Austin Anderson
Hw4 Analysis

See comments in code for detailed explanation of how the programming tasks were accomplished.

1. Hypothesis Class Complexity

From the results of the rademacher.py: rectangle ≈ 0.9 > origin plane ≈ 0.7 > constant classifier ≈ 0. Using this as a guide I posit that from a theoretical standpoint the highest complexity lies with rectangles such that, rectangles>hyper-plane>origin-centered hyper-planes> constant classifier. This intuitively makes sense as this also represents a linear increase in the degree of freedoms used to specify each classifier (4>3>2>0). From a mathematical standpoint, higher degrees of freedom lead to higher numbers of possible hypothesis sets. From our four-point test data we found that the number of hypotheses were 14-rectangles>6-origin-centered-planes>1-constant. The growth functions of these classifiers are given by (from lecture 6a):

$$\forall m \in \mathbb{N}, \Pi_H(m) \equiv \max_{\{x_1,\ldots,x_m\}\in X} \left|\{(h(x_1),\ldots,h(x_m)) : h \in H\}\right| \qquad (33)$$

Where more hypotheses classes result in higher growth functions as there are more ways to classify the same number of points. The Rademacher complexity can be bounded by the growth function(from lecture 6a):

$$\mathscr{R}_m(G) \leq \sqrt{\frac{2 \ln \Pi_G(m)}{m}} \qquad (34)$$

Where G is a function taking values [-1,+1] (ie the classifiers we have considered. Then as we considered the same number of points for each test, we know the growth functions order the complexity in the way posited above. Then as the Rademacher complexity is bounded by these functions it suggests the complexities of the classes are bounded by the number of ways they can classify the same set of points so our ordering holds.

2. Frequency Classifier Proof

From the lecture linked on Piazza, the frequency prediction was selected to be:

$$h(d_i) = sign(\sin(ad_i)), \quad d_i = 2^{-x_i}, \qquad a = \pi\left(1 + \sum_{i=1}^{M} \frac{(1-y_i)}{2} 2^{x_i}\right)$$

Then we know the phase boundaries for positive and negative classification are:

$$y_i = 1 => 0 + 2\pi n < ad_i < \pi + 2\pi n, \;\; y_i = -1 => \pi + 2\pi n < ad_i < 2\pi + 2\pi n$$

For some integer n. The first portion of the summation term in the formula for *a* returns 0 for a label of 1 and 1 otherwise, so we can simply split these terms:

$$h(2^{-x_i}) = sign\left(\sin\left(\frac{\pi}{2^{x_i}}\right)\right), y_i = 1$$

$$h(2^{-x_i}) = sign\left(\sin\left(\frac{\pi}{2^{x_i}} + \frac{\pi}{2^{x_i}} \sum_{i\ for\ yi<0} 2^{x_i}\right)\right), y_i = -1$$

Addressing the first term, from our phase bounds we know for non-zero, positive, integer values of xi, that phase will ALWAYS be bounded by (0,pi/2) as we're simply dividing pi by a power of 2 (for non-zero, positive powers) so that will always return a positive classifier. Then making a similar argument for the next term, in the summation, the negative label being evaluated contributes a phase of +pi. Other negative labels with greater values of xi will only contribute phase shifts 2pi and can be ignored in the total phase calculation. Negative labels with values of xi less than the xi we're evaluating will contribute a phase bounded by (0,pi/2) as we showed in the previous part. So the summation term will be bounded by (pi,pi+pi/2). Then we showed that the first term in that phase term is bounded by (0,pi/2) from our previous result so the total phase is always bounded by (pi,2pi) so that label will always return negative values!

Austin Anderson
Hw4 Analysis

3. Shattering the Sine Classifier

Now if we're classifying real numbers, some sets cannot be shattered. For example, if the set is given by:

$$x = a, 2a, 3a, 4a$$

For some integer a, this set cannot be shattered for a set of labels:

$$y = true, true, false, true$$

From the bounds above, this would suggest that:

$$0 < mod(wa, 2\pi) < \pi\ (1)$$
$$0 < mod(wa + wa, 2\pi) < \pi\ (2)$$
$$\pi < mod(wa + wa + wa, 2\pi) < 2\pi\ (3)$$
$$0 < mod(wa + wa + wa + wa, 2\pi) < \pi\ (4)$$

Assuming wa = phi+m*2pi we can see this reduces to:

$$0 < \emptyset < \pi\ (1)$$
$$0 < 2\emptyset < \pi \ => \ 0 < \emptyset < \frac{\pi}{2}\ (2)$$

The new equation 1 supports equation 2 so looking at 4:

$$0 < 4\emptyset < \pi\ (4)$$
$$(1), (2) => 0 < \ \emptyset < \frac{\pi}{4}$$

If 1 alone was true, then 4 could be ambiguous, but because 2 is also assumed to be true it suggests this result which is still consistent with 1,2 looking at 3 however:

$$\pi \ < 3\emptyset < 2\pi\ (3)\ INCONSISTENT!$$
$$(4)\ \emptyset = \frac{\pi}{4} - \ \varepsilon, \qquad 3\emptyset < \frac{3\pi}{4}\ \ (5)$$

So (4) suggests (5), but this is inconsistent with (3) so given this set and a sin classifier, we cannot shatter it given the 4 points provided.