Basis construction for range estimation by phase unwrapping

Assad Akhlaq, R. G. McKilliam, and R. Subramanian

Abstract—We consider the problem of estimating the distance, or range, between two locations by measuring the phase of a sinusoidal signal transmitted between the locations. This method is only capable of unambiguously measuring range within an interval of length equal to the wavelength of the signal. To address this problem signals of multiple different wavelengths can be transmitted. The range can then be measured within an interval of length equal to the least common multiple of these wavelengths. Estimation of the range requires solution of a problem from computational number theory called the closest lattice point problem. Algorithms to solve this problem require a basis for this lattice. Constructing a basis is non-trivial and an explicit construction has only been given in the case that the wavelengths can be scaled to pairwise relatively prime integers. In this paper we present an explicit construction of a basis without this assumption on the wavelengths. This is important because the accuracy of the range estimator depends upon the wavelengths. Simulations indicate that significant improvement in accuracy can be achieved by using wavelengths that cannot be scaled to pairwise relatively prime integers.

Index Terms—Range estimation, phase unwrapping, closest lattice point

I. INTRODUCTION

Range (or distance) estimation is an important component of modern technologies such as electronic surveying [1, 2] and global positioning [3, 4]. Common methods of range estimation are based upon received signal strength [5, 6], time of flight (or time of arrival) [7, 8], and phase of arrival [1, 2, 10]. This paper focuses on the phase of arrival method which provides the most accurate range estimates in many applications. Phase of arrival has become the technique of choice in modern high precision surveying and global positioning [3, 4, 11].

A difficulty with phase of arrival is that only the principal component of the phase can be observed. This limits the range that can be unambiguously estimated. One approach to address this problem is to utilise signals of multiple different wavelengths and observe the phase at each. Range estimators from such observations have been studied by numerous authors [4, 5, 10, 12]. Least squares maximum likelihood and maximum a posteriori (MAP) estimators of range have been studied by Teunissen [4], Hassibi and Boyd [12], and more recently Li et. al. [10]. A key realisation is that least squares and MAP estimators can be computed by solving a problem from computational number theory known as the closest lattice point problem [13, 14]. Teunissen [4] appears to have been the first to have realised this connection.

Efficient general purpose algorithms for computing a closest lattice point require a basis for the lattice. Constructing a basis for the least squares estimator of range is non-trivial. Based upon the work of Teunissen [4], and under some assumptions about the distribution of phase errors, Hassibi and Boyd [12] construct of a basis for the MAP estimator. Their construction does not apply for the least squares estimator.

This is problematic because the MAP estimator requires sufficiently accurate prior knowledge of the range, whereas the least squares estimator is accurate without this knowledge. An explicit basis construction for the least squares estimator was recently given by Li et. al. [10] under the assumption that the wavelengths can be scaled to pairwise relatively prime integers. In this paper, we remove the need for this assumption and give an explicit construction in the general case. This is important because the accuracy of the range estimator depends upon the wavelengths. Simulations show that a more accurate range estimator can be obtained using wavelengths that are suitable for our basis, but are not suitable for the basis of Li et. al. [10].

The paper is organised as follows. Section II presents the system model and defines the least squares range estimator. Section III introduces some required properties of lattices. Section IV shows how the least squares range estimator is given by computing a closest point in a lattice. An explicit basis construction for these lattices is described. Simulation results are discussed in Section V and the paper is concluded by suggesting some directions for future research.

II. LEAST SQUARES ESTIMATION OF RANGE

Suppose that a transmitter sends a signal \( x(t) = e^{2\pi (ft+\phi)} \) with phase \( \phi \) and frequency \( f \) in Hertz. The signal is assumed to propagate by line of sight to a receiver resulting in the signal

\[
y(t) = \alpha x(t-r_0/c) + w(t) = \alpha e^{2\pi (ft+\theta)} + w(t)
\]

where \( r_0 \) is the distance (or range) in meters between receiver and transmitter, \( c \) is the speed at which the signal propagates in meters per second, \( \alpha > 0 \) is the real valued amplitude of the received signal, \( w(t) \) represents noise, \( \theta = \phi - r_0/\lambda \) is the phase of the received signal, and \( \lambda = c/f \) is the wavelength.

The least squares estimator is also the maximum likelihood estimator under the assumptions made by Hassibi and Boyd [12]. The matrix \( G \) in [13] is rank deficient in the least squares and weighted least squares cases and so \( G \) is not a valid lattice basis. In particular, observe that the determinant of \( G \) [12, p. 2948] goes to zero as the a priori assumed variance \( \sigma_0^2 \) goes to infinity.
The receiver is assumed to be synchronized by which it is meant that the phase $\phi$ and frequency $f$ are known to the receiver.

Our aim is to estimate $r_0$ from the signal $y(t)$. To do this we first calculate an estimate $\hat{\theta}$ of the principal component of the phase $\theta$. In optical ranging applications $\hat{\theta}$ might be given by an interferometer. In sonar or radio frequency ranging applications $\hat{\theta}$ might be obtained from the complex argument of the demodulated signal $y(t)e^{-2\pi ft}$. Whatever the method of phase estimation, the range $r_0$ is related to the phase estimate $\hat{\theta}$ by the phase difference

$$Y = \langle \phi - \hat{\theta} \rangle = \langle r_0 / \lambda + \Phi \rangle$$

where $\Phi$ represents phase noise and $\langle x \rangle = x - \lfloor x + \frac{1}{2} \rfloor$ where $\lfloor \cdot \rfloor$ denotes the greatest integer less than or equal to $x$. For all integers $k$,

$$Y = \langle r_0 / \lambda + \Phi \rangle = \langle (r_0 + k\lambda) / \lambda + \Phi \rangle,$$

and so, the range is identifiable only if $r_0$ is assumed to lie in an interval of length $\lambda$. A natural choice is the interval $[0, \lambda)$. This poses a problem if the range $r_0$ is larger than the wavelength $\lambda$. To alleviate this, a common approach is to transmit multiple signals $x_n(t) = e^{2\pi ft_n}$ for $n = 1, \ldots, N$, each with a different frequency $f_n$. Now $N$ phase estimates $\hat{\theta}_1, \ldots, \hat{\theta}_N$ are computed along with phase differences

$$Y_n = \langle \phi - \hat{\theta}_n \rangle = \langle r_0 / \lambda_n + \Phi_n \rangle \quad n = 1, \ldots, N$$

where $\lambda_n = c / f_n$ is the wavelength of the $n$th signal and $\Phi_1, \ldots, \Phi_N$ represent phase noise. Given $Y_1, \ldots, Y_N$, a pragmatic estimator of the range $r_0$ is a minimiser of the least squares objective function

$$LS(r) = \sum_{n=1}^{N} (Y_n - r / \lambda_n)^2.$$  \hfill (4)

This least squares estimator is also the maximum likelihood estimator under the assumption that the phase noise variables $\Phi_1, \ldots, \Phi_N$ are independent and identically wrapped normally distributed with zero mean [15, p. 50][16, p. 76][17, p. 47].

The objective function $LS$ is periodic with period equal to the smallest positive real number $P$ such that $P / \lambda_n \in \mathbb{Z}$ for all $n = 1, \ldots, N$, that is, $P = \text{lcm}(\lambda_1, \ldots, \lambda_N)$ is the least common multiple of the wavelengths. The range is identifiable if we assume $r_0$ to lie in an interval of length $P$. A natural choice is the interval $[0, P)$ and we correspondingly define the least squares estimator of the range $r_0$ as

$$\hat{r} = \arg \min_{r \in [0, P)} LS(r).$$  \hfill (5)

If $\lambda_n / \lambda_k$ is irrational for some $n$ and $k$ then the period $P$ does not exist and the objective function $LS$ is not periodic. In this paper we assume this is not the case and that a finite period $P$ does exist.

### III. Lattice Theory

Let $B$ be the $m \times n$ matrix with linearly independent column vectors $b_1, \ldots, b_n$ from $m$-dimensional Euclidean space $\mathbb{R}^m$ with $m \geq n$. The set of vectors

$$\Lambda = \{ Bu : u \in \mathbb{Z}^n \}$$

is called an $n$-dimensional lattice. The matrix $B$ is called a basis or generator for $\Lambda$. The basis of a lattice is not unique. If $U$ is an $n \times n$ matrix with integer elements and determinant $\det U = \pm 1$ then $U$ is called a unimodular matrix and $B$ and $BU$ are both bases for $\Lambda$. The set of integers $\mathbb{Z}^n$ is called the integer lattice with the $n \times n$ identity matrix $I$ as a basis. Given a lattice $\Lambda$ its dual lattice, denoted $\Lambda^*$, contains those points that have integral inner product with all points from $\Lambda$, that is,

$$\Lambda^* = \{ x : x' y \in \mathbb{Z} \text{ for all } y \in \Lambda \}.$$  

The following proposition follows as a special case of Proposition 1.3.4 and Corollary 1.3.5 of [18].

**Proposition 1.** Let $v \in \mathbb{Z}^n$, let $H$ be the $n - 1$ dimensional subspace orthogonal to $v$, and let

$$Q = I - \frac{vv'}{v'v} = I - \frac{v}{\|v\|^2}v'$$

be the $n \times n$ orthogonal projection matrix onto $H$. The set of vectors $\mathbb{Z}^n \cap H$ is an $n - 1$ dimensional lattice with dual lattice $(\mathbb{Z}^n \cap H)^* = \{ Qz : z \in \mathbb{Z}^n \}$.

Given a lattice $\Lambda$ in $\mathbb{R}^n$ and a vector $y \in \mathbb{R}^m$, a problem of interest is to find a lattice point $x \in \Lambda$ such that the squared Euclidean norm $\|y - x\|^2 = \sum_{i=1}^{m} (y_i - x_i)^2$ is minimised. This is called the closest lattice point problem (or closest vector problem) and a solution is called a closest lattice point (or simply closest point) to $y$ [14].

The closest lattice point problem is known to be NP-hard [19, 20]. Nevertheless, algorithms exist that can compute a closest lattice point in reasonable time if the dimension is small (less than about 60) [14, 21, 24]. These algorithms have gone by the name “sphere decoder” in the communications engineering and signal processing literature. Although the problem is NP-hard in general, fast algorithms are known for specific highly regular lattices [23, 26]. For the purpose of range estimation the dimension of the lattice will be $n - 1$ where $N$ is the number of frequencies transmitted. The number of frequencies is usually small (less than 10) and, in this case, general purpose algorithms for computing a closest lattice point are fast [14].

### IV. Range Estimation and the Closest Lattice Point Problem

In this section we show how the least squares range estimator $\hat{r}$ can be efficiently computed by choosing a closest point in a lattice of dimension $N - 1$. The derivation is similar to those in [27, 29]. Our notation will be simplified by the change of variable $r = P\beta$, where $P$ is the least common multiple of the wavelengths. Put $v_n = P / \lambda_n \in \mathbb{Z}$ and define the function

$$F(\beta) = LS(P\beta) = \sum_{n=1}^{N} (Y_n - \beta v_n)^2.$$  

Because $LS$ has period $P$ it follows that $F$ has period 1. If $\beta$ minimises $F$ then $P\beta$ minimises $LS$ and, because $\hat{r} \in [0, P)$, we have $\hat{r} = P(\beta - [\beta])$. It is thus sufficient to find a minimiser $\beta \in \mathbb{R}$ of $F$. 

---

1. [15]: Reference to a specific page number.
2. [16]: Reference to another page number.
3. [17]: Yet another page reference.
4. [18]: Mention of a specific source.
5. [19]: Other relevant work.
6. [20]: Further details.
7. [21]: Additional notes.
8. [22]: Related study.
9. [23]: Specific case.
10. [24]: Particular algorithm.
11. [25]: Mentioned in passing.
12. [26]: Note on related topics.
13. [27]: Similar approach.
14. [28]: Specific technique.
15. [29]: Related context.
Observe that \( (Y_n - \beta v_n)^2 = \min_{z \in \mathbb{Z}} (Y_n - \beta v_n - z)^2 \) and so \( F \) may equivalently be written

\[
F(\beta) = \min_{z_1, \ldots, z_N \in \mathbb{Z}} \sum_{n=1}^{N} (Y_n - \beta v_n - z_n)^2.
\]

The integers \( z_1, \ldots, z_N \) are often called wrapping variables and are related to the number of whole wavelengths that occur over the range \( r_0 \) between transmitter and receiver. The minimiser \( \beta \) can be found by jointly minimising the function

\[
F_1(\beta, z_1, \ldots, z_N) = \sum_{n=1}^{N} (Y_n - \beta v_n - z_n)^2
\]

over the real number \( \beta \) and integers \( z_1, \ldots, z_N \). This minimisation problem can be solved by computing a closest point in a lattice. To see this, define column vectors

\[
y = (Y_1, \ldots, Y_N)' \in \mathbb{R}^N,
\]

\[
z = (z_1, \ldots, z_N)' \in \mathbb{Z}^N,
\]

\[
v = (v_1, \ldots, v_N)' = (P/\lambda_1, \ldots, P/\lambda_N)' \in \mathbb{Z}^N.
\]

Now

\[
F_1(\beta, z_1, \ldots, z_N) = F_1(\beta, z) = ||y - \beta v - z||^2.
\]

The minimiser of \( F_1 \) with respect to \( \beta \) as a function of \( z \) is

\[
\hat{\beta}(z) = \frac{y - z}{v'} v.
\]

Substituting this into \( F_1 \) gives

\[
F_2(z) = \min_{\beta \in \mathbb{R}} F_1(\beta, z) = F_1(\hat{\beta}(z), z) = ||Qy - Qz||^2
\]

where \( Q = I - vv' ||v||^2 \) is the orthogonal projection matrix onto the \( N-1 \) dimensional subspace orthogonal to \( v \). Denote this subspace by \( H \). By Proposition \( \ref{prop:projection} \) the set \( \Lambda = \mathbb{Z}^N \cap H \) is an \( N-1 \) dimensional lattice with dual lattice \( \Lambda^* = \{Qz : z \in \mathbb{Z}^N \} \). We see that the problem of minimising \( F_2(z) \) is precisely that of finding a closest point in the lattice \( \Lambda^* \) to \( Qy \in \mathbb{R}^N \). Suppose we find \( \hat{x} \in \Lambda^* \) closest to \( Qy \) and a corresponding \( \hat{z} \in \mathbb{Z}^N \) such that \( \hat{x} = Q\hat{z} \). Then \( \hat{z} \) minimises \( F_2 \) and \( \hat{\beta}(\hat{z}) \) minimises \( F_1 \). The least squares range estimator in the interval \( [0, P) \) is then

\[
\hat{r} = P(\hat{\beta}(\hat{z}) - \hat{\beta}(\hat{z}))_+. \tag{6}
\]

It remains to provide a method to compute a closest point \( \hat{x} \in \Lambda^* \) and a corresponding \( \hat{z} \in \mathbb{Z}^N \). In order to use known general purpose algorithms we must first provide a basis for the lattice \( \Lambda^* \). A basis for \( \Lambda^* \) can be obtained by computing the projection of the lattice \( \Lambda^* \) onto \( H \). That is, \( Q\mathbf{u}_1, \ldots, Q\mathbf{u}_N \) is a basis for \( \Lambda^* \) where \( \mathbf{u}_1, \ldots, \mathbf{u}_N \) are the columns of \( Q \).

Proposition 2. Let \( U \) be an \( N \times N \) unimodular matrix with first column given by \( v \). A basis for the lattice \( \Lambda^* \) is given by the projection of the last \( N-1 \) columns of \( U \) orthogonally onto \( H \). That is, \( Q\mathbf{u}_2, \ldots, Q\mathbf{u}_N \) is a basis for \( \Lambda^* \) where \( u_1, \ldots, u_N \) are the columns of \( U \).

Proof: Because \( U \) is unimodular it is a basis matrix for the integer lattice \( \mathbb{Z}^N \). So, every lattice point \( z \in \mathbb{Z}^N \) can be uniquely written as \( z = c_1u_1 + \cdots + c_Nu_N \) where \( c_1, \ldots, c_N \in \mathbb{Z} \). The lattice

\[
\Lambda^* = \{Qz \in \mathbb{Z}^N \}
\]

\[
= \{Q(c_1u_1 + \cdots + c_Nu_N) \in \mathbb{Z}^N \}
\]

\[
= \{c_2Qu_2 + \cdots + c_Nu_N \in \mathbb{Z}^N \}
\]

because \( Qu_1 = Qv = 0 \) is the origin. It follows that \( Q\mathbf{u}_2, \ldots, Q\mathbf{u}_N \) form a basis for \( \Lambda^* \).

To find a basis for \( \Lambda^* \) we require a matrix \( U \) as described by the previous proposition. Such a matrix is given by Li et al. \cite{li2009estimation} Eq. (76) under the assumption that the wavelengths can be scaled to pairwise relatively prime integers. We do not require this assumption here. Because \( P = \text{lcm}(\lambda_1, \ldots, \lambda_N) \) it follows that the integers \( v_1, \ldots, v_N \) are jointly relatively prime, that is, \( \text{gcd}(v_1, \ldots, v_N) = 1 \). Define integers \( g_1, \ldots, g_N \) by \( g_N = v_N \) and observe that \( g_{k+1}/g_k \) and \( v_k/g_k \) are relatively prime integers. For \( k = 1, \ldots, N - 1 \), define the \( N \) by \( N \) matrix \( A_k \) with \( m, n \)th element

\[
A_{kmn} = \begin{cases} v_k/g_k & m = n = k \\ g_{k+1}/g_k & m = k + 1, n = k \\ a_k & m = k, n = k + 1 \\ b_k & m = n = k + 1 \\ I_{mn} & \text{otherwise} \end{cases}
\]

where \( I_{mn} = 1 \) if \( m = n \) and 0 otherwise. The integers \( a_k \) and \( b_k \) are chosen to satisfy

\[
b_k v_k = a_k g_k g_{k+1} = 1 \tag{7}
\]

and can be computed by the extended Euclidean algorithm. The matrix \( A_k \) is equal to the identity matrix everywhere except at the 2 by 2 block of indices \( k \leq m \leq k + 1 \) and \( k \leq n \leq k + 1 \). The matrix \( A_k \) is unimodular for each \( k \) because it has integer elements and because the determinant of the 2 by 2 matrix

\[
\begin{bmatrix} v_k/g_k & a_k \\ g_{k+1}/g_k & b_k \end{bmatrix} = b_k v_k - a_k g_k g_{k+1} = 1
\]

as a result of (7). A matrix \( U \) satisfying the requirements of Proposition 2 is now given by the product

\[
U = \prod_{k=1}^{N-1} A_k = A_{N-1} \times A_{N-2} \times \cdots \times A_1.
\]
That $U$ is unimodular follows immediately from the unimodularity of $A_1, \ldots, A_{N-1}$. It remains to show that the first column of $U$ is equal to $v$. Let $v_1, \ldots, v_{N-1}$ be column vectors of length $N$ defined as
\[
v_k = (v_1, \ldots, v_k, g_{k+1}, 0, \ldots, 0)', \quad k = 1, \ldots, N - 2
\]
\[
v_{N-1} = (v_1, \ldots, v_{N-1}, g_N)' = v.
\]
One can readily check that $v_{k+1} = A_{k+1}v_k$ for all $k = 1, \ldots, N - 1$. The first column of the matrix $A_1$ is $v_1$ and so, by induction, the first column of the product $\prod_{k=1}^{K} A_k = v_K$ for all $K = 1, \ldots, N - 1$. It follows that the first column of $U$ is $v_{N-1} = v$ as required.

Let $U_2$ be the $N$ by $N - 1$ matrix formed by removing the first column from $U$, that is, $U_2 = (u_2, \ldots, u_N)$. By Proposition \[2\] a basis for $\Lambda^*$ is given by projecting the columns of $U_2$ orthogonally to $v$, that is, a basis matrix for $\Lambda^*$ is the $N$ by $N - 1$ matrix $B = QU_2$. Given $B$ a general purpose algorithm \[14\] can be used to compute $\hat{w} \in \mathbb{Z}^{N-1}$ such that $\hat{x} = B\hat{w}$ is a closest lattice point in $\Lambda^*$ to $Qy \in \mathbb{R}^N$. Now
\[
\hat{x} = B\hat{w} = QU_2\hat{w} = Q\hat{z}
\]
and so $\hat{z} = U_2\hat{w} \in \mathbb{Z}^N$. The least squares range estimator $\hat{r}$ is then given by \[6\].

V. Simulation Results

We present the results of Monte-Carlo simulations with the least squares range estimator. Simulations with $N = 4$ and $N = 5$ wavelengths are performed. For each case we consider two different sets of wavelengths. The first set is suitable for the basis of Li et. al. \[10\] and was used in the simulations in \[10\]. The second set is suitable only for our basis. In each simulation the true range $r_0 = 20$ and the phase noise variables $\Phi_1, \ldots, \Phi_N$ are wrapped normally distributed, that is, $\Phi_n = (X_n)$ where $X_1, \ldots, X_N$ are independent and normally distributed with zero mean and variance $\sigma^2$. In this case, the least squares estimator is also the maximum likelihood estimator. Figure \[1\] shows the sample mean square error for $\sigma^2$ in the range $10^{-5}$ to $10^{-2}$ and $10^7$ Monte-Carlo trials used for each value of $\sigma^2$.

For $N = 4$ the two sets of wavelengths are
\[
A = \{2, 3, 5, 7\}, \quad B = \{\frac{210}{7}, \frac{210}{6}, \frac{210}{41}, \frac{210}{34}\}.
\]
For both sets the wavelengths are contained in the interval $[2, 7]$ and $P = 210 = \text{lcm}(A) = \text{lcm}(B)$ so that the identifiable range is the same. The wavelengths $A$ are relatively prime integers and are suitable for the basis of Li et. al. \[10\] and are used in the simulations in \[10\]. The wavelengths $B$ are not suitable for the basis of \[10\] because they can not be scaled to pairwise relatively prime integers. To see this, observe that the smallest positive number by which we can multiply the elements of $B$ to obtain integers is $c = \frac{7124949}{210}$. Multiplying the elements of $B$ by $c$ we obtain the set
\[
c \times B = \{77531, 100409, 149389, 197579\}
\]
and these elements are not pairwise relatively prime because, for example, $\gcd(77531, 100409) = 1271$. Figure \[1\] shows the results of simulations with both sets $A$ and $B$. When the noise variance $\sigma^2$ is small wavelengths $A$ result in slightly reduced sample mean square error as compared with $B$. As $\sigma^2$ increases the sample mean square error exhibits a ‘threshold’ effect and increases suddenly. The threshold occurs at $\sigma^2 \approx 1.2 \times 10^{-4}$ for wavelengths $A$ and $\sigma^2 \approx 3 \times 10^{-4}$ for wavelength $B$. Wavelengths $B$ are more accurate than $A$ when $\sigma^2$ is greater than approximately $1.2 \times 10^{-4}$.

For $N = 5$ the two sets of wavelengths are
\[
C = \{2, 3, 5, 7, 11\}, \quad D = \{\frac{210}{877}, \frac{210}{523}, \frac{210}{277}, \frac{210}{221}, \frac{210}{211}\}.
\]
For both sets all wavelengths are contained in the interval $[2, 11]$ and $P = 2310 = \text{lcm}(C) = \text{lcm}(D)$ so that the maximum identifiable range is the same. The basis of Li et. al. \[10\] can be used for wavelengths $C$ but not for $D$. The wavelengths $C$ were used in the simulations in \[10\]. Figure \[1\] shows the result of Monte-Carlo simulations with these wavelengths. Wavelengths $C$ result in slightly smaller sampler mean square error than $D$ when $\sigma^2$ is small, but dramatically more error for $\sigma^2$ above the threshold occurring at $\sigma^2 \approx 7 \times 10^{-5}$.

The sets $B$ and $D$ have been selected based on a heuristic optimisation criterion. The properties of this criterion are not yet fully understood and will be the subject of a future paper.

VI. Conclusion

We have considered least squares maximum likelihood estimation of range from observation of phase at multiple wavelengths. The estimator can be computed by finding a closest point in a lattice. This requires a basis for the lattice.

Bases have previously been constructed under the assumption that the wavelengths can be scaled to relatively prime integers. In this paper, we gave a construction in the general case and indicated by simulation that this can dramatically improve range estimates. An open problem is how to select wavelengths to maximise the accuracy of the least squares estimator. We will study this problem in future research.
