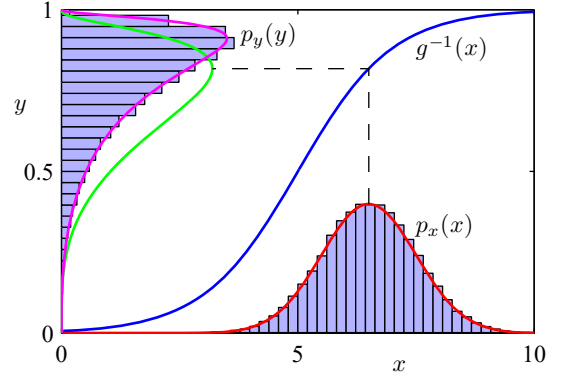


**Figure 2.12** Example of the transformation of the mode of a density under a nonlinear change of variables, illustrating the different behaviour compared to a simple function.



If we simply transform  $p_x(x)$  as a function of  $x$  we obtain the green curve  $p_x(g(y))$  shown in Figure 2.12, and we see that the mode of the density  $p_x(x)$  is transformed via the sigmoid function to the mode of this curve. However, the density over  $y$  transforms instead according to (2.71) and is shown by the magenta curve on the left side of the diagram. Note that this has its mode shifted relative to the mode of the green curve.

To confirm this result, we take our sample of 50,000 values of  $x$ , evaluate the corresponding values of  $y$  using (2.75), and then plot a histogram of their values. We see that this histogram matches the magenta curve in Figure 2.12 and not the green curve.

### 2.4.1 Multivariate distributions

We can extend the result (2.71) to densities defined over multiple variables. Consider a density  $p(\mathbf{x})$  over a  $D$ -dimensional variable  $\mathbf{x} = (x_1, \dots, x_D)^T$ , and suppose we transform to a new variable  $\mathbf{y} = (y_1, \dots, y_D)^T$  where  $\mathbf{x} = \mathbf{g}(\mathbf{y})$ . Here we will limit ourselves to the case where  $\mathbf{x}$  and  $\mathbf{y}$  have the same dimensionality. The transformed density is then given by the generalization of (2.71) in the form

$$p_{\mathbf{y}}(\mathbf{y}) = p_{\mathbf{x}}(\mathbf{x}) |\det \mathbf{J}| \quad (2.76)$$

where  $\mathbf{J}$  is the *Jacobian matrix* whose elements are given by the partial derivatives  $J_{ij} = \partial g_i / \partial y_j$ , so that

$$\mathbf{J} = \begin{bmatrix} \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_1}{\partial y_D} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_D}{\partial y_1} & \cdots & \frac{\partial g_D}{\partial y_D} \end{bmatrix}. \quad (2.77)$$

Intuitively, we can view the change of variables as expanding some regions of space and contracting others, with an infinitesimal region  $\Delta \mathbf{x}$  around a point  $\mathbf{x}$  being transformed to a region  $\Delta \mathbf{y}$  around the point  $\mathbf{y} = \mathbf{g}(\mathbf{x})$ . The absolute value of the determinant of the Jacobian represents the ratio of these volumes and is the same factor