A Markov chain for Steiner triple systems

Peter J. Cameron

CSG notes, February 2002

Abstract

These notes describe what might be a Markov chain method for choosing a random Steiner triple system. Many things are not known, including whether or not the Markov chain is connected! I include a positive result of Grannell and Griggs according to which any two isomorphic Steiner triple systems lie in the same connected component.

1 Choosing at random

Suppose I have a fair coin. How can I choose a random Steiner triple system on 103 points?

A fair coin is a device which can in one step (or toss) produce one bit of information (a 0 or a 1, or informally, “heads” or “tails”), in such a way that the results of different tosses are independent – this means that each of the 2^n sequences of results produced by n tosses occurs with the same probability, namely $1/2^n$.

Given a fair coin, there is a simple algorithm for choosing a random integer x in the range $[0,2^n-1]$. We just toss the coin n times and interpret the sequence of bits as the expansion of x in base 2.

What about choosing an integer in the range $[0,N-1]$, where N is arbitrary? We cannot do this with a bounded number of coin tosses if N is not a power of 2, since the probability of any event defined by n coin tosses is a rational number with denominator 2^n. So we have to make a small compromise, as follows. Choose n to be the least integer such that $2^n \geq N$. Choose an integer in the range $[0,2^n-1]$ as before. If it is smaller than N, we accept this result; otherwise we try agin, and continue until a result is obtained. It is not hard to show that, if $p = N/2^n$, then

- the resulting integer is uniformly distributed in the range $[0,N-1]$;

- the expected number of attempts is $1/p$;
- the probability that more than m attempts are required is q^m, where $q = 1-p$.

(The last two statements follow because the number of attempts is a geometric random variable). Since $p > 1/2$, the expected number of attempts is less than 2 and the probability of needing a long series of tries is exponentially small.

Now we can choose a random structure of a certain type in some situations. If we can count the structures, then we may suppose there are N altogether; choose a random number x from $[0, N-1]$, skip over the first x structures in the count, and take the next one. If each structure is determined by a sequence of choices, and making these choices uniformly gives the uniform distribution, then we can make the choices at random as long as we know how many choices there are at each stage.

For example, how can we choose a random permutation σ of the set $\{1, \ldots, n\}$? The image 1σ of 1 can be any of $1, \ldots, n$; choose this image at random. Then 2σ can be any of $1, \ldots, n$ except 1σ; choose a random number x from $1, \ldots, n-1$ and add one if $x \geq 1\sigma$, then set $2\sigma = x$. Continuing in this way gives the required random permutation.

Choosing a random graph on n vertices is even easier: simply decide with a single coin toss whether each pair of vertices is joined by an edge or not. (Indeed, the number of graphs is $2^{n(n-1)/2}$, and counting them is equivalent to choosing one at random.)

In other cases, e.g. Latin squares, Steiner systems, we don't even know how many structures there are, so choosing one at random cannot be done by these methods; we need a new idea.

2 Markov chains and random walks

We consider only Markov chains with a finite number of states. Let $S = \{s_1, \ldots, s_m\}$ be a finite set of *states*. Suppose we are given a matrix $P = (p_{ij})$ of order m, whose entries are non-negative real numbers satisfying

$$\sum_{j=1}^{m} p_{ij} = 1.$$

The interpretation is that we have a marker on one of the states; at a certain moment it moves to a new state, where the probability of moving from state s_i to state

s_j is p_{ij}. The displayed equation just reflects the fact that the marker is bound to move to some state!

We can now iterate this procedure. The marker starts out on some state, possibly chosen at random from an arbitrary probability distribution. At each postive integer time it makes a transition according to the specification above. We are interested in how it behaves in the long term. There are two extremes of behaviour:

- Suppose that $p_{i\,i+1} = 1$ for $i = 1, \ldots, m-1$ and $p_{m1} = 1$, all other probabilities being zero. Then the marker simply marches around the m-cycle in a mechanical way; if it starts at state s_i then after n steps it is certainly at state s_j, where $j \equiv i + n \pmod{m}$.

- Suppose that $p_{ij} = 1/m$ for all i, j. Then, no matter where the marker starts, after one transition it is in a random state chosen uniformly from $s_1, \ldots, s_m$, and this remains true after any number of transitions.

For most interesting chains, we don't have either of these extremes, but instead, under certain hypotheses the marker's position approaches a limiting distribution as the number of transitions increases.

The displayed equation above can be rewritten as $Pj^\top = j^\top$, where j is the all-one vector, $j = (1, 1, \ldots, 1)$. So 1 is a right eigenvalue of P. Since the left and right eigenvalues of a matrix are the same, there is a vector $q = (q_1, \ldots, q_m)$ such that $qP = q$.

It can be proved that we can choose q to have all its entries non-negative, so we can normalise the entries so that $\sum_{i=1}^m q_i = 1$. Then we can interpret q as a probability distribution on the states. Suppose that the marker starts in this distribution. Then after one transition, its probability of being in state s_j is

$$\sum_{i=1}^m q_i p_{ij} = q_j,$$

that is, the same as before the transition! So if the marker starts in the distribution q, then it remains in this distribution. So q is certainly a candidate for a limiting distribution.

We need a couple of conditions on the chain to guarantee good limiting behaviour and rule out cases like the first example above. Let p_{ij}^n be the probability of moving from state s_i to state s_j after n transititions. (Exercise: this is just the (i, j) entry of the matrix P^n.) The chain is said to be *irreducible* if, for any two states s_i and s_j, there exists n such that $p_{ij}^n > 0$, that is, it is possible to move from

s_i to s_j. The chain is said to be *aperiodic* if, for any state s_i, the greatest common divisor of the set

$$\{n : p_{ii}^n > 0\}$$

is equal to 1.

Theorem 1 *Let P be an irreducible and aperiodic Markov chain, and q the normalised left eigenvector of P with eigenvalue* 1*. Then, starting from an arbitrary initial distribution, the distribution after n steps approaches q as $n \to \infty$.*

The particular type of Markov chain we consider is the random walk on an undirected graph. The states are the vertices of the graph, and a transition consists of choosing an edge through the vertex on which the marker sits (all edges being equally likely) and moving to the other end of this edge. In other words, p_{ij} is the reciprocal of the valency of the ith vertex v_i if v_i and v_j are adjacent, and is zero if they are non-adjacent. It is not hard to see that the random walk is irreducible if and only if the graph is connected, and is aperiodic if and only if the graph is not bipartite. (Since the graph is undirected, we can always return to the start vertex after 2 steps with non-zero probability.)

With our fair coin we can do a random walk on a graph, since we have to choose among a number of edges at each step, giving each edge the same probability.

It is also simple to compute the limiting state. We claim that the vector whose ith component is the valency of v_i is a left eigenvector with eigenvalue 1. This is an easy exercise; here is a heuristic argument. If the probability of starting at v_i is ck_i, where k_i is the valency of v_i and c is a constant, then the probability of passing along any given edge is c, and so the probability of arriving at v_j is ck_j.

In other words, if a graph is connected and non-bipartite, then the random walk on that graph has the property that, in the limit, the probability of being at any vertex is proportional to its valency. In particular, if the graph is regular, then the limiting distribution is uniform.

3 Steiner triple systems

A *Steiner triple system* consists of a set S of n *points* and a set T of *triples* or 3-subsets of S with the property that any two points of S are contained in a unique triple. It is well-known that a Steiner triple system on n points exists if and only if $n = 0$ or $n \equiv 1$ or $3 \pmod 6$.

Steiner triple systems on at most 3 points are trivial. On a set of 7 points, there are precisely 30 different Steiner triple systems, all of which are isomorphic. (An *isomorphism* of Steiner triple systems is a bijection of their point sets which carries triples to triples.) Similarly on 9 points there are 840 Steiner triple systems, all isomorphic. In these cases, it is simple to choose a random Steiner triple system: just start with any Steiner triple system and apply a random permutation to it.

After that, life gets more complicated. There are two non-isomorphic Steiner triple systems on 13 points, and 80 (up to isomorphism) on 15 points. (The total numbers of systems are 1 197 504 000 and 60 281 712 691 200 respectively.) For the next value, $n = 19$, Petteri Kaski and Patric R. J. Östergård [6] have very recently shown that the number of isomorphism classes is 11 084 874 829, while the total number of systems is 1 348 410 350 618 155 344 199 680 000.

Asymptotically, Richard Wilson [7] has shown that the number of isomorphism classes of Steiner triple systems on n points is between $(\mathrm{e}^{-5}n)^{n^2/6}$ and $n^{n^2/6}$. Lászlo Babai [1] showed that almost all of these have trivial automorphism group, so the total number of Steiner triple systems is obtained approximately by multiplying this number by $n!$; but $n!$ is negligible even compared to $c^{n^2/6}$.

In the absence of an exact count, we could attempt to choose a random system by a Markov chain method. The idea is to start with any Steiner triple system and make some random modification of it to produce another system. Unfortunately there isn't an obvious way to do this, so we have to enlarge the space in which we work by including some so-called *improper* Steiner triple systems, as follows.

First, we re-define Steiner triple systems slightly. Instead of taking a set T of triples or 3-subsets of S, we take a function f from the set of 3-subsets of S to $\{0,1\}$ (the characteristic function of T); it has the property that, for any distinct $x,y \in S$, we have

$$\sum_{z\in S} f(\{x,y,z\}) = 1.$$

Now we define an *improper* Steiner triple system on S to be a function f from the set of 3-subsets of S to $\{-1,0,1\}$ satisfying the two conditions:

- there is a unique 3-set $\{x,y,z\}$ with $f(\{x,y,z\}) = -1$;
- $\sum_{z\in S} f(\{x,y,z\}) = 1.$

We call a Steiner triple system *proper* where necessary to avoid confusion.

Now the state space of the Markov chain is defined to be $P \cup I$, where P and I are the sets of proper and improper Steiner triple systems on S, respectively. A transition works as follows. Let f be a state.

(a) If f is proper, choose $\{x,y,z\}$ with $f(\{x,y,z\}) = 0$; if f is improper, start with the unique $\{x,y,z\}$ such that $f(\{x,y,z\}) = -1$.

(b) Let x',y',z' be points such that

$$f(\{x',y,z\}) = f(\{x,y',z\}) = f(\{x,y,z'\}) = 1.$$

(If f is proper, these points are unique; if f is improper, there are two choices for each of them. The points x',y',z' are distinct.)

(c) Now increase the value of f by 1 on $\{x,y,z\}$, $\{x,y',z'\}$, $\{x',y,z'\}$, and $\{x',y',z\}$, and decrease it by 1 on $\{x',y,z\}$, $\{x,y',z\}$, $\{x,y,z'\}$, and $\{x',y',z'\}$. We obtain another proper or improper Steiner triple system, according as in the original system we have $f(\{x',y',z'\}) = 1$ or $f(\{x',y',z'\}) = 0$.

All choices are to be made uniformly.

This definition is a simple modification of one proposed for Latin squares by Jacobson and Matthews [5].

We denote the move above by $\begin{pmatrix} x & y & z \\ x' & y' & z' \end{pmatrix}$. If the bottom row of the symbol is a triple, we obtain a proper STS; otherwise we obtain an improper STS for which the bottom row is the negative block. Note that the number of moves from a proper STS is $\binom{n}{3} - n(n-1)/6 = n(n-1)(n-3)/6$; the number of moves from an improper STS is 8.

The first thing to note is that the moves are reversible: that is, if there is a move from f to f', then there is a move from f' to f. Also, our stipulation of uniformity means that all states which can be reached in one move from f are equally likely. So the Markov chain is just the random walk on the graph G with vertex set $P \cup I$ whose edges are the transitions defined above. We call this the *transition graph* for Steiner triple systems.

In this graph, as we have noted, any proper system has valency $n(n-1)(n-3)/6$, whereas any improper system has valency 8. So we conclude:

> If the graph G is connected and not bipartite, then the unique limiting distribution of the Markov chain has the property that all proper Steiner triple systems are equally likely.

This means that if we run the Markov chain for sufficiently long and then stop once we reach a proper Steiner triple system, it will be approximately uniformly distributed, and the approximation will be better the more steps we take.

So the crucial question is:

Question Is the above graph G connected and non-bipartite?

In fact, connectedness is not essential: all we require is that the proper STSs are all in the same connected component.

In the next section, we outline a result of Grannell and Griggs [4] which shows that any two isomorphic Steiner triple systems lie in the same connected component of the graph G.

If connectedness can be proved, then other questions arise, such as:

- What is the diameter of the transition graph?
- What is the maximum value of the minimum distance of an improper STS from all proper STSs?
- How fast does the probability distribution converge to its limit?

4 Small values of n

For motivation, we look at the case $n = 7$. In this (atypical) case, there are no improper Steiner triple systems. For suppose that $f(\{1,2,3\}) = -1$. Then there are two triples (with $f = +1$) containing $\{1,2\}$, say $\{1,2,4\}$ and $\{1,2,5\}$. Also there are two triples containing $\{1,3\}$; these can have no further points in common with the triples containing $\{1,2\}$, so must be $\{1,3,6\}$ and $\{1,3,7\}$. Now there is no way to choose the triples containing $\{2,3\}$.

In particular, a move in the Markov chain must take a proper system to another proper system. If we begin with the system (written in the obvious short form) $\{123,145,167,246,257,347,356\}$, and take $(x,y,z) = (1,2,4)$, then we make the move $\begin{pmatrix} 1 & 2 & 4 \\ 6 & 5 & 3 \end{pmatrix}$, and we obtain $\{124,135,236,456,167,257,347\}$, which is the image of the first system under the transposition $(3\,4)$. By a sequence of moves, we can apply any permutation of $\{1,\ldots,7\}$ to the points of the original system. Since all Steiner triple systems of order 7 are isomorphic, we have shown that the graph is connected. Also, since the automorphism group of a system contains only even permutations, we see that the graph is bipartite, since every step changes the parity of the permutation applied.

The case $n = 9$ is a bit more typical. In this case, there is a unique improper STS up to isomorphism. For suppose that 123 is the negative triple. Then, without loss, we have positive triples $124, 125, 136, 137, 238, 239$. Each of $1, 2, 3$ lies in just one further triple; these must be $189, 267, 345$. The remaining triples are transversals to $45, 67, 89$, for which there are just two (isomorphic) possibilities, one of which is $468, 479, 569, 578$.

Now we examine the eight possible moves from this improper STS. For the four moves like $\begin{pmatrix} 1 & 2 & 3 \\ 8 & 6 & 4 \end{pmatrix}$, where the bottom row is a (positive) triple, we move to a proper STS (in this case it would have triples $125, 137, 239, 189, 267, 345, 479, 569, 578, 146, 248, 368$). For the four moves like $\begin{pmatrix} 1 & 2 & 3 \\ 9 & 6 & 4 \end{pmatrix}$, where the bottom row is not a triple, we obtain another improper system, in which (in this case) 469 is the negative block.

Hence we can move from a proper STS to another proper STS in three moves.

We will see in the next section that we can move from any proper STS to any isomorphic system in an even number of moves. This shows that, for $n = 9$, the graph is non-bipartite. From this we deduce:

Theorem 2 *The transition graph is non-bipartite for all admissible $n \geq 19$.*

Proof The Doyen–Wilson theorem [3] asserts that, for all admissible m, n with $n \geq 2m + 1$, there is a STS of order n with a subsystem of order m. So, for all admissible $n \geq 19$, there is a STS of order n containing a subsystem of order 9. Now there is a sequence of moves of odd length which starts and ends at this subsystem; since the moves only involve points within the subsystem, the rest of the STS is unaltered. ■

In fact the graph is non-bipartite for all $n > 7$. The remaining cases ($n = 13$ and $n = 15$) were settled by Matt Ollis, as we will describe at the end of the next section.

Connectedness is more difficult. We have seen that the graph is connected for $n = 7$; its connectedness for $n = 9$ is easily shown by similar arguments, or follows from the next section. For larger values, the question is open.

5 Applying a permutation

Grannell and Griggs [4] proved the following result.

Theorem 3 *Any two isomorphic proper Steiner triple systems lie in the same connected component of the graph defined earlier.*

Proof Starting with any fixed STS of order n, we obtain all systems isomorphic to it by applying all permutations of the symmetric group S_n to its points. Since S_n is generated by transpositions, it is enough to show that the result of applying an arbitrary transposition lies in the same connected component.

Let a and b be any two points of the (proper) STS S. There is a unique triple $\{a,b,c\}$ containing them. Consider the graph whose vertices are the points different from a,b,c, in which x and y are joined by a red edge if $\{a,x,y\}$ is a triple, a blue edge if $\{b,x,y\}$ is a triple, and by no edge otherwise. Any vertex lies on one red and one blue edge, so that the graph is the disjoint union of cycles of even length.

Suppose that $(0,1,\ldots,2m-1)$ is a cycle, with the edge 01 red. Thus, we have triples

$$\{a,0,1\},\{b,1,2\},\{a,2,3\},\ldots,\{a,2m-2,2m-1\},\{b,2m-1,0\}.$$

We will replace these by the triples

$$\{b,0,1\},\{a,1,2\},\{b,2,3\},\ldots,\{b,2m-2,2m-1\},\{a,2m-1,0\}.$$

in $m-1$ moves.

The first move is $\begin{pmatrix} a & 1 & 2 \\ b & 3 & 0 \end{pmatrix}$. This changes the first three triples on the list to $\{b,0,1\}$, $\{a,1,2\}$ and $\{b,2,3\}$, and also introduces a positive triple $\{a,0,3\}$ and a negative triple $\{b,0,3\}$.

The second move is $\begin{pmatrix} a & 3 & 4 \\ b & 5 & 0 \end{pmatrix}$, which interchanges a and b in the next two triples in the list and chages 3 to 5 in the positive and negative pair.

It is readily checked that on the $m-1$st move, we have achieved our aim. The last move is $\begin{pmatrix} a & 2m-3 & 2m-2 \\ b & 2m-1 & 0 \end{pmatrix}$. Instead of introducing a positive triple $\{a,0,2m-1\}$ and a negative triple $\{b,0,2m-1\}$, we cancel the existing positive triple $\{b,0,2m-1\}$ and replace it with $\{a,0,2m-1\}$.

Repeating this procedure for each cycle of the graph, we end up with a system having a and b interchanged. ■

The number of steps required to apply the transposition (a,b) is equal to $(n-3)/2-c$, where c is the number of cycles in the above graph. In the case $n=9$, each graph has a single cycle, so two moves are required for each transposition. This justifies the claim at the end of the last section.

Matt Ollis used this observation to show that the transition graph is not bipartite in the remaining cases $n=13$ and $n=15$, as follows.

Consider first $n=13$. There are two STS of order 13; we take the more symmetric one, whose points are the integers mod 13 and whose blocks are translates of $\{0,1,4\}$ and $\{0,2,8\}$. The graph associated with $\{0,1\}$ is a 10-cycle

$$2,8,6,11,10,9,3,12,7,5,2,$$

so the transposition $(0,1)$ is achieved in $5-1=4$ moves. The transposition $(1,2)$ also requires four moves. Similarly, the graph associated with $(0,2)$ has a 4-cycle and a 6-cycle, so the transposition $(1,2)$ is achieved in $5-2=3$ moves. Now $(1,2)(0,1)(0,2)(0,1)$ is the identity permutation, so we return to the original system in $4+4+3+4=15$ moves.

For $n=15$, a similar analysis can be applied to system number 44 in the list in the CRC Handbook [2].

References

[1] L. Babai, Almost all Steiner triple systems are asymmetric, in *Topics in Steiner systems* (ed. C. C. Lindner and A. Rosa), *Ann. Discrete Math.* **7**, Elsevier, Amsterdam, 1979, pp. 37–39.

[2] C. J. Colbourn and J. H. Dinitz (eds.), *The CRC handbook of combinatorial designs*, CRC Press, Boca Raton, 1996.

[3] J. Doyen and R. M. Wilson, Embeddings of Steiner triple systems, *Discrete Math.* **5** (1973), 229–239.

[4] M. J. Grannell and T. S. Griggs, Negative trades in Steiner triple systems, working document.

[5] M. T. Jacobson and P. Matthews, Generating uniformly distributed random Latin squares, *J. Combinatorial Design* **4** (1996), 405–437.

[6] P. Kaski and P. R. J. Östergård, The Steiner triple systems of order 19, available from `http://www.tcs.hut.fi/~pkaski/sts19.ps`

[7] R. M. Wilson, Non-isomorphic Steiner triple systems, *Math. Z.* **135** (1974), 303–313.