# Dérivée à une variable: rappel

## Limite d'une fonction

Soit $f : \mathbb{R} \to \mathbb{R}$ et $a \in \mathbb{R}$. On dit que $f$ tend vers $L$ lorsque $x$ tend vers $a$, et on note

$$\lim_{x \to a} f(x) = L$$

si $f(x)$ devient arbitrairement proche de $L$ dès que $x$ est suffisamment proche de $a$. Formellement: pour tout $\varepsilon > 0$, il existe $\delta > 0$ tel que

$$0 < |x - a| < \delta \implies |f(x) - L| < \varepsilon$$

La valeur $f(a)$ n'intervient pas dans la définition: $f$ ne doit même pas être définie en $a$. Par exemple, $\sin(x)/x$ n'est pas définie en $0$ mais $\lim_{x \to 0} \sin(x)/x = 1$.

## Continuité

Soit $f : D \to \mathbb{R}$ définie sur un domaine $D \subseteq \mathbb{R}$ et $a \in D$. On dit que $f$ est *continue en* $a$ si

$$\lim_{x \to a} f(x) = f(a)$$

Contrairement à la limite, la valeur $f(a)$ intervient ici: la fonction doit être définie en $a$ et la limite doit être égale à cette valeur. Intuitivement, le graphe de $f$ ne fait pas de « saut » en $a$.

On dit que $f$ est *continue* (sur $D$) si elle est continue en tout point $a \in D$. Par exemple, $|x|$ est continue sur $\mathbb{R}$, la fonction signe n'est pas continue en $0$, et $1/x$ est continue sur son domaine $\mathbb{R} \setminus \{0\}$.

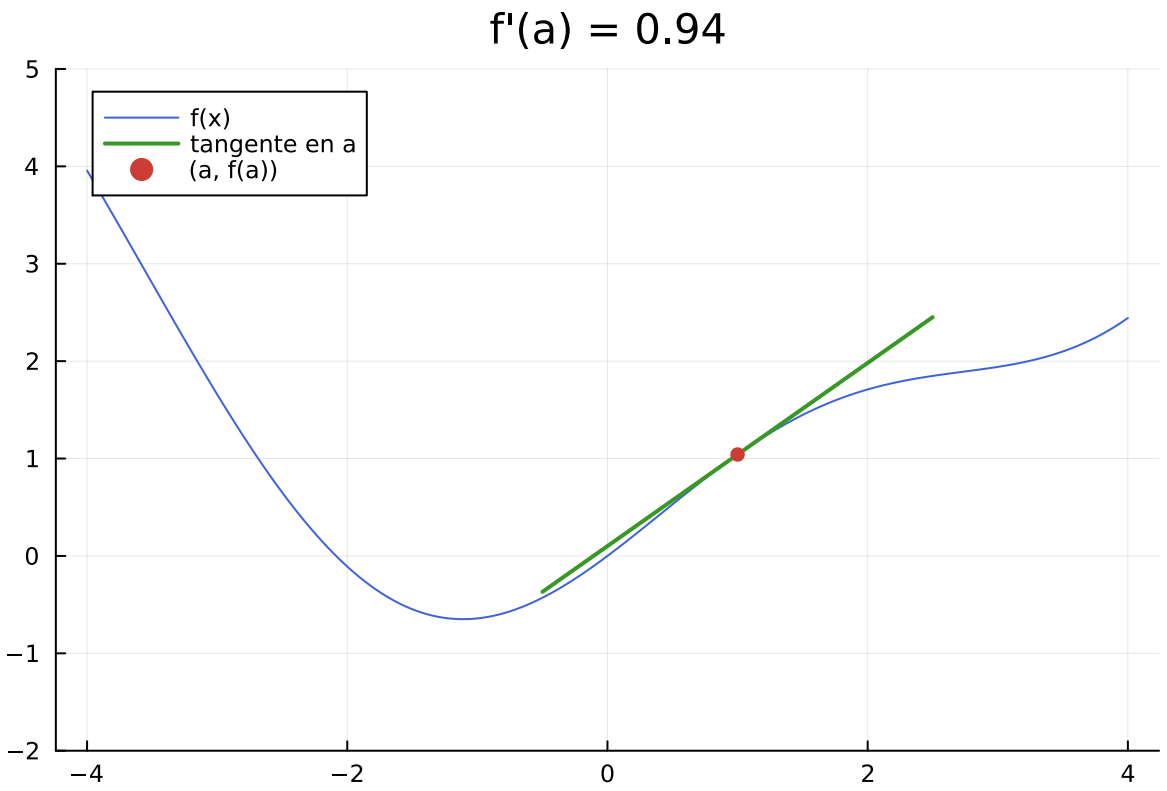
## Dérivée

La dérivée de $f$ en $a$ est la limite, si elle existe, du *taux d'accroissement*:

$$f'(a) = \lim_{h \to 0} \frac{f(a + h) - f(a)}{h}$$

Si cette limite existe, on dit que $f$ est *dérivable* en $a$. Le quotient est la pente de la droite sécante passant par les points $(a, f(a))$ et $(a+h, f(a+h))$. Lorsque $h \to 0$, la sécante tend vers la tangente: la dérivée univariée au point $a$ correspond à la pente de la **droite tangente** à la fonction en $a$.

a = 1.0




La droite tangente (en vert) touche le graphe de $f$ au point $(a, f(a))$ et a pour équation $y = f(a) + f'(a)(x - a)$. La dérivée $f'(a)$ est donc la **pente** de la droite tangente: elle est positive quand $f$ croît en $a$, négative quand $f$ décroît, et nulle quand la tangente est horizontale.

# Règles de dérivation: rappel

En pratique, on ne calcule presque jamais une dérivée avec la limite: on combine des dérivées connues à l'aide des règles suivantes ($f$ et $g$ dérivables, $c \in \mathbb{R}$ une constante).

$$
\begin{aligned}
\textbf{Somme:} &\quad (f+g)' = f' + g' \\
\textbf{Constante:} &\quad (cf)' = cf' \\
\textbf{Produit:} &\quad (fg)' = f'g + fg' \\
\textbf{Quotient:} &\quad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2} \quad (g \neq 0) \\
\textbf{Composition (dérivée en chaîne):} &\quad \big(f(g(x))\big)' = f'(g(x))\, g'(x) \\
\textbf{Puissance:} &\quad \big(f^n\big)' = n f^{n-1} f'
\end{aligned}
$$

Dérivées usuelles:

$$
\begin{aligned}
(c)' &= 0 & (x^n)' &= n x^{n-1} \\
(e^x)' &= e^x & (\ln x)' &= \frac{1}{x} \\
(\sin x)' &= \cos x & (\cos x)' &= -\sin x
\end{aligned}
$$

Par exemple, en combinant produit et composition:

$$
\frac{d}{dx} \sin(\pi x) e^{-x^2/4} = \pi \cos(\pi x) e^{-x^2/4} - \sin(\pi x) \frac{x}{2} e^{-x^2/4}
$$

et avec le quotient:

$$
\frac{d}{dx} \frac{x^2}{1+x^2} = \frac{2x(1+x^2) - x^2 \cdot 2x}{(1+x^2)^2} = \frac{2x}{(1+x^2)^2}
$$

# Dérivée à plusieurs variables

## De droite tangente à plan tangent

La dérivée univariée au point $a$ correspond à la pente de la **droite tangente** à la fonction en $a$. Pour une fonction bivariée, la dérivée dans une direction $d$ correspond à la pente du **plan tangent** à la fonction en $a$ dans la direction $d$.

```
md"""
# Dérivée à plusieurs variables

## De droite tangente à plan tangent

La dérivée univariée au point $a$ correspond à la pente de la **droite tangente** à la fonction en $a$.
Pour une fonction bivariée, la dérivée dans une direction $d$ correspond à la pente du **plan tangent** à la fonction en $a$ dans la direction $d$.
"""
```

```
PlotlyBackend()
```

```
plotly()
```

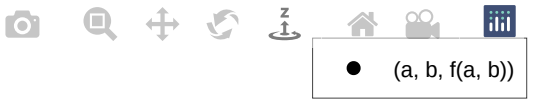
Commençons par un exemple simple: le paraboloïde $f(x, y) = x^2 + y^2$. Ses dérivées partielles sont $\partial f/\partial x = 2x$ et $\partial f/\partial y = 2y$, et le plan tangent (en rouge) au point $(a, b)$ a pour équation

$$z = f(a, b) + 2a(x - a) + 2b(y - b)$$

Le plan tangent est horizontal en $(0, 0)$, le minimum de $f$, et de plus en plus incliné lorsqu'on s'en éloigne.

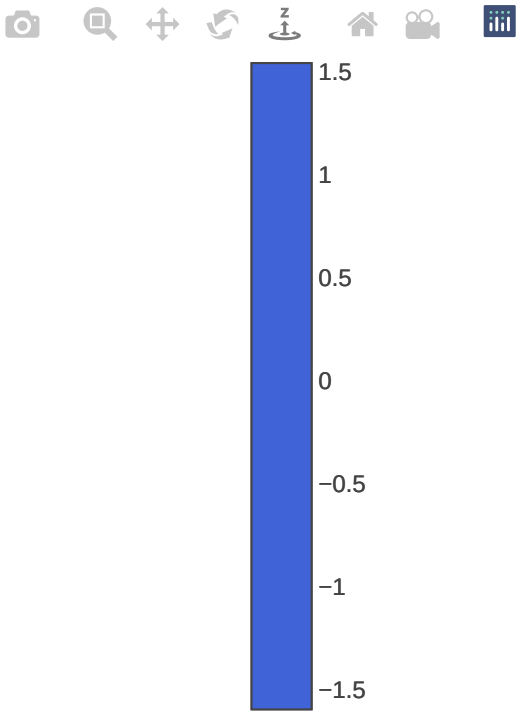
a = 1.0

b = 0.5




-0.01

-0.01




```
tangent_plane((x, y) -> sin(x * π) * exp(-y^2 / 4), x0, x1, y0, y1, cx, cy)
```

# Dérivée directionelle

Que vaut la dérivée dans la direction rouge `(1, 1)` si on connait les dérivée vertes en `x` : $\partial f/\partial x$ et en `y` : $\partial f/\partial y$ ? Un plan est défini par un point et deux vecteurs donc le plan bleu est entièrement défini par le point $(x, f(x))$ et les deux vecteurs $(1, 0, \partial f/\partial x)$ et $(0, 1, \partial f/\partial y)$. En utilisant la linéarité du plan, quelle est la valeur de $\alpha$ pour que le vecteur $(1, 1, \alpha)$ soit dans le plan ?

$$(1, 1, \partial f/\partial x + \partial f/\partial y)$$

En général, pour une direction arbitraire, on a

$$(d_x, d_y, d_x \cdot \partial f/\partial x + d_y \cdot \partial f/\partial y)$$

La dérivée en direction $(d_x, d_y)$ est donc obtenue par produit scalaire entre le vecteur $(d_x, d_y)$ et le *gradient* $(\partial f/\partial x, \partial f/\partial y)$. Voici trois notations équivalentes pour écrire ce produit scalaire:

$$\langle (d_x, d_y), (\partial f/\partial x, \partial f/\partial y) \rangle = \begin{bmatrix} d_x & d_y \end{bmatrix} \begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix} = d_x \cdot \partial f/\partial x + d_y \cdot \partial f/\partial y$$

Le gradient est souvent dénoté avec le symbol $\nabla$:

$$\nabla f = (\partial f/\partial x, \partial f/\partial y)$$

## Calculer le gradient à la main

Comment calculer les valeurs de $\partial f/\partial x$ et $\partial f/\partial y$ ? Commençant par voir comment faire ça à la main. Nous verrons au cours suivant comment le calculer algorithmiquement. L'astuce: pour calculer $\frac{\partial}{\partial x} f(x, y)$, il faut voir $y$ comme constant et donc voir $f$ comme fonction de $x$ uniquement. Par exemple:

$$\frac{\partial}{\partial x} \sin(x\pi) e^{-y^2/4} = \pi \cos(x\pi) e^{-y^2/4}$$

Idem pour $y$:

$$\frac{\partial}{\partial y} \sin(x\pi) e^{-y^2/4} = -\sin(x\pi) \frac{y}{2} e^{-y^2/4}$$

## Quand les dérivées sont-elles dans un plan ?

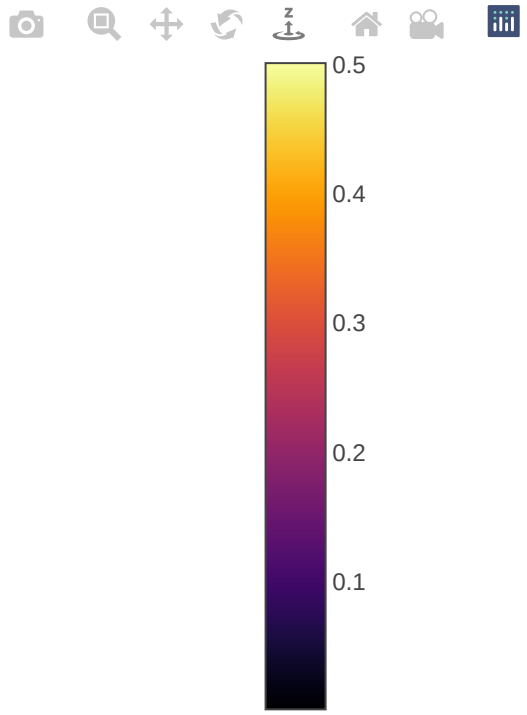
On dit dans ce cas que la fonction est *différentiable* en $(0, 0)$ et les dérivées directionelles sont donnée par produit scalaire avec le gradient. Quels sont les conditions pour la différentiabilité ?

Voyons un exemple:

$$\frac{x^{i_1} y^{i_2}}{x^{i_3} + y^{i_4}}$$

i_1 = 2

i_2 = 2



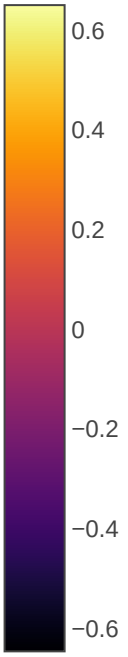

```
let
    x = y = range(-1, stop = 1, length = 32) # nombre pair de points pour ne pas évaluer en (0, 0)
    surface(x, y, (x, y) -> x^i_1 * y^i_2 / (x^i_3 + y^i_4))
end
```

i_3 = 2

i_4 = 2

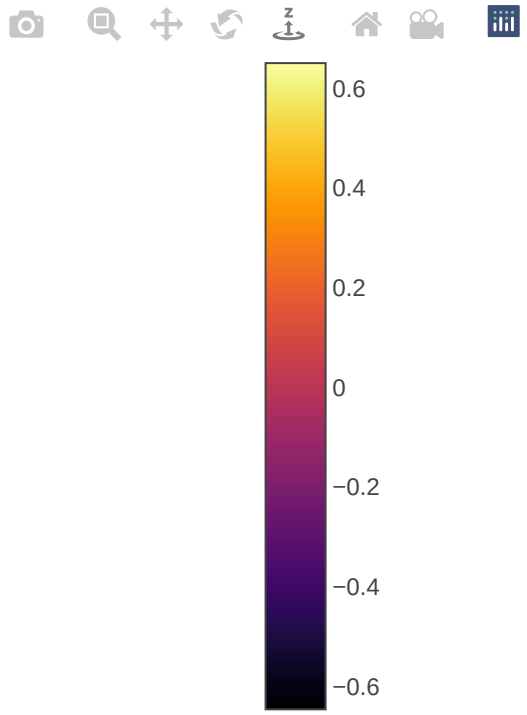
Pour quels valeurs de `i_1`, `i_2`, `i_3`, `i_4` est ce que la fonction est continue en `(0, 0)` ? Vous y répondrez en partie à la séance d'exercices! Ici, on se contentera de se fier au visuel.

Il est nécessaire que la fonction soit continue en $(0,0)$, donc par exemple $xy/(x^2+y^2)$ n'est pas continue. Y a-t-il des conditions sur $\partial f/\partial x$:




```
let
    x = y = range(-1, stop = 1, length = 32)
    function dfdx(x, y)
        d = -x^i_1 * y^i_2 / (x^i_3 + y^i_4)^2 * i_3 * x^(i_3 - 1)
        if i_1 > 0
            d += i_1 * x^(i_1 - 1) * y^i_2 / (x^i_3 + y^i_4)
        end
        return d
    end
    surface(x, y, dfdx)
end
```

Et $\partial f/\partial y$:




```
let
    x = y = range(-1, stop = 1, length = 32)
    function dfdy(x, y)
        d = -x^i_1 * y^i_2 / (x^i_3 + y^i_4)^2 * i_4 * y^(i_4 - 1)
        if i_2 > 0
            d += i_2 * x^i_1 * y^(i_2 - 1) / (x^i_3 + y^i_4)
        end
        return d
    end
    surface(x, y, dfdy)
end
```

**Théorème**: Si les dérivées partielles de $f$ sont définies dans un voisinage de $a$ et continues en $a$ alors $f$ est différentiable en $a$ (et donc en particulier continue en $a$).

# La Hessienne

Pour obtenir la dérivée second, on dérive à nouveau le gradient. On les notations:

$$\frac{\partial}{\partial x}(\frac{\partial}{\partial x}f) = \frac{\partial^2}{\partial x^2}f$$
$$\frac{\partial}{\partial x}(\frac{\partial}{\partial y}f) = \frac{\partial^2}{\partial x \partial y}f$$
$$\frac{\partial}{\partial y}(\frac{\partial}{\partial x}f) = \frac{\partial^2}{\partial y \partial x}f$$
$$\frac{\partial}{\partial y}(\frac{\partial}{\partial y}f) = \frac{\partial^2}{\partial y^2}f$$

**Attention**: $\partial x^2$ ne veut **pas** dire qu'on dérive par $x^2$, c'est juste une notation pour dire qu'on dérive 2 fois par $x$.

Cette notation peut paraître mal choisie ! En effet, dans $\partial xy$ et $\partial yx$, les parties $xy$ et $yx$ ressemblent à des produits. Comme le produit est commutatif, on a à tendance à penser que dériver par $x$ puis $y$ est égal à dériver par $y$ puis $x$... À moins que ça soit vraiment égal ?

**Théorème (Schwarz)** Si les dérivées secondes de $f$ existent et sont continues dans un voisinage de $a$ alors $\frac{\partial^2}{\partial x \partial y}f(a) = \frac{\partial^2}{\partial y \partial x}f(a)$.

Il ne suffit pas que $f$ soit différentiable: il existe des fonctions différentiables pour lesquelles les deux dérivées croisées sont différentes.

On aime mettre ces 4 valeurs dans une matrice nommée *Hessienne*. Sous l'hypothèse de ce théorème, la matrice est donc symmétrique!

$$\mathrm{Hess}(f) = \begin{bmatrix} \frac{\partial^2}{\partial x^2}f & \frac{\partial^2}{\partial x \partial y}f \\ \frac{\partial^2}{\partial x \partial y}f & \frac{\partial^2}{\partial y^2}f \end{bmatrix}$$

# Dérivée à plus de 2 variables

Tous les résultats se généralisent sans broncher à plus de 2 variables. Considérons une fonction $f(x_1, x_2, \ldots, x_n)$ à $n$ variables. Le gradient est $\nabla f = (\partial f/\partial x_1, \partial f/\partial x_2, \ldots, \partial f/\partial x_n)$ La dérivée dans la direction $d = (d_1, d_2, \ldots, d_n)$ est:
$\langle d, \nabla f \rangle = d^\top \nabla f = d_1 \cdot \partial f/\partial x_1 + \cdots + d_n \cdot \partial f/\partial x_n$

La Hessienne est symétrique (sous l'hypothèse du théorème de Schwarz) et devient:

$$\mathbf{Hess}(f) = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_1 \partial x_3} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_1 \partial x_2} & \frac{\partial^2 f}{\partial x_2^2} & \frac{\partial^2 f}{\partial x_2 \partial x_3} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_1 \partial x_{n-1}} & \frac{\partial^2 f}{\partial x_2 \partial x_{n-1}} & \cdots & \frac{\partial^2 f}{\partial x_{n-1}^2} & \frac{\partial^2 f}{\partial x_{n-1} \partial x_n} \\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \frac{\partial^2 f}{\partial x_2 x_n} & \cdots & \frac{\partial^2 f}{\partial x_{n-1} \partial x_n} & \frac{\partial^2 f}{\partial x_n^2} \end{bmatrix}$$

# Extremum d'une fonction

On étudie les points critiques, là où le gradient est nul (la fonction est plate).

### Points critiques et nature des extremums

Un point critique de $f(x, y)$ est un point où les dérivées partielles de $f$ sont toutes nulles : $f_x = 0$ et $f_y = 0$ (où $f_x = \partial f / \partial x$, $f_{xx} = \partial^2 f / \partial x^2$, etc.). Pour déterminer la nature du point critique, on utilise la matrice Hessienne $H(f)$ au point critique $(a, b)$ :

- Si $\det(H(f)(a, b)) > 0$ et $f_{xx}(a, b) > 0$, alors $(a, b)$ est un minimum local.
- Si $\det(H(f)(a, b)) > 0$ et $f_{xx}(a, b) < 0$, alors $(a, b)$ est un maximum local.
- Si $\det(H(f)(a, b)) < 0$, alors $(a, b)$ est un point-selle.
- Si $\det(H(f)(a, b)) = 0$, le test est indéterminé.

Les trois cas en $(0, 0)$, où $\nabla f(0, 0) = 0$ :

- $$f(x, y) = x^2 + y^2$$

  : $H = \mathbf{diag}(2, 2)$, $\det H = 4 > 0$ et $f_{xx} > 0$ → minimum.

- $$f(x, y) = -x^2 - y^2$$

  : $H = \mathbf{diag}(-2, -2)$, $\det H = 4 > 0$ et $f_{xx} < 0$ → maximum.

- $$f(x, y) = x^2 - y^2$$

  : $H = \mathbf{diag}(2, -2)$, $\det H = -4 < 0$ → point-selle.







```
let
    x = y = range(-1, stop = 1, length = 32)
    fonctions = [
        ((x, y) -> x^2 + y^2, "Minimum"),
        ((x, y) -> -x^2 - y^2, "Maximum"),
        ((x, y) -> x^2 - y^2, "Point-selle"),
    ]
    plots = map(fonctions) do (f, titre)
        surface(x, y, f, title = titre, colorbar = false, zlims = (-2, 2))
        scatter!([0], [0], [f(0, 0)], markersize = 4, color =
        Colors.JULIA_LOGO_COLORS.red, label = "")
    end
    plot(plots..., layout = (1, 3), size = (900, 350), legend = false)
end
```

# Utils

```
tangent_plane (generic function with 1 method)
```

```
function tangent_plane(f, x0, x1, y0, y1, cx, cy; Δ = 0.5, n = 32)
    x = range(x0, stop = x1, length = n)
    y = range(y0, stop = y1, length = n)
    df(x, y) = ForwardDiff.gradient(xy -> f(xy...), [x, y])
    surface(x, y, f, label = "f(x)", zlims = (minimum(f.(x, y') + maximum.(df.(x, y'))), maximum(f.(x, y') + maximum.(df.(x, y')))), legend=:bottomleft)
    grad = df(cx, cy)
    fc = f(cx, cy)
    plane(x, y) = fc + grad' * [x - cx, y - cy]
    plot!([cx, cx + 1], [cy, cy], [fc, fc + grad[1]], linewidth = 10, color = Colors.JULIA_LOGO_COLORS.green, label = "")
    plot!([cx, cx], [cy, cy + 1], [fc, fc + grad[2]], linewidth = 10, color = Colors.JULIA_LOGO_COLORS.green, label = "")
    plot!([cx, cx+1], [cy, cy + 1], [fc, fc + grad[1] + grad[2]], linewidth = 10, color = Colors.JULIA_LOGO_COLORS.red, label = "")
    return surface!( # works with plotly but not GR
        range(cx - Δ, stop = cx + Δ, length = 3),
        range(cy - Δ, stop = cy + Δ, length = 3),
        plane,
        color = [Colors.JULIA_LOGO_COLORS.blue],
        label = "",
    )
end
```

```
using LinearAlgebra, Plots, Colors, ForwardDiff, PlutoUI
```

```
import PlotlyBase, PlotlyKaleido
```

```
tangent_line (generic function with 1 method)
```

```
function tangent_line(f, x0, x1, c; n = 100)
    x = range(x0, stop = x1, length = n)
    df(x) = ForwardDiff.derivative(f, x)
    plot(x, f, color = Colors.JULIA_LOGO_COLORS.blue, label = "f(x)", ylims =
    (minimum(f.(x) + df.(x)), maximum(f.(x) + df.(x))), legend=:bottomleft)
    plot!(x, df, color = Colors.JULIA_LOGO_COLORS.purple, label = "f'(x)")
    pente = df(c)
    tangente(x) = pente * (x - c) + f(c)
    x_right = range(c, stop = c + 1, length = 2)
    plot!([c, c + 1], tangente, color = Colors.JULIA_LOGO_COLORS.green, label = "")
    plot!([c, c + 1], [f(c), f(c)], color = Colors.JULIA_LOGO_COLORS.green, label =
    "")
    plot!([c + 1, c + 1], [f(c), tangente(c + 1)], color =
    Colors.JULIA_LOGO_COLORS.red, label = "")
    scatter!([c], [pente], markerstrokewidth = 0, color =
    Colors.JULIA_LOGO_COLORS.red, label = "")
end
```