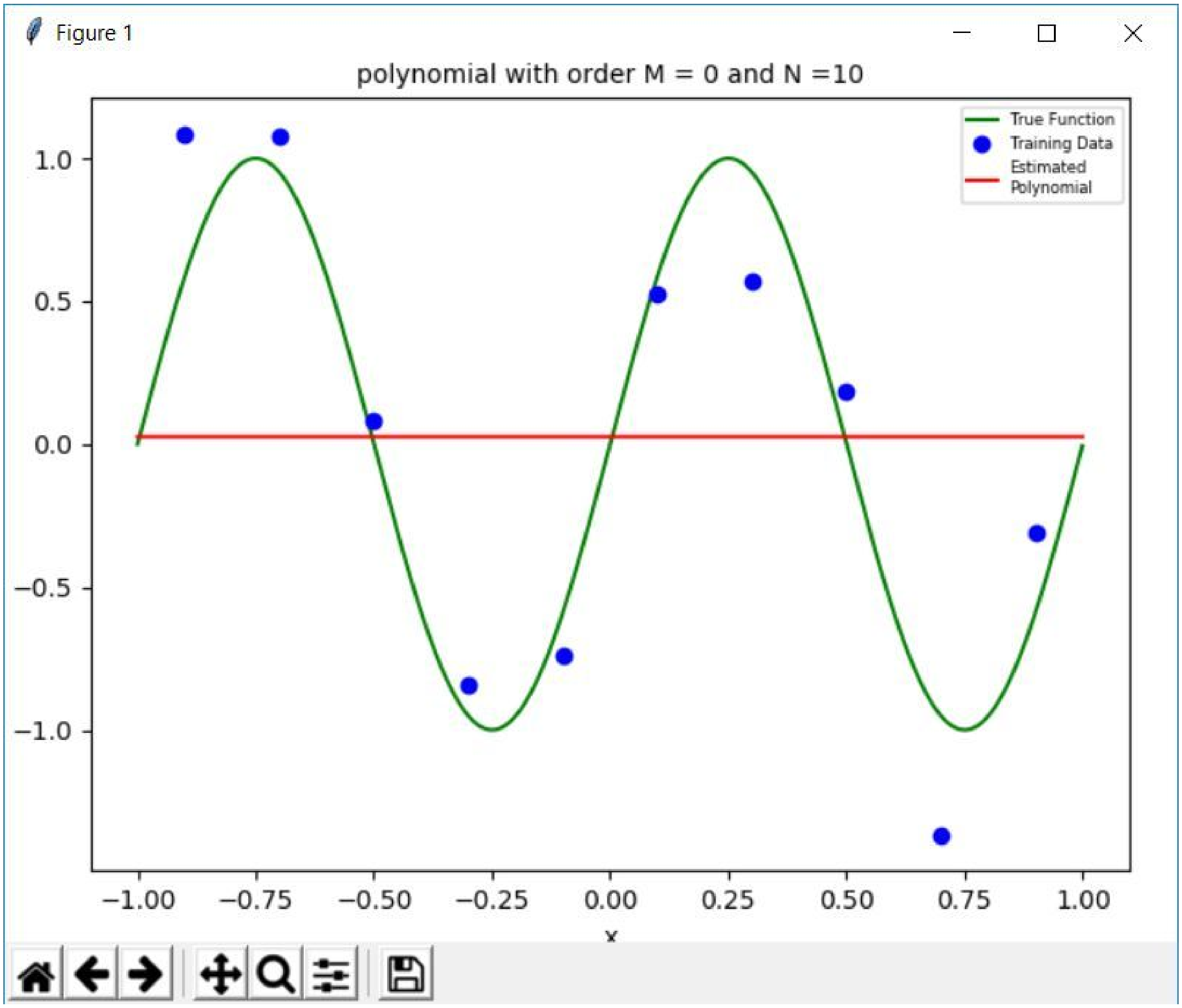

Figure 1: the polynomial of 0th order is a constant value of the mean error

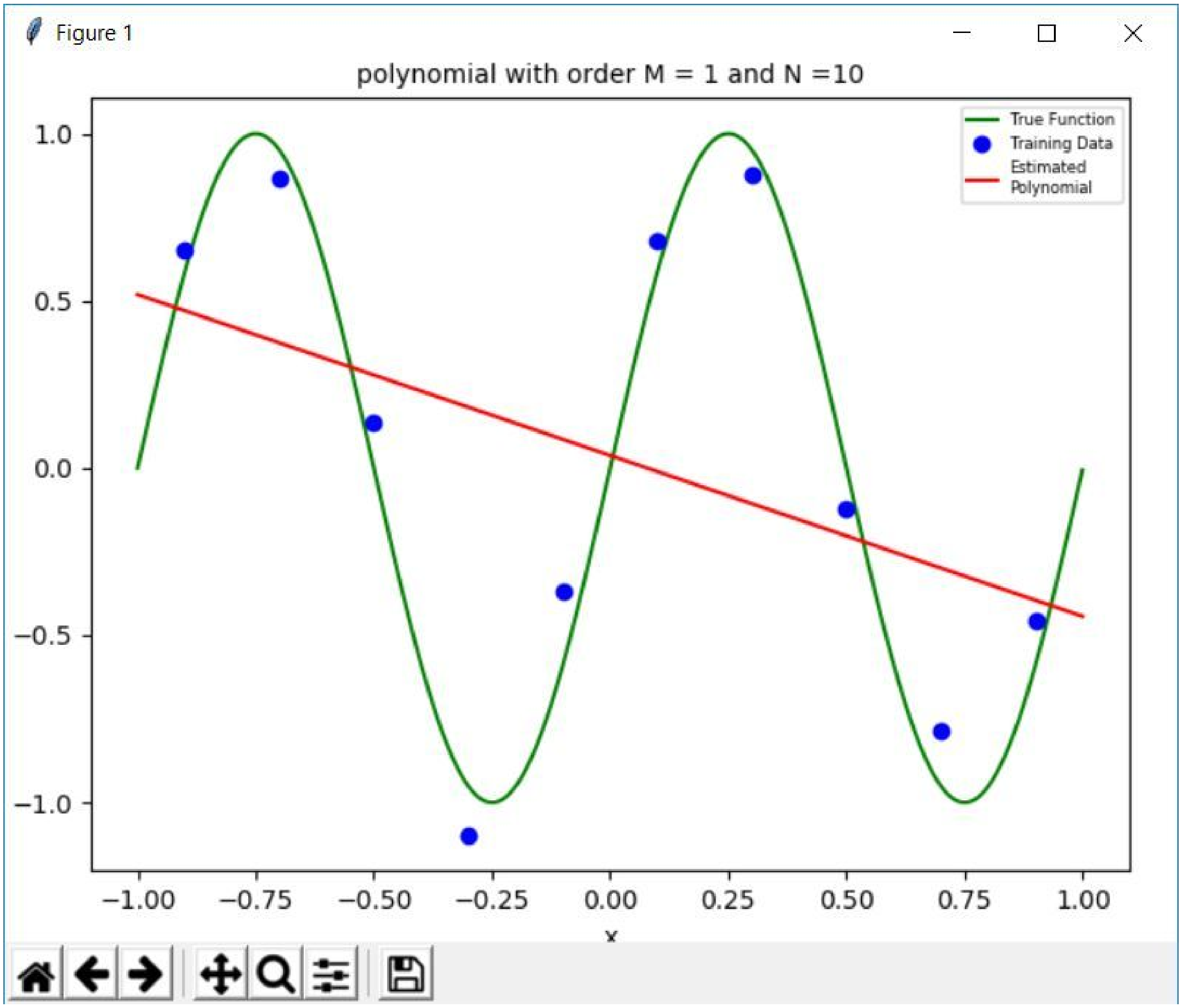

Figure 2: Polynomial approximation of 1st order better approximates the sudden changes in the sinusoidal nature of the data

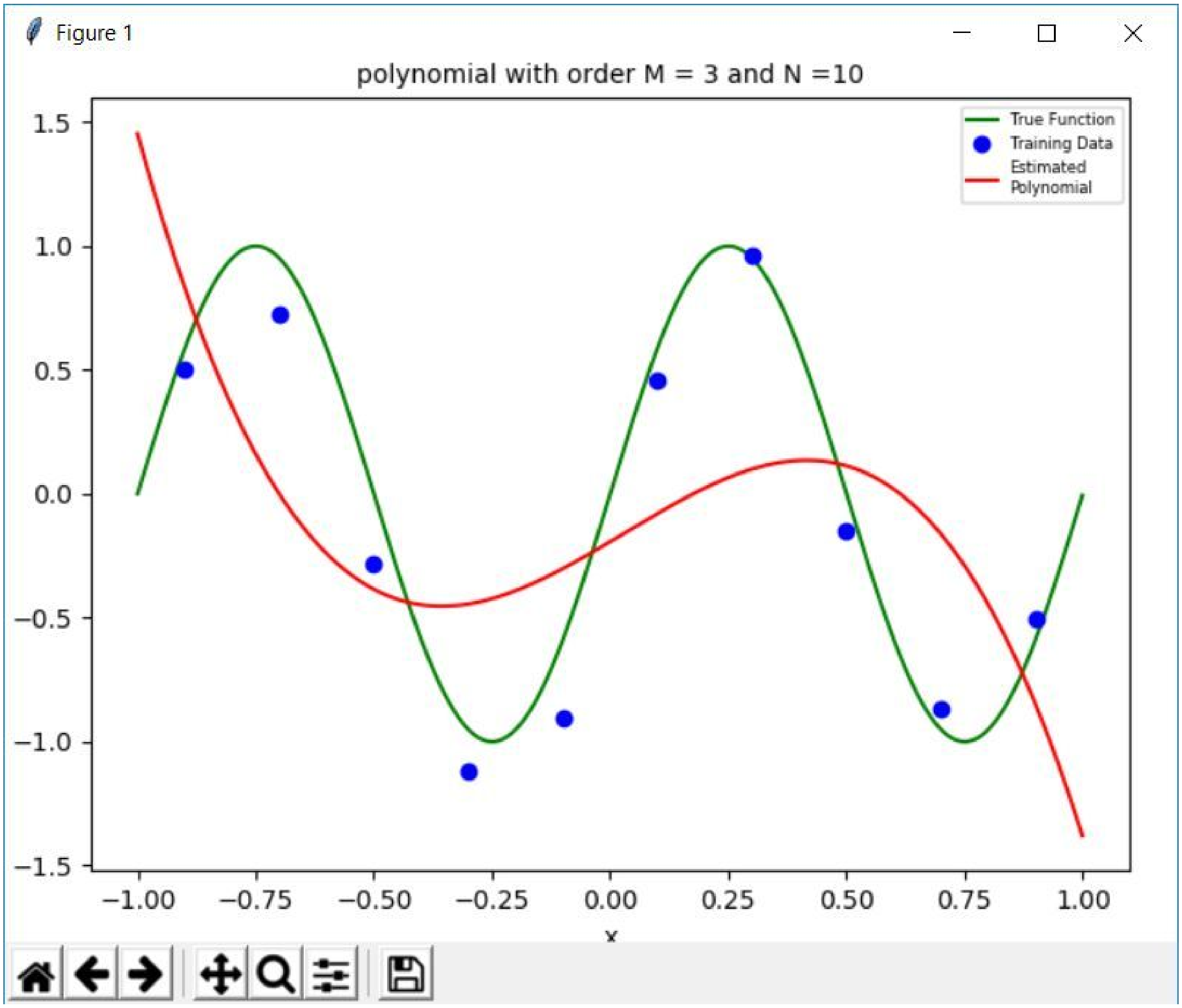

Figure 3: A polynomial of order 3 is more differentiable and can roughly predict the curves of the sin(x) function

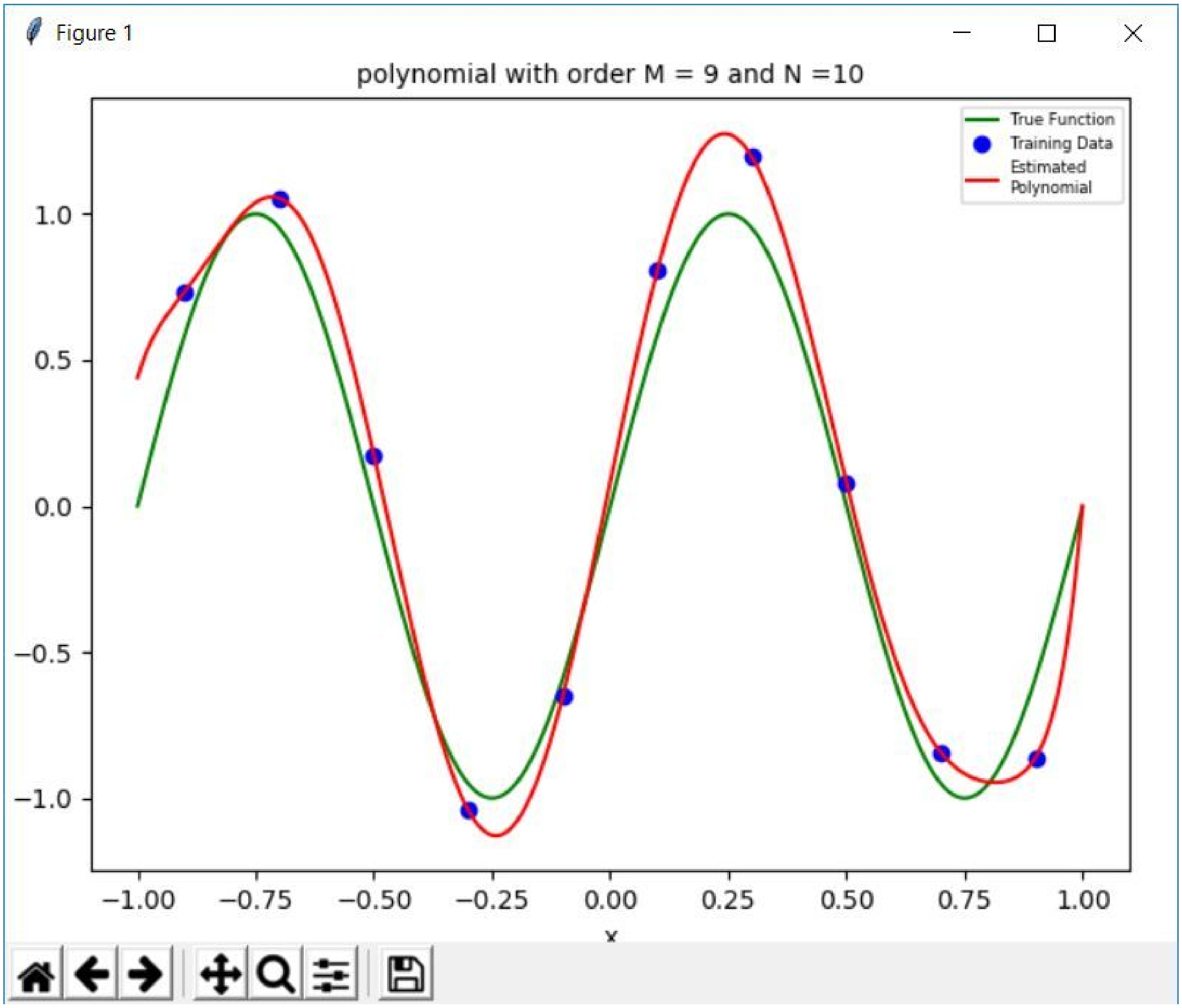

Figure 4: A polynomial of 9th order is fitted almost prefectly because it passes through all of the data points and has essentially “memorized” where the function exists

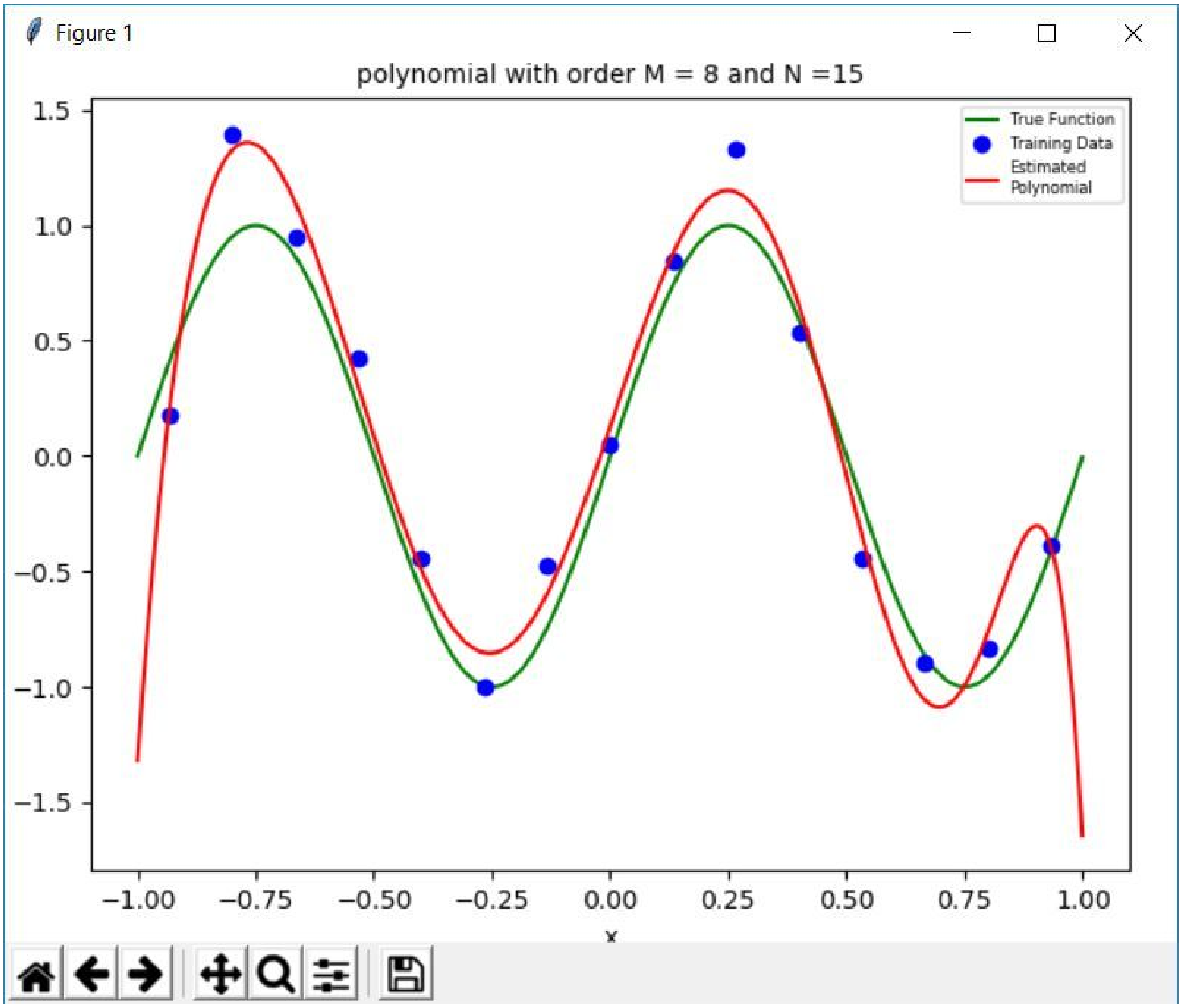

Figure 5: 5 more data points are added and the curve almost fits the data with the exception of a couple outlying points

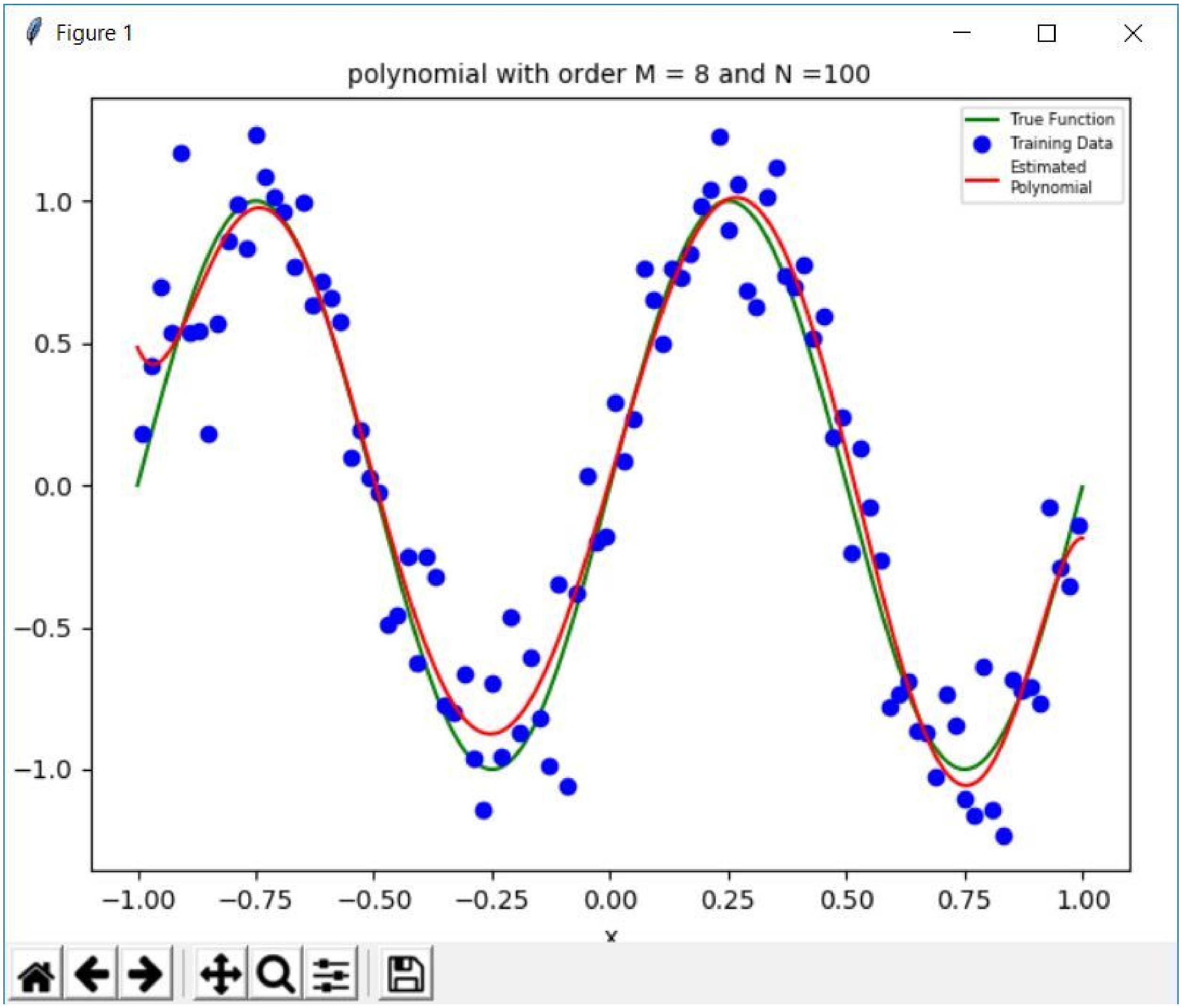

Figure 6: with increased data points the roughness of the curve is minimized because there is less severity to the overfitting phenomenon due to the flexibility of the approximation

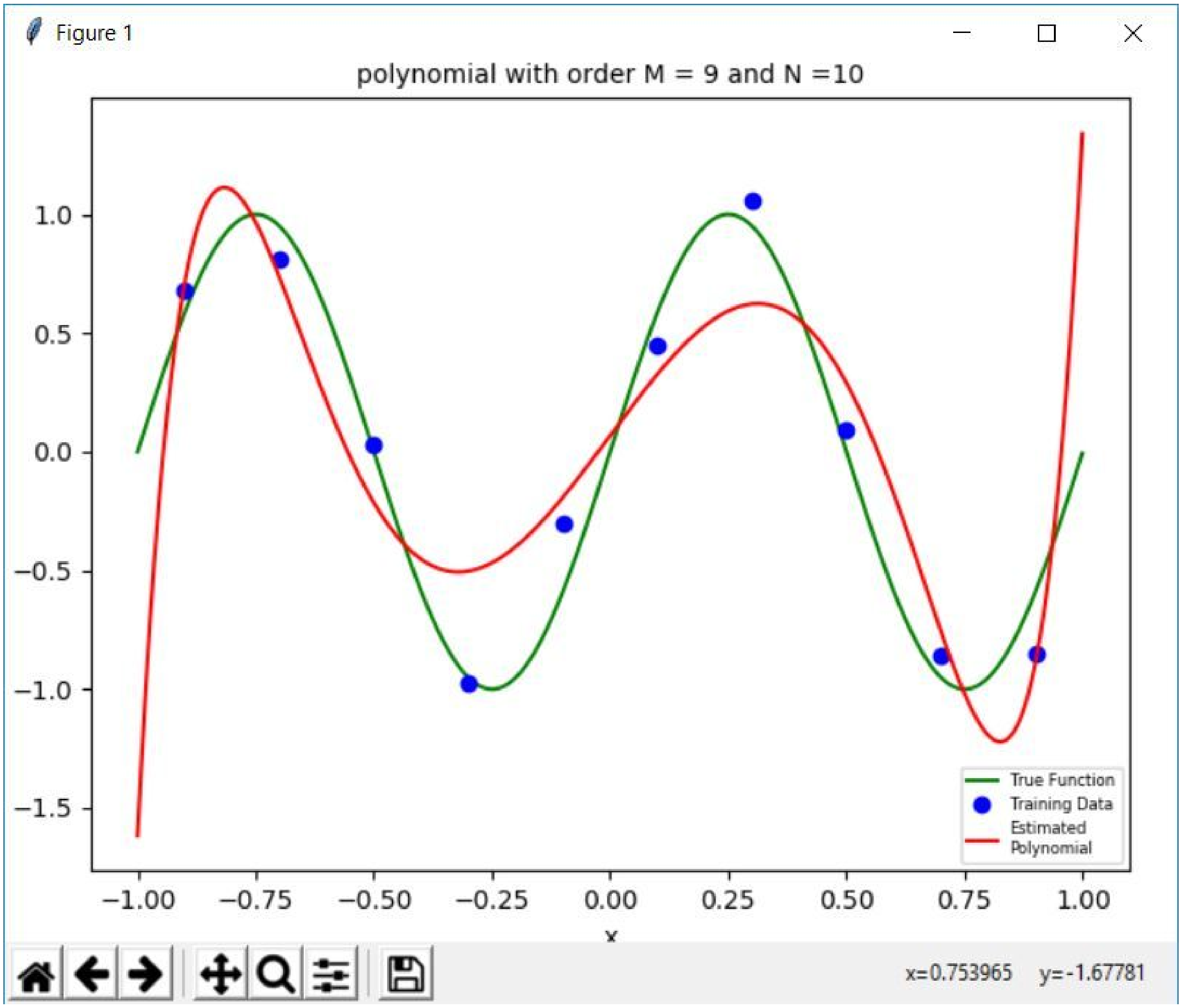

Figure 7: a small lambda value is chosen to regularize the data spread to prevent the model from learning the background noise. In practice the lambda is chosen by tuning to values given by cross-validating the data available

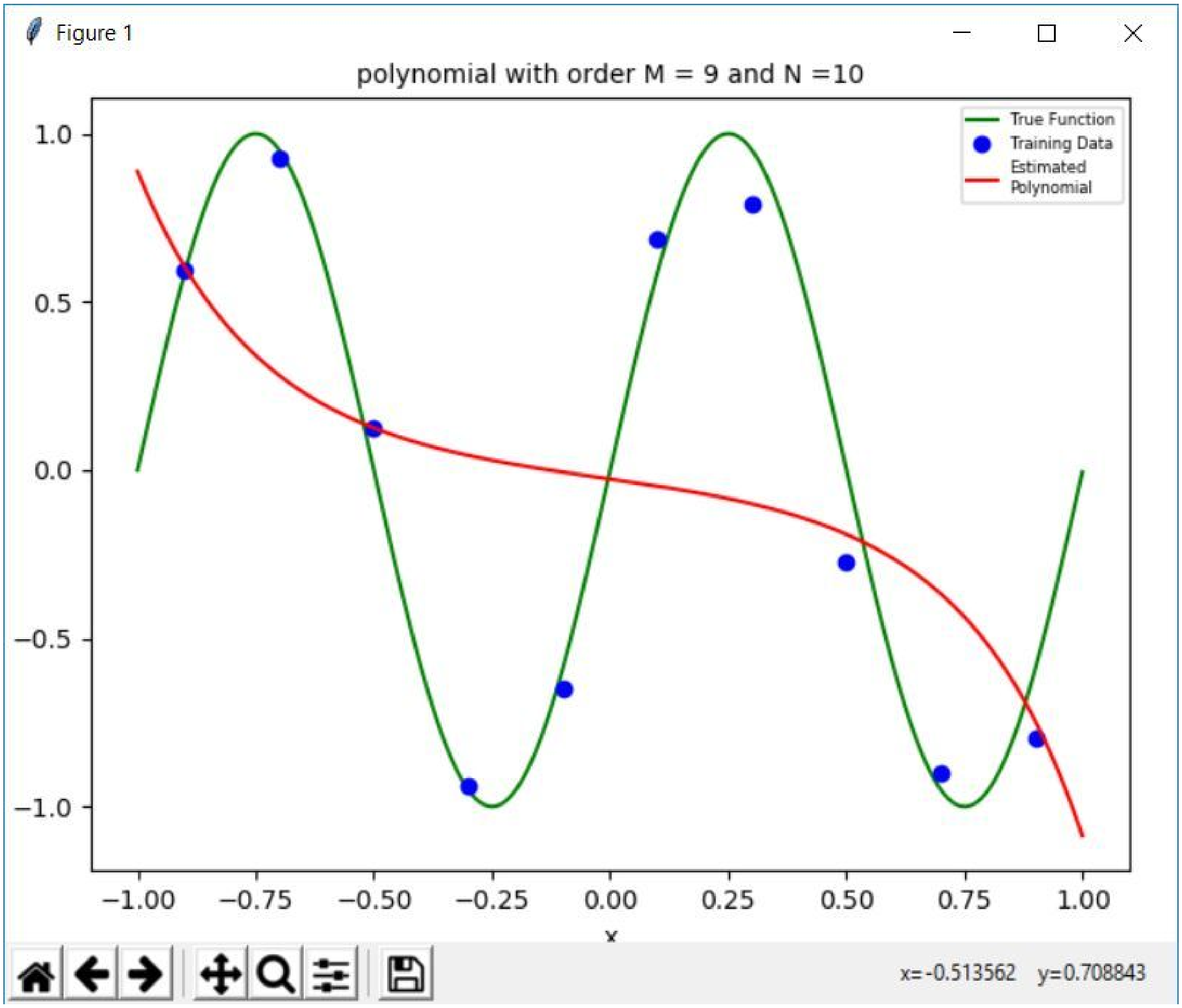

Figure 8: A very large value of lambda = 1 is chosen to exaggerate the contour of the polynomial compared to the previous entry which was closer to the data points

	M = 0	M = 1	M = 6	M = 9
w^*_0	-0.01393686	-0.08355768	1.22748311	-0.13513716
w^*_1		-0.58551107	4.36010494	6.0769382
w^*_2			-0.4.29329626	-2.16614221
w^*_3			-2.14515062	-42.34595292
w^*_4				29.65812765
w^*_5				60.1393377
w^*_6				-75.60046443
w^*_7				11.04340289
w^*_8				52.46490554
w^*_9				-41.88330971

Table 1-1: A table demonstrating the increasing magnitude of values as the order of the Mth polynomial expression increases. This is due to the higher powers in the objective matrix X. These larger values tend to overshoot the data points because they are very large and overfit the problem.