A General Lagrange Theorem

Giovanni Panti

1. INTRODUCTION The ordinary continued fractions expansion of a real number is based on the Euclidean algorithm. Variants of the latter yield variants of the former, all encompassed by a more general dynamical systems framework. For all these variants the Lagrange theorem holds: a number has an eventually periodic expansion if and only if it is a quadratic irrational. This fact is surely known for specific expansions, but the only proof for the general case that I could trace in the literature follows as an implicit corollary from much deeper results by Boshernitzan and Carroll on interval exchange transformations [2]. It may then be useful to have at hand a simple and virtually computation-free proof of a general Lagrange theorem.

Let \( D \) be any unimodular partition of the real unit interval. By this we mean that \( D \) is a family (finite or countable, of cardinality at least 2) of half-open intervals \( \Delta_a = (p/q, r/s] \), with \( a \) varying in a fixed index set \( I \subseteq \mathbb{Z} \), such that:

(i) each interval \( \Delta_a \in D \) is unimodular, i.e., its extrema \( 0 \leq p/q < r/s \leq 1 \) are rational numbers (always written in reduced form), and \( \begin{vmatrix} r & s \\ q & s \end{vmatrix} = -1 \) (this amounts to saying that the column vectors \( (p q)^t \) and \( (r s)^t \) constitute a \( \mathbb{Z} \)-basis for \( \mathbb{Z}^2 \));

(ii) two distinct intervals in \( D \) have empty intersection;

(iii) the set \( X = \bigcup D \setminus \{1\} \) contains all irrational numbers in \([0, 1]\).

Fix an arbitrary function \( \varepsilon : I \to \{-1, +1\} \), let \( \Delta_a = (p/q, r/s] \in D \), and let

\[
G_a = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}^{\frac{(\varepsilon(a)+1)}{2}} \begin{pmatrix} p \\ q \end{pmatrix} \begin{pmatrix} r \\ s \end{pmatrix}^{-1}.
\]

Any matrix \( T = \begin{pmatrix} t & u \\ v & w \end{pmatrix} \) in \( \text{GL}_2\mathbb{Z} \) induces a fractional-linear homeomorphism \( \phi_T \) of the projective real line \( \mathbb{R} \cup \{\infty\} \) onto itself via \( \phi_T(x) = (tx + u)/(vx + w) \). Using projective coordinates \((x, y)^t \) to represent the real number \( x/y \) (with \((1, 0)^t \) representing \( \infty \)), \( \phi_T \) amounts to multiplication on the left by \( T \). We apply the above to \( T = G_a \), observing that \( G_a(p q)^t \) and \( G_a(r s)^t \) are \((0 1)^t \) and \((1 1)^t \) (in one order or the other, depending on the value of \( \varepsilon(a) \)). We can then check easily that \( \phi_{G_a} \mid \Delta_a \) is monotonically increasing with range \((0, 1]\) if \( \varepsilon(a) = -1 \), and monotonically decreasing with range \([0, 1)\) if \( \varepsilon(a) = +1 \). The resulting piecewise-fractional map \( G : X \to [0, 1] \), defined by \( Gx = \phi_{G_a}(x) \) for \( x \in \Delta_a \), is the Gauss map determined by \( D \) and \( \varepsilon \).

Denote by \( \psi_a = \phi_{G_a^{-1}} \mid [0, 1] \) the \( a \)th inverse branch of \( G \); the range of \( \psi_a \) is the topological closure of \( \Delta_a \). Computing \( G_a^{-1} \), we obtain explicitly

\[
\psi_a(x) = \begin{cases} 
\frac{(r - p)x + p}{(s - q)x + q}, & \text{if } \varepsilon(a) = -1; \\
\frac{(p - r)x + r}{(q - s)x + s}, & \text{if } \varepsilon(a) = +1.
\end{cases}
\]

If \( x, Gx, G^2x, \ldots, G^{n-1}x \) are all in \( X \) then, letting \( G^{t-1}x \in \Delta_a \), for \( 1 \leq t \leq n \), we have the identity

\[
x = \psi_{a_1} \psi_{a_2} \cdots \psi_{a_n}(G^n x).
\]
In dynamical systems language, the (finite or infinite) sequence $a_1, a_2, \ldots$ is the symbolic sequence of $x$. By (1), if $x$ has a finite symbolic sequence (i.e., $G^n x \in [0, 1] \ \backslash \ X \subset \mathbb{Q}$ for some $n$), then $x$ is rational. Conversely, let $x = u/v \in \Delta_a$ be rational, and let $p/q < r/s$ be the extrema of $\Delta_a$. Then $(u \ v)^t = l(p \ q)^t + m(r \ s)^t$ for some nonnegative integers $l, m$ with $m \geq 1$. Multiplying by $G_a$ to the left, we see that the denominator of $Gx$ equals $l + m$, which is strictly less than the denominator $v = lq + ms$ of $x$. Hence the sequence of the denominators along the $G$-orbit of $x$ is strictly decreasing, and $x$ must eventually leave $X$ (note that 0 and 1 are the only rational points having denominator 1, and neither of them is in $X$).

**Example 1.** The first and main example is of course given by ordinary continued fractions: $I = \{1, 2, \ldots\}$, $\Delta_a = \{1/(a + 1), 1/a\}$, $X = (0, 1)$, and $\varepsilon = +1$ throughout. The Gauss map is $Gx = 1/x - [1/x]$, and its graph is shown in Figure 1.

![Figure 1.](image1)

We have $\psi_a(x) = 1/(a + x)$, and hence (1) assumes the familiar shape

$$x = \frac{1}{a_1 + \frac{1}{a_2 + \frac{1}{\ddots + \frac{1}{a_n + G^n x}}}}.$$

**Example 2.** In odd continued fractions we have $I$, $X$, and the $\Delta_a$’s as above, and $\varepsilon(a) = (-1)^{a+1}$. The Gauss map is $Gx = |1/x - b|$, where $b$ is the odd integer in $\{[1/x], [1/x] + 1\}$, and its graph is shown in Figure 2.

![Figure 2.](image2)
Setting \((b_t, \varepsilon_t)\) equal to \((a_t, +1)\) for \(a_t = \lfloor 1/G^{t-1}x \rfloor\) odd, and to \((a_t + 1, -1)\) for \(a_t\) even, we have \(\psi_{a_t}(x) = 1/(b_t + \varepsilon_t x)\), and therefore

\[
x = \frac{1}{b_1 + \frac{\varepsilon_1}{b_2 + \frac{\varepsilon_2}{\ddots + \frac{\varepsilon_{n-1}}{b_n + \varepsilon_n G^nx}}}}.
\]

Among the continued fractions algorithms covered by our framework are the even continued fractions, the nearest integer fractions, the Farey fractions, and so on; see [1, 7] for more examples and a detailed analysis of the ergodic properties of the corresponding Gauss maps. Other algorithms—such as various versions of the cyclic method discovered by the Indians in the 12th century for solving the Pell equation—are not covered; see [8, §2-3] and references therein for these algorithms.

2. CONVERGENCE . . . Unimodular intervals have a curious property.

Observation 3. Let \(\Gamma_1 \supset \Gamma_2 \supset \cdots\) be a nested sequence of closed unimodular intervals, each one strictly contained in the previous one. Then

\[
\lim_{n \to \infty} \text{length}(\Gamma_n) = 0.
\]

Proof. The length of \(\Gamma_n = [p_n/q_n, r_n/s_n]\) is \((q_n s_n)^{-1}\), which is less than or equal to \([\max(q_n, s_n)]^{-1}\). If \(u/v\) is a rational in the topological interior of \(\Gamma_n\), then \(v\) is strictly greater than \(\max(q_n, s_n)\) (again, because \((u/v)' = l(p_n q_n)' + m(r_n s_n)'\), for certain positive \(l, m \in \mathbb{Z}\)). It follows that the sequence \(\max(q_1, s_1), \max(q_2, s_2), \ldots\) is strictly increasing and goes to infinity, and therefore the sequence of the reciprocals, which bounds above the sequence of the lengths, goes to 0.

As a consequence, we have:

(i) the map \(k : [0, 1] \setminus \mathbb{Q} \to I^\mathbb{N}\) that associates to each irrational its symbolic sequence is injective, continuous, and open;

(ii) if \(x\) is irrational with symbolic sequence \(a_1, a_2, \ldots\), and \(y \in [0, 1]\), then

\[
\lim_{n \to \infty} \psi_{a_1} \cdots \psi_{a_n}(y) = x.
\]

Indeed, let \(k(x) = a_1, a_2, \ldots\); by §1(1), the set of all numbers whose symbolic sequence agrees with \(k(x)\) up to \(a_n\) is contained in the closed unimodular interval \(\Gamma_n\) whose extrema are \(\psi_{a_1} \cdots \psi_{a_n}(0)\) and \(\psi_{a_1} \cdots \psi_{a_n}(1)\). Since each \(\Delta_n \in D\) is properly contained in \((0, 1]\), the inclusions \(\Gamma_1 \supset \Gamma_2 \supset \cdots\) are proper, and Observation 3 implies (ii) and the injectivity of \(k\). We leave the continuity and openness of \(k\) as an exercise for the reader; \(k\) may be surjective (e.g., in the case of ordinary continued fractions), but is never so if \(I\) is finite (because then \(I^\mathbb{N}\) is compact, while \([0, 1] \setminus \mathbb{Q}\) is not).
3. ... AND PERIODICITY. On August 25th, 1769, Joseph Louis Lagrange read at the Royal Berlin Academy of Science a Mémoire on the resolution of algebraic equations via continued fractions [5]. He had been in Berlin since 1766, as director of the mathematical section of the same Academy, on the recommendation of his predecessor, Leonhard Euler. In this Mémoire he proved that the irrational \( x \) has an eventually periodic continued fractions expansion if and only if it has degree two over the rationals. The “only if” implication is clear even in our general setting. If \( x \) has symbolic sequence \( a_1, \ldots, a_t, a_{t+1}, \ldots, a_{t+r} \) under any Gauss map \( G \), then \( G^t x \) and \( G^{t+r} x \) have the same symbolic sequence, namely \( a_{t+1}, \ldots, a_{t+r} \), and hence are equal. This implies that the vectors

\[
G_{a_t} \cdots G_{a_1} \begin{pmatrix} x \\ 1 \end{pmatrix}
\]

and

\[
G_{a_{t+r}} \cdots G_{a_{t+1}} G_{a_t} \cdots G_{a_1} \begin{pmatrix} x \\ 1 \end{pmatrix}
\]

are projectively equal, i.e., differ by a nonzero multiplicative constant. But then \((x \ 1)^t\) is an eigenvector for the integer matrix \((G_{a_t} \cdots G_{a_1})^{-1} G_{a_{t+r}} \cdots G_{a_{t+1}} G_{a_t} \cdots G_{a_1}\), and hence \( x \) is quadratic.

The “if” direction is trickier. The usual proof [4], [6] is basically Lagrange’s. One considers the minimal polynomial \( c_n x^2 + d_n X + e_n \in \mathbb{Z}[X] \) of \( 1/G^n x \), and shows that, from some \( n \) on, the coefficients \( c_n, d_n, e_n \) must satisfy certain inequalities. Sometimes this proof is reworded by writing \( 1/G^n x \) in reduced form

\[
\frac{1}{G^n x} = \frac{P_n + \sqrt{D}}{Q_n}.
\]

and again bounding \( P_n \) and \( Q_n \) in terms of the common discriminant \( D \) of the above polynomials, for \( n \) large enough. All of this is tightly related to Gauss’s theory of reduced quadratic forms; see [3, §5.7].

It is not clear how to adapt the above proof to the case of arbitrary piecewise-fractional expansions. I will switch to a more geometric vein by formulating another observation.

**Observation 4.** Let \( x \) be a quadratic irrational. Then there exists a matrix \( H \) with integer entries having \((x \ 1)^t\) as an eigenvector, with corresponding eigenvalue \( \lambda > \bar{\lambda} > 0 \) (\( \bar{\lambda} \) being the other eigenvalue).

**Proof.** Let \( cX^2 + dX + e \in \mathbb{Z}[X] \) be the minimal polynomial of \( x \), and let \( \bar{x} \) be the algebraic conjugate. Plainly

\[
\begin{pmatrix} -d & -e \\ c & 0 \end{pmatrix} \begin{pmatrix} x \\ 1 \end{pmatrix} = cx \begin{pmatrix} x \\ 1 \end{pmatrix}.
\]

Let \( K \) be the sum of the matrix displayed above with \( t \) times the \( 2 \times 2 \) identity matrix. For \( t \) a sufficiently large positive integer, the eigenvalues \( cx + t \) and \( c\bar{x} + t \) of \( K \) are both positive. If \( c\bar{x} + t < cx + t \), we take \( H = K \) and we are through. Otherwise, we take \( H = |K|K^{-1} \) and observe that \( 0 < |K|(c\bar{x} + t)^{-1} = cx + t < c\bar{x} + t = |K|(cx + t)^{-1} \).

January 2009] NOTES 73
We may now prove the Lagrange theorem for arbitrary piecewise-fractional expansions. Let \( x \in [0, 1] \) be a quadratic irrational, with symbolic sequence \( a_1, a_2, \ldots \) under some Gauss map \( G \), and let \( H, \lambda, \bar{\lambda}, \bar{x} \) be as above. Let \( h : \mathbb{R}^2 \to \mathbb{R}^2 \) be the linear transformation whose matrix is \( H \), with respect to the standard basis. By \( \S 1(1) \), \( x \) belongs to the interval \( \Gamma_n \) whose extrema are \( \psi_{a_1} \cdots \psi_{a_n}(0) \) and \( \psi_{a_1} \cdots \psi_{a_n}(1) \), for every \( n \geq 0 \). The vectors \( (\psi_{a_1} \cdots \psi_{a_n}(0))' \) and \( (\psi_{a_1} \cdots \psi_{a_n}(1))' \) are positively proportional to the columns of the matrix

\[
B_n = G_{a_1}^{-1} G_{a_2}^{-1} \cdots G_{a_n}^{-1} \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.
\]

Let \( C_n \) be the cone spanned positively by these columns. By \( \S 2(\text{ii}) \), there exists \( m \) such that \( \bar{x} \notin \Gamma_n \), for every \( n \geq m \). Since \( 0 < \bar{\lambda} < \lambda \), \( (x 1)' \in C_n \), and \( (\bar{x} 1)' \notin C_n \cup -C_n \), we easily see, by writing an arbitrary vector in \( C_n \) as a linear combination of the eigenvectors \( (x 1)' \) and \( (\bar{x} 1)' \), that \( h[C_n] \setminus \{0\} \) is contained in the topological interior of \( C_n \) for \( n \geq m \). This implies that the matrix \( H_n = B_n^{-1} H B_n \) (i.e., the matrix of \( h \) with respect to the basis given by the columns of \( B_n \)) has positive entries for \( n \geq m \). Hence \( H_m, H_{m+1}, \ldots \) are positive integer matrices, all conjugate to each other via matrices in \( \text{GL}_2 \mathbb{Z} \). Since the determinant and the trace of a matrix are invariant under conjugation, these matrices are finite in number. Therefore \( H_t = H_{t+r} \), for some \( t \geq m \) and \( r > 0 \). The \( \lambda \)-eigenspace of \( H_t \) is 1-dimensional with basis \( B_t^{-1}(x 1)' \), and analogously for \( H_{t+r} \). Multiplying on the left by \( \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix} \), we see that the vectors

\[
G_{a_1} \cdots G_{a_t} \begin{pmatrix} x \\ 1 \end{pmatrix}
\]

and

\[
G_{a_{t+r}} \cdots G_{a_{t+1}} G_{a_t} \cdots G_{a_1} \begin{pmatrix} x \\ 1 \end{pmatrix}
\]

are projectively equal. This means \( G^t x = G^{t+r} x \); hence \( x \) is preperiodic under the Gauss map, and its symbolic sequence is eventually periodic.

REFERENCES

1. V. Baladi and B. Vallée, Euclidean algorithms are Gaussian, *J. Number Theory* 110 (2005) 331–386.
2. M. D. Boshernitzan and C. R. Carroll, An extension of Lagrange’s theorem to interval exchange transformations over quadratic fields, *J. Anal. Math.* 72 (1997) 21–44.
3. H. Cohen, *A Course in Computational Algebraic Number Theory*, Graduate Texts in Mathematics, vol. 138, Springer-Verlag, Berlin, 1993.
4. G. H. Hardy and E. M. Wright, *An Introduction to the Theory of Numbers*, 5th ed., Oxford University Press, Oxford, 1985.
5. J. L. Lagrange, Additions au Mémorie sur la Résolution des Équations Numerériques, *Mémoires de l’Académie royale des Sciences et Belles-Lettres de Berlin* XXIV (1770) 581–652; also available at the Göttinger Digitalisierungszentrum, http://www.gdz-cms.de.
6. A. M. Rockett and P. Szusz, *Continued Fractions*, World Scientific, River Edge, NJ, 1992.
7. B. Vallée, Euclidean dynamics, *Discrete Contin. Dyn. Syst.* 15 (2006) 281–352.
8. H. C. Williams, Solving the Pell equation, in *Number Theory for the Millennium, III*, M. A. Bennett et al., eds., A K Peters, Natick, MA, 2002, 397–435.