Normal bivariate Birkhoff interpolation schemes and Pell equation

Marius Crainic, Nicolae Crainic

Abstract. Finding the normal Birkhoff interpolation schemes where the interpolation space and the set of derivatives both have a given regular "shape" often amounts to number-theoretic equations. In this paper we discuss the relevance of the Pell equation to the normality of bivariate schemes for different types of "shapes". In particular, when looking at triangular shapes, we will see that the conjecture in Lorentz R.A., *Multivariate Birkhoff Interpolation*, Lecture Notes in Mathematics, 1516, Springer, Berlin-Heidelberg, 1992, is not satisfied, and, at the same time, we will describe the complete solution.

Keywords: Birkhoff interpolation, Pell equation

Classification: 65D05, 11D09

1. Normal schemes

The bivariate Birkhoff interpolation problem depends on a finite set $S \subset \mathbb{N}^2$ which is a lower subset, a finite subset $A \subset \mathbb{N}^2$ (describing the derivatives), and a finite set $Z = \{z_1, \ldots, z_m\} \subset \mathbb{R}^2$ of nodes. Here $\mathbb{N}$ stands for the set of pairs of non-negative integers, while "lower subset" means

$$(i_0, i_1) \in S \Longrightarrow R\,(i_0, i_1) \subset S,$$

where $R(s,t)$ is the rectangle

$$R(s,t) = \left\{(u,v) \in \mathbb{N}^2 : u \leq s, v \leq t\right\}.$$

The role of the set S is to describe the "interpolation space"

$$\mathcal{P}_S = \Big\{ \sum_{(i,j)\in S} a_{i,j} x^i y^j : a_{i,j} \in \mathbb{R} \Big\},$$

and the condition that S is lower ensures that (and is equivalent to) $\mathcal{P}_S$ is invariant under transformations of type $(x,y) \mapsto (ax+b, cy+d)$. The basic examples of lower sets are the rectangles $R(s,t)$ and the triangles $T_n(a,b)$:

$$T_n(a,b) = \left\{(u,v) \in \mathbb{N}^2 : au + bv \leq n\right\}.$$

The regular versions of these are the squares $R_n = R(n,n)$, and the triangles $T_n = T_n(1,1)$.

The interpolation problem associated to (Z, S, A) consists of finding polynomials $P \in \mathcal{P}_S$ satisfying

$$\frac{\partial^{\alpha+\beta} P}{\partial x^\alpha \partial y^\beta}(z_i) = c_i^{\alpha,\beta}, \ \forall\, (\alpha, \beta) \in A, \ 1 \leq i \leq k,$$

where $c_i^{\alpha,\beta} \in \mathbb{R}$ are arbitrary constants. Central to this theory is the notion of *regularity* of (Z, S, A) which means that the problem has a unique solution for all choices of the constants $c_i^{\alpha,\beta}$, and that of *almost regularity* of (S, A) (with respect to sets of m nodes) which means that (Z, S, A) is regular for at least one choice of Z (which, in turn, implies that this happens for almost all choices of $Z \subset \mathbb{R}^{2m}$). The *normality* condition appears as the simplest necessary condition for (almost) regularity. First recall that

Definition 1.1. We say that (Z, S, A) is normal (or that (S, A) is normal with respect to sets of m nodes) if $|S| = |Z||A|$ (or $|S| = m|A|$, respectively).

The relationship with regularity is immediate by a counting dimension argument. Indeed, (Z, S, A) determines a linear map from $\mathcal{P}_S$ into the $|Z||A|$-dimensional euclidean space (which associates to P the evaluation of the derivatives which appear in the interpolation equations), and the regularity condition is equivalent to the bijectivity of this map. In particular, since $\dim(\mathcal{P}_S) = |S|$, one obtains the well known

Lemma 1.2. *If (S, A) is almost regular, then it must be a normal scheme.*

2. The Pell equation

In this section we recall some basic facts about the Pell equation

$$\alpha^2 - d\beta^2 = 1, \tag{2.1}$$

where $d \geq 1$ is a given integer with $\sqrt{d} \notin \mathbb{Z}$, and the unknowns are the pairs $(\alpha, \beta) \in \mathbb{N}^2$ with $\beta \neq 0$. First of all, it is easy to see that starting from a solution (α_0, β_0) one can construct a sequence (α_n, β_n) of new solutions defined by

$$\begin{cases} \alpha_{n+1} = \alpha_0 \alpha_n + d\beta_0 \beta_n \\ \beta_{n+1} = \beta_0 \alpha_n + \alpha_0 \beta_n \end{cases} \tag{2.2}$$

We now assume that (α_0, β_0) is the minimal solution, where "minimal" is with respect to the well known order on $\mathbb{N}^2$:

$$(\alpha, \beta) \leq (\alpha', \beta') \Longleftrightarrow \alpha < \alpha', \ \text{ or } \ \alpha = \alpha', \beta \leq \beta'.$$

The point of starting with the minimal solution is:

Theorem 2.1. *If* (α_0, β_0) *is, as above, the minimal solution of* (2.1), *then*

$$\{(\alpha_n, \beta_n) : n \geq 0\}$$

coincides with the set of all solutions of (2.1).

The proof is elementary and, as it often happens in mathematics, it is sometimes more useful to know the proof then the result itself (since the idea of the proof can be adapted to various other equations). For this reason, we recall (and will use) the proof.

PROOF: Let $\mathcal{S}$ be the set of all solutions of (2.1). We now make two remarks:

(i) if $(\alpha, \beta) \in \mathcal{S}$, then $(\alpha_0\alpha + d\beta_0\beta, \alpha_0\beta + \beta_0\alpha) \in \mathcal{S}$;

(ii) if $(\alpha, \beta) \in \mathcal{S} \setminus \{(\alpha_0, beta_0)\}$, then $(\alpha_0\alpha - d\beta_0\beta, \alpha_0\beta - \beta_0\alpha) \in \mathcal{S}$ is a new solution strictly smaller then (α, β).

We have already mentioned (and used) (i) when constructing the sequence (α_n, β_n) above. Also the proof of (ii) is straightforward. For instance, to check that $\alpha_0\beta - \beta_0\alpha > 0$, we rewrite it as $d^2\alpha_0^2\beta^2 > d^2\beta_0^2\alpha^2$, and this is equivalent to $\alpha_0^2(\alpha^2 - 1) > \alpha^2(\alpha_0^2 - 1)$, or $\alpha^2 > \alpha_0^2$, and this follows from the minimality of (α_0, β_0). We have already remarked that starting from (α_0, β_0) and applying (i) repeatedly, one gets the sequence (α_n, β_n). Similarly, starting from an arbitrary solution (a_0, b_0) and applying (ii) repeatedly one gets a sequence (a_n, b_n) of new solutions of (2.1), which are well defined as long as $(a_n, b_n) \neq (\alpha_0, \beta_0)$. But, since the sequence (a_n, b_n) is strictly decreasing (cf. (ii)), it must stop at some point, hence there is an integer n such that $(a_n, b_n) = (\alpha_0, \beta_0)$. On the other hand, the operations in (i) and (ii) above are inverse to each other, hence we must have $(a_0, b_0) = (\alpha_n, \beta_n)$, i.e. any $(a_0, b_0) \in \mathcal{S}$ belongs to our sequence. □

Remark 2.2. We will also see variations of the Pell equation (2.1) of type

$$ax^2 - by^2 = r,$$

where $a, b, r \in \mathbb{N}$, $\sqrt{ab} \notin \mathbb{Z}$ are given, and the unknown are the pairs $(x, y) \in \mathbb{N}^2$. We associate to this equation the auxiliary Pell equation (2.1) with $d = ab$. It is easy to see that, for any solution (α, β) of the auxiliary equation and any solution (x, y) of our equation, one obtains a new solution (x', y') of our equations by:

$$x' = \alpha x + b\beta y, \; y' = a\beta x + \alpha y.$$

In the case $r = 1$, the previous theorem tells us that starting from the minimal solution and applying this transformation repeatedly one obtains all solutions of (2.1). Such a theorem is no longer valid for the more general equations $ax^2 - by^2 = r$ with $r \geq 2$.

3. Finding normal schemes

In this section we will use the Pell equation for finding the normal schemes (S, A) when S and A have a given regular shape.

An immediate application of the previous section is

Proposition 3.1. *If $\sqrt{2m} \notin \mathbb{Z}$, then the schemes (T_n, R_p) which are normal with respect to sets of m nodes are in one-to-one correspondence with the solutions of the Pell equation $\alpha^2 - d\beta^2 = 1$ with $d = 8m$. The correspondence is given by*

$$n = \frac{\alpha - 3}{2}\,,\ p = \beta - 1.$$

If $\sqrt{2m} \in \mathbb{Z}$, then there are no such normal schemes.

Proof: Since $|T_n| = \frac{(n+1)(n+2)}{2}$ and $|R_p| = (p+1)^2$, the normality condition $|T_n| = m|R_p|$ can be rewritten as $(2n+3)^2 - 8m(p+1)^2 = 1$, hence the statement. □

Proposition 3.2. *Write $2m = a^2 m_0$ where $a \in \mathbb{Z}$, and m_0 is a square-free integer (i.e. it is never divisible by a square of a positive integer other then 1).*

Then the schemes (R_n, T_p) which are normal with respect to sets of m nodes

(i) *if m_0 is even, they correspond to the solutions of the Pell equation $\alpha^2 - d\beta^2 = 1$ with $d = m_0$, where the correspondence is given by*

$$n = \frac{am_0\beta}{4} - 1,\ p = \frac{\alpha - 3}{2}\,;$$

(ii) *if $m_0 \neq 1$ is odd, they correspond to the solutions of the Pell equation $\alpha^2 - d\beta^2 = 1$ with $d = 4m_0$, where the correspondence is given by*

$$n = \frac{am_0\beta}{2} - 1,\ p = \frac{\alpha - 3}{2}\,;$$

(iii) *if $m_0 = 1$, there are no such normal schemes.*

Proof: We rewrite the equation $|R_n| = m|T_p|$ as

$$(4n+4)^2 - 2m(2p+3)^2 = -2m.$$

We then see that $a^2 m_0$ must divide $(4n+4)^2$, hence $4n + 4 = am_0 y$ for some integer y. Denoting $x = (2p+3)$, we obtain the new equation $x^2 - m_0 y^2 = 1$. If m_0 is even, then, starting with any solution (x, y) of this equation, we can go back and get a solution of our equation. To see this, we first remark that $x^2 = m_0 y^2 + 1 \equiv 1$ modulo 2, hence x is odd. On the other hand, this implies $x^2 \equiv 1$ modulo 8, i.e. $m_0 y^2 \equiv 0$ modulo 8, hence, since m_0 is not divisible by 4 (it is square-free), it follows that $y \equiv 0$ modulo 2. This shows that $n = \frac{am_0\beta}{4} - 1,\ p = \frac{\alpha-3}{2}$ are indeed integers. This proves (i). When m_0 is odd, it is not true that any solution of $x^2 - m_0 y^2 = 1$ gives a solution to our equation. The condition is that x is odd and $am_0 y \equiv 0$ modulo 4. One can actually see that the first condition implies the second. Indeed, since $x^2 = m_0 y^2 + 1 \equiv y^2 + 1$ modulo 2, y must be even, and then, since $a^2 m_0 = 2m$ is even, the second condition follows immediately. We then see that the only condition is $y \equiv 0$ modulo 2. In other words, the relevant equation is $x^2 - 4m_0z^2 = 1$, and the correspondence with the initial equation is $x = 2p + 3, z = 2y = \frac{4(n+1)}{am_0}$. This proves (ii), while (iii) is immediate. □

Before going into the general discussion on the scheme (T_n, T_p), let us first consider the particular case $m = 3$ which gives us some idea of what is going on in general, and which is also a good illustration of our Remark 2.2.

Proposition 3.3. *The schemes (T_n, T_p) which are normal with respect to sets of three nodes are only those with $n = n_k, p = p_k$ for some integer k, where the sequence (n_k, p_k) is defined by the recurrence*

$$\begin{cases} n_{k+1} = 2n_k + 3p_k + 6 \\ p_{k+1} = n_k + 2p_k + 3 \end{cases} \tag{3.1}$$

and starts with $n_0 = 8$, $p_0 = 4$.

Proof: The equation $|T_n| = 3|T_p|$ can be written as $(2n+3)^2-1 = 3((2p+3)^2-1)$. Hence we have to consider the solutions (x, y) of the equation

$$x^2 - 3y^2 = -2$$

with the property that $x > 3$, $y > 3$ and x and y are both odd. Let us now look at all positive solutions of this equation. Inspired by Remark 2.2 we see that, as in the proof of Theorem 2.1, one has (using the minimal solution $\alpha = 2$, $\beta = 1$ of the auxiliary equation $\alpha^2 - 3\beta^2 = 1$)

(i) if (x, y) is solution, then so is $(2x + 3y, x + 2y)$,
(ii) if (x, y) is a solution with $y > 2$, then $(2x - 3y, 2y - x)$ is another solution strictly smaller then (x, y),
(iii) there is only one solution with $y \leq 2$, namely $x = y = 1$.

One can then proceed exactly as in the proof of Theorem 2.1 and conclude that the positive solutions of $x^2 - 3y^2 = -2$ are given by the sequence (x_k, y_k) with $x_0 = y_0 = 1$, and $x_{k+1} = 2x_k + 3y_k$, $y_{k+1} = x_k + 2y_k$. We immediately see that the x_k's and the y_k's are all odd, and the first relevant solutions are $x_2 = 19$, $y_2 = 11$. Writing the previous recurrences in terms of n and p, one obtains recurrences in the statement. □

In general, if m is not a square, we have

Proposition 3.4. *If $\sqrt{m} \notin Z$, then (all) the schemes (T_n, T_p) which are normal with respect to sets of m nodes are obtained as follows.*

We denote by (α_0, β_0) the minimal solution of the Pell equation $\alpha^2 - m\beta^2 = 1$, and by $\{(x^1, y^1), \ldots, (x^r, y^r)\}$ all the solutions of the equation $x^2 - my^2 = 1 - m$ satisfying

$$y \leq \alpha_0 \sqrt{\frac{m-1}{m}}.$$

For each such solution (x^i, y^i)*, we define a sequence* $(x^i_k, y^i_k)_{k\geq 0}$ *starting at* (x^i, y^i) *and satisfying*

$$\left\{\begin{array}{l} x^i_{k+1} = \alpha_0 x^i_k + m\beta_0 y^i_k \\ y^i_{k+1} = \alpha_0 y^i_k + \beta_0 x^i_k \end{array}\right. \tag{3.2}$$

Then (T_n, T_p) *is normal with respect to sets of* m *nodes if and only if* $(2n+3,$ $2p+3)$ *belongs to the set*

$$\left\{(x^i_k, y^i_k) : 1 \leq i \leq r, k \geq 0\right\}.$$

Proof: We proceed as in the proof of the previous proposition and of Theorem 2.1, and remark that

(i) if (x, y) is solution, then so is $(\alpha_0 x + m\beta_0 y, \alpha_0 y + m\beta_0 x)$,

(ii) if (x, y) is a solution with $y > \alpha_0\sqrt{\frac{m-1}{m}}$, then $(\alpha_0 x - m\beta_0 y, \alpha_0 y - m\beta_0 x)$ is another solution strictly smaller then (x, y).

The bound on y is the one that ensures that the new solution is positive. Indeed, $\alpha_0 x > m\beta_0 y$ is equivalent to (square it, and use that $m\beta_0^2 = \alpha_0^2 - 1$, $my^2 = x^2 + (m-1)$)

$$\alpha_0^2 x^2 > \alpha_0^2 x^2 + (m-1)\alpha_0^2 - x^2 - (m-1),$$

i.e. $my^2 - (m-1) = x^2 > (m-1)(\alpha^2 - 1)$ which is precisely the bound on y appearing in the statement. All the other verifications are straightforward. □

Example 3.5. The new pattern that this theorem reveals is that there might be several different sequences that control the normality equation, and that is because the equation $x^2 - my^2 = 1 - m$ might have more then one solution with $y \leq \alpha_0\sqrt{\frac{m-1}{m}}$. This has not been present in the case $m = 3$ (Proposition 3.3). The first case when this is present (and relevant for normality) is when $m = 6$. Then the bound on y is $y \leq 4$, and there are two such solutions of the equation corresponding to this case (i.e. $x^2 - 6y^2 = -5$), namely $(1, 1)$ and $(7, 3)$. Note also that the minimal solution of the corresponding Pell equation is $(5, 2)$. We immediately deduce that the schemes (T_n, T_p) which are normal with respect to sets of six nodes correspond to two sequences. Both are defined by the recurrence

$$\left\{\begin{array}{l} n_{k+1} = 5n_k + 12p_k + 24 \\ p_{k+1} = 2n_k + 5p_k + 9 \end{array}\right. \tag{3.3}$$

the first one starts with $(7, 2)$, while the second one starts with $(34, 13)$.

We now have to consider the case where $m = l^2$ with $l \in \mathbb{N}$. The idea is to relate the normality condition to the Pell equation on $(2n+3, l)$ (and not $(2n+3, 2p+3)$ as before). We have

Proposition 3.6. *The scheme (T_n, T_p) is normal with respect to sets of l^2 nodes if and only if (n, p, l) is of type*

$$n = \frac{(\sqrt{p+2}+\sqrt{p+1})^{2k+2} + (\sqrt{p+2}-\sqrt{p+1})^{2k+2} - 6}{4}$$

$$l = \frac{(\sqrt{p+2}+\sqrt{p+1})^{2k+2} - (\sqrt{p+2}-\sqrt{p+1})^{2k+2}}{4\sqrt{(p+1)(p+2)}}$$

with $k \geq 0$ (arbitrary) integer.

Proof: Denoting $\alpha = 2n+3$, $\beta = l$, the normality condition $|T_n| = l^2|T_p|$ is equivalent to the Pell equation $\alpha^2 - d\beta^2 = 1$ with

$$d = 4(p+1)(p+2).$$

Moreover, we do know the minimal solution of this equation, namely $(2p+3, 1)$. It follows that the general solution is induced by the recurrences

$$\left\{ \begin{array}{l} \alpha_{k+1} = (2p+3)\alpha_k + 4(p+1)(p+2)\beta_k \\ \beta_{k+1} = \alpha_k + (2p+3)\beta_k \end{array} \right. \tag{3.4}$$

We then see that both α_k and β_k are given by the recurrence $t_{k+2} = 2(2p+3)t_{k+1} - t_k$, and the initial conditions are $\alpha_0 = 2p+3$, $\alpha_1 = 2(2p+3)^2-1$, $\beta_0 = 1$, $\beta_1 = 2(2p+3)$. It then follows that

$$\alpha_k = \frac{(\sqrt{p+2}+\sqrt{p+1})^{2k+2} + (\sqrt{p+2}-\sqrt{p+1})^{2k+2}}{2}$$

$$\beta_k = \frac{(\sqrt{p+2}+\sqrt{p+1})^{2k+2} - (\sqrt{p+2}-\sqrt{p+1})^{2k+2}}{4\sqrt{(p+1)(p+2)}}$$

and the proposition follows. □

Remark 3.7. The first non-obvious solutions are (n_1, p, l_1) with

$$n_1 = (2p+3)^2 - 2, l_1 = 2(2p+3).$$

When saying "non-obvious", we exclude the solution $(p, p, 1)$ (i.e. schemes with one node only). In [2] it has been conjectured that these are all the non-obvious solutions. Of course, the previous proposition shows that this is not the case.

References

[1] Barbeau E.J., *Pell's Equation*, Springer, New York, 2003.
[2] Lorentz R.A., *Multivariate Birkhoff Interpolation*, Lecture Notes in Mathematics, 1516, Springer, Berlin-Heidelberg, 1992.
[3] Gasca M., Maeztu J.I., *On Lagrange and Hermite interpolation in* $\mathbb{R}^n$, Numer. Math. **39** (1982), 1–14.
[4] Stillwell J., *Elements of number theory*, Undergraduate Texts in Mathematics, Springer, New York, 2003.

Department of Mathematics, Utrecht University, 3508 TA Utrecht, The Netherlands
Email: crainic@math.uu.nl

1 Decembrie 1918 University, Alba Iulia, Romania

(*Received* June 24, 2008, *revised* March 19, 2009)