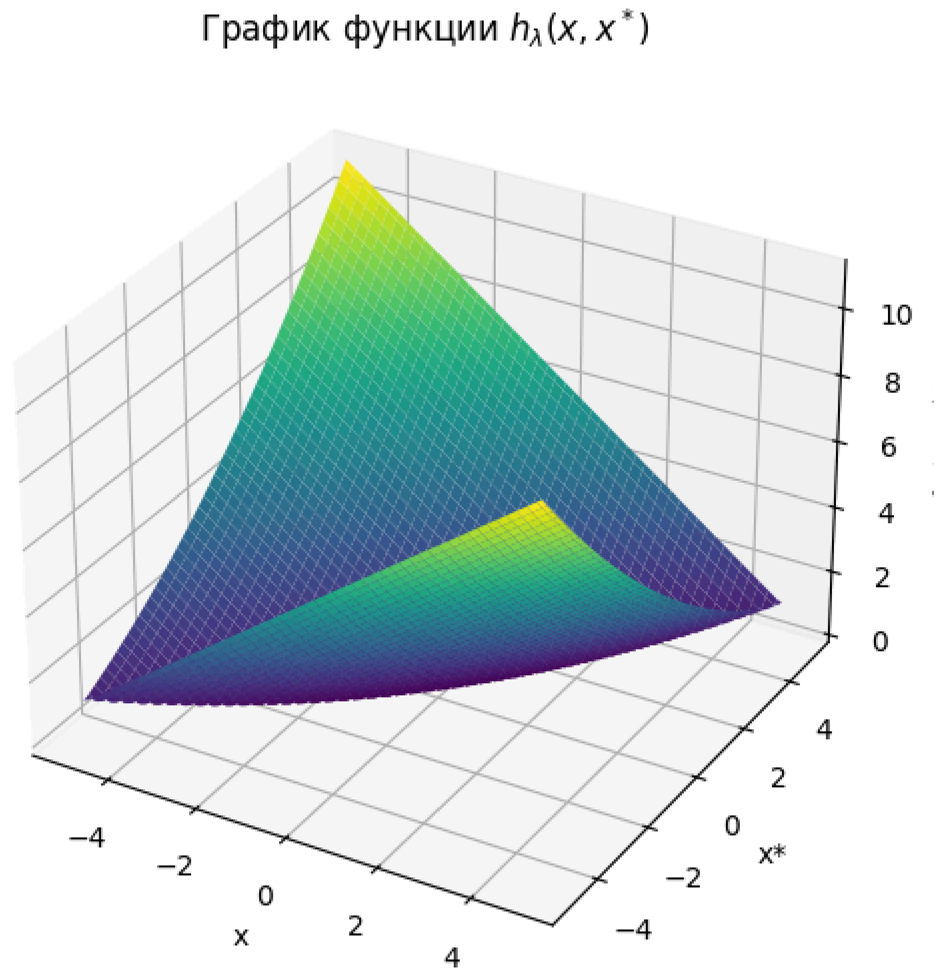
чем больше лямбда тем сильнее "сжатие"

```
def h(z, x_star):
    return np.abs(z - x_star)  # Например, это может быть абсолютная ошибка
# Функция для вычисления h_lambda(x, x^*)
def h_lambda(x, x_star, lambda_):
    # Определяем функцию для нахождения инфимума
    def objective(z):
        return h(z, x_star) + (1 / (2 * lambda_)) * (z - x)**2
    # Минимизируем по z (используем минимизацию)
    result = minimize(objective, x)  # Начальная точка - x
    z_opt = result.x  # Оптимальное значение z
    # Вычисляем значение h_lambda(x, x^*)
    return result.fun + (lambda_ / 2) * (x_star**2)
```

лямбда =2



лямбда=0.1



то:

$$T_\lambda = \{(x, x^*) \in X \times X^* \mid h_\lambda^*(x^*, x) = \langle x, x^* \rangle\},$$

и функции clh_λ и $g(x, x^*) := h_\lambda^*(x^*, x)$ являются выпуклыми представлениями T_λ. Если X рефлексивно, то h_λ также является выпуклым представлением T_λ.
Доказательство. Пусть h — выпуклое представление T. Поскольку h мажорирует двойственное скалярное произведение, для любых $x, z \in X$, $x^* \in X^*$:

$$h(z, x^*) + \frac{1}{2\lambda}\|z - x\|^2 + \frac{\lambda}{2}\|x^*\|^2 \geq \langle z, x^* \rangle + \langle x - z, x^* \rangle = \langle x, x^* \rangle.$$

Следовательно, h_λ мажорирует двойственное скалярное произведение в $X \times X^*$. Используя Теорему 3.3, пункт 2, имеем, что h^* мажорирует двойственное скалярное произведение в $X^* \times X^{**}$. Следовательно, для любых $x^* \in X^*$, $x^{**}, z^{**} \in X^{**}$:

$$h^*(x^*, z^{**}) + \frac{1}{2\lambda}\|z^{**} - x^{**}\|^2 + \frac{\lambda}{2}\|x^*\|^2 \geq \langle z^{**}, x^* \rangle + \langle z^{**} - x^{**}, x^* \rangle = \langle x^{**}, x^* \rangle.$$

Соединяя вышеуказанное уравнение с Леммой 4.2, уравнением (13), заключаем, что (h_λ^*) также мажорирует двойственное скалярное произведение.

Согласно Лемме 4.2, h_λ и h_λ^* выпуклы, а h_λ^* полунепрерывна снизу. Поскольку h_λ и h_λ^* мажорируют двойственное скалярное произведение в своих соответствующих областях, определим:

$$S = \{(x, x^*) \in X \times X^* \mid h_\lambda^*(x^*, x) = \langle x, x^* \rangle\},$$

и, используя Теорему 2.2 и Следствие 2.3, заключаем, что S максимально монотонен типа (D), и $g(x, x^*) = h_\lambda^*(x^*, x)$, $\mathrm{cl}h_\lambda$ — выпуклые представления S.

Чтобы доказать, что $T_\lambda = S$, используем уравнение (13), чтобы заключить, что $(x, x^*) \in S$ тогда и только тогда, когда существует $z^{**} \in X^{**}$ такое, что:

$h^*(x^*, z^{**}) + \frac{1}{2\lambda}\|z^{**} - x\|^2 + \frac{\lambda}{2}\|x^*\|^2 = \langle x, x^* \rangle.$

Используя Теорему 2.2, имеем:

$(z^{**}, x^*) \in \widetilde{T} \Longleftrightarrow h^*(x^*, z^{**}) = \langle z^{**}, x^* \rangle.$
Соединяя два вышеуказанных уравнения с уравнением (15), фактом, что h^* мажорирует двойственное скалярное произведение, уравнениями (9) и (4), заключаем, что уравнение (16) эквивалентно:
$x^* \in \widetilde{T}(z^{**}), \quad -\lambda x^* \in \widetilde{J}(z^{**} - x),$
Следовательно, $T_\lambda = S$ максимально монотонно, и g, $\mathrm{cl}h_\lambda$ — выпуклые представления S. Поскольку $(\mathrm{cl}h_\lambda)^* = h_\lambda^*$ и h_λ^* мажорирует двойственное скалярное произведение, используя снова Теорему 2.2, заключаем, что T_λ типа (D). Наконец, если X рефлексивно, используем Лемму 4.2, чтобы заключить, что h_λ полунепрерывна снизу, и следовательно, $h_\lambda = \mathrm{cl}h_\lambda$.

Мы доказали, что если T максимально монотонно типа (D), то T_λ также (максимально монотонно) типа (D). Следовательно, естественно искать выражения для $\widetilde{T_\lambda}$.

Следствие 4.4. Если $(T : X \rightrightarrows X^*)$ — максимальный монотонный оператор типа (D), то:

$\widetilde{T_\lambda} = \left(\widetilde{T}^{-1} + \lambda^{-1}\widetilde{J}^{-1}\right)^{-1}, \quad T_\lambda = \left(\widetilde{T}^{-1} + \lambda^{-1}\widetilde{J}^{-1}\right)^{-1} \cap X \times X^*.$
Если X рефлексивно, то:

$T_\lambda = \left(T^{-1} + \lambda^{-1}J^{-1}\right)^{-1},$

и если X — гильбертово пространство, то $T_\lambda = \left(T^{-1} + \lambda^{-1}I\right)^{-1}$.

Для $h : X \times X^* \to \overline{\mathbb{R}}$ и $(z, z^*) \in X \times X^*$ определим:

$h_{(z,z^*)}(x, x^*) = h(x + z, x^* + z^*) - [\langle x, z^* \rangle + \langle z, x^* \rangle + \langle z, z^* \rangle].$

Лемма 4.5. Пусть $T : X \rightrightarrows X^*$ — максимальный монотонный оператор типа (D). Тогда для любого $\lambda > 0$, $\mathrm{Dom}(T_\lambda) = X$.

Доказательство. Возьмем $x_0 \in X$. Выберем $h \in \mathcal{F}_T$ и пусть:

$$\beta := \inf_{x^* \in X^*} h_\lambda(x_0, x^*) - \langle x_0, x^* \rangle,$$

где h_λ определено как в уравнении (12). Определим также $f : X \times X^* \to \overline{\mathbb{R}}$:

$$f(y, y^*) = h(y + x_0, y^*) - \langle x_0, y^* \rangle.$$

Отметим, что f выпукла, полунепрерывна снизу и:

$$f^*(y^*, y^{**}) = h^*(y^*, y^{**} + x_0) - \langle x_0, y^* \rangle.$$

Поскольку h и h^* мажорируют двойственное скалярное произведение, f и f^* мажорируют двойственное скалярное произведение в своих соответствующих областях, и $\beta \geq 0$

Прямое использование уравнения (12) дает:

$$\beta = \inf_{z \in X, x^* \in X^*} f(z - x_0, x^*) + \frac{1}{2\lambda}\|z - x_0\|^2 + \frac{\lambda}{2}\|x^*\|^2 = \inf_{x \in X, x^* \in X^*} f(x, x^*) + \frac{1}{2\lambda}\|x\|^2 + \frac{\lambda}{2}\|x^*\|^2.$$

Используя Теорему 2.1, заключаем, что существуют $w^* \in X^*$, $w^{**} \in X^{**}$ такие, что:

$$\beta = -\left[f^*(w^*, w^{**}) + \frac{\lambda}{2}\|w^*\|^2 + \frac{1}{2\lambda}\|w^{**}\|^2\right] \leq 0,$$

где последнее неравенство следует из того, что f^* мажорирует двойственное скалярное произведение. Поскольку $\beta \geq 0$, заключаем, что вышеуказанное неравенство выполняется как равенство. Следовательно, используя также уравнения (9) и (4), имеем:

$$h^*(w^*, w^{**} + x_0) - \langle x_0, w^* \rangle = \langle w^*, w^{**} \rangle, \quad -\lambda w^* \in \widetilde{J}(w^{**}).$$

Определив $z^{**} = w^{**} + x_0$, заключаем, что $h^*(w^*, z^{**}) = \langle z^{**}, w^* \rangle$. Поскольку h^* — выпуклое представление $\widetilde{T}^{-1}$, заключаем, что $w^* \in \widetilde{T}(w^*)$, и:

$$0 \in \lambda w^* + \widetilde{J}(z^{**} - x_0),$$

что означает $w^* \in T_\lambda(x_0)$.

Теорема 4.6. Пусть $T : X \rightrightarrows X^*$ — максимальный монотонный оператор типа (D). Тогда для всех $\lambda > 0$ выполняются следующие утверждения:

1. T_λ — максимальный монотонный оператор типа (D),
2. $\mathrm{Dom}(T_\lambda) = X$,
3. T_λ отображает ограниченные множества в ограниченные множества.

Доказательство. Пункт 1 доказан в Теореме 4.3. Соединяя Теорему 4.3 и Лемму 4.5, получаем пункт 2. Для завершения доказательства остается доказать пункт 3. Зафиксируем $(y, y^*) \in T$. Возьмем $(x, x^*) \in T_\lambda$. По определению T_λ существует $z^{**} \in X^{**}$ такое, что:

$$x^* \in \widetilde{T}(z^{**}) \quad \text{и} \quad -\lambda x^* \in \widetilde{J}(z^{**} - x).$$

```
                                                              (18)
```

Пусть $(y, y^*) \in T$. Используя оба включения в уравнении (18) и тот факт, что $T \subset \widetilde{T}$, получаем:

$$0 = \frac{\lambda}{2}\|x^*\|^2 + \frac{1}{2\lambda}\|z^{**} - x\|^2 + \langle x^*, z^{**} - x\rangle$$

$$= \frac{\lambda}{2}\|x^*\|^2 + \frac{1}{2\lambda}\|z^{**} - x\|^2 + \langle x^* - y^*, z^{**} - y\rangle + \langle x^* - y^*, y - x\rangle + \langle y^*, z^{**} - x\rangle \$\$\$\$ \geq \frac{\lambda}{2}\|x^*\|^2 + $$

$$\geq \frac{\lambda}{2}\|x^*\|^2 - \frac{\lambda}{2}\|y^*\|^2 + \langle x^* - y^*, y - x\rangle$$

$$\geq \frac{\lambda}{2}\|x^*\|^2 - \|x^*\|\|y - x\| - \frac{\lambda}{2}\|y^*\|^2 - \|y^*\|\|y - x\|.$$

Следовательно:

$$0 \geq \|x^*\|^2 - \frac{2}{\lambda}\|y - x\|\|x^*\| - \left(\|y^*\|^2 + \frac{2}{\lambda}\|y^*\|\|y - x\|\right),$$

и, следовательно:

$$\|x^*\| \leq \frac{2}{\lambda}\|y - x\| + \|y^*\| \leq \frac{2}{\lambda}\|x\| + \left(\frac{2}{\lambda}\|y\| + \|y^*\|\right).$$