CHAPTER SIX. Concluding Comments.

In this thesis, we have considered various Markov chains, and in particular their rates of convergence. Many of the results presented suggest possible generalizations and directions for further research.

The cut-off phenomenon presented in Chapter 2 is strikingly similar to those found in certain cases on finite groups, such as the Random Transpositions of Diaconis and Shashahani. This leads to certain obvious questions, such as, how generally does this cut-off phenomenon occur? What are the essential properties of the underlying random walk which cause it? Can a more general proof be constructed, that applies to a wider variety of random walks on different groups? (It appears that a deeper understanding of group representations may be required to achieve this.) Does a similar phenomenon occur on Markov chains that are not random walks on groups? (Here one could not use Fourier Analysis at all. Still, there are some precedents. For example, the "top in at random" shuffle of Aldous and Diaconis can be shown to have a cut-off phenomenon using a coupling argument.)

The main result of Chapter 3 says that a *constant* number of iterations are required to approach stationarity on the circle groups, provided the step distributions are all defined by a fixed global function f. Does this generalize to other, more complicated families of groups? A result of Alan Turing says that the circle groups (and their direct products) are the only groups which "fit" into a compact Lie group in such a nice way. Thus, the construction may have to be modified. But one can still ask, are there "natural" growing families of groups, and "natural" measures defined on them, such that a constant number of iterations suffices to approach uniform? The symmetric groups are an obvious place to start.

Chapters 4 and 5 concern Markov chains which are not random walks on groups. In both chapters, essentially the same Lemma (related to Harris Recurrence) is used to bound the distance to stationarity. But the Lemma is easily seen to be somewhat clumsy, and to throw away a lot of information. Can the Lemma be sharpened in a simple way? Can it be applied to the Gibbs Sampler with models other than those considered here, for example to image processing models? (As mentioned in Chapter 5, this appears to be fairly hopeful, at least for models with lots of data.) Can it be applied to other stochastic algorithms (of which there are many, in Statistics, Physics, and Computer Science)? Does it "always" work in the sense that for any Markov chain which does converge to a unique stationary distribution, the Lemma can at least be used to show eventual convergence even if it can't always get the rate right? Are there other useful approaches to convergence rates for large, complicated Markov chains with no group structure?

Throughout this thesis, total variation distance has been used to measure distance to stationarity. But there are many other natural topologies, including convergence on intervals and on compact subsets. Also, the L^2 metric is often used in place of the L^1 or total variation metric. It is known that with the L^2 metric, the cut-off phenomenon does not occur. But what properties are preserved? What new properties arise? And which metric is the "best" one to use?

Finally, rates of convergence are of course only one area of research related to Markov chains. Other natural areas include hitting times, corrolation between successive iterations, fluctuation theory, large deviations, and so on, and there are many interesting unanswered questions there as well. Markov chain theory is a very active area of research. And, with its many connections to applications and to other areas of mathematics, it promises to continue to be active for a long time to come.