

Why use this resource?

This resource pulls together ideas developed in [Mapping a function](#) and [Mapping a derivative](#). The first of those resources investigated the detailed behaviour of linear functions when represented using mapping diagrams, while the second developed the idea that a derivative is a *local scaling*. In this resource, students explore the effect of composing two functions, first linear functions and then more general functions. The warm-up leads to the idea that a composition of scalings is just a scaling, with an easily specifiable scale factor. This then allows for a straightforward visual and conceptual justification of the chain rule.

A fuller discussion of the potential power of this approach is discussed in the blog post [Percolation, patience and the chain rule](#).

Preparation

Students will need to have thought about derivatives as local scalings to gain the most benefit from this resource, as explored in [Mapping a derivative](#).

Students will either need to be able to draw mapping diagrams or use the GeoGebra applet on the website. Blank mapping diagrams can be downloaded and printed in advance.

Possible approach

The warm-up invites students to think about how they might represent a composition of two linear functions on a mapping diagram. The remainder of the resource assumes a particular approach, but there may be others which work equally well. (It is important that the representation clearly shows f and g as well as the composite.) The second question then points towards finding relationships between the functions.

The main problem extends this idea to a composition of any two (differentiable) functions, and students could be asked to predict the behaviour they might expect to see. How can they use mapping diagrams to find the derivative of h at a specified point? They may need to recall the definition of a derivative obtained from earlier work with mapping diagrams before they can proceed, and the applet in either this resource or [Mapping a derivative](#) could be useful for this purpose. As they pull the pieces together, the meaning of the chain rule will hopefully become clear, and the need to identify how a function can be decomposed as a composition before applying the chain rule will make more sense.

Key questions

- How can we differentiate a composition of functions?
- Why does this work?

Possible extension

- What would happen if we had a composition of three functions, $h(g(f(x)))$?