Inversion of Shortened LDI Matrix for Designing Discrete-time Filters

Dalibor Biolek*
Department of Electrical Engineering, Faculty of Military Technology, UD Brno, Czech Republic

Jaromir Kuben, Jr.
Department of Mathematics and Statistics, Faculty of Science, Masaryk University Brno, Czech Republic

Jaromir Kuben
Department of Mathematics and Physics, Faculty of Military Technology, UD Brno, Czech Republic

* Corresponding author. E-mail: dalibor.biolek@unob.cz
Received: 27 January 2014; Accepted: 28 February 2014; Published online: 6 March 2014
DOI: 10.14416/j.ijast.2014.01.009

Abstract
The paper deals with the problem of the inversion of the so-called shortened LDI matrix for conversion between the coefficients of continuous-time and discrete-time linear filters whose s- and z-domain transfer functions are coupled via LDI transform. The elements of the inverse matrix are expressed by analytic formulae for general order of the matrix. The corresponding mathematical proofs are given.

Keywords: Lossless Discrete Integration, LDI matrix, Matrix Inversion

1 Introduction

In engineering practice, the relationship between the models of continuous-time system and system originating from it by discretizing the time axis is analyzed in two different aspects.

The first one is a numerical solution of differential equations in simulation programs where the solution is searched via approximating difference equation. The corresponding methods, such as Backward Euler, Trapezoidal, Gear, and others, are then universally usable for solving linear and non-linear problems [1,2]. The concrete integration algorithm introduces an error to the solution of the differential equation due to approximation of the derivative by difference relations. The nature and the gravity of this error generally depend on the type of the approximation, on the system model, on the type of its excitation, and on the selection of the integration step and the method of its control. It turns out that some simulation models exhibit large sensitivity of the solution to the way of the time axis discretization, and that then the numerical solution can be far from the reality [3,4]. Such anomalous manifestations of the numerical methods increase the level of uncertainty of the designer about the correctness of the outputs generated by commercial simulation programs. Therefore they are currently analyzed at large, among other with the use of the theory of chaos [5].

The second aspect, falling into the operator domain, deals with the relationship between the transfer functions of linear continuous-time (CT) and discrete-time (DT) systems [6,7]. For linear systems, one-to-one relations exist between both areas which result from the well-known connections between the system behavior in time, operator, and spectral domains.

The paper is concerned with the problem from the second aspect. The bilinear (BL) s-z transform belongs to frequently utilized methods of designing linear DT system from its CT prototype [8,9]. Probably the first attempts at algorithmic solution of the task of the conversion of the coefficients of s-domain transfer function of CT system into coefficients of z-domain transfer function of DT system for BL transform and several other transformations are described in [10] and [11]. It is shown that the vectors of the coefficients are
matched via the so-called Pascal matrix. Its elements as well as the elements of its inverse matrix can be easily generated for an arbitrary order of the system, thus for an arbitrary matrix size [12]. The Pascal matrix is used for digital filter design [13-16], in numerical mathematics [17-22], and for signal processing [23-25]. Some papers indicate that the Pascal matrix also attracts the attention of pure mathematicians [26].

However, the BL transform is only one of more utilized first-order $s$-$z$ transformations for converting $s$-domain transfer function to $z$-domain transfer function of the same order. Among other transforms, let us mention BD (Backward-Difference), FD (Forward-Difference) [8], and parametric transforms, the latter representing combinations of the basic transformations [27,28]. For these reasons, the so-called generalized Pascal matrix is introduced in [29], which includes the classical Pascal matrix as a special case, concurrently enabling conversions of the coefficients of CT and DT systems for all existing first-order $s$-$z$ transforms. The rules for compiling the generalized Pascal matrix and also the inverse matrix are described in [30]. The rigorous mathematical analysis of the generalized Pascal matrix for various $s$-$z$ transforms is given in [31] and [32].

Another useful $s$-$z$ transform, the LDI (Lossless Discrete Integration) [33], plays an important role since the eighties of the last century, either directly or in combination with the BL transform, for designing low-sensitive digital filters and switched-capacitor filters based on analog ladder structures [34,35]. At later stage, it was utilized for designing switched-current filters [36,37] and switched-capacitor sigma-delta AD converters [38]. For example, the LDI transform is used for designing digital filters with low complexity coefficients in [39], Lossless Discrete Differentiator (LDD) in [40], and digital allpass filters in [41].

In contrast with the BL, BD, FD transforms and the parametric transforms derived from them, the LDI is not the first-order $s$-$z$ transformation, such that the generalized Pascal matrix cannot be used for this case. The corresponding computations are provided by the so-called LDI matrix, introduced in [42]. However, unlike Pascal matrix, the LDI matrix is not square because the DT system has more coefficients than the CT system. Owing to a certain symmetry of the vector of the coefficients of the DT system, all the information about the elements of the LDI matrix is contained in its specific part, forming a square submatrix, denoted in [43] as shortened LDI matrix. General properties of the original and shortened LDI matrices are analyzed in [43], and effective procedure for their composition for general order is found therein. However, the analytical formulae of the elements of the matrix, coming from the inversion of shortened LDI matrix for backward DT-CT transformation, have not been known till now.

This paper presents the inverse matrix of the shortened LDI matrix in the analytical form. That completes the missing part of the matrix algorithm for bi-directional transformation between $s$ and $z$ domains, which was formerly built for the first-order $s$-$z$ transform via generalized Pascal matrix.

The paper has the following structure: Section 2, following this Introduction, summarizes the definition of LDI transform and the resultant definitions of complete and shortened LDI matrices. The inverse matrix of the shortened LDI matrix in the analytic form is presented in Section 3. Section 4 presents the respective mathematical proofs. Section 5 describes software experiments, demonstrating the correctness and usefulness of the proposed method of the matrix inversion.

## 2 LDI Transform and Its Matrix Description

The LDI transform is defined by the well-known equation [33]

$$ s = \frac{1}{T}(z^{1/2} - z^{-1/2}) $$

(1)

where $s$ and $z$ are operators of the Laplace and $z$ transforms, and $T$ is the sampling period, utilized for time axis discretization during CT to DT transformation.

Applying the substitution

$$ x = z^{1/2}, \quad f_s = 1/T, $$

(2)

where $f_s$ is the sampling frequency, Eq. (1) is modified to the form

$$ s = f_s \left( x - \frac{1}{x} \right) = f_s \frac{x^2 - 1}{x}, \quad x \neq 0. $$

(3)

Consider $N$-th order CT linear system with the transfer function
Applying the LDI transform (3) to (4) and arrangement yield the transfer function of DT system [43].

\[ K_{DT}(s) = \sum_{k=0}^{2N} \frac{c_k s^k}{d_k s^k}. \]  

The \( c_k \) coefficients in the numerator can be obtained from the \( a_k \) coefficients in the numerator of the transfer function (4) of CT system. Similarly, the \( d_k \) coefficients in the denominator can be derived from the \( b_k \) coefficients in the denominator of (4). The procedure is given by the following uniform algorithm [43]:

\[ C = L \tilde{A}, \; D = L \tilde{B}. \]  

Here \( C \) and \( D \) are vectors of the coefficients of DT system

\[ C = [c_0 \; c_1 \; c_2 \ldots c_{2N}]^T, \; D = [d_0 \; d_1 \; d_2 \ldots d_{2N}]^T, \]  

and \( \tilde{A}, \; \tilde{B} \) are vectors of modified coefficients of the CT system

\[ \tilde{A} = [\tilde{a}_0 \; \tilde{a}_1 \; \tilde{a}_2 \ldots \tilde{a}_N]^T, \; \tilde{B} = [\tilde{b}_0 \; \tilde{b}_1 \; \tilde{b}_2 \ldots \tilde{b}_N]^T, \]  

where

\[ \tilde{a}_i = f_{k}^2 a_i, \; \tilde{b}_i = f_{k}^2 b_i, \; i = 0, 1, ..., N. \]

The transformation matrices \( L \) appearing in both formulae (6) are identical. In [42] and [43], \( L \) is denoted LDI matrix. Its structure is obvious from the following notation of the first equation in (6) [43]:

\[ c_{N+k} = \sum_{i=0}^{N-k} (-1)^{k+i} \binom{k+2i}{i} \tilde{a}_{k+2i}, \; 0 \leq k \leq N, \]  

\[ c_{N+k} = \sum_{i=0}^{N-k} (-1)^{k+i} \binom{k+2i}{i} \tilde{b}_{k+2i}, \; 0 \leq k \leq N, \]  

where the notation \([ \cdot ]\) denotes the integer part of the argument. The proof is given in [43].

Example of the matrix formula (10), (11) for \( N = 5 \) is given below as Eq. (12). It demonstrates specific symmetry of the LDI matrix, which can be described as follows:

\[ c_{N+k} = (-1)^k c_{N-k}, \; 0 \leq k \leq N. \]  

Example of the matrix formula (10), (11) for \( N = 5 \) is given below as Eq. (12). It demonstrates specific symmetry of the LDI matrix, which can be described as follows:
It is obvious that the complete LDI matrix can be unambiguously restored only from its upper square submatrix, which is expressed by Eq. (10). In [43], this matrix \( L_s \) is called shortened LDI matrix. Its analytical notation results from Eq. (10):\
\[
\begin{pmatrix}
0 & \ldots & 0 & 0 & (-1)^s \binom{n}{s} \\
0 & \ldots & 0 & (-1)^{s-1} \binom{n-1}{s-1} & 0 \\
0 & \ldots & 0 & (-1)^{s-2} \binom{n-2}{s-2} & 0 \\
\vdots & \ddots & \ddots & \ddots & \ddots \\
\end{pmatrix}
\]
(14)

Note that the shortened \((N+1)\times(N+1)\) LDI matrix can be used for transforming the coefficients of \(N\)-th-order CT system into a subset of the coefficients of DT system according to the rule
\[
C_S = L_S \tilde{A}, \quad D_S = L_S \tilde{B},
\]
where \(C_S\) and \(D_S\) are shortened vectors of the coefficients of DT system
\[
C_S = [c_0, c_1, c_2, \ldots, c_N]^T, \quad D_S = [d_0, d_1, d_2, \ldots, d_N]^T.
\]
(16)

The remaining coefficients, not included in the shortened vectors, can be easily determined according to the rule of symmetry (13).

Detailed analysis of the properties of shortened LDI matrix is given in [43]. Similar analysis is provided below for the inverse matrix to the matrix \(L_S\).

3 Inversion of Shortened LDI Matrix

Beside the shortened LDI matrix \(L_S\), we will also work with a lower square submatrix of \(L\) of size \((N+1)\times(N+1)\), which we denote \(L'_S\). This matrix is upper triangular with ones on the main diagonal, hence it is invertible. The main reason why we prefer this matrix is that triangular matrices are easier to handle (unlike \(L_S\) which has zero elements above the diagonal).

In the sequel, for two integers \(p, q\) the symbol \(p|q\) means that \(p\) is a divisor of \(q\) and \(p \neq q\) means that \(p\) is not a divisor of \(q\).

Let us denote \(L'_S = \{a_{s,r}\}_{r,s=0}^N\) and \((L'_S)^{-1} = \{b_{s,r}\}_{r,s=0}^N\). Then according to (11)
\[
a_{s,r} = \begin{cases} 
\left(\frac{r+s}{2}\right)^2 \frac{1}{r-s} & \text{for } r \leq s \text{ and } 2|\{r+s\}, \\
0 & \text{otherwise}.
\end{cases}
\]
(17)

In the following section it will be proven that
\[
b_{s,r} = \begin{cases} 
\frac{r+s}{2} + \frac{1}{r-s} & \text{for } r < s \text{ and } 2|\{r+s\}, \\
1 & \text{for } r = s, \\
0 & \text{otherwise}.
\end{cases}
\]
(18)

The proof is quite crucial. It is well-known that if an augmented matrix \((L'_S|E)\), where \(E\) is the identity matrix, is transformed using elementary row operations to a matrix \((E|P)\), then \(P = (L'_S)^{-1}\). Proceeding this way, rather complicated recurrence formulae are obtained from which the closed form (18) can be guessed, although it is not trivial (two partially overlapping Pascal triangles occur there). Nevertheless, once the formula for the inverse matrix is predicted, it is possible to verify its correctness—the product with \(L'_S\) must be equal to the identity matrix.

The inverse matrix \((L'_S)^{-1}\) can then easily be obtained as follows. Consider a matrix
\[
M = \begin{pmatrix}
0 & 0 & \ldots & 0 & (-1)^N \\
0 & 0 & \ldots & (-1)^{N-1} & 0 \\
\vdots & \vdots & \ddots & \ddots & \ddots \\
0 & -1 & \ldots & 0 & 0 \\
1 & 0 & \ldots & 0 & 0
\end{pmatrix}
\]

having zeros outside the antidiagonal. It is orthogonal, i.e., \(M^{-1} = M^T\). From (10) and (11) it is evident that \(L'_S = ML'_S\), hence it is sufficient to rearrange the rows of \(L'_S\) in the reverse order and then multiply them through consequently by \((-1)^N, (-1)^{N-1}, \ldots, -1, 1\). If we denote \(L'_S = \{c_{r,s}\}_{r,s=0}^N\), we obtain
\[
c_{r,s} = \begin{cases} 
\frac{r+s-N}{2} \left(\frac{1}{r-s} \right) & \text{for } 2|r+s-N, \\
\left(-1\right)^{\frac{r+s-N}{2}} \left(\frac{1}{r-s} \right) & \text{for } N \leq r+s \text{ and } 2|\{r+s-N\}, \\
0 & \text{otherwise}.
\end{cases}
\]
(19)
Further, \((L_x)^i = (L_x^i)^N = (L_x^i)^{N-i}\) Thus it is sufficient to rearrange the columns of \((L_x^i)^N\) in the reverse order and then multiply them through in consequence by \((-1)^{N}, (-1)^{N-1}, \ldots, -1, 1\). If we denote \((L_x)^{N-i} = \{d_{x, i}^{N-i}\}\) we obtain
\[
d_{x, i} = \begin{cases} (-1)^{N-i} \left[ \frac{r+N-s}{r} \right] + \left[ \frac{r+N-s-1}{r} \right] & \text{for } r+s < N \text{ and } 2|(N-r-s), \\ (-1)^{r} & \text{for } r+s = N, \\ 0 & \text{otherwise.} \end{cases}
\]

4 Proof of the Formula for the Inverse of \(L_x^i\)

Let us denote \(N_0\) the set of all nonnegative integers. To prove the main result, we will need the following auxiliary statement.

**Lemma** For every \(a, b \in \mathbb{N}_0\), there is
\[
\sum_{r=0}^{b} (-1)^r \binom{a+2r}{t} \binom{b+r+t}{a+2t} = (-1)^b.
\]

**Proof.** Let \(L\) denotes the left-hand side of (21). For every \(t \in \{0, \ldots, b\}\), the equality
\[
\binom{a+2t}{t} \binom{a+b+t}{a+2t} = \frac{(a+2t)!}{t!(a+t)!(a+2t)!(b-t)!} = \frac{b!(a+b+t)!}{t!(b-t)!(a+t)!(a+2t)!(b-t)!} = \binom{b}{a+b+t} \binom{t}{a+t}
\]
holds. Using this formula and the substitution \(t \to b-t\) we get
\[
L = \sum_{r=0}^{b} (-1)^r \binom{b}{a+b+t} \binom{t}{a+t} = \sum_{r=0}^{b} (-1)^r \binom{b}{a+2b-t} \binom{t}{a+b-t} = \sum_{r=0}^{b} (-1)^r \binom{b}{a+2b-t} \binom{t}{a+b-t} = (-1)^b \sum_{r=0}^{b} (-1)^r \binom{b}{a+2b-t} \binom{t}{a+b-t}.
\]

Consider a polynomial \(P(x, y) = (1-x+y)(1+y)^{a+b}\). Applying twice binomial formula we get
\[
P(x, y) = \sum_{t=0}^{b} (-1)^t \binom{b}{t} x^t (1+y)^{2b-t} = \sum_{t=0}^{b} (-1)^t \binom{b}{t} x^t \sum_{s=0}^{2t} \binom{2b-t}{s} \binom{a+2b-t}{s} y^s.
\]
The coefficient at the term \(x^ay^{b-t}\) is thus
\[
(-1)^t \binom{b}{t} \binom{a+2b-t}{t} \binom{a+b-t}{a+b-t}.
\]

Therefore, if we put \(y = x\), the coefficient at \(x^{a+b}\) will be
\[
\sum_{t=0}^{b} (-1)^t \binom{b}{t} \binom{a+b-t}{t} \binom{a+b-t}{a+b-t}.
\]

On the other hand, since \(P(x, x) = (1+x)^{a+b}\), this coefficient is obviously equal to 1. According to (22) this completes the proof of lemma.

Now we are ready to prove the main result.

**Theorem** Let \(A = (a_{i,j})_{i,j=0}^{N} \) and \(B = (b_{i,j})_{i,j=0}^{N}\) be \((N+1) \times (N+1)\) matrices whose entries are given by (17) and (18), respectively. Then \(B = A^{-1}\).

**Proof.** We will be done if we prove that \(AB = E\). Let \(F = AB = (f_{i,j})_{i,j=0}^{N} \). Both \(A\) and \(B\) are upper triangular matrices with ones in the principal diagonal, which implies that matrix \(F\) has this property as well.

So we will be done if we prove that all the entries of \(F\) above the principal diagonal are zeros. For \(r < s\), \(2 \nmid (r+s)\) obviously \(f_{s,r} = 0\). For \(r < s\), \(2 \nmid (r+s)\) we have
\[
f_{r,s} = \sum_{t=0}^{s-r} \binom{r+2t}{s-r} \binom{s}{t} + \sum_{t=0}^{s-r} \binom{s-2t}{r+2t} \binom{r+2t-r}{t} = \binom{r+2t+s}{s-r} \binom{s-r}{2} + \binom{r+2t-s}{2} \binom{s-r}{r+2t} = \binom{s-r}{2} \binom{s-r}{2}.
\]
\[ \sum_{t=0}^{s-r+1} (-1)^t \left( r + 2t + \frac{r + s - r + t}{r + 2t} \right) + \\
\sum_{t=0}^{r} (-1)^t r + 2t + \frac{r + s - r}{2} - 1 + t = 
\]
\[ = (-1)^\frac{r}{2} + (-1)^\frac{s-r}{2} = 0 \]

according to Lemma. This completes the proof.

5 Software-based Experiments

To verify the correctness and usefulness of the above analytical formulae of the inverse of the LDI matrix, the following computations in Maple 16 have been performed on PC with Intel i5-3337U CPU 1.8 GHz with 8 GB RAM.

Note that the computation with the LDI matrix and also with its inverse has two practical aspects, especially for higher matrix sizes, which are typical for designing digital filters of orders exceeding values of hundreds: the computation time and particularly the computation accuracy. The latter aspect is crucial if filter zeros and poles are to be evaluated from the filter coefficients. Since such computations are extremely sensitive to the truncation errors of the coefficients, these coefficients should be represented by higher number of digits than the commonly used format on PCs, double-precision binary floating-point, can provide.

Until explicit formula (20) was known, the only way to find the inverse of \( L_s \) was the use of some standard numerical method for the evaluation of the inverse matrix. Since entries of the inverse of \( L_s \) are integers and the inverse LDI transform is sensitive to rounding errors, symbolic computations should be used. As the following Table 1 shows, this way (see the rows “symbol.”) is very slow for large \( N \). Alternatively, software floating-point numbers with a lot of digits have to be used to obtain accurate results (see the rows “num.”). The entry “Digits” represents a number of digits which are necessary for providing numerically correct results. The conventional numerical calculation utilizing hardware 64-bit IEEE binary floating-point numbers is faster, but it gives incorrect results for \( N=100 \) and 200 and fails for \( N=1000 \). Therefore, formula (20) represents the true progress in the study of the inverse LDI transform.

Table 1: Computational times for various \( N \) (first column). Computing \( L_s \) from Eq. (14) (second column), computing \( L_s^{-1} \) inverse from Eq. (20) (third column) and by numerical inversion (fourth column).

<table>
<thead>
<tr>
<th>N/Digits</th>
<th>( L_s ) (14)</th>
<th>( L_s^{-1} ) (20)</th>
<th>( L_s^{-1} ) inversion</th>
</tr>
</thead>
<tbody>
<tr>
<td>symbol.</td>
<td>100/- 0.125 0.109 1.578</td>
<td>200/- 0.828 0.703 11.125</td>
<td>1000/- 88.875 63.578 1520.453</td>
</tr>
<tr>
<td>numer.</td>
<td>100/33 0.141 0.140 0.218</td>
<td>200/68 0.860 0.703 1.969</td>
<td>1000/347 101.844 65.922 244.360</td>
</tr>
</tbody>
</table>

6 Conclusions

Analytical formulae for the elements of the matrix, which is the inverse matrix of the shortened LDI matrix, are derived in the paper. That completes the last and hitherto missing part of the matrix algorithm for the conversion between the coefficients of CT and DT systems, which are interconnected via the LDI transformation. The experiments from Section 5 demonstrate the fundamental limits of conventional numerical inversion of large LDI matrices. Then the analytic formula for the inverse matrix, proposed in the paper, can overcome the numerical problems associated with the conventional inversion.

Acknowledgments

This work has been supported by the Project for the development of K215 and K217 Departments, UD Brno, Czech Republic.

References


