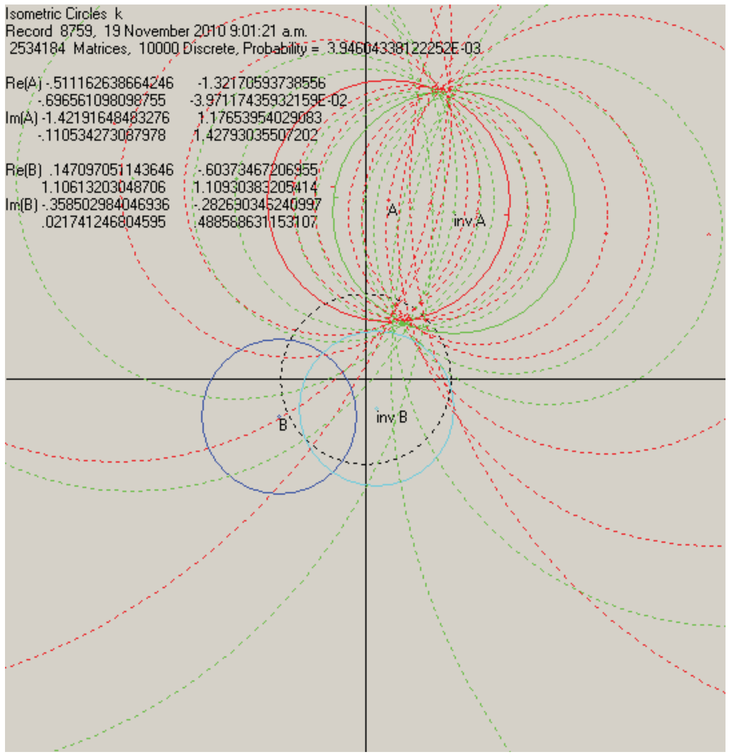
Random Discrete Groups in the Space of Möbius Transformations



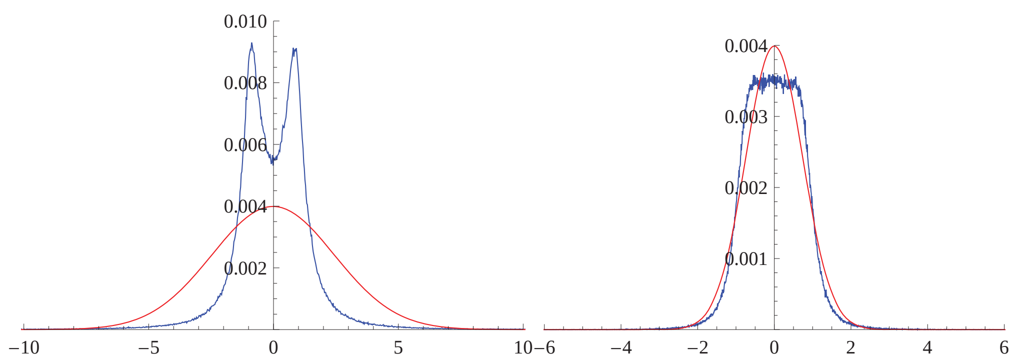
A Thesis presented in partial fulfillment of the requirements for the degree of
Master of Science
in
Mathematics
at Massey University, Albany, New Zealand

Graeme K O'Brien

March 2012

"I do not quite understand you," I said, with an uneasy foreboding as to what she meant...
"Surely a man must do a day's work first!"
I gazed in the white face of the woman, and my heart fluttered. She returned my gaze in silence.
"Let me first go home," I resumed, "and come again after I have found or made, invented, or at least discovered something!"

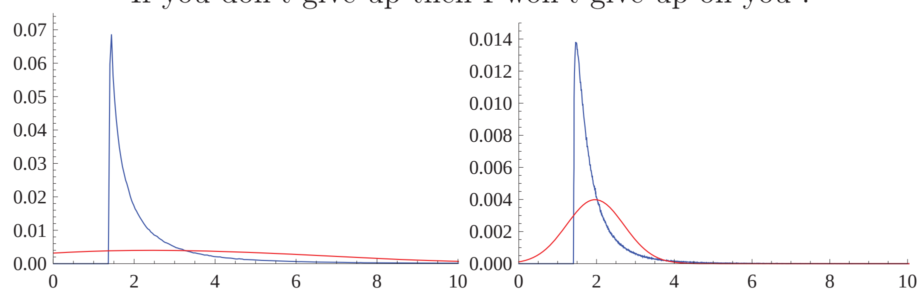
- George MacDonald



Acknowledgements

Distinguished Professor Gaven Martin.

“If you don’t give up then I won’t give up on you”.



ABSTRACT

Random Discrete Groups in the Space of Möbius Transformations

Graeme K O'Brien

Discrete subgroups of random Möbius transformations are investigated using computational methods together with collateral mathematical analysis. The main results include quantification of the likelihood of occurrence of two generator discrete groups and studies of the sharpness of the Hadamard inequality for random matrices and of the scale invariance for the domain of definition for matrix entry distributions derived by normalisation of matrices in $GL(2, \mathbb{C})$ to $SL(2, \mathbb{C})$.

TABLE OF CONTENTS

LIST OF FIGURES

Chapter 1

INTRODUCTION

In this thesis, discrete subgroups of random Möbius transformations are investigated using computational methods together with some collateral mathematical analysis. The main results presented are as follows:

- quantification of the likelihood of occurrence of two-generator discrete groups
- a study of the sharpness of the Hadamard inequality for random matrices
- for normalisation of matrices in $GL(2,\mathbb{C})$ to $SL(2,\mathbb{C})$ accomplished by division of all matrix entries by the square root of the determinant, a study of scale invariance of the resultant distributions with respect to the magnitude of matrix entry domains in $GL(2,\mathbb{C})$
- derivations of some algebraic expressions and determination of methodology for distributions of random variables over domains not restricted to non negative numbers, with particular interest in distributions of determinants of random matrices

We use these results and observations to test the efficacy of standard criteria for discreteness of Möbius groups, further work to prove the statistical inferences and refine experimentally determined features will be the subject of a subsequent thesis.

1.1 Möbius transformation Groups

The complex plane $\mathbb{C}$ can be extended to the Riemann sphere $\hat{\mathbb{C}}$ by the addition of the point at ∞, $\hat{\mathbb{C}} = \mathbb{C} \cup \{\infty\}$, and considering that all automorphisms on the sphere are Möbius transformations, fractional linear mappings of the form $f(z) = \dfrac{az+b}{cz+d}$, representation by matrices A in $SL(2,\mathbb{C})$ of a standard form is possible:

$$f \leftrightarrow A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \qquad a,b,c,d \in \mathbb{C},\ \det(A) = ad - bc = 1 \tag{1.1}$$

and composition of Möbius transformations corresponds to matrix multiplication. The group of Möbius transformations under composition is homomorphic to a subgroup of 2×2 complex matrices under multiplication, unique up to sign, allowing

Möbius transformations to be studied by considering subgroups of the matrix groups GL(2,$\mathbb{C}$) and $SL(2, \mathbb{C})$.

Groups of Möbius transformations can be identified with groups of hyperbolic isometries, hence there are connections to 3-manifold theory, see e.g. Thurston [1], Maskit [2].

Two-generator subgroups of $SL(2, \mathbb{C})$ are important in the context of this thesis as a result of theorems by Jørgensen [3], Gehring and Martin [5] and Klein (see [6]), these theorems are discussed in the next section. Gehring and Martin show that the subgroup $G =< A, B >\subset PSL(2, \mathbb{C})$ is uniquely determined up to conjugacy by the three complex numbers, the parameters of the two-generator subgroup,

$$\left.\begin{array}{rcl} \beta(A) & = & trace^2(A) - 4 \\ \beta(B) & = & trace^2(B) - 4 \\ \gamma(A, B) & = & trace([A, B]) - 2 \end{array}\right\} \tag{1.2}$$

where here the commutator $ABA^{-1}B^{-1}$ of A and B (the generators of the subgroup $\langle A, B\rangle$) is designated by $[A, B]$ and we assume $\gamma(A, B) \neq 0$. The β and γ parameters are dependent only on the traces of the matrices and of the commutator of a matrix pair, and since the trace and determinant of a matrix in $SL(2, \mathbb{C})$ are invariant under conjugation, the study of many properties of two-generator subgroups of $SL(2, \mathbb{C})$ is possible via suitably chosen conjugate matrices. It is noted that conjugacy preserves both geometric and algebraic invariants of a group.

Except where specifically indicated, in this thesis the term *group* refers to a finitely generated subgroup of $SL(2, \mathbb{C})$ (or of $GL(2, \mathbb{C})$ etc), the generators A and B of $< A, B >$ are distinct and a matrix in *standard form* means with complex entries as in (1.1).

Möbius transformations can be represented up to scalar multiple in $GL(2, \mathbb{C})$ and $GL(2, \mathbb{R})$, up to sign in $SL(2, \mathbb{C})$ and $SL(2, \mathbb{R})$ and uniquely in $PSL(2, \mathbb{C})$ and $PSL(2, \mathbb{R})$. The matrix spaces are elements of the sequences:

$$\begin{array}{llll} \mathbb{R}^8 & \supset GL(2, \mathbb{C}) & \supset SL(2, \mathbb{C}) & \supset PSL(2, \mathbb{C}) \\ \cup & \cup & \cup & \cup \\ \mathbb{R}^4 & \supset GL(2, \mathbb{R}) & \supset SL(2, \mathbb{R}) & \supset PSL(2, \mathbb{R}) \end{array} \tag{1.3}$$

The spaces of 2×2 complex and real matrices are $\mathbb{R}^8$ and $\mathbb{R}^4$ respectively. Omission of matrices with zero determinant allows $GL(2, \mathbb{C})$ and $GL(2, \mathbb{R})$ and their subsets to be groups under matrix multiplication.

1.2 Discreteness

A group is discrete if it contains no sequence that tends to the identity (see Beardon, [7]). Theorem 1.1 below (due to Jørgensen, see [7], [3]) shows that the discreteness of the two-generator subgroups of Möbius transformations determines the discreteness of the group. In a later theorem Jørgensen [4] shows that a similar condition on single generator groups obtains for Möbius transformations that can be represented by matrices in $SL(2,\mathbb{R})$. A subgroup G of Möbius transformations is *elementary* if any two of its elements of infinite order have a common fixed point in $\hat{\mathbb{C}}$ (cf [4]), it follows that for a non elementary subgroup G there must exist no finite G-orbits in $\hat{\mathbb{C}}$ (cf [7]).

Theorem 1.1. *A non-elementary group G of Möbius transformations is discrete if and only if for each f and g in G, the group $\langle f, g\rangle$ is discrete*

Jørgensen also proves a theorem [3] which when restated in terms of the trace parameters provides a necessary condition for non elementary two-generator groups to be discrete:

Theorem 1.2. $|\beta(A)| + |\gamma(A,B)| \;\geq\; 1$

The commutator shrinking property of Lie groups (see e.g. [9]) leads to an inequality that we will use in the study of discrete groups as it describes neighbourhoods of the identity:

$$\|[A,B]-I\| \leq \sup\{\|A-I\|\,,\|B-I\|\} \quad \forall A,B \in SL(2,\mathbb{C}) \tag{1.4}$$

Suppose now that a Möbius transformation in the space $GL(2,\mathbb{C})$ maps the point z to the point $f(z)$ in $\hat{\mathbb{C}}$. Write the transformation:

$$f = \frac{az+b}{cz+d} \tag{1.5}$$

hence

$$\frac{df}{dz} = \frac{ad-bc}{(cz+d)^2} \tag{1.6}$$

If the transformations are represented by matrices using the standard form of (1.1), then the numerator of (1.6) is recognised as the determinant of the matrix, and for matrices in $SL(2,\mathbb{C})$ the equation reduces to:

$$df = \frac{dz}{(cz+d)^2} \tag{1.7}$$

Ford [30] equates the absolute values of the differentials to determine the locus of points in whose neighbourhood distances are preserved:

$$|df| = |dz| \Leftrightarrow |cz+d| = 1 \tag{1.8}$$

which is a set of points for which A represents an isometric transformation.

The locus is a circle in $\hat{\mathbb{C}}$, which provided $c \neq 0$ (that is, $f(\infty) \neq \infty$), is given by the equation:

$$\left|z + \frac{d}{c}\right| = \left|\frac{1}{c}\right| \tag{1.9}$$

and this is the *isometric circle* corresponding to the matrix, and is centered at the point $-\dfrac{d}{c}$ with radius $\dfrac{1}{|c|}$.

Ford notes that within the isometric circle, lengths and areas are increased under the transformation A, outside the circle they are decreased, and the circle itself is transformed into the isometric circle of the inverse transformation A^{-1} determined by the equation:

$$\left|z - \frac{a}{c}\right| = \left|\frac{1}{c}\right| \tag{1.10}$$

which is the equation of a circle centered at the point $\dfrac{a}{c}$ with the same radius $\dfrac{1}{|c|}$ as for the isometric circle of A. The three independent complex entries a, c and d of the matrix determine the two isometric circles of a transformation uniquely. The two complex coordinates of the centres of the isometric circles are determined by entries d, a and c but the real radii are determined by the modulus of the entry c. Hence the isometric circles of a matrix A in $SL(2,\mathbb{C})$ characterise the matrix only up to the modulus of the complex entry c, and for any two pairs of circles of equal radii in $\hat{\mathbb{C}}$, there exist an infinite number matrices in $SL(2,\mathbb{C})$ which have these pairs as isometric circles.

For matrices in $GL(2,\mathbb{C})$:

$$dz_1 = \frac{(ad-bc)dz}{(cz+d)^2} \tag{1.11}$$

The locus is a circle in $\mathbb{C}$, which provided $c \neq 0$, is given by:

$$\left|z + \frac{d}{c}\right| = \sqrt{\left|\frac{ad-bc}{c^2}\right|} \tag{1.12}$$

centered at the point $-\dfrac{d}{c}$ with radius $\dfrac{|\sqrt{ad-bc}|}{|c|}$.

Both these quantities are scale invariant, that is, if $A \in GL(2,\mathbb{C})$ then A and λA have the same isometric circles, $\lambda \in \mathbb{C}$. The scaling of the domain of definition for entry distributions of random matrices in $GL(2,\mathbb{C})$ is considered later in this thesis, but it is clear that normalisation of individual matrices in $GL(2,\mathbb{C})$ via division of each entry by the square root of the determinant leaves centres of isometric circles constant but the scales the radii. We note that although the square root function is two-valued, by (1.1) the two resultant matrices represent the same Möbius transformation and both also have determinant 1. It is immaterial to the transformation which root is used for normalisation.

A well known result of Klein (see e.g. Gilman, [6]) shows that if the discs enclosed by the isometric circles corresponding to the matrices (including inverses) of the two-generators A and B of $\langle A, B\rangle \in SL(2,\mathbb{C})$ are disjoint (or at most tangential) then the group $\langle A, B\rangle$ must be discrete. This criterion allows discreteness to be expressed explicitly in terms of an inequality between the parameters of disjoint isometric circles.

In this thesis the main problem is the study of random groups of Möbius transformations in order to assess the efficacy of the discreteness criteria embodied in Theorem 1.2 and Klein's Isometric Circle result by determining the likelyhood of corresponding inequalities being satisfied.

1.3 Random matrices

The nature of random distribution of matrices can be considered both from a computational perspective and from consideration of precise theorems on entry distribution. This thesis focuses on the computational approach (using Monte Carlo methodology), but also addresses some of the issues relating to the required theoretical distributions.

A fuller study of the distributions is to be the subject of future research.

For computational analysis (unless otherwise indicated) the definition of a random complex matrix will be a matrix of the form (1.1) where the eight real and imaginary components of the entries a, b, c and d of a matrix $A \in GL(2, \mathbb{C})$ are selected uniformly from a domain $[-k, k]$ where k is a non-negative real number. Determination of the complex square root $\pm\sqrt{ad - bc}$ of the determinant allows rejection of matrices not in $GL(2, \mathbb{C})$ and also normalisation as required to $SL(2, \mathbb{C})$ by division of every entry by the positive value of this complex square root. For normalisation of matrices in $GL(2, \mathbb{R})$ the mapping to $SL(2, \mathbb{R})$ is defined for positive determinant values only, and matrices in $GL(2, \mathbb{R})$ with negative determinants are rejected and replaced in the normalisation process. We study the results for different values of k and also the limit as $k \to \infty$, together with the implications of the use of domains $[-m, n]$, $m \neq n$. Alternative methods have been investigated, but are either impractical computationally (such as putting a distribution on the non-compact space $SL(2, \mathbb{C}) \subset \mathbb{R}^8$) or result in non homogeneous distributions (e.g. choosing three complex entries from a random distributions then calculating the fourth subject to the determinant constraint). Indeed certain geometric invariants of Möbius transformations can be computed directly from the $GL(2, \mathbb{C})$ matrix without referance to the determinant. The chosen procedure was used because it was efficient and resulted in homogeneous entry distributions in most cases; further justification lies in the identification of unexpected results regarding subgroups of Möbius transformations.

There are other valid approaches however that will be studied in subsequent work.

The generation of random numbers requires some care in choice of algorithm to maintain integrity and efficiency of computation. Almost without exception, random number generators produce near uniform distributions. There is an immediate problem with investigation of geometry in unbounded spaces using uniform distributions in that a finite interval must always be specified, such problems do not exist to the same extent for Gaussian distributions. Recent work by Strang [8] makes this point quite clearly and also concludes that Gaussian rather than uniform distributions seem to give the right answers for many geometric problems. Where required, Box-Muller transformations [10] were used for conversion of uniform distributions to Gaussian. In practice, other considerations affected the choice of distribution.

Chapter 2

COMPUTATIONAL ANALYSES

2.1 Mathematical Issues

We early determined that in generating random matrices in $SL(2,\mathbb{C})$ from those in $GL(2,\mathbb{C})$ both the domain of entry distribution for $GL(2,\mathbb{C})$ and the origin of the $GL(2,\mathbb{C})$ and $SL(2,\mathbb{C})$ matrix spaces are significant as far as the nature of the resultant distributions is concerned. Apart from distributions of matrix entries themselves, the following are important to the assesssment of discreteness of matrix groups in $SL(2,\mathbb{C})$:

- The determinant, or rather the square root of the determinant, involved in the normalisation of matrices to $SL(2,\mathbb{C})$
- The norm $\|A - I\|$, the distance between matrices and the identity used in determining discreteness
- The quantity $\|[A,B] - I\|$ as occurs in the Jørgensen (Theorem 1.2) and other discreteness criteria
- The *isometric distance* between two matrices, here defined as the infimum of the separations between the isometric circles corresponding to the two matrices; if all four such separations are non negative then $\langle A, B\rangle$ is a discrete subgroup of $SL(2,\mathbb{C})$. Circle separation is defined by equation (2.1):

$$\delta_{A,B} = \sqrt{(y_B - y_A)^2 + (x_B - x_A)^2} - r_A - r_B \tag{2.1}$$

(for circles of centres $(x_A, y_A), (x_B, y_B)$ and radii r_A, r_B

With respect to the significance of the origin of the $GL(2,\mathbb{C})$ and $SL(2,\mathbb{C})$ spaces, useful choices with respect to which distances (as defined by Euclidean norms on the $\mathbb{R}^8$ and $\mathbb{R}^4$ embedding spaces) can be measured are 0 (the origin of the embedding space) and $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ (corresponding to the identity of the matrix group). In zero centred space the real and imaginary matrix entry components are distributed over $[-k, k]$. In identity centred space the leading diagonal real components of matrix entries are distributed over $[-k+1, k+1]$ while all other entry components are distributed over $[-k, k]$. It eventuates that both zero and identity centred $SL(2, \mathbb{C})$, $SL(2, \mathbb{R})$, $PSL(2, \mathbb{C})$ and $PSL(2, \mathbb{R})$ spaces require consideration for analysis purposes.

2.2 Computation and Presentation Issues

Programs to define the procedures for the experiments were written in Visual Basic, and actual probabilistic events correspond to the sets of data files generated by program runs. Data analysis and presentation was performed within either the Visual Basic programs or with Mathematica via import of event files.

The standard random number generators that come with most software programs are inadequate when dealing with large amounts of data, the main problem being aliasing, whereby a reccurrance of the sequence of generated random numbers can create spurious artifacts. Such problems can be visible to close inspection of unprocessed data. A much better algorithm is embedded in the CryptoSys API [26] and conforms to a very conservative specification [27] and [28] providing random numbers having higher specification than required by ANSI X9.31 Appendix A for data encryption. This means that fixed point numbers can be provided having an extremely high quality of uniform randomness but with a severe concomittment speed penalty. However the simple and faster algorithm is still highly uniform, and accordingly the higher specification algorithm is used only where aliasing would otherwise be a problem as determined by the size of the set of numbers to be generated (which can be considerably greater than that corresponding to the sample size) and the nature of the random number algorithm specification. IEEE double precision floating point numbers are used for raw number generation and provide precision and integrity far in excess of that required for these statistical analyses.

Experimental parameters are chosen according to the following considerations:

- limitation of numerical resolution: restrict double precision numbers to 9 decimal digits for calculation (to avoid overflow in certain processing routines), use single precision floating point numbers for result arrays

- matrix sample size: dependent on context and the conflicting requirements for speed and resolution, typically chosen to be 1,000,000 which allows use of the faster random number generation algorithm without risk of aliasing

- quantile size: $\frac{1}{100}$ of a standard deviation rather than a fraction of the total range

- matrix space parameters: see later in this chapter for considerations

For distribution analysis we map the elemental events from experiments onto quantiles of a partitioned occurrence frequency domain, the standard deviation of the distribution being precisely the quantile size fraction with the above non-standard definition of quantile. Division of the quantile occurrences by the total of occurrences in the frequency domain gives probabilities. For comparison purposes a possibly constrained normal distribution having the same standard deviation and mean as the event distribution is then created over an identical frequency domain.

A Kolmogorov-Smirnov analysis procedure (see for instance [12]) is used whenever quantitative comparison of distributions is required. Unlike many competing techniques, this is easily extended to non normal distributions. Our use of this technique contains a refinement, in that we sum the computed probability density distributions quantile by quantile in order to derive cumulative probability distributions having substantially fewer but less variable data points than for ordered empirical distributions. Prior to comparison, distributions are synchronised by truncation of the domains so as to ensure that means and domains correlate; use of quantile sizes related to standard deviation ensures the validity of this procedure. The supremum of the set of distances between each of the n quantile values determines the Kolmogorov-Smirnov statistic D_n, and comparison of $\sqrt{n}D_n$ with standard published tables [11] allows for assessment of the degree of correlation. Specifically, if $\sqrt{n}D_n > 1.52$ then the two distributions compared fail to satisfy an equality hypothesis at an 80% level (a lower % figure is more stringent). In this thesis, if $\sqrt{n}D_n$ is of an order of magnitude below this 80% threshold then the compared distributions are adjudged to have non-significant differences.

For graphical presentations we use the following conventions unless otherwise stated. Horizontal axes represent the domain of a p.d.f (sometimes with outliers suppressed) and the vertical axis the probability. Blue is used for event data, red for comparative normal distributions. For overlaid cumulative distribution functions red is used for the first mentioned function, blue for the second and green for a third, then the colours recycled. In annotations, the variable k is the domain bound for entries of matrices in $GL(2,\mathbb{C})$ or $GL(2,\mathbb{R})$.

2.3 Discussion

It might be expected by reason of the isometric circle criterion for discreteness that a random two-generator subgroup of $PSL(2,\mathbb{C})$ would usually be discrete, but it is shown experimentally that this is not the case. The reason is that with entries of matrices A and B sampled from the same restricted distribution then A and B can be expected to be close to the mean of the matrix distribution; we will see that for matrices in $PSL(2,\mathbb{C})$ this mean is the identity and $[A,B]$ in consequence is close to the identity. Iteration of this process in the following fashion:

$$[A,B] \mapsto [[A,B],A] \mapsto [[[A,B],A],A]...$$

will then be expected to result in a sequence converging to the identity and $< A, B >$ will accordingly not be a discrete group.

Similarly, it might be expected on the basis of fractions such as $\frac{a}{\sqrt{ad-bc}}$ that scale invariance of resultant distributions in $SL(2,\mathbb{C})$ would apply with respect to the magnitudes of the entries of the matrix in $GL(2,\mathbb{C})$ under the normalisation process, and indeed experimental results generally support the expectation. We will show in Chapter 3 that this is not a conclusion that can be drawn from such real or complex number algebraic expressions, the algebra of distributions is via Fourier and Mellin convolutions.

Bearing in mind the comments made in the introduction, it might be expected that we could work with matrices in $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$ having entry distributions both identical and normal. However, it was seen from the early computational results that freedom of choice of the nature of entry distributions is constrained, and our objective necessarily becomes somewhat less, namely the generation of matrices in $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$ whose entries are identically randomly distributed. We note that the generation of a set of random matrices with a certain entry distribution does not mean that such distribution will be maintained after matrix operations within subgroups, nor does it mean that identicality of entry distributions over all entries of the matrices will be maintained. This will be demonstrated with computational examples.

2.4 Entry Distributions of Random Matrices in $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$

The procedure of normalisation of random matrices in $GL(2,\mathbb{C})$ and $GL(2,\mathbb{R})$ via division by the square root of the determinant must be demonstrated to be capable of meeting the objective of all matrix entry components having identical distributions.

We accordingly consider the nature of the entry distributions in these spaces in terms of three parameters: uniform or normal entry distributions in $GL(2,\mathbb{C})$, the size of the domain interval, and the spatial origin in $GL(2,\mathbb{C})$; we also consider variation in the choice of particular entry components in $SL(2,\mathbb{C})$ or $SL(2,\mathbb{R})$ (e.g. whether they are real components of leading diagonal entries or otherwise); our objective being a suitable choice of parameters.

2.4.1 Random Matrices in $SL(2,\mathbb{R})$ derived from uniformly and normally distributed matrices in zero centred $GL(2,\mathbb{R})$ space

The generation of random matrices in $SL(2,\mathbb{R})$ from multiple sets of $1,000,000$ matrices in zero centred $GL(2,\mathbb{R})$ is considered first. Kolmogorov-Smirnov analysis shows the differences between the observed event distributions of leading diagonal entries and otherwise of matrices in $SL(2,\mathbb{R})$ generated from matrices in $GL(2,\mathbb{R})$ having positive or negative determinants and with entries uniformly distributed over domains from $[-0.01, 0.01]$ to $[-10^7, 10^7]$, to be non-significant. In Figure (2.1), the left hand graph shows a probability density distribution representative of these identical distributions while the identicality itself is indicated by the right hand graph, which consists of twelve overlaid cumulative distributions corresponding to different parameter selections. Kolmogorov-Smirnov statistics for these distributions range from $\sqrt{n}D_n = 0.07$ to $\sqrt{n}D_n = 0.17$, all at least an order of magnitude below the threshold of 1.52.

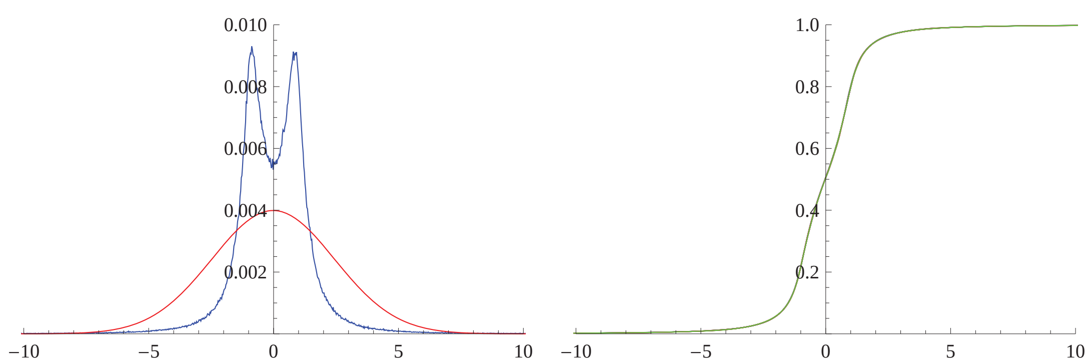


Figure 2.1: Left: Representative entry distributions (see text) for matrices in $SL(2,\mathbb{R})$ generated by normalising 1,000,000 random matrices in $GL(2,\mathbb{R})$ with positive determinants and having entries distributed uniformly about 0, Right: Superimposed cumulative distributions for real components of leading diagonal entries and otherwise over the intervals $[-0.01, 0.01]$ to $[-10^7, 10^7]$, twelve distributions shown

The individual entry distributions can be characterised as identical at all scales and of primarily bimodal symmetrical form about 0 and with peak occurrences centred at ± 1. As an indication of the size of the envelope that matrices so generated occupy, we note that the probability that an entry of a random matrix in $SL(2,\mathbb{R})$ generated as specified is within the interval $[-2,2]$ is 0.9875. The distribution is such that there is only a very small probability of any entry of a random matrix in $SL(2,\mathbb{R})$ generated as specified being outside a single digit radius of the centre 0.

The distributions of entries of matrices in $SL(2,\mathbb{R})$ generated from normally distributed matrices in $GL(2,\mathbb{R})$ differ only qualitatively, and similar conclusions apply.

2.4.2 Random Matrices in $SL(2,\mathbb{R})$ derived from uniformly and normally distributed matrices in identity centred $GL(2,\mathbb{R})$ space

For random Matrices in $SL(2,\mathbb{R})$ (generated from both uniformly and normally distributed matrices in identity centred $GL(2,\mathbb{R})$ space) the situation is somewhat different. The qualitative nature of the distributions changes with entry domain size in $GL(2,\mathbb{R})$ and also differs for leading diagonal entries and otherwise, but the distributions in all cases approach those for the generation from matrices in zero centred $GL(2,\mathbb{R})$ as the domain bound $k \to \infty$. Figure (2.2) indicates distributions for real components of leading diagonal entries for a selection of domain bounds.

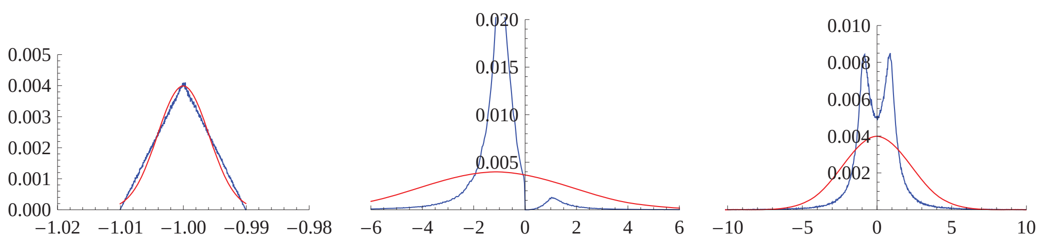


Figure 2.2: Entry distributions for real components of leading diagonals of matrices in $SL(2,\mathbb{R})$ generated by normalising 1,000,000 random matrices with entries from identity centred $GL(2,\mathbb{R})$ space, entries being uniformly distributed over the intervals; Left to Right: $[-0.01, 0.01]$, $[-1,1]$, $[-10^7, 10^7]$

2.4.3 Choice of experimental parameters for the generation of random matrices in $SL(2,\mathbb{R})$

We recall the objective of generated matrices in $SL(2,\mathbb{R})$ having identical entry distributions. It is clear that this can be achieved with mimumum processing overhead by using zero centred matrices in $GL(2,\mathbb{R})$ with entries uniformly distributed over any domain and normalising the random matrices by suitable division by the square root of the determinant. The resultant random matrices in $SL(2,\mathbb{R})$ have identical zero centred entry distributions which are however non normal, being composed of bimodal distribution components about 0 and with peak occurrences about ± 1.

2.4.4 Random matrices in $SL(2,\mathbb{C})$

We repeat the exercise for matrices in $SL(2,\mathbb{C})$, first with uniformly distributed entries then with normally distributed entries, and again find in both cases (using the same procedure, generating from random $GL(2,\mathbb{C})$ matrices in zero centred space) that identical entry distributions together with scale invariance of the distributions with respect to the domain of entries of matrices in $GL(2,\mathbb{C})$ result. In Figure (2.3) the left hand graph shows a representative probability density distribution as before, while the identicality of all these entry distributions is indicated by the right hand graph, which consists of twelve overlaid cumulative distributions corresponding to different parameter selections.

The distributions of entries of matrices in $SL(2,\mathbb{R})$ generated from normally distributed matrices in $GL(2,\mathbb{R})$ differ only qualitatively, and similar conclusions apply.

Kolmogorov-Smirnov analysis shows differences between the distributions to be non-significant for all the above cases, values of $\sqrt{n}D_n$ ranging about from 0.05 to 0.15, again at least an order of magnitude below the threshold of 1.52.

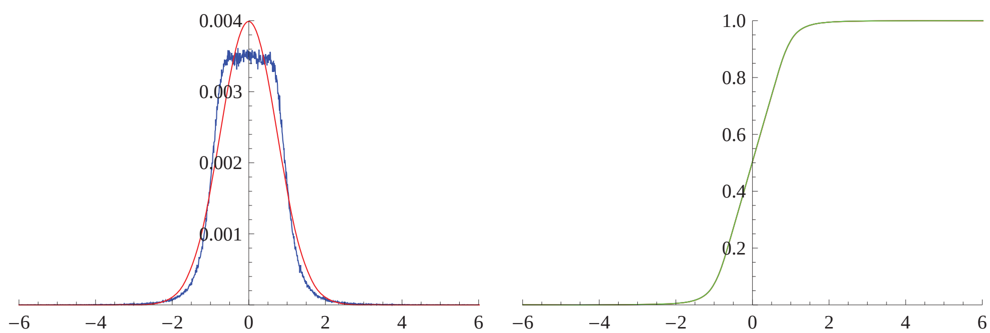


Figure 2.3: For matrices in $SL(2,\mathbb{C})$ generated by normalising 1,000,000 random matrices in $GL(2,\mathbb{C})$ having entries distributed uniformly about 0, Left: Representative entry distribution (see text) Right: Superimposed cumulative distributions for real components of leading diagonal entries and otherwise over the intervals $[-0.01, 0.01]$ to $[-10^7, 10^7]$, twelve distributions shown

Our conclusion is that the same choice of parameters will suit generation of suitable matrices in $SL(2,\mathbb{C})$ as for $SL(2,\mathbb{R})$, the only difference is that the resultant entry distributions in $SL(2,\mathbb{C})$ are unimodal about 0 (with distribution close to uniform between approximately ± 1). Generation of matrices in $SL(2,\mathbb{C})$ from identity centred $GL(2,\mathbb{C})$ matrices again results in distributions that are qualitatively different for lower domain bounds (and with differences between distributions for real components of leading diagonal entries and otherwise), but the distributions tend to that for the zero centred case as the domain bound $k \to \infty$.

We summarise here our reasons for choosing to generate matrices in $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$ from uniformly distributed zero centred matrices in $GL(2,\mathbb{C})$ and $GL(2,\mathbb{R})$ respectively and accept the resultant entry distributions rather than requiring matrices to have normally distributed entries:

- The construction of large quantities of random matrices in $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$ with normally distributed entries with reasonable efficiency presents serious problems at present
- There is no point in starting with normal distributions of the $GL(2,\mathbb{C})$ entries as the resultant distributions in $SL(2,\mathbb{C})$ are qualitatively similar and generation of normally distributed numbers takes more proccessing time
- The procedure does result in identical distributions for all matrix entry components

The entry component distributions of the resultant $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$ matrices are a natural result of the normalisation process as described, are symmetric about 0, are identical for all components, and yield useful results.

2.5 Random matrices in $GL(2,\mathbb{R})$ and $GL(2,\mathbb{C})$

Although our focus is on matrices in $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$, we consider briefly the distributions of traces, entry products and determinants in $GL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$. The reasons are for comparison with existing literature, correlation with some mathematical derivations in Chapter 3, and in the case of entry products, familiarisation with the concept of non-invariance of distributions under some matrix operations.

2.5.1 Traces of Random Matrices

Figure (2.4) shows an experimentally determined distribution of traces of random matrices in $GL(2,\mathbb{R})$ with uniform entry distributions. The results for $k >> 1$ are indistinguishable for identity and zero centred spaces and for $GL(2, \mathbb{C})$ and real or imaginary components. This computational distribution has been determined mathematically in Chapter 3, and Figure (3.1) is identical with Figure (2.4) except for the scales; frequency of occurrence is generally used in this thesis except for purposes of comparison with published work.

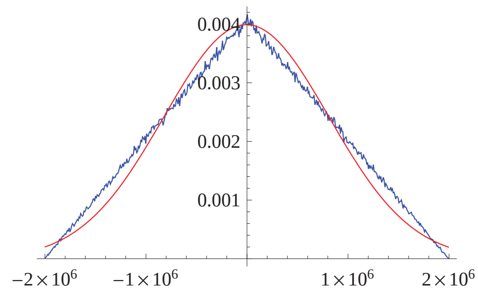


Figure 2.4: Distribution of trace for 1,000,000 random matrices in $GL(2, \mathbb{R})$, entries from a zero centred space with entries uniformly distributed over the interval $[-10^6, 10^6]$

2.5.2 Products of entries in $GL(2, \mathbb{R})$

Figure (2.5) shows for a matrix in the standard form (but with real entries) the distribution of products ad of the leading diagonal entries entries a, d for 1,000,000 uniform random matrices in identity centred $GL(2, \mathbb{R})$ space, the distribution is identical to that for zero centred space and distributions of leading diagonal entry products ad are identical to those for non leading diagonal entry products bc. As matrices in $GL(2, \mathbb{R})$ (and $GL(2, \mathbb{C})$) are multiplied successively, resultant entry distributions narrow markedly. The evolution of distributions in single and two-generator subgroups will be developed further in a subsequent thesis.

2.5.3 Determinants of matrices in $GL(2, \mathbb{R})$ and $GL(2, \mathbb{C})$

The distributions of the real and imaginary components of determinants of uniform random matrices in $GL(2, \mathbb{C})$ for $k >> 1$ are very close to identical to each other and also to the distributions of determinants of matrices in $GL(2, \mathbb{R})$ and to those for matrices in $GL(2, \mathbb{C})$ or $GL(2, \mathbb{R})$ spaces whether identity or zero centred. Figure (2.6) shows an experimentally observed distribution for a domain bound $k = 10^6$ and $1,000,000$ uniform random matrices in $GL(2, \mathbb{R})$. The mathematical derivation is only straightfoward for distributions with either all non negative or all non positive domains. The nature of algebra with convolutions ensures that results for distributions over domains that cross 0 are qualitatively different to the simpler case, in the

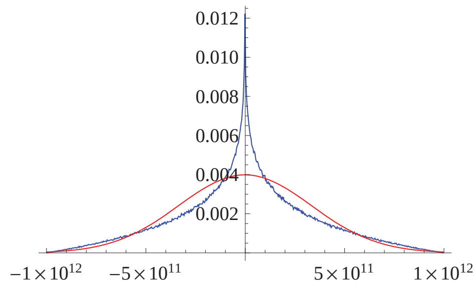


Figure 2.5: Distribution of entry products ad for 1,000,000 random matrices in $GL(2,\mathbb{R})$, entries from a identity centred space with entries uniformly distributed over the interval $[-10^6, 10^6]$

case of determinants the mathematical derivation has complications, hence our inability to provide correlation at this stage.

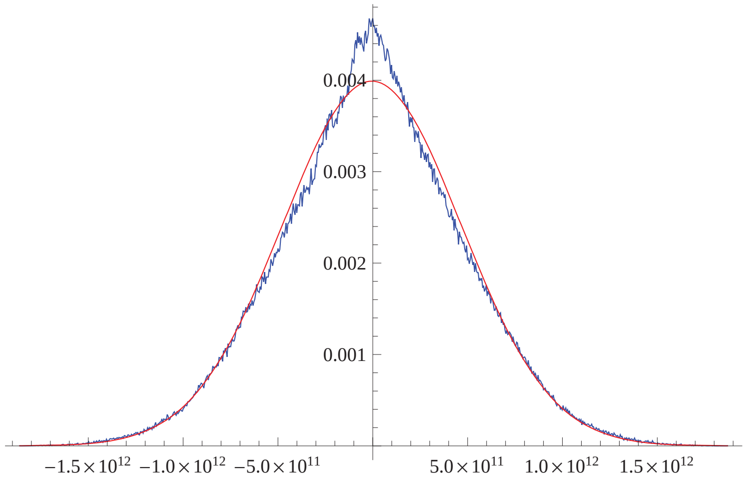


Figure 2.6: Distribution of the determinant for 1,000,000 random matrices in $GL(2,\mathbb{R})$, entries from a zero centred space with entries uniformly distributed over the interval $[-10^6, 10^6]$

2.6 $PSL(2,\mathbb{C})$ and $PSL(2,\mathbb{R})$

While matrices in the spaces $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$ serve well in general to represent composition, functions, and parameters of Möbius transformations they are inadequate whenever a metric (as defined earlier on the embedding spaces) is required, because the spaces represent double covers for the complex and real spaces of Möbius transformations respectively. As a result distances between Möbius transformations are not well defined when represented by matrices in $SL(2,\mathbb{C})$ or $SL(2,\mathbb{R})$. $PSL(2,\mathbb{C})$ and $PSL(2,\mathbb{R})$ are the quotient spaces of $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$ respectively for a homomorphism whose kernal is $\{-I, I\}$, and so an algorithm to convert matrices $SL(2,\mathbb{C}) \mapsto PSL(2,\mathbb{C})$ and $SL(2,\mathbb{R}) \mapsto PSL(2,\mathbb{R})$ is sought. The following formulations assume that all non diagonal entries of the matrices remain invariant.

For a matrix A in $SL(2,\mathbb{C})$ in the form (1.1) embedded in $\mathbb{R}^8$, we consider the space defined by the complex diagonal entries a and d to be a pseudo-plane with the complex numbers a along the x axis and d along the y axis. If we apply the following mapping to the quadrants of this pseudo-plane:

$$\begin{array}{llll} (-a,-d) & \in Q_3 \mapsto & (a,d) & \in Q_1 \quad \text{third quadrant to first} \\ (a,-d) & \in Q_4 \mapsto & (-a,d) & \in Q_2 \quad \text{fourth quadrant to second} \end{array} \tag{2.2}$$

then this does successfully map $-I$ to I, leaves the determinant invariant, partitions $SL(2,\mathbb{C})$ into two cosets, and is a projection that allows antipodal points on zero centred hyperspherical shells to be not distinguished; and accordingly is a mapping function (which we will call ϕ) that takes the entire ad pseudo-plane onto the positive half pseudo-plane as follows:

$$\phi(a,d) = \begin{cases} 1 & Re(d) \geq 0 \\ -1 & d < 0 \end{cases} \tag{2.3}$$

A second mapping function ψ that would perform a projection with the required properties equally well is:

$$\psi(a,d) = \begin{cases} 1 & Re(a) \geq 0 \\ -1 & a < 0 \end{cases} \tag{2.4}$$

These mappings meet all the requirements for the projective space $PSL(2,\mathbb{C})$, and mapping functions of precisely the same form apply for $SL(2,\mathbb{R}) \mapsto PSL(2,\mathbb{R})$ (where a and d are now real numbers and the ad plane is real rather than a pseudo-plane), and it can be seen that the resultant spaces represent the projective subgroups $PSL(2,\mathbb{C})$ and $PSL(2,\mathbb{R})$ of $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$ respectively. The particular subsets of $PSL(2,\mathbb{C})$ or $PSL(2,\mathbb{R})$ generated by experimental events will have different members depending on whether ϕ or ψ is used, we will compare distributions for random matrices generated via both mappings. It is clear that to ensure that distance between Möbius transformations is well defined, one of these two mappings must first be applied to matrices in $SL(2,\mathbb{C})$ or $SL(2,\mathbb{R})$.

2.7 Random matrices in $PSL(2,\mathbb{R})$ and $PSL(2,\mathbb{C})$

The use of $PSL(2,\mathbb{R})$ and $PSL(2,\mathbb{C})$ spaces rather than $SL(2,\mathbb{R})$ and $SL(2,\mathbb{C})$ changes entry and parameter distributions in ways that are at times significant. For all entries and parameters investigated distributions are indistinguishable for either of the two possible mappings $SL(2,\mathbb{C}) \mapsto PSL(2,\mathbb{C})$ and also $SL(2,\mathbb{R}) \mapsto PSL(2,\mathbb{R})$; scale invariance of distributions under normalisation is again seen with respect to the domain of matrix entries in $GL(2,\mathbb{R})$ and $GL(2,\mathbb{C})$ in all cases. However, distributions for real components of leading diagonal entries and otherwise differ in that while entry distributions for other than real components of leading diagonal entries are centred about 0, those for real components of leading diagonal entries are centred asymmetrically about 1. This means that while matrices in $SL(2,\mathbb{R})$ and $SL(2,\mathbb{C})$ are distributed about a null matrix in the embedding space of $GL(2,\mathbb{R})$ or $GL(2,\mathbb{C})$ respectively, those in $PSL(2,\mathbb{R})$ and $PSL(2,\mathbb{C})$ are distributed about the identity. This can be seen in the Figures (2.7) and (2.8), which are representative of the identical distributions. In both cases ($PSL(2,\mathbb{R})$ and $PSL(2,\mathbb{C})$) the distributions of real componens of other than leading diagonal entries are as for $SL(2,\mathbb{R})$ and $PSL(2,\mathbb{C})$ respectively, but the real components of leading diagonal entries are distributed asymmetrically about 1 and show a discontinuity at 0.

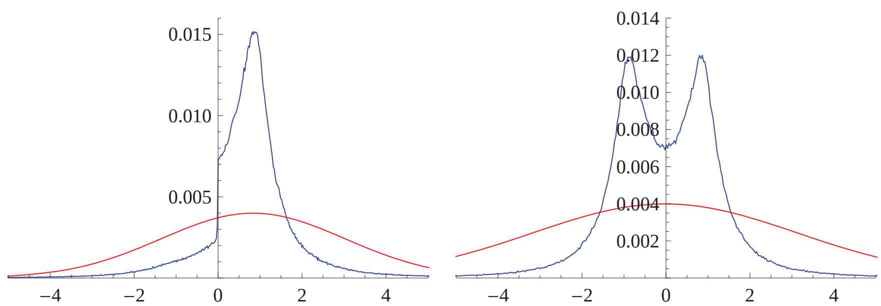


Figure 2.7: Distribution of real components of leading diagonal entries (left) and for all other components of entries (right) of 1,000,000 random matrices in $PSL(2,\mathbb{R})$

2.8 Distributions of Norms

It is not immediately obvious what form of norm should be used in the spaces $GL(2,\mathbb{C})$ or $SL(2,\mathbb{C})$; inner product and Euclidean norms, norms based on the identity, norms based on the adjugate, operator norms and chordal norms have all been considered.

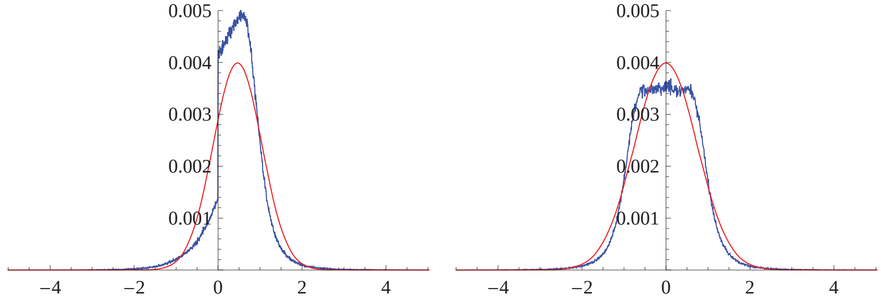


Figure 2.8: Distribution of real components of leading diagonal entries (left) and otherwise (right) of 1,000,000 random matrices in $PSL(2, \mathbb{C})$

2.8.1 The Chordal Norm

Beardon[7] and Gehring and Martin [5] use the chordal norm on the extended complex plane $\hat{\mathbb{C}}$ as defined in Copson [29]:

Definition of *Chordal Distance* between z_1 and z_2 in $\mathbb{C}$:

$$q(z_1, z_2) = \frac{2\,|z_1 - z_2|}{(|z_1|^2 + 1)^{1/2}(|z_2|^2 + 1)^{1/2}} \tag{2.5}$$

A metric on the group of Möbius transformations $\mathfrak{M}$:

$$d(f, g) = sup\{q(f(z), g(z)\ :\ z \in \mathbb{C}\} \tag{2.6}$$

A norm based on this metric, defined on $\mathfrak{M}$ with identity e:

$$d(f) = d(f, e) = sup\{q(f(z), z)\ :\ z \in \mathbb{C}\} \tag{2.7}$$

and finally, if functions f in $\mathfrak{M}$ are represented by matrices A in $SL(2, \mathbb{C})$, the norm is:

$$d(A) = d(A, I) = sup\{q(Az, z)\ :\ z \in \mathbb{C}\} \tag{2.8}$$

or

$$d^2(A) = sup\left\{\frac{4\,|(A - I)z|^2}{(|Az|^2 + 1)(|z|^2 + 1)}\ :\ z \in \mathbb{C}\right\} \tag{2.9}$$

2.8.2 Inner product and Euclidean Norms

Beardon [7] shows that the inner (scalar) product “on the vector space of all 2×2 matrices” satisfies the requirements of a norm, and it is easy to show that this is equivalent to the Euclidean norm in $\mathbb{R}^8$:

$$\|A\| = \sqrt{a^2 + b^2 + c^2 + d^2} \quad A \in \mathbb{C}^{2\times 2} \tag{2.10}$$

Here we investigate the distributions of Euclidean norms in the four and eight dimensional embedding spaces. From Hadamard's inequality, we have that for any 2×2 complex matrix A in the standard form of (1.1):

$$\|A\|^2 = |a|^2 + |b|^2 + |c|^2 + |d|^2 \geq 2\,|ad - bc| = 2\,|det(A)| \tag{2.11}$$

and since the norm is by definition positive:

$$\|A\| \geq \sqrt{2} \qquad \forall\, A \in SL(2, \mathbb{C}) \tag{2.12}$$

Further, we have for the distance between a matrix and the identity in the embedding spaces:

$$\|A - I\|^2 = \|A\|^2 - 2\ \mathrm{trace(A)} + 2 \tag{2.13}$$

hence

$$\|A - I\| \geq \sqrt{4 - 2\ \mathrm{trace(A)}} \qquad \forall\, A \in SL(2, \mathbb{C}) \tag{2.14}$$

The computational results that follow later in this section show that for random matrices in $SL(2, \mathbb{R})$ and $SL(2, \mathbb{C})$ these inequalities seem to have a high probability of holding with "near equality".

2.8.3 The Operator Norm

An *operator norm* is defined in terms of the magnification of unit entries under the action of an operation. Here operators are represented by matrices in $SL(2, \mathbb{C})$ which operate on complex quantities in $\hat{\mathbb{C}}$, and unit entries z in $\hat{\mathbb{C}}$ have magnitudes $|z| = 1$; they lie on the unit circle in $\hat{\mathbb{C}}$. That is,

$$\|A\|_{op} = sup_{(|z|=1)}\{\|Az\|\}, \quad z \in \hat{\mathbb{C}} \tag{2.15}$$

this form based on [31]. The following expression for the operator norm is then derived for $SL(2, \mathbb{C})$:

$$\|A\|_{op}^2 = \frac{\|A\|^2 + \sqrt{\|A\|^4 - 4}}{2}, \ A \in SL(2, \mathbb{C}) \tag{2.16}$$

As $\|A\| \to \infty,\ \|A\|_{op} \to \|A\|$ and $\|I\|_{op} = 1$ as would be expected.

2.8.4 Euclidean norms for matrices in $GL(2,\mathbb{R})$ and $GL(2,\mathbb{C})$

Distributions of standard Euclidean norms for uniformly distributed matrices in $GL(2,\mathbb{R})$ and $GL(2,\mathbb{C})$ identity centred spaces are shown in Figure (2.9). There are low probabilities of random matrices in $GL(2,\mathbb{R})$ or $GL(2,\mathbb{C})$ being close to the identity, and for an interval bound of 10^6 the expected value of the distance of a random matrix in $GL(2,\mathbb{R})$ to the identity is 1.1×10^6 while for $GL(2,\mathbb{C})$ it is 1.65×10^6.

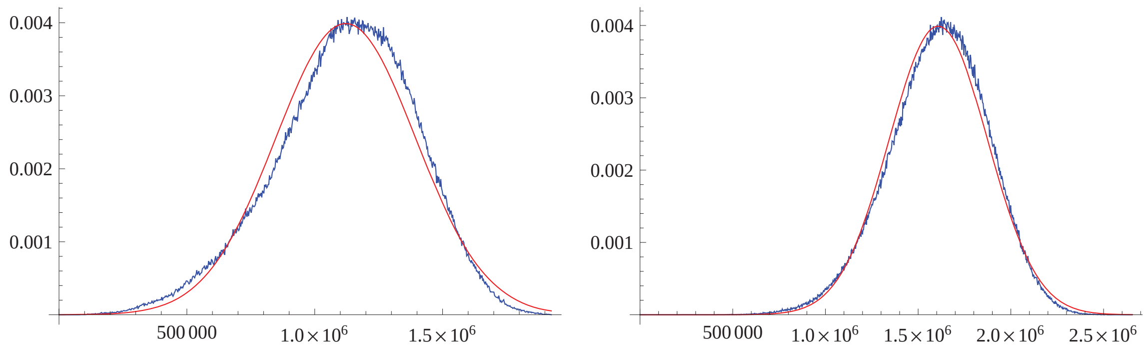


Figure 2.9: Distributions of $\|A - I\|$ for 1,000,000 uniformly distributed matrices A in, Left: $GL(2,\mathbb{R})$ and Right: $GL(2,\mathbb{C})$

2.8.5 Euclidean norms for matrices in $SL(2,\mathbb{R})$ and $SL(2,\mathbb{C})$

Distributions of Euclidean norms for uniformly distributed matrices in $SL(2,\mathbb{R})$ and $SL(2,\mathbb{C})$ spaces are shown in figure (2.10). The two distributions are far from identical, but agreement with Hadamard's identity is clear in both cases. Of interest is that the lower bound on the norms of $\sqrt{2}$ is sharp, with concentration of occurrences close to the bound. Distributions of Euclidean norms for $PSL(2,\mathbb{R})$ and $PSL(2,\mathbb{C})$ are identical respectively to those for $SL(2,\mathbb{R})$ and $SL(2,\mathbb{C})$.

Noting that the norm of the identity is $\sqrt{2}$, we consider now the distributions of distances between matrices in $SL(2,\mathbb{R})$ and the identity. Figure (2.11) shows the distributions in question for $SL(2,\mathbb{R})$, Figure (2.12) for $SL(2,\mathbb{C})$. It is clear that the probability is very high that matrices in $SL(2,\mathbb{C})$ are separated from the identity, but the qualifier "very" does not apply to matrices in $SL(2,\mathbb{R})$.

By inspection of the frequency data files for the events, the expectations of distances to the zero and identity origins are determined as:

- $E(\|A\|),\ A \in SL(2,\mathbb{R}) \quad = \ 2.63$

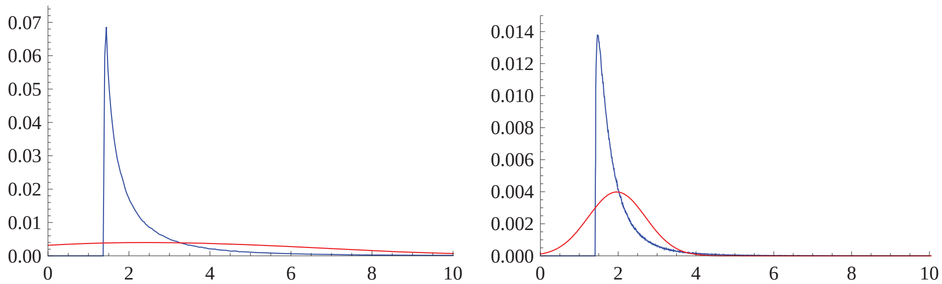


Figure 2.10: Distributions of $\|A\|$ for 1,000,000 uniformly distributed matrices A in, Left: $SL(2,\mathbb{R})$ and Right: $SL(2,\mathbb{C})$

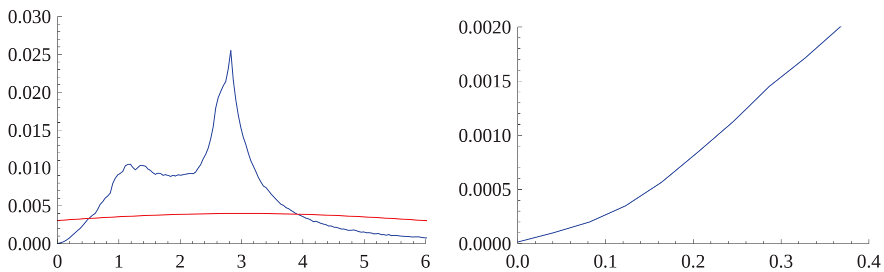


Figure 2.11: Distribution of the standard Euclidean norm, matrices in $SL(2,\mathbb{R})$ identity centred space, magnified on the right

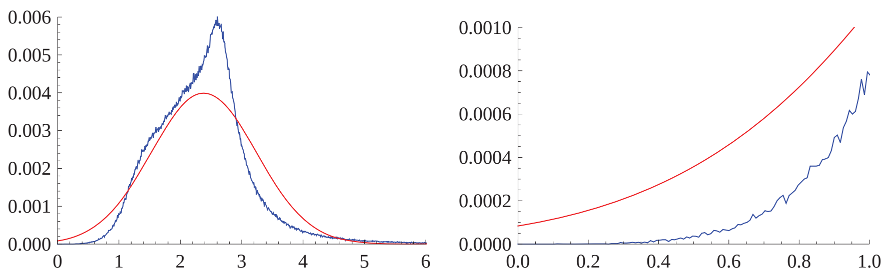


Figure 2.12: Distribution of the standard Euclidean norm, matrices in $SL(2,\mathbb{C})$ identity centred space, magnified on the right

- $E(\|A\|),\ A \in SL(2,\mathbb{C}) \quad = \ 2.94$
- $E(\|A - I\|),\ A \in SL(2,\mathbb{R}) \quad = \ 2.94$
- $E(\|A - I\|),\ A \in SL(2,\mathbb{C}) \quad = \ 2.36$

2.9 Distributions of Traces of matrices in $SL(2,\mathbb{C})$, $SL(2,\mathbb{R})$, $PSL(2,\mathbb{C})$ and $PSL(2,\mathbb{R})$

Distributions of traces of random matrices in $SL(2,\mathbb{C})$ and $PSL(2,\mathbb{C})$ are shown in figures (2.13) and (2.14), noting that the traces in $SL(2,\mathbb{R})$ and $PSL(2,\mathbb{R})$ are identical to those of the real components of the trace in the corresponding complex groups.

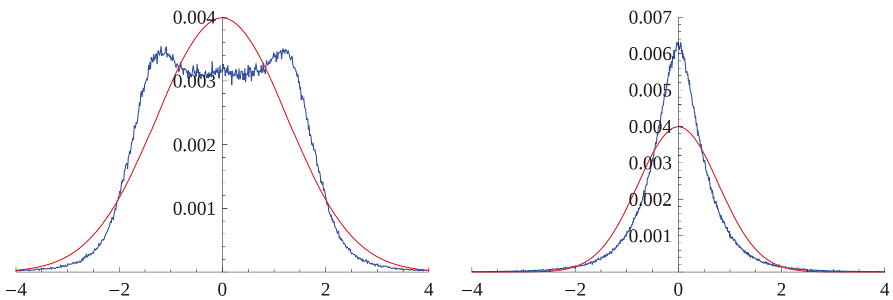


Figure 2.13: Distribution of (Left:) real and (Right:) imaginary components of the trace of a matrix in $SL(2,\mathbb{C})$

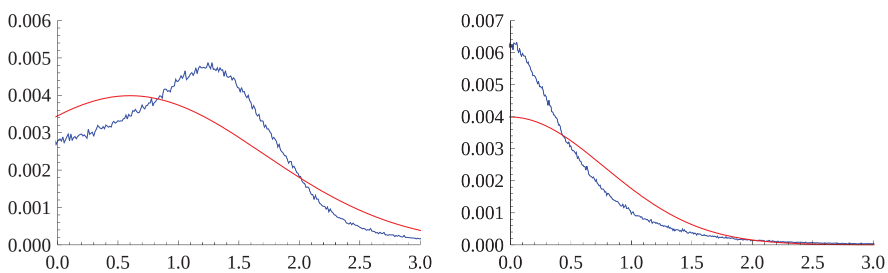


Figure 2.14: Distribution of (Left:) real and (Right:) imaginary components of the trace a matrix in $PSL(2,\mathbb{C})$

From Equation (2.14) we see that the trace of a matrix in $SL(2,\mathbb{C})$ determines the lower bound of the distance between the matrix and the identity. The mean trace in $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$ is 0 and we conclude that this inequality is often sharp.

2.10 Inequalities

Here we derive and test some inequalities that might provide a means of assessing the nature of subgroups of matrices close to the identity, further work will be reported in a subsequent thesis.

2.10.1 Commutator Inequalities

For matrices in single generator subgroups, the commutator is the identity since all entries commute. For matrices in two-generator subgroups we first make the assumption that $A \neq B$; then if $A, B \approx I$ (according to 1.4), since $[A, B] = ABA^{-1}B^{-1}$ is in the subgroup it is closer to I than either A or B, and for any such pair of matrices a matrix can always be found within the subgroup that is closer to I. This would imply that no two-generator subgroups are discrete whenever the generators $A, B \approx I$, which would make it important to know under what conditions $A, B \approx I$.

Figures 2.15 and 2.16 show the distributions of $sup\{\|A - I\|, \|B - I\|\}$ for $PSL(2, \mathbb{R})$ and $PSL(2, \mathbb{C})$ respectively where $A, B \approx I$ (according to 1.4), and inspection of the experimental event files reveals that while for $PSL(2, \mathbb{R})$ there are no recorded occurrences for the first 10 quantiles (≤ 0.1), there are none for the first 91 quantiles (≤ 0.5) for $PSL(2, \mathbb{C})$. Most matrix pairs in $PSL(2, \mathbb{R})$ which are close to the identity have one component matrix at a distance 3 from the identity, while those in $PSL(2, \mathbb{C})$ have one component at a distance 2.55 from the identity. We note that in both cases the distributions for $A, B \approx I$ and $A, B \not\approx I$ are not disjoint.

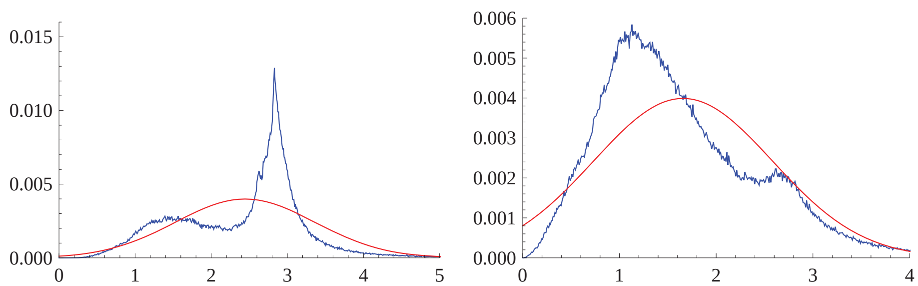


Figure 2.15: Left: Distribution of $\sup\{\|A - I\|, \|B - I\|\}$ for pairs of matrices $A, B \approx I$ (according to 1.4) in $PSL(2, \mathbb{R})$, Right: Distribution of $\inf\{\|A - I\|, \|B - I\|\}$ for pairs of matrices $A, B \not\approx I$ in $PSL(2, \mathbb{R})$

It would have been nice to have established that a simple criterion for the discreteness of two-generator subgroups based on the magnitudes of $\|A - I\|$ and $\|B - I\|$ applied, but unfortunately this is not the case. Accordingly we now look at derivation of inequalities relating to the distance between the commutator and the identity.

We now derive two inequalities in $SL(2, \mathbb{C})$. First:

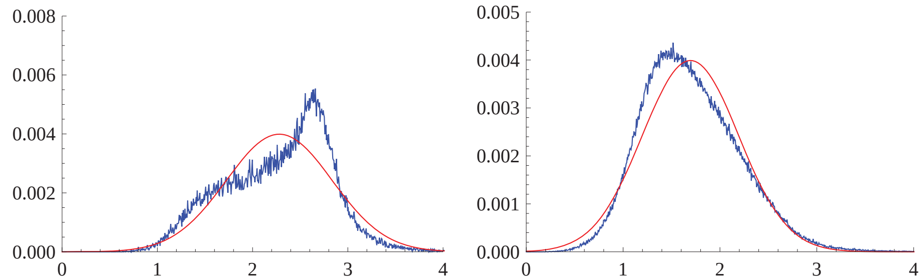


Figure 2.16: Left: Distribution of $\sup\{\|A-I\|,\|B-I\|\}$ for pairs of matrices $A, B \approx I$ (according to 1.4) in $PSL(2,\mathbb{C})$, Right: Distribution of $\inf\{\|A-I\|,\|B-I\|\}$ for pairs of matrices $A, B \not\approx I$ in $PSL(2,\mathbb{C})$

$$\begin{aligned}(B-I)(A^{-1}-I) &= \quad BA^{-1}-A^{-1}-B+I \\ (A^{-1}-I)(B-I) &= \quad A^{-1}B-A^{-1}-B+I \\ \Rightarrow BA^{-1}-A^{-1}B &= \quad (B-I)(A^{-1}-I)-(A^{-1}-I)(B-I)\end{aligned} \tag{2.17}$$

and since

$$[A,B]-I = ABA^{-1}B^{-1}-I = A(BA^{-1}-A^{-1}B)B^{-1} \tag{2.18}$$

then

$$[A,B]-I = A(B-I)(A^{-1}-I)B^{-1}-A(A^{-1}-I)(B-I)B^{-1} \tag{2.19}$$

Taking norms on both sides and applying the Cauchy-Schwarz inequality and the triangle inequality, since $\|A\| = \|A^{-1}\|$ and $\|A-I\| = \|A^{-1}-I\| \ \forall \ A \in \ GL(2,\mathbb{C})$ and norms are scalars, we have:

$$\|[A,B]-I\| \leq 2\,\|A\|\,\|B\|\,\|A-I\|\,\|B-I\| \tag{2.20}$$

The second inequality is derived from (2.20) using the reverse triangle inequality,

$$\begin{gathered}\|A-I\| \geq |\|A\|-\|I\|| \\ \Rightarrow \|A-I\|^2 \geq \|A\|^2+\|I\|^2-2\,\|A\|\,\|I\| = \|A\|^2-2\sqrt{2}\,\|A\|+2 \\ \Rightarrow \|A\|^2 \leq \|A-I\|^2+2\sqrt{2}\,\|A\|-2\end{gathered} \tag{2.21}$$

we square both sides of (2.20):

$$\|[A,B]-I\|^2 \leq 4\,\|A\|^2\,\|B\|^2\,\|A-I\|^2\,\|B-I\|^2 \tag{2.22}$$

and substitute for the square of the norms as derived in (2.21):

$$\|[A,B]-I\|^2 \;\leq\; 4(\|A-I\|^2{+}2\sqrt{2}\,\|A\|{-}2)(\|B-I\|^2{+}2\sqrt{2}\,\|B\|{-}2)\,\|A-I\|^2\,\|B-I\|^2 \tag{2.23}$$

For $\|A-I\|,\|B-I\|$ small, we can neglect the squares of these terms inside the brackets:

$$\|[A,B]-I\|^2 \;\leq\; 4(2\sqrt{2}\,\|A\|-2)(2\sqrt{2}\,\|B\|-2)\,\|A-I\|^2\,\|B-I\|^2 \tag{2.24}$$

and finally, by simplifying and taking positive square roots we arrive at the inequality:

$$\|[A,B]-I\| \;\leq\; 4\,\|A-I\|\,\|B-I\|\,\sqrt{(\sqrt{2}\,\|A\|-1)(\sqrt{2}\,\|B\|-1)} \tag{2.25}$$

Experimental analysis (see Figure 2.17) shows (2.25) to be a sharper inequality than (2.20), especially for random matrices in $PSL(2,\mathbb{R})$ (2.25).

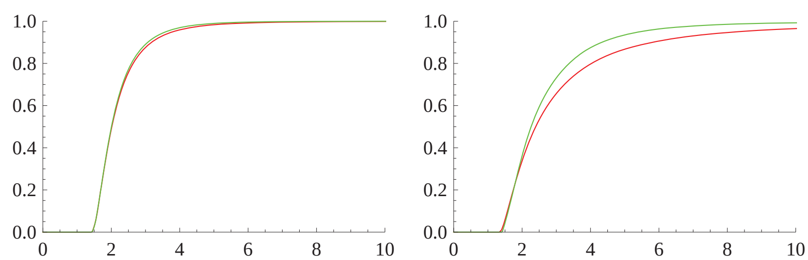


Figure 2.17: Comparison of inequalities in (2.20) and (2.25)for both $PSL(2,\mathbb{C})$ and $PSL(2,\mathbb{R})$: Overlaid cumulative distributions for $sup\{\|A\|,\|B\|\}$, red = inequality(2.20), green = inequality (2.25), Left: $PSL(2,\mathbb{C})$, Right: $PSL(2,\mathbb{R})$

2.10.2 The K Inequality

The inequality in Equation (1.4) can also be tested by defining a parameter K:

$$K = \frac{\|[A,B]-I\|}{\|A-I\|\,\|B-I\|} \tag{2.26}$$

We note from Figure (2.18) that the distributions for $PSL(2,\mathbb{R})$ and $PSL(2,\mathbb{C})$ are qualitatively different. For matrices in $PSL(2,\mathbb{R})$ the expectation is $K = 2.45$

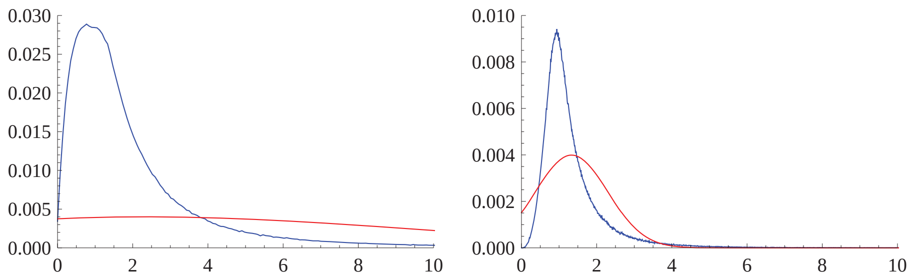


Figure 2.18: Distribution of parameter K for $1,000,000$ pairs of uniformly random matrices A, B in, Left: $PSL(2,\mathbb{R})$, Right: $PSL(2,\mathbb{C})$

with probability peaking at $K = 0.94$ and falling towards 0 with non zero occurrences in each quantile. For matrices in $PSL(2,\mathbb{C})$ the expectation is $K = 1.133$ with probability peaking at $K = 0.77$ and falling towards 0 with zero observed occurrences in the first three quantiles.

If we consider K values for just those A, B pairs for which $A, B \approx I$ (according to 1.4), then for matrices in $PSL(2,\mathbb{R})$, 0.861 of occurrences are in the quantile containing 0 while for matrices in $PSL(2,\mathbb{C})$ 0.868 of occurrences are in the quantile containing 0. However, for those A, B pairs for which $A, B \not\approx I$, 0.267 of occurrences for matrices in $PSL(2,\mathbb{R})$ are in the quantile containing 0 while only 0.49×10^{-2} of occurrences for matrices in $PSL(2,\mathbb{C})$ are in the quantile containing 0.

While these observations are of interest, further study is required to assess the usefulness of the K parameter in determining discreteness.

2.11 Discrete Groups

This section though short contains probably the most important result, that discrete groups of Möbius transformations seem to be much rarer than might be supposed.

Distributions are analysed for experimental events derived as occurrences of random matrices in $PSL(2,\mathbb{R})$ and $PSL(2,\mathbb{C})$ generated via by either of the two mappings ϕ and ψ (discussed earlier) from matrices in $SL(2,\mathbb{R})$ and $SL(2,\mathbb{C})$, these matrices derived via normalisation as described from zero centred matrices in $GL(2,\mathbb{R})$ and $GL(2,\mathbb{C})$ having entries uniformly distributed over $[-10^6, 10^6]$.

Inspection of event data files for processed experimental data reveals the following:

- The probability that matrices in $PSL(2, \mathbb{C})$ have disjoint isometric circle discs: 0.00394
- The probability that matrices in $PSL(2, \mathbb{R})$ have disjoint isometric circle discs: 0.00844
- The probability that matrices in $PSL(2, \mathbb{C})$ are discrete by Jørgensen's inequality (Theorem 1.2): 0.993
- The probability that matrices in $PSL(2, \mathbb{R})$ are discrete by Jørgensen's inequality (Theorem 1.2) 0.998

What was not really expected was that the experimentally determined probabilities of occurrence of discrete groups in $SL(2, \mathbb{C})$ meeting the isometric circle criterion be extremely low (0.004) and of occurrence of groups meeting Jørgensen's criterion (Theorem 1.2) would be extremely high (0.99).

Chapter 3

MATHEMATICAL ANALYSIS OF DISTRIBUTIONS

Of central importance to the algebra of distributions (where the algebraic entity is a distribution of a random variable rather than the variable itself) is the idea of convolution of two piecewise continuous functions f and g, which can be defined on an interval (possibly infinite) $[a, b]$ as:

$$f * g = \int_a^b f(x)\, g(t - x)\, dx \tag{3.1}$$

where the convolution $f * g$ is a function of t and the interval limits a and b are possibly functions of t. The Convolution Theorem referred to later in this section can found in [21].

Mathematical analyses of convolutions of distributions, particularly with regard to sums, differences, products and quotients have been performed in the past and limited results have been applied to the calculation of distributions of determinants of real matrices and hence of inverse matrices. However, the problem is significantly greater for the normalisation of matrices representing Möbius transformations for two reasons. Firstly, if the matrices are complex then what is involved is the division of a random distribution by the complex square root of distributions. Secondly, the convolutions of distributions of random variables that can take negative values as well as positive is vastly more complicated than for the non negative case, and the mathematics for the more general case has not been previously derived. A high proportion of the published work involves distributions of random variables over the real domain $[0, 1]$. We commence the analysis of these more general convolutions here (including some comparisons with experimental events) and a mathematical outline of what is required for a more complete solution is given.

The prime interest in the distribution of the determinant here is because it determines the homomorphisms between the general linear and special linear matrix groups, and this provides motivation for the following literature review.

3.1 A brief review of the literature on distributions of determinants

Williamson and Downs [13] have accessed many of the papers and provide information that allows their scope and relevance to be assessed, this is of importance since a significant number of the relevant papers are old and unattainable. Williamson and Downs (in 1989) state that “There are very few results on the distribution of random determinants” and unfortunately this opinion seems still valid. That the literature review below shows the information published to date to be not highly relevant to the task in hand is supported by closer inspection of the algebra of distributions of random numbers on finite intervals.

Fortet([14] (1951), “Random Determinants” published in a National Bureau of Standards (US) research journal, the particular volume is missing from the online collection; Williamson and Downs imply that “a few special cases” are covered.

Nyquist, Rice and Riordan ([15] 1954), “On the Distribution of Random Determinants” published in the Quarterly of Applied Mathematics of Brown University. This journal is now distributed through the American Mathematical Society, but no early issues are available online. According to Williamson and Downs, the authors derive precise expressions for 2×2 matrices with “normally distributed entries with zero means”.

Komlos ([16] 1967) published several papers in a Hungarian journal, no early editions of which are available online; Williamson and Downs again imply that “a few special cases” are covered.

Alagar ([17] 1978) presents results for determinants with exponential distributions.

Williamson and Downs ([13] 1988) themselves consider the determinant of a matrix with uniformly distributed entries over $[0, 1]$, their motivation being to derive the distribution of the inverse of such a matrix.

Wise and Hall ([18] 1991) write in order to criticise Williamson and Downs’s use of a series expansion and question the conclusions.

More recent papers appear to concentrate on particular scientific rather than pure mathematical applications (e.g. [19]), while [13] still appears to be the most relevant despite relating only to determinant entry distributions over $[0, 1]$. The paper does not address complex determinants and the focus is on the determinant rather than the square root of the determinant. What is really needed is an expression for the square root of the determinant of a complex 2×2 matrix with uniformly distributed zero centred entries over a more general domain in $\mathbb{R}$, and then $\mathbb{C}$. No attempt to derive such an expression has been found in the literature, and a foundation will be established here, with a more complete determination to follow in a later thesis.

3.2 Square Root of a Complex Determinant

Let

$$A = \begin{pmatrix} a+i\alpha & b+i\beta \\ c+i\gamma & d+i\delta \end{pmatrix} \tag{3.2}$$

where the matrix entry components are $a, b, c, d, \alpha, \beta, \gamma, \delta \in \mathbb{R}$, and then:

$$\mathrm{det(A) = (ad - bc) - (\alpha\delta - \beta\gamma) + i[(a\delta - \beta c) + (\alpha d - b\gamma)]} \tag{3.3}$$

The following expression then needs to be evaluated:

$$\sqrt{\mathrm{det(A)}} = \sqrt{(ad - bc) - (\alpha\delta - \beta\gamma) + i[(a\delta - \beta c) + (\alpha d - b\gamma)]} \tag{3.4}$$

before performing a further convolution for complex division. As a first step then, it is noted that in (3.4) the descriminant contains sums and differences of four real component expressions of a form similar to $ad - bc$ (which is the determinant of a real matrix expressed in the standard form (1.1).

3.3 Random Variables and Distributions

Statistical terminology will be based on Papoulis and Pillai [23] and Springer [20].

Concept	**symbol**	**comment**
Experiment	S	−e.g. uniformly random matrices in $\mathrm{SL}(2, \mathbb{C})$
Event	T	−e.g. $1,000,000$ specific matrices A
Outcome	ξ	−e.g. index $\mathrm{i} \in [0, 999999]$
Random variable	$X(\xi)$	−e.g. complex matrix $\mathrm{A_i}$
Functional map	X	−analytic representation of a random variable
Set of bounded outcomes	$\{X \leq x\}$	−all experimental outcomes ξ for which $\mathrm{X}(\xi) <$ the number x
	$\{x_1 \leq X \leq x_2\}$	−...$x_1 \leq X(\xi) \leq x_2$

In this thesis an experiment is defined by a computer program which encapsulates an iteration of a random number generating algorithm that supplies arguments for specific defined functions. Of the set S of all possible outcomes of an experiment, actual events T are recorded from specific program runs. S is the domain of a random variable function X, a particular instance of which from the event $T \subset S$ is $X(\xi)$, for example a random matrix in $SL(2, \mathbb{C})$ or a random real number depending on the experimental context. Note that the term *distribution* is unqualified, we will express distributions as both cumulative probability functions (*c.d.f*) and probability density functions (*p.d.f*) as required. Every experiment has imposed conditions, but there is a minimum requirement for a random variable that the set of bounded outcomes exist (which bounds may include $-\infty$ and ∞) and that whatever the bounds the probability of occurrence of events outside those bounds be zero, with the concomitant requirement that the probability be unity within the bounds. The c.d.f of the random variable X is defined over the domain S as a function of a moveable upper bound x:

$$F_x(x) = P(\{X \leq x\}) \tag{3.5}$$

and is necessarily non negative, and its derivative is the p.d.f. If the c.d.f $F_x(x)$ is linear over the domain then the p.d.f $f_x(x)$ is constant and the distribution is *uniform.*

Given an expression $f(u)$ for a non negative monotonically increasing function, then treated as a p.d.f its integral is the cumulative probability:

$$F_x(x) = \int_{-\infty}^{x} f(u)du \tag{3.6}$$

provided that the total integral of $F_x(x)$ over the entire domain of S evaluates to unity. For two random variables X and Y distributed over $S \times S$ and moveable bounds x and y, the *joint p.d.f* is $f_{x,y}(x, y)$ s.t.

$$F_{x,y}(x, y) = P(\{(X, Y\}) \in B = \int \int_B f(x, y)dxdy \tag{3.7}$$

where $B \subset S \times S$ is a rectangle bounded above by x and y. If $S = [-k, k] \subset \mathbb{R}$, then $B = [-k, x] \times [-k, y]$.

3.4 Algebra of Random Distributions

The algebraic calculations on real number components of matrix entries involved in normalising a $GL(2, \mathbb{C})$ matrix and calculating trace, determinant, norms and other parameters for functions of matrices determine the elements that are required in an adequate algebra of distributions.

Much of the work in deriving analytical algebraic expressions for distributions of random numbers was not completed until the 1960s (Springer [20] has a good review), and again no reference is found to the derivation of the square root of a distribution, while other published results are in the main for real intervals between 0 and 1.

The reason that most of the literature concentrates on non negative domains is that the nature of the calculation changes qualitatively whenever the domain crosses 0. For just two variables x and y distributed over a real domain $\Delta = [-k, k]$ mapped by a binary operation $\oplus$ to a two dimensional domain $\Delta \times \Delta$, the resultant distribution is over four quartiles. Whenever portions of domains in the four quartiles are disjoint then those portitions may be summed algebraically, otherwise the summation must be achieved via convolution.

Springer describes the historical work and uses Fourier and Mellin convolutions to derive expressions for sums, differences, products and quotients for some specific cases for positive domains, his work in chapters three and four is very instructive of the basic techniques. What is required further is a generalisation of of the intervals used as well as derivation of the square root of a determinant in order to normalise a matrix. All the work that follows in this chapter is independant of Springer's contributions, he has other objectives.

3.5 Convolutions of Random Distributions

Suitable integral transforms and the convolution theorem can be applied to statistical frequency analysis as long as the distributions of interest are either defined on finite intervals or have 'tails' that converge rapidly to zero. Hence from (3.7) the convolution theorem allows the joint p.d.f of two such independent identical distributions to be calculated:

$$f(x, y) = f(x)f(y) \tag{3.8}$$

Convolution is plainly a binary operation which is closed on the set of integrable functions. Considering the integration by components rule, since $d(t-x) = -dx$ it is easy to show from (3.1) that convolution is commutative, associative and distributive. In the case of probability density functions, the integral transforms involved have inverses and the Dirac delta function acts as an identity for convolution (refer [20]). Hence probability density functions form a commutative group under convolution and expressions do exist for such derived functions as the square root of a distribution.

All distributions analysed in this chapter will be generally presumed to be positive and with even symmetry about a mean, with the understanding that not all matrix entry distributions will meet this criterion. The purpose is to provide some analysis of matrix parameter distributions where such is lacking at present, and all marginal (as opposed to joint) pre-convolution distribution functions that are of interest will be of independent and identically distributed variables unless otherwise stated. Though the initial theory is based on Springer [20], he does not derive any of these detailed distributions.

3.6 Sums and differences of Distributions

Let the probability density function of the sum of two identical distributions $f(x), f(y)$ over $\mathbb{R}$ be $g(w)$ where $w = x + y$, then from (3.1) and (3.8) $g(w)$ can be expressed by the Fourier convolution:

$$g(w) = \int_{-\infty}^{\infty} f(w-y)f(y)dy \tag{3.9}$$

But we require x and y to be distributed over finite domains, so let $x, y \in [-k, k] \subset \mathbb{R},\ k > 0$, then $-2k \leq w = x + y \leq 2k$, and

$$-k \leq x \leq k \Rightarrow -k \leq w - y \leq k \Rightarrow w - k \leq y \leq w + k \tag{3.10}$$

which allows derivation of the finite integration limits for (3.9):

(i) if $w \geq 0$: then $w - k \leq y \leq w + k \Rightarrow w - k \leq y \leq k$
(ii) if $w \leq 0$: then $w - k \leq y \leq w + k \Rightarrow -k \leq y \leq w + k$

An expression for the sum of two identical distributions as a piecewise function follows:

$$g(w) = \begin{cases} \int_{-k}^{w+k} f(w-y)f(y)dy & -2k \leq w \leq 0 \\ \\ \int_{w-k}^{k} f(w-y)f(y)dy & 0 \leq w \leq 2k \end{cases} \tag{3.11}$$

For $g(w)$ to be a valid p.d.f the functional requirements stated earlier must be satisfied, it should be noted that these do not preclude the existence of a singularity at $w = 0$.

The probability density function $g(w)$ for the difference between two identical distributions $f(x), f(y)$ over $\mathbb{R}$ where $w = x - y$ can be determined in a similar fashion. From (3.8) $g(w)$ can be expressed by the Fourier convolution:

$$g(w) = \int_{-\infty}^{\infty} f(w+y)f(y)dy \tag{3.12}$$

But x must be distributed over a finite domain, so let $x, y \in [-k, k] \subset \mathbb{R},\ k > 0$, then $-2k \leq w \leq 2k$, and

$$-k \leq x = w + y \leq k \Rightarrow -k - w \leq y \leq k - w$$

The integration limits for (3.12) are required:

(i) if $w \geq 0$: then $-k \leq y \leq k - w$
(ii) if $w \leq 0$: then $-k - w \leq y \leq k$

hence an expression for the difference between two identical distributions over $[-k, k]$ as a piecewise function can be determined as:

$$g(w) = \begin{cases} \int_{-k-w}^{k} f(w+y)f(y)dy & -2k \leq w \leq 0 \\ \\ \int_{-k}^{k-w} f(w+y)f(y)dy & 0 \leq w \leq 2k \end{cases} \tag{3.13}$$

We now define a uniform distribution over the interval $[-k, k]$ as:

$$\mathfrak{D}_U(x) = \begin{cases} \frac{1}{2k} & \forall x \in [-k, k] \\ \\ 0 & \text{otherwise} \end{cases} \tag{3.14}$$

but whereas the domain of $f(x)$ and $f(y)$ is $[-k, k]$, the domain of $g(w)$ in (3.11) and (3.13) is $[-2k, 2k]$; integration and evaluation over the established bounds yields identical expressions for the sum and difference of two uniform distributions over $[-k, k]$ respectively:

$$\mathfrak{D}_{US}(w) = \mathfrak{D}_{UD}(w) = \frac{2k - |w|}{4k^2} \qquad -2k \leq w \leq 2k \tag{3.15}$$

Then for uniform functions $\mathfrak{D}_U(x)$ over $[-k, k]$ the resultant distributions for sum and difference are identically triangular over $[-2k, 2k]$, and it can be seen in figure (3.1) that the total area under the piecewise union of the lines between $-2k$ and $2k$ for $k = 1$ is unity and that $g(w)$ in (3.15) is indeed a valid p.d.f. This result is consistant with Springer's more limited example, which is for functions over $[0, 1]$.

The situation with identical normal distributions is more complicated, not least because of the constraint of necessarily finite domains. If σ is the standard deviation then an expression for a standard normal p.d.f which has a mean of 0 can be written:

$$\mathfrak{D}_N(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{x^2}{2\sigma^2}} \tag{3.16}$$

We note that if this distribution is constrained to the interval $[-k, k]$, then the integral of $\mathfrak{D}_N(x)$ over that interval is $Erf(\frac{k}{\sqrt{2}})$, later discussion in this section shows this only tends to unity under certain conditions.

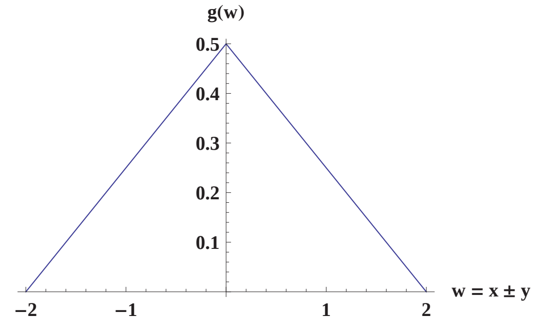


Figure 3.1: Convolution of Sum or Difference of Uniform Distributions over $[-1, 1]$

An expression for the sum of two identical normal distributions of variables which are constrained to domains $[-k, k]$ can be derived from (3.11):

$$g(w) = \begin{cases} \frac{1}{2\pi\sigma^2}\int_{-k}^{w+k} e^{-\frac{2y^2-2yw+w^2}{2\sigma^2}} dy & -2k \leq w \leq 0 \\ \\ \frac{1}{2\pi\sigma^2}\int_{w-k}^{k} e^{-\frac{2y^2-2yw+w^2}{2\sigma^2}} dy & 0 \leq w \leq 2k \end{cases} \tag{3.17}$$

Evaluation of the integrals results in the following p.d.f for the sum of constrained identical normal distributions:

$$\mathfrak{D}_{NS}(w) = \begin{cases} \frac{1}{2\sigma\sqrt{\pi}} e^{-\frac{w^2}{4\sigma^2}} Erf(\frac{2k+w}{2\sigma}) & -2k \leq w \leq 0 \\ \\ \frac{1}{2\sigma\sqrt{\pi}} e^{-\frac{w^2}{4^2}} Erf(\frac{2k-w}{2\sigma}) & 0 \leq w \leq 2k \end{cases} \tag{3.18}$$

It is seen that $\mathfrak{D}_{NS}(w)$ is a function of k as well as of σ, the extent of this multi-variable relationship can be seen by inspection of the 3-dimensional plot for $\sigma = 1$ in Figure (3.2).

A conclusion is that the sum of constrained normal distributions is only normal for a domain sufficiently large with respect to the standard deviation. This contention is at apparent odds with the oft quoted result that the sum of gaussian distributions is gaussian, see for example [24] where Weisstein states “Amazingly, the distribution of a sum of two normally distributed independent variates... is another normal distribution”. All the proofs that this author has been able to peruse have in common the assumption that the distributions run from $-\infty$ to $+\infty$; this is impossible to accomplish for distributions generated over finite intervals, the best that can be done is to attempt to use large enough initial intervals and keep track of the effect of successive convolutions. These do not preserve distributions for any of the groups $SL(2, \mathbb{C})$,

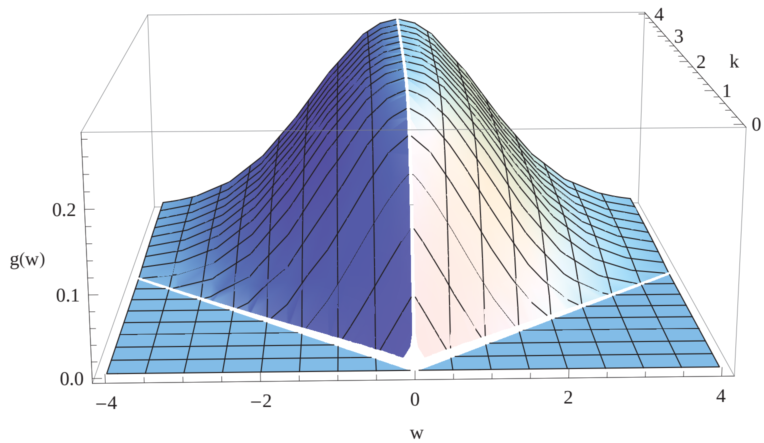


Figure 3.2: Sum of Normal Distributions over the domain interval $[-k, k]$ for $\sigma = 1$

$SL(2, \mathbb{R})$, $PSL(2, \mathbb{C})$ or $PSL(2, \mathbb{R})$, but it is much more difficult to generate normally distributed matrices in these groups anyway.

We need an expression which allows us to determine the resultant of the sum of two identical constrained normal distributions over an interval, and it is apparent from figure (3.2) that for a standard deviation of 1, the resultant distribution depends critically on the width of the interval. Given such a distribution, the total probability within the range $[-2k, 2k]$ only approaches unity for values of the interval bound $k \approx 3$ or greater, and similar considerations apply to all values of σ. Figure (3.3) shows how the calculated probability given by (3.18) varies with the width of the interval $[-k, k]$ for standard deviation $\sigma = 1$ and $\sigma = 3$. The sum of identical constrained normal distributions also depends critically on the standard deviation of those distributions, the interval bound k must be ≈ 9 or more to assure total probability of close to 1 in $[-2k, 2k]$.

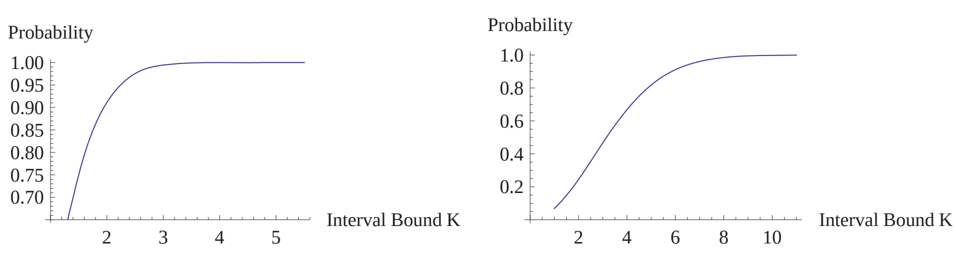


Figure 3.3: Variation of calculated probability for the sum of two normal distributions, Left: $\sigma = 1$, Right: $\sigma = 3$

The p.d.f for the sum of two identical constrained normal distributions for constant k and σ is now considered by setting $k = 1$, $\sigma = 1$. Figure (3.4) shows the left and right component convolution functions and the resultant piecewise p.d.f for the sum of two constrained normal distributions with standard deviation 1 and interval bound $k = 1$. The portion of the curve (Left) between $w = -2$ and $w = 0$ is combined with the portion of the curve (Centre) between $w = 0$ and $w = 2$, to get the final piecewise function $g(w)$ (Right).

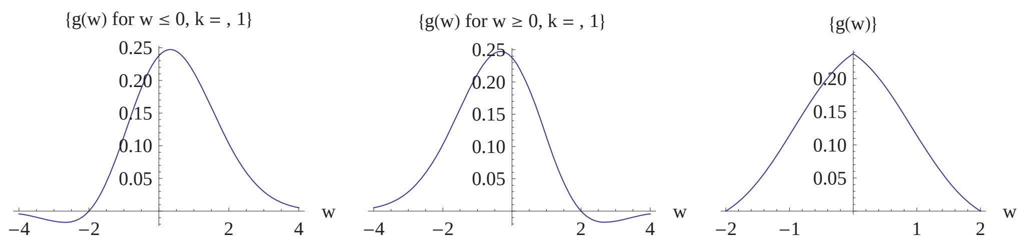


Figure 3.4: Convolution functions, Left to Right: left, right then piecewise sum, $g(w) = \frac{1}{2\sqrt{\pi}} e^{-\frac{w^2}{4}} Erf(k + \frac{w}{2})$, $\sigma = 1$, $k = 1$

For the first two portions of the figure the domain has been extended to show clearly how the final function over $[-2k, 2k]$ (in this case $[-2, 2]$) is obtained from its left and right components. The combined piecewise function is everywhere positive over $[-2, 2]$ but the integrated area under the curve for this domain is 0.466 for $(k, \sigma) = (1, 1)$, not unity since the convolution is of constrained normal functions, the resultant function otherwise meeting all the requirements for a p.d.f. As the domain bound k is increased, the total integrated area tends to unity and the resultant function to a normally distributed p.d.f. For $(k, \sigma) = (1, 1)$ the p.d.f of the sum of two normal distributions is most definitely not normal. The error function is able to be regarded as a measure of the departure of a derived probability function for normal distributions constrained to finite domains from a true p.d.f, and accordingly the term p.d.f will be used even for constrained domain results. This is in effect a generalisation of the usual definition of p.d.f.

Now from (3.13) the difference betweeen two identical normal distributions constrained to $[-k, k]$ is given by:

$$g(w) = \begin{cases} \frac{1}{2\pi\sigma^2} \int_{-k-w}^{k} e^{-\frac{2y^2+2yw+w^2}{2\sigma^2}} dy & -2k \le w \le 0 \\ \\ \frac{1}{2\pi\sigma^2} \int_{-k}^{k-w} e^{-\frac{2y^2+2yw+w^2}{2\sigma^2}} dy & 0 \le w \le 2k \end{cases} \tag{3.19}$$

Evaluation of these integrals yields a p.d.f for the difference between two identical normal distributions which is precisely the same as that for the sum of two identical normal distributions in (3.18):

$$\mathfrak{D}_{ND}(w) = \mathfrak{D}_{NS}(w) \tag{3.20}$$

Figure (3.5) which shows the error function component of equation (3.18) clarifies how closely the derived probability for sum and difference of constrained normal distributions approaches a Gaussian curve.

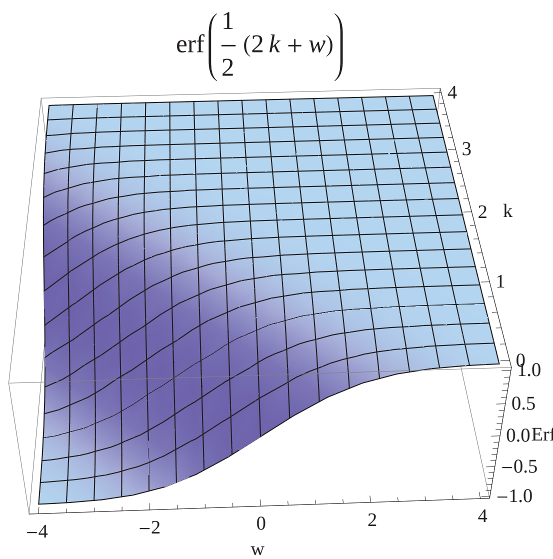


Figure 3.5: Error function $Erf(\frac{2k+w}{2})$ as in the left component (in Figure 3.4) of of the p.d.f expression for the sum of identical constrained normal distributions

It can be seen that for sufficiently large k the error function factor of the left component (in Figure 3.4) of the p.d.f for the sum or difference of two normal distributions tends to 1 (as does that of the right component (in Figure 3.4)), the result also applies qualitatively for larger standard deviations. Hence for sums and differences of normal distributions, from Equation (3.18):

$$\mathfrak{D}_{ND}(w) = \mathfrak{D}_{NS}(w) \rightarrow \frac{1}{2\sigma\sqrt{\pi}} e^{-\frac{w^2}{4\sigma^2}} \text{ for sufficiently large k} \tag{3.21}$$

Since this is an expression for a Gaussian distribution, we conclude that the distributions of sums and differences of variables from identically normal constrained distributions can be regarded as normal for domain intervals sufficiently large compared to the standard deviations of the distributions, but definitely not otherwise.

As has been noted earlier, the result for the domain bound k not large appears contrary to general perception, but distributions derived from finitely bounded domains are commonly encountered, especially with experimentally determined data. An a priori assumption cannot be made that resultant sum or difference distributions are normal given pre-image functions with normal distributions restricted to finite domains. Even with judicious choice of domain size compared to standard deviation, the convolutions involved in generating matrix parameters can result in distributions that deviate markedly from an expected Gaussian form.

3.7 Products and quotients of Distributions

Here some important results are derived, most up to formulation of the integrals, and some are evaluated completely. A method of evaluation is established, with the more complicated examples left to a subsequent thesis.

The following variable and functional transformations are made in the expression (3.9) for the sum of two identical distributions, noting that logarithmic transforms can be applied to the positive intervals only and that w is on the way to becoming xy:

$$\begin{array}{lcll} \mathrm{x} & \rightarrow & \mathrm{log(x)} & x \geq 0 \\ \mathrm{y} & \rightarrow & \mathrm{log(y)} & y \geq 0 \\ \mathrm{w = x + y} & \rightarrow & \mathrm{log(x) + log(y) = log(xy) = log(w)} & w \geq 0 \end{array} \tag{3.22}$$

then by substituting into (3.9):

$$\begin{aligned} g(log(w)) &= \int_0^\infty f(log(w) - log(y)) f(log(y)) d\,log(y) \qquad x, y, w \geq 0 \\ &= \int_0^\infty \frac{1}{y} f\left(log\left(\frac{w}{y}\right)\right) f(log(y)) dy \qquad x, y, w \geq 0 \end{aligned} \tag{3.23}$$

we now simply transform the logarithmic functions:

$$\begin{array}{lcl} \mathrm{f(log(.))} & \rightarrow & \mathrm{f(.)} \\ \mathrm{g(log(.))} & \rightarrow & \mathrm{q(.)} \end{array} \tag{3.24}$$

and the resultant expression is the probability density function of the product of two identical distributions of non negative numbers, which can be seen to be a Mellin convolution:

$$q(w) = \int_0^\infty \frac{1}{y}\, f\left(\frac{w}{y}\right) f(y) dy \qquad x, y, w \geq 0 \tag{3.25}$$

Here the variables x and y are probability distribution functions, and the probability distribution of the product of identical distributions over a domain $[-k, k] \times [-k, k]$ is obtained by summation of the infinitesimal probabilities $\delta x \delta y$ over the level curves $y = w\frac{1}{x}$ for $w \in [-k^2, k^2]$, since on the level curves $w = xy$. It is clear that both the shape of the p.d.f curve and the total probability can be assessed by considering only the first quadrant (where $x, y, w \geq 0$). Hence the full p.d.f for products of all identical symmetrical distributions not limited to non negative numbers can be obtained by summation of two copies of the function (since $w \geq 0$ in two quadrants, noting that the process is not the same as doubling the function values) and then including the reflection of the curve about the w axis (to take account of $w \leq 0$).

In a similar fashion, the quotient of two identical distributions of non negative numbers is derived as the Mellin convolution:

$$q(w) = \int_0^\infty y\, f(wy) f(y)\, dy \qquad x, y, w \geq 0 \tag{3.26}$$

Consideration of the lines $y = wx$ over the domain $[-k, k] \times [-k, k]$ leads us to a conclusion similar to that for products in that all the information we require can be obtained by considering only the first quadrant (where $x, y, w \geq 0$) then proceeding via summation of two copies of the function (since $w \geq 0$ in two quadrants) and inclusion of the reflection of the curve about the w axis to get the p.d.f. In both these expressions for $q(w)$ the constituent p.d.f's $f(.)$ are presumed to be even functions with symmetry about 0.

The partial convolution functions as described will be called *quarter p.d.f*'s and *half p.d.f*'s, and for any half p.d.f $h(w)$ the *full p.d.f* is given by:

$$g(w) = h\left(|w|\right) \quad -\infty \leq w \leq \infty \tag{3.27}$$

The process of integration, summation of two copies and reflection, based on a quarter p.d.f will be termed the *quarter p.d.f method*.

The method can then be used for any product or quotient of symmetrical distributions about 0.

To consider the quarter p.d.f corresponding to the product of identical distributions, let $x, y \in [0, k] \subset \mathbb{R}$, $k > 0$ and $w = xy$, then $0 \leq w \leq k^2$ while

$x = \dfrac{w}{y}$, $0 \leq x \leq k$, $0 \leq y \leq k$. Hence the integration limits for y can be determined as:

$$0 \leq x = \frac{w}{y} \leq k \Rightarrow \; 0 \leq w \leq yk \Rightarrow \frac{w}{k} \leq y \leq k$$

and the required quarter p.d.f is given by:

$$q(w) = \int_{\frac{w}{k}}^{k} \frac{1}{y} f\left(\frac{w}{y}\right) f(y)dy \qquad 0 \leq w \leq k^2 \tag{3.28}$$

The domains of the p.d.f's for the product of two distributions are then:

$$\begin{array}{ll} 0 \leq w \leq k^2 & \text{quarter p.d.f} \\ 0 \leq w \leq k^2 & \text{half p.d.f} \\ -k^2 \leq w \leq k^2 & \text{full p.d.f} \end{array} \tag{3.29}$$

In the case of products of uniform distributions (3.14) over $[-k, k]$, the quarter p.d.f method maps the original uniform distribution progressively as follows:

$$\begin{array}{lrl} \mathfrak{D}_U(w) = \frac{1}{2k} & -k \leq w \leq k & \text{uniform p.d.f} \\ q(w) = \frac{1}{4k^2} log\frac{k^2}{w} & 0 \leq w \leq k^2 & \text{quarter p.d.f for products} \\ h(w) & 0 \leq w \leq k^2 & \text{half p.d.f for products} \\ g(w) & -k^2 \leq w \leq k^2 & \text{full p.d.f for products} \end{array} \tag{3.30}$$

This is where we begin to see the complexity of the integrals increasing. Since the functions $q(w)$ to be summed are not over disjoint domains, to calculate $h(w)$ we must convolute two copies of the quarter p.d.f:

$$\begin{aligned} h(w) &= \int_0^{k^2} \frac{1}{4k^2} log\left(\frac{k^2}{w-y}\right) \frac{1}{4k^2} log\left(\frac{k^2}{y}\right) dy \\ &= \frac{1}{16k^4} \int_0^{k^2} log\left(\frac{k^2}{w-y}\right) log\left(\frac{k^2}{y}\right) dy \end{aligned} \tag{3.31}$$

and the indefinate integral before evaluation is:

$$\begin{aligned} h(w) = y + y\; log\left(\frac{k^2}{w-y}\right)\Bigg(1 + log\left(\frac{k^2}{y}\right) - y(-1 + log(y)) + w\; log(-w+y) \\ + \left(log\left(\frac{k^2}{y}\right) + log(y)\right)(y + w\; log(-w+y) - w(log(y) log\left(1 - \frac{y}{w}\right) + Li_2\left(\frac{y}{w}\right) \end{aligned} \tag{3.32}$$

where $Li_2(.)$ represents the dilogarithm of x [25], here defined as the integral function:

$$Li_2(x) = \int_0^x \frac{log(1-u)}{u} du, \quad x \in (-\infty, 1] \tag{3.33}$$

Further evaluation to obtain the full distribution $g(w)$ (which in this case will be called the *Uniform Product Distribution*, $\mathfrak{D}_{UP}$) will be left to a later thesis, as will the remaining full evaluations in this thesis.

The quarter product of normal distributions from (3.25) and (3.16 by similar arguments is given by:

$$q(w) = \frac{1}{2\pi\sigma^2} \int_{\frac{w}{k}}^{k} \frac{1}{y} e^{-\frac{1}{2\sigma^2}\left(\frac{w^2}{y^2}+y^2\right)} dy \qquad 0 \leq w \leq k^2 \tag{3.34}$$

The quotient of identical distributions is now considered. Let $x, y \in [0, k] \subset \mathbb{R},\ y \neq 0,\ \ k > 0$ and $w = \frac{x}{y}$, then $0 < w < \infty$ while $x = wy,\ 0 \leq x \leq k,\ 0 \leq y \leq k$. Hence the integration limits for y can be determined:

$$\begin{array}{llll} (a) & 0 \leq w \leq 1, & \Rightarrow\ 0 \leq x = wy \leq k & \Rightarrow\ 0 \leq y \leq k \\ (b) & 1 \leq w < \infty, & \Rightarrow\ 0 \leq x = wy \leq k & \Rightarrow\ 0 \leq y \leq \frac{k}{w} \end{array}$$

and the quarter p.d.f over $[0, k]$ is given by:

$$q(w) = \begin{cases} \int_0^k y\ f(wy)f(y)dy & 0 \leq w \leq 1 \\ \\ \int_0^{\frac{k}{w}} y\ f(wy)f(y)dy & 1 \leq w < \infty \end{cases} \tag{3.35}$$

Quotients of distributions about 0 are distributed over $[-\infty, \infty]$.

The distribution of the quotient of two identical uniform distributions is k-invariant, and has tails extending to $\pm\infty$, unlike the distribution of the product of uniform distributions which has finite bounds of $\pm 2k^2$.

3.8 Difference between two Uniform Product Distributions

The distribution of the difference between two $\mathfrak{D}_{UP}$ distributions is required in order to calculate (in the first instance) determinants of matrices in $GL(2, \mathbb{R})$ with uniformly distributed entries. Suppose a matrix A in $GL(2, \mathbb{R})$ is of the standard form and that entries $a, b, c, d \in [-k, k] \subset \mathbb{R}$ are uniformly distributed over that interval.

The general distribution of the difference between two identical distributions each over $[-k, k]$ has been derived in (3.13), that expression is now applied to two $\mathfrak{D}_{UP}$ distributions over $[-2k^2, 2k^2]$ to determine the quarter p.d.f:

$$q(w) = \begin{cases} \int_{-2k^2-w}^{2k^2} \mathfrak{D}_{UP}(w+y)\mathfrak{D}_{UP}(y)dy & 4k^2 \le w \le 0 \\ \\ \int_{-2k^2}^{2k^2-w} \mathfrak{D}_{UP}(w+y)\mathfrak{D}_{UP}(y)dy & 0 \le w \le 4k^2 \end{cases} \tag{3.36}$$

3.9 Unequal Domain Limits

Since almost all the published work involves the restriction to intervals $[0, 1]$, investigation of a generalisation to uniformly distributed variables over $[-m, n]$ for $m, n \in \mathbb{R}^+$ is made. The uniform distribution function for such unequal domains is:

$$\mathfrak{D}_{Umn}(x) = \begin{cases} \frac{1}{m+n} & \forall x \in [-m, n] \\ \\ 0 & \text{otherwise} \end{cases} \tag{3.37}$$

Equation (3.25) gives the integral representation of the product of two identical distributions of non negative numbers, but the argument as to the total distribution with the non-negative restriction removed requires modification. It is observed that the areas enclosed by the limits in quadrants $1, 2, 3$ and 4 respectively are n^2, mn, m^2 and mn. Hence calculation and modification of quarter p.d.f's cannot be performed; instead integral components must be calculated separately for all four quadrants and the p.d.f components combined, in a piecewise fashion where the function domains are disjoint but with the use of convolution otherwise.

- First quadrant: Let $x, y \in [0, n] \subset \mathbb{R},\ n > 0$ and $w = xy$, then $0 \le w \le n^2$ while $0 \le x = \frac{w}{y} \le n,\ 0 \le y \le n$. Hence the integration limits for y can be determined as:

$$0 \le x = \tfrac{w}{y} \le n \ \Rightarrow\ 0 \le w \le yn \ \Rightarrow\ \tfrac{w}{n} \le y \le n$$

  and the first quadrant p.d.f is given by:

$$h_{Q1}(w) = \int_{\frac{w}{n}}^{n} \tfrac{1}{y} f\left(\tfrac{w}{y}\right) f(y)dy \qquad 0 \le w \le n^2 \tag{3.38}$$

  Here, for uniform product distributions:

$$h_1(w) = \tfrac{1}{(m+n)^2} \log \tfrac{n^2}{w} \qquad 0 \le w \le n^2 \tag{3.39}$$

and for uniform product distributions over $[0, 1]$:

$$h(w) = -\log(w) \quad 0 \le w \le 1 \tag{3.40}$$

which latter is a function which integrates to a value of 1 over the interval $0, 1]$ and is a p.d.f, and is shown in Figure (3.6).

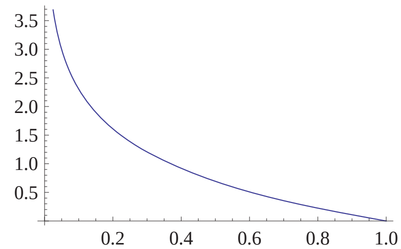


Figure 3.6: Distribution of products of entries uniformly distributed over the interval $[0, 1]$

- Second quadrant: Let $x \in [-m, 0],\ y \in [0, n] \subset \mathbb{R},\ m, n > 0$ and $w = xy$, then $-mn \le w \le 0$ while $-m \le x = \frac{w}{y} \le 0,\ 0 \le y \le n$. Hence the integration limits for y can be determined as:

$$-m \le x = \tfrac{w}{y} \le 0 \Rightarrow\ -my \le w \le 0 \Rightarrow -\tfrac{w}{m} \le y \le n$$

and the second quadrant p.d.f is given by:

$$h_{Q2}(w) = \int_{-\frac{w}{m}}^{n} \tfrac{1}{y} f\left(\tfrac{w}{y}\right) f(y) dy \qquad -mn \le w \le 0 \tag{3.41}$$

For uniform product distributions:

$$h_2(w) = \tfrac{1}{(m+n)^2} \log \tfrac{-mn}{w} \qquad -mn \le w \le 0 \tag{3.42}$$

- Third quadrant: Let $x, y \in [-m, 0] \subset \mathbb{R},\ m > 0$ and $w = xy$, then $0 \le w \le m^2$ while $-m \le x = \frac{w}{y} \le 0,\ -m \le y \le 0$. Hence the integration limits for y can be determined as:

$$-m \le x = \tfrac{w}{y} \le 0 \Rightarrow\ -ym \ge w \ge 0 \Rightarrow -m \le y \le -\tfrac{w}{m}$$

and the third quadrant p.d.f is given by:

$$h_{Q3}(w) = \int_{-m}^{-\frac{w}{m}} \tfrac{1}{y} f\left(\tfrac{w}{y}\right) f(y)dy \qquad 0 \leq w \leq m^2 \tag{3.43}$$

For uniform product distributions:

$$h_3(w) = \tfrac{-1}{(m+n)^2} \log \tfrac{m^2}{w} \qquad 0 \leq w \leq m^2 \tag{3.44}$$

- Fourth quadrant: Let $x \in [0, n],\ y \in [-m, 0] \subset \mathbb{R},\ m, n > 0$ and $w = xy$, then $-mn \leq w \leq 0$ while $0 \leq x = \frac{w}{y} \leq n,\ -m \leq y \leq 0$. Hence the integration limits for y can be determined as:

$$0 \leq x = \tfrac{w}{y} \leq n \Rightarrow\ 0 \geq w \geq yn \Rightarrow -m \leq y \leq \tfrac{w}{n}$$

and the fourth quadrant p.d.f is given by:

$$h_{Q4}(w) = \int_{-m}^{\frac{w}{n}} \tfrac{1}{y} f\left(\tfrac{w}{y}\right) f(y)dy \qquad -mn \leq w \leq 0 \tag{3.45}$$

For uniform product distributions:

$$h_4(w) = \tfrac{-1}{(m+n)^2} \log \tfrac{-mn}{w} \qquad -mn \leq w \leq 0 \tag{3.46}$$

Suppose $n < m$, then:

$$-m^2 < -mn < -n^2 \leq 0 \leq n^2 < mn < m^2 \tag{3.47}$$

Since $h_2(w)$ and $h_4(w)$are both distributed over $[-mn, 0]$ and are identical but opposite signed functions, the sum of these partial p.d.f's can be obtained by Fourier convolution of the difference between identical distributions. Working from (3.12):

$$-mn \leq x = w + y \leq 0 \Rightarrow -mn - w \leq y \leq 0$$

where $x = h_2(z),\ y = h_4(z)$, and then

$$h_{2,4}(w) = \frac{1}{(m+n)^4} \int_{-mn-w}^{0} log(\tfrac{-mn}{w+y}) log(\tfrac{-mn}{y})dy \qquad -mn \leq w \leq mn \tag{3.48}$$

Over the domain $[0, n^2]$ the functions $h_1(w)$ and $h_3(w)$ are identical but opposite signed, the sum of these partial p.d.f's can be obtained by Fourier convolution of the difference between identical distributions. Again working from (3.12):

$$0 \leq x = w + y \leq n^2 \Rightarrow 0 \leq y \leq n^2 - w$$

where $x = h_1(z),\ y = \{h_3(z) \mid 0 \leq z \leq n^2\}$, and then

$$h_{1,3}(w) = \frac{1}{(m+n)^4} \int_0^{n^2-w} log(\tfrac{n^2}{w+y}) log(\tfrac{n^2}{y}) dy \qquad -n^2 \leq w \leq n^2 \tag{3.49}$$

Noting that the domains of $h_{2,4}(w)$ and $h_{1,3}(w)$ are not disjoint, and that the remainder of $h_3(w)$, namely $\{h_3(w) \mid n^2 \leq w \leq m^2\}$ has domain not disjoint from either of these, there is in fact a considerable amount of work required to complete this determination.

3.10 The Square Root of a Distribution

The distributions developed so far are based on independent random variables, but suppose now that x and y are such that $x = y$. We apply the expression for the quarter product of two distributions each over $[-k, k]$ (3.28) with $w = xy = y^2$:

$$q(w) = \int_{\frac{w}{k}}^{k} \frac{1}{y} f^2(y)\, dy \tag{3.50}$$

which is the quarter product of the square of the single distribution$f(y)$. Instead of carrying through the integration, we substitute y^2 for w and differentiate:

$$\frac{1}{y} f^2(y) = \frac{d(q(y^2)}{dy} \tag{3.51}$$

hence

$$f(y) = \pm \sqrt{y \frac{d(q(y^2)}{dy}} \tag{3.52}$$

where f(y) is the p.d.f of $\sqrt{q(y)}$ with $q(y) = \sqrt{\frac{1}{2} h(|y|)}$ being a quarter product corresponding to a distribution $h(y)$.

Chapter 4

ISOMETRIC CIRCLE ANALYSIS

4.1 Disjoint Isometric Circles in $GL(2,\mathbb{C})$ and $SL(2,\mathbb{C})$

For the pair of isometric circles corresponding to a matrix A in $GL(2,\mathbb{C})$ to be disjoint or at the most tangential would require the separation to be greater than twice the radius, or:

$$\left|\frac{a}{c}+\frac{d}{c}\right| \geq 2\frac{\left|\sqrt{ad-bc}\right|}{|c|} \tag{4.1}$$

which condition reduces to

$$|trace(A)|^2 \geq 4\,|\det(A)| \tag{4.2}$$

Of the four complex coordinates of the $GL(2,\mathbb{C})$ matrix space, the two complex coordinates of the centres of the isometric circles are determined by entries d, a and c; but the real radii are determined by the modulus of a composition of all entries. Hence again for each pair of isometric circles there exist an infinite number of corresponding matrices in $GL(2,\mathbb{C})$.

For the pair of isometric circles corresponding to a matrix in $SL(2,\mathbb{C})$ the condition further reduces to

$$|trace(A)|^2 \geq 4 \tag{4.3}$$

and whenever the square of the trace is real the disjoint isometric circles of matrices in $SL(2,\mathbb{C})$ are precisely those that correspond to parabolic or hyperbolic transformations.

It is easy to extend the disjointness criteria to pairs of isometric circles corresponding to the matrices A, B in $SL(2,\mathbb{C})$.

4.2 Equations and Centres of Isometric Circles in $SL(2,\mathbb{C})$

We derive an expression for the equation of an isometric circle for a complex matrix in the form

$$A = \begin{pmatrix} a+i\alpha & b+i\beta \\ c+i\gamma & d+i\delta \end{pmatrix} \qquad a,b,c,d,\alpha,\beta,\gamma,\delta \in \mathbb{C} \tag{4.4}$$

as:

$$\left(x + \frac{cd+\gamma\delta}{c^2+\gamma^2}\right)^2 + \left(y + \frac{c\delta - d\gamma}{c^2+\gamma^2}\right)^2 = \frac{1}{c^2+\gamma^2} \tag{4.5}$$

which is clearly the equation of a circle in $\mathbb{C}$, of centre $-\frac{cd+\gamma\delta}{c^2+\gamma^2} - i\frac{c\delta - d\gamma}{c^2+\gamma^2}$ and radius $\sqrt{\frac{1}{(c^2+\gamma^2)}}$. All terms in the complex expression are real, hence the radius of the isometric circle is real and positive whenever $c+i\gamma \neq 0$, and isometric circles exist for all matrices in $SL(2,\mathbb{C})$ provided $c+i\gamma \neq 0$. As $c+i\gamma \to 0$, both radii and centres tend to ∞ and in the limit both isometric circles have as centre the point at infinity on the Riemann sphere and as circumference the point $(0,0)$ and correspond to matrices representing transformations of the form $f = \frac{a}{d}z + \frac{b}{d}$.

For $A \in SL(2,\mathbb{R})$ (4.5) reduces to:

$$\left(x + \frac{d}{c}\right)^2 + y^2 = \frac{1}{c^2} \tag{4.6}$$

In $SL(2,\mathbb{C})$, if the radius is denoted by r_A the centre of the isometric circle of A in $\mathbb{C}$, $x_A + iy_A$ (and similarly for B), then from (4.5), for the isometric circles of matrices A and B:

$$r_A = \sqrt{\frac{1}{c^2+\gamma^2}}, \quad r_B = \sqrt{\frac{1}{g^2+\varphi^2}} \tag{4.7}$$

$$x_A = -\frac{cd+\gamma\delta}{c^2+\gamma^2} = -r_A^2(cd+\gamma\delta), \quad x_B = -\frac{gh+\varphi\eta}{g^2+\varphi^2} = -r_B^2(gh+\varphi\eta) \tag{4.8}$$

$$y_A = -\frac{c\delta - d\gamma}{c^2+\gamma^2} = -r_A^2(c\delta - d\gamma), \quad y_B = -\frac{g\eta - h\varphi}{g^2+\varphi^2} = -r_B^2(g\eta - h\varphi) \tag{4.9}$$

and the equations for the isometric circles of A and B can be written:

$$(x-x_A)^2 + (y-y_A)^2 = r_A^2, \quad (x-x_B)^2 + (y-y_B)^2 = r_B^2 \tag{4.10}$$

From these results the equations for lines of symmetry and conditions for intersection are derived for use in isometric circle analysis of the discrete group criterion.

Chapter 5

CONCLUSIONS

The following are considered to be significant results related to the methodology of this thesis, namely computations based on random matrices in $(P)SL(2,\mathbb{C})$ and $(P)SL(2,\mathbb{R})$ derived from matrices in $GL(2,\mathbb{C})$ and $GL(2,\mathbb{R})$ respectively with uniformly distributed entries divided by the (appropriate) square root of the determinant as described in this thesis, the use of Euclidean norms on the embedding spaces, and further development of Springer's methods for analytical determination of composition of distributions.

The distributions of matrices in $SL(2,\mathbb{C})$, $SL(2,\mathbb{R})$, $PSL(2,\mathbb{C})$ and $PSL(2,\mathbb{R})$ exhibit invariance with respect to the size of the domain of definition for entries in $GL(2,\mathbb{C})$.

The likelihood of occurrence of two-generator discrete groups with respect to the criteria has been quantified, and though over 99% of matrices in $PSL(2,\mathbb{C})$ and $PSL(2,\mathbb{R})$ meet Jørgensen's criterion, only 0.4% in $PSL(2,\mathbb{C})$ and 0.8% in $PSL(2,\mathbb{R})$ meet the isometric circle criterion.

The entries of $SL(2,\mathbb{C})$ and $SL(2,\mathbb{R})$ appear to be distributed about 0 but those of $PSL(2,\mathbb{C})$ and $PSL(2,\mathbb{R})$ appear to be distributed about the identity.

Derivations have been made of some algebraic expressions for distributions of random variables over domains not restricted to non negative numbers.

Chapter 6

FURTHER WORK

The following are some areas in which preliminary results have been determined but for which there is insufficient scope for further work in this thesis.

The evolution of single generator and two-generator subgroups, results fall into two categories. Firstly, work on isometric circles shows that after the initial product has been formed in a two-generator free group, then in most cases the convergence of the isometric circles is determined from just the iteration of a single generator. Secondly, by following through iterations of single and two-generator subgroups, it is noted that whereas the qualitative nature of distributions can be maintained through the evolution of a subgroup of $GL(2,\mathbb{R})$ or $GL(2,\mathbb{C})$, in the evolution of subgroups of $SL(2,\mathbb{R})$, $SL(2,\mathbb{C})$, $PSL(2,\mathbb{R})$ and $PCL(2,\mathbb{R})$ the width of the distributions is progressively narrowed about quantiles containing the identity.

It is also noted that the centres of isometric circles necessarily discrete by the isometric circle criterion are distributed about the unit circle, while those possibly discrete by Jørgensen's criterion are distributed about the origin in $\mathbb{C}$.

The application of the mathematics of convolution to the algebra of random distributions needs to be taken further along the lines suggested in Chapter 3, with an ultimate objective of proof of the computationally determined distributions.

Further analysis is required to determine how realistic the constrained entry distributions are in the subgroups of $SL(2,\mathbb{C})$.

Further computational analysis of inequalities may prove fruitful in investigating sharpness and distribution near bounds.

REFERENCES

[1] William P. Thurston, *The Geometry and Topology of Three-Manifolds*, Princeton University, 1980

[2] Bernard Maskit, *Kleinian Groups*, Springer-Verlag, 1988

[3] Troels Jørgensen, *On discrete groups of Mobius Transformations*, American Journal of Mathematics, Vol 98 No. 3, 739-749, 1973

[4] Troels Jørgensen and Peter Klein, *Algebraic convergence of finitely generated Kleinian groups*, Quart. J. Math. Oxford (2), 33 325-332, 1982

[5] F W Gehring and G J Martin, *Inequalities for Möbius transformations and discrete groups*, J. reine angew. Math. 418, 31 - 76, 1991

[6] J Gilman, *A Discreteness condition for subgroups of PSL(2,*$\mathbb{C}$), Proceedings of the Bers Colloquium, Contemporary Math Series, 211, AMS 261-267, 1997

[7] Alan F. Beardon, *The Geometry of Discrete Groups*, Springer-Verlag, 1983

[8] Gilbert Strang and Alan Edelman, *Random triangle theory with geometry and applications*, in preparation 2010

[9] Emmanuel Breuillard, *An exposition of Jordan's original proof of his theorem on finite subgroups of* $GL_n(\mathbb{C})$, Laboratoire de Mathematiques, University Paris Sud

[10] G E P Box and Mervin E Muller, *A note on the generation of random normal deviates*, The Annals of Mathematical Statistics, Vol 29, No. 2, 610-611, 1958

[11] Z W Birnbaum and R A Hall, *Small Sample Distributions for Multi-sample Statisticsof the Smirnov type*, Springer-Verlag, 1960

[12] Oliver Thas, *Comparing Distributions*, Springer-Verlag, 2010

[13] R. C. Williamson and T. Downs, *The Inverse and Determinant of a* 2×2 *Uniformly Distributed Random Matrix,* Statistics and Probability Letters 7, 167-170, 1989.

[14] R. Fortet, *Random Determinants*, Journal of Research of the National Bureau of Standards 47, 465-470, 1951.

[15] H. Nyquist, S. O. Rice and J. Riordan, *The Distribution of Random Determinants,* Quarterly of Applied Mathematics 12, 97-104, 1954.

[16] J Komlos, *On the determination of random matrices*, Studia Scientiarium Mathematicarum Hungarica 3, 387-399, 1968

[17] V. S. Alagar, *The Distributions of Random Determinants,* The Canadian Journal of Statistics, 6, 1-9, 1978.

[18] Gary L. Wise and Eric B. Hall, *A note on the Distribution of the Determinant of a Random Matrix,* Statistics and Probability Letters 11, 147-148, 1991.

[19] R Delannay and G Le Caer *Distribution of the determinant of a random real-symmetric matrix from the Gaussian orthogonal ensemble*, Physical Review E, 62 no. 2, 2000

[20] M D Springer, *The Algebra of Random Variables*, John Wiley Sons, 1979

[21] Erwin Kreyszig, *Advanced Engineering Mathematics*, John Wiley Sons, 7 ed 1993

[22] W. L. Nicholson, *On the Distribution of* 2×2 *Random Normal Determinants,*Annals of Mathematical Statistics 29, 575-580, 1958.

[23] Athanasios Papoulis and S. Unnikrishna Pillai, *Probability, Random Variables and Stochastic Processes*, McGraw Hill, 4ed 2002

[24] Weisstein,Eric W, *Normal Sum Distribution*,From MathWorld–A Wolfram Web Resource. http://mathworld.wolfram.com/NormalSumDistribution.html, 2011

[25] Don Zagier, *The Dilogarithm Function*, Max-Plank-Institut fur Mathematic, Paper still in preparation

[26] DI Management Services, Sydney, Australia *Cryptosys API*, di-mgt.com.au

[27] NIST, *Recommendation for Random Number Generation using Deterministic Random Bit Generators*, National Institute of Standards and Technology, SP 800-90, 2010

[28] FIPS PUB, *Security Requirements for Cryptographic Modules*, Federal Information Processing Standards Publication 140-2, 2001

[29] E T Copson, *Metric Spaces*, Cambridge University Press, 1968

[30] Lester R. Ford, *On the foundations of the theory of discontinuous groups of linear equations*, Proc. National Academy of Sciences, March 1927

[31] Rowland Todd, *Operator Norm*, Wolfram Mathworld 2010