1. Let . Note that
   1. We wish to find a cubic polynomial that satisfies for all . Since in this case, we have

Therefore, the polynomial with , , , is a solution.

* 1. We wish to find a cubic polynomial that satisfies for all . Since in this case, we have

Therefore, the polynomial with is a solution. We have established that the function is a piecewise.

* 1. We have

Thus .

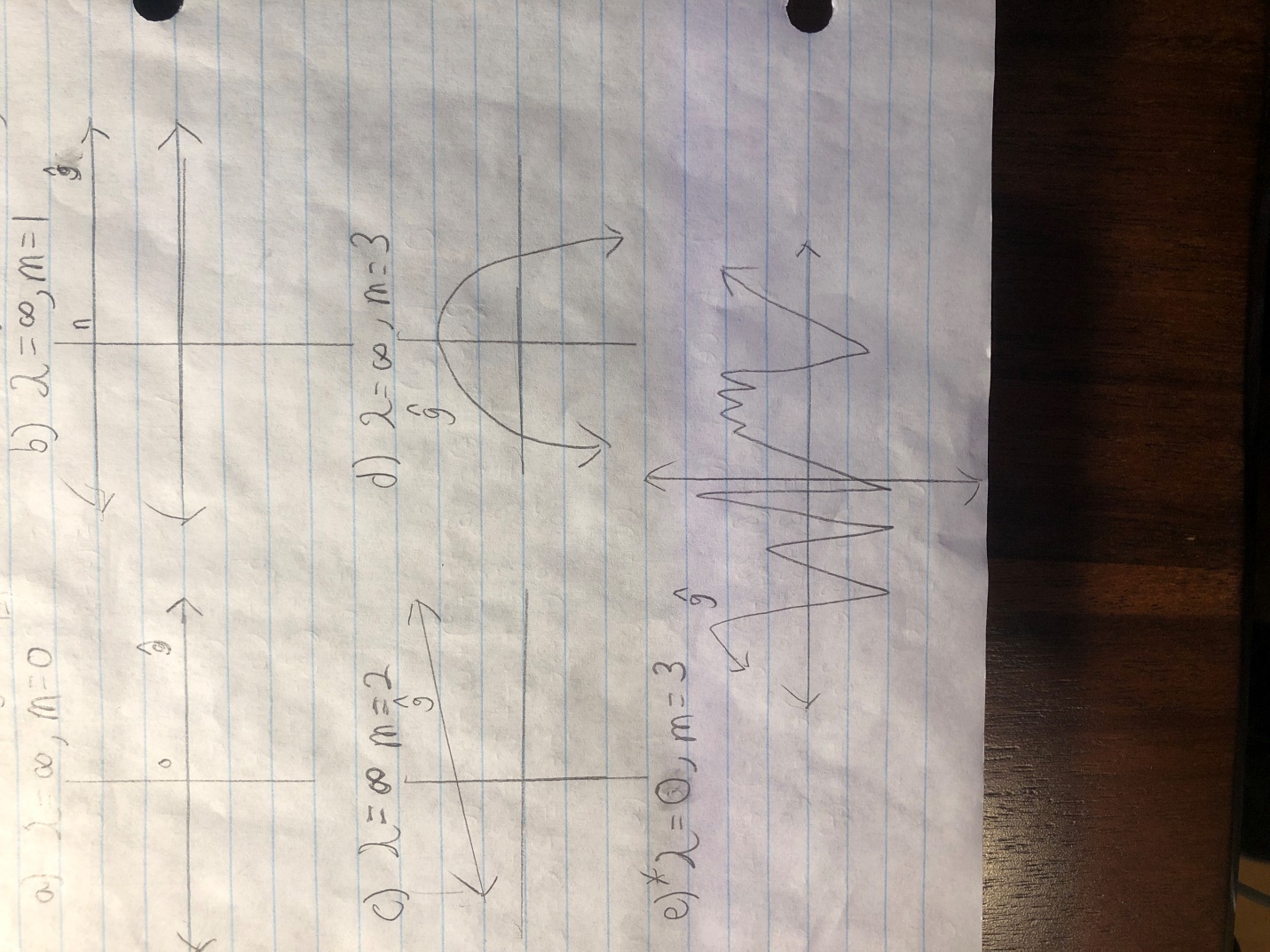
* 1. We have

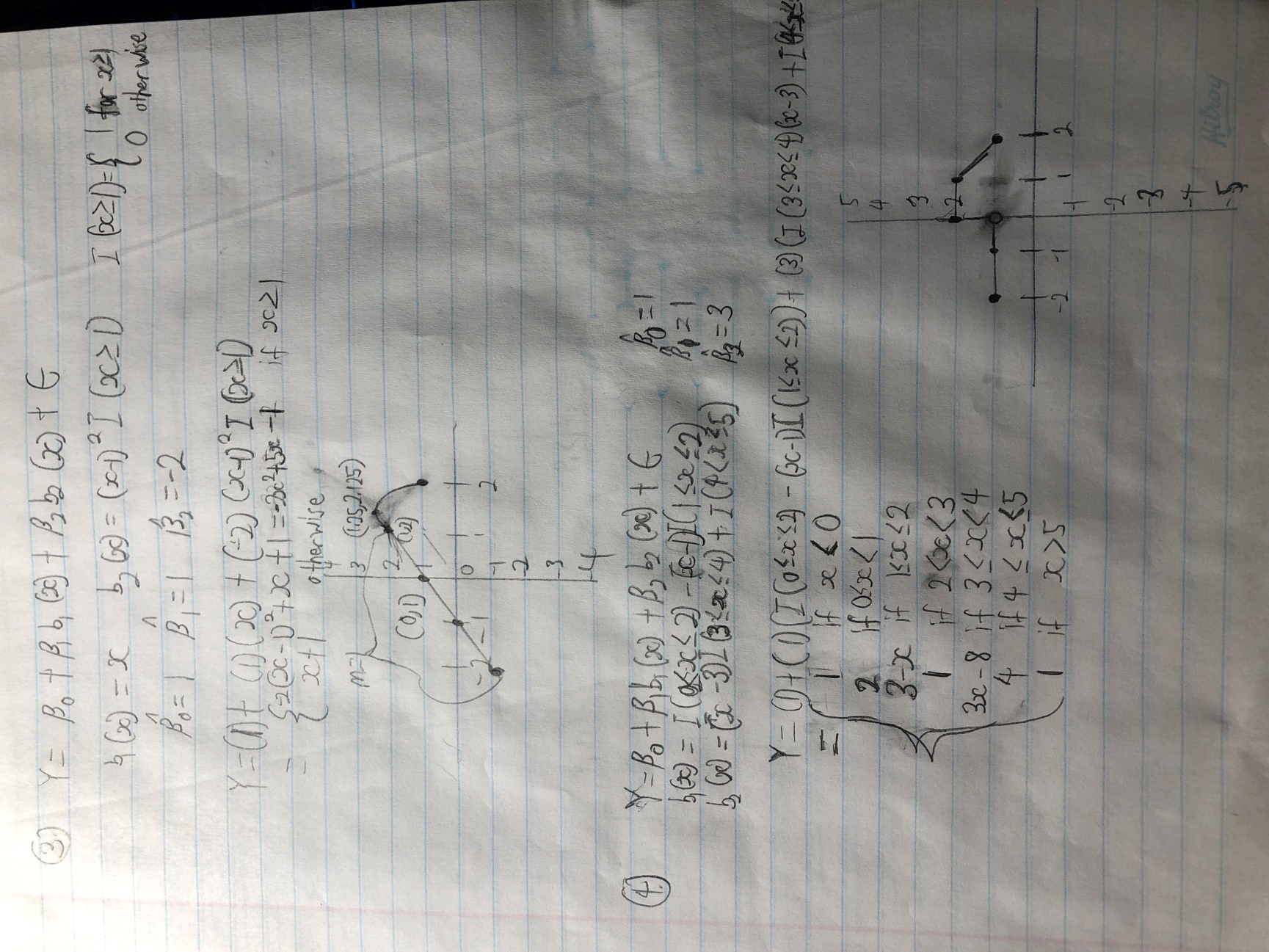
Thus .

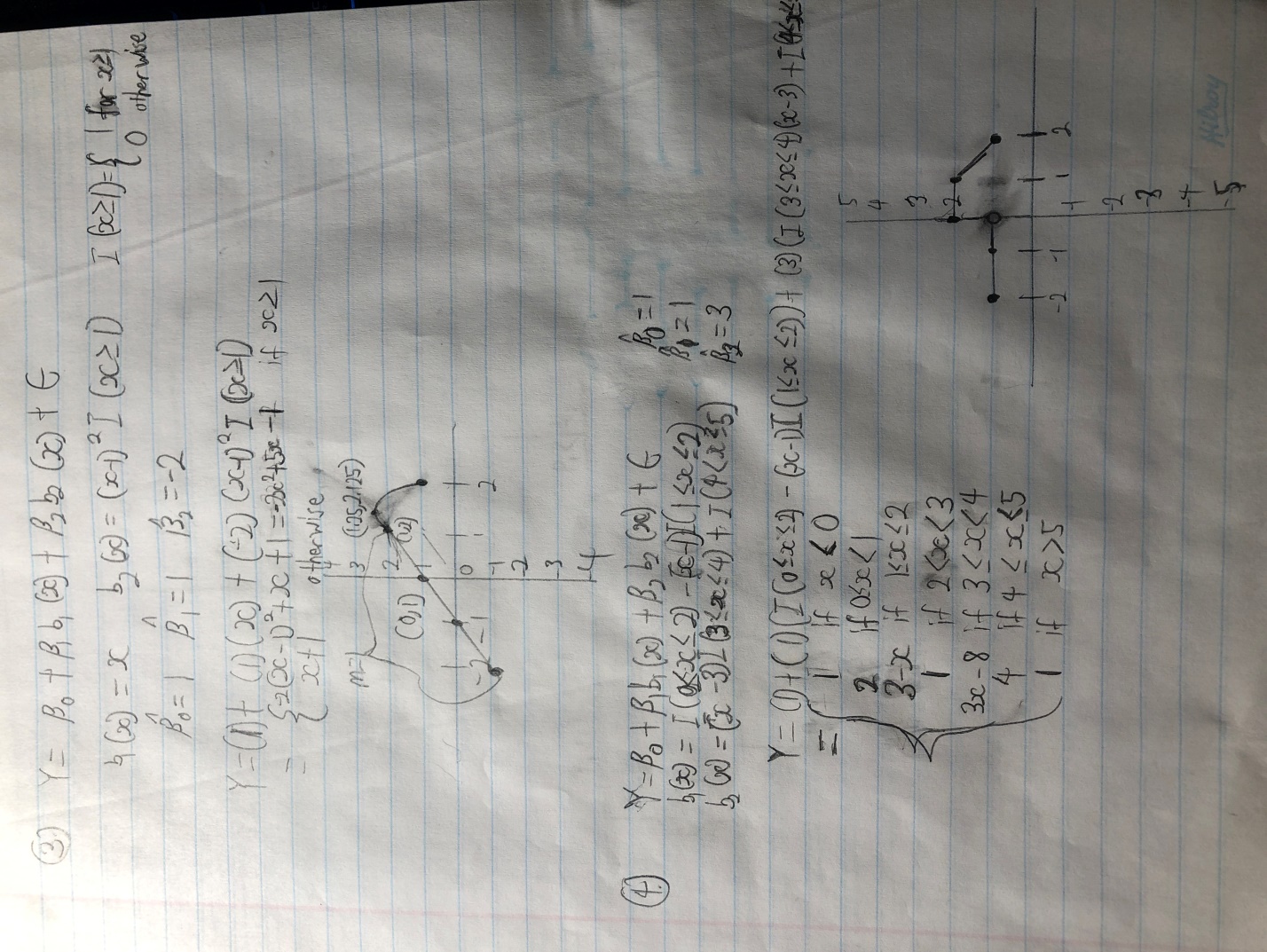
* 1. We have

Thus . Therefore, we have shown that is indeed a cubic spline regardless of the values of its coefficients.

1. .



\*If , then regardless of the value of , is simply the least squares fit.

1. Fitting the model with estimates between and .
2. Fitting the model with estimates between and .
3. Two curves
   1. As lambda approaches infinity, g2 will likely have the smaller training RSS. This is because g2 forces the 4th derivative of g to be 0 rather than the 3rd derivative of g, allowing the estimate of g to be a higher order polynomial and thus be more flexible.
   2. As lambda approaches infinity, determining whichever model has the smaller test RSS depends on the true relationship between the predictor and the response. If their true relationship is quadratic or even-degree, then g1 would likely perform better and have a smaller test RSS. If the relationship is cubic or odd-degree, then g2 would perform better. It is also possible that, due to g2’s flexibility, it is likely to overfit the data, leaving g1 to perform better on more datasets in general.
   3. For lambda=0, the penalty terms in both become irrelevant, leaving both g1 and g2 to simply take the form of the least squares fit and making them the same curve. Therefore, they will have the same test and training RSS in this case.