1 On the discrete convolution.

Definition 1. *Define a* ***centered image domain*** *as a subset $D \subseteq \mathbb{R}$ of the form $D = [d, u] \times [l, r] \cap \mathbb{Z}^2$, where $d, u, l, r \in \mathbb{Z}$, $d \leq 0 \leq u$ $l \leq 0 \leq r$, and both $-1 \leq d + u \leq 0$ and $-1 \leq l + r \leq 0$.*

This means taking the usual representation for the domains of (grayscale) digital images (from 0 to the height or the width minus one), and centering it around the origin, with the negative parts of each dimension being one pixel larger than the strictly positive parts if the number of pixels in that dimension is even, and equal in size if it is odd.
Consider a grayscale image I represented as a 2D centered function taking real values:

$$I \colon D(I) =\!\!\rightarrow \mathbb{R},$$

where $D(I)$ is the unique centered image domain determined by the resolution of I. Namely, if I has a height of H and a width of W (in pixels), then

$$D(I) = \left[-\left\lceil \frac{H-1}{2} \right\rceil, \left\lfloor \frac{H-1}{2} \right\rfloor \right] \times \left[-\left\lceil \frac{W-1}{2} \right\rceil, \left\lfloor \frac{W-1}{2} \right\rfloor \right] \tag{1}$$

Conversely, if an image I has a centered domain $D(I) = [d, u] \times [l, r] \cap \mathbb{Z}^2$, its height H and width W (in total pixels) are given by

$$H = u - d + 1, \qquad W = r - l + 1 \tag{2}$$

It is clear that one may extend this representation to an arbitrarily larger domain using different methods. This is usually referred to as **padding**. For instance, one could add zeros or mirror I around its edges forming an infinite pattern. It is also possible to do the converse: going from a given domain to an arbitrarily smaller domain by simply restricting it.

Definition 2. *Given a* ***kernel*** *g and an* ***image*** *I, we define the* ***convolution*** *$g * I$ in the domain D as follows: for each $x \in D$,*

$$g * I(x) = \sum_{y \in D(g)} g(y) I(x - y),$$

where I has been extended if necessary to accommodate for out-of-bounds values $x - y$.

Note how one may **choose** the domain of $g * I$ by simply padding I as necessary. However, the case where I needs not be extended and the domain is maximal is of particular interest. It is simple to check that this happens whenever $x = (x_1, x_2)$ satisfies the constraints

$$\begin{array}{c} d_I + u_g \leq x_1 \leq d_g + u_I \\ l_I + r_g \leq x_2 \leq l_g + r_I \end{array},$$

meaning that the output image would have height $(u_I - u_g) - (u_g - d_g) + 1 = H_I - H_g + 1$ and width $(r_I - l_I) - (r_g - l_g) + 1 = W_I - W_g + 1$, where H_I, W_I and H_g, W_g are the height and width (in pixels) of I and g, respectively.
If one wishes to obtain a domain $D = [d, u] \cap [l, r] \cap \mathbb{Z}^2$ with height $H \geq H_I - W_g + 1$ and width $W \geq W_I - W_g + 1$, it is possible to characterize the padding needed in I: call d_p, u_p, l_p and r_p the

bottom, top, left and right padding required, respectively (all are nonnegative quantities). Then, the following equalities should be satisfied:

$$\begin{cases} d_I - d_p = d - u_g \\ u_I + u_p = u - d_g \\ l_I - l_p = l - r_g \\ r_I + r_p = r - l_g \end{cases},$$

from which it is trivial to obtain the padding amounts. Note that this may be combined with eq. (1) in order to obtain the padding amounts from the (more intuitive) height and width of the desired output.

Theorem 1. *For a fixed domain D, zero-padded convolution is commutative:*

$$g * I\big|_D = I * g\big|_D.$$

Proof. For any given $x \in D$, since I is being zero-padded,

$$g * I(x) = \sum_{y \in D(g) \cap (x - D(I))} g(y) I(x - y),$$

where $x - D(I)$ is the usual set addition notation $b + A = \{b + a | a \in A\}$. Now introduce a variable change $z = x - y$ (invertible via $y = x - z$) and note that $y \in D(g) \cap (x - D(I))$ if, and only if, $z \in (x - D(g)) \cap D(I)$. Therefore,

$$\sum_{y \in D(g) \cap (x - D(I))} g(y) I(x - y) = \sum_{z \in (x - D(g)) \cap D(I)} g(x - z) I(z),$$

and the last term is clearly $I * g(x)$. □

1.1 Derivative of convolution.

Consider the following functional:

$$\tilde{F}(g) = \frac{1}{2} \|g * J - I_2\|_2^2,$$

where J and I_2 are images of the same resolution and the convolution domain is chosen so as to have the same resolution as them.
Claim: $\tilde{F}$ is Fréchet differentiable and $\nabla F(g) = (g * J - I_2) * \overline{J}$, where the inner convolution has the same domain as I_2, the outer convolution has the same domain as g, and $\overline{J}$ is given by $\overline{J}(x) = J(-x)$.
Proof. First note that, for a given $k \in D(g)$

$$\frac{\partial (g * J)}{\partial g(k)} = \frac{\partial \left(\sum_{y \in D(g)} g(y) J(x - y) \right)}{\partial g(k)} = J(x - k).$$

Now, fully "unroll" $\tilde{F}$ into a summatory

$$\tilde{F}(g) = \frac{1}{2} \sum_{x \in D(J)} \left(g * J(x) - I_2(x) \right)^2.$$

Combining the two previous equalities and the chain rule, we obtain

$$\frac{\partial \tilde{F}}{\partial g(k)} = \sum_{x \in D(J)} (g * J(x) - I_2(x)) \cdot J(x-k) = \sum_{x \in D(J)} \big(g * J - I_2)(x) \cdot \overline{J}(k-x) = \big((g * J - I_2) * \overline{J}\big)\,(k).$$

Therefore, $\nabla \tilde{F}(g) = \big(g * J - I_2\big) * \overline{J}$, as claimed. $\square$