

Intuitive explanation of the Riemann hypothesis

I. Characterisation of the nontrivial zeroes of ζ .

There is a unique (canonical) one-form α on \mathbb{H} invariant under $\Gamma(2)$ with a pole of residue 1 at the image of $i\infty$ and a pole of residue -1 at the image of 1. Under the embedding $\mathbb{H} \rightarrow \mathbb{C}$ with τ the coordinate on \mathbb{C} the ratio $[\alpha : i\pi d\tau]$ tends to one at the upper end of the interval $(0, i\infty)$. Let T be the connected real multiplicative group and consider the multiplication actions

$$\begin{cases} \mu_+ : T \times \mathbb{H} \rightarrow \mathbb{H} \\ (g, z) \mapsto \sqrt{g}z \end{cases}$$

$$\begin{cases} \mu_- : T \times \mathbb{H} \rightarrow \mathbb{H} \\ (g, z) \mapsto \frac{1}{\sqrt{g}}z \end{cases}$$

1. Theorem. For each unitary character ω of T and each real number c with $0 < c < 1$, the differential two-form

$$g^{2c-2}\omega(g)\mu_+^*(\alpha - i\pi d\tau) \wedge \mu_-^*(\alpha - i\pi d\tau)$$

is real and integrable (rapidly decreasing, that is ‘Schwartz’ with respect to an invariant vector-field) on $T \times (0, i\infty)$. Among rapidly decreasing forms, it is exact if and only if $\zeta(c + i\omega_0)$ is zero where ζ is Riemann’s zeta function and ω_0 is the real number corresponding to ω under the rule $\omega(g) = g^{i\omega_0}$.

Proof. It is real because the factors besides $\omega(g)$ are anti-symmetric with respect to interchanging μ_+ and μ_- which matches the reversal of orientation of T . The two-form integrates to the squared absolute value of a holomorphic integral, namely $\int g^{c-1}\omega(g)(\alpha - i\pi d\tau)$. In turn, it is easy to calculate the holomorphic *definite* integral; it is $-L(s, \chi)\Gamma(s)\pi^{1-s}$ L is the L series for sums of four squares, χ is the Dirichlet sign character and $s = c + i\omega_0$. The rule $\omega(g_1g_2^{-1}) = \omega(g_1)\omega(g_2)^{-1}$ is all that is needed.

II. Remark about the dynamical interpretation.

Here is an intuitive way of integrating the two-form let, us call it A_s for $s = c + i\omega_0$. Let $\tau = ie^t$. By ‘integration by parts’

$$\int e^{(c-1)t+i\omega_0t} d \log\left(\frac{\lambda}{q}\right) = -(c-1+i\omega_0) \int e^{(c-1)t+i\omega_0t} \log\left(\frac{\lambda}{q}(ie^t)\right) dt.$$

Therefore

$$\int \int A_s = |(s-1)|^2 \left| \int_{-\infty}^{\infty} e^{i\omega_0t} e^{(c-1)t} \log\left(\frac{\lambda}{q}(ie^t)\right) dt \right|^2.$$

The second term on the right is the squared magnitude of the Fourier transform value at frequency ω_0 of the real function

$$e^{(c-1)t} \log\left(\frac{\lambda}{q}(ie^t)\right).$$

A disk spinning with angular rate ω_0 with pivot point held by a pair of opposing movable bearings, if we move the bearings in a line according to this function (of time), the limiting radius of the circle traced by the initial pivot point will be the magnitude and

2. Theorem.

$$\frac{\pi}{|s-1|^2} \int A_s = \text{area inside final circle.}$$

III. Lie actions.

Whenever A_s is a Lie derivative $A_s = \delta B$ under the action of a vector-field δ , then A_s can be obtained by multiplying B by a suitable divergence ratio; put differently $A_s = d i_\delta B$ which is an exact form. If the contraction $i_\delta B$ is Schwartz then $\zeta(s) = 0$ (still assuming $0 < 1 < c$).

IV. The action of $\frac{\partial}{\partial c}$.

Conversely, a vector field which does not preserve $T \times (0, i\infty)$ is the partial derivative with respect to c . If $g = e^t$ it sends A_s to $2tA_s$.

3. Question. For $0 < c < 1/2$, is the partial derivative $\frac{\partial}{\partial c} \int A_{c+i\omega_0}$ non-positive?

An affirmative answer would imply that A_s is non-exact, and $\zeta(c + i\omega_0) \neq 0$, for all c in the same range. The reason is that for each value of ω_0 the dependence on c would be a non-increasing real analytic function $(0, 1/2) \rightarrow [0, \infty)$. Such a function cannot take the value of zero.

Let's attempt to estimate the partial derivative to see if we can start to answer the question. Let

$$h(r, v) = e^{2(c-1)v} \log\left(\frac{\lambda}{q}(v + r/2)\right) \log\left(\frac{\lambda}{q}(v - r/2)\right)$$

This has the properties that for $0 < c < 1/2$

$$\begin{cases} h(r, v) > 0 & \text{for all } r, v \\ h(r, v) - h(r, -v) < 0 & \text{for all } r \text{ and all } v > 0 \end{cases}$$

For each fixed c and r let

$$\gamma(c, r) = \frac{\int v h(r, v) dv}{\int h(r, v) dv}$$

This is the mean value of $h(r, v)$ as a function of v .

Now

$$\begin{aligned}
\frac{\partial}{\partial c} \int A_s &= \frac{\partial}{\partial c} ((c-1)^2 + \omega_0^2) \int \int \cos(r\omega_0) h(r, v) dv dr. \\
&= (2c-2) \int \cos(r\omega_0) \int h(r, v) + ((c-1)^2 + \omega_0^2) (2v) h(r, v) dv dr \\
&= (2c-2) \int \cos(r\omega_0) \int h(r, v) dv dr \\
&\quad + ((c-1)^2 + \omega_0^2) \int 2\gamma(c, r) \cos(r\omega_0) \int h(r, v) dv dr
\end{aligned}$$

The integral $\int \int \cos(r\omega_0) h(r, v) dv dr$ is semi-positive since it is the squared magnitude of a complex number. Each of the coefficients $2c-2$ and $\gamma(c, r)$ are negative when $0 < c < 1/2$.

Let

$$\rho(c + i\omega_0) = \frac{\int \gamma(c, r) \cos(\omega_0 r) \int h(r, v) dv dr}{\int \cos(\omega_0 r) \int h(r, v) dv dr}$$

so our integral is

$$\begin{aligned}
&= ((2c-2) + 2((c-1)^2 + \omega_0^2) \rho(s)) \int \cos(r\omega_0) \int h(r, v) dv dr \\
&= 2 \left(\frac{c-1}{(c-1)^2 + \omega_0^2} + \rho(s) \right) \int A_s.
\end{aligned}$$

Removing the leading factor of -1 in $-L(s, \chi) \Gamma(s) \pi^{1-s}$ which has no effect, and removing our leading factor of 2 which relates the real part of the logarithmic derivative with our integral, we obtain

$$\operatorname{Re} \frac{d}{ds} \log (L(s, \chi) \Gamma(s) \pi^{1-s}) = \frac{\operatorname{Re}(s-1)}{|s-1|^2} + \rho(s).$$

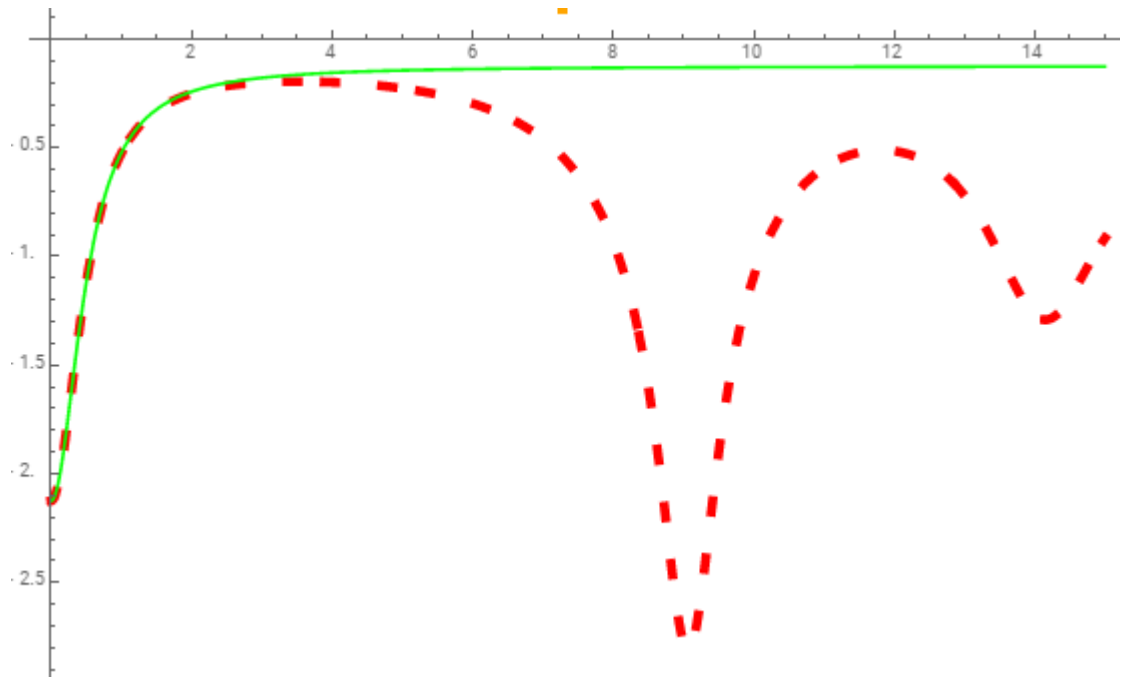
The logarithmic derivative of the gamma function is the digamma function $\Psi(s)$. With this notation then

4. Theorem.

$$\operatorname{Re} \frac{d}{ds} \log L(s, \chi) = \frac{\operatorname{Re}(s-1)}{|s-1|^2} + \rho(s) + \log(\pi) - \operatorname{Re} \Psi(s).$$

From this we can write an expression for the real part of the logarithmic derivative of $L(s)$ itself and $\zeta(s)\zeta(s-1)$.

When $c = 1/2$ a first approximation of $\rho(s)$ would just be the constant $\log(\frac{\log(16)}{\pi})$. Here the real part of the logarithmic derivative of $L(s, \chi)\Gamma(s)\pi^{1-s}$ as a red graph, and the $\frac{\text{Re}(s-1)}{|s-1|^2} + \log(\frac{\log(16)}{\pi})$ as a green graph as a function of ω_0 when $c = 1/2$.



This graph is not much evidence as we don't know why the actual value departs from the approximation.

V. Holomorphic interpretation

Before we begin, it makes sense to notice that the Cauchy integral theorem is a statement about functions which a single-valued function times a logarithmic form. In our situation, when we look at a cusp expansion, and a multiple-valued function times a logarithmic form, what takes place is that one can still define a residue and the statement of the Cauchy theorem regarding arcs which pass through a pole of the form, and it is still true. The residue proven to equal the limit of horocycle integrals. Thus

5. Remark. We may speak of ‘residues’ of one-form even if there is a cusp expansion instead of a Laurent expansion.

Recall we let α be the unique $\Gamma(2)$ invariant holomorphic one-form on \mathbb{H} which has a simple pole at the image of $i\infty$ of residue $+1$ and a simple pole at the image of 1 of residue -1 , and we consider for each constant s with $0 < \text{Re}(s) < 1$ the holomorphic one-form

$$(-i\tau)^{s-1}(\alpha - i\pi d\tau) = u^{s-1}(\alpha + \pi du)$$

where $u = -i\tau$. The restriction to $u \in (0, \infty)$ is rapidly decreasing or ‘Schwartz’ once one makes the definition in an appropriately symmetric way (e.g. on the Lie algebra of the multiplicative group).

Our earlier arguments gave equivalent conditions for the existence of f rapidly decreasing. Since they were based on integration by parts, we lost the actual underlying symmetry on the level of forms.

Denote by $0, 1, \infty$ in order be the images in the Riemann sphere of the boundary points $0, 1, i\infty$ of \mathbb{H} , so that α can be interpreted as the meromorphic one-form on the Riemann sphere with a pole of residue $1, -1$ at ∞ and 1 .

Let Z be the rational function which takes the value 0 at 0 with a simple pole of residue 1 at ∞ . The permutations of the three-element set $\{0, 1, \infty\}$ extend uniquely to automorphisms, for instance the transform of α under interchanging 0 and ∞ is the same as the multiple of α by $\frac{1}{Z}$. That is to say, $\frac{1}{Z}$ is the eigenfunction for this involution. In fact $\alpha = \frac{-dZ}{Z-1}$.

The real points of the projective line form a real circle inside the Riemann sphere, which contains all three of the points 0, 1, ∞ . Writing $\tau = iu$, the line where u takes positive real values actually covers the line where Z takes *negative* real values.

The ideal triangle in \mathbb{H} which contains the boundary points 0, 1, $i\infty$ maps topologically isomorphically to the circle of real points of the projective line.

Suppose now that f is a function on the \mathbb{H} such that

$$df = \frac{-u^{s-1}dZ}{Z-1} + \pi u^{s-1}du.$$

We can restrict f to the ideal triangle and therefore to the real points of the projective line.

The form df is rapidly vanishing at the points 0 and ∞ . At the point 1 the form $\alpha = \frac{-dZ}{Z-1}$ has a simple pole of residue -1 (as we know) and the function u^{s-1} takes the value $(-i)^{s-1}$. Since the point $\tau = 1$ is invariant under

$$\frac{-1}{\tau-1} \mapsto \frac{-1}{\tau-1} + 2$$

the function $Z = \frac{\lambda-1}{\lambda}$ can be expressed by a ‘ q -expansion’ in terms of $e^{i\pi/(1-\tau)}$ about $Z = 1$

$$Z = 1 + 16e^{i\pi/(1-\tau)} \dots$$

Here $i\pi/(1-\tau) = -\pi/(u+i)$. We can expand u in terms of Z giving

$$u = -i - \frac{\pi}{\log(\frac{Z-1}{16})} \dots$$

To the first approximation

$$\pi du = (u+i)^2 \frac{dZ}{Z-1}.$$

This means our form is to the first approximation

$$(-1 + (u + i)^2)u^{s-1} \frac{dZ}{Z - 1}.$$

We already knew that including the terms $(u+i)^2$ in the first bracket would not affect the ‘residue.’

Then our form is approximately merely

$$-\left(-i - \frac{\pi}{\log\left(\frac{Z-1}{16}\right)}\right)^{s-1} \frac{dZ}{Z - 1}$$

which agrees closely with

$$-(-i)^{s-1} \frac{dZ}{Z - 1}.$$

This means that we may use the residue calculation, we may work as if the form has a simple pole at the point 1 of the Riemann sphere, and integrate up to a point near 1, then on the other side from a corresponding point, and as the gap is made smaller, the omitted difference will approach $-i\pi$ times $-(-i)^{s-1}$. That is, the counter-clockwise integral about the ideal triangle makes sense, and evaluates to

$$(i\pi)((-i)^{s-1}) = -\pi e^{-\pi is/2}$$

The condition $Re(s) = 1/2$ corresponds to the semilinear eigenvalue of complex conjugation on the residue being the imaginary unit i .

The relation with the condition that $\zeta(s) = 0$ is that this is equivalent to the value at 0 and ∞ being equal, in which case the residue equals the integral over the arc from 0 to ∞ which passes through 1 (now where Z runs over the positive real numbers). Here we must interpret this as an improper integral, leaving a gap about 1 which is symmetrical with respect to the involution, and taking the limit as the gap tends to zero. Thus

6. Theorem. The form $u^{s-1}(\alpha + \pi du)$ corresponds to a well-defined one-form on the real projective line except the point 1, where the cusp expansion has residue $-(-i)^{s-1}$. Thus the integral of the form over the real projective line evaluates to $-\pi e^{-\pi i s/2}$. The condition for $\zeta(s) = 0$ is that the integral takes equal value at 0 and ∞ while the condition for $Re(s) = 1/2$ is that the eigenvalue of complex conjugation on the whole integral is the imaginary unit i . Thus the Riemann hypothesis is equivalent to asserting that when the integral from 0 to ∞ which does not pass 1 is trivial, complex conjugation acts by i on the integral along the arc which passes through 1.

7. Example. Using Wolfram Alpha with $s = .1 + 5i$ to integrate over the ideal triangle omitting the gap from 1 to $1 + e^{-6}i$ and the corresponding gap near 1 on the circle with endpoints at 0 and 1 gives the approximate sum over the three arcs of $-511350 - 3024940i + 503458 + 3026210i + 0.003 - 0.039i$ reading counter-clockwise from 0. These add to $-7891 + 1269i$; while the completely precise actual value, the integral of the natural form over the real points $\mathbb{P}^1(\mathbb{R})$ as we have calculated it, is more precisely

$$-\pi e^{-i\pi s/2} = -7993.016037640268... + 1265.9693716266...i$$

exposing an error of roughly five, due to the missing gap.

The integral over what is here the third arc is zero if and only if $\zeta(s) = 0$, while the real and imaginary parts of our residue, at the unique ‘pole’ on the Riemann sphere, proven here equal to the real and imaginary parts of the full integral, add to zero if and only if $Re(s) = 1/2$.

I should remark, it isn’t actually possible to have a meromorphic form on the Riemann sphere with just one pole (the sum of the residues being zero would prohibit that). The extra care was due to the fact that the form is actually multi-valued, since we never insisted that the coefficient u needs to descend to the quotient modulo the action of $\Gamma(2)$. We really have a multi-valued meromorphic form; however it is single valued on the real points and a single residue determines the whole integral.

VI. Some calculation details.

The interval $(0, i\infty)$ maps to one of our arcs under $\tau \mapsto \tau + 1$ and to the other under $\tau \mapsto 1/(1 - \tau)$, the second being the circle meeting the cusps at 0 and 1. We have oriented this one clockwise so that images of the interval $[a, \infty)$ together with a gap, formed by the images of $[0, a)$, cover the union.

Our integral on the first arc, correctly oriented, is

$$\begin{aligned} & - \int_{ia}^{i\infty} \left(\frac{-i}{1-\tau}\right)^{s-1} d \log \lambda\left(\frac{1}{1-\tau}\right) - i\pi \left(\frac{-i}{1-\tau}\right)^{s-1} d \frac{1}{1-\tau} \\ & = - \int_a^\infty (i+u)^{1-s} d \log \lambda\left(\frac{1}{1-\tau}\right) - \frac{\pi}{s} e^{-i\pi s/2} \end{aligned}$$

where the limits now refer to values of u .

The second integral is

$$\begin{aligned} & \int_{ia}^{i\infty} (-i(1+\tau))^{s-1} d \log\left(\frac{\lambda}{q}(1+\tau)\right) \\ & = \int_a^\infty (u-i)^{s-1} d \log\frac{\lambda}{q}(1+\tau). \end{aligned}$$

As we know that the residue term is $-\pi e^{-i\pi s/2}$, the integrals along all three arcs add to this residue term.

8. Remark. Since the residue term is the limit of the horocycle integral, each of the two integrals shown above must be asymptotic in every case to a purely imaginary multiple of the residue, and the multiplier magnitudes must be approximately negative since no other term depends on a .

9. Remark. We are allowed to truncate the upper half-plane, cutting away a neighbourhood of the ideal point which is in the $\Gamma(2)$ orbit of $0, 1, \infty$, or we can just imagine the universal cover of the Riemann sphere with three disks deleted. Our three arcs now connect a chain of three circles, making a tri-partite graph with nine edges and six vertices, and the inverse image of this graph in \mathbb{H} is such that each circle unwraps to a horocycle, and the two types of edges coming from it alternate. The integral of our form along any edge path is the sum of $-i\pi(-ia)^{(s-1)(b-c)}$ over cusps in the $\Gamma(2)$ orbit of 1 where b and c are the half-integer indices of edges meeting the cusp with rational coordinate a . When we lifted the ‘fundamental cycle’ of $\mathbb{P}^1(\mathbb{R})$ containing $(0, i\infty)$, we might have chosen to use the pair of arcs meeting at the cusp at -1 . Our sum of $-i\pi(-ia)^{(s-1)(b-c)}$ would still have just one term, but since it is the cusp at -1 we would set $a = -1$. The interchange of b and c negates the half-integer $b - c$ and the resulting integral would be the same. Note that it is the same rather than conjugate since the same value of s is used. However, the symmetry here, of negating the real axis, does act by complex conjugation on the variable u such that $iu = \tau$.

10. Remark. Since our form is not invariant under covering transformations, even for defining the integrals along homotopically trivial paths, the integral will depend on the choice of lift of the starting point. In the case when $\zeta(s) = 0$ we choose two lifts, and these we choose to be 0 and $i\infty$ and we know that the integral along any path in \mathbb{H} connecting these will be zero; the route of entry into either of those cusps does not matter since the form is zero there. In our CW decomposition of the sphere, we traced two of the components of the real projective line with $0, 1, \infty$ deleted, to reach the cusp at 1. In our CW decomposition we see a pair of semicircles, which is where we had cut off the cusp.

But if we traverse the wrong one, the lifted path of three arcs will not connect our chosen elements in the orbit of 0 and $i\infty$. If we merely integrate our form along the two arc components of the real projective line, up until they meet the vertex at the horocycle, depending on how small we make it, the sum will converge to what we have called the residue. Also the integral along the semicircular arc will converge to the same number. In this way, integrating ‘straight through’ the cusp downstairs, without making any choices, really means taking the limit of a horocycle integral where we have chosen a particular one of the two semicircular arcs, implicitly, when we decided to represent the beginning and end points of our arc as the points 0 and $i\infty$. But the zeta function implicitly refers to these.

VIII. Functional equations and magnitudes.

The residue term $R(s) = -\pi e^{-i\pi s/2}$ satisfies the functional equations

$$\begin{aligned} R(2-s)R(s) &= -\pi^2 \\ R(2-s) + \overline{R(\overline{s})} &= 0. \end{aligned}$$

At the same time, the product

$$\frac{2^s + 2}{2^s - 4} \int_0^\infty u^{s-1} d \log \frac{\lambda}{q}$$

commutes with complex conjugation acting on s and invariant by $s \mapsto 2 - s$.

Let

$$I(s) = \int_0^\infty (u-i)^{s-1} (d \log \lambda(i(u-i)) + \pi du) - \left(\frac{1}{u+i}\right)^{s-1} d \log \lambda\left(\frac{i}{u+i}\right)$$

and

$$J(s) = \int_0^\infty (u+i)^{s-1} (d \log \lambda(i(u-i)) + \pi du) - \left(\frac{1}{u-i}\right)^{s-1} d \log \lambda\left(\frac{i}{u+i}\right)$$

It is worth looking carefully at these definitions. In both, the πdu term is missing from the factor that has $u-i$ or $u+i$ in the denominator. Also, the only difference between the two is that only the *first* occurrence of $u-i$ or $u+i$ in each of the two terms, is i negated; but s which is not real is not conjugated, nor is the differential in any term affected by the change.

If $\zeta(s) = 0$ then our calculation is

$$-\pi e^{-i\pi s/2} = I(s) - \frac{\pi}{s} e^{-i\pi s/2},$$

which we can simplify as

$$-\pi \left(\frac{s-1}{s} \right) e^{-i\pi s/2} = I(s).$$

Since also $\zeta(\bar{s}) = 0$ we also have

$$-\pi \left(\frac{\bar{s}-1}{\bar{s}} \right) e^{-i\pi \bar{s}/2} = I(\bar{s}).$$

Because of our careful choices in the definition of $J(s)$ we have

$$I(\bar{s}) = \overline{J(s)}.$$

Applying complex conjugation to both sides of the equation then yields

$$-\pi \left(\frac{s-1}{s} \right) e^{i\pi s/2} = J(s).$$

Now note, if we could prove

$$|I(s)J(s)| = \pi^2$$

then we would have

$$\begin{aligned} \pi^2 &= \left| \left(-\pi \frac{s-1}{s} e^{-i\pi s/2} \right) \left(-\pi \frac{s-1}{s} e^{i\pi s/2} \right) \right| \\ &= \pi^2 \left| \frac{s-1}{s} \right|^2 \end{aligned}$$

which would yield

$$|s-1|^2 = |s|^2.$$

Writing $s = c + i\omega_0$ this would give

$$c^2 + \omega_0^2 = (c-1)^2 + \omega_0^2$$

so

$$\begin{aligned} c^2 &= c^2 - 2c + 1 \\ 0 &= 2c - 1 \end{aligned}$$

and we could conclude $Re(s) = 1/2$.

References

1. Marcus Du Sautoy, The music of the primes: why an unsolved problem in Mathematics matters, 2004
2. Keith Ball, Rational approximations to the zeta function, June, 2017
3. Robert Mackay, email, 19 October, 2017
4. Robert Mackay, email, 20 October, 2017
5. Robert Mackay, email, 22 October, 2017
6. Robert Mackay, email, 14 December, 2017
7. Nayo Reid, conversation, Euston Place, Leamington, 23 January, 2018
8. Robert Mackay, email, 28 February, 2018
9. Robert Mackay, email, 20 March, 2018

John Atwell Moody
20 April 2018