Order and symmetry in birational difference equations and their signatures over finite phase spaces

John A G Roberts
University of New South Wales, Australia
jag.roberts@unsw.edu.au

Abstract
We consider two classes of birational maps, or birational difference equations, that have structural properties defined by algebraic relations. The properties are possession of a rational integral and having a discrete time-reversal symmetry. We then consider how these algebraic structures constrain the distribution of orbit lengths of such maps when they are reduced over finite fields, giving a characteristic signature for each property. This can be exploited to test for such properties over the continuum.

1 Introduction
In this contribution, I will review very briefly various aspects of our research into algebraic and arithmetic properties of birational difference equations or birational maps. I refer to the original references for more details and for extended bibliographies.

It has long been recognized that dynamical systems can be classified according to certain structural properties they might possess. Such structures strongly affect the dynamics, and an overarching theme of modern dynamical systems is to identify, characterise and exploit dynamical structures [4, 7]. Here we deal with discrete-time dynamical systems (birational maps or difference equations), and the structures we consider are possession of integrals or invariants of motion and possession of reversing (time-reversal) symmetries. Both structures are defined by algebraic relations which make sense over fields other than the real or complex numbers. By considering birational maps with these structures over finite fields, we ask what signatures do these structures leave on such finite phase spaces.

2 Integrable birational maps
Integrable birational maps of the plane are prominent examples of Discrete Integrable Systems ([1, 21] and references therein). They are birational maps $L : (x, y) \mapsto (x', y')$ with $(x, y) \in \mathbb{R}^2$ (or $\mathbb{C}^2$) that are measure-preserving1 and preserve a rational integral of motion $I(x, y) = \frac{1}{m(x, y)}$.

Recall [15, chapter 2.2] that $L$ is (anti) measure-preserving with density $m(x, y)$ if the Jacobian determinant $J(x, y) := \det dL(x, y)$ can be written $J(x, y) = \left(-\frac{m(x, y)}{m(x', y')}\right)$. The latter is equivalent to $\int_V m(x, y) \, dx\, dy = \left(-\int_{L(V)} m(x', y') \, dx'\, dy'\right)$ for any region $V$ in $\mathbb{R}^2$ (anti measure-preservation corresponds to $L$ being measure-preserving and orientation-reversing).
$n(x, y)/d(x, y)$, with $n(x, y)$ and $d(x, y)$ coprime polynomials, i.e.
\[ (I \circ L)(x, y) = I(x', y') = \frac{n(x', y')}{d(x', y')} = \frac{n(x, y)}{d(x, y)} = I(x, y). \] (1)

A large 18-parameter class of these were given in [12] and take the form:
\[
\begin{align*}
x' &= \frac{f_1(y) - x f_2(y)}{f_2(y) - x f_3(y)}, \\
y' &= \frac{g_1(x') - y g_2(x')}{g_2(x') - y g_3(x')},
\end{align*}
\] (2)

where $f_i$ and $g_i$ are particular quartic polynomials. Let
\[
X := \begin{pmatrix} x^2 \\ x \\ 1 \end{pmatrix}, \quad Y := \begin{pmatrix} y^2 \\ y \\ 1 \end{pmatrix}, \quad A_i := \begin{pmatrix} \alpha_i & \beta_i & \gamma_i \\ \delta_i & \epsilon_i & \xi_i \\ \kappa_i & \lambda_i & \mu_i \end{pmatrix}, \quad i = 0, 1. \tag{3}
\]

Then the polynomials $f_i$ and $g_i$ can be neatly expressed as components of cross products:
\[
(f_1, f_2, f_3)(y) = (A_0 Y) \times (A_1 Y), \quad (g_1, g_2, g_3)(x') = (A_0^T X') \times (A_1^T X'). \tag{4}
\]

while the integral is the ratio of two biquadratics
\[
I(x, y) = B_0(x, y)/B_1(x, y) \tag{5}
\]

where
\[
B_i(x, y) = X \cdot A_i Y = \alpha_i x^2 y^2 + \beta_i x^2 y + \delta_i x y^2 + \gamma_i x^2 + \epsilon_i x y + \xi_i x + \lambda_i y + \mu_i, \quad i = 0, 1. \tag{6}
\]

It follows that the map (2) with (4) preserves the one-parameter family (or fibration) of algebraic curves
\[
B_0(x, y) + t B_1(x, y) = 0, \tag{7}
\]

by leaving invariant each curve for a given $t$ (i.e. it preserves the family curve-wise) [see Figure 3 (Left) for an example where 3 invariant curves are shown]. One can check that $1/B_1(x, y)$ is a preserved density for the dynamics and $1/B_0(x, y)$ is another one.2

The original proof of the integrability of (2)-(4) builds upon the fact that the map (2) is the composition $H \circ G$ with $H$ and $G$ being involutions (i.e. $H^2 = G^2 = \text{id}$):
\[
H: x' = x, \quad y' = \frac{g_1(x) - y g_2(x)}{g_2(x) - y g_3(x)}, \quad G: x' = \frac{f_1(y) - x f_2(y)}{f_2(y) - x f_3(y)}, \quad y' = y. \tag{8}
\]

This is the property of reversibility that we will discuss more in the next section. One checks that $H$ and $G$ preserve $I(x, y)$ of (5) [which is simplified in each case because $H$ and $G$ preserve one of the coordinates] and so does their composition. As a byproduct, one realises that if the matrices $A_i$ of (3) are taken to be symmetric, then the biquadratics of (6) are symmetric in $x$ and $y$ and so is $I(x, y)$. The latter can then be preserved by a composition of one of $H$ or $G$ with the switch $S: x' = y, \ y' = x$. In particular, we find $H \circ S$ gives the birational integrable map
\[
\begin{align*}
x' &= y, \quad y' = \frac{g_1(y) - x g_2(y)}{g_2(y) - x g_3(y)},
\end{align*}
\] (9)

2In fact, for a measure-preserving map with an integral, multiplying a density by any function of the integral gives another density. Conversely, for any measure-preserving map which has two densities that are not multiples of one another, the ratio of the two densities furnishes an integral.
In the literature, (2) and (9) with (4) are known, respectively, as the asymmetric and symmetric QRT maps [taking \(A_i\) as symmetric matrices in the latter case]. The symmetric maps (9) can be written as a second order difference equation by identifying \((x, y)\) with \((x_{n-1}, x_n)\) and \((x', y')\) with \((x_n, x_{n+1})\). A particular example is the Lyness map or difference equation:

\[
x' = y, \quad y' = \frac{y + \alpha}{x} \quad \Leftrightarrow \quad x_{n+1} x_{n-1} = x_n + \alpha, \tag{10}
\]

which is well-known in the difference equation community.

Another way to derive the symmetric and asymmetric QRT maps was provided in [2, 13], via a process of interchanging a parameter in a map with (the value) of an integral it preserves. To illustrate, an early example of an integrable planar map (which can also be written as a difference equation) is McMillan’s map [8]:

\[
x' = y \quad y' = -x - \frac{\beta y^2 + \epsilon y + \xi}{\alpha y^2 + \beta y + \gamma}, \tag{11}
\]

which is area-preserving and has the symmetric biquadratic integral

\[
B(x, y) = \alpha x^2 y^2 + \beta(x^2 y + xy^2) + \gamma(x^2 + y^2) + \epsilon xy + \xi(x + y) + \mu. \tag{12}
\]

McMillan’s map is contained as the special case of (9) when all entries of the matrix \(A_1\) are 0 except for \(\mu_1\). On the other hand, [2] shows how to obtain (9) directly from (11). Firstly, reparametrize the parameters in (11) and (12) by making them affine in a new parameter \(t\), so \(\alpha \mapsto \alpha_0 + \alpha_1 t\) etc. This corresponds to forming a one-parameter family of McMillan maps. Since \(B(x', y', t) = B(x, y, t)\), we can solve \(B(x, y, t) = 0\) for \(t\) and substitute it into the reparametrized version of (11), knowing that the resulting map will preserve the new integral ‘\(t\)’. The resulting map turns out to be (9) and the new integral is (the negative of) (5). This technique can be used to create new integrable maps from existing ones (e.g. in 3 dimensions, see [16]).

In [3], we gave an algebraic geometric description of birational maps, such as the symmetric and asymmetric QRT maps; that preserve a family of algebraic curves – they are conjugate to translations on particular elliptic surfaces (see also [1]). More precisely, let \(\mathbb{C}(t)\) be the field of rational functions over \(\mathbb{C}\) in the indeterminate \(t\), called the function field of the complex line, and let \(L\) be an infinite order birational map defined over \(\mathbb{C}(t)\) that preserves an elliptic curve over \(\mathbb{C}(t)\) (like (7)). Then on the associated Weierstrass curve over \(\mathbb{C}(t)\), \(L\) is conjugate to

\[
P \mapsto P + \Omega(t) \tag{13}
\]

where “+” is the associated group law on the Weierstrass cubic and \(P\) and \(\Omega(t) = (\omega_1(t), \omega_2(t))\), \(\omega_i(t) \in \mathbb{C}(t)\) are points on the cubic (for (2)-(4), the associated \(\Omega(t)\) can be deduced from [1, Proposition 2.5.6]). This is a strong characterization of the underlying ordered dynamics with various consequences, including that the dynamics on compact real level sets is conjugate to rotation. Studies of the rotation number as a function of the height of the level set can be found in [1]. Periodic orbits of period \(k\) occur on curves whose parameter value \(t\) satisfies \(k \Omega(t) = O\), where \(O\) is the identity element on the Weierstrass cubic. Generically, there will be a finite number of curves in the preserved fibration that are conics with genus 0 rather than genus 1. A characterisation of these singular curves and their geometry has been given in [11].
3 Reversible maps

A map $L$ is called reversible if there exists a map $G$ that conjugates it to its inverse:

$$G \circ L \circ G^{-1} = L^{-1}. \quad (14)$$

The map $G$ is called a reversing symmetry of $L$. The conjugacy (14) is a discrete version of time-reversal symmetry [5]. It implies that a trajectory of $L$ leads to another trajectory, its image under $G$, followed in the reverse time direction (possibly this is one and the same trajectory, then called symmetric, i.e. invariant under $G$). A particular case of (14) is when $G$ is an involution – in this case, reversibility is equivalent to saying that $L$ is the composition of the two involutions $H := L \circ G$ and $G$. We saw above in (8) that the QRT integrable maps were reversible in this sense. Indeed, birational maps preserving algebraic fibrations that are conjugate to the translation (13) via the result of [3] will always be reversible since (13) is the composition of the two involutions $P \mapsto -P + \Omega(t)$ and $P \mapsto -P$.

However, many non-integrable maps are reversible, as the product of involutions, particularly paradigm examples of symplectic maps like the standard map of the cylinder and Hénon’s area-preserving map:

$$L_{\text{hen}}: x' = y, \quad y' = -x + y^2 + \epsilon, \quad (15)$$

which is the composition $H_{\text{hen}} \circ G_{\text{hen}}$ with

$$H_{\text{hen}}: x' = x, \quad y' = -y + x^2 + \epsilon, \quad G_{\text{hen}}: x' = y, \quad y' = x. \quad (16)$$

Symmetries and time-reversal symmetries have played a fundamental role in dynamical systems with discrete time ([15, 5] and references therein), starting with Birkhoff’s foundational studies into the restricted three-body problem. One advantage of reversibility in a map is that symmetric periodic orbits must contain two points in $\text{Fix}(G)$, the fixed point set of $G$, two points in $\text{Fix}(H)$, the corresponding fixed point set of $H = L \circ G$ or one point in each of these fixed sets (the first two possibilities lead to the period being even and the last possibility to the period being odd). This characterisation often means that many such periodic orbits can be found by searching on these fixed sets rather than in the dimension of the full phase space. For instance, the fixed sets from (16) are evidently one-dimensional – $\text{Fix}(G_{\text{hen}})$ is the line $y = x$ and $\text{Fix}(H_{\text{hen}})$ is the parabola $y = (x^2 + \epsilon)/2$.

4 Reductions of birational maps over finite fields and the signatures of integrals and reversibility

There has been a recent interdisciplinary development of potentially great significance: the emergence of a dynamical systems perspective in algebra and arithmetic, and the parallel growth in the use of algebraic and arithmetical methods in dynamics. This interaction is loosely labelled arithmetic dynamics [20]. A particular aspect of arithmetic dynamics concerns dynamics over finite fields. The simplest example of a finite field is the set $\mathbb{F}_p$ of integers modulo a prime number $p$, which contains $p$ points [6]. Rational or birational maps can be represented over any mathematical field in which their coefficients can be so represented. When this representation can be achieved over finite fields, the $n$-dimensional birational map is said to be reduced over the finite space $\mathbb{F}_p^n$ of $p^n$ points. So, for instance, $L_{\text{hen}}$ gives an invertible map, a permutation, of the $p^2$ points of $\mathbb{F}_p^2$. 
for any rational value of $\epsilon$ whose denominator does not contain the prime $p$. Reduction of polynomials has been an important part of number theory for hundreds of years, whereas the reduction of the dynamics of rational maps is in its infancy. A big advantage is that over a finite field, the dynamics of a rational map can be described exactly and in full by a complete decomposition of the finite phase space into its constituent orbits. An important subclass of such orbits are the periodic orbits. The main dynamical questions relate to (periodic) orbit statistics (their number, length and distribution). What is lost in reduction is any retention of the topology of the continuum, i.e. the reduced map is not a spatial approximation of its continuum version.

In recent years, Franco Vivaldi and I have studied the implications upon reduction of key dynamical structures like integrability and symmetry. Starting in [17], we asked: If a rational map of the real or complex numbers has a structural property expressible in algebraic terms, like (1) or (14), does the inherited dynamics of this map over a finite field leave a signature of the property in the finite phase space? The ensuing program of research is summarised in Figure 1. Knowing the signature of a property after the reduction to finite fields can be used as a necessary condition to test for the property in a parametrized family of maps (discussed further below for the property of possession of a rational integral).

Our results suggest that there are different distributions of periods over finite fields for reductions of birational maps, dependent upon the algebraic property present and with universal (i.e. map-independent) aspects. We build period distribution functions via

$$D_p(x) = \frac{\# \{ z : T(z) \leq \kappa x \}}{\# \text{periodic points}}$$

(17)

where $T(z)$ is the minimal period of the periodic point $z$, and the constant $\kappa$ is a normalisation parameter to be determined below. Thus $D_p(x)$ is the proportion of periodic points of the reduction of the map to $\mathbb{F}_p^n$ that have normalised period $T(z)/\kappa$ less than or equal to $x$. One coarse measure is to consider the number of periodic orbits or cycles. In two dimensions, in the case that the birational map induces a permutation of the $p^2$ points, Figure 2 (Left) highlights the different behaviours found. There is a grading of number of cycles from integrable through reversible non-integrable through to irreversible non-integrable. So the presence of reversibility forces there to be more cycles than when this property is absent, whereas the presence of an integral forces even more cycles to arise than when reversibility is present on its own. Of course, the average length of cycles follows the opposite grading to the number of cycles.
Figure 2: Left: Growth of the number of cycles of the permutation of $p^2$ points with prime $p$, for the reduction of an integrable map of the plane (top curve), a reversible non-integrable map of the plane (middle), and an irreversible non-integrable planar map (bottom). The curves are conjectured to be asymptotic to $p \log(p)$, $p$, and $2 \log(p)$, respectively. Right: Period distributions for a reversible map. Shown is the experimental data for the Hénon map (15) at the prime $p = 997$ for a single choice of $\epsilon$ (the irregular curve), the experimental average over all $\epsilon$-values, and the theoretical distribution (19) (the smooth curves). At this resolution, the last two curves are indistinguishable.

Summarising our results in more detail:

(i) Irreversible non-integrable birational maps that induce permutations over finite fields appear to have period statistics consistent with a random permutation of the finite phase space [18]. The expected number of cycles in a random permutation of cardinality $N$ is $\log(N) + \gamma$, where $\gamma$ is Euler’s constant, so $2 \log(p) + \gamma$ when $N = p^2$ as in 2 dimensions. On the other hand, the period distribution (17) with $\kappa = N = p^2$ can be shown to be $D_p(x) = x$.

(ii) Birational reversible maps have a lower bound on the number of cycles on their reduction to finite fields. If the map $L$ satisfies (14) for some birational involution $G$ and does not commute with maps other than the powers $L^i$, the number of symmetric cycles can be proved [18] to be $[\#Fix(G) + \#Fix(L \circ G)]/2$, where $\#Fix(G)$ and $\#Fix(L \circ G)$ are the cardinalities of the reductions of the component involutions. Therefore

$$\#Cycles \geq \#SymmetricCycles = [\#Fix(G) + \#Fix(L \circ G)]/2.$$  \hspace{1cm} (18)

For example, for the reduction of $L_{Hen}$ of (15), one sees from (16) and the discussion following that $\#Fix(G_{Hen}) = \#Fix(H_{Hen}) = p$ so there are $p$ symmetric cycles, and so at least $p$ cycles. The result of [18] is one about reversible permutations so it is applicable to reductions of birational maps in any dimension that yield permutations. Furthermore, we have found that (18) is a very tight bound, as illustrated for (15) in Figure 2 (Left) which generates the number of cycles curve very close to the diagonal. There are very few asymmetric cycles, asymptotically none in percentage in the large $p$ limit so that the inequality in (18) then becomes an equality. The period distribution for reductions of the reversible maps considered is found via (17) to be

$$D_p(x) = R(x) = 1 - e^{-x} (1 + x),$$  \hspace{1cm} (19)
in the large $p$ limit, with $\kappa$ taken to be the average cycle length (see Figure 2 (Right)). Although we have not proved that (19) is the period distribution for a specific reversible map, [10] has recently shown experimentally that a particularly simple reversible map of the torus which is piecewise constant has period distribution $R(x)$ and that this distribution arises as a limit of singular period distributions. Furthermore, [19] has taken a probabilistic approach to the problem in line with Topic 3 of Figure 1. There it is shown that $R(x)$ is the expected period distribution of the composition of a pair of random involutory permutations of a set of cardinality $N$ under mild conditions on the growth of the cardinality of the fixed sets of the involutions as $N \to \infty$. Finally, in [9] we have shown that (19) also governs the period distributions in reversible rational maps, where singularities arise in the denominators leading to aperiodic orbits as well.

(iii) In two dimensions, if a birational map has a rational integral as in (1) which reduces over a finite field, the level set becomes, typically, an algebraic curve of genus 1 over $\mathbb{F}_p^2$ (more correctly, over its projective version). The number of points on such a curve is restricted by the Hasse-Weil bound [14]

$$p + 1 + 2\sqrt{p},$$

which also represents a bound on the length of orbits of the integrable map over a finite field. Indeed, orbit length observing the bound (20) provides a very effective necessary condition for the existence of an algebraic invariant as indicated in Figure 3. A naive idea in $\mathbb{F}_p^2$ might be that one algebraic integral divides the $p^2$ points into $p$ level sets of approximately $p$ points, so that there are at least $p$ orbits. Computationally, we find more: $p \log p$ cycles when the integrable map induces a permutation – see Figure 2 (Left). This exceeds the $p$ (symmetric) cycles expected from the reversibility of the integrable birational map. This excess comes from asymmetric cycles so these are much more plentiful in two dimensional integrable reversible maps than in their non-integrable reversible perturbations. This can be understood from the algebraic geometry we discussed in Section 2 and the conjugacy to the translation (13). The period on each level set is a reflection of the order of the translative element $\Omega(t)$ reduced over the finite field. In the case of non-maximal order, the level curve is filled with copies of cycles of the same period, most of which are asymmetric. The bound (20) and these considerations explain why the period distribution function (17) for reduced integrable maps (with $\kappa$ taken as (20)) is found to be singular with steps at the reciprocals of the integers [3].

A current research interest of ours is to study reductions over finite fields of birational maps in higher dimensions with integrals and to develop an understanding of the contraints on orbit length and the number of orbits (with a view to being able to test birational maps for the possession of one or more integrals).

Acknowledgements

The more recent part of the work described here, since 2007, has been supported by the Australian Research Council.

References

Figure 3: Results for birational maps that reduce to permutations over finite fields $\mathbb{F}_p$. Left: Part of the real phase portrait of a planar birational map with a rational integral, showing the orbits of the initial conditions $(1,0), (1,1)$ and $(1,2)$. Middle: Lengths of the orbit of $(1,1)$ in $\mathbb{P}(\mathbb{F}_p^2)$, divided by the Hasse-Weil bound (20), plotted as a function of $p$ for the reduction of a one-parameter family of birational maps in the case of the integrable parameter value $\epsilon = 0$, whose phase portrait was shown at left. Right: Normalised length plot again for $(1,1)$ as for the Middle figure but now for a non-integrable perturbation $\epsilon = 10^{-4}$. Whilst the phase portrait of this perturbation is virtually indistinguishable from the Left figure, the generic orbit lengths are now well in excess of the Hasse-Weil bound [there is no contradiction since the invariant KAM curves in the perturbation are not algebraic curves].


