Journal of Humanistic Mathematics

Volume 6 | Issue 2 July 2016

Combinatorics of the Sonnet

Terry S. Griggs
The Open University

Recommended Citation

Griggs, T. S. "Combinatorics of the Sonnet," *Journal of Humanistic Mathematics*, Volume 6 Issue 2 (July 2016), pages 38-46. DOI: 10.5642/jhummath.201602.05 . Available at: https://scholarship.claremont.edu/jhm/vol6/iss2/5

Combinatorics of the Sonnet

Terry S. Griggs

Department of Mathematics and Statistics, The Open University, Milton Keynes, UK
t.s.griggs@open.ac.uk

Abstract

Using a definition of a sonnet, the number of basic rhyming schemes is enumerated. This is then used to discuss the 86 sonnets which appear in John Clare's *The Rural Muse*.

1. Introduction

The Oxford Companion to English Literature [3] defines a sonnet as "a poem consisting of 14 lines ... with rhymes arranged according to one or other of certain definite schemes". As is well known, the form originated in Italy and was introduced into England by Sir Thomas Wyatt sometime in the 16$^{\text{th}}$ century.

The Companion is not specific concerning the exact form of the rhyming schemes, just giving the two main examples. One of these is the *Shakespearean sonnet* which has the rhyming scheme

$$\mathcal{ABAB} - \mathcal{CDCD} - \mathcal{EFEF} - \mathcal{GG}.$$

The reason for the insertion of hyphens will be explained later, but for readers familiar with the form of a sonnet, it is no doubt already clear. The other example given is the *Petrachean sonnet*, the first eight lines of which have the rhyming scheme

$$\mathcal{ABBA} - \mathcal{ABBA}$$

"followed by two, or three, other rhymes in the remaining six lines" [3]. `Wikipedia` (http://en.wikipedia.org/wiki/Sonnet) is more specific. It lists two variations,

$$\mathcal{CDE} - \mathcal{CDE} \quad \text{and} \quad \mathcal{CDC} - \mathcal{CDC}$$

for the final six lines. However, further schemes are also used; Wyatt in particular employed the rhyming scheme

$$\mathcal{CCD} - \mathcal{CCD}.$$

The scheme

$$\mathcal{CD} - \mathcal{CD} - \mathcal{CD}$$

is another option. To a combinatorial mathematician, an obvious question is to classify what all of these rhyming schemes might be and to enumerate them. This we do in Sections 2-3 and is one of the aims of this short paper. Then we consider, in Section 4, a collection of sonnets written over two hundred years after Shakespeare: the 86 sonnets which appear in John Clare's *The Rural Muse* [1], and discuss their rhyming schemes in relation to our classification.

However, before we attempt the enumeration we need to recall the basic structure of the sonnet. The fourteen lines are divided into two parts; the first eight lines, known as the *octave* and the last six lines, known as the *sestet*. Moreover the octave is further divided into two *quatrains*, each consisting of four lines. The hyphens in the rhyming schemes above reflect this partitioning.

2. Structure of the octave

In a rhyming scheme, any variable must appear at least twice (otherwise there is no rhyme). So for a quatrain, the only possibilities are two rhymes each occurring twice or one rhyme occurring four times. We have already seen two of the former. In a Shakespearean sonnet the rhyming scheme is $\mathcal{XYXY}$, i.e. the rhymes are *alternate*. In a Petrachean sonnet it is $\mathcal{XYYX}$, i.e. the rhymes are *symmetric*. Still there is a third possibility, $\mathcal{XXYY}$ which we will call *consecutive*. As we will see later this form does occur and is a perfectly legitimate form to use as a quatrain. Thus there are just 3 possibilities. Alternatively, we can adopt a more mathematical approach. Define the *combination function*

$$C(n, m), \qquad n \geq m \geq 1,$$

to be the number of ways of choosing a subset of m elements from a set of n elements. Then the number of rhyming schemes of four lines where two rhymes each occur twice is $C(4, 2)/2 = 3$.

We now return to the latter possibility, one rhyme occurring four times, i.e. $\mathcal{XXXX}$. Call such a scheme *perfect*. Again, this form certainly occurs. But there is a possible difference here between what might be understood in literature and in mathematics. In the three rhyming schemes $\mathcal{XYXY}$, $\mathcal{XYYX}$ and $\mathcal{XXYY}$, a poet would probably assume that $\mathcal{X}$ and $\mathcal{Y}$ are different rhymes, or, in mathematical terms, that $\mathcal{X} \neq \mathcal{Y}$. Thus the form $\mathcal{XXXX}$ would be a fourth distinct possibility. However a mathematician would take a different viewpoint, regarding the perfect rhyming scheme as a special case of each of the other three forms, i.e. where $\mathcal{X} = \mathcal{Y}$. As we will see later this is more appropriate for our purposes. Readers worried about this approach can consider that what we are doing is entirely analogous to regarding a square with sides of length x as a special case of a rectangle with sides of length x and y in which $x = y$. Thus there are 3 possible rhyming schemes for a quatrain and since the octave consists of two quatrains, the number of possibilities is simply $3 \times 3 = 9$. Again we make no assumptions as to whether there are any rhymes between the two quatrains such as for example in a *Spenserian sonnet* with rhyming scheme

$$\mathcal{ABAB} - \mathcal{BCBC} - \mathcal{CDCD} - \mathcal{EE}.$$

This has the same basic structure as the Shakespearean sonnet and in mathematical terms would be considered as an important subclass of the latter.

3. Structure of the sestet

If the structure of the octave is quite tightly controlled, then the structure of the sestet is much less so. In all of Shakespeare's 154 sonnets the sestet consists of a further quatrain and a rhyming couplet. In Petrachean sonnets the sestet is generally divided into two *tercets* of three lines with specific rhyming schemes. We will take a more general view following the definition quoted above from [3], i.e. "two, or three, other rhymes in the remaining six lines". As we observed above, to be a rhyme each ending must appear at least twice and so there are four possibilities.

1. Three rhymes each occurring twice.
2. Two rhymes, one occurring twice and the other four times.
3. Two rhymes each occurring three times.
4. One rhyme occurring six times.

As with the octave we regard the last case as a special case of all of the other three schemes, which we will now consider in detail.

Case 1: For three rhymes each occurring twice there are 15 possibilities. These rhyming schemes include those of a quatrain followed by a rhyming couplet, (as well as the opposite, a rhyming couplet followed by a quatrain), regular schemes such as the two tercets $\mathcal{XYZ}-\mathcal{XYZ}$ as well as various irregular forms such as $\mathcal{XYZYXZ}$.

$\mathcal{XXYYZZ}$	$\mathcal{XYXYZZ}$	$\mathcal{XYYXZZ}$	$\mathcal{XYYZXZ}$	$\mathcal{XYYZZX}$
$\mathcal{XXYZYZ}$	$\mathcal{XYXZYZ}$	$\mathcal{XYZXYZ}$	$\mathcal{XYZYXZ}$	$\mathcal{XYZYZX}$
$\mathcal{XXYZZY}$	$\mathcal{XYXZZY}$	$\mathcal{XYZXZY}$	$\mathcal{XYZZXY}$	$\mathcal{XYZZYX}$

Table 1: Three rhymes each occurring twice.

Case 2: Considering next the possibility of two rhymes, one occurring twice and the other four times, there are again 15 possibilities.

$\mathcal{XXYYYY}$	$\mathcal{XYXYYY}$	$\mathcal{XYYXYY}$	$\mathcal{XYYYXY}$	$\mathcal{XYYYYX}$
$\mathcal{YXXYYY}$	$\mathcal{YXYXYY}$	$\mathcal{YXYYXY}$	$\mathcal{YXYYYX}$	$\mathcal{YYXXYY}$
$\mathcal{YYXYXY}$	$\mathcal{YYXYYX}$	$\mathcal{YYYXXY}$	$\mathcal{YYYXYX}$	$\mathcal{YYYYXX}$

Table 2: Two rhymes, one occurring twice and the other four times.

Case 3: Finally for the possibility of two rhymes each occurring three times there are 10 possibilities. Of these, the schemes $\mathcal{XXXYYY}$ and $\mathcal{XYXYXY}$ might be thought of as regular but the others seem to be irregular.

$\mathcal{XXXYYY}$	$\mathcal{XXYXYY}$	$\mathcal{XXYYXY}$	$\mathcal{XXYYYX}$	$\mathcal{XYXXYY}$
$\mathcal{XYXYXY}$	$\mathcal{XYXYYX}$	$\mathcal{XYYXYX}$	$\mathcal{XYYXXY}$	$\mathcal{XYYYXX}$

Table 3: Two rhymes each occurring three times.

Again, the entries in the tables above can be counted using the combination function. In Table 2 the number is $C(6,2) = C(6,4) = 15$ and in Table 3 the number is $C(6,3)/2 = 10$. The counting in Table 1 is more complex, but the number of rhymes can be evaluated as $(C(6,2)/3) \times (C(4,2)/2) = 15$.

The fact that the number of rhyming schemes in Tables 1 and 2 is the same is not coincidental. Those in the latter can be obtained from those in the former as special cases. For example the scheme $\mathcal{YXXYYY}$ can be obtained from the schemes $\mathcal{XYYXZZ}$, $\mathcal{XYYZXZ}$ and $\mathcal{XYYZZX}$ by simultaneously making all of the switches

$$\mathcal{X} = \mathcal{Y}, \qquad \mathcal{Z} = \mathcal{Y} \qquad \text{and} \qquad \mathcal{Y} = \mathcal{X}.$$

In fact all of the rhyming schemes in Table 2 can be obtained from three schemes in Table 1. Moreover each scheme in Table 1 gives rise to three schemes in Table 2. For example the scheme $\mathcal{XXYYZZ}$ becomes the schemes $\mathcal{XXYYYY}$, $\mathcal{YYXXYY}$ and $\mathcal{YYYYXX}$.

In mathematical terms this defines a structure known as a *graph*. Formally this is defined as a set or collection of entities called *vertices*, some of which are joined by further entities called *edges*. In our case, the vertices are the 30 rhyming schemes, partitioned into two classes of 15 schemes corresponding to Table 1 and Table 2 respectively. We join two vertices by an edge if one rhyming scheme is related to another as described above. Thus every vertex in each class is joined to three vertices in the other class. In mathematical terminology we have a *cubic bipartite graph*.

In fact this graph is very well-known and has interesting properties. It is known as *Tutte's* $(3, 8)$-*cage* and is shown on page 271 of [4]. The number 3 here refers to the fact, stated above, that every vertex is adjacent to three others and the number 8 to the fact that the shortest cycle in the graph is of that length. It is the unique graph on the minimum number of vertices which has these two properties. We include a picture of the graph below (see Figure 1) with the vertices numbered together with a table (see Table 4) indicating the rhyming scheme corresponding to each vertex. This description is essentially equivalent to that given in a paper by H. S. M. Coxeter [2]. An excellent Introduction to Graph Theory is the book having the same title by R. J. Wilson [5].

1	$\mathcal{XYYZZX}$	2	$\mathcal{XYYYYX}$	3	$\mathcal{XYZYZX}$	4	$\mathcal{YYXYXY}$
5	$\mathcal{XXYZYZ}$	6	$\mathcal{XXYYYY}$	7	$\mathcal{XXYYZZ}$	8	$\mathcal{YYYYXX}$
9	$\mathcal{XYXYZZ}$	10	$\mathcal{YXYXYY}$	11	$\mathcal{XYZYXZ}$	12	$\mathcal{XYYYXY}$
13	$\mathcal{XYYZXZ}$	14	$\mathcal{YYYXYX}$	15	$\mathcal{XYXZYZ}$	16	$\mathcal{XYXYYY}$
17	$\mathcal{XYXZZY}$	18	$\mathcal{YYYXXY}$	19	$\mathcal{XXYZZY}$	20	$\mathcal{YYXYYX}$
21	$\mathcal{XYZXYZ}$	22	$\mathcal{YXYYXY}$	23	$\mathcal{XYZZYX}$	24	$\mathcal{YYXXYY}$
25	$\mathcal{XYZZXY}$	26	$\mathcal{YXYYYX}$	27	$\mathcal{XYZXZY}$	28	$\mathcal{XYYXYY}$
29	$\mathcal{XYYXZZ}$	30	$\mathcal{YXXYYY}$				

Table 4: Assignment of rhyming schemes to vertices.

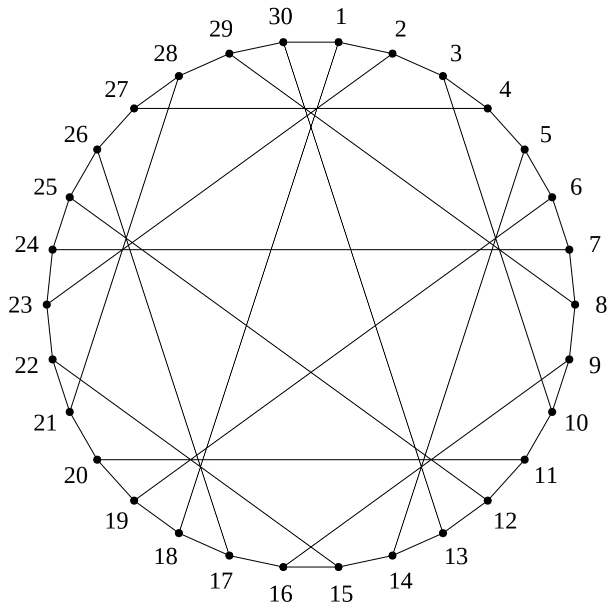

Figure 1: Tutte's $(3, 8)$ cage

Thus for the sestet, there are 15 possibilities from Case 1 and 10 possibilities from Case 3, giving a total of 25 possibilities overall and a total of $9 \times 25 = 225$ different rhyming schemes for a sonnet.

4. The Rural Muse

As noted above, Shakespeare's sonnets all exhibit the same rhyming scheme. However those by John Clare in [1] are much more varied. Our analysis shows that in fact only just over half (47 out of 86) of the sonnets satisfy one of the 225 enumerated possibilities, which might be termed *basic sonnet forms*. John Clare is more flexible in his use of rhymes and in the next paragraph we consider some of these.

As an example, consider Sonnet #63 where the rhyming scheme used

$$\mathcal{ABAB} - \mathcal{ACAC} - \mathcal{DDAD} - \mathcal{EE}$$

is still essentially Shakespearean but with the rhyme $\mathcal{D}$ introduced in the ninth line. Sonnet #50 has the scheme

$$\mathcal{ABAB}\text{–}\mathcal{CBCD} - \mathcal{DADADA},$$

again with the rhyme $\mathcal{D}$ introduced early but otherwise satisfying one of the basic sonnet forms. Sonnet #13 has the scheme

$$\mathcal{ABAB} - \mathcal{CDCD} - \mathcal{CDCECE}.$$

But perhaps the most interesting of the sonnets which formally violate the rules is Sonnet #21 "Fame". Here the rhyming scheme is very tightly controlled,

$$\mathcal{AAAA} - \mathcal{AABA} - \mathcal{ABACAC}.$$

Thus we have an example of a quatrain with a perfect rhyming scheme, $\mathcal{A}$ being the ending "ing". Unusually, John Clare also includes a "sonnet" with 16 lines! This is #57 and consists of four quatrains

$$\mathcal{AABB} - \mathcal{CCDD} - \mathcal{EEFF} - \mathcal{GGHH}.$$

We now consider the 47 sonnets which are of basic form. Of these, 20 are Shakespearean sonnets, sometimes with equality of rhymes such as in Sonnet #52 with $\mathcal{E} = \mathcal{D}$ and Sonnet #60 with $\mathcal{C} = \mathcal{A}$ and $\mathcal{E} = \mathcal{A}$. Just one of the sonnets #24 "The Sycamore" is Spenserian. A further 4 of the sonnets have the rhyming scheme

$$\mathcal{ABAB} - \mathcal{CDCD} - \mathcal{EE} - \mathcal{FGFG},$$

i.e. in the sestet, the rhyming couplet precedes the quatrain. Another 17 sonnets have the scheme

$$\mathcal{AABB} - \mathcal{CCDD} - \mathcal{EEFF} - \mathcal{GG},$$

clearly a form favoured by John Clare. This just leaves 6 sonnets, #7, #10, #12, #32, #70 and #79, the first 3 of which might be termed *quasi-Shakespearean* since they have the form of three quatrains with alternate rhyming schemes but a concluding couplet which is not rhyming. The most interesting of the other 3 sonnets is #32, "The Milking Shed" which has the rhyming scheme

$$\mathcal{ABBA} - \mathcal{CDCD} - \mathcal{EFEF} - \mathcal{GG},$$

the only example to include quatrains with differering rhyming schemes.

5. Concluding remarks

In this paper, we have enumerated all possible rhyming schemes in a sonnet based on rhymes occurring within both quatrains of the octave and within the sestet. The analysis could be extended to consider rhymes between the two quatrains and between the octave and sestet. However this would be a much more complex undertaking and to some extent would obscure the essential structure of the basic sonnet forms. But it would certainly be of interest to apply the same ideas as presented in this paper to other literary forms.

Our illustrations of sonnets have all been taken from one poet, John Clare, and we have compared his use of rhyming schemes with that of Shakespeare. If we just look at the statistics of the first quatrains of Clare's 86 sonnets, we find that, not including the perfect scheme which occurs just once, (in Sonnet #21, see above), the most numerous is the alternate scheme $\mathcal{ABAB}$ occurring 60 times. The consecutive scheme $\mathcal{AABB}$ occurs 18 times and the symmetric scheme $\mathcal{ABBA}$ just twice. The remaining 5 sonnets have non-basic forms. It would be interesting to know how this compares with the work of other poets.

Acknowledgement

I wish to thank my colleagues, Jozef Širáň, Rob Lewis and especially Grahame Erskine from the Open University for conversations about this work which led to identifying the graph in Section 3 as Tutte's $(3,8)$-cage. I also thank Tony Forbes, also from the Open University, for producing the picture of the graph.

References

[1] J. Clare, *The Rural Muse*, Mid Northumberland Arts Group and Carcanet New Press, Ashington, Northumberland, 1982.

[2] H. S. M. Coxeter, Twelve points in PG(5,3) with 95040 self-transformations, *Proceedings of the Royal Society, London A*, Volume **247** (1958), pages 279–293.

[3] M. Drabble (editor), *The Oxford Companion to English Literature*, 5th ed., Oxford University Press, Milton Keynes, 1985.

[4] R. C. Read and R. J. Wilson, *An Atlas of Graphs*, Oxford University Press / Clarendon Press, Oxford, 1998.

[5] R. J. Wilson, *Introduction to Graph Theory*, 2nd ed., Longman Group Limited, London, 1979.