Towards a unified theory of GNSS ambiguity resolution

P.J.G. Teunissen

Department of Mathematical Geodesy and Positioning, Delft University of Technology, Kluyverweg 1, 2629 HS Delft

Received: 18 November 2003 / Accepted: 20 November 2003

Abstract. In this invited contribution a brief review will be presented of the integer estimation theory as developed by the author over the last decade and which started with the introduction of the LAMBDA method in 1993. The review discusses three different, but closely related classes of ambiguity estimators. They are the integer estimators, the integer aperture estimators and the integer equivariant estimators. Integer estimators are integer aperture estimators and integer aperture estimators are integer equivariant estimators. The reverse is not necessarily true however. Thus of the three types of estimators the integer estimators are the most restrictive. Their pull-in regions are translational invariant, disjunct and they cover the ambiguity space completely. Well-known examples are integer rounding, integer bootstrapping and integer least-squares.

A less restrictive class of estimators is the class of integer aperture estimators. Their pull-in regions only obey two of the three conditions. They are still translational invariant and disjunct, but they do not need to cover the ambiguity space completely. As a consequence the integer aperture estimators are of a hybrid nature having either integer or non-integer outcomes. Examples of integer aperture estimators are the ratio-testimator and the difference-testimator. The class of integer equivariant estimators is the less restrictive of the three classes. These estimators only obey one of the three conditions, namely the condition of being translational invariant. As a consequence the outcomes of integer equivariant estimators are always real-valued.

For each of the three classes of estimators we also present the optimal estimator. Although the Gaussian case is usually assumed, the results are presented for an arbitrary probability density function of the float solution. The optimal integer estimator in the Gaussian case is the integer least-squares estimator. The optimality criterion used is that of maximizing the probability of correct integer estimation, the so-called success rate. The optimal integer aperture estimator in the Gaussian case is the one which only returns the integer least-squares solution when the integer least-squares residual resides in the optimal aperture pull-in region. This region is governed by the probability density function of the float solution and by the probability density function of the integer least-squares residual. The aperture of the pull-in region is governed by a user-defined aperture parameter. The optimality criterion used is that of maximizing the probability of correct integer estimation given a fixed, user-defined, probability of incorrect integer estimation. The optimal integer aperture estimator becomes identical to the optimal integer estimator in case the success rate and the fail rate sum up to one.

The best integer equivariant estimator is an infinite weighted sum of all integers. The weights are determined as ratios of the probability density function of the float solution with its train of integer shifted copies. The optimality criterion used is that of minimizing the mean squared error. The best integer equivariant estimator therefore always outperforms the float solution in terms of precision.

Keywords: GNSS ambiguity resolution, integer estimation, integer aperture estimation, integer equivariant estimation

Contents
1 Introduction
2 Integer Estimation
2.1 Pull-in regions
2.2 Three integer estimators
2.3 Optimal integer estimation
3 Integer Equivariant Estimation
3.1 A larger class of estimators
3.2 Best integer equivariant estimation
4 Integer Aperture Estimation
4.1 Aperture pull-in regions
4.2 Three examples of integer aperture estimators
4.3 Optimal integer aperture estimation
5 Summary
6 References
1 Introduction

Global Navigation Satellite System (GNSS) carrier phase ambiguity resolution is the process of resolving the carrier phase ambiguities as integers. It is the key to fast and high precision GNSS positioning and it therefore applies to a great variety of GNSS models which are currently in use in navigation, surveying, geodesy and geophysics. It will be clear that a rigorous theoretical framework is needed in order to understand, execute and validate carrier phase ambiguity resolution properly. This theory was lacking in the early days of ambiguity resolution. In the last decade however the contours of a rich and rigorous theory has emerged from the contributions of many. The current invited contribution will focus on a part of this theoretical framework. It presents a brief review of the integer estimation theory as developed by the author and which started with the introduction of the LAMBDA method in 1993. The theory presented is non-Bayesian throughout. Although the theory is applicable to any linear(ized) model with integer parameters, the motivation for its development stems first and foremost from the desire to have a rigorous theory available for solving the very significant ‘carrier phase ambiguity problem’ of GNSS. Although the first applications of the theory have been for single baseline positioning with GPS, the applications have grown to cover network positioning, attitude determination, formation flying, other systems than GNSS, such as Glonass and Galileo, augmented and integrated GNSS, and other systems than GNSS, such as in the field of interferometric synthetic aperture radar (InSAR).

As our point of departure we take the following system of linear observation equations

\[ E\{y\} = Aa + Bb, \quad a \in \mathbb{Z}^n, \quad b \in \mathbb{R}^p \]  

with \( E\{\cdot\} \) the mathematical expectation operator, \( y \) the \( m \)-vector of observables, \( a \) the \( n \)-vector of unknown integer parameters and \( b \) the \( p \)-vector of unknown real-valued parameters. All the linear(ized) GNSS models can in principle be cast in the above frame of observation equations. The data vector \( y \) will then usually consist of the ‘observed minus computed’ single-, dual- or multi-frequency double-difference (DD) phase and/or pseudorange (code) observations accumulated over all observation epochs. The entries of vector \( a \) are then the DD carrier phase ambiguities, expressed in units of cycles rather than range, while the entries of the vector \( b \) will consist of the remaining unknown parameters, such as for instance baseline components (coordinates) and possibly atmospheric delay parameters (troposphere, ionosphere).

The procedure for solving the above GNSS model can be divided conceptually into three steps. In the first step one simply discards the integer constraints \( a \in \mathbb{Z}^n \) and performs a standard adjustment. As a result one obtains the so-called ‘float’ solution \( \tilde{a} \) and \( \tilde{b} \). This solution is real-valued. Then in the second step the ‘float’ solution \( \tilde{a} \) is further adjusted so as to take in some pre-defined way the integerness of the ambiguities into account. This gives

\[ \hat{a}_S = S(\tilde{a}) \]  

in which \( S \) is an \( n \)-dimensional mapping that in some way takes the integerness of the ambiguities into account. This estimator is then used in the final step to adjust the ‘float’ estimator \( \tilde{b} \). As a result one obtains the so-called ‘fixed’ estimator of \( b \) as

\[ \hat{b}_S = \tilde{b} - Q_{ba} Q_a^{-1}(\tilde{a} - \hat{a}_S) \]  

in which \( Q_a \) denotes the \( \nu \)-matrix of \( \tilde{a} \) and \( Q_{ba} \) denotes the covariance matrix of \( \tilde{b} \) and \( \tilde{a} \).

The above three-step procedure is still ambiguous in the sense that it leaves room for choosing the \( n \)-dimensional map \( S \). Different choices for \( S \) will lead to different ambiguity estimators and thus also to different baseline estimators \( \hat{b}_S \). One can therefore now think of constructing family of maps \( S \) with certain desirable properties. In this contribution we will review three such classes of ambiguity estimators. They are the integer estimators, the integer equivariant estimators and the integer aperture estimators. These classes were introduced by the author in respectively (Teunissen, 1999), (Teunissen, 2002) and (Teunissen, 2003). The three classes of estimators are subsets of one another. The first class is the most restrictive class. This is due to the fact that the outcomes of any estimator within this class are required to be integer. The most relaxed class is the class of integer equivariant estimators. These estimators are real valued and they only obey the integer remove-restore principle. The class of integer aperture estimators is a subset of the integer equivariant estimators but it encompasses the class of integer estimators. The integer aperture estimators are of a hybrid nature in the sense that their outcomes are either integer or non-integer. We also present the optimal estimator for each of the three classes of estimators. The optimality criterion for the first class is the maximization of the probability of correct integer estimation. For the second class it is the minimization of the mean squared error and for the third class it is the maximization of the probability of correct integer estimation given a user-defined fixed probability level of incorrect integer estimation. The theory presented is non-Bayesian throughout. There is therefore no need to work with priors. All the results presented in this review are given without proof. The proofs can be found in the references cited.

In this review we focus only on the estimation of the integer ambiguities and therefore refrain from discussing the probabilistic consequences for the baseline estimator. For the probabilistly distribution of the GNSS baseline we refer to (Teunissen, 1999b).
2 Integer Estimation

2.1 The pull-in regions

We will start with the requirement that the estimator \( \hat{a}_S \) needs to be integer. In that case \( S : R^n \rightarrow Z^n \). Integer estimators will be denoted as \( \hat{a} \). Thus \( \hat{a} = \hat{a}_S \) in case \( S : R^n \rightarrow Z^n \). The map \( S \) will not be one-to-one due to the discrete nature of \( Z^n \). Instead it will be a many-to-one map. This implies that different real-valued vectors will be mapped to one and the same integer vector. One can therefore assign a subset \( S_z \subset R^n \) to each integer vector \( z \in Z^n \):

\[
S_z = \{ x \in R^n | z = S(x) \}, \quad z \in Z^n
\]

(4)

The subset \( S_z \) contains all real-valued vectors that will be mapped by \( S \) to the same integer vector \( z \in Z^n \). This subset is referred to as the pull-in region of \( z \). It is the region in which all vectors are pulled to the same integer vector \( z \).

Since the pull-in regions define the integer estimator completely, one can define classes of integer estimators by imposing various conditions on the pull-in regions. The following class of integer estimators was introduced by Teunissen (1999).

**Definition 1 (Integer estimators)**
The mapping \( \hat{a} = S(\hat{a}) \) is said to be an integer estimator if its pull-in regions satisfy

1. \[ \bigcup_{z \in Z^n} S_z = R^n \]
2. \[ \text{Int}(S_{z_1}) \cap \text{Int}(S_{z_2}) = \emptyset, \quad \forall z_1, z_2 \in Z^n, z_1 \neq z_2 \]
3. \[ S_z = z + S_0, \quad \forall z \in Z^n \]

(iii)

This definition is motivated as follows. Each one of the above three conditions describes a property of which it seems reasonable that it is possessed by an arbitrary integer estimator. The first condition states that the pull-in regions should not leave any gaps and the second that they should not overlap. The absence of gaps is needed in order to be able to map any float solution \( \hat{a} \in R^n \) to \( Z^n \), while the absence of overlaps is needed to guarantee that the float solution is mapped to just one integer vector. Note that we allow the pull-in regions to have common boundaries. This is permitted if we assume to have zero probability that \( \hat{a} \) lies on one of the boundaries. This will be the case when the probability density function (pdf) of \( \hat{a} \) is continuous.

The third and last condition of the definition follows from the requirement that \( S(x + z) = S(x) + z, \forall x \in R^n, z \in Z^n \). Also this condition is a reasonable one to ask for. It states that when the float solution \( \hat{a} \) is perturbed by \( z \in Z^n \), the corresponding integer solution is perturbed by the same amount. This property allows one to apply the integer remove-restore technique: \( S(\hat{a} - z) + z = S(\hat{a}) \).

It therefore allows one to work with the fractional parts of the entries of \( \hat{a} \), instead of with its complete entries.

Using the pull-in regions, one can give an explicit expression for the corresponding integer estimator \( \hat{a} \). It reads

\[
\hat{a} = \sum_{z \in Z^n} z s_z(\hat{a}) \quad \text{with} \quad s_z(\hat{a}) = \begin{cases} 
1 & \text{if} \quad \hat{a} \in S_z \\
0 & \text{if} \quad \hat{a} \not\in S_z
\end{cases}
\]

(5)

Note that the \( s_z(\hat{a}) \) can be interpreted as weights, since \( \sum_{z \in Z^n} s_z(\hat{a}) = 1 \). The integer estimator \( \hat{a} \) is therefore equal to a weighted sum of integer vectors with binary weights.

2.2 Three integer estimators

**Integer rounding:** The three best known integer estimators are integer rounding, integer bootstrapping and integer least-squares. The simplest way to obtain an integer vector from the real-valued float solution is to round each of the entries of \( \hat{a} \) to its nearest integer. The corresponding integer estimator reads therefore

\[
\hat{a}_R = ([\hat{a}_1], \ldots, [\hat{a}_n])^T
\]

(6)

where \([.] \) denotes rounding to the nearest integer. The pull-in region of this integer estimator equals the multivariate version of the unit-square. It is given as

\[
S_{R,z} = \{ x \in R^n | | c_i^T (x - z) | \leq \frac{1}{2}, i = 1, \ldots, n \}, \quad \forall z \in Z^n
\]

(7)

where \( c_i \) denotes the \( i \)th canonical unit vector having a 1 as its \( i \)th entry and zeros otherwise.

**Integer bootstrapping:** Another relatively simple integer ambiguity estimator is the bootstrapped estimator. The bootstrapped estimator can be seen as a generalization of the previous estimator. It still makes use of integer rounding, but it also takes some of the correlation between the ambiguities into account. The bootstrapped estimator follows from a sequential conditional least-squares adjustment and it is computed as follows. If \( n \) ambiguities are available, one starts with the first ambiguity \( \hat{a}_1 \), and rounds its value to the nearest integer. Having obtained the integer value of this first ambiguity, the real-valued estimates of all remaining ambiguities are then corrected by virtue of their correlation with the first ambiguity. Then the second, but now corrected, real-valued ambiguity estimate is rounded to its nearest integer. Having obtained the integer value of the second ambiguity, the real-valued estimates of all remaining \( n - 2 \) ambiguities are then again corrected, but now by virtue of their correlation with the second ambiguity. This process is continued until all ambiguities are considered. The entries of the bootstrapped
estimator \( \hat{a}_B = (\hat{a}_{B,1}, \ldots, \hat{a}_{B,n})^T \in \mathbb{Z}^n \) are thus given as

\[
\begin{align*}
\hat{a}_{B,1} &= [\hat{a}_1] \\
\hat{a}_{B,2} &= [\hat{a}_{2,1}] = [\hat{a}_2 - \sigma_{21} \sigma_1^{-2} (\hat{a}_1 - \hat{a}_{B,1})] \\
& \vdots \\
\hat{a}_{B,n} &= [\hat{a}_{n,N}] = [\hat{a}_n - \sum_{j=1}^{n-1} \sigma_{n,j} \sigma_1^{-2} (\hat{a}_{j,1} - \hat{a}_{B,j})]
\end{align*}
\] (8)

where \( \sigma_{i,j} \) denotes the covariance between \( \hat{a}_i \) and \( \hat{a}_{j,1} \), and \( \sigma_1^{-2} \) is the variance of \( \hat{a}_{j,1} \). The shorthand notation \( \hat{a}_{j,1} \) stands for the \( j \)th least-squares ambiguity obtained through a conditioning on the previous \( I = \{1, \ldots, (i - 1)\} \) sequentially rounded ambiguities.

Because of the close relationship that exists between sequential conditional least-squares estimation and the triangular factorization of the vc-matrix, \( Q_a = LDL^T \), the unit lower triangular matrix \( L \) can be used to describe the bootstrapped pull-in regions. They are given as

\[
S_{B,z} = \{ x \in \mathbb{R}^n \mid | c_i^T L^{-1}(x - z) | \leq \frac{1}{2}, \forall z \in \mathbb{Z}^n \}
\] (9)

The pull-in regions of integer rounding are unit-cubes, while those of integer bootstrapping are multivariate versions of parallelograms. The bootstrapped pull-in regions reduce to multivariate unit-cubes in case the vc-matrix is a diagonal matrix, \( L = I_n \). Bootstrapping reduces namely to rounding in the absence of any correlation between the ambiguities.

Note that the bootstrapped estimator is not unique. The outcome of bootstrapping and its performance depend on the chosen ambiguity parameterization. Thus although the principle of bootstrapping remains the same, every choice of ambiguity parameterization has its own bootstrapped estimator. Bootstrapping of DD-ambiguities, for instance, will generally perform poorly due to the high correlation and poor precision of DD ambiguities when short observation time spans are used. One should therefore make use of an appropriate parameterization when using bootstrapping. This can be done by applying the decorrelating \( Z \)-transformation of the LAMBDA (Least-squares AMBiguity Decorrelation Adjustment) method. When this transformation is applied, one works with the more precise and decorrelated ambiguity vector \( \hat{z} = Z\hat{a} \) instead of with the original ambiguity vector \( \hat{a} \). For more information on the LAMBDA method, we refer to e.g. (Teunissen, 1993), (Teunissen, 1995) and (de Jonge and Tiberius, 1996) or to the textbooks (Hofmann-Wellenhof et al., 2002), (Strang and Borre, 1997), (Teunissen and Kleusberg, 1998), (Misra and Enge, 2001) and (Seeber, 2003).

**Integer least-squares:** The integer least-squares estimator minimizes the weighted squared norm of the ambiguity residual over all integers. It is defined as

\[
\hat{a}_{LS} = \arg \min_{z \in \mathbb{Z}^n} \| \hat{a} - z \|_Q^2
\] (10)

with \( \| . \|_Q^2 = (.)^T Q_a^{-1} (.) \). In contrast to integer rounding and integer bootstrapping, an integer search is needed to compute \( \hat{a}_{LS} \). The integer least-squares (ILS) procedure is mechanized in the LAMBDA method, which is currently one of the most applied methods for GNSS carrier phase ambiguity resolution. Practical results obtained with it can be found, for example, in (Boon and Ambrosius, 1997), (Boon et al., 1997), (Cox and Brading, 1999), (de Jonge and Tiberius, 1996a), (de Jonge et al., 1996), (Han, 1995), (Jonkman, 1998), (Peng et al., 1999), (Tiberius and de Jonge, 1995), (Tiberius et al., 1997).

To determine the ILS pull-in regions we need to know the set of float solutions \( \hat{a} \in \mathbb{R}^n \) that are mapped to the same integer vector \( z \in \mathbb{Z}^n \). This set is described by all \( x \in \mathbb{R}^n \) that satisfy \( z = \arg \min_{u \in \mathbb{Z}^n} \| x - u \|_Q^2 \). The ILS pull-in-region that belongs to the integer vector \( z \) follows therefore as

\[
S_{LS,z} = \{ x \in \mathbb{R}^n \mid \| x - z \|_Q^2 \leq \| x - u \|_Q^2, \forall u \in \mathbb{Z}^n \}
\] (11)

It consists of all those points which are closer to \( z \) than to any other integer point in \( \mathbb{R}^n \). The metric used for measuring these distances is determined by the vc-matrix \( Q_a \).

It is possible to give a representation of the ILS pull-in regions that resembles the representation of the bootstrapped pull-in regions. This representation is given as

\[
S_{LS,z} = \bigcap_{c_i \in \mathbb{Z}^n} \{ x \in \mathbb{R}^n \mid | c_i^T Q_a^{-1}(x - z) | \leq \frac{1}{2} | c_i \|_Q^2 \}, \forall z \in \mathbb{Z}^n
\] (12)

This shows that the ILS pull-in regions are constructed from intersecting half-spaces. One can show that at most \( 2^n - 1 \) pairs of such half spaces are needed for constructing the pull-in region. The ILS pull-in regions are convex, symmetric sets of volume 1, which satisfy the conditions of Definition 1. They are hexagons in the two-dimensional case. Two-dimensional examples of the pull-in regions of integer rounding, integer bootstrapping and integer least-squares are given in Figure 1.

### 2.3 Optimal integer estimation

For the evaluation of the integer ambiguity estimator one needs the distribution of \( \hat{a} \). This distribution is of the discrete type and it will be denoted as \( P(\hat{a} = z) \). It is a probability mass function (pmf), having zero masses at nongrid points and nonzero masses at some or all grid points. In order to obtain this pmf we need the probability density function of the ‘float’ solution \( \hat{a} \). This pdf will be denoted
as \( f_a(x \mid a) \), in which we explicitly show the dependence on the unknown but integer vector \( a \). In the Gaussian case we therefore have

\[
f_a(x \mid a) = \frac{1}{\sqrt{\det(Q_a)(2\pi)^n}} \exp\left(-\frac{1}{2}||x - a||_Q^2\right)
\]

The pmf \( P(\hat{a} = z) \) follows from integrating \( f_a(x \mid a) \) over the pull-in regions \( S_z \):

\[
P(\hat{a} = z) = \int_{S_z} f_a(x \mid a) dx, \quad z \in \mathbb{Z}^n
\]

This distribution is of course dependent on the pull-in regions \( S_z \) and thus on the chosen integer estimator. Since various integer estimators exist which are admissible, some may be better than others. Having the problem of GNSS ambiguity resolution in mind, one is particularly interested in the estimator which maximizes the probability of correct integer estimation. This probability equals \( P(\hat{a} = a) \), but it will differ for different ambiguity estimators. The answer to the question which estimator maximizes the probability of correct integer estimation was given by Teunissen (1999a).

**Theorem 1 (Optimal integer estimation)**

Let \( f_a(x \mid a) \) be the pdf of the float solution \( \hat{a} \) and let

\[
\hat{a}_{ML} = \arg \max_{a \in \mathbb{Z}^n} f_a(\hat{a} \mid a)
\]

be an integer estimator. Then

\[
P(\hat{a}_{ML} = a) \geq P(\hat{a} = a)
\]

for any arbitrary integer estimator \( \hat{a} \).

The above theorem holds true for an arbitrary pdf of the float ambiguities \( \hat{a} \). In most GNSS applications however, one assumes the data to be normally distributed. The estimator \( \hat{a} \) will then be normally distributed too, with mean \( a \in \mathbb{Z}^n \) and vc-matrix \( Q_a \), \( \hat{a} \sim \mathcal{N}(a, Q_a) \). In this case the optimal estimator becomes identical to the integer least squares estimator

\[
\hat{a}_{LS} = \arg \min_{a \in \mathbb{Z}^n} ||x - a||_{Q_a}^2
\]

The above theorem therefore gives a probabilistic justification for using the integer least-squares estimator when the pdf is Gaussian. For GNSS ambiguity resolution one is thus better off using the integer least-squares estimator than any other admissible integer estimator.

### 3 Integer Equivariant Estimation

#### 3.1 A larger class of estimators

The result of the above theorem holds true for the defined class of integer (I) estimators. One may now wonder what happens if the conditions of Definition 1 are relaxed. Would it then still be possible to find an ambiguity estimator which in some sense outperforms the float solution? In order to answer this question the class of integer equivariant (IE) estimators was introduced in (Teunissen, 2002). This class is larger than the class of integer estimators and it is defined as follows.

**Definition 2 (Integer equivariant estimators)**

The estimator \( \hat{\theta}_{IE} = F_\theta(\hat{a}) \), with \( F_\theta : \mathbb{R}^n \to \mathbb{R} \), is said to be an integer equivariant estimator of the linear function \( \theta = a^T z \) if

\[
F_\theta(x + z) = F_\theta(x) + a^T z, \quad \forall x \in \mathbb{R}^n, z \in \mathbb{Z}^n
\]

This definition was motivated by the fact that of the conditions of Definition 1 one should at least retain the property that the integer remove-restore principle applies. It will be clear that integer (I) estimators are also IE-estimators. Simply check that the above condition is indeed fulfilled by the estimator \( \hat{\theta} = a^T \hat{a} \). The converse, however, is not necessarily true. The class of IE-estimators is therefore indeed a larger class than the class of I-estimators.

The class of IE-estimators is also a larger class than the class of linear unbiased estimators, assuming that the float solution is unbiased. Let \( F_{\theta, a} \), for some \( F_\theta \in \mathbb{R}^n \), be the linear estimator of \( \theta = a^T z \). For it to be unbiased one
needs, using $E\{\hat{a}\} = a$, that $F_0^T E\{\hat{a}\} = t^T a$, $\forall a \in R^n$ holds true, or that $F_0 = l$. But this is equivalent to stating that $F_0^T (\hat{a} + a) = F_0^T \hat{a} + t^T a$, $\forall a \in R^n, a \in R^n$. Comparison with (17) shows that the condition of linear unbiasedness is more restrictive than the condition of integer equivariance. The class of linear unbiased estimators is therefore a subset of the class of integer equivariant estimators. This result implies that IE-estimators exist which are unbiased. Thus, if one denotes the class of IE-estimators as $IE$, the class of unbiased estimators as $U$, the class of unbiased IE-estimators as $IEU$, the class of unbiased integer estimators as $IU$, and the class of linear unbiased estimators as $LU$, one may summarize their relationships as: $IEU = IE \cap U \neq \emptyset, LU \subset IEU$ and $IU \subset IEU$ (see Figure 2).

3.2 Best integer equivariant estimation

Having defined the class of IE-estimators one may now look for an IE-estimator which is ‘best’ in a certain sense. Assuming that such an estimator exists, it immediately follows that it must be better or at least as good as the float solution. After all the float solution is an IE-estimator as well. We will denote the best integer equivariant (BIE) estimator as $\hat{\theta}_{BIE}$ and use the mean squared error (MSE) as our criterion of ‘best’. The best integer equivariant estimator of $\theta = t^T a$ is therefore defined as:

$$\hat{\theta}_{BIE} = \text{arg min}_{F_0 \in IE} E\{(F_0(\hat{a}) - \theta)^2\}$$

(18)

in which $IE$ stands for the class of IE-estimators. The minimization is thus taken over all integer equivariant functions that satisfy the condition of Definition 2.

The reason for choosing the MSE-criterion is twofold. First, it is a well-known probabilistic criterion for measuring the closeness of an estimator to its target value, in our case $\theta = t^T a$. Second, the MSE-criterion is also often used as measure for the quality of the float solution itself. It should be kept in mind however that the MSE-criterion is a weaker criterion that the probabilistic criterion used in the previous section for determining the optimal integer estimator. The following theorem, due to Teunissen (2002a), gives the solution to the above minimization problem.

**Theorem 2 (Best integer equivariant estimation)**

Let $f_0(x \mid a)$ be the pdf of the float solution and let $\hat{\theta}_{BIE}$ be the best integer equivariant estimator of $\theta = t^T a$. Then $\hat{\theta}_{BIE} = t^T \hat{\theta}_{BIE}$, with

$$\hat{\theta}_{BIE} = \sum_{z \in Z^n} z w_z(\hat{a})$$

and

$$w_z(x) = \frac{f_0(x + a - z \mid a)}{\sum_{u \in Z^n} f_0(x + a - u \mid a)}$$

(19)

Note the resemblance in structure between the BIE-estimator and an arbitrary integer estimator, see (5). The BIE-estimator is also a weighted sum of all integer vectors in $Z^n$. In the present case, however, the weights are not binary. They vary between zero and one, and their values are determined by the float solution and its pdf. As a consequence the BIE-estimator will be real-valued in general, instead of integer-valued.

An important consequence of the above theorem is that the BIE-estimator is always better than or at least as good as any integer estimator as well as any linear unbiased estimator. After all the class of integer estimators and the class of linear unbiased estimators are both subsets of the class of IE-estimators. The nonlinear BIE-estimator is therefore also better than the best linear unbiased (BLU) estimator. The BLU-estimator is the minimum variance estimator of the class of linear unbiased estimators and it is given by the well-known Gauss-Markov theorem. One therefore has

$$\text{MSE}(\hat{\theta}_{BIE}) \leq \text{MSE}(\hat{\theta}_{BLU})$$

(20)

The two estimators $\hat{\theta}_{BIE}$ and $\hat{\theta}_{BLU}$ both minimize the mean squared error within their class, see (Teunissen, 2002a).

The above results hold true for any pdf the ‘float’ solution $\hat{a}$ might have. In many applications however it is assumed that the pdf of $\hat{a}$ is Gaussian. In that case we have the following corollary.

**Corollary (BIE in the Gaussian case)**

Let $\hat{a}$ be distributed as $\hat{a} \sim N(a, Q_a)$ and let $\hat{\theta}_{BIE}$ be the best integer equivariant estimator of $\theta = t^T a$. Then

$$\hat{\theta}_{BIE} = t^T \hat{\theta}_{BIE}$$

(21)

with

$$\hat{\theta}_{BIE} = \sum_{z \in Z^n} z \frac{\exp\{-\frac{1}{2} \| \hat{a} - z \|^2_{Q_a}\}}{\sum_{z \in Z^n} \exp\{-\frac{1}{2} \| \hat{a} - z \|^2_{Q_a}\}}$$

For the proof see (Teunissen, 2003a). Since the space of integers $Z^n$ can be seen as a certain discretized version of the space of real numbers $R^n$, one would expect if the integer grid size gets smaller in relation to the size and extend
of the pdf, that the difference between the two estimators ̂a_{BIE} and ̂a gets smaller as well. Similarly, if the pdf gets more peaked in relation to the integer grid size, one would expect that the BIE-estimator ̂a_{BIE} tends to an integer estimator. This is made precise in the following lemma.

Lemma (limits of the integer grid)

(i) If we replace \( \sum_{z \in \mathbb{Z}^n} \) by \( \int_{R^n} dz \) in (21), then
\[
\lim_{\sigma \to 0} ̂a_{BIE} = ̂a_{ILS}
\]

(ii) Let the vc-matrix of ̂a be factored as \( Q_α = σ^2G \). Then
\[
\lim_{\sigma \to 0} ̂a_{BIE} = ̂a_{ILS}
\]

It is interesting to observe that the above expression given for ̂a_{BIE} is identical to its Bayesian counterpart as given in (Betti et al., 1993) (Gundlich and Koch, 2001), and (Gundlich and Teunissen, 2002). This is not quite true for the general case however. Still, the above equivalence nicely bridges the gap which existed so far between the current theory of integer inference and the Bayesian approach. Despite the similarity in the above case however, there are important differences in the probabilistic evaluation of the solutions. Like the BLU-estimator, the BIE-estimator is a random variable with the property of being unbiased and of minimum variance. In the Bayesian framework the solution is considered to be nonrandom due to the conditioning that takes place. Furthermore, in the Bayesian framework the unknown parameters are assumed to be random variables for which probability distributions need to be specified a priori. The theory presented in the present contribution is non-Bayesian throughout with no need at all to make assumptions about prior distributions. As a final note we remark that the theory of integer equivariant estimation can be extended to linear functions of both the ambiguities and the baseline components, see (Teunissen, 2003a). A probabilistic performance comparison between the BIE-estimator and the ‘float’ and ‘fixed’ ambiguity estimators can be found in (Verhagen and Teunissen, 2003).

4 Integer Aperture Estimation

4.1 Aperture pull-in regions

The two classes of ambiguity estimators discussed so far are related as \( I \subset IE \). That is, integer estimators are integer equivariant, but integer equivariant estimators are not necessarily integer. We will now discuss a third class of ambiguity estimators. This class was introduced in (Teunissen, 2003) and is referred to as the class of integer aperture (IA) estimators. This class will be larger than the I-class, but smaller than the IE-class, \( I \subset IA \subset IE \). Whereas the IE-class was defined by dropping two of the three conditions of Definition 1, the IA-class will be defined by dropping only one of the three conditions, namely the condition that the pull-in regions should cover \( R^n \) completely. We will therefore allow the pull-in regions of the IA-estimators to have gaps.

In order to introduce the new class of ambiguity estimators from first principles, let \( Ω \subset R^n \) be the region of \( R^n \) for which ̂a is mapped to an integer if ̂a ∈ Ω. It seems reasonable to ask of the region \( Ω \) that it has the property that if ̂a ∈ Ω then also ̂a + z ∈ Ω, for all \( z \in \mathbb{Z}^n \). If this property would not hold, then float solutions could be mapped to integers whereas their fractional parts would not. We thus require \( Ω \) to be translational invariant with respect to an arbitrary integer vector: \( Ω + z = Ω \), for all \( z \in \mathbb{Z}^n \).

Knowing \( Ω \) is however not sufficient for defining our estimator. \( Ω \) only determines whether or not the float solution is mapped to an integer, but it does not tell us yet to which integer the float solution is mapped. We therefore define
\[
Ω_z = Ω \cap S_z, \forall z \in \mathbb{Z}^n
\]

where \( S_z \) is a pull-in region satisfying the conditions of Definition 1. Then
\[
(i) ∪ z Ω_z = Ω \cap (∪ z S_z) = Ω \cap R^n = Ω
\]
\[
(ii) Ω_z \cap Ω_{z_2} = (Ω \cap Ω_{z_1}) \cap (Ω \cap Ω_{z_2}) = Ω \cap (S_{z_1} \cap S_{z_2}) = ∅, \forall z_1, z_2 \in \mathbb{Z}^n, z_1 ≠ z_2
\]
\[
(iii) Ω_0 + z = (Ω \cap S_0) + z = (Ω + z) \cap (S_0 + z) = Ω \cap S_z = Ω_z, \forall z \in \mathbb{Z}^n
\]

This shows that the subsets \( Ω_z \subset S_z \) satisfy the same conditions as those of Definition 1, be it that \( R^n \) has now been replaced by \( Ω \subset R^n \). Hence, the mapping of the IA-estimator can now be defined as follows. The IA-estimator maps the float solution ̂a to the integer vector \( z \) when \( ̂a ∈ Ω_z \) and it maps the float solution to itself when \( ̂a ∉ Ω \).

The class of IA-estimators can therefore be defined as follows.

Definition 3 (Integer aperture estimators)

Integer aperture estimators are defined as
\[
̂a_{IA} = ̂a + \sum_{z \in \mathbb{Z}^n} (z - ̂a_0)ω_z(̂a)
\]

with \( ω_z(x) \) the indicator function of \( Ω_z = Ω \cap S_z \) and \( Ω \subset R^n \) translational invariant.

Note that an IA-estimator is indeed also an IE-estimator, just like an I-estimator is. There is also resemblance between an IA-estimator and an I-estimator. Since the indicator functions \( ω_z(x) \) of the pull-in regions \( S_z \) sum up to
unity, $\sum_{z \in Z^n} s_z(x) = 1$, the I-estimator (5) may be written as

$$\hat{a} = \hat{a} + \sum_{z \in Z^n} (z - \hat{a})s_z(x) \quad (24)$$

Comparing this expression with that of (23) shows that the difference between the two estimators lies in their binary weights, $s_z(x)$ versus $\omega_z(x)$. Since the $s_z(x)$ sum up to unity for all $x \in R^n$, the outcome of an I-estimator will always be integer. This is not true for an IA-estimator, since the binary weights $\omega_z(x)$ do not sum up to unity for all $x \in R^n$. The IA-estimator is therefore an hybrid estimator having as outcome either the real-valued float solution $\hat{a}$ or an integer solution. The IA-estimator returns the float solution if $\hat{a} \not\in \Omega$ and it will be equal to $z$ when $\hat{a} \in \Omega$. Note, since $\Omega$ is the collection of all $\Omega_z = \Omega_0 + z$, that the IA-estimator is completely determined once $\Omega_0$ is known. Thus $\Omega_0 \subset S_0$ plays the same role for the IA-estimators as $S_0$ does for the I-estimators. By changing the size and shape of $\Omega_0$ one changes the outcome of the IA-estimator. The subset $\Omega_0$ can therefore be seen as an adjustable pull-in region with two limiting cases. The limiting case in which $\Omega_0$ is empty and the limiting case when $\Omega_0$ equals $S_0$. In the first case the IA-estimator becomes identical to the float solution $\hat{a}$, and in the second case the IA-estimator becomes identical to an I-estimator. The subset $\Omega_0$ therefore determines the aperture of the pull-in region.

### 4.2 Three examples of IA-estimators

Various examples can be given of IA-estimators. In fact one can devise one’s own IA-estimator by specifying the aperture pull-in region $\Omega_0$. Here we will give three examples of IA-estimators.

**The ratio-testimator:** In the practice of GPS carrier phase ambiguity resolution various tests are in use for discriminating between the 'best' and the so-called 'second-best' solution. These tests are usually referred to as discriminability tests. In (Verhagen, 2003) the probabilistic characteristics and performance of the test statistics were evaluated. One such a discernibility test is the popular ratio-test. The ratio-test is defined as follows. Let $\hat{a}$ be the float solution, $\hat{a} = \arg \min_{z \in Z^n} || z - \hat{a} ||^2_{Q_0}$ the integer least-squares solution and $\hat{a}' = \arg \min_{z \in Z^n \cap \Omega_0} || z - \hat{a} ||^2_{Q_0}$ the so-called 'second-best' solution. Then $\hat{a}$ is accepted as the fixed solution if

$$\frac{|| \hat{a} - \hat{a}' ||^2_{Q_0}}{|| \hat{a} - \hat{a} ||^2_{Q_0}} \leq \rho \quad (25)$$

This test has been used in e.g. (Euler and Schaffrin, 1990), (Wei and Schwarz, 1995) and (Han and Rizos, 1996). Thus with the ratio-test $\hat{a}$ is accepted as the fixed solution if the float solution $\hat{a}$ is sufficiently more closer to $\hat{a}$ than to the 'second-best' solution $\hat{a}'$. The non-negative scalar $\rho$ is a user-defined tolerance level.

In (Teunissen, 2003b) it was shown that the procedure underlying the above test is actually that of an IA-estimator. The rejection region of the above test is integer translational invariant and thus an example of $R^n \setminus \Omega$. For this region the outcome will be $\hat{a}$. The outcome will be the integer $z \in Z^n$ however, when the test is passed and $\hat{a}$ lies in the least-squares pull-in region of $z$.

The aperture of the pull-in region of the ratio-test is governed by the choice of the single parameter $\rho$. One has a zero aperture in case $\rho = 0$ and a maximum aperture in case $\rho = 1$. In the first case the procedure of the ratio-test will always output the float solution, while in the second case it will always output the integer least-squares solution $\hat{a}$. Changing the value of the aperture parameter $\rho$ will thus change the performance of the ratio-test.

**The difference-testimator:** Although perhaps less popular, tests other than the ratio-test have been proposed in the GPS literature as well. One such test is the difference-test. This test was introduced in (Tiberius and de Jonge, 1995). This test also makes use of the integer least-squares solution and the 'second best' solution. It is defined as follows. The integer least-squares solution $\hat{a}$ is accepted as the fixed solution with the difference-test if

$$|| \hat{a} - \hat{a}' ||^2_{Q_0} - || \hat{a} - \hat{a} ||^2_{Q_0} \geq \delta \quad (26)$$

where the non-negative scalar $\delta$ is a user-defined tolerance level. As with the ratio-test, the difference-test accepts $\hat{a}$ as the fixed solution if the float solution is sufficiently more closer to $\hat{a}$ than to the 'second best' solution $\hat{a}'$. 'Closeness' is however measured differently. The procedure underlying the difference-test can also be shown to be that of an IA-estimator, see (Teunissen, 2003b).

**The ellipsoidal IA-estimator:** The procedures currently in place for GPS ambiguity resolution all make use of comparing, in some pre-defined sense, the 'best' solution with the so-called 'second-best' solution. But when one thinks of the concept of the aperture region, there is in principle no need to compute or to make use of the 'second-best' solution. That is, one can do without the 'second-best' solution, as long as one is able to measure and evaluate the closeness of the float solution to an integer. The ellipsoidal IA-estimator is one such IA-estimator. The aperture pull-in regions of the ellipsoidal integer aperture (EIA) estimator are defined as

$$E_z = E_0 + z, \quad E_0 = S_0 \cap C_\varepsilon, \quad \forall z \in Z^n \quad (27)$$

with $S_0$ being the least-squares pull-in region and $C_\varepsilon = \{ x \in R^n | || x ||^2_{Q_0} \leq \varepsilon^2 \}$, an origin-centred ellipsoidal region of which the size is controlled by the aperture parameter $\varepsilon$.

Thus the EIA-estimator equals $\hat{a}_{EIA} = z$ if $\hat{a} \in E_z$ and $\hat{a}_{EIA} = \hat{a}$ otherwise. From the definition follows that
$E_z = \{ x \in S_z \mid \| x - z \|_Q^2 \leq \epsilon^2 \}$. This shows that the procedure for computing the EIA-estimator is rather straightforward. Using the float solution $\hat{a}$, its vc-matrix $Q_a$, and the aperture parameter $e$ as input, one only needs to compute the integer least-squares solution $\hat{a}$ and verify whether or not the inequality

$$\| \hat{a} - \hat{a} \|_Q^2 \leq \epsilon^2$$

(28) is satisfied. If the inequality is satisfied then $\hat{a}_{EIA} = \hat{a}$, otherwise $\hat{a}_{EIA} = \hat{a}$. A comparison with the ratio-test (25) and with the difference-test (26) shows that (28) is indeed the simplest of the three inequalities. Instead of (25) and with the difference-test (26) shows that (28) is straightforward. Using the float solution $\hat{a}$, the aperture pull-in region which uniquely defines the I-estimator, is always integer and therefore $P_U = 0$, the fail-rate of an I-estimator equals one minus its success-rate. Thus although the optimal I-estimator has the largest possible success-rate, one can not exercise any control over its fail-rate. That is, the fail-rate of an I-estimator is determined completely by the strength, or the lack of strength for that matter, of the underlying mathematical model. It can not be fixed a priori independently of the model. This situation changes however in case of IA-estimation. Due to the fact that IA-estimators allow one to exercise control over the aperture of the pull-in region, it also gives one the possibility to exercise control over the fail-rate. The idea is therefore to constrain the fail-rate to a user-defined fixed value and then to find the size and shape of the pull-in region which maximizes the success-rate. In (Teunissen, 2003b) it was shown that the aperture pull-in region of this optimal IA-estimator is given as follows.

**Theorem 3 (optimal integer aperture estimation)**

Let $f_a(x \mid a)$ be the pdf of the float solution $\hat{a}$, and let $P_S$ and $P_F$ be respectively the success-rate and the fail-rate of the IA-estimator. Then the solution to

$$\max_{\Omega_0 \subseteq S_0} P_S \text{ subject to given } P_F$$

(30) is given by the aperture pull-in region

$$\Omega_0 = \{ x \in S_0 \mid \sum_{z \in \mathbb{Z}^n} f_a(x + z \mid a) \leq \lambda f_a(x + a \mid a) \}$$

(31)

with

$$S_0 = \{ x \in \mathbb{R}^n \mid 0 = \arg \max_{z \in \mathbb{Z}^n} f_a(x \mid z) \}$$

and with the aperture parameter $\lambda$ chosen so as to satisfy the a priori fixed fail-rate $P_F$.

With this result we are now able to make a connection with the theory of integer estimation of Section 2. First note that the above $S_0$ equals the pull-in region of the optimal integer estimator of Theorem 1. In the Gaussian case it reduces to the ILS pull-in region. The size of the aperture pull-in region $\Omega_0 \subseteq S_0$ is governed by the aperture parameter $\lambda$, which on its turn is governed by the user-defined fixed fail-rate $P_F$. $\Omega_0$ is empty in case $\lambda = 1$ and its size will get larger when $\lambda$ gets larger. For a large enough value of $\lambda$, $\Omega_0$ will become identical to $S_0$, in which case the optimal integer aperture estimator becomes identical to the optimal integer estimator.
For most applications one would like the fail-rate $P_F$ to be smaller than the success-rate $P_S$. The result of the above theorem can be used to determine the value of the aperture parameter that guarantees this to be the case. Since $P_F = \sum_{z \in \mathbb{Z}^n \setminus \{0\}} \int_{\Omega_0} f_\lambda(x + z \mid a)dx$ and $P_S = \int_{\Omega_0} f_\lambda(x + 2 \mid a)dx$ it follows that $P_F \leq (\lambda - 1) P_S$. This shows that the fail-rate will be smaller than the success-rate when $\lambda < 2$.

The above result applies to an arbitrary pdf of $\hat{a}$. In most cases however the pdf of the 'float' solution is assumed to be Gaussian. In case the 'float' solution is normally distributed as $\hat{a} \sim N(a, Q_a)$, the optimal aperture pull-in region becomes

$$
\Omega_0 = \{ x \in S_0 \mid \sum_{z \in \mathbb{Z}^n \setminus \{0\}} \exp\left\{-\frac{1}{2} \| x - z \|^2_{Q_a}\right\} \\
\leq (\lambda - 1) \exp\left\{-\frac{1}{2} \| x \|^2_{Q_a}\right\} \tag{32}
$$

with $S_0$ being the ILS pull-in region. The computational steps involved in computing the optimal integer aperture estimator are now as follows. First compute the integer least-squares solution $\hat{a}_{LS} = \arg\min_{z \in \mathbb{Z}^n} \| \hat{a} - a \|^2_{Q_a}$. Then form the ambiguity residual $\epsilon = \hat{a} - \hat{a}_{LS}$ and check whether $\epsilon \in \Omega_0$. If this is the case then the outcome of the optimal estimator is $\hat{a}_{LS}$, otherwise the outcome is $\hat{a}$. For the purpose of computational efficiency it is advised to compute $\hat{a}_{LS}$ with the LAMBDA method and use the LAMBDA-transformed ambiguities also for the evaluation of $\epsilon \in \Omega_0$.

Note that the contribution of the exponentials in the sum of (32) gets smaller the more peaked the pdf of the 'float' solution is. The aperture pull-in region $\Omega_0$ will therefore get larger the more peaked the pdf is. This is also what one would expect. With (32) we are now also in a position to make an interesting link with one of the IA-estimators presented in the previous section, namely the difference-testimator. If we approximate $\Omega_0$ by retaining only the largest term in the sum of the inequality of (32) we obtain the inequality of the difference test. This shows that the difference-testimator is a close to optimal IA-estimator in case the pdf is peaked. Including an extension of the theory of integer aperture estimation which starts form predefined integer estimators, can be found in (Teunissen, 2003b+c). There is furthermore an interesting link with the concept of penalized ambiguity resolution as introduced in (Teunissen, 2003d).

5 Summary

In this contribution a brief review was presented of the estimation theory which has been developed over the last decade for estimating integer parameters such as the carrier phase ambiguities of GNSS. Three different classes of estimators were discussed, the integer (I) estimators, the integer aperture (IA) estimators and integer equivariant (IE) estimators. These three classes of estimators are related as $I \subset IA \subset IE$, see Figure 3.

Of the three type of estimators the integer estimators are the most restrictive. Their pull-in regions are translational invariant, disjunct and cover $\mathbb{R}^n$ completely. Every integer estimator can be represented as

$$
\hat{a} = \hat{a} + \sum_{z \in \mathbb{Z}^n} (z - \hat{a}) s_z(\hat{a})
$$

with $s_z(x)$ being the indicator function of the pull-in region $S_z$. Well-known examples of integer estimators are integer rounding, integer bootstrapping and integer least-squares. In the Gaussian case, the integer least-squares estimator is the optimal estimator. The optimality criterion used is that of maximizing the probability of correct integer estimation, the so-called success rate.

A less restrictive class of estimators is the class of integer aperture estimators. The pull-in regions of these estimators only need to obey two of the three conditions. They are still translational invariant and disjunct, but they do not need to cover $\mathbb{R}^n$ completely. Hence gaps are allowed. As a consequence the outcomes of integer aperture estimators are either integer or non-integer, this in contrast to integer estimators which always have integer outcomes. Every integer aperture estimator can be represented as

$$
\hat{a}_{IA} = \hat{a} + \sum_{z \in \mathbb{Z}^n} (z - \hat{a}) \omega_z(\hat{a})
$$

with $\omega_z(x)$ being the indicator function of the aperture pull-in region $\Omega_z$. The outcome of an integer aperture estimator equals therefore $z$ in case $\hat{a} \in \Omega_z$ and $\hat{a}$ otherwise. Well-known examples of integer aperture estimators are the ratio-testimator and the difference-testimator. The fact that the aperture pull-in regions do not need to cover $\mathbb{R}^n$ completely allows one to excercize control over the probability of incorrect integer estimation, the so-called fail rate. In the Gaussian case, the optimal integer aperture estimator is the one which returns the integer least-squares solution $\hat{a}_{LS}$ in case $f_{\lambda}(\hat{a}) \leq \lambda f_{\lambda}(\hat{a} + a \mid a)$ in which $f_{\lambda}(x)$ denotes...
the pdf of the residual $\xi = \hat{a} - \hat{a}_{LS}$. The optimality criterion used is that of maximizing the probability of correct integer estimation given a fixed, user-defined, probability of incorrect integer estimation. The optimal integer aperture estimator becomes identical to the optimal integer estimator in case the success rate and the fail rate sum up to one.

The class of integer equivariant estimators is the less restrictive of the three classes. These estimators only obey one of the three conditions, namely the condition of being translational invariant. That is, they only obey the integer remove-restore principle. As a consequence the outcomes of integer equivariant estimators are always real-valued. In the Gaussian case, the best integer equivariant estimator therefore outperforms the 'float' estimator.

A comparison between the representations of the three types of estimators shows that they only differ in the weighting functions used. These weighting functions are respectively $s_2(x)$, $\omega_2(x)$ and $w_2(x)$. The functions $s_2(x)$ are binary and sum up to unity for any $x$, the functions $\omega_2(x)$ are also binary but do not sum up to unity for any $x$, and the functions $w_2(x)$ do sum up to unity for any $x$, but are not binary.

Acknowledgement

The formatting and the figures of this contribution were made by Mrs. ir. S. Verhagen.

6 References


