The mathematical model for LTI systems is,

$$\frac{\mathrm{d}^n y(t)}{\mathrm{d}t^n} + a_{n-1}\frac{\mathrm{d}^{n-1} y(t)}{\mathrm{d}t^{n-1}} + \cdots + a_0 y(t) = b_m \frac{\mathrm{d}^m r(t)}{\mathrm{d}t^m} + b_{m-1}\frac{\mathrm{d}^{m-1} r(t)}{\mathrm{d}t^{m-1}} + \cdots + b_0 r(t) \tag{1}$$

and n initial conditions,

$$\frac{\mathrm{d}^{n-1} y(t_0)}{\mathrm{d}t^{n-1}} = \frac{\mathrm{d}^{n-2} y(t_0)}{\mathrm{d}t^{n-2}} = \cdots = y(t_0) = 0 \tag{2}$$

Its transfer function is defined as,

$$G(s) = \frac{b(s)}{a(s)} = \frac{b_m s^m + \ldots + b_1 s + b_0}{s^n + a_{n-1}s^{n-1} + \ldots + a_1 s + a_0} \tag{3}$$

In all books it is stated that a causal system should obey $n \geq m$.

But with no rigorous proof.