

---

# Value Function Approximation using the Fourier Basis

---

George Konidaris  
Sarah Osentoski

GDK@CS.UMASS.EDU  
SOSENTOS@CS.UMASS.EDU

Autonomous Learning Laboratory, Computer Science Dept., University of Massachusetts Amherst, 01003 USA

## 1. Introduction

Reinforcement learning researchers have employed a wide variety of linear function approximation schemes for continuous domains, most commonly Radial Basis Functions (RBFs) and CMACs (Sutton & Barto, 1998). For most problems, choosing the right basis function set is critical for successful learning. Unfortunately, most approaches require significant design effort or problem insight, and no basis function set is both sufficiently simple and sufficiently reliable to be generally satisfactory. Recent work (Mahadevan, 2005) has focused on learning basis function sets from experience, removing the need to design a function approximator but introducing the need to gather data to create one before learning can take place.

In the applied sciences the most common continuous function approximation method is the Fourier Series. Although the Fourier Series is a simple and effective function approximator with solid theoretical underpinnings, it is almost never used in for value function approximation.

We describe the Fourier Basis, a simple fixed linear function approximation scheme using the terms of the Fourier Series as basis functions. We present empirical results showing that it performs well compared to RBFs and the Polynomial Basis, and is competitive with Proto-Value Functions.

## 2. The Fourier Basis

We define the univariate  $k$ th order Fourier Basis as:

$$\phi_i(x) = \cos(i\pi x), \quad (1)$$

for  $i = 0, \dots, k$ , and the  $k$ th order multivariate Fourier Basis for  $m$  variables as:

$$\phi_i(\mathbf{x}) = \cos(\pi \mathbf{c}^i \cdot \mathbf{x}), \quad (2)$$

where  $\mathbf{c}^i = [c_1, \dots, c_m]$ ,  $c_j \in [0, \dots, k]$ ,  $0 \leq j \leq m$  is a coefficient vector which assigns integer weight  $c_j$  for each input variable  $x_j$ . The full basis includes terms for all such weight vectors (resulting in  $(k+1)^m$  terms for  $m$  input variables), although prior knowledge of the domain can be used to rule some out.

Although the Fourier Basis seems like a natural choice for value function approximation, we know of very few instances (e.g., (Kolter & Ng, 2007)) of its use in reinforcement learning, and no empirical study of its properties.

## 3. Empirical Evaluation

### 3.1. The Swing-Up Acrobot

In the swing-up Acrobot (Sutton & Barto, 1998), we employed Sarsa( $\lambda$ ) ( $\gamma = 1.0$ ,  $\lambda = 0.9$ ,  $\epsilon = 0$ ) for Fourier Bases of orders 3 (resulting in 256 basis functions), 5 (resulting in 1296 basis functions) and 7 (resulting in 4096 basis functions) and RBF, Polynomial and (normalized Laplacian) PVF bases of equivalent sizes. (We did not run PVFs with 4096 basis functions because the nearest-neighbour calculations for a graph that size proved too expensive) We systematically varied  $\alpha$  (the gradient descent term) to obtain the best performance for each basis function type and order combination.

The results (Figure 1) demonstrate that the Fourier Basis learners outperformed all other types of learning agents for all sizes of function approximators. In particular, the Fourier Basis performs better initially than the Polynomial Basis (which generalizes broadly) and converges to a better policy than the RBF Basis (which generalizes locally). It also performs slightly better than PVFs, even though the Fourier Basis is a fixed, rather than learned, basis.

### 3.2. The Discontinuous Room

In the Discontinuous Room, show in Figure 2, an agent in a room must reach a target; however the direct path to the target is blocked by a wall, which the agent must go around. This domain is intended as a simple continuous domain with a large discontinuity, to empirically determine how the Fourier Basis handles such a discontinuity compared to Proto Value Functions, which were specifically designed to handle discontinuities.

In order to make a reliable comparison, agents using PVFs were given a perfect adjacency matrix containing every legal state and transition, thus avoiding sampling issues and

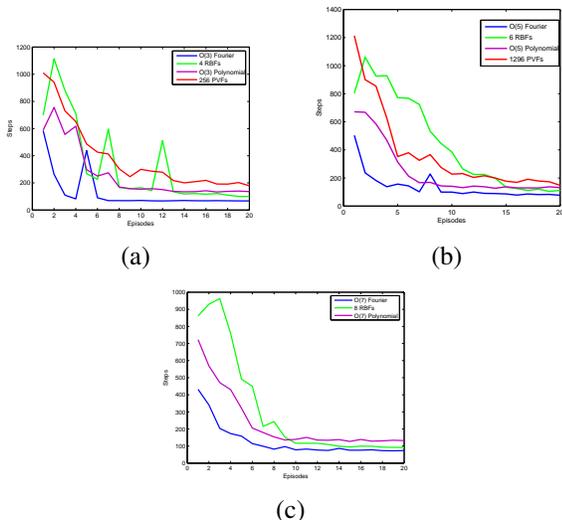


Figure 1. Learning curves for agents using (a) order 3 (b) order 5 and (c) order 7 Fourier Bases, and RBFs and PVFs with corresponding number of basis functions.

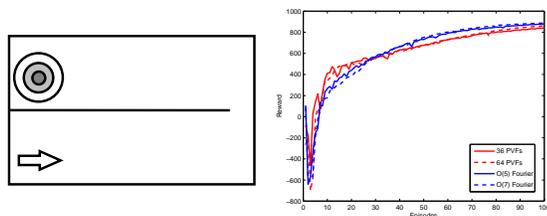


Figure 2. The Discontinuous Room and its learning curves for agents using  $O(5)$  and  $O(7)$  Fourier Bases and agents using 36 and 64 PVFs.

removing the need for an out-of-sample extension method. We used  $Sarsa(\lambda)$  ( $\gamma = 0.9$ ,  $\lambda = 0.9$ ,  $\epsilon = 0$ ) using Fourier Bases of order 5 and 7 and agents using 36 and 64 PVFs. The agents using the Fourier Basis do not initially perform as well as the agents using PVFs, probably because of the discontinuity, but this effect is transient and the Fourier Basis agents eventually perform better.

### 3.3. Mountain Car

For Mountain Car (Sutton & Barto, 1998), we employed  $Sarsa(\lambda)$  ( $\gamma = 1.0$ ,  $\lambda = 0.9$ ,  $\epsilon = 0$ ), Fourier Bases of orders 3 and 5, and RBFs and PVFs of equivalent sizes (we were unable to learn with the Polynomial Basis). Here we found it best to scale the  $\alpha$  values for the Fourier Basis by  $\frac{1}{1+m}$ , where  $m$  was the maximum degree of the basis function. This allocated lower learning rates to higher frequency basis functions.

The results (shown in Figure 3) indicate that for low orders,

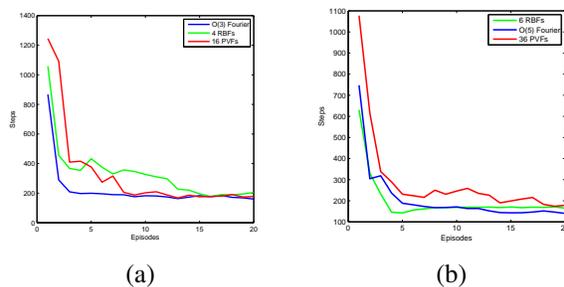


Figure 3. Learning curves for agents using (a) order 3 (b) order 5 Fourier Bases, and RBFs and PVFs with corresponding number of basis functions.

the Fourier Basis outperforms RBFs and PVFs. For higher orders, we see a repetition of the learning curve in Figure 2, where the Fourier Basis initially performs worse (because it does not model the discontinuity in the value function well) but converges to a better solution.

## 4. Summary

Our results show that the Fourier Basis provides a simple and reliable basis for value function approximation. We expect that for many smaller problems, the Fourier Basis will be sufficient to learn a good value function without any extra work.

Our experiences with the Fourier Basis show that it is a simple and easy to use basis set that reliably performs well on a range of different problems. As such, although it may not perform well for domains with several significant discontinuities, its simplicity and reliability suggest that the Fourier Basis should be the first function approximator used when reinforcement learning is applied to a new continuous problem.

For the full version of this paper, please see Konidaris and Osentoski (2008).

## References

Kolter, J., & Ng, A. (2007). Learning omnidirectional path following using dimensionality reduction. *Proceedings of Robotics: Science and Systems*.

Konidaris, G., & Osentoski, S. (2008). *Value function approximation in reinforcement learning using the fourier basis* (Technical Report UM-CS-2008-19). Department of Computer Science, University of Massachusetts, Amherst.

Mahadevan, S. (2005). Proto-value functions: Developmental reinforcement learning. *Proceedings of the Twenty Second International Conference on Machine Learning*.

Sutton, R., & Barto, A. (1998). *Reinforcement learning: An introduction*. Cambridge, MA: MIT Press.