Relaxed Hensel lifting of triangular sets
Romain Lebreton

To cite this version:
Relaxed Hensel lifting of triangular sets

Extended abstract

Abstract

In this paper, we present a new lifting algorithm for triangular sets over power series. Our contribution is to give, for any power series triangular set, a shifted algorithm of which the triangular set is a fixed point. Then we can apply the relaxed recursive power series framework and deduce a relaxed lifting algorithm for this triangular set.

We compare our algorithm to the existing techniques. Our algorithm always improves the asymptotic cost in the precision for the special case of univariate representations.

Our goal in this paper is to extend a growing body of work on relaxed algorithms to the context of lifting techniques for univariate representations and triangular sets.

It is well-known that, under some regularity conditions, techniques such as Newton iteration can be used to compute a power series root of an equation such as $f(t, x(t)) = 0$, with $f$ in $k[t, X]$, or a $p$-adic root of an equation of the form $f(x) = 0$ with $f$ in $\mathbb{Z}[X]$.

Relaxed methods, introduced by van der Hoeven [Hoe02, Hoe11], offer an alternative to Newton iteration. The case of computing one power series root, or one $p$-adic root, of a system of polynomial equations was worked out in [BL12]; for this problem, the relaxed algorithm was seen to behave better than Newton iteration in some cases, for instance for multivariate systems with a large number of equations.

In this paper, we go beyond the case of lifting a single root of a multivariate system: we deal with all roots at once, introducing relaxed algorithms that deal with objects such as univariate and triangular representations.

Example 1. We consider the polynomial system $f = (f_1, f_2)$ in $\mathbb{Z}[X_1, X_2]$ with

$$f_1 := 33X_1^3 + 14699X_2^2 + 276148X_1 + 6761112X_2 - 11842820$$
$$f_2 := 66X_1X_2 + X_2^3 - 94X_1 - 75X_2 - 22.$$

Let $m = (7)$, let $k = \mathbb{Z}/7\mathbb{Z}$ and let $T_1$ be the triangular set of $k[X_1, X_2]$, that is a lexicographical Gröbner basis for $X_1 < X_2$, given by

$$T_1 := (X_1^2 + 5X_1, \quad X_2^3 + 3X_1X_2 + 2X_2 + 4X_1 + 6).$$

We lift the triangular set $T_1$ defined modulo 7 to triangular sets $t$ defined modulo $7^2, 7^3, \ldots$. At the first step, we have

$$T_2 = (X_1^2 + (5 + 5 \cdot 7)X_1 + 7, X_2^3 + (3 + 2 \cdot 7)X_1X_2 + (2 + 3 \cdot 7)X_2 + 4X_1 + (6 + 3 \cdot 7))$$

in $(\mathbb{Z}_7/7^2\mathbb{Z}_7)[X_1, X_2]$. We iterate again and find

$$T_3 = (X_1^2 + (5 + 5 \cdot 7 + 6 \cdot 7^2)X_1 + (7 + 7^2),$$
$$X_2^3 + (3 + 2 \cdot 7 + 7^2)X_1X_2 + (2 + 3 \cdot 7 + 5 \cdot 7^2)X_2 + (4 + 5 \cdot 7^2)X_1 +$$
$$(6 + 3 \cdot 7 + 6 \cdot 7^2))$
in \((\mathbb{Z}_7/7^3\mathbb{Z}_7)[X_1, X_2]\). The precision is enough to recover the triangular set
\[
T := (X_1^2 - 9 X_1 + 56, \ X_2^2 + 66 X_1 X_2 - 75 X_2 - 94 X_1 - 22) \in \mathbb{Z}[X_1, X_2].
\]

For simplicity, we deal only with lifting power series (so our equations will involve a parameter \(t\)). Similar techniques apply as well to \(p\)-adic lifting, but the simpler case of power series is useful in itself (see below) and shows all useful ideas. Consider a system of polynomial equations \(f = (f_1, ..., f_n) \in k[t][X_1, ..., X_n]\), where \(k\) is a field. Denote by \(I\) the ideal generated by \((f_1, ..., f_n)\) in \(k(t)[X_1, ..., X_n]\). In what follows, we make the following assumptions, denoted (H):

- the algebraic set \(V = V(I) \subset \overline{k(t)}^n\) has dimension zero;
- the Jacobian determinant of \((f_1, ..., f_n)\) vanishes nowhere on \(V\).

Equivalently, one may consider the curve defined by \(f\) in \(\mathbb{F}^{n+1}\); the zero-dimensional set \(V\) describes its generic points over \(k(t)\).

Let \(d\) be cardinality of \(V\); due to our non-vanishing assumption on the Jacobian of \(f\), the extension \(k(t) \rightarrow A = k(t)[X_1, ..., X_n]/I\) is a product of separable field extensions. As a result, \(A\) has dimension \(d\) over \(k(t)\).

In order to describe \(V\), we will consider two data structures, which we briefly describe now.

An element \(P\) of \(A = k(t)[X_1, ..., X_n]/I\) will be called primitive if the \(K\)-algebra \(k(t)[P] \subset A\) spanned by \(P\) in \(A\) is equal to \(A\) itself. If \(\Lambda\) is a primitive linear form in \(A\), a univariate representation of \(A\) consists of polynomials \(P = (Q, S_1, ..., S_n)\) in \(k(t)[Z]\), where \(Z\) is a new variable, with \(\deg(S_i) < \deg(Q)\) for all \(i\) and such that we have a \(K\)-algebra isomorphism
\[
A = k(t)[X_1, ..., X_n]/I \rightarrow k(t)[Z]/(Q)
\]
\[
\{X_1, ..., X_n\} \mapsto \{S_1, ..., S_n\}
\]
\[
\Lambda \mapsto Z.
\]

In particular, \(Q\) has degree \(d\) in \(Z\); we will say that \(P\) has degree \(d\).

Such representations, or slight modifications thereof (using for instance a rational form for the \(S_i\)'s) have been used for decades in computer algebra, under the names of Shape Lemma, Rational Univariate Representation, Geometric Resolution, etc [GM89, GH91, GHH+97, Rou99, GLS01, HMW01].

On the other side, one finds triangular representations. A triangular set is a set of \(n\) polynomials \(T = (T_1, ..., T_n) \subset k(t)[X_1, ..., X_n]\) such that for all \(i\), \(T_i\) is in \(k(t)[X_1, ..., X_i]\), monic and reduced with respect to \((T_1, ..., T_{i-1})\). This is thus a reduced Gröbner basis for the lexicographic order \(X_1 < \cdots < X_n\). The notion of triangular set comes from [Rit66] in the context of differential algebra. Many similar notions were introduced afterwards [Wu84, Laz91, Kal93, ALM99]; although all these notions do not coincide in general, they are the same for zero-dimensional ideals.

Assume that \(I\) admits a triangular family of generators \(T\), and write \(d_i = \deg(T_i, X_i)\) for all \(i\). Then, we have the equality \(d = d_1 \cdots d_n\); we will say that \(T\) has multidegree \((d_1, ..., d_n)\).
Both kinds of data structures have found numerous applications, and it is not our purpose here to compare their respective merits. Univariate representations always exist, provided the base field is large enough. On the other hand, the ideal \( \mathcal{I} \) may not admit a triangular family of generators: such issues are handled using triangular decompositions of \( \mathcal{I} \) [DMIS+05]. We will not enter this discussion here, and we will simply suppose when needed that \( \mathcal{I} \) admits one or the other such representation.

Let us for instance suppose that \( \mathcal{I} \) admits a univariate representation \( P \). Let further \( m = (t - t_0) \) be a maximal ideal in \( k[t] \), with residual field \( k \), for some \( t_0 \in k \). We make the following assumptions, denoted by \((H')_{P,t_0}\):

- none of the denominators appearing in \( P \) vanishes modulo \( m \); we write \( P_1 \) for the reduction of \( P \) modulo \( m \);
- the polynomials \( f \mod m \) still satisfy \((H)\);
- \( P_1 \) is a univariate representation of \( k[X_1, \ldots, X_n]/(f \mod m) \).

Then, given \( P_1 = P \mod m \) and \( f \), our objective is to compute objects of the form \( P_N = P \mod m^N \), for higher powers of \( N \). Note that, without loss of generality, we could assume \( t_0 = 0 \) (by shifting the origin in \( t \) of the input polynomials).

Similar questions can be asked for triangular representations: suppose that \( \mathcal{I} \) admits a triangular representation \( T \), and that the natural analogue \((H')_T, t_0\) of \((H')_P\) holds; we will show how to compute \( T_N = T \mod m^N \), for high powers of \( N \), starting from \( T_1 = T \mod m \) and \( f \).

Remark that for both the univariate and the triangular case, assumption \((H')_{P,t_0}\), resp. \((H')_T, t_0\), holds for all values of \( t_0 \) except finitely many: this is analyzed in detail in [Sch03a, Sch03b]; thus, these are very mild assumptions.

Let us say a few words about applications of this kind of techniques. One direct application is naturally to solve polynomial systems that depend on a parameter \( t \): the techniques discussed here allow one to compute a description of \( V \) over \( k(t) \) by computing it modulo \( (t - t_0) \) (which we expect to be an easier problem) and lifting it to a sufficient precision (and possibly applying rational reconstruction to recover coefficients in \( k(t) \) from their power series expansion).

Such lifting techniques are also at the core of many further algorithms: for instance, it is used in the geometric resolution algorithm [GLS01, HMW01] that computes univariate representations. This algorithm relies on an iterative lifting / intersection process: the lifting part involves lifting univariate representation, as explained in this paper, while the intersection part uses resultant computations. On the triangular side, a similar lifting / intersection process is used in the change of order algorithm of [DJMS08].

In order to analyze the cost of our lifting algorithms, we will need the following notations for polynomial and power series arithmetic.

- We denote by \( M: \mathbb{N} \to \mathbb{N} \) a function such that one can multiply polynomials in \( k[t] \) of degree at most \( N \) in \( M(N) \) operations in \( k \); we also ask that \( M \) satisfies the super-linearity conditions of [GG03]. Using FFT techniques, one can take \( M(N) = \mathcal{O}(N \log(N) \log\log(N)) \).
• We denote by \( R: \mathbb{N} \to \mathbb{N} \) a function such that online (or equivalently, relaxed) multiplication of power series over \( k \) can be done at precision \( N \) using \( R(N) \) operations in \( k \). Using the algorithm of [FS74, Hoe02], one can take \( R(N) = \mathcal{O}(M(N) \log (N)) \).

Recall that online power series multiplication computes the product of two power series \( \sum_{i \geq 0} c_i t^i := \sum_{i \geq 0} a_i t^i \times \sum_{i \geq 0} b_i t^i \), with the constraint that for all \( i \), the coefficient \( c_i \) of the output is written before the coefficients \( a_{i+1}, a_{i+2}, \ldots \) or \( b_{i+1}, b_{i+2}, \ldots \) of the inputs are read [Hen66, FS74, Hoe02].

We will further need linear algebra operations. We denote by \( \omega \) the constant such that one can multiply matrices of size \( n \) over a ring \( R \) using \( \mathcal{O}(n^\omega) \) operations \( \{+,-,\times\} \) in \( R \). Similarly, we denote \( \Omega \) the constant such that one can compute the determinant and the adjoint matrix using \( \mathcal{O}(n^\Omega) \) operations \( \{+,-,\times\} \) in \( R \). Using the algorithms of [CW90, Sto10, Vas11] for the multiplication and of [Ber84, Kal92, KV04] for the determinant and the adjoint matrix, one can take \( \omega \leq 2.38 \) and \( \Omega \leq 2.70 \). The computation of the determinant and the adjoint matrix is used to compute an unevaluated inverse of a matrix \( A^{-1} \) via the formula \( A^{-1} = 1/\det(A) \text{Adj}(A) \), i.e. the division by the determinant has yet to be performed.

The bulk of the computations will rely on modular arithmetic: we will need to perform arithmetic operations, mostly additions / multiplications, modulo either a univariate polynomial (in the case of univariate representations) or a whole triangular set.

• In the former case, it is known that if \( Q \in R[Z] \) has degree \( d \) (for some ring \( R \)), additions and multiplications modulo \( Q \) can be done using \( \mathcal{O}(M(d)) \) operations in \( R \) [GG03]. When \( R \) is a field, inversion modulo \( Q \), when possible, can be done in time \( \mathcal{O}(M(d) \log (d)) \) [GG03].

• If \( T = (T_1, \ldots, T_n) \) is a triangular set in \( R[X_1, \ldots, X_n] \) with degrees \( d_1, \ldots, d_n \), where \( R \) is a ring, we will let \( MT(d_1, \ldots, d_n) \) denote an upper bound on the number of operations in \( R \) needed to perform multiplication modulo \( T \). If \( R \) is a field, we let \( IT(d_1, \ldots, d_n) \) denote an upper bound for the cost of inversion modulo \( T \) (when possible).

These operations are less straightforward than in the univariate case. Known upper bounds are \( MT(d_1, \ldots, d_n) = \mathcal{O}(4^n d \log (d) \log \log (d)) \) where \( d = d_1 \cdots d_n \) and \( IT(d_1, \ldots, d_n) = \mathcal{O}(c^3 d \log (d)^3) \), for some (large) constant \( c \) [DMSX06, DJMS08, LMS09].

We seize the occasion to state our first result. We propose an algorithm, which modifies slightly the algorithm of [LMS09], that saves a recursive call.

**Proposition 2.** The multiplication modulo a triangular set \( T \) of multidegree \( (d_1, \ldots, d_n) \) can be done in \( MT(d_1, \ldots, d_n) = \mathcal{O}(3^n d \log (d) \log \log (d)) \) operations in \( R \).

Finally, we will make the following assumption on the representation of the input system: we assume that \( f \) is given by a straight-line program with inputs \( t, X_1, \ldots, X_n \) and \( n \) outputs corresponding to \( f_1, \ldots, f_n \), using operations in \( \{+,-,\times\} \). We let \( L \) be the size of this straight-line program.
With this notation being introduced, we can state our first result, which deals with lifting univariate representations.

**Theorem 3.** Let $\mathbf{f} = (f_1, \ldots, f_n)$ be a polynomial system in $k[t][X_1, \ldots, X_n]$ given by a straight-line program $\Gamma$ of size $L$. Suppose that $\mathbf{f}$ satisfies assumption (H) and admits a univariate representation $\mathbf{P}$ with coefficients in $k(t)$ and degree $d$.

Let further $t_0$ be in $k$, let $\mathbf{m} = (t - t_0)$, and suppose that assumption $(H')_{\mathbf{P}, t_0}$ holds. Given $\mathbf{P}_1 = \mathbf{P} \mod \mathbf{m}$, one can compute $\mathbf{P}_N = \mathbf{P} \mod \mathbf{m}^N$ in time

$$O((M(d) \log(d) + n^{\Omega} M(d)) + (LR(N) + (n^2 + nL)N)M(d)).$$

In the running time, the first term corresponds to the inversion of the Jacobian matrix of $\mathbf{f}$ modulo $\mathbf{P}_1$, with coefficients in $k$; this allows us to initialize the lifting. Then, to reach precision $N$, most of the work consists in relaxed multiplications for power series in $k[[t]]$ at precision $N$, coupled with polynomial multiplication in the variable $Z$ in degree $d$. The overall cost is quasi-linear in $N$.

The other known algorithm to perform this kind of task is a suitable version of Newton iteration, introduced in [GLS01, HMW01]. To compute the same output, it runs in time

$$O((M(d) \log(d) + n^{\Omega} M(d)) + (nL + n^{\omega})M(N)M(d)).$$

As the term $n^{\omega} M(N)M(d)$ suggests, this algorithm requires one to do matrix multiplications, with entries that are univariate polynomials of degree $d$, having themselves power series coefficients of precision $N$. Our solution requires no such matrix multiplication.

On the other hand, Newton’s iteration can use plain power series multiplication, which is slightly faster than the online product used in our algorithm. Thus, there exists a trade-off between the dependencies in $n$ and $N$ for both algorithms.

Let us next turn to the case of triangular representations. Our main result for this problem is the following.

**Theorem 4.** Let $\mathbf{f} = (f_1, \ldots, f_n)$ be a polynomial system in $k[t][X_1, \ldots, X_n]$ given by a straight-line program $\Gamma$ of size $L$. Suppose that $\mathbf{f}$ satisfies assumption (H) and admits a triangular representation $\mathbf{T}$ with coefficients in $k(t)$ and multidegree $(d_1, \ldots, d_n)$.

Let further $t_0$ be in $k$, let $\mathbf{m} = (t - t_0)$, and suppose that assumption $(H')_{\mathbf{T}, t_0}$ holds. Given $\mathbf{T}_1 = \mathbf{T} \mod \mathbf{m}$, one can compute $\mathbf{T}_N = \mathbf{T} \mod \mathbf{m}^N$ in time

$$O((|\mathbf{T}|(d_1, \ldots, d_n) + n^{\Omega} \mathbf{M}(d_1, \ldots, d_n)) + (nL \mathbf{R}(N) + n^2N)\mathbf{M}(d_1, \ldots, d_n)).$$

Once more, this is quasi-linear in the precision $N$. For comparison, an adaptation of Newton’s iteration for triangular sets [Sch02] computes the same output in time

$$O((|\mathbf{T}|(d_1, \ldots, d_n) + n^{\Omega} \mathbf{M}(d_1, \ldots, d_n)) + (nL + n^{\omega})\mathbf{M}(N)\mathbf{M}(d_1, \ldots, d_n)).$$

Thus, as in the univariate case, the relaxed approach avoids matrix multiplication with multivariate entries; the price to pay is that we have to replace classical power series multiplication by its relaxed variant.
A univariate representation can be seen as a particular case of a triangular one, through the introduction of an extra unknown $Z$, and the equation expressing $Z$ as a linear form in $X_1, \ldots, X_n$. However, the result in our first theorem is not obtained by specializing the triangular case to the univariate one; further considerations are actually needed.

Let us now give the outline of the proofs of these two theorems. We apply the fundamental theorem concerning relaxed recursive power series that states that the computation of a recursive power series $y = \Phi(y)$ at precision $N$ takes the time necessary to evaluate $\Phi$ at $y$ at the same precision by a shifted algorithm [Wat89, Hoe02, BL12]. The idea behind this theorem is simple: at the $N$th step, we know the coefficients $y_0, \ldots, y_{N-1}$ of $y$ and we have already computed the coefficients $\Phi(y)_i$ for $0 \leq i < N$. Because a shifted algorithm for $\Phi$ reads at most the coefficients $y_0, \ldots, y_{N-1}$ to compute $\Phi(y)_N$, we can compute one more coefficient of $\Phi(y)$. Then we write the new coefficient $y_N = \Phi(y)_N$ of the input. And we can continue the process.

Our contribution consists in finding a recursive equation of which the triangular set we seek to lift is a fixed point. More precisely, the challenge was to find a shifted algorithm that computes an operator $\Phi$ satisfying $y = \Phi(y)$.

The computer algebra software MATHEMAGIX [HLM+02] provides a C++ library named ALGEBRAMIX implementing the relaxed recursive power series framework [Hoe02, Hoe07, BHL11, BL12]. This representation allows the computation of recursive power series. The efficiency of this approach is partially due to the relaxed multiplication algorithm, which gives a fast method for multiplication in this representation. On this basis, we implemented the lifting of univariate representations over the power series ring $\mathbb{F}_p[[t]]$ for both our relaxed approach and Newton iteration.

Our implementation is available in the files lift_series.hpp and lifting_fiber_relaxed.hpp in the C++ library GEOMSOLVEX of MATHEMAGIX, with the help of G. Lecerf.

We now give some implementation details:

- For the multiplication of polynomials of power series in $\mathbb{F}_p[[t]][Z]$, we first converted them to power series of polynomials in $\mathbb{F}_p[Z][[t]]$. Then, the relaxed multiplication algorithm reduces to multiplications of finite precision power series of polynomials, that is polynomials of polynomials in $\mathbb{F}_p[Z][[t]]$. We classically used a Kronecker substitution to reduce these products to multiplications of univariate polynomials in $\mathbb{F}_p[Z]$. Univariate polynomial multiplication uses FFT techniques.

- For the matrix multiplication used in Newton iteration, whose entries are polynomials reduced modulo a univariate $Q$, we delayed the reductions until after the matrix multiplication to reduce their numbers from $n^3$ to $n^2$. 
We report here timings of our implementation (in milliseconds), which are measured using one core of an Intel Xeon X5650 at 2.67 GHz running Linux 64 bits, GMP 5.0.2 [G+91] and setting $p = 16411$ (a 15-bit prime number).

Before considering lifting itself, we start by giving some comparison of timings between the relaxed and classical (zealous) product in $\mathbb{F}_p[[t]][Z]$, depending on the degree in $Z$ and the precision $N$ of power series.

<table>
<thead>
<tr>
<th>$N$</th>
<th>Degree 32 in $Z$</th>
<th>Degree 64 in $Z$</th>
<th>Degree 128 in $Z$</th>
<th>Degree 256 in $Z$</th>
</tr>
</thead>
<tbody>
<tr>
<td>zealous</td>
<td>relaxed</td>
<td>zealous</td>
<td>relaxed</td>
<td>zealous</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>0</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>32</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>64</td>
<td>1</td>
<td>8</td>
<td>3</td>
<td>18</td>
</tr>
<tr>
<td>128</td>
<td>3</td>
<td>21</td>
<td>6</td>
<td>49</td>
</tr>
<tr>
<td>256</td>
<td>6</td>
<td>56</td>
<td>12</td>
<td>130</td>
</tr>
<tr>
<td>512</td>
<td>12</td>
<td>150</td>
<td>30</td>
<td>340</td>
</tr>
<tr>
<td>1024</td>
<td>29</td>
<td>370</td>
<td>71</td>
<td>860</td>
</tr>
<tr>
<td>2048</td>
<td>72</td>
<td>920</td>
<td>170</td>
<td>2100</td>
</tr>
</tbody>
</table>

Table 1. Timings of zealous and relaxed multiplication in $(\mathbb{F}_p[[t]])[Z]$

We observe that the ratio of the timings between the relaxed and zealous algorithms grows as $\log(N)$, as predicted by the estimate $R(N) = O(M(N) \log(N))$ [FS74, Hoe02].

We applied our algorithm on two families of polynomial systems of low complexity of evaluation. The KATSURA polynomials systems comes from a problem of magnetism in physics [Kat90]. We have added some power series coefficients to the original system in order to have a non trivial power series univariate representation. The system KATSURA-$n$ has $n+1$ unknowns $X_0, ..., X_n$ and $n+1$ equations:

$$\text{for } 0 \leq m < n, \quad \sum_{\ell=-n}^{n} X_{[\ell]} X_{[m-\ell]} = \alpha_m X_m$$

and $X_0 + 2 \sum_{\ell=1}^{n} X_{\ell} = \alpha_n$ where the $\alpha_i$ are random coefficients in $\mathbb{F}_p[[t]]$.

The other family of polynomial system MULLINFORM-$n$ has $n$ unknowns and $n$ equations of the form

$$(\lambda_1 X_1 + \cdots + \lambda_n X_n) (\mu_1 X_1 + \cdots + \mu_n X_n) = \alpha$$

where the $\lambda_i, \mu_i$ and $\alpha$ are random coefficients in $\mathbb{F}_p[[t]]$.

For both these examples, we lift a univariate representation on power series at different precision $N$. We indicate with a bold font the theoretical bound for the precision of power series required in the lifting step of the geometric resolution algorithm.
We remark that, on these examples, relaxed algorithms always perform better than zealous algorithms, especially when they are many variables. However, for a given polynomials system, the gap between the two algorithms reduces as the precision gets bigger and we expect zealous algorithms to be faster when the precision tends to infinity.

Bibliography


