for a function $f(x)$, if ~~we change~~ we define a transformation $x = g(y)$. this defines a new function of $y$ given by.

$$\hat{f}(y) = f(g(y))$$

assuming the mode of $f(x)$ occurs at $x = \hat{x}$

its easy to verify that the mode of $\hat{f}(y)$ occurs at $y = \hat{y} = g^{-1}(\hat{x})$

Therefore, finding a mode w.r.t. $x$ is equivalent to first transforming to the variable $y$, then finding the mode w.r.t. $y$, and transforming back to $x$.

For probability density, the behavior is different because we cannot simply write $P_Y(y) = P_X(g(y))$.

according to (1.27), $P_Y(y) = P_X(g(y))\,|g'(y)|$

Let $|g'(y)| = s\,g'(y)$ where $s \in \{-1, 1\}$, so

$$P_Y(y) = P_X(g(y))\, s\, g'(y)$$

$$\Rightarrow P_Y'(y) = s P_X'(g(y)) \{g'(y)\}^2 + s\, P_X(g(y))\, g''(y)$$

if the second term on the right hand side is 0, then clearly the mode of $P_Y(y)$ occurs at $y = \hat{y} = g^{-1}(x)$

however, if the second term is not zero, $\hat{y} \neq g^{-1}(x)$

When the second term is zero? if $x = g(y)$ is a linar function, the second term will be zero.

if not, the second term will not be zero.