$$f(x, y) = (x + y)(y + 1) \tag{1}$$

Let's define intermediate variables a and b such that

$$a = x + y \tag{2}$$

$$b = y + 1 \tag{3}$$

$$f = a \times b \tag{4}$$

A computation graph for the "forward pass" through f is shown in Fig. 1.

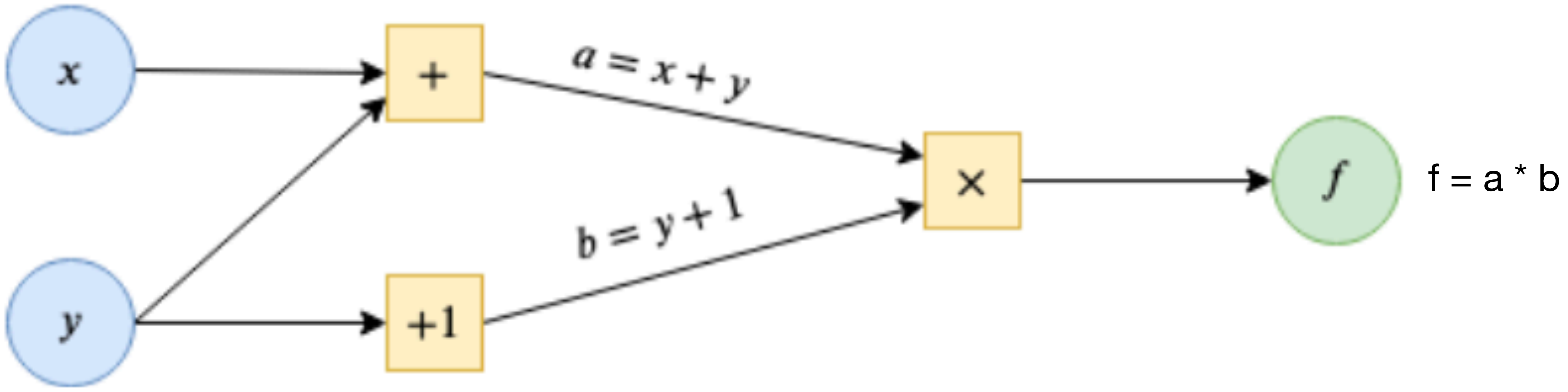


Figure 1

$$f(x,y) = (x+y)(y+1) \tag{1}$$

Let's define intermediate variables a and b such that

$$a = x + y \tag{2}$$

$$b = y + 1 \tag{3}$$

$$f = a \times b \tag{4}$$

A computation graph for the "forward pass" through f is shown in Fig. 1.

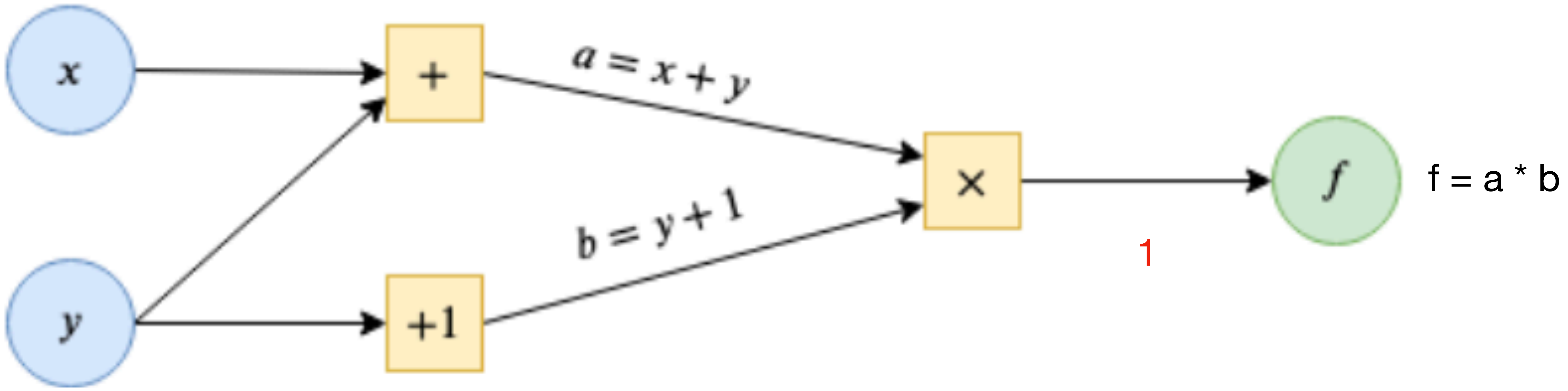


Figure 1

$$f(x, y) = (x + y)(y + 1) \tag{1}$$

Let's define intermediate variables a and b such that

$$a = x + y \tag{2}$$

$$b = y + 1 \tag{3}$$

$$f = a \times b \tag{4}$$

A computation graph for the "forward pass" through f is shown in Fig. 1.

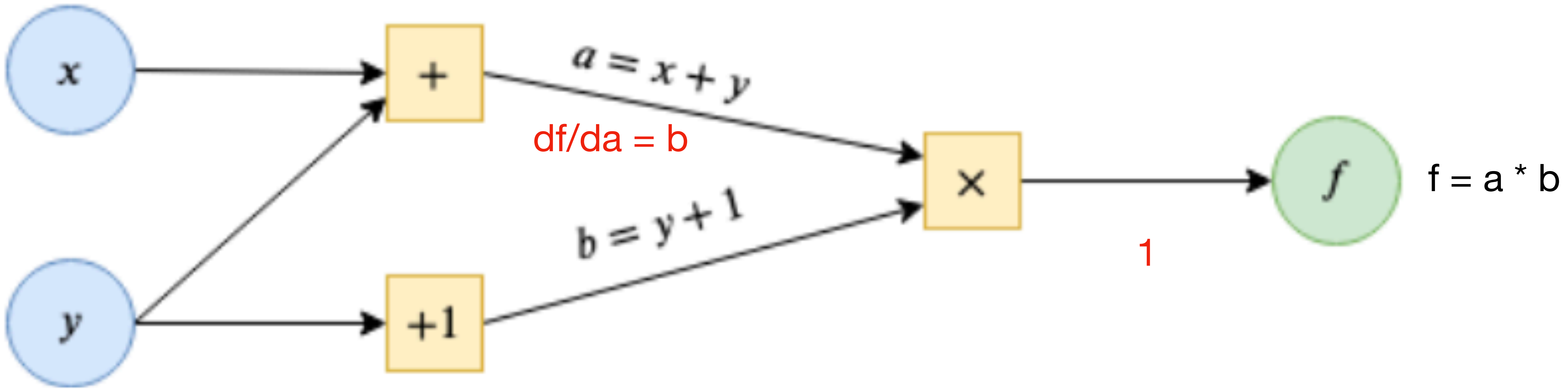


Figure 1

$$f(x, y) = (x + y)(y + 1) \tag{1}$$

Let's define intermediate variables a and b such that

$$a = x + y \tag{2}$$

$$b = y + 1 \tag{3}$$

$$f = a \times b \tag{4}$$

A computation graph for the "forward pass" through f is shown in Fig. 1.

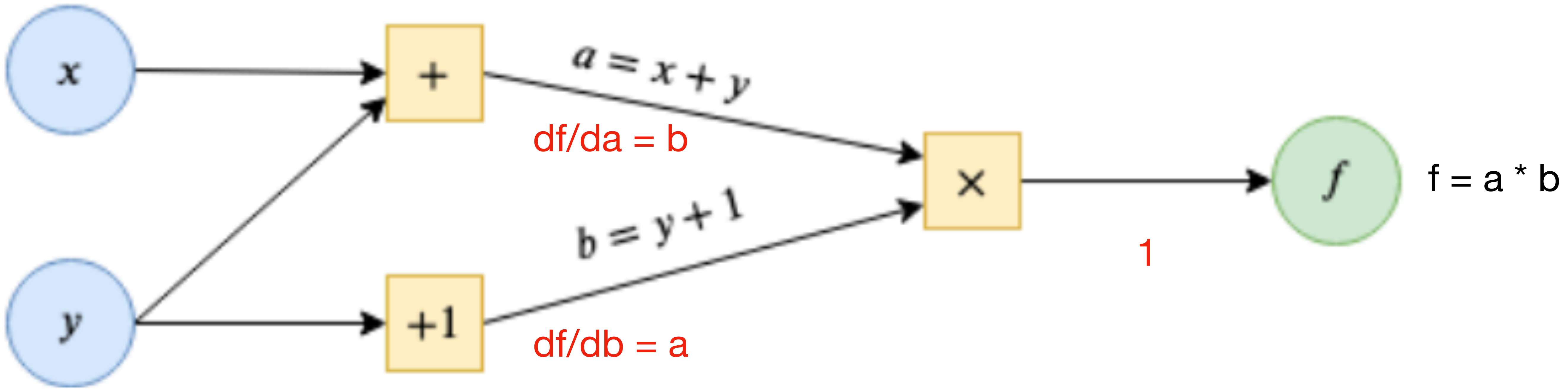


Figure 1

$$f(x, y) = (x + y)(y + 1) \quad (1)$$

Let's define intermediate variables a and b such that

$$a = x + y \quad (2)$$

$$b = y + 1 \quad (3)$$

$$f = a \times b \quad (4)$$

A computation graph for the "forward pass" through f is shown in Fig. 1.

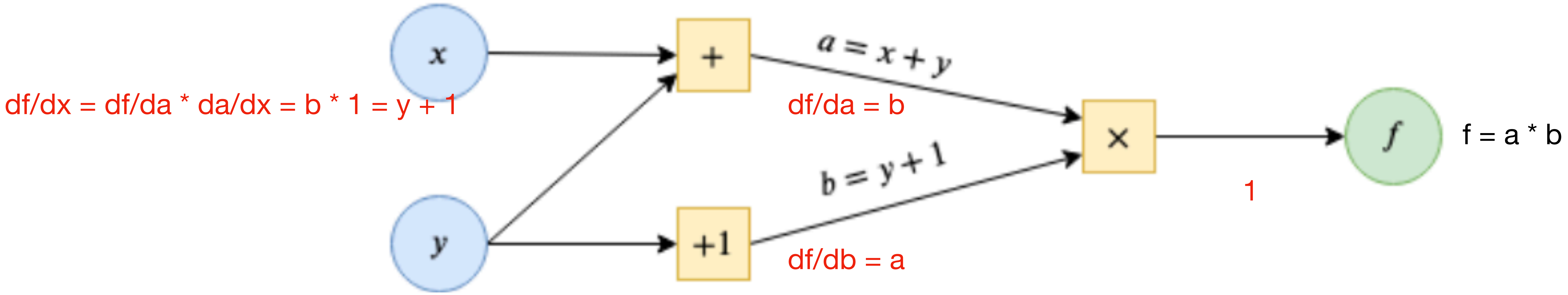


Figure 1

$$f(x, y) = (x + y)(y + 1) \tag{1}$$

Let's define intermediate variables a and b such that

$$a = x + y \tag{2}$$

$$b = y + 1 \tag{3}$$

$$f = a \times b \tag{4}$$

A computation graph for the "forward pass" through f is shown in Fig. 1.

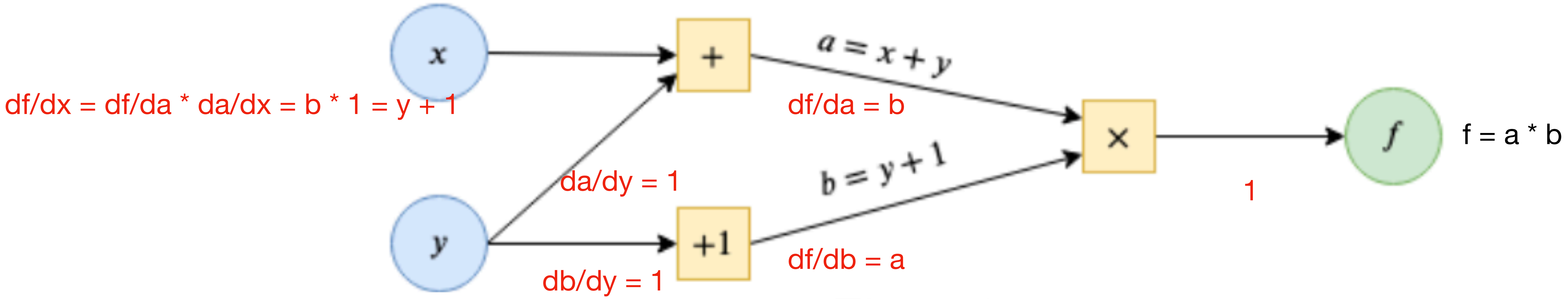


Figure 1

$$f(x, y) = (x + y)(y + 1) \tag{1}$$

Let's define intermediate variables a and b such that

$$a = x + y \tag{2}$$

$$b = y + 1 \tag{3}$$

$$f = a \times b \tag{4}$$

A computation graph for the "forward pass" through f is shown in Fig. 1.

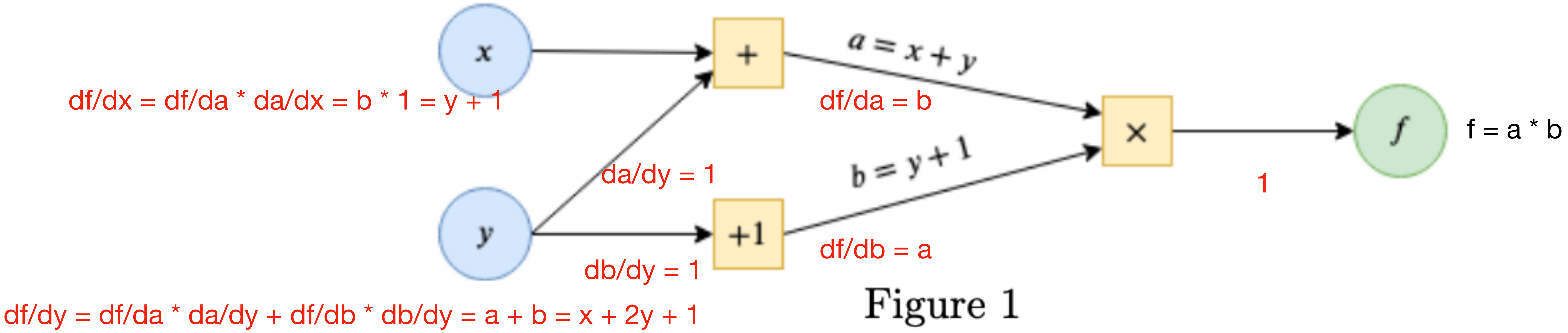


Figure 1

$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(a) Draw the computation graph. Compute the value of f at $\vec{w} = (1, -1)$.

$$f_1(w_1, w_2) = e^{e^{w_1}+e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$
$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(a) Draw the computation graph. Compute the value of f at $\vec{w} = (1, -1)$.

$$f = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

$$f(w_1, w_2) = \begin{bmatrix} f_1(w_1, w_2) \\ f_2(w_1, w_2) \end{bmatrix}$$

$$f(1,-1) = \begin{bmatrix} 18.296 \\ 0.0 \end{bmatrix}$$

$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

At this $\vec{w}$, compute the Jacobian $\frac{\partial \vec{f}}{\partial \vec{w}}$ using numerical differentiation (using $\Delta w = 0.01$).

$$f_1(w_1, w_2) = e^{e^{w_1 + e^{2w_2}}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

At this $\vec{w}$, compute the Jacobian $\frac{\partial \vec{f}}{\partial \vec{w}}$ using numerical differentiation (using $\Delta w = 0.01$).

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$$

$$f_1(w_1, w_2) = e^{e^{w_1 + e^{2w_2}}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

At this $\vec{w}$, compute the Jacobian $\frac{\partial \vec{f}}{\partial \vec{w}}$ using numerical differentiation (using $\Delta w = 0.01$).

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$$

$$\frac{\partial \overline{f}}{\partial \overline{w}} = \begin{bmatrix} \frac{\partial f_1}{\partial w_1} & \frac{\partial f_1}{\partial w_2} \\ \frac{\partial f_2}{\partial w_1} & \frac{\partial f_2}{\partial w_2} \end{bmatrix}$$

$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

At this $\vec{w}$, compute the Jacobian $\frac{\partial \vec{f}}{\partial \vec{w}}$ using numerical differentiation (using $\Delta w = 0.01$).

$$f'(x) \approx \frac{f(x+h) - f(x)}{h}$$

$$f'(x) \approx \frac{f(x+h) - f(x-h)}{2h}$$

$$\frac{\partial \overline{f}}{\partial \overline{w}} = \begin{bmatrix} \frac{\partial f_1}{\partial w_1} & \frac{\partial f_1}{\partial w_2} \\ \frac{\partial f_2}{\partial w_1} & \frac{\partial f_2}{\partial w_2} \end{bmatrix}$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{w}} \approx \left[\frac{\mathbf{f}(1+\Delta w, -1) - \mathbf{f}(1, -1)}{\Delta w}, \quad \frac{\mathbf{f}(1, -1+\Delta w) - \mathbf{f}(1, -1)}{\Delta w} \right]$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{w}} \approx \begin{bmatrix} 48.192 & 4.764 \\ 0.00 & 1.00 \end{bmatrix}$$

Or, using central differences:

$$\frac{\partial \mathbf{f}}{\partial \mathbf{w}} \approx \left[\frac{\mathbf{f}(1+\Delta w, -1) - \mathbf{f}(1-\Delta w, -1)}{2\Delta w}, \quad \frac{\mathbf{f}(1, -1+\Delta w) - \mathbf{f}(1, -1-\Delta w)}{2\Delta w} \right]$$

$$\frac{\partial \mathbf{f}}{\partial \mathbf{w}} \approx \begin{bmatrix} 47.316 & 4.711 \\ 0.0 & 1.00 \end{bmatrix}$$

$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

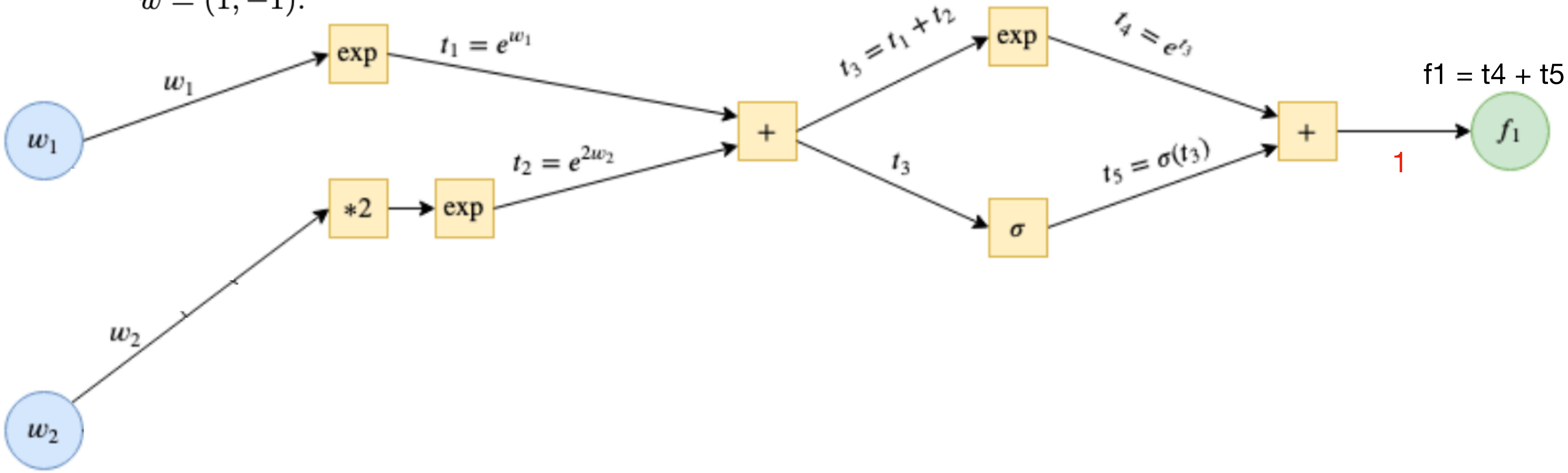
(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

$\vec{w} = (1, -1)$.

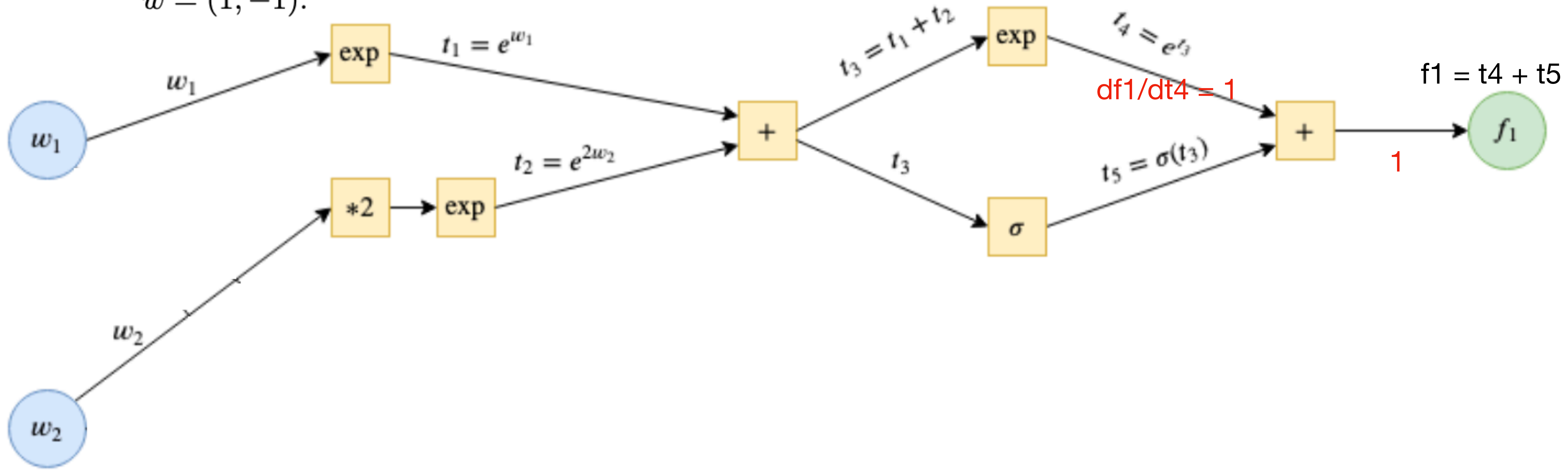


$$f_1(w_1, w_2) = e^{e^{w_1}+e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

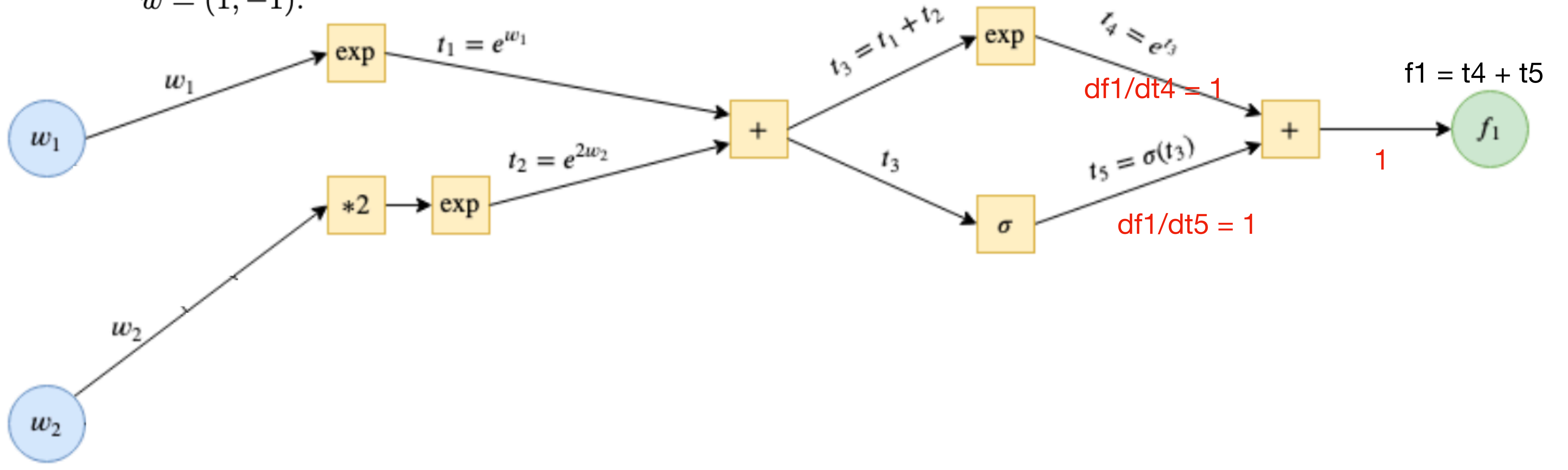
$\vec{w} = (1, -1)$.



$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$
$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

$\vec{w} = (1, -1)$.

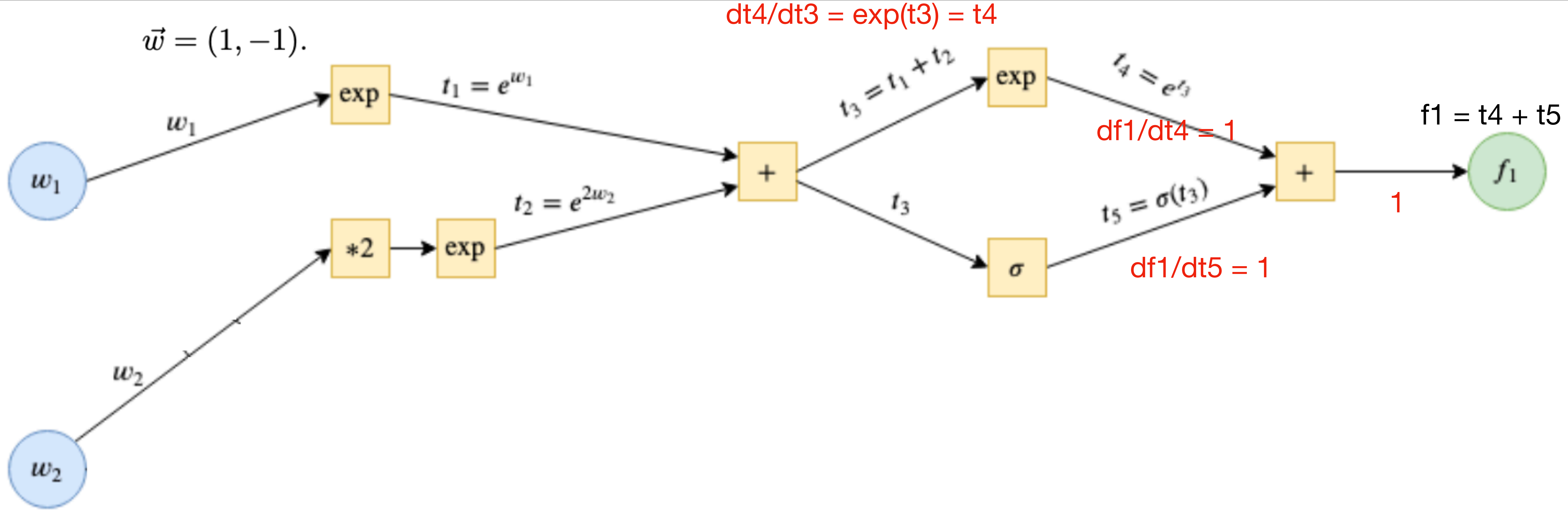


$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

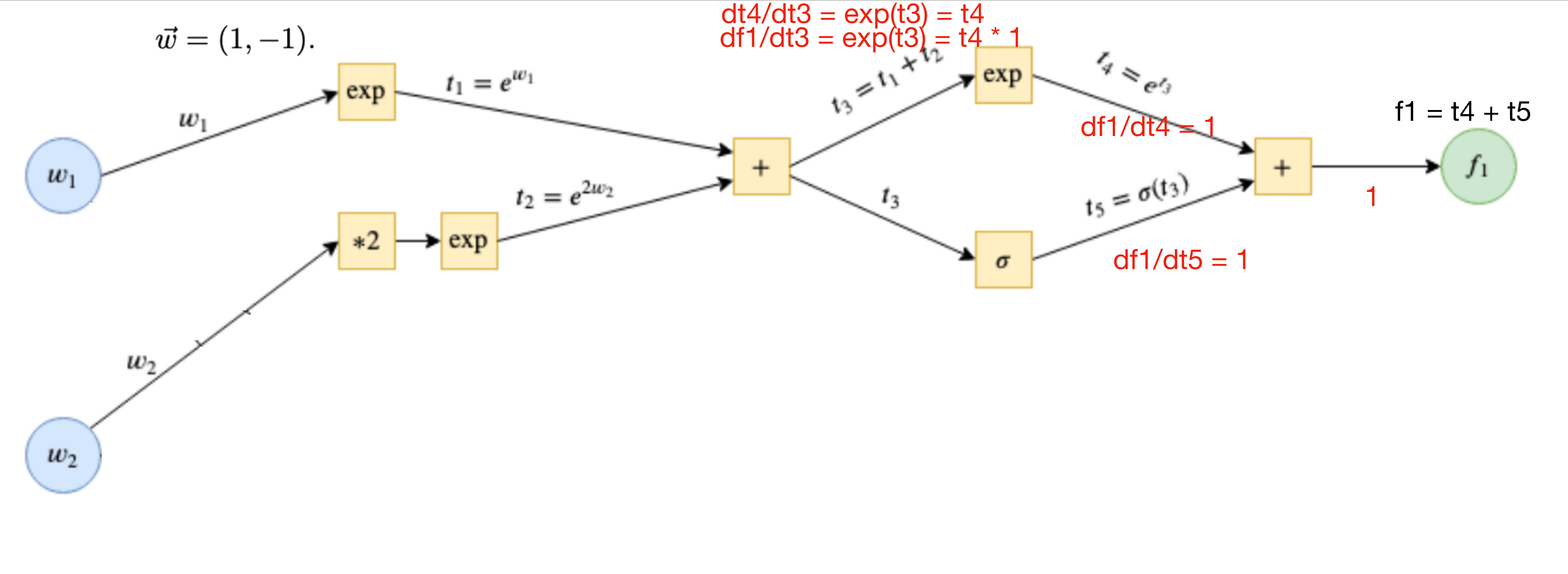
$\vec{w} = (1, -1)$.



$$f_1(w_1, w_2) = e^{e^{w_1}+e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

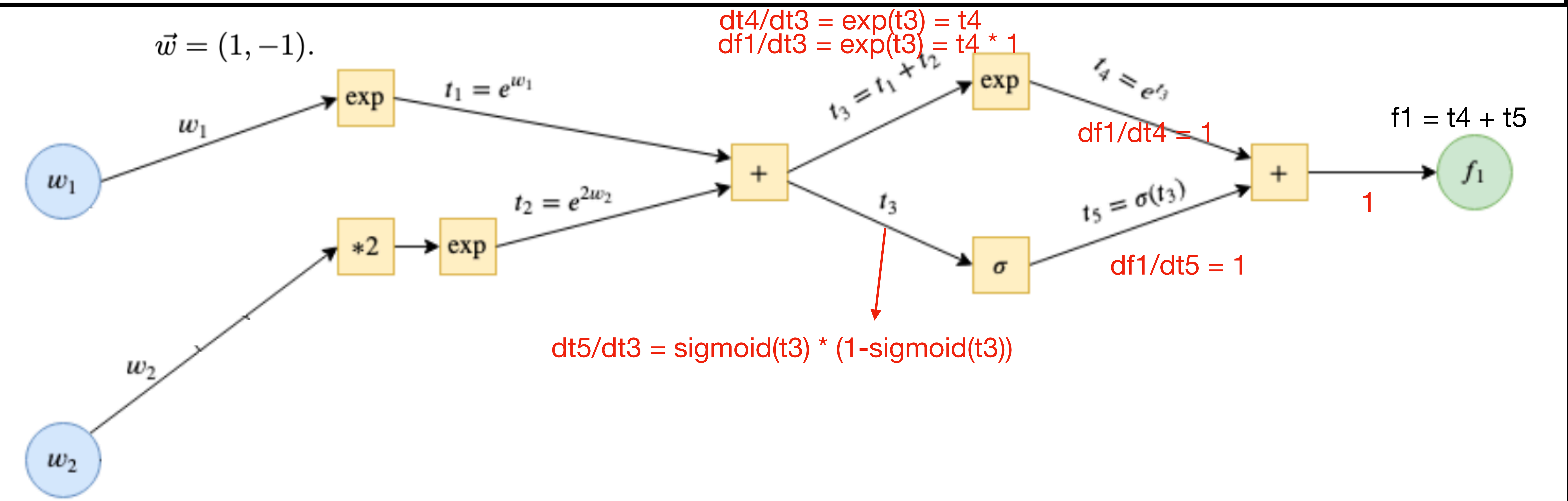
(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.



$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

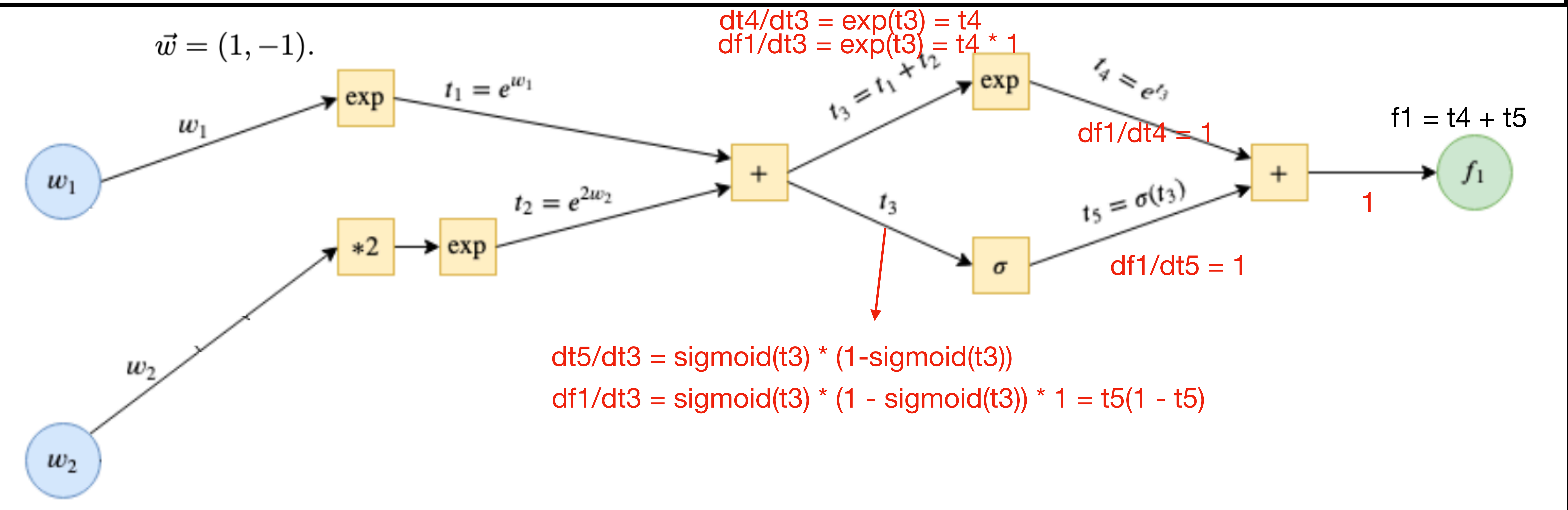
(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.



$$f_1(w_1, w_2) = e^{e^{w_1}+e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

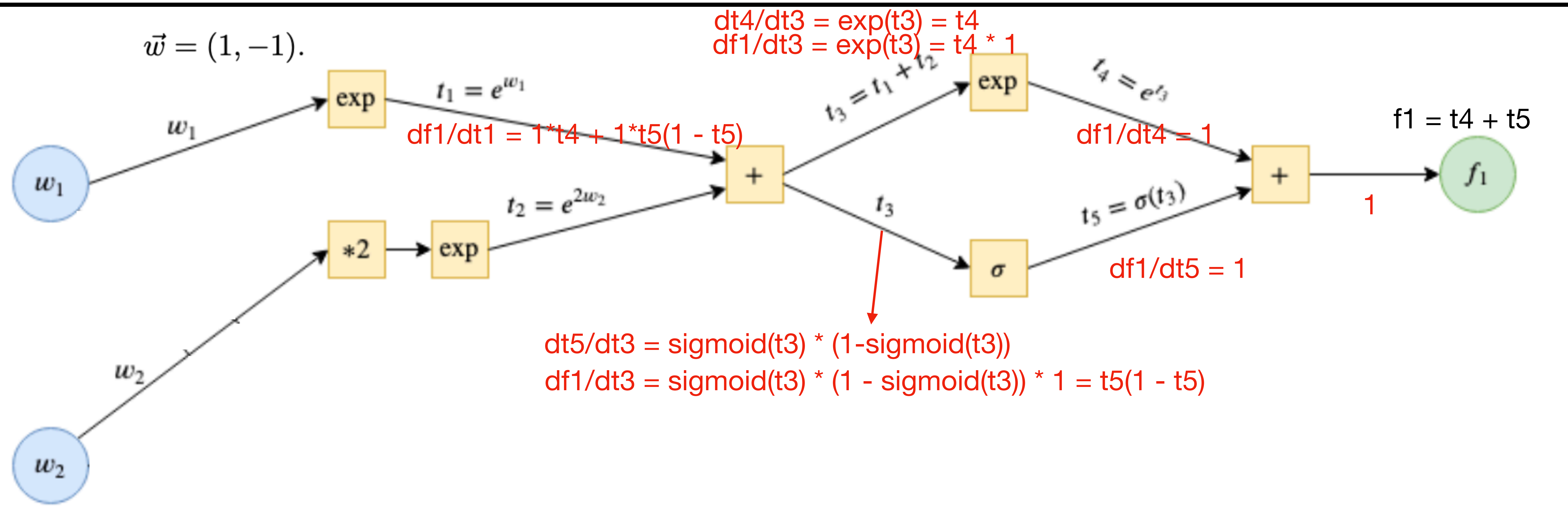


$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \quad (5)$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \quad (6)$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

$\vec{w} = (1, -1)$.

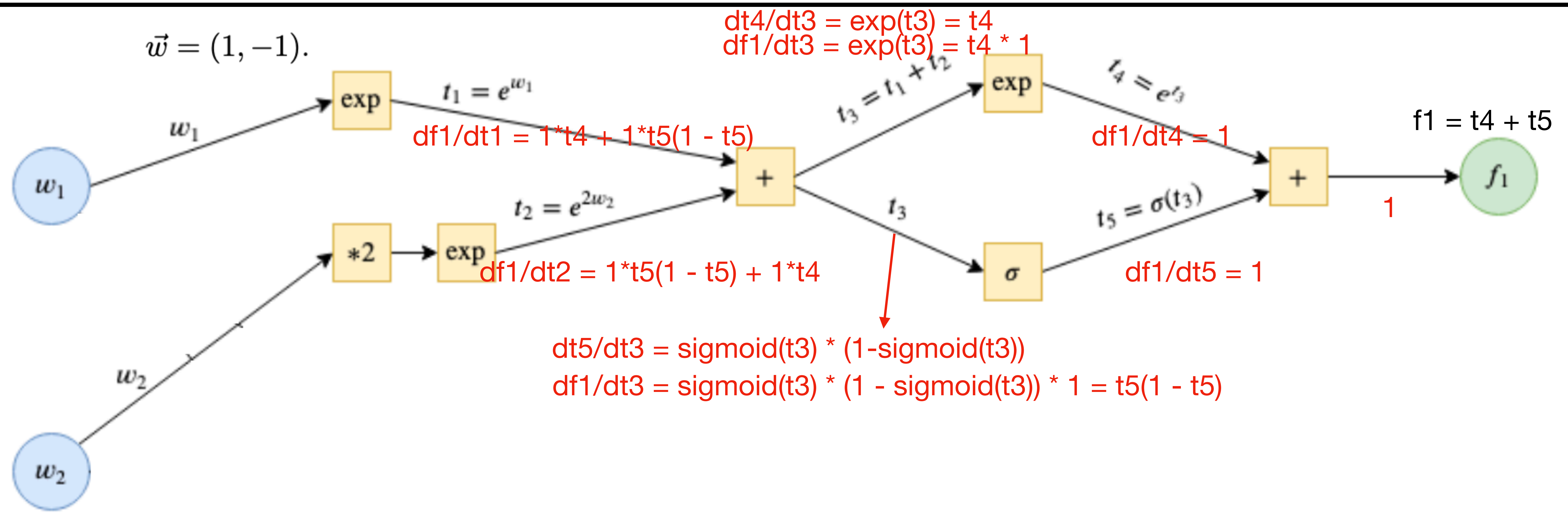


$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

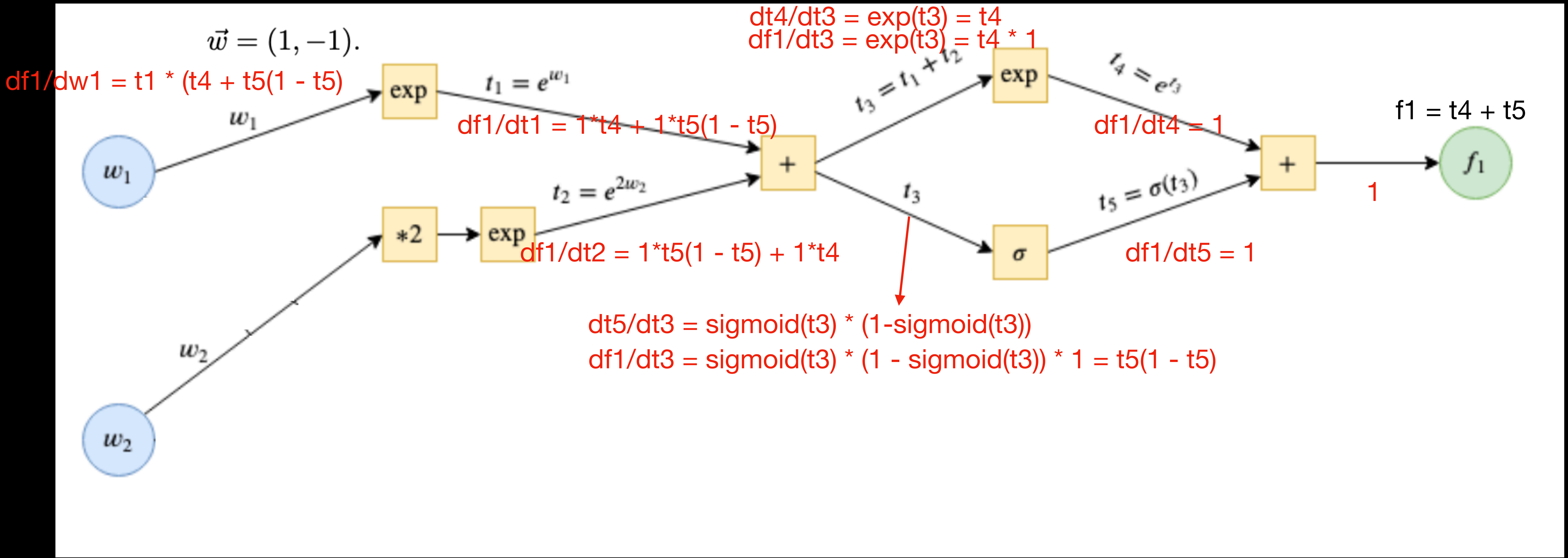
$\vec{w} = (1, -1)$.



$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \quad (5)$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \quad (6)$$

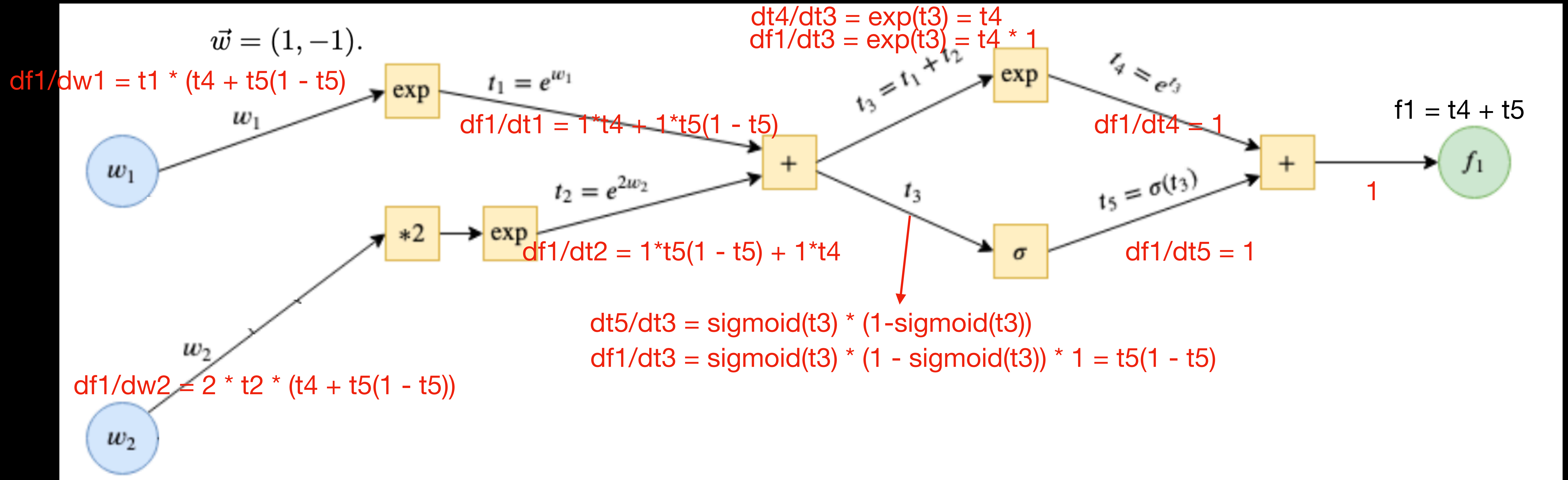
(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.



$$f_1(w_1, w_2) = e^{e^{w_1}+e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

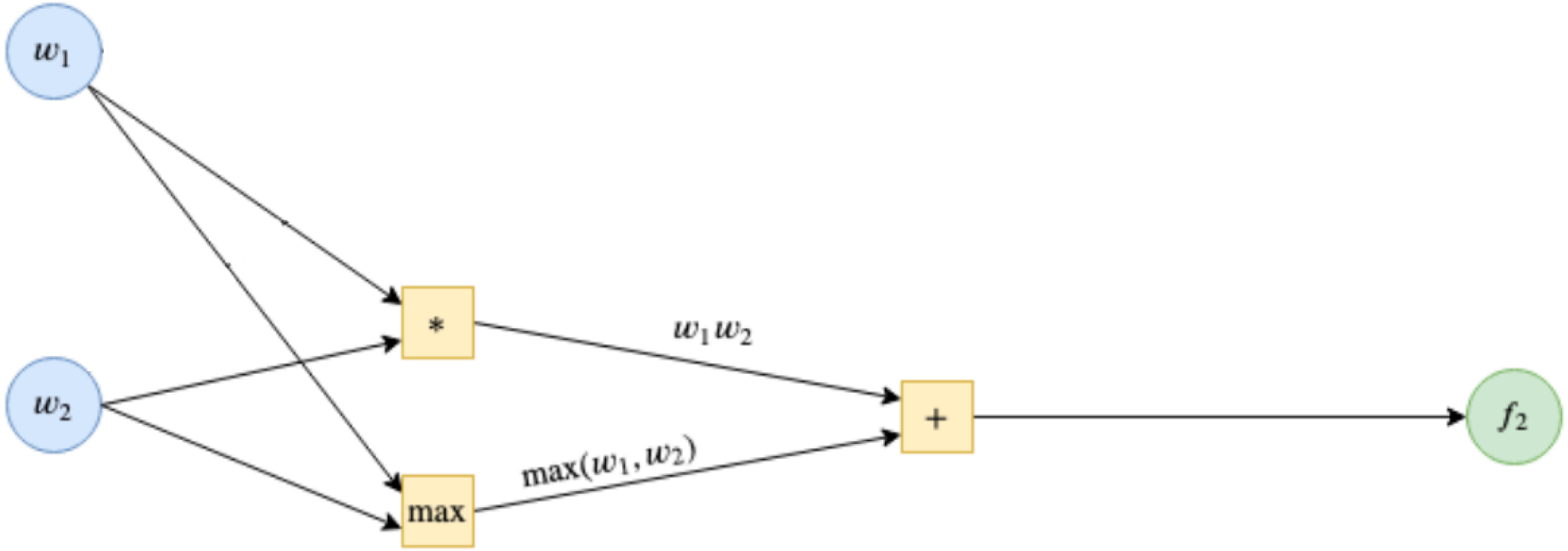


$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

$\vec{w} = (1, -1)$.

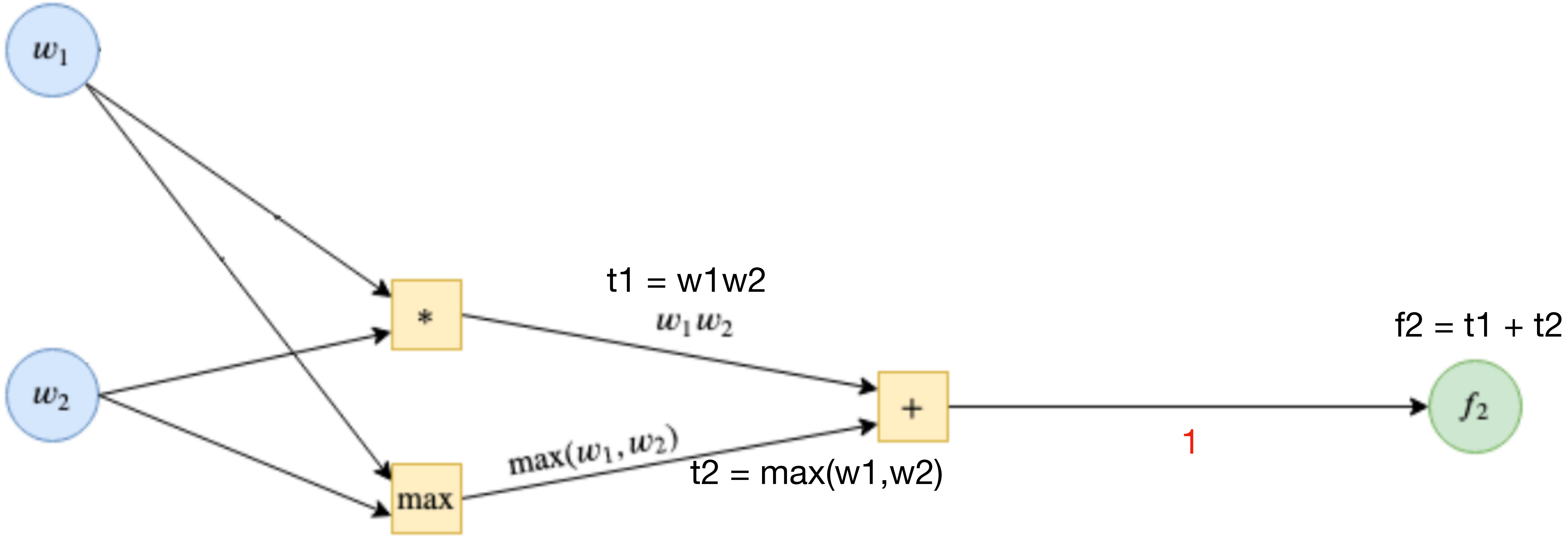


$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \quad (5)$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \quad (6)$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

$\vec{w} = (1, -1)$.

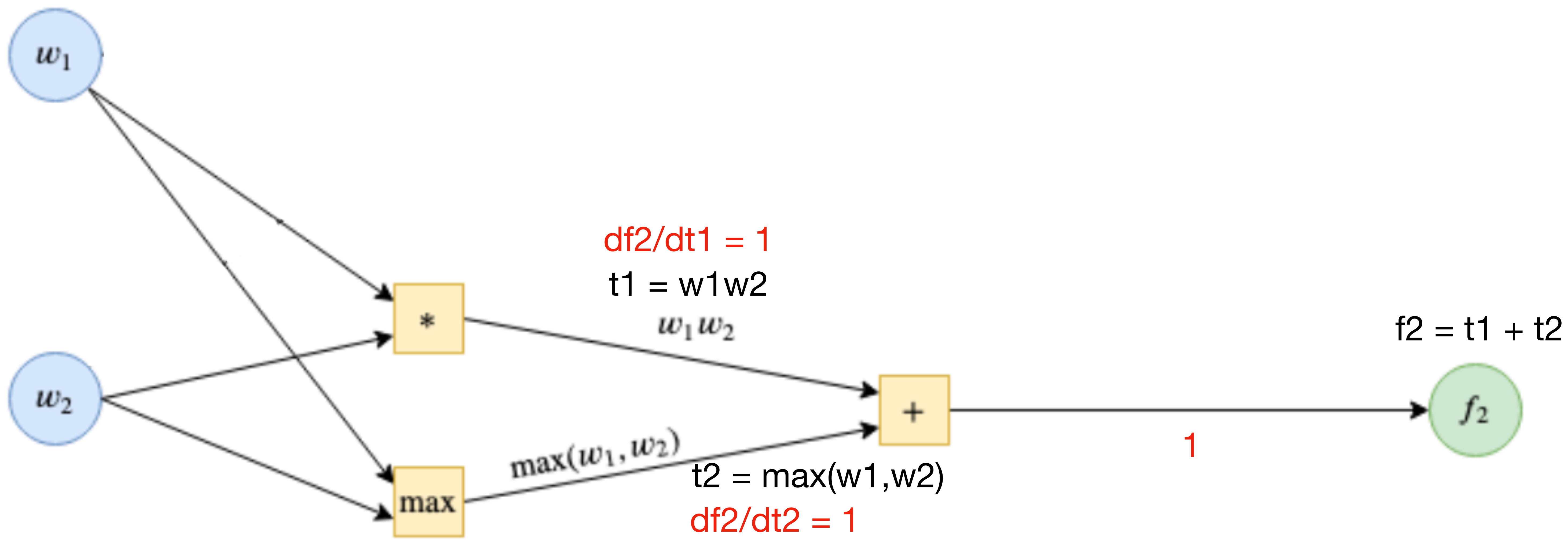


$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

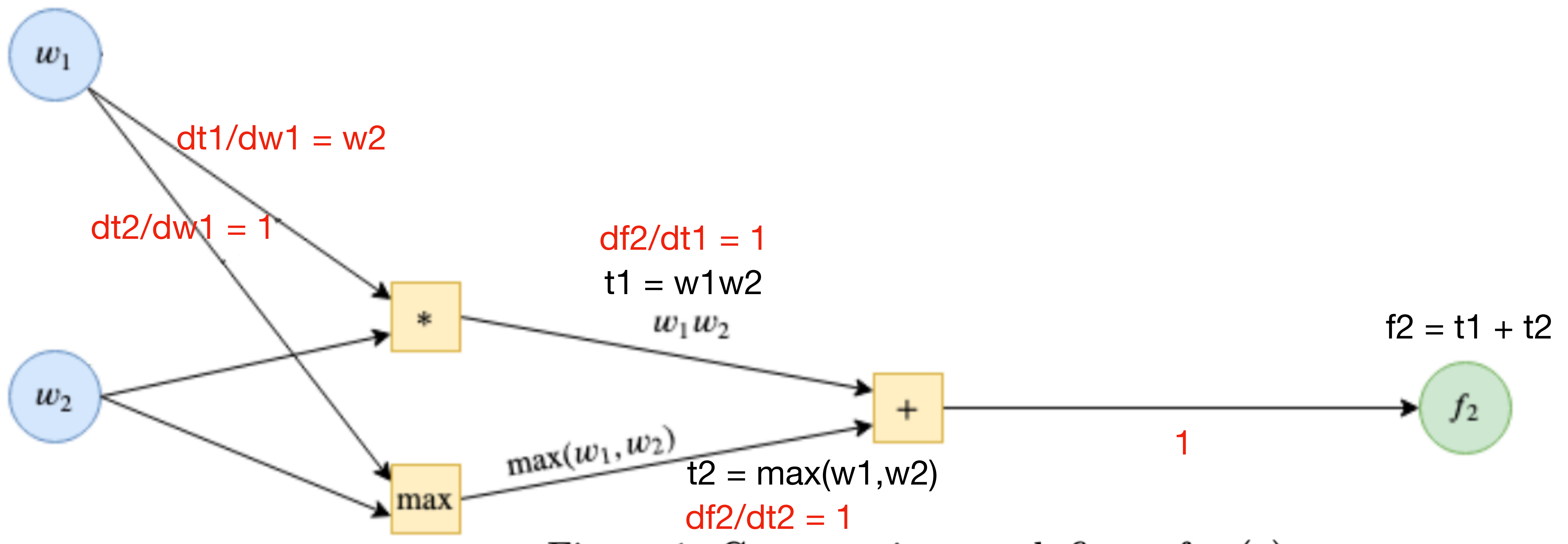
$\vec{w} = (1, -1)$.



$$f_1(w_1, w_2) = e^{e^{w_1}+e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \quad (5)$$
$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \quad (6)$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

$\vec{w} = (1, -1)$.

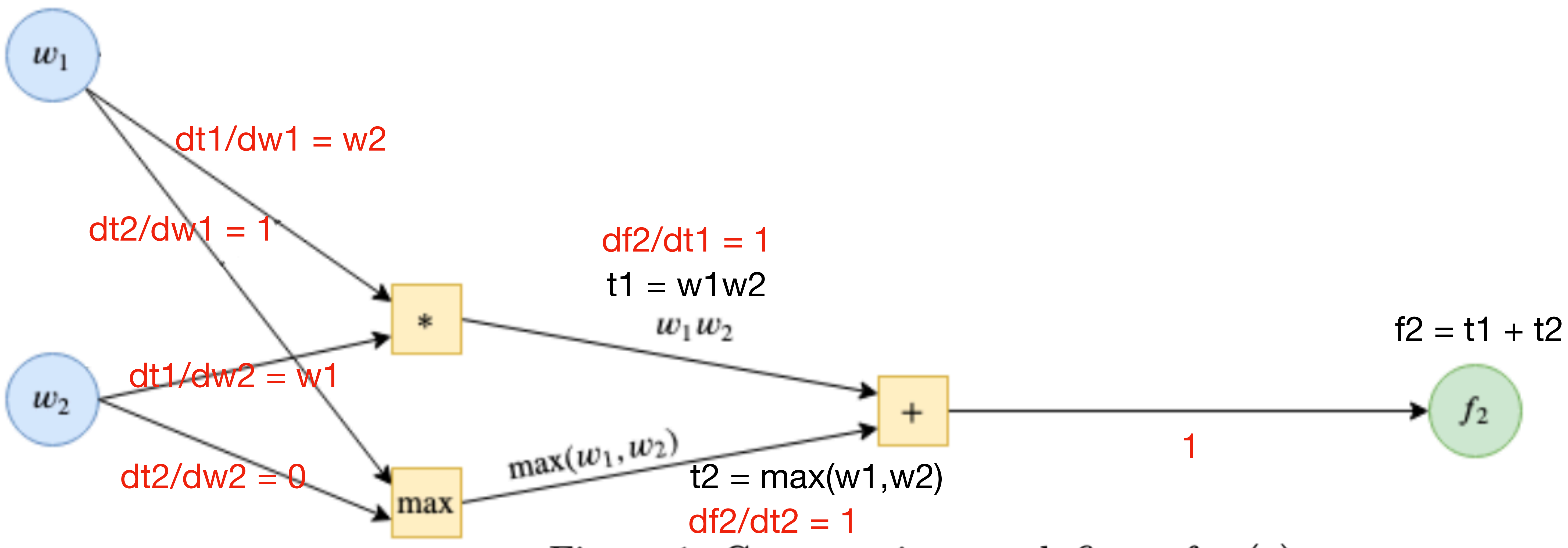


$$f_1(w_1, w_2) = e^{e^{w_1}+e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

$\vec{w} = (1, -1)$.



$$f_1(w_1, w_2) = e^{e^{w_1} + e^{2w_2}} + \sigma(e^{w_1} + e^{2w_2}) \tag{5}$$

$$f_2(w_1, w_2) = w_1 w_2 + \max(w_1, w_2) \tag{6}$$

(d) At this $\vec{w}$, compute the Jacobian using backward mode auto-differentiation.

$\vec{w} = (1, -1)$.