Theorem 1. *Given X, N_Y, and Y that satisfy an ANM with a function ϕ, if there is a backward mechanism of the same form, then ϕ, P_X, P_{N_Y} must satisfy the following differential equation:*

$$\xi''' = \xi''\left(-\frac{\nu'''\phi'}{\nu''} + \frac{\phi''}{\phi'}\right) - 2\nu''\phi''\phi' + \nu'\phi''' + \frac{\nu'\nu'''\phi''\phi'}{\nu''} - \frac{\nu'(\phi'')^2}{\phi'}$$

where $\nu := \log P_{N_Y}$ and $\xi := \log P_X$, and we also have that $\nu''(y - \phi(x))\phi'(x) \neq 0$.

Also, we have that if these conditions hold, then if there is a y for which $\nu''(y - \phi(x))\phi'(x) \neq 0$ is true for every x aside from a countable set, then the set of all P_X which admit a backward model is 3-dimensional (i.e. can be contained in a 3-dimensional affine space).

Proof. Let $\pi(X, Y) := \log P(X, Y)$. Then we have that

$$\pi(x, y) = \log P(x, y) = \log(P_{N_Y}(y - \phi(x)) \cdot P_X(x)) = log(P_{N_Y}(y - \phi(x))) + log(P_X(x))$$

$$= \nu(y - \phi(x)) + \xi(x) \tag{1}$$

If there existed a backward model, then similar to our prior reasoning, it would have the form

$$P(x, y) = P_{N_X}(x - \psi(y)) \cdot P_Y(y)$$

for some function ψ. So, similar to above, we get

$$\pi(x, y) = \bar{\nu}(x - \psi(y)) + \eta(y) \tag{2}$$

where $\bar{\nu} := \log P_{N_X}$ and $\eta := \log P_Y$. Now, taking partial derivatives of (2), we get that

$$\frac{\partial \pi}{\partial y} = -\psi'(y)\bar{\nu}'(x - \psi(y)) + \eta'(y) \implies \frac{\partial^2 \pi}{\partial x \partial y} = -\psi'(y)\bar{\nu}''(x - \psi(y))$$

Similarly,

$$\frac{\partial^2 \pi}{\partial x^2} = \bar{\nu}''(x - \psi(y)).$$

Notice that since

$$\frac{\frac{\partial^2 \pi}{\partial x^2}}{\frac{\partial^2 \pi}{\partial x \partial y}} = \frac{\bar{\nu}''(x - \psi(y))}{-\psi'(y)\bar{\nu}''(x - \psi(y))} = \frac{1}{-\psi'(y)}$$

we have that

$$\frac{\partial}{\partial x}\left(\frac{\frac{\partial^2 \pi}{\partial x^2}}{\frac{\partial^2 \pi}{\partial x \partial y}}\right) = 0 \tag{3}$$

Now similarly to the above results, taking the same partial derivatives of (1), we get

$$\frac{\partial^2 \pi}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial}{\partial y}\left[\nu(y - \phi(x)) + \xi(x)\right]\right) = \frac{\partial}{\partial x}\left[\nu'(y - \phi(x))\right]$$

$$= -\phi'(x)\nu''(y - \phi(x))$$

and

$$\frac{\partial^2 \pi}{\partial x^2} = \frac{\partial}{\partial x}\left(-\nu'(y - \phi(x)\phi'(x) + \xi'(x)\right) = \phi'(x)\nu''(y - \phi(x)\phi'(x) - \nu'(y - \phi(x))\phi''(x) + \xi''(x)$$

This gives us that

$$\frac{\frac{\partial^2 \pi}{\partial x^2}}{\frac{\partial^2 \pi}{\partial x \partial y}} = \frac{\phi'^2(x)\nu''(y-\phi(x)) - \nu'(y-\phi(x))\phi''(x) + \xi''(x)}{-\phi'(x)\nu''(y-\phi(x))}$$

In the following work, we'll leave out the inputs of $\nu, \xi, \bar{\nu}, \eta$ and their derivatives for the sake of readability. Taking the partial derivative of this with respect to x, we get

$$\frac{\partial}{\partial x}\left(\frac{\frac{\partial^2 \pi}{\partial x^2}}{\frac{\partial^2 \pi}{\partial x \partial y}}\right) = \frac{\partial}{\partial x}\left(\frac{\phi'^2\nu'' - \nu'\phi'' + \xi''}{-\phi'\nu''}\right)$$

$$= \frac{\partial}{\partial x}\left(-\phi' + \frac{\phi''\nu'}{\phi'\nu''} - \frac{\xi'}{\phi'\nu''}\right) = -\phi'' + \frac{[\phi'''\nu' - \phi''\phi'\nu''](\phi'\nu'') - [\phi''\nu'' - \phi'^2\nu'''](\phi''\nu')}{(\phi')^2(\nu'')^2} - \frac{\xi'''[\phi'\nu''] - \xi''[\phi''\nu'' - \phi'^2\nu'''}{(\phi')^2(\nu'')^2}$$

$$= -\phi'' + \frac{\phi'''\nu'}{\phi'\nu''} - \phi'' - \frac{(\phi'')^2\nu'}{(\phi')^2\nu''} + \frac{\nu'''\phi''\nu'}{(\nu'')^2} - \frac{\xi'''}{\phi'\nu''} + \frac{\xi''\phi''}{(\phi')^2\nu''} - \frac{\xi''\nu'''}{(\nu'')^2} \tag{4}$$

By (3), we know that the expression obtained in (4) is equal to 0. Thus, equating (4) to 0 and reordering terms, we get the differential equation obtained in the theorem.

Now, we will prove the second statement in the theorem. To do this, notice that our differential equation

$$\xi''' = \xi''\left(-\frac{\nu'''\phi'}{\nu''} + \frac{\phi''}{\phi'}\right) - 2\nu''\phi''\phi' + \nu'\phi''' + \frac{\nu'\nu'''\phi''\phi'}{\nu''} - \frac{\nu'(\phi'')^2}{\phi'}$$

has the form of a linear equation

$$z'(x) = z(x)G(x,y) + H(x,y) \tag{5}$$

where

$$G(x,y) = \frac{\nu'''\phi'}{\nu''} + \frac{\phi''}{\phi'} \text{ and } H(x,y) = -2\nu''\phi''\phi' + \nu'\phi''' + \frac{\nu'\nu'''\phi''\phi'}{\nu''} - \frac{\nu(\phi'')^2}{\phi'}$$

To solve linear ODEs, first, let us assume an initial condition $z_0 = z(x_0)$. Then we use the theory of integrating factors from elementary differential equations to get that the integrating factor is

$$\text{I. F} = e^{\int_{x_0}^{x} G(\hat{x},y)d\hat{x}}$$

This, along with the initial condition z_0, gives us that

$$z(x) = z_0 e^{\int_{x_0}^{x} G(\hat{x},y)d\hat{x}} + \int_{x_0}^{x} e^{\int_{x_0}^{x} G(\hat{x},y)d\hat{x}} H(\hat{x},y)d\hat{x}$$

Fix y such that $\nu''(y-\phi(x))\phi'(x) \neq 0$ for all but countably many x. We know the general solution to (5), without any initial conditions, is

$$z(x) = \frac{1}{\text{I. F}\left[\int \text{I. F}(\hat{x})H(x)d\hat{x} + C\right]}$$

So, clearly we have that given a linear 1st order ODE, z is determined by z_0. So, in our case,

$$\xi'' = z \implies \xi''(x_0) = z_0$$

Thus fixing $\xi_0 = \xi''(x_0)$ determines ξ''. Let F be a second antiderivative of z (that is, $F'' = z$). Then we have that

$$\xi'' = F'' \implies \xi' = F' + c_1$$

Thus for $\xi_0' = \xi'(x_0)$, we have $c_1 = \xi_0' - F''(x_0)$ and so fixing ξ_0' determines ξ'. Following this, we have that

$$\xi(x) = F(x) + c_1(x) + c_2$$

and similarly fixing $\xi_0 = \xi(0)$ determines c_2 and therefore ξ. Thus, we get that ξ is uniquely determined by $\xi(x_0), \xi'(x_0), \xi''(x_0)$ and so the solution space is of dimension 3.

□