Taylor Series-based Tracking Algorithm for Through-Wall Tracking of a Moving Person

Mária Švecová, Dušan Kocur

Department of Electronics and Multimedia Communications
Faculty of Electrical Engineering and Informatics
Technical University of Košice
Park Komenského 13, 041 20 Košice, Slovak Republic
Maria.Svecova@tuke.sk, Dusan.Kocur@tuke.sk

Abstract: Target tracking using time of arrival measurements belongs to the primary tasks solved within radar signal processing. In this paper, the Taylor series-based tracking algorithm that uses time of arrival measurements for through wall tracking of the moving target is introduced. The proposed algorithm is derived from the Taylor series method applied for target localization. In contrast to the Taylor series method, the Taylor series-based tracking algorithm exploits for target positioning not only actual time of arrival measurements, but also the target position estimated in the previous time instant. In order to improve the tracking ability of the proposed algorithm, a suitable weighting of the input data of the algorithm is applied. The performance of the Taylor series-based tracking algorithm will be compared with the performance of the direct calculation method and linear Kalman filtering. For that purpose, two real scenarios of through wall tracking of a moving person will be analysed. The obtained results will show very clearly that the new Taylor series based tracking algorithm introduced in this paper can provide the better estimate of the target trajectory than the other tested localization and tracking algorithms.

Keywords: localization, target tracking, Taylor series, TOA, UWB radar system

1 Introduction

Electromagnetic waves occupying the spectral band below 4-5 GHz show reasonable penetration through most typical building materials, such as brick, wood, dry wall, concrete and reinforced concrete. This electromagnetic wave penetration property can be exploited with advantage by UWB radar systems operating in this frequency band for through wall detection and tracking of moving and breathing persons [3]. There are a number of practical applications where such radar systems can be very helpful, e.g. for through wall tracking of moving people during security operations, through wall imaging during fire, through rubble localization of trapped people following an emergency (e.g.

earthquake or explosion) or through snow detection of trapped people after an avalanche, etc. [3], [7].

Moving target tracking, i.e. determining target coordinates as the continuous function of the time, is a complex process that includes four phases [10]: target detection, the distance estimation between transmitting antenna (*Tx*), target (*T*) and receiving antenna (*Rx*), target localization and finally, tracking itself. The decision of whether the target is or is not present in a scanned area is the detection phase output. If some target is detected, the distance between *Tx*, *T* and *Rx* can be estimated. For that purpose, the measurement of the time of arrival (TOA) corresponding to the target to be tracked and electromagnetic wave propagation velocity between *Tx*, *T* and *Rx* can be used [1]. If the radar system is equipped with more than one receiving antennas, the target position can be estimated by using the suitable localization algorithm exploiting distances between *Tx*, *T* and *Rx* [12], [13]. Tracking is the last phase of the radar signal processing. Target tracking provides a new estimation of the target position based on the actual and foregoing estimations of the target positions. Usually, the target tracking will result in the target trajectory estimation error decreasing, including the smoothing of the target trajectory obtained as the localization phase output.

In the case of through wall tracking of moving targets, the estimation of the moving target position obtained by the localization phase is usually characterized by large estimation error (e.g. [10]). This effect is due to the extremely small power of signals scattered by the moving target and subsequently received by *Rx*, by the increased attenuation of the power of the electromagnetic waves at their transmission through the wall, and at the same time, by the complex and strong clutter due to multiple reflections of the electromagnetic waves emitted by the radar system from static objects located within the scanned area. With regard to these effects, target tracking algorithms must be applied in order to get moving target position estimation with acceptable error.

There are two basic approaches how to track a moving target. The former consists of localization followed by tracking itself. The target locations estimated in consecutive time instants create the target trajectory. Here, several iterative and non-iterative methods can be applied for the target localization [12], [13], [14]. On the other hand, most tracking systems utilize a number of basic and advanced modifications of Kalman filters, such as linear, extended and unscented Kalman filters [6]. In addition to Kalman filter theory, further methods of target tracking have been proposed [8]. These are usually based on smoothing of the target trajectory obtained by the target localization methods.

The latter approach to tracking the moving target is to join the localization and tracking phases into one phase of the radar signal processing. In this paper, a new tracking algorithm of that kind will be introduced for the radar system equipped with one transmitting and two receiving antennas. For the target coordinates estimation in the time instant t, the proposed algorithm exploits the TOA

corresponding to the target to be tracked for the actual time instant t and the distance between *Tx*, *T* and *Rx* determined for the previous time instant t-1. Then, these data are used as the target localization inputs for the Taylor series method (TSM) [13]. Because for the target localization, according to the outlined approach, the data concerning the actual position of the target (TOA for the time instant t) as well as the estimation of the previous position of the target (more precisely, the distance between *Tx*, *T* and *Rx* for the time instant $t-1$) are used, the proposed algorithm includes the localization and tracking phases. The proposed algorithm allows for the control of the influence of the reliability of the estimation of the TOA and the previous position of the target to the final estimation of the target position. Here, a weighting matrix $\mathbf{W}$ is used as the controlling parameter. This new algorithm, joining the phases of the target localization and target tracking based on the TSM application, will be referred to as the Taylor Series Based Tracking Algorithm (TST) in this paper.

In order to illustrate the TST performance, two real scenarios of through wall tracking of a moving person by the UWB radar system will be analyzed. Within both scenarios the M-sequence UWB radar system equipped with one transmitting and two receiving antennas will be used [4], [11]. The first scenario will focus on the comparison of the tracking ability and accuracy of the new TST with the traditional approach of the target localization and tracking represented by the direct calculation method [1], [13] and linear Kalman filtering [6]. The results obtained for this scenario will show that the TST is able to provide a better estimation of the target trajectory than the mentioned traditional approach. On the other hand, the latter scenario is devoted to bringing out the dependence of the TST tracking accuracy on the mentioned weighting matrix $\mathbf{W}$. In this case, the obtained results will outline how to choose the weighting matrix $\mathbf{W}$ for the different reliability of the estimation of the TST input quantities.

The structure of the paper is as follows. The problem statement concerning the target localization by the UWB radar system will be outlined in the next Section. Then, in the Section 3, the TST for the join target localization and tracking will be introduced. The illustration of the TST properties will be given in the Section 4. For that purpose, two scenarios of through wall tracking of a moving person will be used. Finally, conclusions and final remarks to this contribution are drawn in Section 5.

2 Target Localization: Problem Statement

Let us consider a UWB radar system equipped with one transmitting antenna *Tx* and two receiving antennas Rx_i, $i = 1,2$. The antenna positions are known and they are given by $Tx=(x_t,y_t)$ and $Rx_i=(x_i,y_i)$ for $i=1,2$. Let $TOA_i(t)$ for $i=1,2$ represent the estimation of TOA of the electromagnetic wave transmitted by *Tx*, reflected by the target and received by the *i*-th *Rx* in the time instant *t*. The goal is to determine the unknown target coordinates $T(t)=(x(t),y(t))$ in 2D for every observed time instant *t*.

In the next, the function

$$D(A,B,C) = \|AB\| + \|BC\| \tag{1}$$

will be used. In this expression, the symbol $\|XY\|$ is set for the Euclidean distance between the points X and Y.

The unknown target coordinates $T(t)=(x(t),y(t))$ for every observed time instant *t* can be computed by using the distances $d_i(t)$ given by

$$d_i(t) = D(Tx, T(t), Rx_i) = c.TOA_i(t), \quad i = 1,2 \tag{2}$$

where *c* is the electromagnetic wave propagation velocity. In our consideration, *c* is set to the electromagnetic wave propagation velocity in air, i.e. $c = 3.10^8$ ms^{-1}.

Under the real conditions, the distances $d_i(t)$ are estimated with an error represented by the additive noise components $e_i(t)$. Then, the estimated distances $d_i(t)$ can be modeled as

$$\begin{aligned} d_i(t) &= r_i(t) + e_i(t) = D(Tx, T(t), Rx_i) + e_i(t) = \|Tx\,T(t)\| + \|T(t)\,Rx_i\| + e_i(t) = \\ &= \sqrt{(x(t)-x_t)^2 + (y(t)-y_t)^2} + \sqrt{(x(t)-x_i)^2 + (y(t)-y_i)^2} + e_i(t) \end{aligned} \tag{3}$$

where $r_i(t)$ are true distances $D(Tx, T(t), Rx_i)$. Then, the target localization task consists in determining the target coordinates $x(t)$ and $y(t)$ by the solution of the nonlinear equation set (3). For that purpose, the TSM and the direct calculation method can also be used [13].

The target localization by the TSM was originally proposed for one way propagation time measurements in [5]. Then in [13], the original TSM was modified for the target localization based on $TOA_i(t)$ estimation. However, this method gives for the radar system equipped with one transmitting antenna and two receiving antennas the same results as simple localization based on the direct calculation method [1], [13]. For a radar system of that kind, the direct calculation method determines the target position under the condition that $e_i(t) \equiv 0$. This means that, according to the direct calculation method, the target coordinates correspond to the intersections of two ellipses defined by (3) for $i = 1,2$.

If the above mentioned simplification is not correct or at least acceptable, a large error in the target position estimation usually arises for the direct calculation method. In order to improve the target coordinates estimation accuracy, the simplified assumption $e_i(t) \equiv 0$ cannot be accepted. It could be done only if the number of the equations of (3) is greater than two. If this condition is fulfilled, the TSM can be used with advantage for the solution of this equation set. Then, improved accuracy of the target localization can be reached by an extension of the equation set (3) by new suitable equations and subsequently by the application of the TSM for the solution for such a set of nonlinear equations. By extending this fundamental idea, the TST can be derived, which is done in the next section of this contribution.

3 Taylor Series-based Tracking Algorithm

As was outlined in the previous sections, the key idea of the TST is the combination of the TSM and the extension of (3) by new equations. The conventional TSM uses for the target localization in the time instant t, $TOA_i(t)$ or the estimated distances $d_i(t)$. In contrast to the TSM, the TST is based on the idea that for the target localization in the time instant t, $TOA_i(t)$ or $d_i(t)$ as well as $\hat{T}(t) = \left(\hat{x}(t\text{-}1), \hat{y}(t\text{-}1)\right)$ i.e. the target coordinates estimated for the time instant $t-1$, will be used. Just the application of $\hat{T}(t) = \left(\hat{x}(t\text{-}1), \hat{y}(t\text{-}1)\right)$ enables us to extend the equation set (3) by two new equations. Then, the TST applied for the calculation of $\hat{T}(t) = \left(\hat{x}(t), \hat{y}(t)\right)$ i.e. the estimation of the target coordinates in the time instant t, consists of the following steps:

1 Let $\hat{d}_i(t\text{-}1) = D(Tx, \hat{T}(t-1), Rx_i)$. Then, the following equations can be obtained by using (1):

$$\hat{d}_i(t\text{-}1) = \sqrt{\left(\hat{x}(t\text{-}1) - x_t\right)^2 + \left(\hat{y}(t\text{-}1) - y_t\right)^2} + \sqrt{\left(\hat{x}(t\text{-}1) - x_i\right)^2 + \left(\hat{y}(t\text{-}1) - y_i\right)^2}, \quad i = 1, 2. \tag{4}$$

It can be found very easily, that for the distances $d_i(t)$ expressed by the target position $\hat{T}(t\text{-}1)$ estimated in the previous time instant *t*-1, the expression

$$d_i(t) = \hat{d}_i(t\text{-}1) + \varepsilon_i(t), \quad i = 1, 2 \tag{5}$$

holds. In this expression, $\varepsilon_i(t)$ is the difference between $d_i(t) = D(Tx, T(t), Rx_i)$ and $\hat{d}_i(t\text{-}1) = D(Tx, \hat{T}(t-1), Rx_i)$. For a relatively slowly moving target, the approximation

$$d_i(t) \approx \hat{d}_i(t\text{-}1), \quad i = 1,2 \tag{6}$$

can be accepted. Then, the distances $\hat{d}_i(t\text{-}1)$ can be expressed by using (3) and (6) as

$$\begin{aligned}\hat{d}_i(t\text{-}1) = &\sqrt{(x(t)\text{-}x_t)^2 + (y(t)\text{-}y_t)^2} + \\ &\sqrt{(x(t)\text{-}x_i)^2 + (y(t)\text{-}y_i)^2} + \sigma_i(t), \quad i = 1,2\end{aligned} \tag{7}$$

where $\sigma_i(t)$ is caused by the errors $e_i(t)$ from the equations (3) and $\varepsilon_i(t)$ from the equations (5).

The connection of the equations (3) and (7) gives a new set of the four non-linear equations with the unknown target coordinates $T(t) = (x(t), y(t))$ in the time instant t:

$$\begin{aligned}d_i(t) &= \sqrt{(x(t)\text{-}x_t)^2 + (y(t)\text{-}y_t)^2} + \sqrt{(x(t)\text{-}x_i)^2 + (y(t)\text{-}y_i)^2} + e_i(t), \\ \hat{d}_i(t\text{-}1) &= \sqrt{(x(t)\text{-}x_t)^2 + (y(t)\text{-}y_t)^2} + \sqrt{(x(t)\text{-}x_i)^2 + (y(t)\text{-}y_i)^2} + \sigma_i(t), \\ &i = 1,2.\end{aligned} \tag{8}$$

The unknown coordinates of the target $(x(t), y(t))$ are estimated from the equations (8) by the TSM where the $e_i(t)$ and $\sigma_i(t)$ are unknown components.

2 The non-linear equations (8) are linearized by their expanding in the Taylor series [2] around the point corresponding to the target position, which is subsequently estimated within the particular iterations of the iteration process and keeping only terms below second order. Let us set the initial estimate (x_v, y_v) of the target coordinates for the time instant t as the target position estimated in the time instant t-1, i.e. $(x_v, y_v) = (\hat{x}(t\text{-}1), \hat{y}(t\text{-}1))$ and define new functions

$$\begin{aligned}f_i(x(t), y(t)) = &\sqrt{(x(t)\text{-}x_t)^2 + (y(t)\text{-}y_t)^2} + \\ &\sqrt{(x(t)\text{-}x_i)^2 + (y(t)\text{-}y_i)^2}, \quad i = 1,2.\end{aligned} \tag{9}$$

Then, (9) can be rewritten as

$$\begin{aligned} f_i\left(x(t),y(t)\right) &= d_i(t) - e_i(t), \\ f_i\left(x(t),y(t)\right) &= \hat{d}_i(t-1) - \sigma_i(t), \quad i=1,2. \end{aligned} \tag{10}$$

If x_v and y_v are the initial estimates of the target coordinates, then

$$x = x_v + \delta_x, \quad y = y_v + \delta_y \tag{21}$$

where x and y are the true coordinates of the target in the time instant t and δ_x and δ_y are the target localization errors to be determined.

Expanding f_i in the Taylor series and retaining the first two terms produces

$$\begin{aligned} f_{iv} + a_{i1}\delta_x + a_{i2}\delta_y &\approx d_i(t) - e_i(t), \\ f_{iv} + a_{i1}\delta_x + a_{i2}\delta_y &\approx \hat{d}_i(t-1) - \sigma_i(t), \quad i=1,2 \end{aligned} \tag{32}$$

where

$$\begin{aligned} f_{iv} &= f_i(x_v, y_v), \\ a_{i1} &= \frac{\partial f_i}{\partial x} \mid x_v, y_v = \frac{x_v - x_t}{r_{tv}} + \frac{x_v - x_i}{r_{iv}}, \\ a_{i2} &= \frac{\partial f_i}{\partial y} \mid x_v, y_v = \frac{y_v - y_t}{r_{tv}} + \frac{y_v - y_i}{r_{iv}}, \\ r_{tv} &= \sqrt{(x_v - x_t)^2 + (y_v - y_t)^2}, \\ r_{iv} &= \sqrt{(x_v - x_i)^2 + (y_v - y_i)^2}. \end{aligned} \tag{43}$$

Then, (12) can be rewritten in the matrix form as

$$\mathbf{A}\delta = \mathbf{D} + \mathbf{e} \tag{54}$$

where

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}, \quad \delta = \begin{bmatrix} \delta_x \\ \delta_y \end{bmatrix}, \quad \mathbf{D} = \begin{bmatrix} d_1(t) - f_{1v} \\ d_2(t) - f_{2v} \\ \hat{d}_1(t-1) - f_{1v} \\ \hat{d}_2(t-1) - f_{2v} \end{bmatrix}, \quad \mathbf{e} = \begin{bmatrix} e_1(t) \\ e_2(t) \\ \sigma_1(t) \\ \sigma_2(t) \end{bmatrix}. \tag{65}$$

3 The set of the linear equations written by the matrix form (14) is solved by the weighted least-squares method [2][1] to produce a new estimate of the target coordinates to be applied in the next iteration of the iteration process. The iteration process continues until a pre-defined criterion is satisfied. Then δ obtained as the weighted least-squares solution of (14) with a weighting matrix $\boldsymbol{W}$ can be expressed in the form

$$\delta = \left[\mathbf{A}^T \mathbf{W} \mathbf{A} \right]^{-1} \mathbf{A}^T \mathbf{W} \mathbf{D}. \tag{76}$$

For the new estimate of the target position, the updated coordinates of the target according to

$$\begin{aligned} x_v &\leftarrow x_v + \delta_x \\ y_v &\leftarrow y_v + \delta_y \end{aligned} \tag{87}$$

is used, where δ_x and δ_y are given by (16). The iteration process is repeated until δ is sufficiently small.

Generally, the TSM and the TST require a good initial guess of the target coordinates. If the initial guess is not close to the true solution, divergence of the iteration process may occur. However, the divergence of the iteration process is easily detectable. For that purpose, we must check if $||\delta||$ at current iteration is larger than that of $||\delta||$ in the previous iteration. If this is the case, the iteration process does not converge and it must be started once again with a new initial guess of the target coordinates.

Generally, the particular weights of the weighting matrix $\boldsymbol{W}$ can be used to characterize the reliability of $TOA_i(t)$ and $\hat{T}(t-1)$ estimations. According to this idea, the more accurate measurements are placed with larger weights to stress the importance of the more reliable observations. In the TST, the weighting matrix $\boldsymbol{W}$ used in (16) is set to

$$\mathbf{W} = diag\left\{1, 1, w, w\right\} \tag{98}$$

where $w \epsilon R$ is the weighting factor. That matrix $\boldsymbol{W}$ is used to weight the actual estimated distances $d_i(t)$ with the added distances $\hat{d}_i(t-1)$ estimated based on using the target position at the time instant $t-1$. If $w < 1$, then $d_i(t)$ are used by the TST with greater reliability than that of $\hat{d}_i(t-1)$. In contrast, if $w > 1$, then $d_i(t)$ are used with lesser reliability than that of $\hat{d}_i(t-1)$. If $w = 1$, then $d_i(t)$ and $\hat{d}_i(t-1)$ are used by the TST with the same level of reliability.

4 When the iteration process has finished, the estimated position of the target in the actual time instant t is

$$\hat{T}(t) = \left(\hat{x}(t), \hat{y}(t)\right) = \left(x_v, y_v\right). \tag{109}$$

The computation flow of the TST is shown in the Fig. 1.

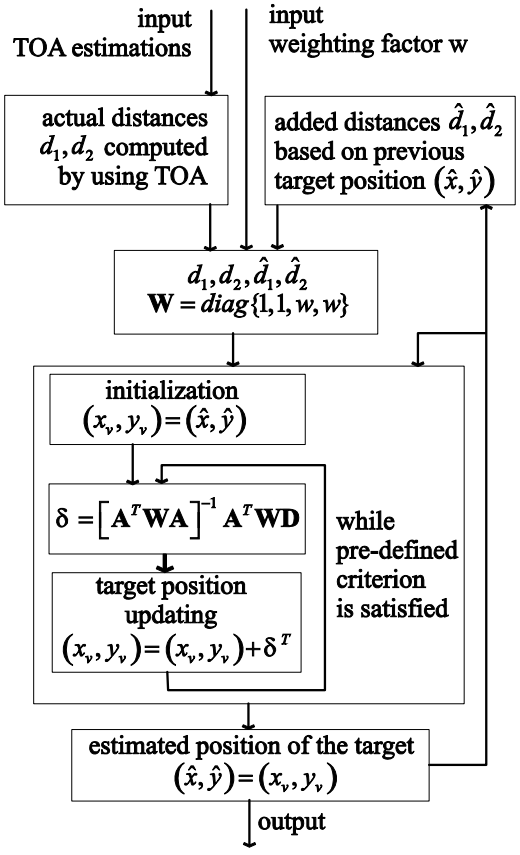


Figure 1
The TST computation flow

4 Performance of TST

In order to illustrate the TST performance, the TST has been applied for the target localization at two scenarios of through wall tracking of a moving person. In the both scenarios, the raw radar data analyzed in this contribution were acquired by means of the M-sequence UWB radar system with one transmitting and two receiving channels [3], [9], [11]. The system clock frequency for the radar device is about 4.5 GHz, which results in an operational bandwidth of about DC-2.25 GHz. The order of the M-sequence emitted by the radar is 9 [4], i.e. the impulse response covers 511 samples regularly spread over 114 ns. This corresponds to an observation window of 114 ns leading to an unambiguous range of about 16 m.

256 hardware averages of the environment impulse responses are always computed within the radar FPGA to provide a reasonable data throughput and to improve the SNR by 24 dB. The basic software of the radar device can provide the additional software averaging. In our measurements, the radar system was set in such a way as to provide approximately 10 impulse responses per second. The total power transmitted by the radar was about 1mW. The radar was equipped by the three double-ridged horn antennas placed along line. At the particular measurements, the transmitting antenna were located in the middle of two receiving antennas.

In order to get $TOA_i(t)$ estimation, raw radar data were processed by the radar signal processing procedure consisting of such signal processing phases as raw radar data pre-processing, background subtraction, detection and trace estimation. Then, $TOA_i(t)$ estimation was obtained as the output of the trace estimation phase. A detailed description of the mentioned radar signal processing procedure is beyond this contribution. It can be found e.g. in [10].

4.1 Scenario 1. Comparison of the TST with the Direct Calculation Method and Linear Kalman Filtering

This scenario was focused on the comparison of the tracking ability and accuracy of the new TST with the traditional approach to the target localization and tracking represented by the combination of the direct calculation method [13] and linear Kalman filtering [6].

The scheme of the analyzed scenario is outlined in Fig. 2. Scenario 1 is represented by moving person tracking through a light concrete wall with a thickness of 18 cm. The person is moving from position P(1), through positions P(2), P(3), and P(4) and back to position P(1) (Fig. 2). The distance between adjacent Rx_1 and Rx_2 was set to 76 cm. The Tx was located in the center between the Rx_1 and Rx_2. All antennas were placed 78 cm above the floor. Other distances are schematically depicted in Fig. 2.

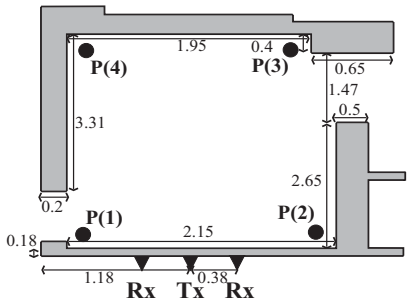


Figure 2
Scenario 1: measurement scheme

The $TOA_i(t)$ for i=1,2 corresponding to the target to be tracked have been obtained by using the above mentioned radar signal processing procedure. By using estimated $TOA_i(t)$, the target position of the moving person for every observation time instant has been determined. For that purpose, the direct calculation method was applied within the target localization phase. The target localization results are depictured in Fig. 3 by the thin curve. In order to improve the target trajectory estimation, target tracking by using linear Kalman filtering applied to the localization phase output given by the direct localization method has been used. For the join target localization and tracking, the TST proposed in this contribution has been also applied. The results of target tracking by these methods are given also in Fig. 3. Here, the target trajectory estimation by linear Kalman filtering and the TST are represented by a dotted and a thick curve, respectively. For this scenario, the weighting factor w has been set to w = 5. It can be seen from Fig. 3, that the positions of the moving person obtained by the TST are estimated more precisely than that of the positions estimated by the direct calculation method or by the direct calculation method followed with linear Kalman filtering. The largest difference between the true and estimated trajectories are visible around the positions P(3) and P(4), which is due to the wall niche (Fig. 2).

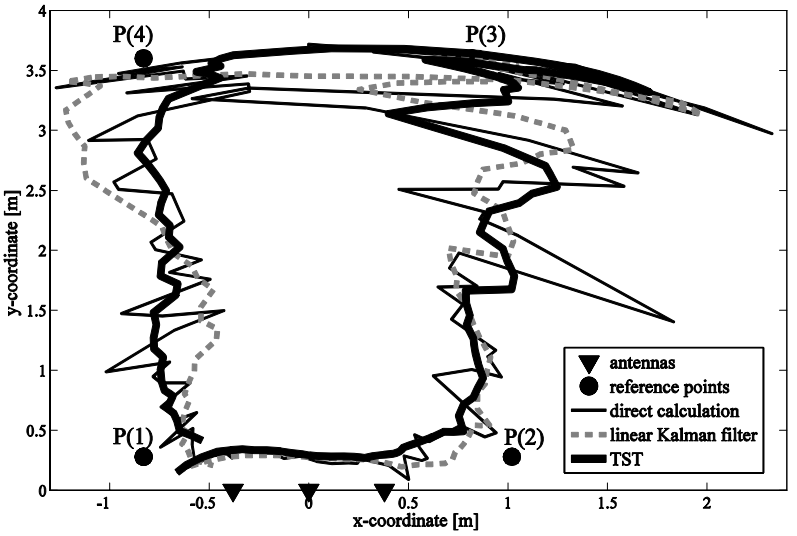


Figure 3

Scenario 1: Target trajectories estimated by the direct calculation method, linear Kalman filtering and the TST

The alternative form of the visualization of the target position estimation accuracy is shown in Fig. 4. In this figure, two tetragons sketched by the dashed lines are outlined. The part of the scanned area bordered by these tetragons, where the true trajectory of the target is located (i.e. the positions P(1) - P(4)), will be referred to as the region of the true positions of the target. The width of this region along the

x- and *y*-coordinates is set to 40 cm, which corresponds approximately to the effective width of a human body. In Fig. 4, the target position is represented by a single point. Since the real width of the target is non-zero and the resolution of the radar used for measurement is approximately 3 cm, we can accept the estimated position of the target as the true one, if the target is located inside the region of the true positions of the target. This approach allows us to evaluate the target position estimation accuracy as the percentage of the "true" estimates of the target positions. This quantity can be evaluated as the ratio of the number of the target positions inside the region of the true positions of the target to the total number of the estimated target positions. The percentages of the "true" estimates of the target positions for the Scenario 1 for the tested localization and tracking methods are given in Table 1. The results presented in this table confirm the results provided by Fig. 3, i.e. the TST in light of the target positioning accuracy overcomes very clearly the other tested methods.

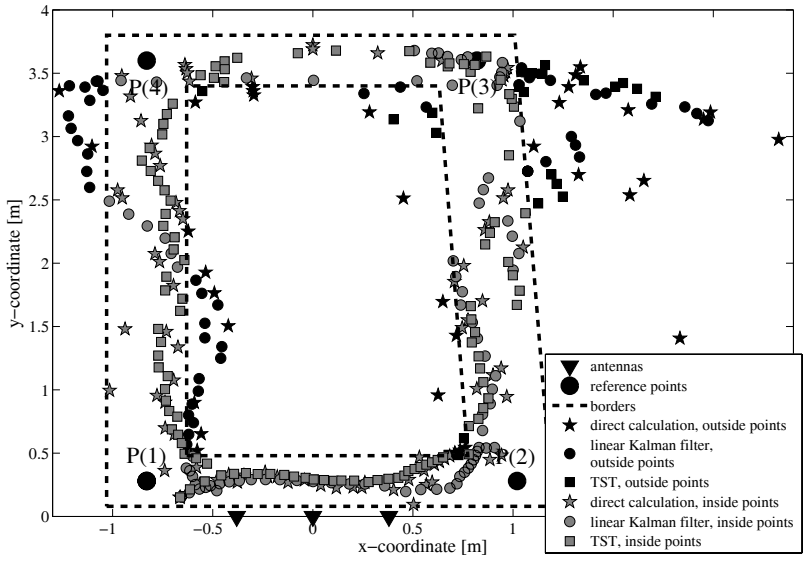


Figure 4

Scenario 1: Visualization of the target position estimation accuracy

In Table 1, the so-called average time of calculation is also brought out. This quantity represents the average time of calculation of the target positions based on TOA for the tested localization and tracking methods at their implementation in MATLAB environment. Therefore, this quantity can be taken as an approximated measure of the implemented algorithm complexity. Then, it can be identified from Table 1 that good TST performance is reached at a lower complexity than the complexity of linear Kalman filtering.

Table 1
Scenario 1: Comparison of the direct calculation method, linear Kalman fitlering and the TST

Localization methods	Percentage of the "true" estimates of the target positions	Average time of calculation [ms]
Direct calculation method	70%	0.07339
Linear Kalman filtering	63%	0.70462
TST	83%	0.29672

4.2 Scenario 2. The TST Performance Depending on the Selection of the Weighting Factor

Scenario 2 is devoted to the illustration of the dependence of the TST tracking accuracy on the selection of the above mentioned weighting factor w. In this case, the obtained result will outline how to choose the weighting factor w for the different reliability of the estimation of the TST input quantities.

The scheme of Scenario 2 is outlined in Fig. 5. This scenario is represented by a moving person tracking through a brick wall covered by tiles with a total thickness of 24 cm. The person was moving from position P(1), through positions P(2), P(3), and P(4) and back to position P(1) (Fig. 5). The distance between adjacent Rx_1 and Rx_2 was set to 260 cm. The Tx was located in the center between the Rx_1 and Rx_2. All antennas were placed 120 cm above the floor. Other distances are schematically depicted in Fig. 5.

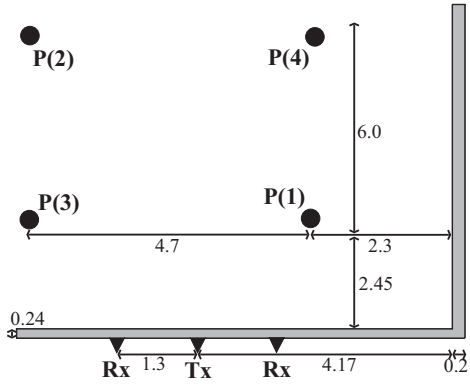


Figure 5
Scenario 2: measurement scheme

The $TOA_i(t)$ for i=1,2 corresponding to the target to be tracked have been obtained by using the radar signal processing procedure mentioned above. By using estimated $TOA_i(t)$, the target position of the moving person for every observation time instant has been determined. For that purpose, the direct calculation method and the TST method have been used. For the TST, the weighting factor w has been

set subsequently to w = 1, 5, 10. Similarly to in Scenario 1, the results of the target positioning for Scenario 2 are represented by the target trajectory estimations (Fig. 6), target positions inside and outside the region of the true positions of the target (Fig. 7) and by the percentage of the "true" estimates of the target positions (Table 2). In Scenario 2, the distance between the parallel lines representing the inside and outside border of the region of the true positions of the target (Fig. 7), has been set to 50 cm.

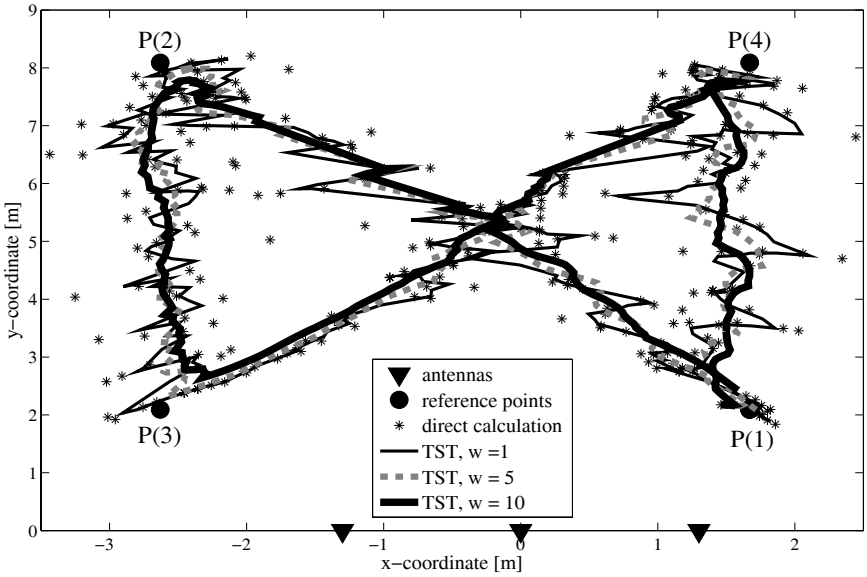


Figure 6

Scenario 2: Target trajectories estimated by using the direct calculation method and the TST for different values of the weighting factor w

The analyses of these results show that the TST surpasses the direct calculation methods. It can also be observed that the TST controlled by the higher weighting factor w can provide a more smoothed estimation of the target trajectories than if the lower values of the weighting factor are used. As can be identified from Figs. 6 and 7, the higher weighting factor w can be used with advantage especially in the case of the straight-line motion of the target. In constrast, if the person changes the motion direction sharply (e.g. in the positions P(1), P(2), P(3), P(4)), the TST using the higher weighting factor gives less precise results of the target positioning. These results indicate that the weighting factor w should be adapted according to the shape of the estimated trajectory of the target. However, a detailed solution of this task is beyond the scope of this paper and will be the object of our next research.

Table 2
Scenario 2: Comparison of the TST for the different values of the weighting factor *w*

Weighting factor *w*	Percentage of the "true" estimates of the target positions
w = 1	81%
w = 5	89%
w = 10	91%

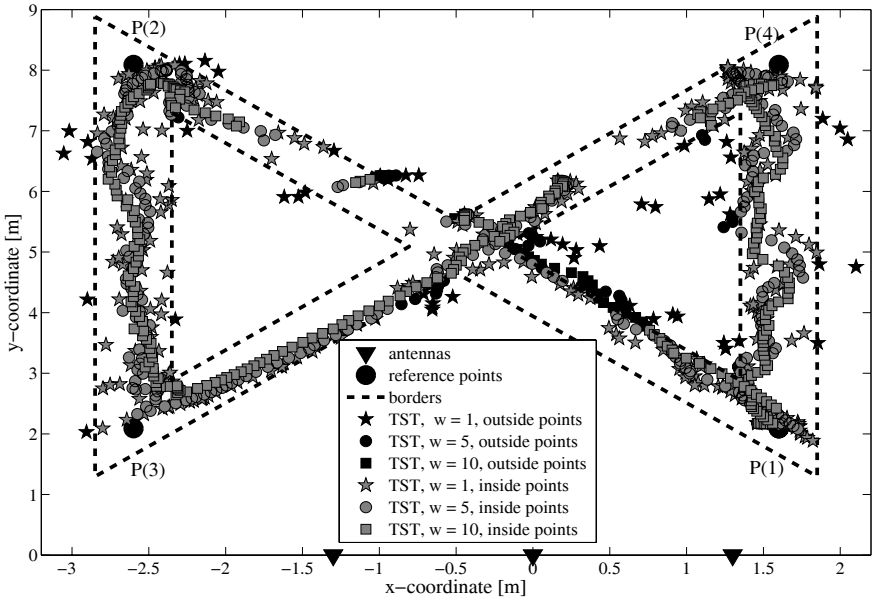


Figure 7
Scenario 2: Visualization of the target position estimation accuracy

Conclusions

In this paper, the TST has been introduced for the purpose of joining the localization and tracking of a moving target. The proposed algorithm uses for the target localization in the time instant t not only actual TOA estimates, but also the target coordinates estimated for the time instant $t-1$. The TST tracking ability has been illustrated and compared with the direct calculation method and linear Kalman filtering for through wall tracking of a moving person. The obtained results confirm that the TST provides a greater accuracy of target positioning at lower computational complexity than linear Kalman filtering. It has also been shown that the smoothness and accuracy of the target trajectory estimates by the TST depend on the selection of its weighting factor *w*. It has been outlined that an improvement in the TST performance could be reached by the adaptation of the weighting factor *w* according to the target motion style.

Acknowledgement

This work was supported by the Slovak Research and Development Agency under contract No. LPP-0080-09.

This work is the result of the project implementation of the Centre of Information and Communication Technologies for Knowledge Systems (project number: 26220120020), supported by the Research & Development Operational Programme funded by the ERDF.

References

[1] Aftanas, M., Rovňáková, J., Rišková, M., Kocur, D., Drutarovský, M.: An Analysis of 2D Target Positioning Accuracy for M-sequence UWB Radar System under Ideal Conditions, in Proceedings of the 17th International Conference Radioelektronika, Brno, Czech Republic, April 24-25, 2007, pp. 189-194

[2] Chong, E. K. P., Zak, S. H.: An Introduction to Optimization (Wiley-Interscience Series in Discrete Mathematics and Optimization), 3rd Edition, Wiley-Interscience, February, 2008

[3] Crabbe, S. et al.: Ultra Wideband Radar for through Wall Detection from the RADIOTECT Project, Fraunhofer Symposium, Future Security, 3rd Security Research Conference Karlsruhe, Congress Center Karlsruhe, Germany, September 10-11, 2008, p. 299

[4] Daniels, D.: M-Sequence Radar, in Ground Penetrating Radar, the Institution of Electrical Engineers, London, United Kingdom, 2004

[5] Foy, W. H.: Position-Location Solutions by Taylor-Series Estimation, in IEEE Transaction on Aerospace and Electronic Systems, Vol. AES-12, No. 2, March, 1976, pp. 187-194

[6] Grewal, M. S., Andrews, A. P.: Kalman Filtering: Theory and Practice Using MATLAB, 3rd Edition, Wiley-IEEE Press, September, 2008

[7] Oppermann, I., Hämäläinen, M., Iinatti, J.: UWB Theory and Applications, John Wiley & Sons, Ltd, England, November, 2004

[8] Ristic, B., Arulampalam, S., Gordon, N.: Beyond the Kalman Filter: Particle Filters for Tracking Applications, Artech House, 2004

[9] Rišková, M., Rovňáková, J., Aftanas, M.: M-Sequence UWB Radar Architecture for Throughwall Detection and Localisation, in Proceedings of the 7th PhD Student Conference and Scientific and Technical Competition of Students of Faculty of Electrical Engineering and Informatics, Technical University of Košice, Košice, Slovak Republik, May 25, 2007

[10] Rovňáková, J., Švecová, M., Kocur, D., Nguyen, T. T., Sachs, J.: Signal Processing for Through Wall Moving Target Tracking by M-sequence UWB Radar, in Proceedings of the 18th International Conference Radioelektronika, Prague, Czech Republic, April 24-25, 2008, pp. 65-68

[11] Sachs, J. et al.: Detection and Tracking of Moving or Trapped People Hidden by Obstacles using Ultra-Wideband Pseudo-noise Radar, in Processing of the 5th European Radar Conference (EuRAD 2008), Amsterdam, Netherlands, October 27-31, 2008, pp. 408-411

[12] Švecová, M.: Node Localization Methods in UWB Wireless Sensor Networks: A Review, in Proceedings of the 8th Scientific Conference of Young Researchers – FEI TU of Košice (SCYR 2008), Košice, Slovak Republic, May 28, 2008

[13] Švecová, M., Kocur, D., Zetik, R.: Object Localization Using Round Trip Propagation Time Measurements, in Proceedings of the 18th International Conference Radioelektronika, Prague, Czech Republic, April 24-25, 2008, pp. 41-44

[14] Yu, K., Montillet, J.-p., Rabbachin, A., Cheong, P., Oppermann, I.: UWB Location and Tracking for Wireless Embedded Networks, Signal Processing, Vol. 86, No. 9, September, 2006, pp. 2153-2171

Design of a Solar Tracker System for PV Power Plants

Tiberiu Tudorache[1], Liviu Kreindler[1, 2]

[1] Electrical Engineering Faculty, University Politehnica of Bucharest, 313 Splaiul Independentei, Sect. 6, Bucharest, Romania, e-mail: tudorach@amotion.pub.ro

[2] Technosoft, 266-268 Calea Rahovei, Sect. 5, Bucharest, Romania, e-mail: l_kreindler@tehnosoftmotion.com

Abstract: This paper deals with the design and execution of a solar tracker system dedicated to the PV conversion panels. The proposed single axis solar tracker device ensures the optimization of the conversion of solar energy into electricity by properly orienting the PV panel in accordance with the real position of the sun. The operation of the experimental model of the device is based on a DC motor intelligently controlled by a dedicated drive unit that moves a mini PV panel according to the signals received from two simple but efficient light sensors. The performance and characteristics of the solar tracker are experimentally analyzed.

Keywords: Solar tracker system, design and execution, experimental investigations

1 Introduction

The increasing demand for energy, the continuous reduction in existing sources of fossil fuels and the growing concern regarding environment pollution, have pushed mankind to explore new technologies for the production of electrical energy using clean, renewable sources, such as solar energy, wind energy, etc. Among the non-conventional, renewable energy sources, solar energy affords great potential for conversion into electric power, able to ensure an important part of the electrical energy needs of the planet.

The conversion of solar light into electrical energy represents one of the most promising and challenging energetic technologies, in continuous development, being clean, silent and reliable, with very low maintenance costs and minimal ecological impact. Solar energy is free, practically inexhaustible, and involves no polluting residues or greenhouse gases emissions.

The conversion principle of solar light into electricity, called Photo-Voltaic or PV conversion, is not very new, but the efficiency improvement of the PV conversion

equipment is still one of top priorities for many academic and/or industrial research groups all over the world.

Among the proposed solutions for improving the efficiency of PV conversion, we can mention solar tracking [1]-[3], the optimization of solar cell configuration and geometry [4]-[5], new materials and technologies [7]-[9], etc.

The global market for PV conversion equipment has shown an exponential increase over the last years, showing a good tendency for the years to come.

Physically, a PV panel consists of a flat surface on which numerous p-n junctions are placed, connected together through electrically conducting strips. The PV panel ensures the conversion of light radiation into electricity and it is characterized by a strong dependence of the output power on the incident light radiation.

As technology has evolved, the conversion efficiency of the PV panels has increased steadily, but still it does not exceed 13% for the common ones. The PV panels exhibits a strongly non-linear I-V (current - voltage) characteristic and a power output that is also non-linearly dependant on the surface insolation.

In the case of solar light conversion into electricity, due to the continuous change in the relative positions of the sun and the earth, the incident radiation on a fixed PV panel is continuously changing, reaching a maximum point when the direction of solar radiation is perpendicular to the panel surface. In this context, for maximal energy efficiency of a PV panel, it is necessary to have it equipped with a solar tracking system.

The topic proposed in this paper refers to the design of a single axis solar tracker system that automatically searches the optimum PV panel position with respect to the sun by means of a DC motor controlled by an intelligent drive unit that receives input signals from a light intensity sensor.

2 Solar Tracking and PV Panel Efficiency

Compared to a fixed panel, a mobile PV panel driven by a solar tracker is kept under the best possible insolation for all positions of the Sun, as the light falls close to the geometric normal incidence angle. Automatic solar tracking systems (using light intensity sensing) may boost consistently the conversion efficiency of a PV panel, thus in this way deriving more energy from the sun.

Technical reports in the USA have shown solar tracking to be particularly effective in summer, when the increases in output energy may reach over 50%, while in autumn they may be higher than 20%, depending on the technology used.

Solar tracking systems are of several types and can be classified according to several criteria. A first classification can be made depending on the number of rotation axes. Thus we can distinguish solar tracking systems with a rotation axis, respectively with two rotation axes. Since solar tracking implies moving parts and control systems that tend to be expensive, single-axis tracking systems seem to be the best solution for small PV power plants. Single axis trackers will usually have a manual elevation (axis tilt) adjustment on the second axis which is adjusted at regular intervals throughout the year.

A single-axis solar tracking system uses a tilted PV panel mount, Fig. 1, and a single electric motor to move the panel on an approximate trajectory relative to the Sun's position.

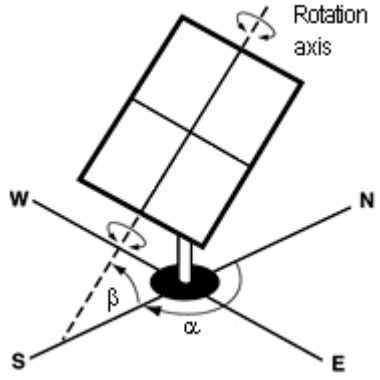


Figure 1
Principle of the single-axis solar tracking system

Another classification of solar tracking systems can be made depending on the orientation type. According to this criterion, we can identify solar tracking systems that orient the PV panels based on a previously computed sun trajectory, in comparison with panels with an on-line orientation system that reacts to the instantaneous solar light radiation. The later solution is more efficient and it was chosen for the solar tracking system proposed in this paper.

Another criterion for the solar tracker classification refers to its activity type. According to this criterion, we can distinguish active or passive solar trackers.

3 Design of the Proposed Solar Tracking System

The proposed solar tracking system should satisfy certain technical requirements specific to the studied application, as follows:

- minimum energy consumption, for the maximization of global efficiency of the installation and optimum performance-cost ratio;
- reliability in operation, under different perturbation conditions (wind, dust, rain, important temperature variations);
- simplicity of movement solution (motor, gears, sensors), to diminish the cost and to increase the viability;
- possibility of system integration in a monitoring and control centralized structure, which means a digital control solution.

Taking into account these implicitly necessary technical requirements, the chosen solution to drive the PV panel is based on the following components:

- a DC electric motor, voltage mode driven, with current monitoring, without movement sensors (speed or position);
- a motor control system of intelligent drive type, completely digital, that allows the implementation of the digital control of the motor as well as the implementation in a dedicated motion control language of the PV panel orientation application;
- a measurement system for light intensity applied to the PV panel, representing the sensor that commands the solar panel movement.

The chosen technical solution offers the following important advantages:

- simplicity of power scheme: DC motor and H bridge converter (4 transistors) for the motor drive;
- use of a compact drive equipment, with a high degree of integration and intelligence, that incorporates in a single module both the power converter and its command system, motion command unit (motion controller), and specific automation elements (of PLC type);
- use of an innovative solution, simple and reliable for the measurement system of light signal intensity.

3.1 Motor Drive System

DC motor. The parameters of the DC motor used as the movement execution element are: rated voltage 24 V, rated current 3 A, maximum speed 3000 rpm, gear box with a speed reduction ratio of 1: 20.

Intelligent drive unit IBL2403 [5]. To command the DC motor, we used the intelligent drive unit IBL2403, designed and executed by Technosoft ® [10]. This component is a completely digital drive system, executed using DSP technology, dedicated to the command of DC electric motors, sinusoidal or trapezoidal commutated brushless motors or stepper motors. It accepts as position sensors, incremental encoders, digital or linear Hall sensors.

IBL2403 allows the command of the motor in voltage mode, current loop, speed or position loop. Being a control system in distributed architecture, it will be placed close to the electric motor, removing the distance problem, connecting wires and perturbations appearing for centralized control solutions.

The IBL2403 drive unit can be used in multi-axes structures; thus the proposed solution can be extended to an array of PV panels that can communicate with each other and with central computation unit by a CAN-bus communication line.

The IBL2403 unit is a drive unit programmable in the high level language TML (Technosoft Motion Language), which allows:

- the setting of various movement modes: profiles, PT, PVT, electronic cam, extern, etc.,
- on-line modification of movement mode,
- the execution of "homing" type movements,
- the execution of decisions in the program, jumps or TML function calls;
- handling of digital and analogical I/O ports of the drive unit;
- the execution of arithmetic and logical operations;
- communication among axes and the control of the movement of other axes;
- synchronization with other axes from a multi-axis system.

To access the drive unit, to set the drive parameters and to implement the motion application, one can use the EasyMotion Studio program, an integrated graphical platform that simplifies the application development.

The main characteristics of the IBL2403 'drive unit are the following:

- Completely digital drive, multi-motor (the same unit can control DC motors, sinusoidal or trapezoidal commutated brushless motors, and stepper motors);
- Voltage control, torque control, speed control of position control of the machine;
- Programmable motion modes: trapezoidal profile, S-curve, PT, PVT, electronic cam or gear box, external or analog reference, 33 Home modes;
- 5 programmable digital inputs;

- 2 programmable digital outputs;
- Communication protocols RS-232 and CAN;
- 1.5 kB RAM memory;
- 8 kB x 16 EEPROM memory;
- Rated frequency of PWM: 20 kHz;
- Supply voltage: 12-28 V, rated current 3 A, peak current 6 A,
- Compact design.

The layout of the IBL2403 drive unit and the connection of this drive to the DC motor and to the supply source are presented in Figs. 2-4.

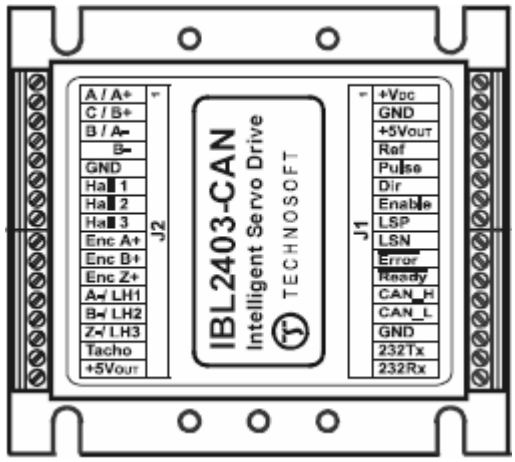


Figure 2
Layout of IBL2403 intelligent drive unit

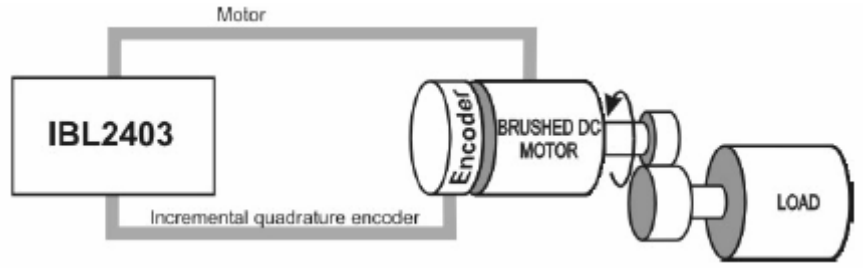


Figure 3
Dc motor drive scheme using IBL2403 unit

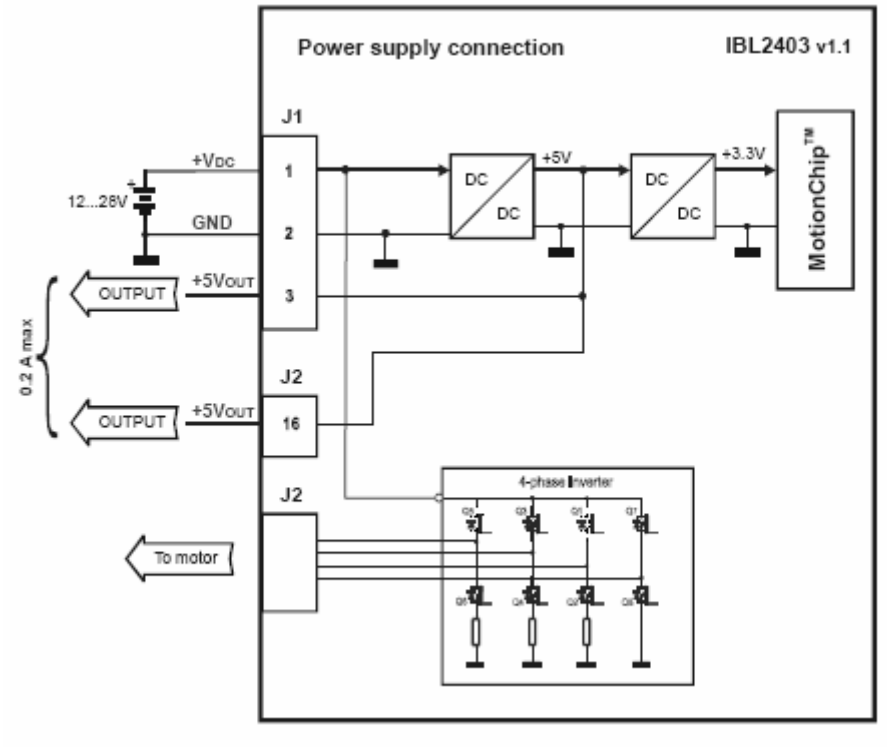


Figure 4
Connection of supply source and motor to the IBL2403 drive unit

3.2 Light Intensity Sensor

To command the PV panel motion, we used two light intensity sensors, executed using two luminescent diodes of LED type, placed so that the signal they generate is correlated with the light intensity applied to the PV panel, as in Fig. 5. The two LEDs are placed normal to the panel surface and are separated by an opaque plate. Thus, depending on the relative position between the PV panel and the solar light direction, one of the two LEDs will generate a stronger signal and the other LED a weaker one. In principle, the stronger signal will indicate the movement direction of the PV panel, so as to be normally oriented to the incident sun light rays and thus to have a maximum conversion efficiency of light into electricity. The operation scheme of the experimental model of the solar tracking system supposed the amplification of the two signals generated by the two LEDs up to the value range of the analog inputs of IBL2403 drive unit, at which were connected the two signals.

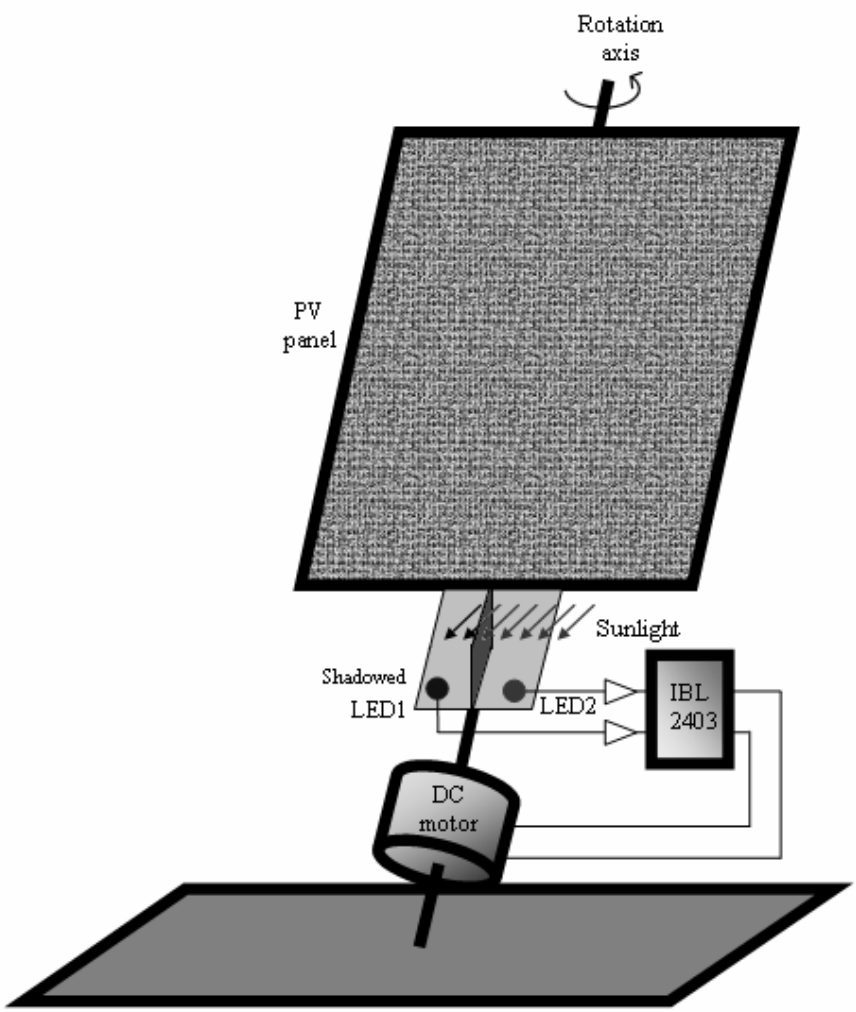


Figure 5
Principle of light sensors and motion control of PV panel

The proposed solution, which uses two independent light intensity sensors and the separate measurement of their output signals, has the following advantages:

- by measuring each separate signal we can detect, for the maximum signal intensity of the signal, if the light intensity is strong enough to justify the panel movement, which means additional energy consumption,
- in the case that the maximum signal intensity is higher than the minimum admitted values, an additional decision criterion of movement will be determined by the difference between the signals from the two LEDs. When

this difference becomes greater than an imposed limit, the command for a new movement of the PV panel will be triggered and thus the panel will be oriented again in order to achieve a better position with respect to the sun.

Thus, the proposed solution will remove the intermittent, frequent and unnecessary movements of the PV panel that would entail a higher consumption of energy, and thus decreasing in this way the conversion efficiency of the system.

We should point out that when the motor is not active, the power circuit of the IBL2403 drive unit can be completely deactivated, thus minimizing in this way the energy consumption of the system.

4 Experimental Model and Results

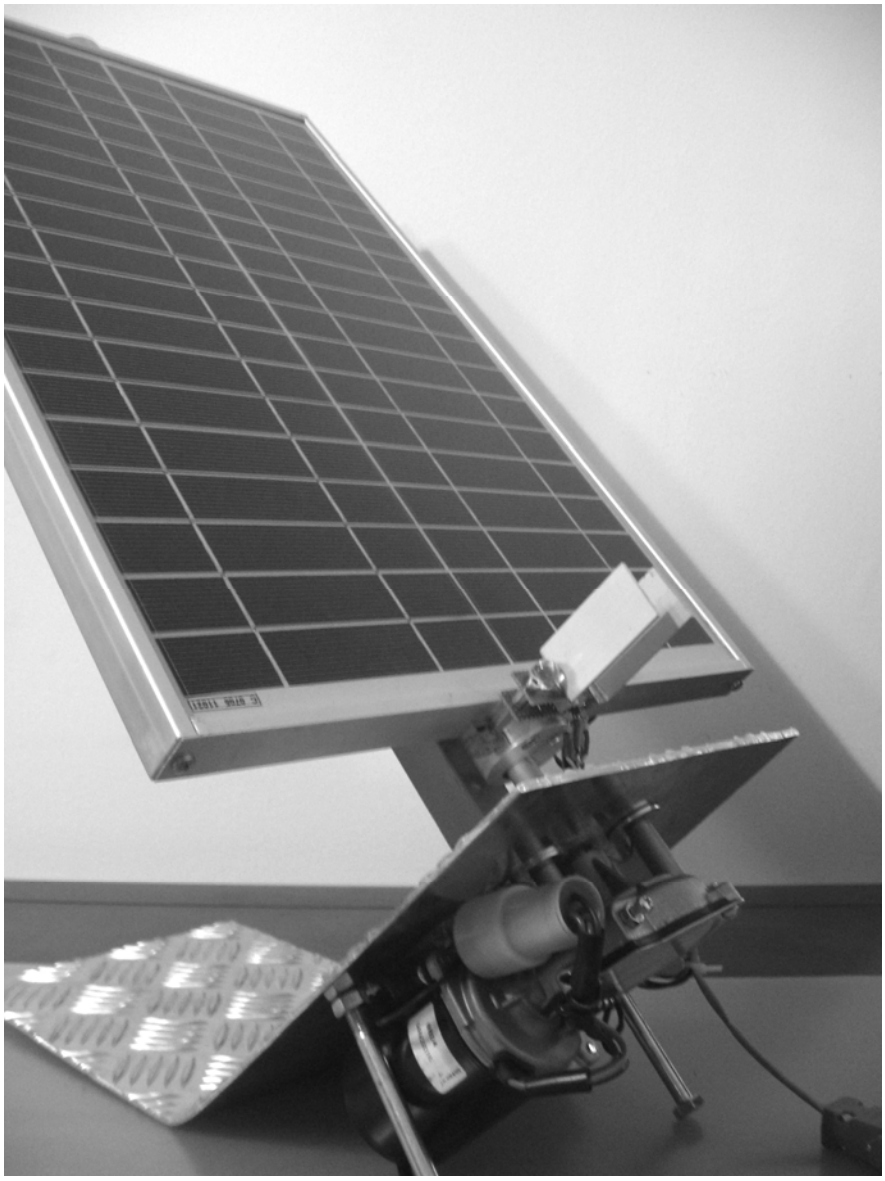
Based on the design configuration presented in the previous section, we executed a completely operational small scale experimental model of a single-axis solar tracking system. With this model we were able to evaluate the specific control components using the Technosoft EasyMotion Studio platform. Starting from this small scale equipment, we can go forward and implement the proposed technology within larger power PV systems.

The experimental model, along with the light intensity measurement device and the two LEDs, are presented in Fig. 6.

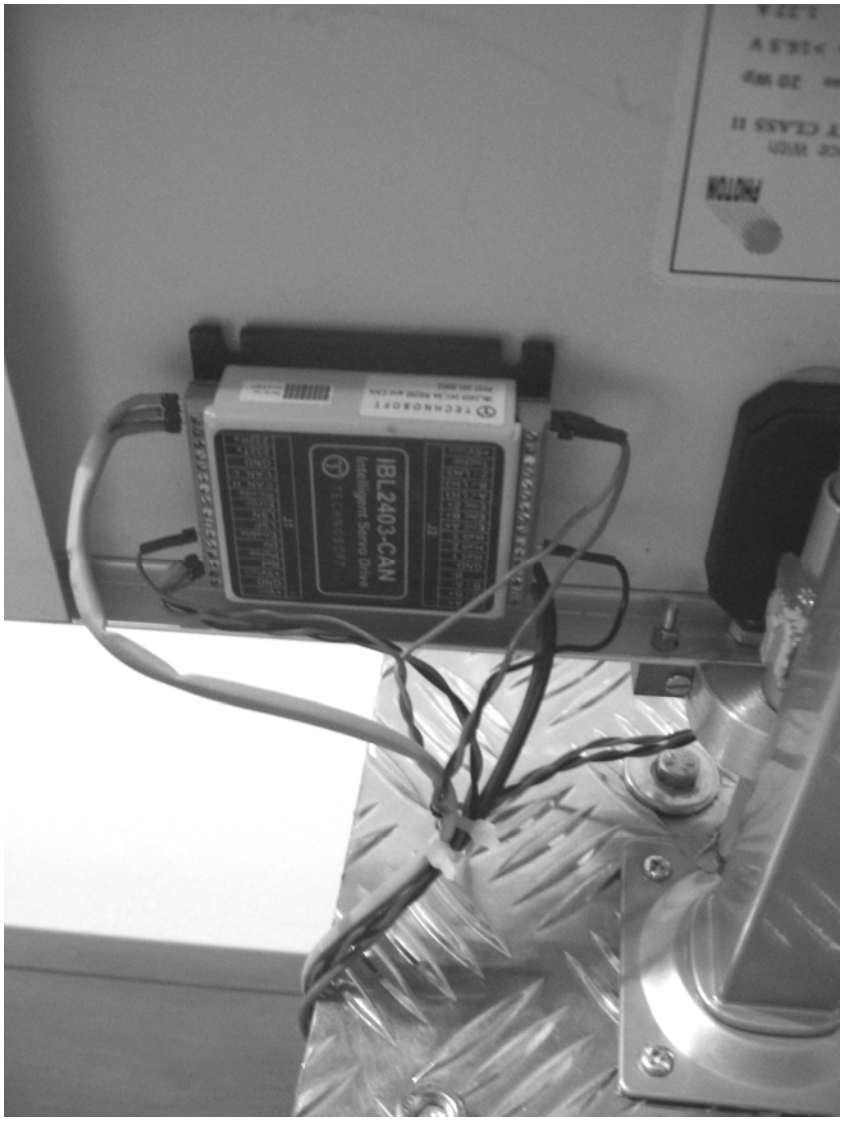
The start-up of the designed system and the setting of the specific parameters of the motor and of the IBL2403 drive unit were executed using EasyMotion Studio.

In Fig. 7 is presented the dialog window for the DC motor parameter settings and in Fig. 8 the dialog window for the IBL2403 drive unit parameter settings.

In Figs. 9 and 10 are presented the variations of the received signals from the two LEDs, when a light source moves in front of the PV panel and the difference signal that can be used for the decision taking module to move the panel to the right or to the left.



a)



b)

Figure 6
Experimental model of a single-axis solar tracking system;
a) system overview; b) detail of the IBL2403 drive unit

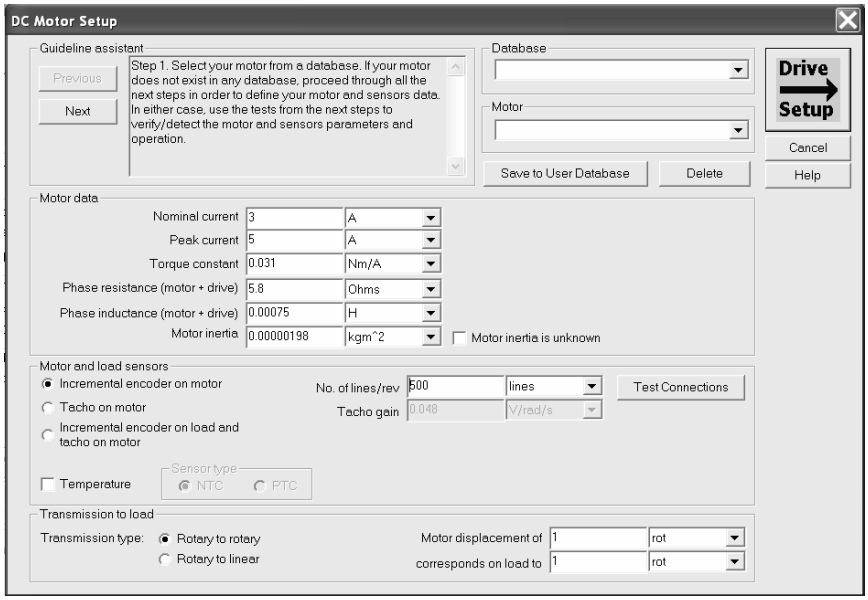


Figure 7
Dialog window for DC motor parameter settings

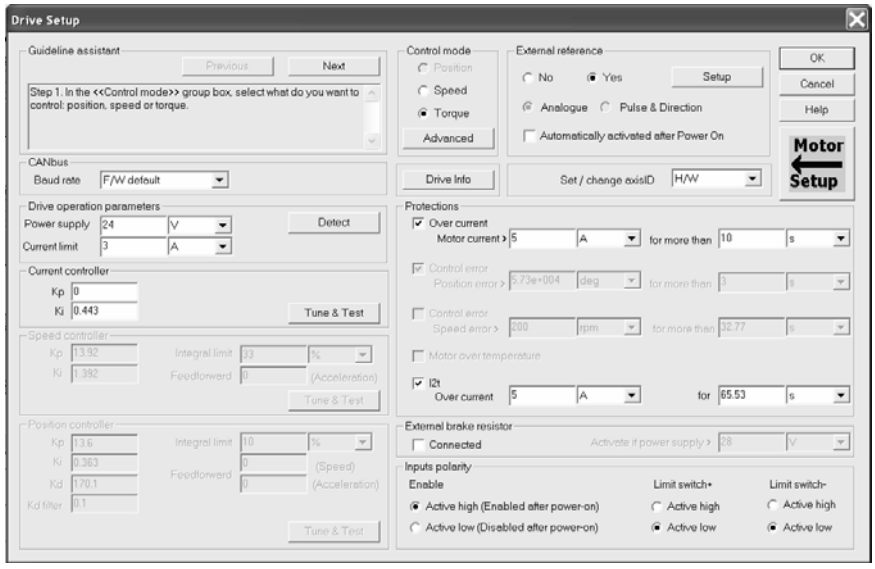


Figure 8
Dialog window for IBL2403 drive unit parameter settings

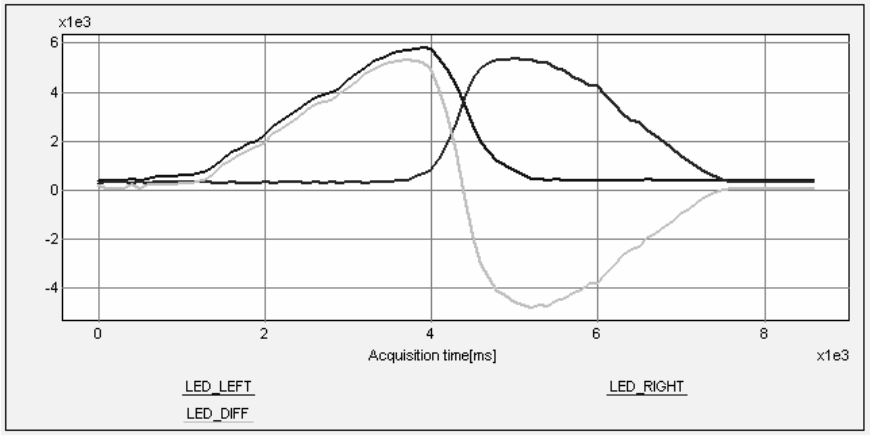


Figure 9
Signals received from the left and right LEDs and their difference when moving a light source from the left to the right, in front of the PV panel

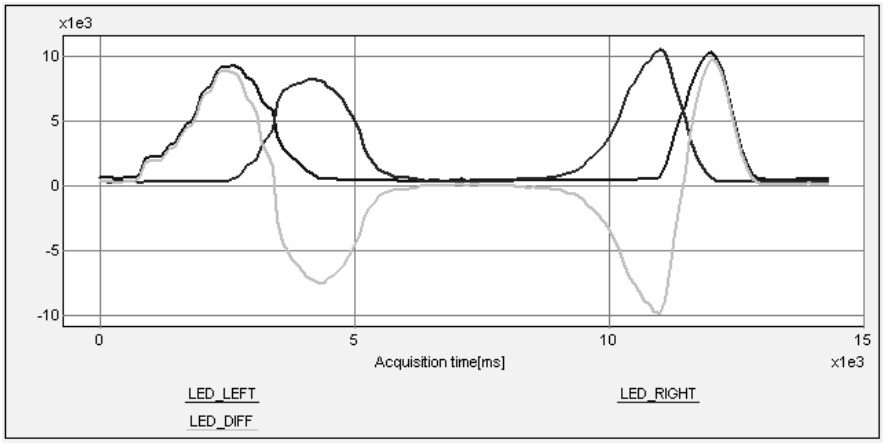


Figure 10
Signals received from the left and right LEDs and their difference when moving from the left to the right and back to the left, the light source, in front of the PV panel

Based on the difference between the two signals received from the two LEDs, compared with two imposed trigger values (programmable by the user), the command signals for the movement of the PV panel to the left or to the right were generated as in Figs. 11 and 12.

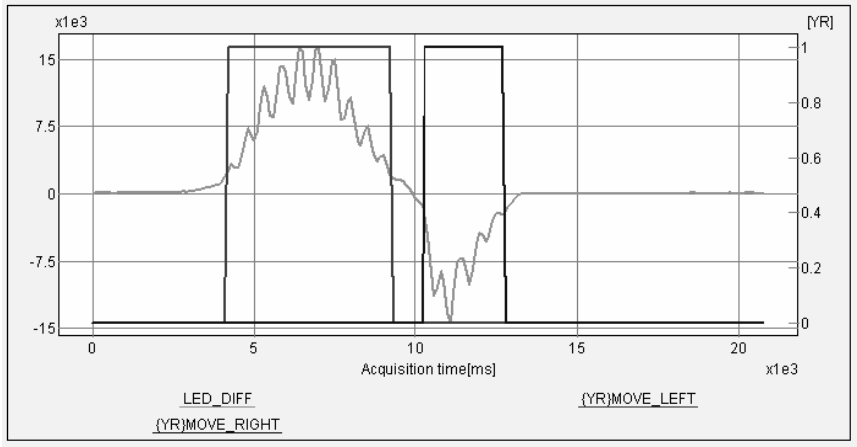


Figure 11

Generation of movement command of the PV panel to the right or to the left when moving a light source from the left to the right in front of the PV panel

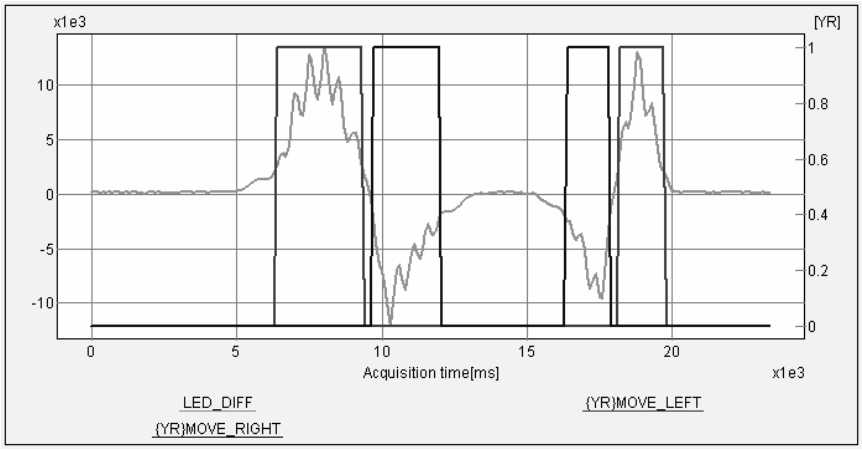


Figure 12

Generation of movement command of the PV panel to the right or to the left when moving a light source from the left to the right and back to the left, in front of the PV panel

In Fig. 13 are presented the variations of the signals received from the LEDs in stand-by state, when the difference between the two signals is lower than the trigger value for executing a movement command. We should notice that in the stand-by state no movement command is generated.

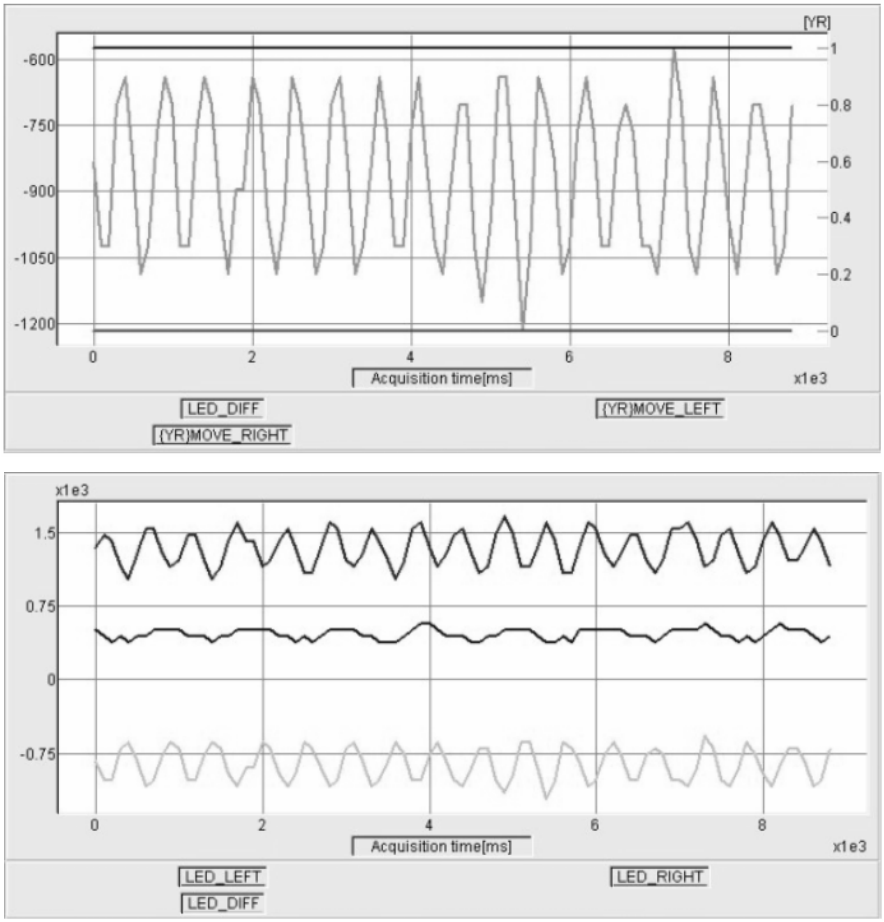


Figure 13
LED signals in the stand-by zone

A global view of the EasyMotion Studio platform, used for the command application of the PV panel movement is shown in Fig. 14.

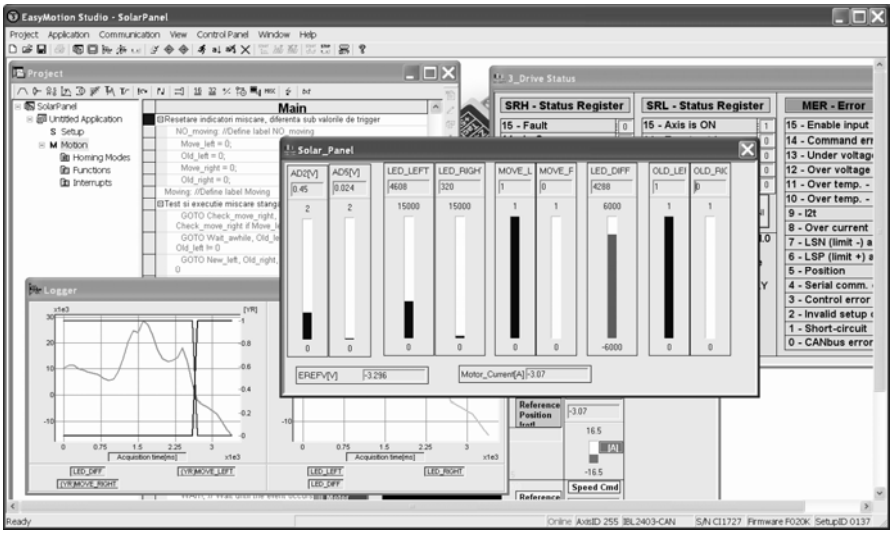


Figure 14

Global view of EasyMotion Studio platform developed by Technosoft

Conclusions

Based on the obtained results we can conclude that the proposed solution for a solar tracking system offers several advantages concerning the movement command of the PV panel:

- an optimum cost/performance ratio, which is achieved via the simplicity of the adopted mechanical solution and the flexibility of the intelligent command strategy;
- a minimum of energy consumption, due to the fact that the panel movement is carried out only in justified cases, eliminating unnecessary consumption of energy, and due to the cutting of the power circuits supply between the movement periods of the PV panel;
- a maximization of output energy produced by the PV panel, through an optimal positioning executed only for sufficient values of light signal intensity;
- a guarantee of the panel positioning starting from any initial position of the PV panel;
- the elimination of unnecessary movements, at too small intensities of the light signals or at too small differences between the signals received from the two LEDs;

- the possibility of extending this solution to an array of PV panels, connected to each other, with inter-connected operability by CAN protocol communication among the panels and managed by a central computation unit for monitoring and control;
- the possibility of centralized monitoring and diagnosis of the system operation.

Based on the obtained results we can affirm that proposed solution is effective and presents interesting advantages from the point of view of practical applicability to larger power PV structures.

Acknowledgement

This work was supported by the Romanian Ministry of Education and Research, in the frame of the research grant PNII, 21-075/2007.

References

[1] J. A. Beltran, J. L. S. Gonzalez Rubio, C.D. Garcia-Beltran: Design, Manufacturing and Performance Test of a Solar Tracker Made by an Embedded Control, CERMA 2007, Mexico

[2] O. Stalter, B. Burger, S. Bacha, D. Roye: Integrated Solar Tracker Positioning Unit in Distributed Grid-Feeding Inverters for CPV Power Plants, ICIT 2009, Australia

[3] M. A. Panait, T. Tudorache: A Simple Neural Network Solar Tracker for Optimizing Conversion Efficiency in Off-Grid Solar Generators, ICREPQ 2008, Spain

[4] A. M. Morega, J. C. Ordonez, P. A. Negoias, R. Hovsapian: Spherical Photovoltaic Cells – A Constructal Approach to Their Optimization, OPTIM 2006, Romania

[5] A. M. Morega, A. Bejan: A Constructal Approach to the Optimal Design of Photovoltaic Cells, Int. Journal of Green Energy, pp. 233-242, 2005

[6] J. Horzel, K. De Clerq: Advantages of a New Metallization Structure for the Front Side of Solar Cells, 13th EC Photovoltaic Solar Energy Conference, France, 1995

[7] P. I. Widenborg, G. Aberle: Polycrystalline Silicon Thin-Film Solar Cells on AIT-Textured Glass Superstrates, Advances in OptoElectronics Journal, Vol. 2007

[8] P. A. Basore: Manufacturing a New Polycrystalline Silicon PV Technology, Conference Record of the 2006 IEEE 4th World Conference on Photovoltaic Energy Conversion, pp. 2089-2093, 2006

[9] P. Turmezei: Chalcogenide Materials for Solar Energy Conversion, Acta Polytechnica Hungarica, Vol. 1, No. 2, pp. 13-16, 2004

[10] Technosoft: IBL2403 Intelligent Drive User Manual

Effect of the Measured Pulses Count on the Methodical Error of the Air Radio Altimeter

Ján Labun, František Adamčík, Ján Piľa

Faculty of Aeronautics, Technical University of Košice
Rampova 7, 041 21 Košice, Slovakia
e-mails: jan.labun@tuke.sk, frantisek.adamcik@tuke.sk, jan.pila@tuke.sk

Ladislav Madarász

Faculty of Electrical Engineering and Informatics, Technical University of Košice,
Letná 9, 042 00 Košice, Slovakia
e-mail: ladislav.madarasz@tuke.sk

Abstract: Radio altimeters are based on the principle of radio location of the earth's surface using a frequency-modulated standing wave. The relatively simple method of measurement consists in the evaluation of the number of pulses generated as resulting from the mixing of the transmitted and received signals. Such a change in the number of modulated pulses within a certain altitude interval, however, is not so simple and is a determinant issue in defining the precision of the radio altimeter. Being knowledgeable of this law in a wider context enables us to enter into discussion on the possibilities of further increasing the precision of measuring low altitudes. The article deals with the law underlying the change in the number of radio altimeter pulses with the changing altitude measured.

Keywords: radio altimeter, precision of measuring altitude, frequency modulation, the law of pulse number changes, critical altitude, methodical error

1 Introduction

The frequency modulated radio altimeter is understood as a radio altimeter that makes uses of a harmonic, frequency-modulated high-frequency signal to measure aircraft altitude over terrain. The altitude measured usually falls between 0 and 1500 m depending on the type of aircraft and the radio altimeter used. Radio altimeters operate on the principle of generating a high-frequency, frequency m-modulated signal which is then radiated via a transmitter antenna towards the

surface of the earth. The reflected signal subsequently received by a receiver antenna is then fed into the balanced mixer, into which the so-called direct signal from the transmitter has already been sent simultaneously with the transmission. Based on the fact that the reflected signal is received with a time delay, the direct signal frequency is quite different, in every instant, from that of the received one.

On the output of the balanced mixer there arises, following the separation of the undesired components, a differential (beat) signal at a frequency of F_r. The signal from the balanced mixer is then amplified and the evaluator circuit determines the magnitude of this differential, which corresponds to the aircraft's true altitude. Currently, radio altimeters employ various means of frequency measurement employed as evaluator circuits which are manufactured as spectrum analyzers or analogue and digital devices that ensure direct or indirect measuring of the differential frequency. The principle of these evaluator circuits is usually based on counting the pulses shaped from the differential signal throughout the entire period of modulation. The magnitude of the differential frequency, depending on the time delay of the reflected signal, is expressed as:

$$F_r = \frac{8 \Delta f F_M}{c} H \tag{1}$$

where: F_r – differential frequency,

Δf – frequency swing,

F_M – modulation frequency,

H – altitude.

The equation (1) implies that the measured altitude is directly proportional to the differential frequency F_r. Actually, it is not that simple, as both the transmitted and the received signal is frequency modulated, and the modulation signal is the periodic function of time. As is given in [1] and [2], the differential signal spectrum is discrete and consists of only those frequency components that are multiples of the F_M modulation frequency. Consequently, it follows that measuring the differential frequency enables the recording of only those changes in altitude ΔH which ensure generation (termination) of the spectral line within the differential signal spectrum. According to [2], the value of ΔH is given by expression:

$$\Delta H = \frac{c}{8 \Delta f} \tag{2}$$

This means that the data of the altitude indicator at continuous changes of the altitude are not changing continuously but in jumps.

2 Changes in the Pulses Count with the Changing Altitude of Flight

Pulses are generated or terminated at the moment when the envelope of the resulting signal passes through zero. This conclusion can be mathematically illustrated as:

$$U(t_1) = U_v + U_p \cos\left[\varphi_0 + \varphi_M \cos\Omega_M t_1\right] = U_v \tag{3}$$

This holds if:

$$\cos\left(\varphi_0 + \varphi_M \cos\Omega_M t_1\right) = 0 \tag{4}$$

This condition is met under the condition that:

$$\varphi(t) = \varphi_0 + \varphi_M \cos\Omega_M t_1 = (2k-1)\frac{\pi}{2} \tag{5}$$

where: k – arbitrary whole number,

$\varphi(t)$ – phase of a LF differential signal.

From the theoretical description above, one can assume that if the measured altitude changes by a value lower than the ΔH, the data on the radio altimeter do not change at all, as the number of steady pulses for the modulation period remains unchanged. Reality, however, is a bit more complex. The conflict exists in that fact that our way of thinking is about the generation and existence of a single steady pulse at changes in the altitude by a value of ΔH. At changes in altitude within a range smaller than ΔH, one pulse is generated and terminated alternatively.

This alternating generation and termination of a pulse is caused by the unequal phase change of values $\varphi(t)_{max}$ - $\varphi(t)_{min}$, representing the sum and the difference of the initial phase φ_0 and the variable phase φ_M, at changes of the altitude within the interval of ΔH. The phase of the $\varphi(t)$ signal is periodically changing within the interval from the value of:

$$\varphi(t)_{\max} = \varphi_0 + \varphi_M = \frac{4\pi}{\lambda_0}H + \frac{4\pi\xi}{\lambda_0} = \frac{4\pi}{\lambda_0}H(1+\xi) \tag{6}$$

to the value of:

$$\varphi(t)_{\min} = \varphi_0 - \varphi_M = \frac{4\pi}{\lambda_0}H - \frac{4\pi\xi}{\lambda_0} = \frac{4\pi}{\lambda_0}H(1-\xi) \tag{7}$$

Around the mean value of:

$$\varphi(t) = \frac{4\pi}{\lambda_0}H \tag{8}$$

With the modulation frequency of Ω_M.

As pulses are generated or terminated alternatively, when the modulation envelope is passing through zero, and the condition of the pulse generation

$$\frac{4\pi}{\lambda_0} H_V (1+\xi) = (2k-1)\frac{\pi}{2} \tag{9}$$

Similarly, the condition of pulse termination is given by the expression of:

$$\frac{4\pi}{\lambda_0} H_Z (1-\xi) = (2k-1)\frac{\pi}{2} \tag{10}$$

From the above one may conclude that pulse generating altitude H_V is given by the expression:

$$H_V = \frac{\lambda_0}{8}\frac{2k-1}{1+\xi} \tag{11}$$

The pulse termination altitude H_Z is given by the expression:

$$H_Z = \frac{\lambda_0}{8}\frac{2k-1}{1-\xi} \tag{12}$$

Analyzing expressions (12) and (13) one can determine the scope of altitude under which the individual signal may exist. Of the given expressions one can determine that the first pulse is generated at an altitude close to the value of $\lambda/4$ and is terminated on surpassing this altitude, whereas the altitude scope of the existence of this pulse is rather small. With a further increase in altitude, the fist pulse is generated and terminated alternatively, but the altitude interval of pulse duration L is increasing. The interval of duration of individual pulses L is calculated as the difference in the altitudes at which the pulses are generated and terminated:

$$L = H_Z - H_V = \frac{\lambda_0}{8}\frac{2k-1}{1-\xi} - \frac{\lambda_0}{8}\frac{2k-1}{1+\xi} = \frac{\lambda_0}{4}(2k-1)\frac{\xi}{1+\xi^2} \tag{13}$$

The altitude interval of generating the individual pulses V, is calculated as the difference between the altitudes, e.g. between the generation of the first and the second pulse:

$$V = H_{V2} - H_{V1} = \frac{\lambda_0}{8}\frac{2(k+1)-1}{(1+\xi)} - \frac{\lambda_0}{8}\frac{2k-1}{1+\xi} = \frac{\lambda_0}{4(1+\xi)} \tag{14}$$

Where: H_{V1} – altitude of the first pulse generation,

H_{V2} – altitude of the second pulse generation.

The expression (15) makes it apparent that the altitude interval of generating the first and further pulses V is of a constant value, which is only a bit smaller than the value of the $\lambda_0/4$.

The altitude interval of terminating the individual pulses Z is calculated as the difference in the altitudes, e.g. between the termination of the first and the second pulse.

Expression (16) makes it clear that the altitude interval when the first or the further pulses are terminated Z is also of constant value, and is only a bit larger than the value of $\lambda/4$.

If based on expressions (15) and (16) we develop a graphic function expressing dependence of the number of pulses on altitude, it becomes evident that the values of the pulse generating altitude intervals are lower than those of the altitudes terminating pulses. Consequently, we experience growth throughout the duration of the individual pulses with the altitude L, until the dependence of the pulses on the minimal altitude ΔH start to overlap.

For practical purposes of altitude measurement, it is important that at changes of altitude there exists a steady pulse for the modulation period. Therefore, the change in the altitude must remain within the scope of intervals equalling roughly to $\lambda_0/4$, within the pulses generated and terminated alternatively, until the altitude is achieved at which the pulse will exist for good. That altitude corresponds to the critical altitude.

3 The Value of the Critical Altitude ΔH

In expressions (14) and (15) it is apparent that the maximum value of the initial phase φ and that of the variable phase $\varphi_M$is dependent on the altitude H. The instant value of the φ_M phase is, however, changing with time in the rhythm of the modulation frequency $\Omega_{M,}$ even at constant altitude as expressed in (12). The number of pulses N generated in this way during a single modulation period, at constant altitude, will be given by the number of the modulation envelope of the φ_M phase passing through zero. This envelope passes through zero at every π radians.

$$N = \frac{\varphi_M}{\pi} = \frac{\frac{8\pi\xi}{\lambda_0}H}{\pi} = \frac{8\pi}{\lambda_0}H \tag{15}$$

Expression (16) leads us to conclude that the number of pulses during a single modulation period is directly proportional to the altitude at which the aircraft is flying.

Further, we analyze by what value is it necessary to change the altitude of flight to increase the number of pulses during a modulation period by a single steady pulse, which, in the end, increases the value on the altitude indicator.

The concrete number of pulses N_k for a concrete altitude H_k is given by the expression:

$$N_k = \frac{8\xi}{\lambda_0} H_k \tag{16}$$

Suppose that this number of pulses increases by a single steady pulse if the altitude H_k increases by a value of ΔH.

Then we are able to write it down as:

$$N_{k+1} = \frac{8\xi(H_k + \Delta H)}{\lambda_0} \tag{17}$$

The altitude interval ΔH corresponds to the change in the number of pulses by a single pulse, during a single modulation period T_M. This value of ΔH is termed as the critical altitude, representing the minimum change in the altitude recorded by the altimeter. It then follows that if the aircraft changes its value by a magnitude smaller than the value of the ΔH, the radio altimeter data remain unchanged, as the number of pulses for a modulation period have not changed either.

This means that the selected principle of the radio altimeter with frequency modulation only allows us to measure altitude on the basis of discrete changes (jumps). Consequently, the altitude indicator reading continuous changes in the measured altitude are actually not changing continuously, but in rather jumps, and the measured altitude can be read only at a precision of $\pm \Delta H$.

This critical altitude is to be understood as the methodical error of the radio altimeter operating with frequency modulation. Decreasing the critical altitude, i.e. the methodical error of the altimeter with frequency modulator, is possible only through increasing the value of the frequency swing Δf, as implied by expression (16).

Increasing the frequency swing Δf is only possible by simultaneously increasing the mean value of the carrier frequency f_0. Thereby, at each value of the carrier frequency, it is possible to realize a limited frequency swing. These boundary frequency limits allows generation of parasite amplitude modulation, negatively affecting the operation of the radio altimeter.

Within the entire scope of the measured altitude H the pulses are generated and terminated alternatively. As the altitude scope of the duration of pulses L increases with the altitude, at certain altitudes mutual overlapping occurs. The first overlapping of two pulses takes place at an altitude that immediately surpasses the value of ΔH, which is termed as a boundary altitude. The second overlapping of three pulses occurs at an altitude that surpasses the value of $2\Delta H$. Each of the subsequent pulse overlappings occurs at altitudes changing by the value of ΔH. Overlapping of a certain amount of generated and terminated pulses – unsteady pulses "$N_{\Delta H}$" – generate steady pulses. The number of steady pulses N

corresponds to the number of overlappings corresponding to the number of overlappings of generated and terminated pulses. It is the number of steady pulses N that is of importance for altitude measurement.

At altitude scope between 0 as ΔH, an $N_{\Delta H}$ number of unsteady pulses is generated and an $N_{\Delta H}$ -1 number of pulses are terminated instead. The fact that the number of terminated pulses is lower than that of the generated ones enables us to ensure overlapping pulses between the first and the second altitude intervals known as ΔH. The number of generated and terminated pulses $N_{\Delta H}$ within the altitude scope of ΔH can be computed from the equality of the altitude scopes in both cases. The altitude scope of the generating of the individual pulses Vφ is given by expression (17). The altitude scope ΔH achieved by whole number multiples of the numbers of unstable pulses generated ΔH_V is:

$$\Delta H_V = N_{\Delta H} . V \tag{18}$$

The altitude scope of the terminating of the individual pulses Z is given by (19). The altitude scope of ΔH derived by whole number multiples of unstable pulses ΔH_Z will be:

$$\Delta H_Z = (N_{\Delta H} - 1)\ Z \tag{19}$$

Based on the equality of the altitudes, we can write:

$$N_{\Delta H} V = (N_{\Delta H} - 1)\ Z \tag{20}$$

Introducing expressions (18) and (19) into (20) and making some arrangements, we can determine the number of unstable pulses $N_{\Delta H}$ within the scope of the boundary altitude ΔH.

Substituting the concrete values of the relative frequency swing ξ of some of the types of RA we can determine the number of unstable pulses $N_{\Delta H}$ within the altitude scope ΔH. At the same time, the last line of the table comprises the serial number of the overlapping pulse N, where one steady pulse is generated within the next altitude scope ΔH.

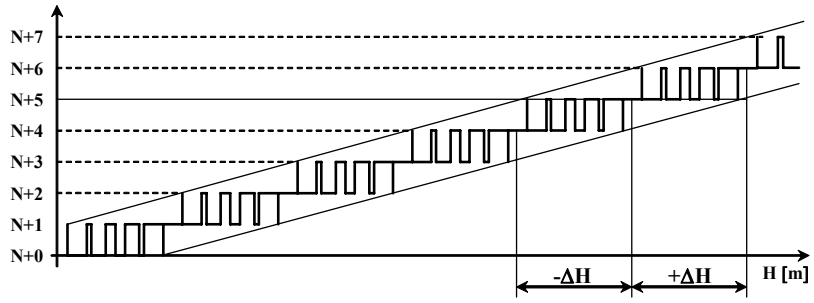


Figure 1
Marking the methodical error of the radio altimeter

Conclusions

Analysis of the change in the number of pulses at measured aircraft altitude changes has led to the development of a mathematical procedure revealing how this law influences the precision of altitude. The law of the generation and termination of altitude pulses has led to the introduction of the notion of "critical altitude", which represents the methodical error of the radio altimeter. On the basis of the analysis presented above, it is possible to quantify the number of unsteady pulses within the scope of critical altitude for all types of radio altimeters currently in use. The knowledge of the laws has led to the drawing up of new methods of evaluating the measured altitude of aircraft and then filing several patents to increase the precision of measuring low altitudes.

References

[1] Adamčík, F.: Avionics Systems Integration of Military Aircraft. In: III. Medzinárodná vedecká konferencia „Nové trendy v rozvoji letectva" (International Scientific Conference, "New Trends in Aviation Development), VLA Košice, 1998, pp. 5-8

[2] Madarász, L., Andoga, R., Főző, L.: Situational Control, Modeling and Diagnostics of Large Scale Systems, In: Towards Intelligent Engineering and Information Technology, ISSN 1860-949X, No. 143 (2009), pp. 153-164

[3] Bréda, R.: Modernization of Avionic Systems on MiG-29A Fighter, In: MOSATT, September 22-24, 2009, Zlata Idka, ISBN 978-80-970202-0-0

[4] Džunda, M.: Accuracy and Resistance of the Radio Communications Systém against Chaotic Impulse Interference. In: MOSATT Košice, 2005

[5] Grabisch, M., Marichal, J. L., Mesiar, R., Pap, E.: Monograph: Aggregation Functions, In: Acta Polytechnica Hungarica, Vol. 6, No. 1, 2009

[6] Györök, G., Makó, M., Lakner, J.: Combinatorics at Electronic Circuit Realization in FPAA, In: Acta Polytechnica Hungarica, Vol. 6, No. 1, 2009

[7] Kováčová I., Kováč, D.: Inductive Coupling of Power Converter's – EMC, In: Acta Polytechnica Hungarica, Vol. 6, No. 2, 2009

[8] Labun, J.: Teória leteckých rádiovýškomerov malých výšok. 1. vyd.. (Theory of aircraft radio-altimeters of low altitudes 1. edition) In: Košice: VLA, 2003. 100 s.

[9] Meziane, S., Toufouti, R., Benalla H.: Nonlinear Control of Induction Machines Using an Extended Kalman Filter, In: Acta Polytechnica Hungarica, Vol. 5, No. 4, 2008

[10] Ochodnický, J.: Sledovanie pohybu vzdušných cieľov v multiradarovej sieti. (Observation of movement of aerial targets in a multi-radar network)

In: Zborník z vojensko-odbornej konferencie PVO (Proceedings of army conference of anti aircraft defense), 25.-26.2.1998, L. Mikuláš, str. 92-99

[11] Olejník, F., Mislivcová, V.: Matematika 1 (Mathematics 1). In: Elfa, Košice 2006, 203 s. ISBN 80-8073-647-2

[12] Sopata, M., Bréda, R.: Možnosti využitia netradičných systémov pri navádzaní nepilotovaných prostriedkov na pristátie. (Possibilities of use of non-traditional systems by guidance of UAV by landing) In: 7. Medzinárodná vedecká konferencia: Nové trendy rozvoja letectva (International scientific conference, "New trends in aviation development), September 6-8, Košice 2006, ISBN 80-8073-520-4

Dynamic Modeling and Simulation of a Switched Reluctance Motor in a Series Hybrid Electric Vehicle

Siavash Sadeghi, Mojtaba Mirsalim

Department of Electrical Engineering
Amirkabir University of Technology
Tehran, Iran
E-mail: sadegs3@rpi.edu, mirsalim@aut.ac.ir

Arash Hassanpour Isfahani

Islamic Azad University, Khomeinishar Branch
Isfahan, Iran
Email: ahassanpour@ieee.org

Abstract: Dynamic behavior analysis of electric motors is required in order to accurately evaluate the performance, energy consumption and pollution level of hybrid electric vehicles. Simulation tools for hybrid electric vehicles are divided into steady state and dynamic models. Tools with steady-state models are useful for system-level analysis whereas tools that utilize dynamic models give in-depth information about the behavior of sublevel components. For the accurate prediction of hybrid electric vehicle performance, dynamic modeling of the motor and other components is necessary. Whereas the switched reluctance machine is well suited for electric and hybrid electric vehicles, due to the simple and rugged construction, low cost, and ability to operate over a wide speed range at constant power, in this paper dynamic performance of the switched reluctance motor for e series hybrid electric vehicles is investigated. For this purpose a switched reluctance motor with its electrical drive is modeld and simulated first, and then the other components of a series hybrid electric vehicle, such as battery, generator, internal combusion engine, and gearbox, are designed and linked with the electric motor. Finally a typical series hybrid electric vehicle is simulated for different drive cycles. The extensive simulation results show the dynamic performance of SRM, battery, fuel consumption, and emissions.

Keywords: Dynamic Modeling, Switched Reluctance Motor, Series Hybrid Electric Vehicle, ADVISOR, Vehicle Performance

1 Introduction

Environmental and economical issues are the major driving forces in the development of electric vehicles (EVs) and hybrid electric vehicles (HEVs) for urban transportation. The selection of the right electric motor is of primary importance to the EVs and HEVs designer [1]. Several types of electrical machines, such as DC, induction, and permanent magnet machines, have been used for HEV applications [1, 2, 3, 4]. These days switched reluctance motors (SRM) have become more attractive for these applications due to the simple and rugged construction, low cost, and ability to operate over a wide speed range at constant power [5].

Several computer programs have since been developed to simulate and analyze HEVs. The advanced vehicle simulator (ADVISOR) is the most widely used computer simulation tool. The structure of ADVISOR makes it ideal for interchanging a variety of components, vehicle configurations, and control strategies. Also, the capability to quickly perform parametric and sensitivity studies for specific vehicles is a unique and invaluable feature of ADVISOR [6, 7]. However, this software only models the steady-state mode, and one cannot see the dynamic performance of electric motors in the simulation. Furthermore, the model for a SRM hasn't been included in the software library of ADVISOR yet [6, 7]. Since the dynamic behavior analysis of electric machines is required in order to accurately evaluate the performance, energy consumption and pollution level of HEVs, it is necessary to model the electric machine dynamically.

Among the various types of HEVs, series hybrid electric vehicles (SHEVs) attract special attention due to their simple structure and uncomplicated energy management strategy. As there is no mechanical connection between the internal combustion engine (ICE) and load, the engine speed and torque is completely independent of vehicle load. Therefore, the engine can run in its optimal efficiency region all the time and minimize engine emissions [8]. In this paper, a co-simulation method is implemented to dynamically model and simulate SHEVs. For this purpose, utilizing MATLAB/SIMULINK software, dynamic modeling of a switched reluctance motor is investigated and simulated separately. In doing so, mechanical components of a SHEV are also called from the software library of ADVISOR and then linked with the electric motor. Finally a typical SHEV is modeled and simulated for different drive cycles using the proposed method. The extensive simulation results show the dynamic performance of SRM, battery, fuel consumption, and emissions for the drive cycle.

2 Vehicle Performance

The first step in vehicle performance modeling is to obtain an equation for calculating the required vehicle propelling force. This force must overcome the road load and accelerate the vehicle. The tractive force, F_{tot}, available from the propulsion system can be written as [9, 10]:

$$F_{tot} = f_{roll} + f_{AD} + f_{grade} + f_{acc} \tag{1}$$

The rolling resistance, aerodynamic drag, and climbing resistance are known as the road load. The rolling resistance is due to the friction of the vehicle tires on the road and can be written as [9, 10]:

$$f_{roll} = f_r \cdot M \cdot g \tag{2}$$

where, M , f_r and g are the vehicle mass, rolling resistance coefficient and gravity acceleration, respectively.

The aerodynamic drag is due to the friction of the vehicle body moving through the air. The formula for this component is [9, 10]:

$$f_{AD} = \frac{1}{2} \cdot \xi \cdot C_D \cdot A \cdot V^2 \tag{3}$$

where, V , ξ , C_D and A are the vehicle speed, air mass density, aerodynamic coefficient and the frontal area of the vehicle, respectively.

The climbing resistance is due to the slope of the road and is expressed by [9, 10]:

$$f_{grade} = M \cdot g \cdot \sin \alpha \tag{4}$$

where, α is the angle of the slope.

If the velocity of the vehicle is changing, then clearly a force will be needed to be applied in addition to the above forces. This force will provide the linear acceleration of the vehicle, and is given by:

$$f_{acc} = M \cdot a = M \cdot \frac{dV}{dt} \tag{5}$$

A typical road load characteristic as a function of vehicle speed and vehicle mass is shown in Fig. 1.

Wheels and axle convert F_{tot} and the vehicle speed to torque and angular speed requirement to the differential as follows:

$$T_{wheel} = F_{tot} \cdot R_{wheel} + I_{wheel} \cdot \frac{d\omega_{wheel}}{dt} + T_{Loss} \tag{6}$$

$$\omega_{wheel} = \frac{V}{R_{wheel}}(s+1) \tag{7}$$

where T_{wheel}, R_{wheel}, I_{wheel}, T_{Loss}, ω_{wheel}, and s are the tractive torque, radius of the wheel, wheel inertia, wheel loss torque, angular velocity of the wheels and wheel slip wheels, respectively.

The torque at the wheels and the wheel angular velocity are converted to engine torque and engine speed requirements using the gears ratio at differential and gearbox as in the following:

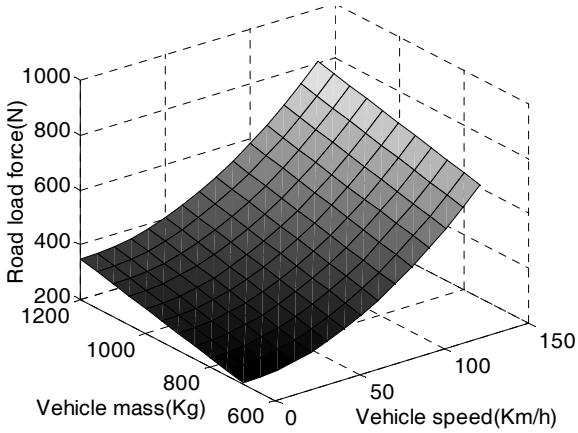


Figure 1
The road load profile as a function of vehicle speed and vehicle mass ($\alpha = 0$ deg.)

$$\begin{aligned} T_{fd\_in} &= T_{gb\_out} = \left(T_{fd\_out} + T_{fd\_Loss} + I_{fd} \cdot \frac{d\omega_{gb}}{dt} \right) \frac{1}{G_{fd}} \\ T_{engine} &= T_{gb\_in} = \left(T_{gb\_out} + T_{gb\_Loss} + I_{gb} \cdot \frac{d\omega_{engine}}{dt} \right) \frac{1}{G_{gb}} \end{aligned} \tag{8}$$

$$\begin{aligned} \omega_{gb} &= G_{fd} \cdot \omega_{wheel} \\ \omega_{engine} &= G_{gb} \cdot \omega_{gb} \end{aligned} \tag{9}$$

where the subscripts fd and gb stand for differential and gear box and T_{fd_in}, $T_{fd\_out}\left(=T_{wheel}\right)$, T_{fd_Loss}, I_{fd}, G_{fd}, T_{engine}, ω_{gb} and ω_{engine}, are the input and output differential torques, differential loss torque, differential inertia, differential gear, engine torque, gearbox and engine angular velocities, respectively.

3 Components Modeling of a Series Hybrid Vehicle

In addition to design studies of overall hybrid power-train systems that tend to use relatively simple models for vehicle components, the authors have looked into the integration of system simulations with high-fidelity component models, such as models of motors, and engines.

The structure of a series hybrid vehicle shown in Fig. 2 consists of two major groups of components:

1) Mechanical components (Wheels and Axle, Transmission box and ICE).
2) Electrical components (battery, electric motor and generator).

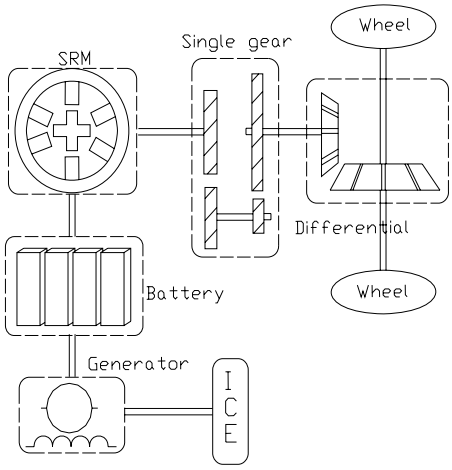


Figure 2
The model for a SHEV

3.1 Transmission Box

In SHEVs the differential box can be eliminated by using two electric motors. Because of the wide speed range operation of electric motors enabled by power electronics control, it is possible for SHEVs to use a single gear ratio for instantaneous matching of the available motor torque (T_m) with the desired tractive torque (T_{wheel}). The single gear ratio and its size will depend on the maximum motor speed, maximum vehicle speed, and the wheel radius as in the following [11]:

$$GR = G_{fd} \cdot G_{gb} = \frac{(\omega_m)_{\max}}{V_{\max}} \cdot R_{wheel} \tag{10}$$

where $(\omega_m)_{\max}$, $V_{\max}$ and GR are maximum motor angular velocity, maximum vehicle speed, and the single gear ratio, respectively.

3.2 Battery

The NiMH type battery is selected for the system and the complex battery model reported in [8] is simplified and used in the simulations. This model can be described by an equivalent circuit which consists of an open circuit voltage, V_{oc}, in series with an internal resistance, R_{int} varying with the state of charge (SOC). The relation between V_{oc}, R_{int}, SOC, and temperature can be easily obtained from the data available from the battery manufacturer, or simple experimental tests. The voltage of the battery tends to drop as the SOC decreases and as the amount of current drawn from the battery increases.

SOC can be calculated from the maximum capacity of the battery as in the following:

$$Soc = \left(Capacity_{Max} - Ah_{used}\right) / Capacity_{Max}$$

$$Ah_{used} = \int \frac{I_{batt}}{3600} \cdot dt \tag{11}$$

where Ah and I_{batt} stand for Ampere-hour and the current, respectively. The terminal voltage of the battery, V_{batt}, can be determined as follows:

$$V_{batt} = V_{oc} - R_{\text{int}} \cdot I_{batt} \tag{12}$$

It is apparent that during the motoring phase of a drive cycle, the electric machine draws current from the battery pack (discharging) and in the regenerative braking mode the machine supplies current to the pack (charging).

3.3 Electric Machine

One important component that can be introduced in greater detail is the electric machine. Electric machines are the key components of SHEVs. Some motor types have inherent property of operation with large extended speeds. The direct current motor is an example, but it has low efficiency, a high need of maintenance and low reliability due to the presence of the commutators. An induction motor (IM) can achieve a large speed range with field-oriented control. However, due to nonlinearity of the dynamic IM model and the dependency on motor parameters, the control is complex [5]. A permanent magnet motor (PM) has the highest torque density and therefore will potentially have the lowest weight for given torque and power ratings [10, 12]. However, the fixed amount of flux for magnets

limits its extended speed range. An SRM has certain advantages, such as rugged construction, fault-tolerant operation, and low cost, which makes it well suited with many applications such as vehicle propulsion. Design and simulation shows a wide speed range of 5.7 (and 6.7 times the based speed), much higher than other kinds of electric motors [10]. Therefore, it can be considered an excellent candidate for SHEVs. In this paper the dynamic model of an SRM is investigated in a series hybrid vehicle.

The model of an SRM is described by its flux linkage and torque characteristics. Some approaches have been used to obtain these characteristics from finite element (FE) analysis to complex nonlinear magnetic circuit models [13]. Since FE analysis is time consuming, in this paper an analytical model is used to model the flux linkage and torque characteristics. In the analytical method, the flux linkage is described with an analytical equation in two positions: unaligned position and poles overlap position. Then, the current can be calculated as in the following [13]:

$$V_{ph} = \frac{d\lambda_{ph}}{dt} + i_{ph}R_{ph} \tag{13}$$

$$\frac{d\lambda_{ph}}{dt} = \frac{\partial\lambda_{ph}}{\partial i_{ph}} \cdot \frac{di_{ph}}{dt} + \frac{\partial\lambda_{ph}}{\partial\theta} \cdot \frac{d\theta}{dt} \tag{14}$$

$$\frac{di_{ph}}{dt} = \frac{1}{\partial\lambda_{ph}/\partial i_{ph}} \left(V_{ph} - i_{ph}R_{ph} - \frac{\partial\lambda_{ph}}{\partial\theta}\omega \right) \tag{15}$$

where:

$$\frac{\partial\lambda_{ph}}{\partial\theta} \cdot \frac{d\theta}{dt} = \omega \cdot \frac{\partial\lambda_{ph}}{\partial\theta} = BEMF$$

The phase torque in SRM is obtained from the co-energy method which can be calculated from the phase flux linkage as follows:

$$T = \frac{\partial W'(i_{ph},\theta)}{\partial\theta} = \frac{\partial}{\partial\theta}\int_0^i \lambda_{ph}(i,\theta)di \tag{16}$$

The total electromagnetic torque is obtained by:

$$T_{total} = \sum_{j=1}^{q} T_{phase}(i_j,\theta_j) \tag{17}$$

where q is number of phases.

Fig. 3 illustrates the phase static torque versus different phase currents and rotor positions for an SRM with parameters given in Tables I-II.

Table I
SRM Specifications

Number of turns per pole	30
Rotor pole width	31^o
Number of pole windings in series	2
Number of pole windings in parallel	1
Peak current	400 A
Stack length	4.26 in
Air gap	0.01 in
Radius to the air gap at the rotor	3.895 in
Radius to the outside rotor yoke	2.42 in
Radius to the inside stator yoke	4.845 in
Stacking factor	0.9
Stator phase resistance	0.0042 ohms
Relative permeability of magnetic material	3300
Saturation flux density	1.9 T

TABLE II
Vehicle Specifications

Vehicle maximum speed	44.4 m/s
Vehicle total mass	1020 kg
Air drag coefficient	0.3
Rolling resistance coefficient	0.008
Frontal area	1.965 m^2
Wheel radius	0.305 m
Height of CG,h_g	0.57 m
Power	40 Kw
voltage	240 V
Stator outer diameter	13.58 in;
Rotor outer diameter	7.4694 in
Stack length	7.4694 in
Air gap	0.0373 in
Shaft diameter	2.8135
Rotor pole width	31 degree
M19 steel	

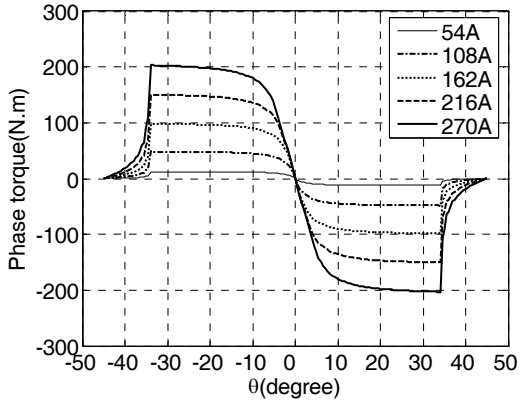


Figure 3
Phase static torque versus rotor position at different phase currents

The most flexible and versatile four-quadrant SRM converter is the bridge converter as shown in Fig. 4, which requires two switches and two diodes per phase. The switches and diodes must be rated to withstand the supply voltage plus any transient overload. Here in the circuit, an IGBT switch is used because of its high input impedance as in a MOSFET, and low conduction losses similar to BJTs [14]. This circuit is capable of operating the machine as a motor and a generator.

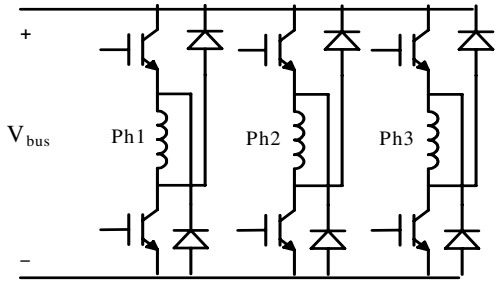


Figure 4
Converter topology

3.4 SRM Controller

The speed controller for an SRM is the same as for other electric motors and can be shown with a PI controller. However, the torque control is important for this motor and is different from other electric motors. In the torque control of a switched reluctance machine, the torque command is executed by regulating the current, as shown in Fig. 5. The control method alters for different speeds of the motor. Below the base speed, the torque is controlled by the pulse width

modulation (PWM) control of current where the phase voltage is switched between zero and Vbus (soft-chopping). Above the base speed with constant power, due to the high back electromotive force (BEMF), which cannot be field weakened, PWM control of current is not possible. Fig. 6 shows the BEMF of the SRM for different speeds. Operation in constant power is made possible in this motor by the phase advancing of the stator phase current, until overlapping between the successive phases occurs [15, 16]. Therefore, the torque of the SRM is controlled with current chopping or PWM control of current under the base speed, and with the phase turn-on and turn-off angles above the base speed. The control of the SRM in generating mode is identical to motoring mode. The only difference is the timing of current pulses. In motoring mode, the windings are excited during the rise of inductance (approaching of poles), while in generating mode each winding is excited when the poles separate from each other. At lower speeds, the current pulses are controlled by chopping the voltage between zero and -Vbus with fixed commutation angles. At higher speeds the generator is operated in single-pulse mode [15, 16].

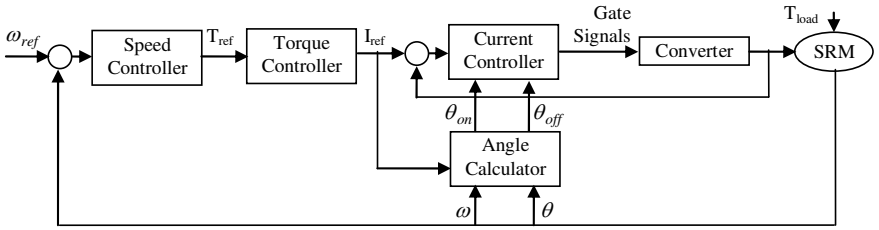


Figure 5
SRM Controller drive

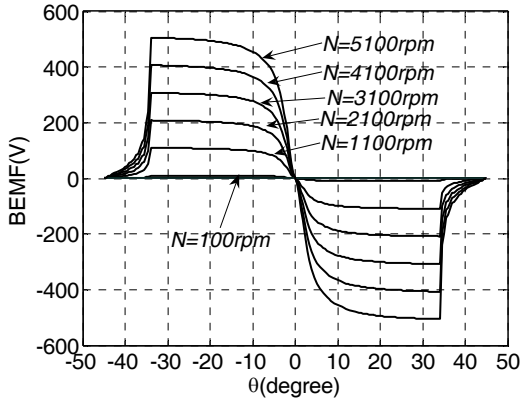


Figure 6
BEMF voltage of an SRM versus rotor position at different speeds (I_{ph}=150A)

Figs. 7 and 8 show the SRM characteristics in low speed motoring and generating modes, respectively. Since the speed is low, control of the motor is done by chopping the phase current around the reference current. The power supply current is positive and shows that the motor is absorbing energy from the battery in an electric propulsion system. In the generating mode, the power supply current is negative; therefore the machine is supplying energy to charge the battery. It is seen from Fig. 9 that the SRM current cannot be chopped completely due to the high back EMF at high speeds, and thus, control of the motor is done by the phase advancing of the stator phase current. Fig. 10 illustrates the SRM torque and power versus motor speed.

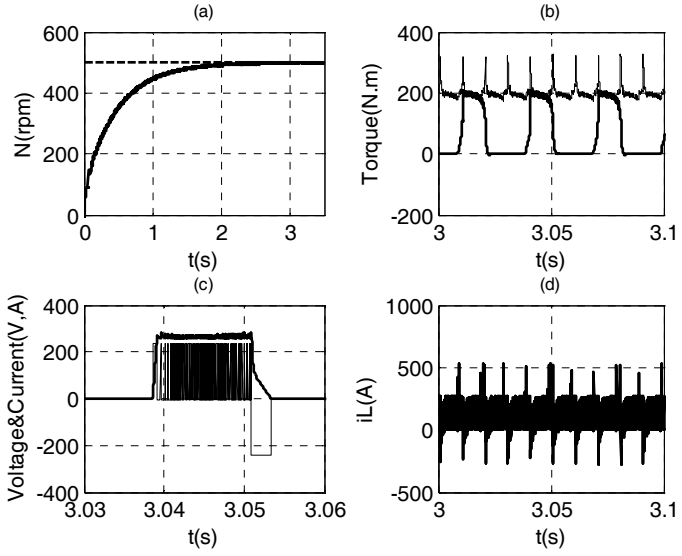


Figure 7
SRM characteristics for $\omega_{ref} = 500rpm, T_{load} = 200N.m$

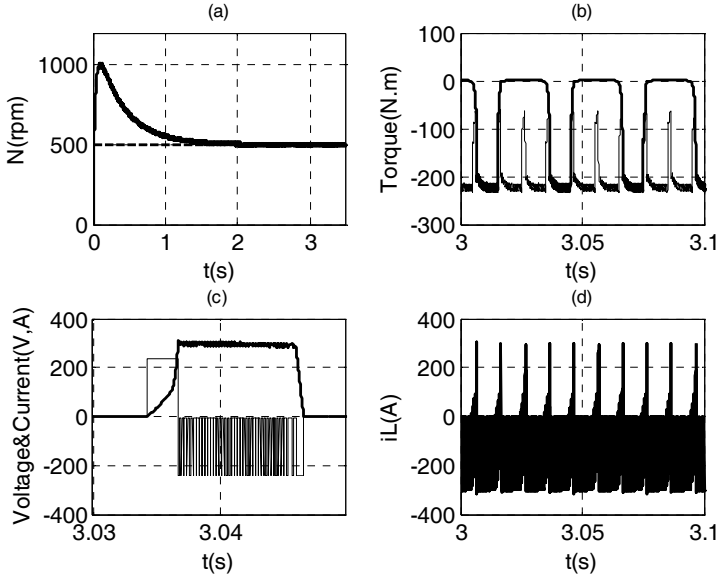


Figure 8
SRM characteristics for $\omega_{ref} = 500rpm, T_{load} = -200N.m$

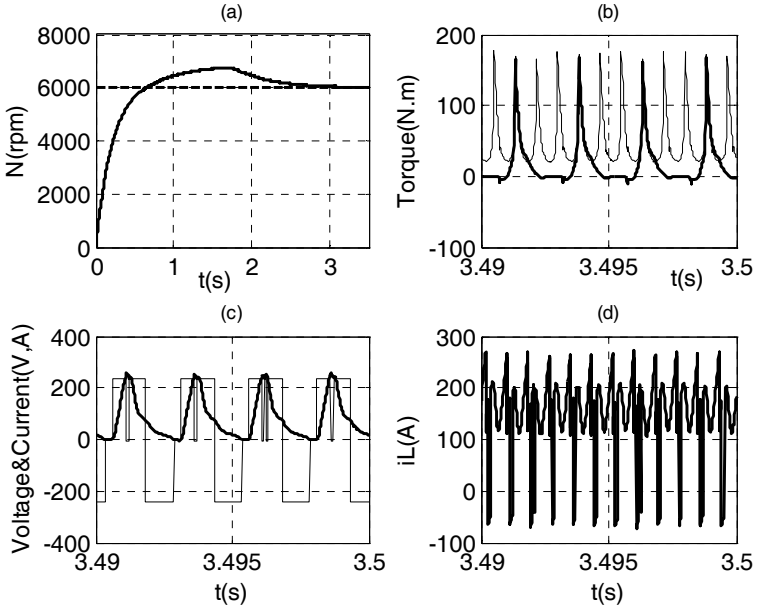


Figure 9
SRM characteristics for $\omega_{ref} = 6000rpm, T_{load} = 50N.m$

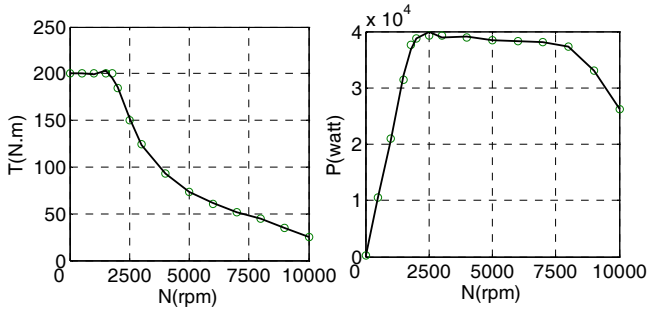


Figure 10
SRM torque and power versus speed

3.5 Generator

The model for the generator is chosen from the library in ADVISOR. This model is unidirectional; i.e. it has one input (the applied torque to the shaft and the rpm), and one output (electrical power output from the generator). From the characteristics of efficiency versus revolution per minute (rpm) and torque, and by the use of a look-up table, the model is able to calculate the output power of the generator and the energy losses at any instant.

3.6 ICE

The model of the internal combustion engine is also chosen from the ADVISOR. The input and output of the model are the desired torque and speed and the developed torque and speed, respectively. If the developed torque and speed are within the operating limits and are obtainable at the specified time step, then the response values are exactly equal to the desired values. In addition to dynamic calculations, the model for the ICE can predict important quantities such as fuel consumption and the amounts of air pollutants in the simulation. This model includes the effects of the inertia of the ICE on operation and the operational boundaries of the engine, and also the effects of temperature on fuel consumption and the pollutants. The modeling of an ICE is complicated, and by the use of engine maps the complexity is reduced to some extant.

The governing output torque and speed equations, or in other words, the desired quantities for the gearbox, are written as in the following:

$$\begin{aligned} T_{engine\_out} &= T_{MAP} - I_{engine} \cdot \alpha_{engine} \\ \omega_{engine\_out} &= \omega_{MAP} \end{aligned} \tag{18}$$

where, α_{engine}, I_{engine}, T_{MAP} and ω_{MAP} are the acceleration and the inertia of the ICE, and the torque and speed determined from the engine maps, respectively.

The fuel consumption and the amount of pollutants are stored in tables indexed by fuel converter speed and torque. The primary output of the exhaust system model is the tailpipe emissions (HC, CO, NOx, and PM) in g/s, as a function of time. Other outputs include the temperature of various exhaust system components and of the exhaust gas temperature into and out of each system component.

4 Linking between the Components and the Control method of SHEVs

The integration of the models of different electrical and mechanical components, where each block consists of two groups of inputs and two groups of the desired outputs, is shown in Fig. 11. One can see that the modeling of each block is a feed-backward.

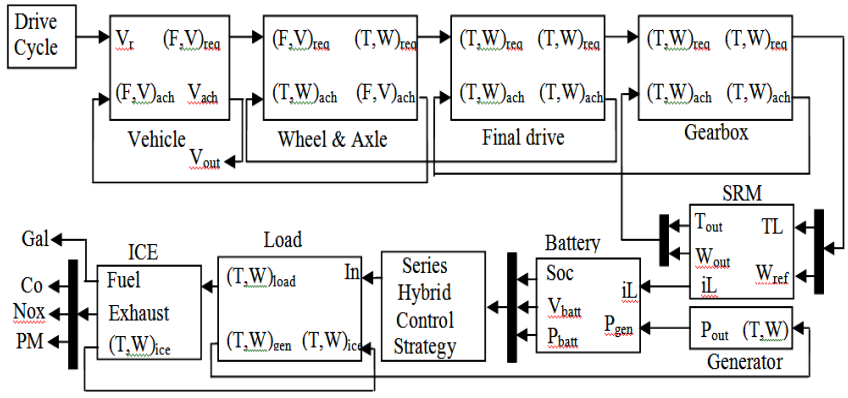


Figure 11
SHEV Model

As for the control strategy of a hybrid vehicle, once again the ADVISOR software is implemented. The main strategy of the control system in a series hybrid is to keep the level of charge of the batteries at a specified range, while charging the batteries at maximum efficiency. In this research, the following control scheme is implemented:

- If the SOC of batteries goes below a specified level, the ICE is turned on to supply energy to the motor and the batteries via an electric generator.
- If the ICE is already turned on, it must stay on till the SOC reaches a high specified level, and as soon the level is reached, the ICE is turned off.
- While staying turned on, the ICE must operate at its maximum efficiency point (best operating point).

5 Simulation Results

Simulation results presented here focus on the dynamic behavior of SRM and the battery of the vehicle. The system is simulated for a 12% road slope on the ECE+EUDC test cycle. The cycle is used for emission certification of light duty vehicles in Europe. Due to dynamic modeling of the electric motor in SIMULINK, the mentioned drive cycle requires about 25 hours of simulation time on a computer with a Pentium 4 processor and 512 MB of memory. The data for the SRM and EV is in the Appendix. The SRM used in the simulation is rated at 240 V and 40 kW. Now, the required number of modules of 12- volt NiMH batteries with the V_{oc}, and R_{int} values shown versus SOC in Fig.12 would be twenty.

Since the SRM used in this simulation has a maximum speed of 8000 rpm, the single gear ratio can be calculated from (10) as follows:

$$GR = \frac{(\omega_m)_{\max}}{V_{\max}} \cdot r_{wheel} = \frac{\left(8000 \cdot \frac{2\pi}{60}\right)}{44.4} \cdot 0.305 \approx 5.75$$

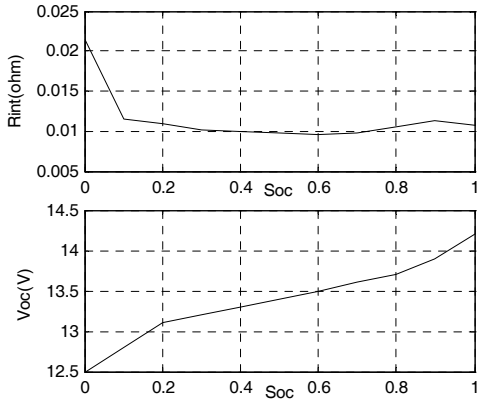


Figure 12
Voc and *Rint* versus SOC

The dynamic behaviors of the SHEV in the ECE+EUDC+12% drive cycle (Fig. 13) are shown in Figs. 14 to 20. The SRM reference and real speeds are shown in Figs. 16 and 17. The SRM load torque and developed torque are also shown in Figs. 16 and 17. It can be seen from Fig. 15 that the torque ripple has no significant effect on the motor speed. The vehicle reference and real speeds are depicted in Fig. 18. Figs. 19 and 20 show the performance of the battery, i.e., the current versus time. The Figure shows the negative SRM current which can be returned to the battery and which therefore increases the overall efficiency of the vehicle.

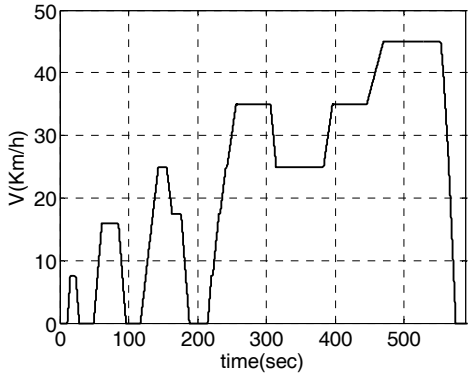


Figure 13
ECE-EUDC Drive cycle (with 12% slope)

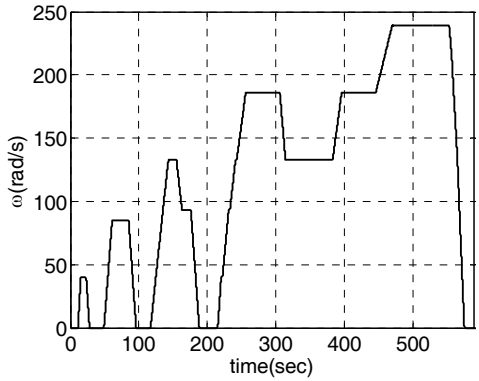


Figure 14
SRM reference speed

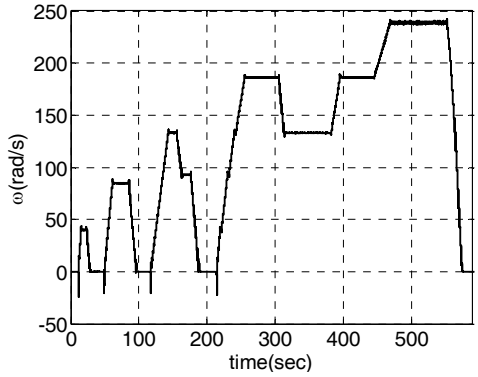


Figure 15
SRM actual speed

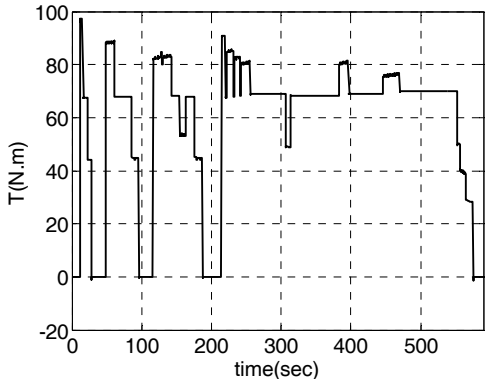


Figure 16
SRM load torque

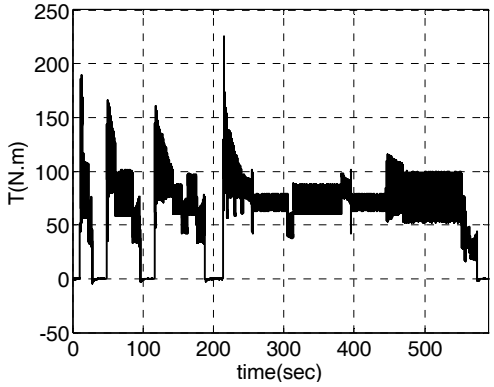


Figure 17
SRM torque

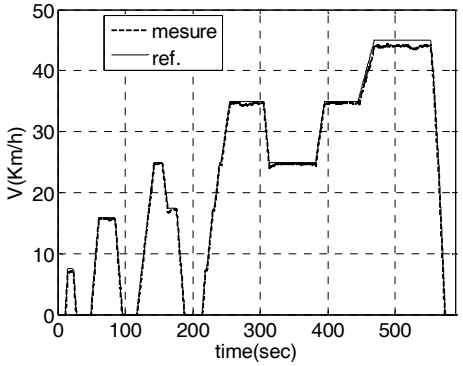


Figure 18
Vehicle speed and drive cycle

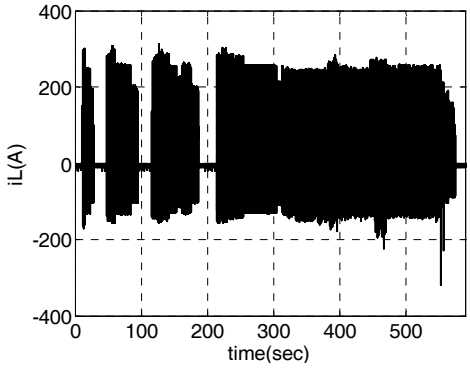


Figure 19
Battery current

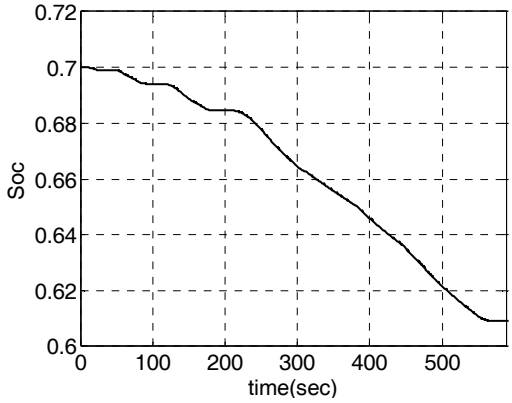


Figure 20
State of charge

Fuel consumption, the state of charge of batteries, and the level of exhaust emissions for the mentioned drive cycle are investigated in Figs. 21 to 24. In these figures, the drive cycle repeats for 15 times and the state of battery charge is kept between 0.4 and 0.7. The rate of pollution caused by ICE is high at the beginning of ICE work period because the engine is still cold. However, the amounts of pollutants decrease as the motor becomes warm enough. The total amount of exhaust emissions for 15 repetitions of the cycle is shown in Fig. 23. The variation of output voltage of battery is as seen in Fig. 24.

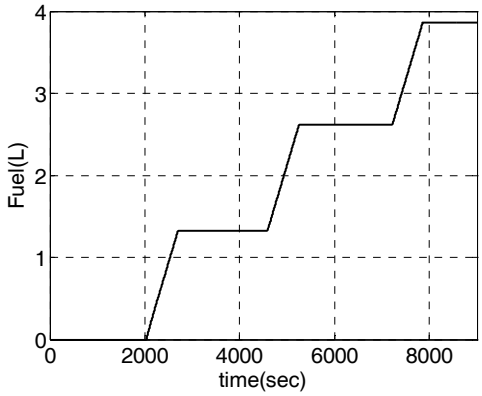


Figure 21
Fuel consumption

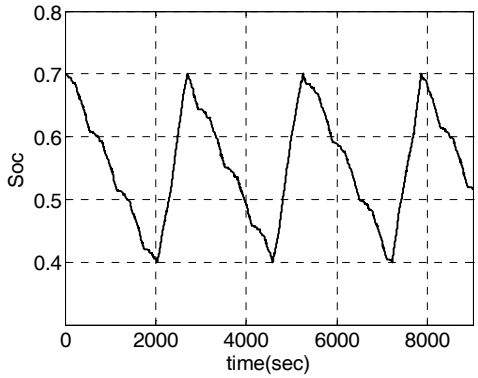


Figure 22
State of charge

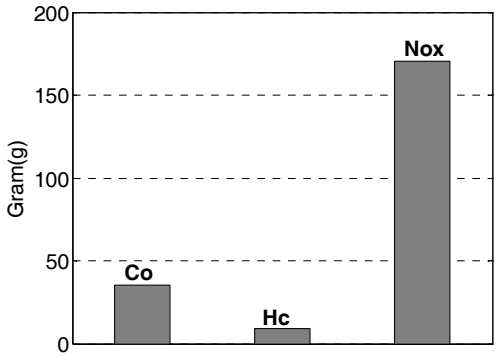


Figure 23
Exhaust output

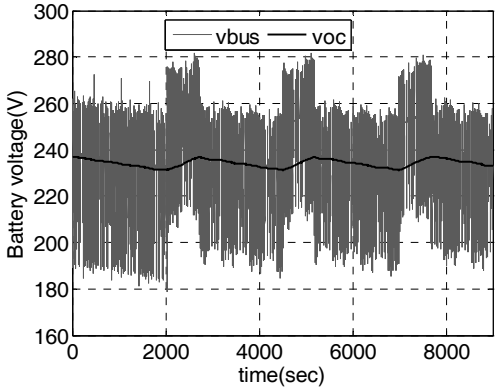


Figure 24
Battery voltage

Conclusions

A co-simulation method has been implemented to dynamically model and simulate the performance of a switch reluctance motor integrated into SHEVs. For this purpose, utilizing MATLAB/SIMULINK software, a dynamic model of a switched reluctance motor has been implemented. A close loop control system with PI controller is employed to control the speed of the motor. The mechanical components of a SHEV are then called from the software library of ADVISOR and then linked with the switch reluctance motor drive in MATLAB/SIMULINK. To evaluate the dynamic performance of switch reluctance motor drive, a typical SHEV has been modeled and simulated for an ECE+EUDC drive cycle with 12% of road slope. The extensive simulation results show the proper dynamic performance of SRM, battery, fuel consumption, and emissions for drive cycles.

References

[1] M. Ehsani, Yimin Gao, S. Gay, "Characterization of Electric Motor Drives for Traction Applications," in proc. 29th IEEE Industrial Electronics Society Annual Conference, IECON'03, Nov. 2-6, 2003, Volume 1, pp. 891- 896

[2] M. Zeraoulia et al, "Electric Motor Drive Selection Issues for HEV Propulsion Systems: A Comparative Study," IEEE Trans. Vehicular Tech., Vol. 55, pp. 1756-1763, Nov. 2006

[3] L. Chang, "Comparison of AC Drives for Electric Vehicles - A report on Experts' Opinion Survey," IEEE AES Systems Magz. pp. 7-10, Aug. 1994

[4] Z. Preitl, P. Bauer, J. Bokor, "Cascade Control Solution for Traction Motor for Hybrid Electric Vehicles", Acta Polytechnica Hungarica, Journal of Applied Sciences at Budapest Tech, Hungary, Vol. 4, Issue 3, pp. 75-88, 2007

[5] K. M. Rahman, B. Fahimi, G. Suresh, A. V. Rajarathnam, M. Ehsani, "Advantages of Switched Reluctance Motor Applications to EV and HEV: Design and Control Issues", IEEE Trans. on Industry Applications, Vol. 36, No. 1, Jan./Feb., 2000, pp. 111-121

[6] T. Markel, A. Brooker, T. Hendricks, V. Johnson, K. Kelly, B. Kramer, M. O' Keefe, S. Sprik, K. Wipke, "ADVISOR: a Systems Analysis Tool for Advanced Vehicle Modeling", ELSEVIER Journal of Power Sources, Vol. 110, pp. 255-266, 2002

[7] M. Amrhein, P. T. Krein," Dynamic Simulation for Analysis of Hybrid Electric Vehicle System and Subsystem Interactions, Including Power Electronics," IEEE Trans. on Vehicular Technology, Vol. 54, No. 3, May 2005

[8] S. Barsali, C. Miulli, A. Possenti, "A Control Strategy to Minimize Fuel Consumption of Series Hybrid Electric Vehicles", IEEE Trans. on Energy Conversion, Vol. 19, No. 1, March 2004

[9] James Larminie, John Lowry, Electric Vehicle Technology Explained, John Wiley, England, 2003

[10] M. Ehsani, Y. Gao, S. Gay, A. Emadi, Modern Electric, Hybrid Electric, and Fuel Cell Vehicles, CRC press, USA, 2005

[11] K. M. Rahman, H. A. Toliyat, M. Ehsani, "Propulsion System Design of Electric and Hybrid Vehicles," IEEE Trans. on Industrial Electronics, Vol. 44, No. 1, pp. 19-27, Feb. 1997

[12] C. C. Chen, K. T. Chau, Modern Electric Vehicle Technology, Published by the US by Oxford University Press, Inc., New York, 2001, pp. 122-133

[13] A. V. Radun, "Design Considerations for the Switched Reluctance Motor," IEEE Trans. on Industry Applications, Vol. 31, No. 5, Sep. 1995

[14] I. Husain, M. S. Islam, "Design, Modeling and Simulation of an Electric Vehicle System," SAE Technical Paper Series. 1999-01-1149

[15] Do-Hyun Jang, "The Converter Topology with Half Bridge Inverter for Switched Reluctance Motor Drive," ISIE Conference, Pusan, Korea, 2001

[16] T. J. E. Miller, Switched Reluctance Motors and their Control. Lebanon, Ohio: Magna Physics Publishing, 1993

[17] A. V. Radun, "Generating with the Switched Reluctance Motor," Proceedings of the Ninth Annual Applied Power Electronics Conference and Exposition, Vol. 1, p. 41, 1994

Single Input Operators of the DF KPI System

Norbert Ádám

Faculty of Electrical Engineering and Informatics
Technical University of Košice
Letná 9, 042 00 Košice, Slovak Republic
norbert.adam@tuke.sk

Abstract: The DF-KPI system is a multiprocessor system based on the dynamic data flow computing model, where each dyiadic instruction requires dynamic matching of its operands. This process is associated with the processing units of the DF-KPI system. The processing units are designed as a dynamic multi-function pipelined unit with LOAD, FETCH, OPERATE, MATCHING and COPY segments. In this article, an architecture design at a logical level of the dynamic multi-function pipelined unit, which handles processing of operand matching for single input data flow operators is suggested.

Keywords: data flow, operand matching, data flow graph, single input data flow operators

1 Introduction

One of the solutions to the problem of reaching higher computer system performance is the concept of high performance parallel computer systems' architecture. Computer systems based on the von Neumann computer principle try to solve these requirements by increasing the performance of particular computer segments. However, possibilities of this kind of speed-up are conditioned by the technological potential [5], [6], [7], [12]. Better results in increasing computer system performance can be achieved by computers based on the data flow computing model. In contrast to control flow architectures, such as the von Neumann model, data flow architectures use the availability of data to fetch instructions, rather than the availability of instructions to fetch data [2]. The data flow computation model enables us to use the program's natural parallelism, which, in turn, shortens the time required to perform a calculation. The advantage of data flow computing (or computers) is that it allows the detection of parallelism at the level of machine instructions, just like the application of data-level parallelism. The disadvantage is the complicated operand-matching control mechanism, increased communication between processing units and the relatively large program storage requirements. But it is possible to bypass all disadvantages of this architecture and get huge computing capacity systems by means of

appropriate architecture design, for example: reconfigurable chip area for processing units, interconnection networks, as are the hypercube, pyramid etc., and by elimination of redundant calculations at the level of program to data flow graph translation, making use of an efficient computation management micro-program with a well designed operand matching algorithm.

2 Data Flow Computers

In the field of the development of high performance new generation computers, architectures based on the data flow computation model are in a parallel computer class of their own. Computational process control in the data flow computation model is implemented by the stream of operands (data) prepared to execute program instructions.

Even though theories exist about data flow models [2], [3], [8], [11], many architectures have been proposed [2], [10], [11]. These can be classified as static and dynamic (or tagged-token) architectures according to the related approach to the model.

The characteristic property of data flow computers is that the data flow program instructions are waiting passively for the arrival of a certain combination of arguments, which are made available as a data control stream in a data-driven sense [3]. The waiting interval, where a program instruction waits for the arrival of operands, represents its selection phase, during which the allocation of computing resources occurs. The fundamental idea behind the data flow computational model is the mapping of tasks (Fig. 1) to the computing elements, which can increase the rate of parallelism. In general, it is necessary to decompose the computational process into smaller communicating processes represented by the data flow program [4].

The data flow program, which uses the data flow computational model, is represented by its machine representation, called a data flow graph. Data flow graphs are directed graphs that show the data dependencies between the respective instructions (operators). Their nodes represent instructions and the arcs connecting these nodes represent the operands of these instructions. An instruction can be executed if all the tokens on the incoming arcs are available: that is, if all the operands of the instruction are available.

The implementation of the data flow computer architecture depends on the form of execution of the data flow program instructions, which is implemented as a process of receiving, processing and transmission of activation symbols (data tokens).

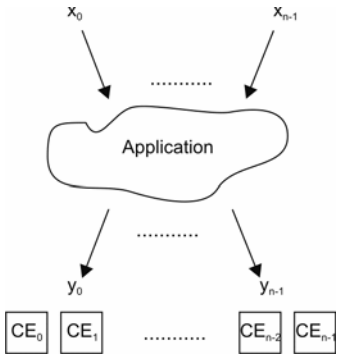


Figure 1
Tasks mapping to the computing elements of the computer system

The following are the types of pure data flow architectures, depending on the concept of processing of data tokens or on the scope of architectural support for its implementation:

- Static models (Figure 2)
- Dynamic models (Figure 3)

The static data flow model was proposed by Dennis and his research group at MIT [3]. In the static approach a node (instruction) can be executed only when all of the tokens are available on its input arcs and no tokens exist on any of its output arcs [8]. This model is capable of using data parallelism and pipelining techniques; therefore this model has found use in applications with regular computing structures. The general organization of the static data flow machine is depicted in Fig. 2.

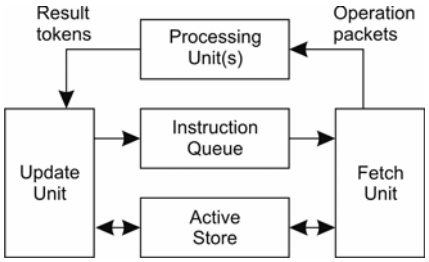


Figure 2
The general organization of the static data flow model

The *Active Store* contains the activation frames of instructions, representing the nodes in data flow graph. Each activation frame contains the operation code, slots for the operands and the destination address. An operand slot contains also a presence (check) bit to determine the availability of the operand.

The *Update Unit* performs the update of the activation tokens and checks whether the instruction is executable. If the feasibility condition is fulfilled, the unit sends the instruction through the instruction front to the instruction fetch unit.

The *Fetch Unit* fetches and sends a complete operation packet containing the corresponding operation code, data, and destination list to the Processing Unit and also clears the presence bits.

The *Processing Unit* performs the operation, forms the result packets and sends them to the Update Unit.

The dynamic data flow model was proposed by Arvind at MIT [1] and by Gurd and Watson at the University of Manchester [4]. The operator represented by a node is executable in the dynamic data flow model (Fig. 3) if all the input edges contain tokens, the symbols of which are identical. Each edge may contain more than one labeled token in this model. When executing a node, the tokens belonging together are removed from the input edges and a token with the corresponding symbol of the output edge is generated. The dynamic dataflow model uses both loop parallelism and recursive parallelism, which appear dynamically during program run-time. Such an architecture must support the process of operand matching (merging).

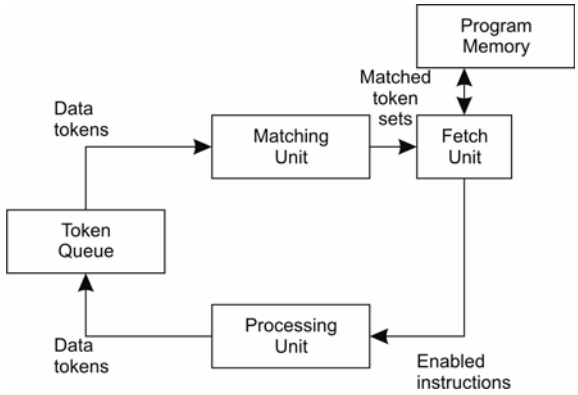


Figure 3
The general organization of the dynamic data flow model

The *Matching Unit* represents a memory which performs matching of tokens at the operator's input. The tokens are merged only if the token is destined for a double input operator, which initiates the operand matching process. If a match exists, the corresponding token is extracted from the Token Queue and the matched token set is passed on to the Fetch Unit. If no match is found, the token is stored in the Matching Unit and waits for a a partner.

The *Fetch Unit* performs a selection of the prepared instructions from the Program Memory and generates executable packages for the Processing Unit.

The *Program memory* is the instruction memory of the data flow program.

The *Processing Unit* processes operations defined in the executable packages and produces result tokens, then sends them to the Matching Unit via the Token Queue.

The *Token Queue* transports the input of tokens from the Processing Unit to the Matching Unit.

3 The DF KPI System

The DF-KPI system [6], being developed at the Department of Computers and Informatics at the Faculty of Electrical Engineering and Informatics of the Technical University of Košice, has been designed as a dynamic system with direct operand matching. The combination of the local control flow model (von Neumann's principle) with the global data flow model allows us to effectively organize the parallel implementation of functional program. The architecture model of the DF-KPI computer is a part of a data flow complex system, which includes support components for data flow computing environment for the implementation of the defined application targets.

The structural organization (Fig. 4) of the DF-KPI computer architecture model consists of the following components:

Coordinating Processors (CP) are intended to manage, coordinate and process instructions of the data flow program, based on the presence of their operands, which are enabled at the CP.DI input port of the coordinating processor - either from its CP.DO output port or from the CP.DO output ports of other CPs through an interconnection network, or from a Data Queue Unit and from the Frame Store. The structure of the CP is a dynamic pipelined multiple-function system, composed of LOAD, FETCH, OPERATE, MATCHING and COPY segments.

The *Data Queue Unit* (DQU) is a unit designed to store the activation symbols (data tokens), which represent operands waiting for matching during program execution.

The *Instruction Store* (IS) is a memory of instructions of the data flow program, in the form of a proper data flow graph.

The *Frame Store* (FS) is a memory of matching (pairing) vectors, by means of which the CP detects the presence of operands to perform the operation defined by the operator (node) in the data flow graph. The short description of the item format of MV matching vector in the FS is <FS>::= <AF><V>, where AF is a flag of the operand's presence (Affiliation Flag) and V is the value of the given operand.

Supporting components of the data flow system are needed to create a realistic computing environment. In the given architecture they are formed by the following:

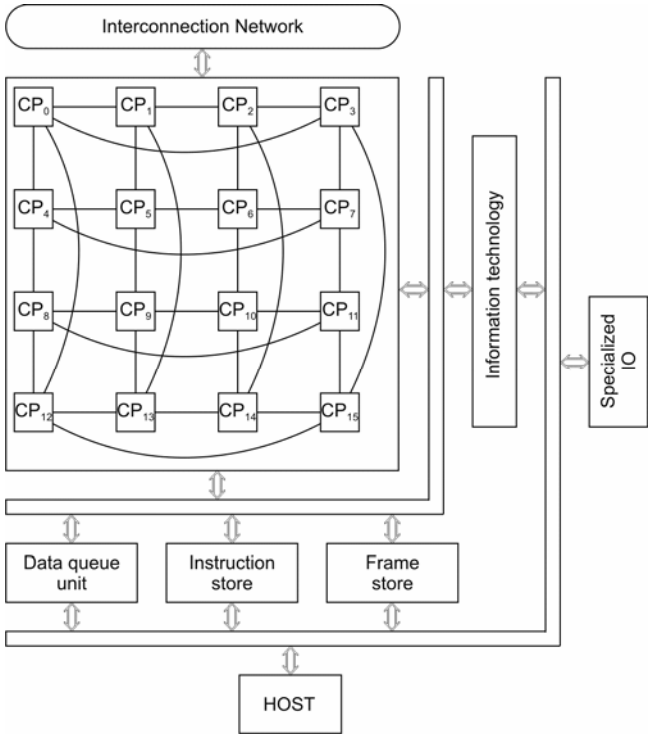


Figure 4
The DF-KPI system

The *Main computer* (HOST) provides standard functions of the computer system during data flow computing process.

The *Information Technology* unit is a unit used to create dedicated application environments (virtual reality, diagnostics, e-learning).

The *I / O processors* for fast direct inputs/outputs into the data flow module (standard I/Os are implemented by the main computer).

3.1 Instruction Format

Each instruction is represented by an operator of the data flow program. Data flow operators are generally classified as single input, double input and N ($N \geq 3$) input operators. In the case of single input operators, the input is formed by one operand, which meets the condition of the data availability firing rule, meaning

operand matching is not necessary. In the case of N input operators, the operator will be executed if all input operands are present, which are then merged in matching segment, based on the matching vector MV. The instruction set of the DF-KPI architecture consists of instructions represented by the single input operators (ACCEPT, IF, KILL, OUT, RET, SEL, UN_OP), double input (BIN_OP, CASE, DEF, GATE, LOAD, SEND, TUP) and N ($N \geq 3$) input (APPLY, CONSTR). The ACCEPT operator is an entry operator of the data flow program and subprograms. The IF operator represents a two-way branch and the CASE operator represents a multi-way switch. The KILL operator is used for consumption of input operand without any response. The OUT operator is the last operator and RET is a return operator of the data flow program or subprograms. The SEL operator selects data from the data structure defined by the CONSTR operator. The UN_OP single input operator and the BIN_OP double input operator represent unary and binary arithmetic operations. The DEF operator defines the program constant. While LOAD creates copies of the data and address part of the input operand, the TUP operator creates copies only of the data section of a data token. Running the program or subprogram is launched by the APPLY operator. A more detailed description of these operators is in [6].

The format of the data flow instructions is as follows:

$$<\text{DFI}>::= <\text{OC}> <\text{LI}> <\{\text{DST}, [\text{IX}]\}^{n}>$$

where OC is the operation code; LI is a literal (e.g. number of copies of the result); DST represents the target address for operation result; IX is a matching index for the operations.

The data flow program instruction represented by a data token is stored in the Instruction Store at the address defined by DST field. The data token has the following format:

$$<\text{DT}>:: = <\text{P}> <\text{T, V}> <\text{MVB}> <\{\text{DST}, [\text{IX}]\}>$$

where P is the priority of the data token; T represents the data type of operand with a value V; MVB defines a base address of matching vector in the Frame Store and DST specifies a destination address of the resulting DT data token. The structure of the DST field is the following:

$$<\text{DST}>:: = <\text{MF}> <\text{IP}> <\text{ADR}>$$

where MF is a matching function, with a defined set of labels {M, B}, M stands for matching (of two DTs), B stands for bypass (without DT matching); IP defines an input port {L(eft), R(ight)}; ADR is the address of the operator or function.

If the operands enter the two-input or multi-input operators, operand matching occurs. The DF-KPI architecture uses the direct operand matching control mechanism. It is based on the allocation of a Matching Vector in the Frame Store according to the activation code (procedure, call). Allocated Matching Vectors are

represented as a matching record in the Frame Store. The format of the Matching Vector in the Frame Store is as follows:

$$<SS\ [B_{ACT} + H + IX + 1]> :: =$$

$$<RC, MVS> <B_{OLD}> < DST_{RET}> <D\{[B_{NEW}]\ \{D\}\} >$$

where B_{ACT} is a pointer to the current top record; H is the size of a header of record; MVS defines the size of a matching vector; RC is the reference counter; B_{OLD} is a pointer to the previous token; DST_{RET} specifies the return address; B_{NEW} defines the base address for new matching record and D represents an operand value.

The RC field is set according to the size of the matching vector at compile-time. After the function associated with the operator has fired, the value of RC is decremented. If RC = 0, the Matching Vector in the frame store is released.

3.2 Operand Matching

As mentioned above, data tokens convey information about the status of the calculation. Their location at the input edges of a node (operators) in the data flow graph means the presence of the operand for a defined and enforceable instruction. The presence of a data token at the output edge of a node reflects the presence of the outcome of operations performed by the operator defined instructions.

The processing of a data flow graph is subject to the rules of firing an instruction (an instruction is executable if all of its operands are available) and activation (an instruction is activated when it is fired and the resources required for activation are available).

One of the most important steps based on the dynamic data flow model is direct operand matching [4], [8]. The concept of direct operator matching represents the elimination of the costly process (in terms of computing time) related to associative searching of the operands. In this scheme, a matching vector is dynamically allocated in the Frame Store memory for each token generated during the execution of the data flow graph. The current location of a matching vector in the Frame Store is determined at compile time, while the Frame Store location is determined after the program starts. Each calculation can be described using an instruction address (ADR) and the pointer to the matching vector MVB in the Frame Store. The <MVB, ADR> value pair is part of the token. A typical action is the searching for the operands pair in the Frame Store. The matching function provides searching for the tokens marked identically. After the operand has arrived to the Coordinating Processor, the matching function detects if a commonly entered operand is present in the Frame Store. Detection is performed according to matching IX index. If the operand is not there yet, it is stored in the Frame Store, in the Matching Vector specified by base address of the MVB operand, into the item specified by index IX.

The operand matching process control at the operator input is influenced by the process of matching, instruction execution and generation of a result at its output. Using a compiler producing DFG output with forward searching that allows for the detecting and eliminating of redundant computations and change order of token processing, process control can be defined as the transition of activation signs along the edges of the data flow graph (Fig. 5), between the "producer" (P) operator and the "consumer" (C) operator.

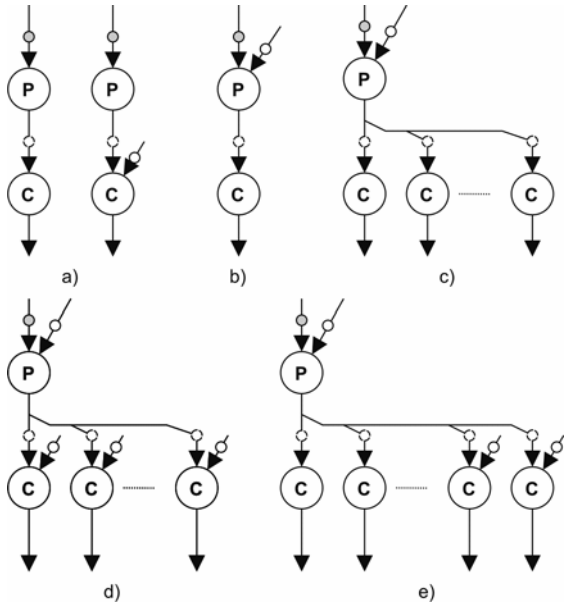


Figure 5
The operand matching (a – P-single input, b – P-double input, c – P-double input/C-single input, d – P-double input/C-double input, e – P-double input/C-u-single,v-double input)

The P and C operators can be single input, double input and management operators respectively. Varying configurations of P and C operators are shown in Fig. 5.

By processing the data flow graph the operators can be configured and connected as follows:

P-single input, C-single input (PS/CS). This configuration doesn't require operand matching. The token produced by the P producer is absorbed by the C consumers (Fig. 5a).

P-single input, C-double input (PS/CD). It requires operand matching on operator C (Fig. 5a).

P-double input, C-single input (PD /CS). Operand matching occurres only on operator P (Fig. 5b).

P-double input u×C-single input (PD /uCS). After processing the instruction defined by double input operator P, the result is distributed among u single input C operators. Operand matching occurs only with the operator P (Fig. 5c).

P-double input, v×C-double input (PD/vCD). The result is distributed between the v double input C operators. Operand matching is activated with operator P and operator C, too (Fig. 5d).

P-double input, u×C-single input, v×C-double input (PD/uCSvCD). The result of the activated double input operator P is sent to u single input and v double input C operators. Operand matching occurs with operator P as well as with double input C operators (Fig. 5e).

In this article we describe the operand matching control for configuration shown in Figure 5a.

3.3 Implementation of PS/CS and PS/CD Operators

The coordinating processor represents a dynamic pipeline system, which allows us to switch between the Load (L), Matching (M), Copy (C), Fetch (F) and Operate (O) states in a different order. The transitions between the states using a micro-management program is shown by means of a state diagram (Fig. 6). The micro-program manages the operand matching process.

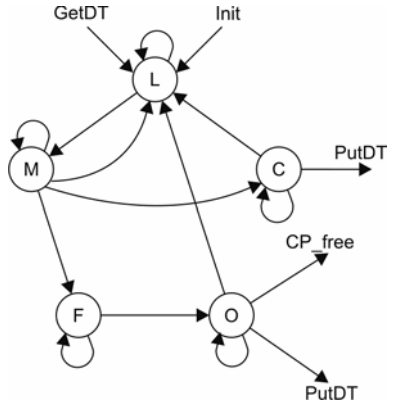


Figure 6
The state diagram of the operand matching control process

The control signals shown in Fig. 6 have the following meanings: CP_free – indicates the occupancy or availability of the coordinating processor; GetDT – read token from DQU; PutDT – write token to DQU; Init – initialization of pipelined system. The proposed architecture at a logical level of operand matching control is show in Fig. 8. FIFO registers with the following specifications have

been inserted to increase the throughput coefficient between the various stages of coordinating processor (Fig. 8):

- Between the stages L and F → register LFR
- Between the stages F and O → register FOR

Single input operator processing is done by means of micro-operations defined for each state (segment) of the dynamic pipelined multi-function unit of the current coordinating processor (Fig. 8).

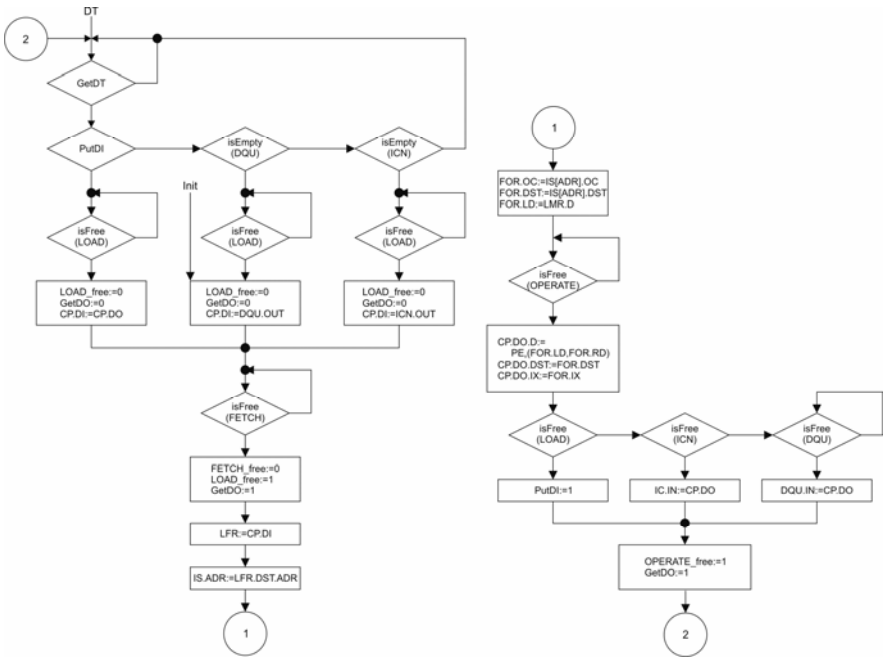


Figure 7
Micro-program for single input operators matching

The function isFree (X) tests the busy state of segment X. Micro-operations, which can be executed in parallel, are placed in a single command block (processing steps) in the program flowchart. Initialization of the coordinating processor is done by sign Init = 1. The boot command of the data flow program loads the data token from the DQU to the LOAD segment, sets the busy flag for the LOAD segment to 1 (i.e. the LOAD segment is occupied) and blocks the processing of the following tokens (GetDT = 0). If the next segment, the Fetch segment, is free, the token is loaded into the Load/Fetch register. After that, the micro program releases the LOAD segment and activates the loading of other tokens into the coordinating processor.

The control mechanism determines the DF address operator based on the LFR.DST.ADR address. The operator will be loaded from the instruction store into the Fetch/Operate register. If the Operate segment is not busy (isFree (Operate) = true), the operator is fetched from the Fetch/Operate register and processed. In the next step, if the CP is not busy, the result of the operator consumption and processing is available for processing in the same CP. Otherwise, the result is to another CP through the interconnection network. If all CPs are busy, the token is stored in the DQU.

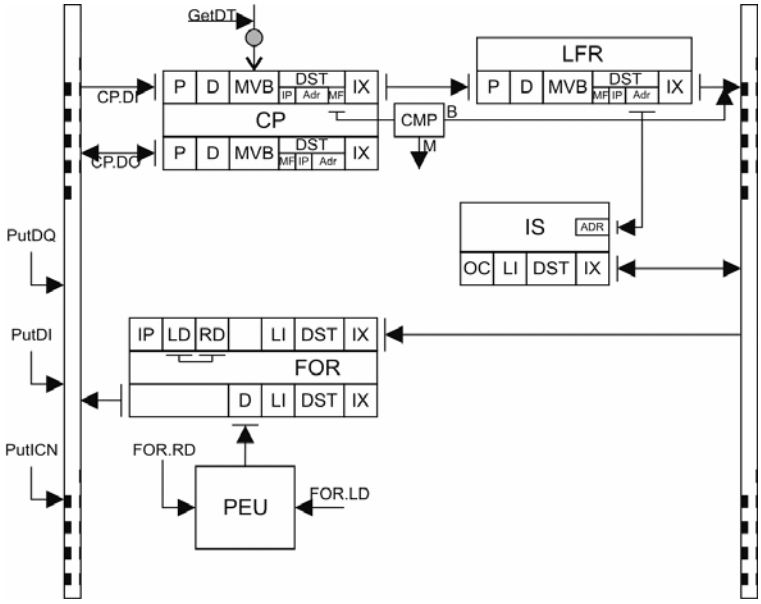


Figure 8
Pipeline system for processing of single input operators

Conclusions

This article presents the architecture design of the dynamic pipelined multiple-function system at a logical level, which handles processing of operand matching for single input data flow operators. Verification of proposal correctness can be realized by simulation or technical realization using statistical formulas and tools. At the Department of Computer and Informatics, a demand for realization of simulation tool for DF-KPI (being developed by this department) system has been recognized, within the scope of APVV-0073-07 and VEGA grant project No. 1/0646/09. Simulations are still to be performed at the level of the interconnection network (matching memory and tables), as well as at the level of the entire system. For system-level simulation compiler from a higher-level programming language to DFG is to be used (functional languages are preferred). The currently developed DF-KPI system with its principle of data processing and its parameters is intended

for solving tasks requiring brute force. The development of the DF-KPI system is focused on the fields of virtual reality [9] and computer security [13].

Acknowledgment

This work was supported by the Slovak Research and Development Agency under the contract No. APVV-0073-07 and VEGA grant project No. 1/0646/09: "Tasks solution for large graphical data processing in the environment of parallel, distributed and network computer systems" and by Agency of the Ministry of Education of the Slovak Republic for the Structural Funds of the EU under the project Centre of Information and Communication Technologies for Knowledge Systems (project number: 26220120020).

References

[1] Arvind, D. E. Culler: Dataflow Architectures. Annual Review in Computer Science, 1986, Vol. 1, pp. 225-253

[2] J. Carlström, T. Bodén: Synchronous Dataflow Architecture for Network Processors, Micro IEEE, Volume 24, Issue 5, Sept.-Oct. 2004, pp. 10-18, ISSN 0272-1732

[3] J. B. Dennis: Data-Flow Supercomputers. Computer, Nov. 1980, pp. 48-56

[4] J. R. Gurd, C. C. Kirkham, I. Watson: The Manchester Prototype Data-Flow Computer. Commun. ACM, Vol. 28, pp. 34-52, Jan. 1985

[5] Gy. Györök, M. Makó, J. Lakner: Combinatorics at Electronic Circuit Realization in FPAA, Acta Polytechnica Hungarica, Vol. 6, No. 1, pp. 151-160, 2009

[6] M. Jelšina: Design of Data Flow KPI Computer System (in Slovak). elfa s.r.o., Košice, 2004, ISBN 80-89066-86-0

[7] M. Jelšina: Computer system architectures (in Slovak), elfa s.r.o., Košice, 2002. ISBN 80-89066-40-2

[8] B. Lee, A. R. Hurson: Issues in Dataflow Computing. Advances in Computers, Vol. 37, Academic Press, Inc., San Diego, CA, 1993, pp. 285-333

[9] B. Sobota, J. Perháč, M. Straka, Cs. Szabó: The Applications of Parallel, Distributed and Network Computer Systems to Solve Computational Processes in an Area of Large Graphical Data Volumes Processing (in Slovak); elfa s.r.o., Košice, 2009, ps. 180, ISBN 978-80-8086-103-2

[10] S. Swanson, K. Michelson, A. Schwerin, M. Oskin: WaveScalar, Proc. of the 36th International Symposium on Microarchitecture (MICRO-36 2003) 2003, pp. 291-302, ISBN 0-7695-2043-X

[11] B. Verdoscia, R. Vacarro: ALFA: A Static Data Flow Architecture, In Proceedings of Fourth Symposium on the Frontiers of Massively Parallel Computation, McLean, VA, USA, 1992, pp. 318-325

[12] L. Vokorokos: Data Flow Computer Principles (in Slovak), Copycenter, spol. s.r.o., Košice, 2002. ISBN 80-7099-824-5

[13] L. Vokorokos, N. Ádám, A. Baláž, J. Perháč: High-Performance Intrusion Detection System for Security Threats Identification in Computer Networks, Computer Science and Technology Research Survey, Košice, elfa, s.r.o., Letná 9, 042 00, Košice, Slovak Republic, 2009, 4, 4, pp. 54-61, ISBN 978-80-8086-131-5

Autonomous Navigation and Landing Tasks for Fixed Wing Small Unmanned Aerial Vehicles

Sefer Kurnaz, Omer Çetin

Turkish Air Force Academy, ASTIN Yesilyurt, Istanbul, 34807 Turkey
s.kurnaz@hho.edu.tr, o.cetin@hho.edu.tr

Abstract: Autonomous control of UAVs has become a popular research topic in recent years. This paper is concerned with the flight of UAVs (Unmanned Aerial Vehicles) and proposes fuzzy logic based autonomous flight and landing system controllers. Three fuzzy logic modules are developed under the main navigation control system and three more for the autonomous landing control system to control of the altitude, the speed, and the position against the runway, through which the global position (latitude-longitude) of the air vehicle is controlled. A SID (Standard Instrument Departure) and TACAN (Tactical Air Navigation) approach is used and the performance of the fuzzy-based controllers is evaluated with time based diagrams under MATLAB's standard configuration and the Aerosim Aeronautical Simulation Block Set which provides a complete set of tools for rapid development of detailed 6 degree-of-freedom nonlinear generic manned/unmanned aerial vehicle models. The Aerosonde UAV model is used in the simulations in order to demonstrate the performance and the potential of the controllers. Additionally, some visual tools are deployed in order to get visual outputs that aid the designer in the evaluation of the controllers. Despite the simple design procedure, the simulated test flights indicate the capability of the approach in achieving the desired performance.

Keywords: UAV, Fuzzy Logic Controller, Autonomous Flight Control, Autonomous Landing System

1 Introduction

In the literature, it can be seen that the research interests in control and navigation of UAVs has increased tremendously in recent years. This may be due to the fact that UAVs increasingly find their way into military and law enforcement applications (e.g., reconnaissance, remote delivery of urgent equipment/material, resource assessment, environmental monitoring, battlefield monitoring, ordnance delivery, etc.). This trend will continue in the future, as UAVs are poised to replace the human-in-the-loop during dangerous missions. Civilian applications of UAVs are also envisioned such as crop dusting, geological surveying, search and rescue operations, etc.

One of the important endeavors in UAV related research is the completion of a mission completely autonomously, i.e. to fly without human support from take off to land on. For unmanned aerial vehicle systems to achieve full autonomy, smarter airplanes need to be developed. Full autonomy means performing takeoffs, autonomous waypoint navigation and, especially landings while the craft is hardest to control under computer control autonomously. The ground station control operator plans the mission and the target destination for reconnaissance and surveillance. UAV then takes off, reaches the target destination, completes the surveillance mission, and turns back to the base and lands on autonomously. Navigation is the topic which is most studied about. In literature, many different approaches can be seen related to the autonomous control of UAVs; some of the techniques proposed include fuzzy control [1], adaptive control [2], [3], neural networks [4], [5], genetic algorithms [6]. Capabilities of autopilot systems are important to successfully complete the mission of an UAV. A number of different autonomous capabilities may be required to be exhibited during a flight, like autonomous take off, navigation and autonomous landing.

In Section 2 of the paper the design of the navigation system with fuzzy controllers used for the autonomous control of the UAV is described and the autonomous landing model is defined by using the landing parameters given. Section 3 starts with the definition of the basic flight pattern for a UAV and then explain a sample mission plan which includes SID (Standard Instrument Departure) and TACAN (Tactical Air Navigation) procedures. A number of simulation studies are carried out with the fuzzy logic based autonomous navigation and landing system and some typical results are presented in Section 4. Finally the concluding remarks and some plans about future work are given in Section 5.

2 Fuzzy Logic-based System Design

As shown in Fig. 1, there basically are two separate computers used in an autonomous UAV. One of them is the flight computer and the other is the mission computer. UAV flight computer basically sets the control surfaces to the desired positions by managing the servo controllers in the defined flight envelope supplied by the UAV navigation computer as a command, reads sensors and communicates with the mission computer and also checks the other systems in the UAV (engine systems, cooling systems, etc.). The navigation computer and the landing system are a part of the mission computer. The mission computer carries out many other duties beside navigation, like payload control, communication with GCS, etc., the navigation computer is used for flying over a pattern which is designed before the flight or while flying in LOS (Line Of Sight). When GCS (Ground Control Station) gets control of UAV, the navigation computer goes into a passive state. During autonomous flight, the navigation computer gets the position values from

sensors (GPS receiver, altimeter, INS (Internal navigation System), etc.) and then matches these values (the current position) with the desired position values (the waypoint values). The navigation computer then determines the required maneuvers of the UAV to reach the goal position and sends these data to the flight computer to apply them to control surfaces.

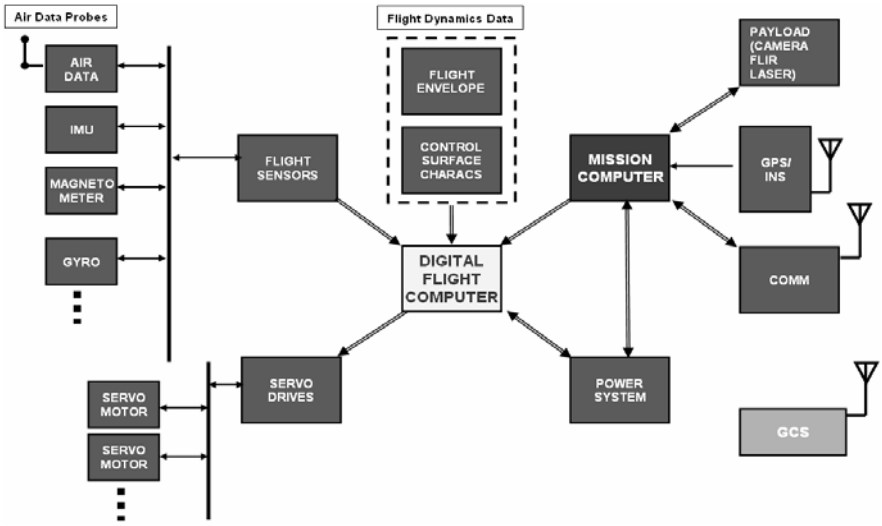


Figure 1
UAVs electronics architecture

Landing is one of the most critical parts of a flight, because, like in traditional aircrafts, UAVs aim to land at minimum air speed, consequently the stability conditions are severe and the maneuvering abilities are limited. A total of 65 Predators have crashed to date and thirty-six of the crashes were attributed to human error, and half of those occurred during landing [7]. The difficulties in taking off and landing arise mostly from instinctual factors, because pilots use their feelings in these periods of flight, such as feeling the ground rush and having peripheral vision. For a kite pilot, landing is a process which aims to see the wings inside of runway, but when a pilot is not in the cockpit, such feelings do not exist any longer. Because of these reasons, the manual control of an UAV from the ground is not a good alternative in the case of an emergency, especially during takeoff and landing.

The architecture used by the authors in [1] for their work on a Fuzzy Logic Based Navigation Control System (FLBNCS) forms the basis of the architecture for the Fuzzy Logic Based Autonomous Landing System (FLBANS). The mission computer has two main parts in this work, FLBNCS and FLBALS, as shown in Fig. 2. After getting the sensor values from the sensor interface, both FLBNCS and FLBANS calculate the desired attitude of the UAV attitudes which must be achieved by the flight computer. Then flight computer selects the correct

commands between the navigation c and the landing system commands. If UAV is in the final approach pattern it uses the landing systems commands, else it uses the navigation computer commands as inputs. The flight computer then calculates the control surfaces and the throttle positions by using its direct sensor inputs and the command inputs to reach the desired attitudes. The flow of this process can be seen in Fig. 2.

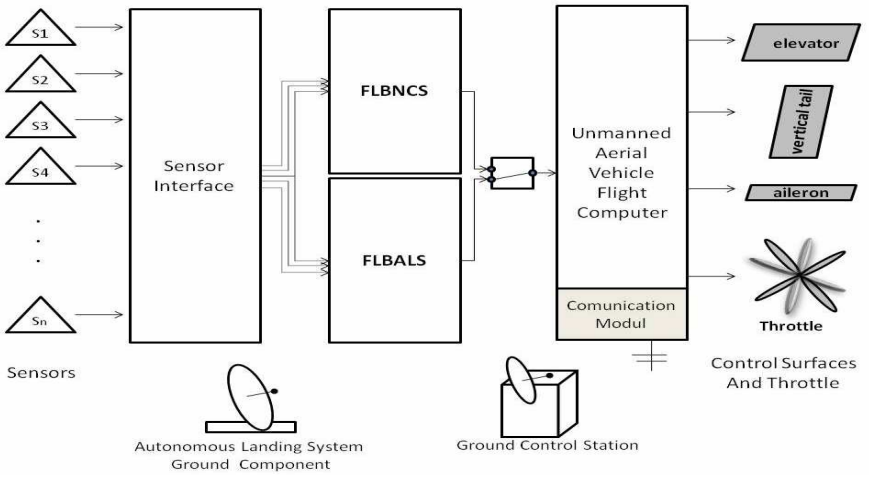


Figure 2
Autonomous System Design

The operation of the navigation computer proposed in this paper is fuzzy logic based. Basically, a fuzzy logic system consists of three main parts: the fuzzifier, the fuzzy inference engine and the defuzzifier. The fuzzifier maps a crisp input into some fuzzy sets. The fuzzy inference engine uses fuzzy IF-THEN rules from a rule base to reason for the fuzzy output. The output in fuzzy terms is converted back to a crisp value by the defuzzifier.

In this paper, Mamdani-type fuzzy rules are used to synthesize the fuzzy logic controllers, which adopt the following fuzzy IF-THEN rules:

$$R^{(i)} : If\,(x_1 isX_1^i)AND...AND(x_n isX_n^i)THENy_1 isY_1^i,....,y_k isY_k^i \qquad (1)$$

where R^l is the l^{th} rule $x=(x_i,...,x_n)^T \in U$ and $y=(y_i,...,y_n)^T \in V$ are the input and output state linguistic variables of the controller respectively, $U,V \subset \Re^n$ are the universe of discourse of the input and output variables respectively, $(x_1,...,x_k)^T \subset U$ and $(y_1,...,y_k)^T \subset V$ are the labels in linguistic terms of input and output fuzzy sets, and n and k are the numbers of input and output states respectively.

We consider a multi-input and single-output (MISO) fuzzy logic controller $(k=1)$, which has singleton fuzzifier. Using triangular membership function, algebraic product for logical AND operation, product-sum inference and Centroid defuzzification method, the output of the fuzzy controller has the following form:

$$y_j = \frac{\sum_{l=1}^{M}(\prod_{i=1}^{N}\mu x_i^l(x_i))\,y_i}{\sum_{l=1}^{M}\prod_{i=1}^{N}\mu x_i^l(x_i)} \tag{2}$$

Where N and M represent the number of input variables and total number of rules respectively. μx_i^l denote the membership function of the l^{th} input fuzzy set for the i^{th} input variable.

If the fuzzy controller types in literature are reviewed, it can be seen that there are two main classes of fuzzy controllers: one is position-type fuzzy controller which generates control input (u) from error (e) and error rate (Δe), and the other is velocity-type fuzzy logic controller. The former is called PD Fuzzy Logic Controller and the latter is called PI Fuzzy Logic Controller according to the characteristics of information that they process and system has two inputs, the error $e(t)$ and change of error $\Delta e(t)$, which are defined by

$$e(t) = y_{ref} - y \tag{3}$$

$$\Delta e(t) = e(t) - e(t-1) \tag{4}$$

Where y_{ref} and y denote the applied set point input and plant output, respectively. The output of the Fuzzy Logic Controller is the incremental change in the control signal $\Delta u(t)$. PD type fuzzy logic controller can be seen in Fig. 3. Then, the control signal is obtained by

$$\Delta u(t) = u(t-1) + \Delta u(t) \tag{5}$$

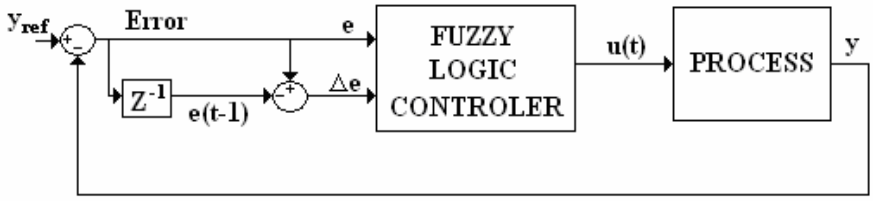


Figure 3
PD Type Fuzzy Logic Controller

As the second system in the architecture, FLBALS system uses the position inputs to calculate the exact location against the runway. It determines the error and calculates the corrective maneuvers by using three additional fuzzy logic subsystem blocks (in addition to the three blocks used for navigation). First fuzzy block is the lateral fuzzy logic controller which resolves the lateral errors in FLBALS. The second block is the vertical fuzzy logic controller which resolves the altitude errors and the last one is the speed fuzzy logic controller which tries to achieve the desired speed for the current conditions in FLBALS.

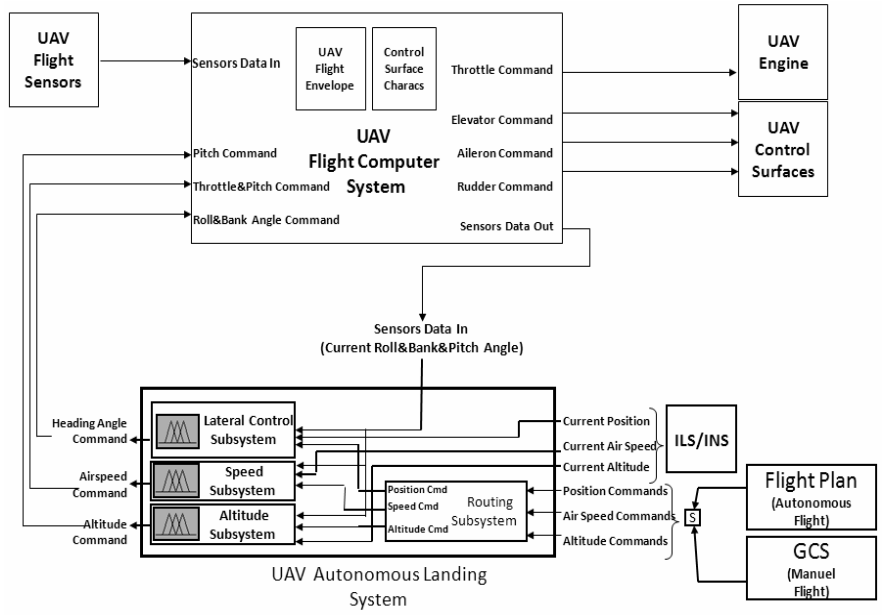


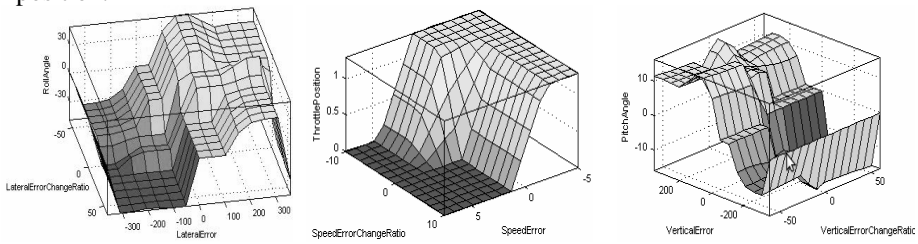
Figure 4

Fuzzy Logic Based Autonomous Landing System Design

The inputs to these fuzzy logic blocks are provided by different systems like ILS/INS and GPS [8, 9], laser based systems [10] or by vision based algorithms [11]. Other inputs of these blocks are the landing pattern flight plan or the manual commands issued by the Ground Control Station (GCS). The inputs of the fuzzy blocks can be seen in Fig. 4 and the surface diagrams of the blocks can be seen in Fig. 5.

3 Autonomous Flight Model

The test pattern used in this study includes a box pattern of Yalova Airport (LTBP) 18-26 Runway to show that UAV can fly autonomously a pattern which is designed for aircrafts if it has enough performance parameters. In classical SID and TACAN maps, a waypoint is defined with the radial angle and the distance between the VOR (VHF Omni-directional Radio Range) station and the waypoint. After transformation of the waypoints as GPS coordinates, UAV can apply SID departure and TACAN approach as a mission plan without a VOR receiver. In Fig. 6, the top and the side views of the test flight pattern is shown. This is a kind of box pattern. There are some important points which must be defined as GPS coordinates, like the initial approach point (IAP), the last turn point (LTP), the last approach point (LAP), the minimum altitude point (MIN) and the downwind turn point (DWTP). The UAV must reach the minimum altitude before the MIN point after takeoff. Then the UAV continues to the MIN point and starts to turn to reach DWTP. The particular set of these points that is used in the simulation studies is shown in Table 2. Each point of the pattern is represented by three values, the latitude and the longitude as the GPS position and the altitude as the vertical position.



a) Lateral Fuzzy Control Surface

b) Speed Fuzzy Control Surface

c) Vertical Fuzzy Control Surface

Figure 5
Control Surfaces

In this study, the UAV is considered to take off manually. Autonomous navigation starts when the UAV reaches the MIN point. It then reaches the way points in order and finishes when the UAV reaches the REP. The definition of each point includes speed, altitude and position (longitude and longitude coordinates) values.

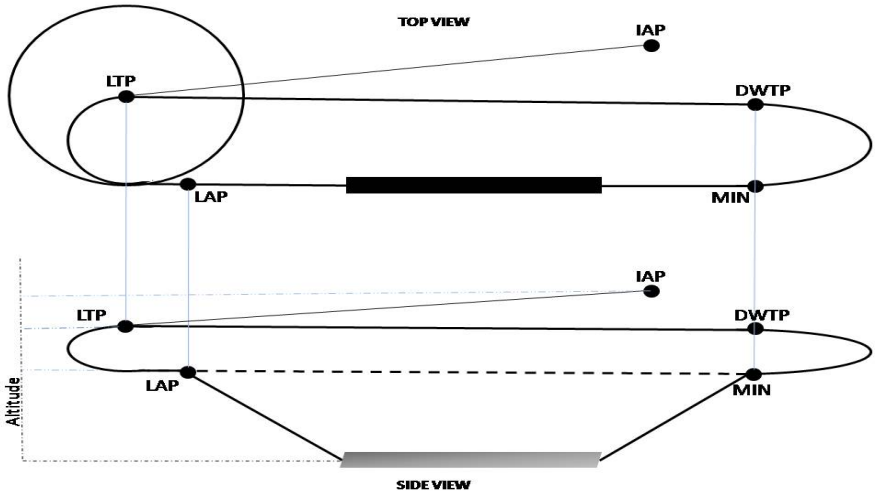


Figure 6
The test pattern of the UAV autonomous flight

To land, the aircraft must reach to the IAP and then it aims to reach the LTP and the LAP in order. After reaching the LAP, if airfield is not suitable for landing, it goes into a holding pattern. When the airfield becomes ready to land, the UAV completes the turn until the LAP is reached and goes into the final approach stage.

Table 1
Definitions of test pattern waypoints.

Point Name	Coordinate (GPS)	Altitude (feet)
Runway Starting Point (RSP)	N40 41 39.20 E29 22 34.81	6
Minimum Altitude Point (MIN)	N40 49 56.67 E29 22 34.81	1500
Down Wind Turn Point (DWTP)	N40 49 56.67 E28 18 43.00	1500
Initial Approach Point (IAP)	N40 41 50.00 E29 17 50.00	1700
Last Turn Point (LTP)	N40 32 26.55 E29 18 43.00	1500
Last Approach Point (LAP)	N40 32 26.55 E29 22 34.81	1200
Runway End Point (REP)	N40 40 56.54 E29 22 34.81	6

The values in Table 1 are contour a huge left side box pattern for Yalova (LTBP) Airfield 18-36 runway. First turn point in pattern is MIN point and it is 9 nm away from runway. Downwind leg is 3 nm away from runway and parallel to the 18-36 runways. Also LAP is 9 nm away from runway too. All the altitude values are mean sea level (MSL) in table.

To accomplish a successful landing, there are three main attributes which must be under control. First of them is the lateral position of the UAV with reference to the runway. As has already been stated, the goal is to touchdown on the lateral middle point of the runway like in Fig. 7. The second attribute is the vertical position, which is the AGL (above ground level) altitude of the UAV. It is a dynamic value since it changes according to the distance to the runway, but the usual glide path angle is 3 degree in aviation literature as in Fig. 7. The glide path angle is 3 degrees in nearly all the airfields in the world if there is no obstacle in this 3o path. The last main attribute is the speed. The speed value is a static value and it depends on the aircraft characteristics. The main aim is keep the desired speed value during the period of the final approach.

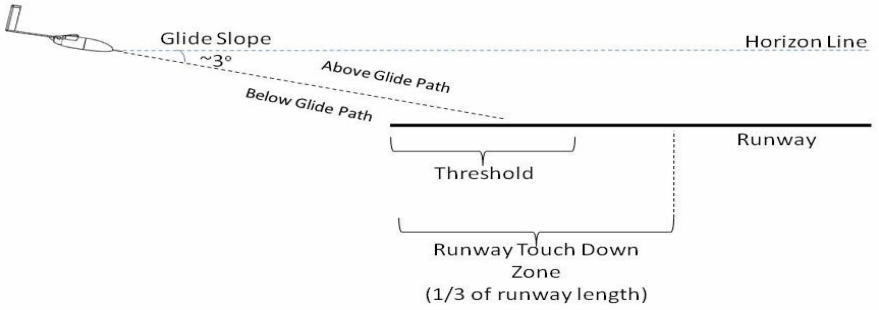


Figure 7
Desired Downward Velocity

In order to obtain the lateral position of the UAV with reference to the runway, different techniques can be used, like image processing [11] or radio based position calculators [8] and ILS (instrument landing systems) [9]. To measure the altitude and the speed of the UAV, laser altimeters and pito systems can be used respectively [10]. In this work, it is assumed that accurate measurements of these three parameters are available.

4 Simulation and Simulation Results

The performance of the proposed system is evaluated by simulating a number of test flights, using the standard configuration of MATLAB and the Aerosim Aeronautical Simulation Block Set [12], which provides a complete set of tools for rapid development of detailed 6 degree-of-freedom nonlinear generic manned/unmanned aerial vehicle models. As a test air vehicle, a model which is called Aerosonde UAV [13], shown in Fig. 8 together with its characteristics is utilized. The great flexibility of the Aerosonde, combined with a sophisticated command and control system, enables deployment and command from virtually any location.

Weight	27-30 lb,
Wing Span	10 ft
Engine	24 cc, 1.2 kw
Flight	Fully Autonomous / Base Command
Speed	18 – 32 m/s
Range	>1800 miles
Altitude Range	Up to 20,000 ft
Payload	Maximum 5 lb with full fuel

Figure 8
The Aerosonde and its specifications

In order to get visual outputs that aid the designer in the evaluation of the controllers, a number of aircraft instruments which are developed by using Delphi programming Active X components are deployed as shown in Fig. 9. Additionally, Flightgear open source flight simulator [14] is used to visualize the flight, like shown in Fig. 10. The details of these visual aids can be found in [15]. In order to be able to visualize the position of the UAV in GPS coordinate system, diagrams like the one shown in Fig. 15 are also plotted.

Figure 9
UAV Aircraft Instruments to get visual outputs of UAV parameters and mission planning

Figure 10
Visualization of landing by the use of FlightGear

The autonomous landing system test pattern can successfully be achieved by using the fuzzy logic based navigation computer system. In this work, the final approach period of landing pattern is handled by FLBALS. Final approach begins with the LAP and finishes at the touchdown point of the runway. The coordinates of these points and the elevations are given in Table 1 and navigation computer Simulink architecture can be seen in Fig. 11.

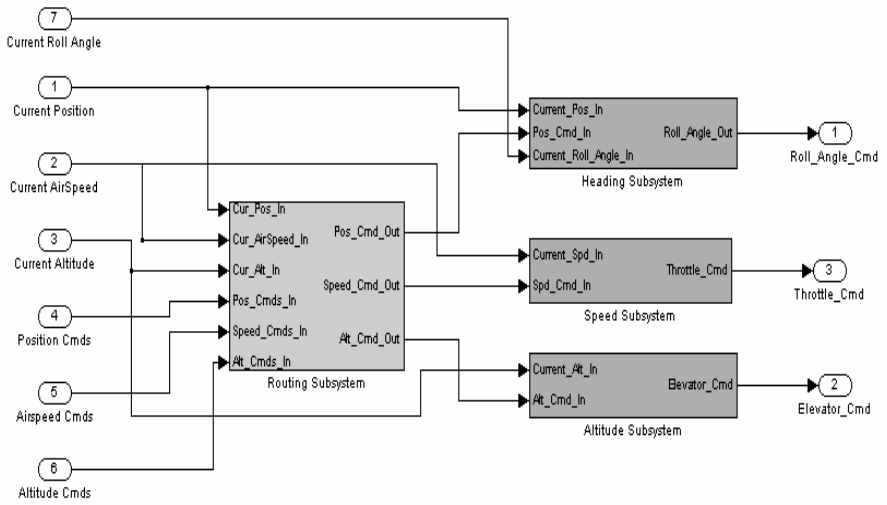


Figure 11
Navigation Computer in Simulink

5 Discussions on the Simulation Results

The UAV must reach exact altitude values during the flying pattern as shown in Fig. 6. There are some levels which depend on the distance from the runway. The dashed line shows the altitude command and the other one shows the current altitude at that simulation time. As we can see in Fig. 12, fuzzy logic-based autonomous landing system gets the desired altitude values in desired time. Also it manages not to sway too much from the 3o glide path angle throughout the pattern.

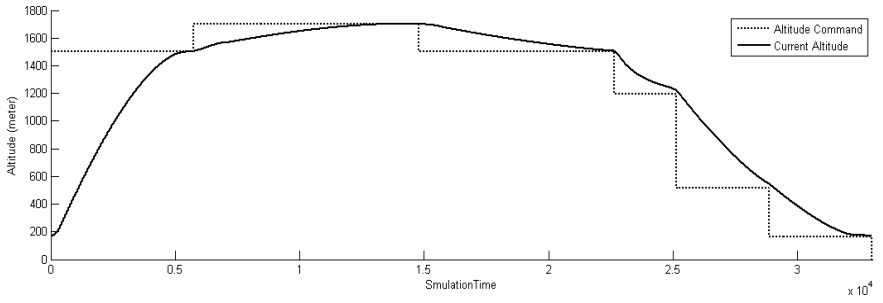


Figure 12
Current and Command Altitude – Simulation Time Diagram

In Fig. 13, UAV's current and command pitch degrees can be seen. Pitch commands produced by FLBNCS and FLBALS during simulation. UAV flight computer manage to hold desired pitch values during test flight. Pitch commands represents the desired values and dashed line shows the instant pitch value of UAV.

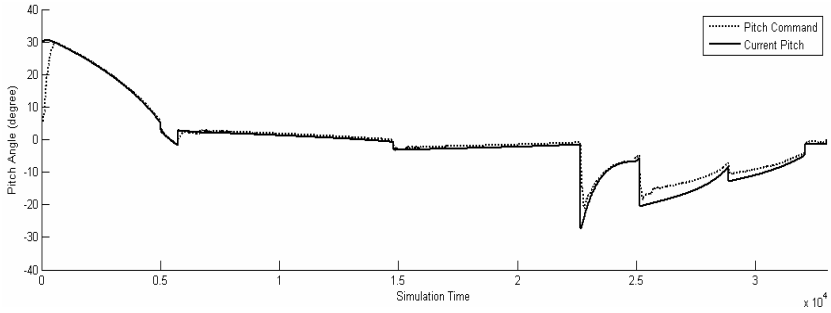


Figure 13
Current and Command Pitch Degree – Simulation Time Diagram

The last approach air speed of Aerosonde UAV is approximately 60 knot. The fuzzy logic-based autonomous landing system therefore tries to hold 60 knots during approach as shown in Fig. 14. The dashed line shows the desired air speed value and continues one indicates the current air speed of UAV in that simulation time.

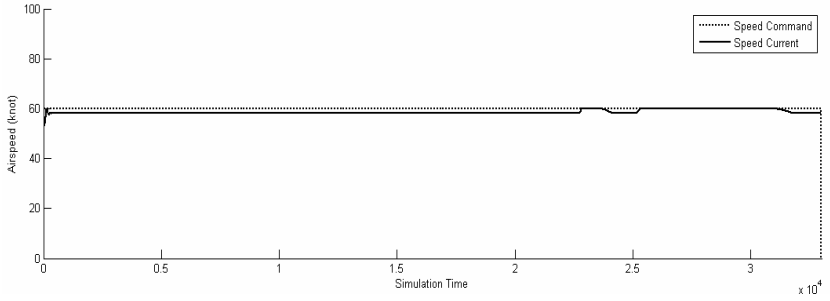


Figure 14
Current and Command Air Speed – Simulation Time Diagram

The vertical control and the air speed UAV are parameters that are related to each other. When the UAV pitches up, its speed decreases in parallel. The opposite of this is true also, when the UAV pitches down its speed increase. However, in this work there is no control relation between the air speed and the vertical control. The control of the air speed is provided by just using the throttle. The Aerosonde UAV is a kind of small fixed wing UAV. So this technique works to get airspeed of UAV under control. But major UAVs air speed must be controlled by using pitch angle and throttle together. So the architecture of control must be definitely different one.

In Fig. 15, the instant position of the UAV is depicted in the two dimensional space during the test pattern. The two dimensions of space are the GPS coordinate frames to show the exact position of the UAV. The other dimension is the altitude of UAV in meters, shown in Fig. 15. By looking at these diagrams, it can be stated that fuzzy logic based autonomous landing system manages to hold UAV in the correct position during the test pattern.

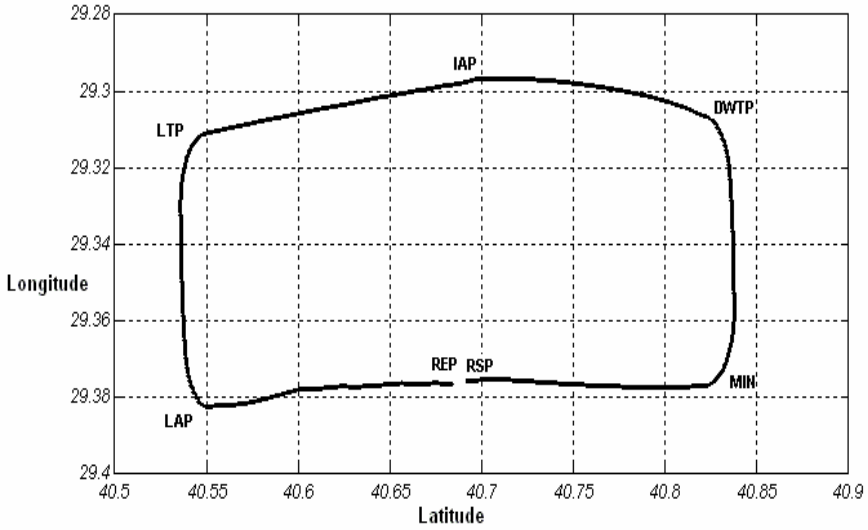


Figure 15
UAV Position in 2D GPS Diagram

As shown in Fig. 16, the UAV reaches the waypoints which have been defined in test pattern waypoints table (Table 1). After manual take off, the fuzzy logic based navigation computer system (FLBNCS) takes the control of the UAV until it reaches the IAP. After reaching the IAP, the UAV starts to be controlled by the fuzzy logic based autonomous landing system (FLBALS). The figure indicates that both of the fuzzy logic based systems successfully manage to navigate UAV in the test pattern.

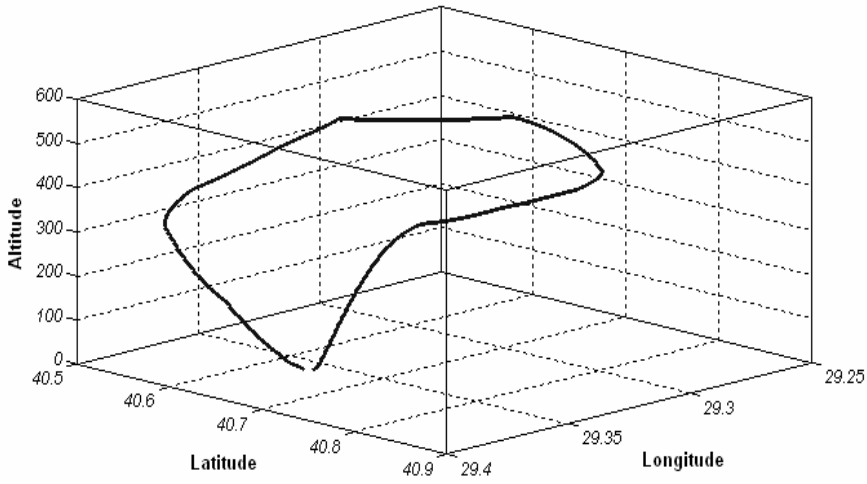


Figure 16
Current Position of UAV in Test Pattern Diagram

Conclusions

The mission computer design described in this paper provides autonomy to the UAV in all phases of a typical mission except take off. The main purpose of the navigation computer is to enable the UAVs to accomplish its mission autonomously, without any (or with minimal) input from the operator In such a way the airplane is endowed with with improved dynamic stability by the regulation of the flight parameters within limited ranges and with, at the same time, a successful tracking of the UAV mission plan.

The main contribution of this paper has been to demonstrate that the navigation and the landing tasks of small, fixed wing aerial vehicles can autonomously be achieved by fuzzy logic based systems. The simulation studies have shown adequate overall performance of the controllers. The main objective of this work is to keep the UAV in a pattern and also in a frame which is critical to hold the correct position during the final approach. This frame will be smaller when the UAV gets closer to the runway. The controllers must therefore show high performance against disruptive effects like wind. In our future work we will demonstrate the performance of fuzzy logic based autonomous landing system under disruptive effects.

Although there are many control law architectures, the classic PID control approach augmented with online gain scheduling provides the ideal mix of robustness and performance for typical aircraft dynamics. The stability and control loops can be tuned to provide the desired performance and robustness specifications by adjusting a set of autopilot parameters or gains. But this is done through linear analysis – the nonlinear aircraft model is linearized for a

representative set of flight conditions that cover the operating envelope of the aircraft. The linear dynamics of the closed-loop system (aircraft + autopilot) are analyzed in terms of stability and control responses (overshoot, settling time). By using fuzzy controllers, this difficult design process is avoided; nevertheless stable control and fast reaction time over conventional autonomous UAVs can be achieved as shown in this paper. The capability to do a dynamic planning of the desirable flight pattern is also important and this is done in this paper by using the current position of the moving UAV and the stationary target positions.

This paper also demonstrates that an UAV can apply a SID and TACAN approach without human control if it has enough performance like an aircraft. A series of SID for an airport can be planned just before the mission and the SID allowed by ATC (Air traffic controller) and meteorology can be applied autonomously. UAVs can to apply a TACAN approach too if they are properly instrumented with VOR and TACAN receivers. Without a VOR receiver, a SID can be applied by using GPS. Based on these attributes, fuzzy logic has been identified as a useful tool for developing controllers for UAVs so that they will be able to perform autonomous navigation.

In this paper, a fuzzy logic based autonomous flight controller for UAVs is proposed. The simulation studies presented verify that the UAV can follow the pre-defined trajectories despite the simplicity of the controllers. In future studies, the goals are to develop algorithms for a better tuning of membership functions which utilize the well known ANFIS (Adaptive Neuro Fuzzy Inference Systems) approach or possibly type-2 fuzzy sets. Autonomous take off will also be tried.

References

[1] S. Kurnaz, O. Çetin, O. Kaynak, “Fuzzy Logic Based Approach to Design Flight Control and Navigation Tasks for Autonomous UAVs”, Journal of Intelligent and Robotic Systems, v. 54, 2009, pp. 229-244

[2] C. J. Schumacher, Rajeeva Kumar: Adaptive Control of UAVs in Close-coupled Formation Flight. Proceedings of the American Control Conference, Volume 2, 2000, pp. 849-853

[3] B. Andrievsky, A. Fradkov: Combined Adaptive Autopilot for an UAV Flight Control. Proceedings of the 2002 International Conference on Control Applications, Volume 1, 2002, pp. 290-291

[4] W. R. Dufrene, Jr,: Application of Artificial Intelligence Techniques in Uninhabited Aerial Vehicle Flight. The 22nd Digital Avionics Systems Conference, Vol. 2, 2003, pp. 8.C.3 - 8.1-6

[5] Y. Li, N. Sundararajan, P. Sratchandran: Neuro-Controller Design for Nonlinear Fighter Aircraft Maneuver using Fully Tuned RBF Networks. Automatica, Vol. 37, 2001, pp. 1293-1301

[6] J. A. Marin, R. Radtke, D. Innis, D. R. Barr, A. C. Schultz: Using a Genetic Algorithm to Develop Rules to Guide Unmanned Aerial Vehicles. Proceedings of the IEEE International Conference on Systems, Man, and Cybernetics, Vol. 1, 1999, pp. 1055-1060

[7] J. Hodges, "Computer Co-pilots", C4ISR Journal, September 2009, pp. 30-32

[8] K. L. Koester, W. Vaillancourt, "TALONS 95 GHz Radar Sensor for Autonomous Landing Guidance" Aerospace and Electronic Systems Magazine, IEEE Volume 7, Issue 7, July 1992, pp. 40-44

[9] E. F. Roy, J. W. Davison, "Autonomous Landing Guidance Systems" Aerospace and Electronic Systems Magazine, IEEE Volume 1, Issue 5, May 1986, pp. 10-15

[10] J. Geske, M. MacDougal, R. Stahl, J. Wagener, D. R. Snyder, "Miniature Laser Rangefinders and Laser Altimeters" Avionics, Fiber-Optics and Photonics Technology Conference, IEEE, Sept. 30-Oct. 2, 2008, pp. 53-54

[11] Y. Fan, S. Haiqing, W. Hong, "A Vision-based Algorithm for Landing Unmanned Aerial Vehicles" Computer Science and Software Engineering, 2008 International Conference on Volume 1, Dec. 12-14, 2008, pp. 993-996

[12] Unmanned Dynamics, Aerosim Aeronautical Simulation Block Set Version 1.2 User's Guide: http://www.u-dynamics.com/aerosim/default.htm [Last reached 12/27/2009]

[13] FlightGear Open-Source Flight Simulator www.flightgear.org [Last reached 12/27/2009]

[14] Global Robotic Observation System, Definition of Aerosonde UAV Specifications http://www.aerosonde.com [Last reached 12/27/2009]

[15] O. Cetin, M. Ç. Kaplan, A. Aydın; "Fuzzy Logic-based Autonomous UAV Ground Control Station Simulator Design" Proceedings of 3rd National Defense Applications Modeling and Simulation Conference proceedings, USMOS METU, 2009

Iterative Algorithm for Channel Re-Estimation and Data Recovery in Nonlinearly Distorted OFDM Systems

Ján Šterba[*], Juraj Gazda[*], Marc Deumal[**], Dušan Kocur[*]

[*] Faculty of Electrical Engineering and Informatics,
Technical University of Košice
Letná 9, 042 00 Košice, Slovakia
sterba.jan@gmail.com, juraj.gazda@tuke.sk, dusan.kocur@tuke.sk

[**] Greco, La Salle - Universitat Ramon Llull,
Quatre Camins 2, 08022 Barcelona, Spain
mdeumal@salle.url.edu

Abstract: This paper presents a new algorithm for the nonlinear distortion cancellation and channel re-estimation in an iterative manner for the OFDM transmission systems. Whereas in practical OFDM systems, where nonlinear amplifiers are used in the transmitter, data as well as pilot symbols are nonlinearly distorted. As perfect channel state information is often assumed in many papers, the effect of the nonlinear distortion of the pilot symbols is usually not accounted for, and the data symbols are considered only. However, as the pilot symbols are used to obtain the channel state information, the distortion caused to pilot symbols by the nonlinear amplification is of the great importance, and for this reason this effect should be taken into account. To address this issue, in this paper, an efficient algorithm will be introduced for joint data recovery and channel re-estimation in an iterative manner for OFDM transmission system. This algorithm iteratively estimates the nonlinear distortion due to nonlinear high power amplifier of the transmitter and then it uses it to compensate for this nonlinear distortion of the data and the pilot symbols, which are subsequently used for subsequent and more precise channel re-estimation. As is shown in this paper, the proposed algorithm can be used to reduce nonlinear distortions even in situations when highly nonlinear amplifiers are applied.

Keywords: OFDM, iterative receiver, channel re-estimaiton, PAPR, nonlinear distortion

1 Introduction

The demands for high data rate transmission over multipath radio channels have increased rapidly. To fulfill this requirement, the utilization of multi-carrier-based transmission techniques seems to be an inevitable solution. In former methods, the signals received through the multipath channel suffer from severe intersymbol interference (ISI) since the delay spread becomes much larger than the symbol duration. In order to overcome this fact, the Orthogonal Frequency Division Multiplex (OFDM) transmission scheme, with very promising potential, is justifiably one major candidate for Beyond 3G and 4G wireless communication systems [1]. OFDM-based transmission systems are characterized by a large number of benefits in comparison with traditional schemes. Among others, easy time invariant multi-path channel equalization has received a great amount of interest recently. On the other hand, the high Peak-to-average power ratio (PAPR) of the OFDM signal makes it very sensitive to nonlinear amplification, which results in a high Bit Error Rate (BER) penalty, as well as to enormous out-of-band radiation. These effects have harmful impact on the overall OFDM transmission system performance, and therefore strict requirements to mitigate these effects must be taken.

High sensitivity to nonlinear amplification has greatly limited the practical applications of OFDM transmission systems. In order to alleviate this fact, many approaches based on different techniques have been introduced. The conventional solution is to back-off the operating point of the nonlinear amplifier, but this approach results in a significant power efficiency penalty. Alternative approaches for OFDM performance improvement are realized by the application of other usually computationally demanding methods of the PAPR reduction at the transmitter side, e.g. by the active constellation extension [2], tone reservation [3] or selected mapping [4]. Another well known and promising solution for reducing the BER of nonlinearly distorted multicarrier systems is to use nonlinear detection at the receiver side. The first contribution to this topic was proposed by Kim and Stuber in [5]. This technique reduces the clipping noise of OFDM symbols by decision-aided reconstruction at the receiver. In [6], Declerq *et al.* proposed reducing the clipping noise in OFDM by introducing a Bayesian interference to the received signal. Finally Chen *et al.* [8] and Tellado *et al.* [9] proposed iterative techniques to estimate and eliminate the clipping noise in OFDM.

However, besides the well-known and investigated BER degradation and out-of-band radiation, OFDM signal sensitivity to nonlinear amplification has in addition another important consequence. The pilot symbols based channel estimation inserts so-called pilot symbols into the transmission signal to carefully select either frequency and/or time domain patterns for the purpose of acquiring almost perfect channel state information (CSI). In the case of nonlinear amplification, the pilot symbols are also nonlinearly distorted and therefore CSI acquired by traditional estimation techniques is inaccurate [10]. As follows from the

aforementioned discussion, the joint effect of the nonlinear amplification and the multi-path propagation is inevitably required for the improvement of overall BER performance. Note that most papers assume only linear communication channels and linear channel equalization, and unfortunately the nonlinear distortion resulting from the large OFDM envelope fluctuation is in this case not taken into account.

The nonlinear equalization introduced in [11] is often described as computationally demanding and the optimal decisions are difficult to implement in practice due to the high degree of complexity. Because of this fact, one needs to find another, simpler approach to combat the joint effect of the nonlinear distortion and multi-path propagation. The promising solution is described in [12], where an algorithm of the nonlinear equalization applied in the OFDM system that estimates separately a linear multi-path channel and transmitted data is introduced.

In this paper, we introduce a similar idea based on the technique introduced in [9] but reformulated for OFDM based transmission systems with imperfect CSI due to the nonlinear distortion. The joint data and pilot symbol recovery is based on the nonlinear noise cancellation that is carried out in an iterative manner. Simulation results show that the proposed scheme reduces significantly the BER degradation of the OFDM systems undergoing nonlinear multi-path propagation.

This paper is organized as follows: The next Section describes the OFDM system model including the basic models of nonlinear high power amplifiers (Subsection 2.1) and the channel model (Subsection 2.2). In the Section 3, the proposed iterative receiver for the iterative nonlinear distortion cancellation and channel re-estimation is introduced, and subsequently, the simulation results are presented in the Section 4. Section 5 contains the conclusion and some considerations concerning future research.

2 OFDM System Model

The block scheme of the investigated OFDM system is presented in Figure 1. As can be observed from the figure, the bits assigned for the transmission are first mapped into the complex-valued vector of QAM constellation points. Then, every $m-th$ block of the mapped symbols is put into parallel streams using the serial to parallel converter, resulting in a $\mathrm{X}_s^m=[\mathrm{X}_0^m, \ldots, \mathrm{X}_{\mathrm{N-1}}^m]^{\mathrm{T}}$.

In the next step, the pilot symbols are inserted into the OFDM frame with a uniform distribution among all N sub-carriers using the block-type pilot arrangement (Figure 2). Then, the signal obtained by that approach is OFDM modulated. Since the OFDM signal is the sum of N independent tones with equal bandwidth and the frequency separation $1/T$, the complex baseband OFDM

signals are usually generated by using the inverse Fast Fourier transform (IFFT) as follows:

$$x_f^m(n) = IFFT(X_k^m) = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} X_k^m \exp\left(\frac{j2\pi kn}{N}\right), \quad (1)$$

where $x_f^m(n)$ is the resultant T/N - spaced discrete-time vector, and m denotes the m-th block of the input symbols.

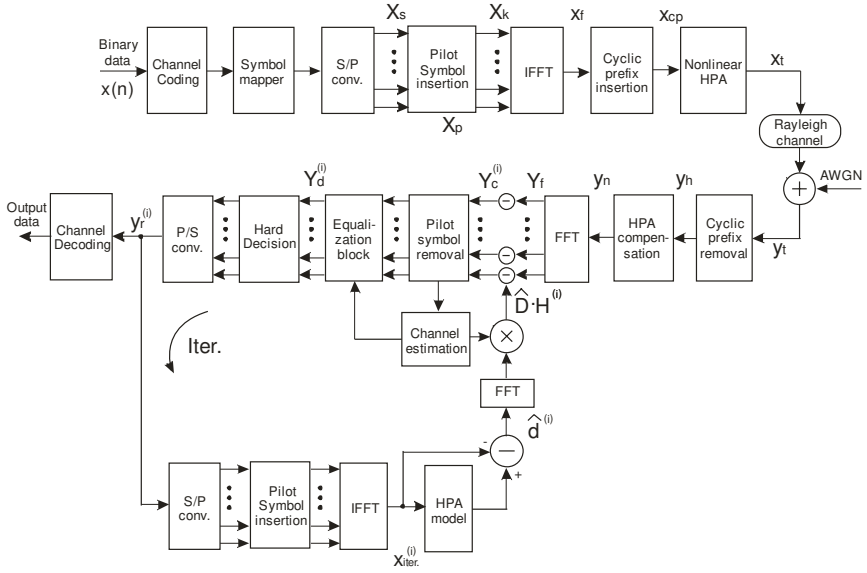


Figure 1
The block scheme of the OFDM system with iterative receiver

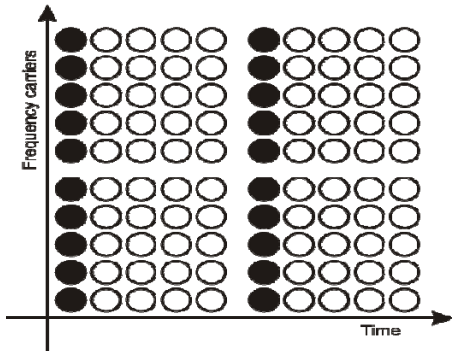


Figure 2
The block type of the pilot symbol arrangement

With the aim to mitigate the ISI introduced by the multi-path propagation of the transmitted signal, a carefully selected cyclic prefix (CP) is used. CP consists of a cyclically extended part of the OFDM symbol over the time interval $[0, T_{CP}]$. Its length is set longer than the delay spread assumed during the signal transmission. The resulting OFDM symbol with CP is afterwards given as

$$x_{cp}(n) = \begin{cases} x_f(N+n), & n = -N_g, -N_g+1, ..., -1 \\ x_f(n), & n = 0, 1, ..., N-1 \end{cases}, \tag{2}$$

where N_g is the length of a guard interval and N is the length of the OFDM symbol. This CP significantly simplifies the design of the equalization circuits because the equalization can be realized as a simple one-tap equalizer, as long as the length of the channel multi-path during the transmission is shorter than the designed CP. Finally, the signal $x_{cp}(n)$ is amplified by a high-power amplifier (HPA).

In this paper, the discrete time indexing $x(n)$ denotes Nyquist rate samples. In order to avoid aliasing and the out-of-band radiation to data bearing tones, the oversampling of the original signal $x(n)$ was implemented in simulations as described in [6].

2.1 Description of HPA Nonlinearity

In this paper, three types of HPA models are used, namely the Saleh model, the Soft Limit model and the Rapp model of nonlinearity. The Saleh model can be described by the following amplitude-to-amplitude modulation (AM/AM) and amplitude-to-phase modulation (AM/PM) characteristics:

$$G(u) = \frac{\kappa_G . u}{1 + \chi_G . u^2}, \quad \Phi(u) = \frac{\kappa_\Phi . u^2}{1 + \chi_\Phi u^2}. \tag{3}$$

In the expression above, we have chosen $\kappa_G = 2, \chi_G = \chi_\Phi = 1$ and $\kappa_\Phi = \pi / 3$. The Saleh model is used to evaluate the performance of a highly nonlinear HPA, which introduces not only amplitude but also phase distortion.

The Rapp model of the HPA is used to model solid state power amplifiers (SSPA) and is described with following AM/AM and AM/PM characteristics:

$$G(u) = \frac{\kappa_G u}{\left(1 + \left(\frac{u_x}{O_{sat}}\right)^{2s}\right)^{\frac{1}{2s}}}, \quad \Phi(u) = 0. \tag{4}$$

And finally, the Soft Limiter of HPA can be described as:

$$G(u) = \begin{cases} u & if \quad u \le 1 \\ 1 & otherwise \end{cases}, \qquad \Phi(u) = 0. \tag{5}$$

The Soft Limiter is used to model the case when pre-distortion is done at the transmitter. The operation point of the HPA is defined by the so-called input back-off (IBO), which corresponds to the ratio between the saturated and average input powers [3].

Let us now analyze the signal at the output of HPA. The following formulation exploits the fact that the OFDM signals with a large number of subcarriers are complex Gaussian distributed. By using the well-known Bussgang theorem [16], the output $y(t)$ of a memory-less nonlinearity with a Gaussian input $x(t)$ can be written as the sum of a scaled input replica of input signal and an uncorrelated distortion term $d(t)$ as

$$y(n) = \alpha . x(n) + d(n), \quad \text{where} \quad \alpha = \frac{R_{xy}(\tau_1)}{R_{xx}(\tau_1)}, \tag{6}$$

where R_{xy} is the cross-correlation function of the signals $x(n)$ and $y(n)$, and R_{xx} is the autocorrelation function of the signal $x(n)$. τ_1 is any possible value of τ, usually $\tau_1 = 0$ is chosen. As can be observed from the Bussgang theorem, the output signal of a nonlinearity with an OFDM-based input signal is equal to a scaled version of the input signal plus an uncorrelated distortion term. The complex scaling term α introduces a uniform attenuation and rotation to the data bearing tones and, therefore, it is responsible for the uniform attenuation and rotation of the signal constellation points. However, this can be easily compensated for at the receiver by introducing a correcting factor $\alpha^* / |\alpha|^2$. On the other hand, the distortion term $d(n)$, also often denoted as nonlinear noise, is responsible for clouding of the constellation and out-of-band radiation and cannot be compensated for by conventional receivers [17].

2.2 Channel Model

The output signal of the transmitter x_t is finally transmitted through the frequency-selective multi-path Rayleigh fading channel with additive white Gaussian noise (AWGN). In our simulations, the transmission channel was modelled as a 4-tapped COST 207 RA (Rural Area) channel [18], with characteristics given in Table 1.

Table 1
The tap coefficients of the COST 207 RA Rayleigh fading channel

Tap number	Delay [μs]	Power [dB]
1	0	-2.408
2	0.2083	-4.408
3	0.4167	-12.41
4	0.625	-22.41

The signal at the receiver antenna can be expressed as:

$$\begin{aligned} y_t(n) &= x_t(n) * h_t(n) + w(n) \\ &= (\alpha \cdot x_t(n) + d(n))h_t(n) + w(n), \end{aligned} \tag{7}$$

where $w(n)$ is the additive white Gaussian noise with zero mean and variance N_0 and $h_t(n)$ is the channel impulse response.

3 Structure of Iterative Receiver

Using the nonlinear amplifier, data symbols as well as pilot symbols are distorted by the nonlinear noise. In a system where the estimated CSI is used, the deterioration of the pilot symbols results in a significant degradation of CSI, and therefore, a less reliable equalization and subsequent data detection. The proposed iterative receiver decreases the nonlinear distortion caused to data and pilot symbols by a nonlinear cancellation, firstly proposed in [12] for clipping noise mitigation, in an iterative manner. Then, using the corrected pilot symbols, which are less distorted by nonlinear noise, the channel is in every iteration re-estimated and used for the more reliable equalization and subsequent data detection. Using this approach, the performance of the iterative receiver operating with the real CSI is significantly improved. The operation of the proposed iterative receiver is described in more detail in the following text.

Following Figure 1 at the receiver side, the CP is removed and the received signal is compensated by the complex scaling term α in a block of HPA compensation, as suggested in [17]. This compensation is realized as the simple channel equalization block set up with the correcting factor $\alpha^* / |\alpha|^2$. The output signal of the block of HPA compensation is sent to the FFT block for OFDM demodulation realizing the following operation:

$$\begin{aligned} Y_f(k) &= FFT\{y(n)\} \qquad k = 0,1,2...,N-1 \\ &= \frac{1}{N}\sum_{n=0}^{N-1} y(n)\exp\left(-\frac{j2\pi kn}{N}\right). \end{aligned} \tag{8}$$

Following the FFT of transmitted signal, the pilot symbols are removed from the received sequence of the data and CSI is estimated in the channel estimation block. Because the pilot symbols, similarly to the data symbols, are subject to the nonlinear amplification, the performance of the channel state estimation is degraded not only by the AWGN noise and channel fluctuations, but also by the nonlinear distortion due to the transmitter HPA.

The estimated CSI is further used in the channel equalization block, which has been implemented utilizing zero forcing method.

After equalization, the obtained signals are converted from the parallel streams to a serial sequence by the parallel-to-serial converter. The QAM constellation points are estimated using the hard decision block.

These hard estimates are subsequently fed into the backward loop of the iterative receiver, the simplified version of which has been firstly used in [6] for the mitigation of the clipping noise.

In the backward loop the nonlinear noise $d(n)$ can be estimated from received vector $y_r(n)$ and further subtracted from the signal $Y_f(k)$, thus decreasing the distortion effect of the nonlinear HPA. Although this estimate of the nonlinear noise denoted as $\hat{d}(n)$ is degraded by the incorrect estimate of the transmitted signal $x(n)$ and by CSI degraded by the channel noise and the nonlinearly distorted pilot symbols, the overall performance will be improved if:

$$E[|\, d(n) - \hat{d}^{(i)}(n)\,|^2] < E[|\, d(n)\,|^2]. \tag{9}$$

The resulting signal $Y_c^{(i)} = Y_f - \hat{D} = Y_f - FFT(\hat{d})$ will be less distorted by the nonlinear noise, and subsequently decoding will result in a better and more reliable estimate of $y_r(n)$. The symbol (i) denotes the order of the loop in the iterative process, and in this manner it will be used throughout the following text.

Newly estimated $y_r^{(i)}$ can be further fed into the feedback of the iterative receiver again, thus increasing the overall performance with every iteration of the iterative receiver.

The outlined iterative process of decreasing the nonlinear noise inflicted by the HPA to the transmitted pilot symbols and the data can be described as follows:

1) From the signal Y_f obtained from the output of FFT, the pilot symbols are extracted. Then, the CSI is obtained in channel estimation block at the pilot symbol positions by performing

$$H_p^{(i)}(k) = \frac{Y_p(k)}{X_p(k)} \qquad k = 0,1,...,N-1 \tag{10}$$
$$= \frac{H(k)\cdot X_p(k) + H(k)\cdot D + n}{X_p(k)},$$

where $Y_p(k)$ and $X_p(k)$ are the received and transmitted pilot signals at the $k-th$ pilot symbol position, D is FFT of the nonlinear distortion d inflicted by the nonlinear amplifier, and (i) denotes the iteration number.

2) Using this estimated CSI, the data symbols are equalized by utilizing zero forcing, and hard decisions of the constellation points are made by

$$\overline{X}_n^{(i)} = \min_{\{X\}} |Y_d^{(i)} - \eta_n X|, \quad 0 \le n \le N-1. \tag{11}$$

3) The hard decision of the constellation points are then OFDM modulated by using IFFT:

$$x_{iter.}^{(i)} = \frac{1}{\sqrt{N}} \sum_{n=0}^{N\text{-}1} \overline{X}_n^{(i)} \exp\left(\frac{j2\pi kn}{N}\right), \quad 0 \le k \le N\text{-}1. \tag{12}$$

4) Resulting signal $x_{iter}^{(i)}$ is sent to the model of HPA given by q(·) as a nonlinear function. Nonlinearly distorted output of the model is then subtracted from its original version, thus creating an estimate of the nonlinear noise $\hat{d}(n)$ inflicted to the data as well as the pilot symbols by the nonlinear HPA:

$$\hat{d}^{(i)}(n) = q(x_{iter.}^{(i)}) - x_{iter.}^{(i)}$$
$$= (x_{iter.}^{(i)}(n) + d_{x_{iter}}^{(i)}(n)) - x_{iter.}^{(i)}(n). \tag{13}$$
$$= d_{x_{iter}}^{(i)}(n)$$

5) The nonlinear noise is then transformed into the frequency domain and multiplied by the estimated CSI:

$$\hat{D}^{(i)} \cdot H(k)^{(i)} = FFT(\hat{d}^{(i)}) \cdot H(k)^{(i)}. \tag{14}$$

6) Finally, the resulting signal is subtracted from the output $Y_f(k)$ of FFT, thus decreasing the nonlinear noise in an iterative manner:

$$Y_c^{(i+1)}(k) = Y_f(k) - \hat{D}^{(i)} \cdot H(k)^{(i)}. \tag{15}$$

Then, the iterative process goes back to the step 1, where the pilot symbols are extracted from the data stream and a new process of the obtaining estimate of the CSI and the detection of the data symbol begins, but now less distorted by the nonlinear noise. In our simulations, the maximum number of the examined

iterations was set to 7; however, in most of the simulations, the system achieved very good performance in a fewer number of the iterations, and in that case, the results with most important iterations are only shown in the simulation results. Although resulting signals will still be distorted by the nonlinear noise and by the imperfect CSI, which is degraded by a cumulative effect of the imperfect estimate and the nonlinear distortion of the pilot symbols, they will be now much less affected by the nonlinear noise. The resulting data are finally decoded in channel decoding block using the Viterbi decoding algorithm.

4 Simulation Results

The performance results of the OFDM system with the proposed iterative receiver for joint nonlinear distortion cancellation and channel re-estimation have been obtained using Monte Carlo computer simulations. Simulations were performed using 256 and 512 subcarriers with 64-QAM modulation. We considered OFDM transmission system with the Saleh, Soft limit and Rapp models of nonlinear amplifier. Transmitted signals were generated by using 4-multiple oversampling. The CP length was set to 3.33 μs. Subcarrier spacing was set to 18.74 kHz. The summary of all important simulation parameters is given in Table 2.

Table 2
The summary of the important simulation system parameters

Parameter	**Specification**
Number of subcarriers	256 and 512
Base-band modulation	64-QAM
Frame length	7
Channel model	COST207 RA (Rural Area) 4-tap channel
Noise	AWGN
Cyclic prefix length	3.33 μs, 6.66 μs
Equalization method	Zero forcing
Pilot symbols arrangement	Block-type
Pilot symbols per OFDM symbol	29
Data symbols per OFDM symbol	227
Oversampling factor	4

The simulation results outlined in the Figure 3 correspond to the Soft limit model of nonlinearity with IBO level = 1 dB for the OFDM system with 256 subcarriers and 64 QAM. The top-most curve is the bit-error rate for the coded system using convolution codes with a conventional, non-iterative receiver (e.g. without backward loop depicted in Fig. 1) with zero forcing equalization. The three curves below correspond to the BER after the first, second and third iteration of the iterative algorithm without improvement of CSI in the loops of the iterative receiver. The following three curves correspond to the BER after the first three iterations of the proposed iterative algorithm. Finally, the bottom-most curve is the BER for the linear HPA at the transmitter. The simulation results show that for Soft limit model of the nonlinearity, three and even two iterations are sufficient for the compensation of transmitter nonlinearity. From this figure, it is easily observable that the difference between the ideal, linear BER curve and the BER curve after third iteration of the proposed algorithm is negligible, and the iterative receiver significantly counteracts the nonlinear distortion inflicted to transmission by HPA.

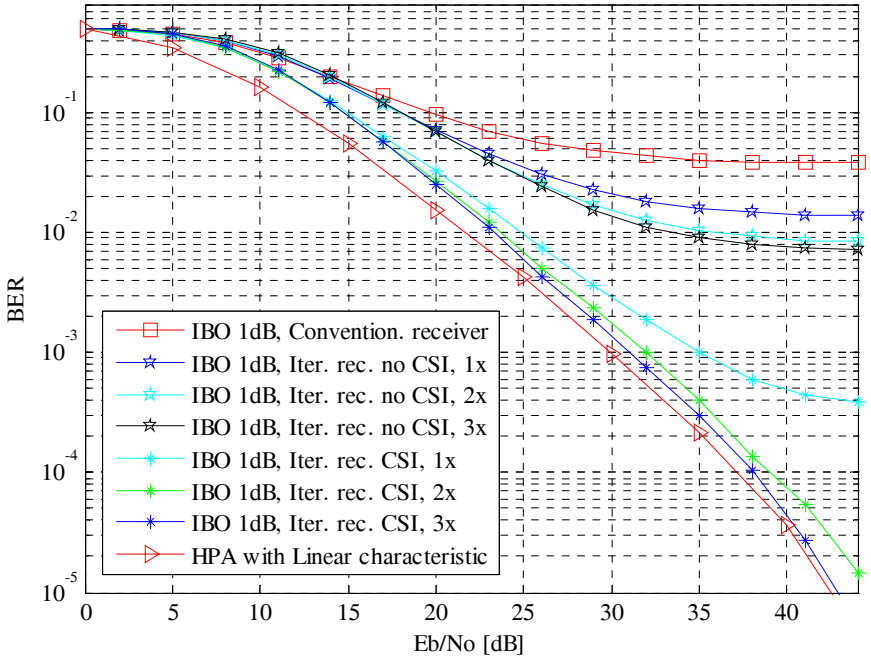


Figure 3
Performance of the iterative algorithm for a Soft Limit nonlinearity model with IBO = 1, 256 subcarriers, 16-QAM

Figure 4 presents the simulation results for the Saleh model of nonlinearity with IBO = 5 dB. The top-most curve corresponds to a typical non-iterative conventional receiver with zero forcing equalization. The 3 subsequent curves (denoted as 'Iter.rec.noCSI' in figure legend) correspond to an iterative receiver without improvement of the CSI with the estimated nonlinear distortion. The next seven curves correspond to the proposed iterative receiver. Their marks in the figure legend, '1x', '2x' and so on, correspond to the number of iterations of the investigated iterative receiver, which resulted in the creation of corresponding BER curves. With every following iteration, the channel is re-estimated and the nonlinear noise is more precisely estimated and then compensated, thus further increasing the performance. As can be seen from Figure 4, after 7 iterations the proposed algorithm can significantly compensate for nonlinear distortion, with a difference of less than 1 dB at BER = 5.10e-5 from the ideal case corresponding to a linear amplifier.

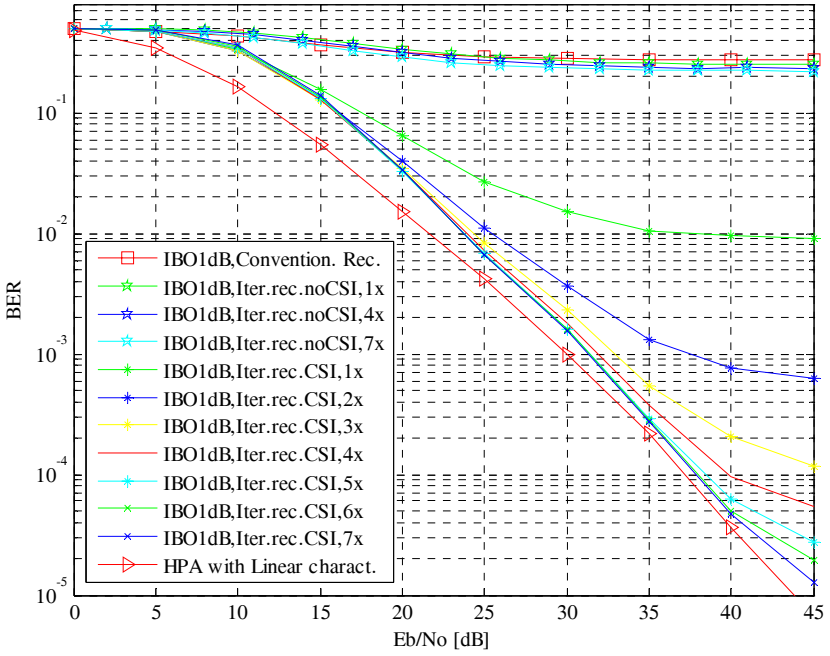


Figure 4
Performance of the iterative algorithm for the Saleh nonlinearity model with IBO = 1, 256 subcarriers, 16-QAM

In the case of the Rapp model of nonlinearity (Figure 5) with severe nonlinear distortions (IBO=1dB), the proposed iterative algorithm of the nonlinear noise cancellation and channel re-estimation provides a significant BER performance improvement (from $3*10^{-1}$ to $4*10^{-5}$ at E_b/N_0 = 40 dB), and in four or five iterations counteracts the nonlinear distortion almost completely.

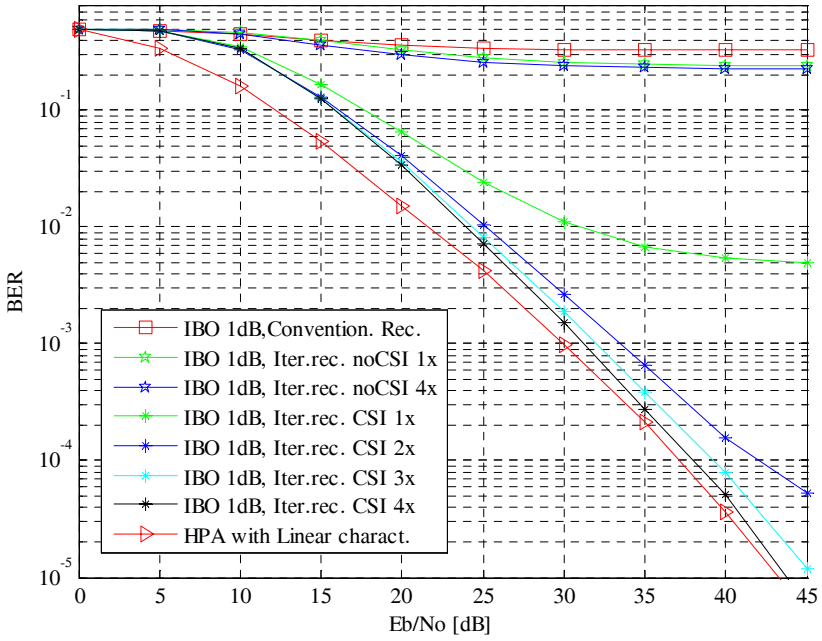


Figure 5
Performance of the iterative algorithm for the Rapp model of nonlinearity
with IBO = 1, for OFDM system with 256 subcarriers, 16-QAM modulation

Figure 6 presents the results for the nonlinearly distorted OFDM system with 512 subcarriers and 64-QAM modulation with the Saleh model of nonlinearity with IBO = 5 dB. The top-most curve is BER for the conventional, non-iterative receiver, e.g. a receiver without backward loop for nonlinear noise cancellation and channel re-estimation. As in the previous figures, the iterative algorithm improves the BER significantly. From the figure, it is easily observable that the difference between the ideal linear BER curve and the BER curve after the fourth iteration of the proposed algorithm is negligible, and the investigated iterative receiver can be used effectively for the compensation of nonlinear distortion inflicted to data and pilot symbols by nonlinear HPA. This would allow for the utilization of HPA in their nonlinear part of characteristic, thus resulting in more efficient energy consumption of OFDM transmitters.

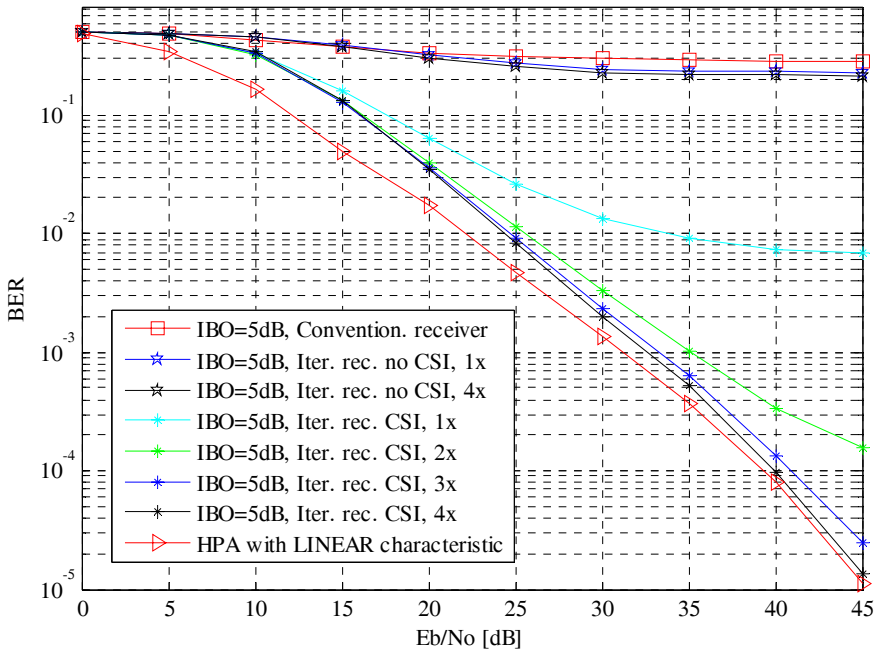


Figure 6
Performance of the iterative algorithm for the Saleh model
of nonlinearity model with IBO = 10, 512 subcarriers, 64-QAM

Conclusion

In this paper, the iterative algorithm for nonlinear noise estimation and subsequent noise cancellation and channel re-estimation is introduced. This algorithm estimates nonlinear distortion inflicted on data as well as pilot symbols, and subsequently uses it for data recovery and channel re-estimation in an iterative manner. With every following iteration, the data symbols and CSI are less distorted by the nonlinear noise, which results in a more reliable detection and improved BER performance. The application of the proposed algorithm in nonlinearly distorted OFDM symbols improves BER significantly for all tested models of the nonlinear HPA. The presented results have shown that each iteration of the presented algorithm improved the BER performance until the nonlinear distortion was significantly negligible. For the Saleh model of the HPA with 512 subcarriers and 64-QAM, the first iteration of the proposed algorithm improved BER from $3.10e^{-1}$ to $7.10e^{-3}$ at E_b/N_0 = 45 dB. The next iteration provided additional gain. The results have shown that the proposed algorithm can be used effectively to reduce the nonlinear distortion of the OFDM symbols with the real CSI obtained through pilot symbol channel estimations.

Acknowledgement

This work is the result of the project implementation of the Centre of Information and Communication Technologies for Knowledge Systems (project number: 26220120020), supported by the Research & Development Operational Program funded by the ERDF. This work was supported also by the project COST Action IC0803: RF/Microwave Communication Subsystems for Emerging Wireless Technologies. This work has been also funded by Grant Agency SPP Hlavička.

References

[1] Dahlman E., Parkvall S., Skold J., Beming P.: 3G Evolution. HSPA and LTE for Mobile Broadband, Academic Press, Elsevier, 2007

[2] Krongold B. S., Jones D. L.: PAR Reduction in OFDM via Active Constellation Extension. IEEE Transactions on Broadcasting, Vol. 49, No. 3, pp. 258-268, Sep. 2003

[3] Duemal M., Behravan A., Eriksson T., Pijoan J. L.: Evaluation of Performance Improvement Capabilities of PAPR Reductig Methods. Wireless Personal Communications, Vol. 47, No. 1, pp. 137-147, Oct. 2008

[4] Cimini L. J., Sollenberger N. R.: Peak-to-Average Power Ratio Reduction of an OFDM Signal Using Partial Transmit Sequences, IEEE Communications Letters, Vol. 4, No. 3, pp. 86-88, Mar. 2000

[5] Declercq D., Giannakis G. B.: Recovering Clipped OFDM Symbols with Bayesian Inference, in Proc. IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 157-160, Jun. 2000

[6] Kim D., Stuber G. L.: Clipping Noise Mitigation for OFDM by Decision-aided Reconstruction, IEEE Communication Letters, Vol. 3, pp. 4-6, 1999

[7] Gifford S., Kleider J. E., Chuprun S.: OFDM Transmitter Power Amplifier and PAR Reduction Performance: Measurement and Simulation, Proceedings on MILCOM 2002, Decision Syst., Gen. Dynamics, Scottsdale, AZ, USA, Oct. 7-10, 2002, pp. 591-595

[8] Chen H., Haimovich A. M.: Iterative Estimation and Cancellation of Clipping Noise for OFDM Signals, IEEE Communications Letters, Vol. 7, pp. 305-307, Jul. 2003

[9] Tellado J., Hoo L., Cio J. M.: Maximum-Likelihood Detection of Nonlinearly Distorted Multicarrier Symbols by Iterative Decoding, IEEE Transactions on Communications, Vol. 51, pp. 218-228, Feb. 2003

[10] Sterba J., Kocur D.: Deterioration of Channel State Information due to Nonlinearity Effect in OFDM systems, 11th International Conference RTT 2009, pp. 1-4, 2009

[11] AI-Naffouri T. Y., AI-Rawi G., Bahai A., Paulraj A.: A Least/Mean-Squres Approach to Channel Identification and Equalization in OFDM, in Proc. IEEE lCASSP'O2, Orlandoge, Florida, USA, pp. 2577-2580, 2002

[12] Ermolova N. Y., Nefedov N., Haggman S. G.: An iterative method for non-linear channel equalization in OFDM systems, 15th IEEE International Symposium on Personal, Indoor and Mobile Radio Communications, 2004. PIMRC 2004, Helsinki Univ. of Technol., Finland, pp. 484-488, 2004

[13] Dukhyun Kim, Stuber G. L: Clipping Noise Mitigation for OFDM by Decision-aided Reconstruction, IEEE Communications Letters, Sch. of Electr. & Comput. Eng., Georgia Inst. of Technol., Atlanta, GA, USA, Jan. 1999, pp. 4-6

[14] Coleri S., Ergen M., Puri A., Bahai A.: Channel Estimation Techniques Based on Pilot Arrangement in OFDM Systems, IEEE Transactions on Broadcasting, Sept. 2002, pp. 223-229

[15] Dardari D., Tralli V.,: A Theoretical Characterization of Nonlinear Distortion Effects in OFDM Systems, IEEE Transactions on Communications, Vol. 48, No. 10, Oct. 2000

[16] Bussgang, J.: Crosscorrelation Function of Amplitude-distorted Gaussian Signals, Research Laboratory of Electronics, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA, [Online]. Available: http://dspace.mit.edu, Tech. Rep. No. 216, Mar. 1952

[17] Deumal, M. H.: Multicarrier Communication Systems with Low Sensitivity to Nonlinear Amplification, Enginyeria i Arquitectura La Salle, Universitat Ramon Llull, Barcelona, 2008

[18] COST 207, Digital Land Mobile Radio Communications, Office for Official Publications of the European Communities, Luxembourg, 1989

Historical Origin of the Fine Structure Constant Subtilis Structurae Constan$\overset{U}{\underset{T}{N}}$s[1]

Part I. St. Stephen's Crowning Achievement

Péter Várlaki[1,2], Imre J. Rudas[3], Lászó T. Kóczy[1,2]

[1] Széchenyi István University
Egyetem tér 1, H-9026 Győr, Hungary, varlaki@sze.hu, koczy@sze.hu

[2] Budapest University of Technology and Economics
Műegyetem rkp. 1-3, H-1111 Budapest, Hungary

[3] Óbuda University
Bécsi út 96/B, H-1034 Budapest, Hungary, rudas@uni-obuda.hu

Abstract: This paper deals with the historical origin of the primordial image of the fine structure constant (FSC) and the number 137 as the Self's own number archetype, which plays a central role in the poetic-hermeneutic system of the Holy Crown of Hungary. We intend to show that the allegorical and symbolical images which are observed as the manifestation of unconscious and interpreted and analyzed in the Pauli-Jung collaboration have a lengthy synchronistic relationship with the "archetypal model" of the FSC found in the marvellous enamel pictures of St. Stephen's Crown.

Keywords: fine structure constant, number archetypes, hermeneutical systems, creative background process

1 Introduction

It is well known that the collaboration between Wolfgang Pauli and Carl Gustav JUNG in physical, psychological and idea-historical topics has led many authors from the 60s onwards (starting with Heisenberg [6]) to discuss and analyze these issues. A special feature of these discussions is that they are directed, in addition to psychological and physical aspects, to Jung's interpretation of Pauli's dreams,

[1] The special "decompositional" feature in the title would like to represent the double meaning of the expression "constans unit" (or "uni T" in Latin i.e. the famous Sign of Constantine) and the name "Constantinus".

with a manifest declaration of the crucial role of the Number archetypes and, at the same time, with a permanent hidden presence of the number 137, as well as the usually "allegorical concepts" of the (inverse) fine structure constant (FSC). [14, 17, 18]

In our earlier papers we dealt in detail with these interesting and potentially fruitful "results" of the Pauli-Jung "correspondence", focusing mainly on the problem of the identification and interpretation of the FSC. [30, 34-37] Our early considerations mainly focused on the germs of a background control and system theory closely connected with the interpretation of the FSC and number archetype 137, strongly linked with the interpretation and possible identification of a so called unconscious creative background processes and their "cognition" by the so called (4+1) symbolic background languages introduced and discussed in the Pauli-Jung correspondence [14, 17]. (The concept was introduced by Pauli in his scientific cooperation with Jung, as he explained a possible collective unconscious background of scientific discoveries.)[2]

It was also discussed in our earlier papers that Pauli took the not-directly-measurable archetypical influences as regulating and ordering processes [17, 20]. Along with the manifestation of the unconscious in the consciousness in the form of images and ideas, he considered the whole "system" more or less as a control system, according to the general view of this concept prior to 1958. Therefore, the conceptualization of the background processes as input/output dynamic systems reinterprets Pauli's view of regulating dynamic systems applied between the unconscious (input) and conscious (output) processes for the treatment of sophisticated cognitive problems (partly on the basis of the probabilistic concept of archetype introduced by Pauli and Jung [14]).

[2] Pauli explains several times in his letters and essays that there are at least three different type of mediator "languages" of cognition corresponding to the usual (e.g. scientific discovery) and the unusual (e.g. active imagination, dreams) ways of "knowledge transformations": theological/metaphysical language; the physical/ symbolic language; and the language of the psyche, or analytical psychology. In addition to these, he postulates a fourth symbolic mediator language that is based on number archetypes, but he considers this language "unknown" so far. These four languages can be related to the four ontical layers of Jung and Pauli: spiritual, psychical, material, and transcendent. In addition to the competence and language of the three fields mentioned (mythological-theological, psychological and physical background theory), the fourth is, in our opinion, not a future "neutral" new language connected to the number archetype, but an artistic competence (and language) in the widest sense: the aesthetical category [30, 34]. The mentioned fourth neutral language appears in the focus as a fifth one, with the number archetype '137', and maybe with the paradigm of the background system and control theory. This scheme corresponds to Jung's famous four functions of the consciousness, where there is a (fifth) transcendent function unifying the other four. Pauli intuitively grasped the significance of this perspective. [17]

According to the geometric and arithmetic Self (*Selbst*) representation, i.e. the Mandala, using a circle with divisions of natural numbers, we introduced a concept of the joint geometric and arithmetic number archetype. It can be a simple function consisting of the basic geometric number archetype π and natural numbers. These approximating and meaningfully interpretable formulas of the FSC, together with the number 137, can be considered as a representation of the generative-receptive Number archetype of the Self (*Selbst*).[3] [30, 34]

The mysterious nature of the FSC[4] is also a well known subject of many varied discussions, and it has partly been discussed by us as well. However, here we intend to select some important meaningful details from this vast literature, ones which are closely related to the actual problems we discuss in this paper.

As is well known, in addition to the mystery of the origin of elementary particle masses, there is the mystery of the dimensionless fine structure constant ... 1/137. In his book *Elementary Particles: Building Blocks of Matter*, Harald Fritzsch describes it as follows:

"*Let us emphasize that this number is the most prominent number in all of the natural sciences. Ever since its first introduction it has caused a lot of speculation. After all, FSC gives the strength of the electromagnetic interaction, which gives it fundamental importance for all the natural sciences and for all technology".* [15]

Citing Fritzsch in his marvellous book, *The Power of Alpha*, MacGregor states:

"The mystery about α is actually a double mystery. The first mystery – the origin of its numerical a value α = 1/137 – has been recognized and discussed for decades. The second mystery – the range of its domain – is generally unrecognized."

Furthermore, he writes that:

"In the current paradigm of elementary particle physics the assumption is made that the masses are created from the Higgs downward. However, the phenomenology of α-quantized lifetimes and masses suggests that the (FSC) coupling constant α acting on the electron ***creates*** *these masses from the bottom up" [15]*

[3] As is well known, synchronicity, according to Jung, is the para-psychological equivalent of the psychological manifestation of the Self (*Selbst*), i.e. the Mandala phenomena [24].

[4] It is worth mentioning MacGregor's opinion: "*The puzzlement over the derivation of the "pure number" 137, or 1/137, has persisted for more than half a century. Eddington's attempts to explain it are legendary. In the present studies we argue the case that, in addition to its numerical value, there is in fact a second mystery involved with the constant α, which is the question: What is the extent of the domain in which α (FSC), operates as a coupling constant?" [15]*

Thus, the *"... experimental evidence is displayed to substantiate this "power of FSC" viewpoint."* Feynmann's opinion is also frequently cited: "*It's one of the* ***greatest*** *damn mysteries of physics: a* ***magic number*** *that comes to us with no understanding by man".* [30]

Werner Heisenberg once proclaimed that all the quandaries of quantum mechanics would shrivel up when 137 was finally explained. [17] Pauli's opinion is well known too: "The concept and interpretation of FSC leads us out from the physics." [20]

And let us note MacGregor again:

"*Thus we have a mathematics that works, but we do not know just what it is calculating. This in a sense is pure mathology, with no interpretive underpinning. The various infinities, which are equally mysterious, are evidently being handled properly, but their origin is left up in the air. As we will see, elementary particle mathology may be able to provide some insight into these enigmas"... "If we really want a solution the mystery of elementary particle masses, it seems that this is the path we must follow. It is a path that is guided by The Power of FSC". [15]*

Now let us consider some important but less known and rarely cited thoughts from Jung related to these kind of topics, concerning the enormous role of number archetypes in physics and psychology: One of the future problems for the human cognition is to explain,

"*…the connection of the psychology of the unconscious with the properties of whole numbers on the one hand and the properties of matter on the other." ... "Here psychological reactions take place which the epistemologist has hardly considered until now, if at all. The reason for this is that he does not know of the existence of an unconscious psyche. If, therefore, my cognitive process comes to a stop at one point or another, this does not mean that the underlying psychological process* (e.g. excited and controlled by an acausal background creative intelligence – author comment*) has also stopped. Experience shows that it continues regardless. When the physicist, for instance, can form no picture of the structure of the atom from the data at hand, something suddenly fleshes into his mind – a model, perhaps the planetary model – as a product of unconscious associative activity. This flash or "hunch" must be considered a psychic statement, which is ordinarily called intuition and is a common product of the external data and psychological apperception. Wherever the inquiring mind comes up against a darkness, in which objects are only dimly discernible, it fills the gap with previous experiences or, if these are lacking, with imaginative, that is with archetypal or mythic, material. In the construction of physical theories you will therefore find the closest analogies with the psychology of the unconscious, since this too is up against the same difficulties. Our psychic foundations are shrouded in such great and inchoate darkness that, as you peer into it, it is instantly compensated by mythic forms. When these compensations become too obvious, we naturally try to obviate them and replace them by "logical" concepts. But this is justified only*

when these concepts really do give adequate expression to what we have dimly discerned. Generally they don't. Hence the borderline concepts in both sciences are partly mythological. This would be a good reason for an epistemological-cum-psychological examination of their fundamental concepts". [12]

Later in this letter he continues:

"When we are confronted with this dark aspect of numbers, the unconscious gives an answer, that is, it compensates their darkness by statements which I call 'indispensable' or 'inescapable.'" ... "In this respect Number proves to be a fundamental element not only of physics but also of the objective psyche". (to F. Lerch 2. 09. 1956. [12])

In the Pauli-Jung relationship, the particular curiosity of the number 137 is its manifestation as a crown symbol in various dreams and allegorical images. Thus, in Pauli's initial dream (from his famous dream series in Jung's *Psychology and Alchemy*), we see a simple start, as "the strange hat" in the dream is treated as a crown, while in later dreams the 8 system and 32 system division of the circle (of the hat as crown) leads to a 137-structured interpretation of a crown – or to what is the same, a wheel or wreath (see dreams 1, 22, 35, 39, 51, and 59 [18]). In Jung's 1927 dream, as we show later and in detail in another article, the symbolism of the mandala appears in a 137 system as well. For Pauli and Jung this crown symbolism and the number 137 which represents it are first of all interwoven in their imaginations and interpretations with Kabbalistic meaning and structures. The crown interpretation of the number 137, the number value of the Hebrew word *Kabbalah (***137** = **קבלה***),* we introduced in an earlier work. [34] The best concise description is as follows (see Dan on the *"Fragment of Eleazar" [3, 4]):*

"Yufiel (the angel priest, the number of his name 137 = **יופיאל***) is then identified with the Crown ('atarah,* **עטרה***) on the head of the Creator, a crown which is comprised of the holy name of forty-two characters, and which was revealed to Ezekiel in the chariot (through the ofen,* **אופן** *=137). This Crown is the divine power which Moses could not see (Ex. 33, 23). Then the author continues this series of identification and states "where the crown is on the head of the Creator…it is the Crown Keter."* [3]

According to this concept, the number 137 means always the archetypal number representation of the Crown, and of the double masculine and feminine crown or crowns, as well[5]. Jung himself pointedly interpreted the Sephiroth structure, the

[5] In the original text „**אופן גימטרא יופיאל אופן אלף וו פה נון גימטרייא שדי**„ [4]. This means that the God (Saddai) is in the wheel (ofen), i. e. in 137, and because the other name for Ofen is Galgal **(גלגל)** whose other meaning is incarnation or reincarnation as Gilgul **(גלגול)**. Thus, this interpretation of **137** could be a "Biblical" basis for the structure and number of the incarnation of God. This "Fragment" (written in artificial old style with a quite new meaning) could have originated from the hypothetical Proto-Kabbalistic circle of St Stephen's court. [35]

essence of the Kabbalistic system, as a mandala-like dual masculine and feminine unified crown:

"With the statements of the Cabala (137), which as we have seen found their way into alchemy, our interpretation of Adam attains a scope and a depth that can hardly be surpassed. This interpretation includes ... the feminine principle itself. She appears chiefly as the "lower," as Malchuth (kingdom), Shekinah (the Indwelling of God), or as Atarah (Crown), the equivalent Below of Kether, the upper crown. She is also present in the „hermaphroditic" Sefiroth system, the right half of which is masculine and the left half feminine. Hence Adam Kadmon, as a personification of the whole "inverted tree," is androgynous, but the system itself is a highly differentiated contiunctio symbol, and, as such, divided into three parts (three columns of three Sefiroth each)". [9]

Concerning these interpretations, it might be interesting to note some spontaneous crown symbolisms from MacGregor related to the FSC. Here the author speaks in a manner closely resembling that of the crown symbolism about the perpetual creative power of the (inverse) fine structure constant, that is, the number archetype 137:

*"We begin this sojourn into the mathology of the elementary particle with a brief discussion of the mathological aspects of QED (*quantum electrodynamics*), which is surely the "crown jewel" of all mathologies" ... "In the lifetime and mass studies..., the PS pseudoscolar mesons – the "PS octet" of quantum chromodydamics plus the "singlet" – exhibit the clearest and most accurate examples of (FSC)* ***α****-quantization. They are the "****crown jewels****" – the best results we can offer".* Furthermore once again *"(The FSC) ... acting on the electron* ***creates*** *these masses from the bottom up" "... experimental evidence is displayed to substantiate this "power of FSC" viewpoint." [15]*

This symbolic concept of perpetual creation (*creatio continua*) and the perpetual incarnation (*incarnatio continua*) in connection with the number 137 can be connected with Pauli as well as with Jung. As Miller writes, the "doublet" lines are always linked with the fine structure constant; meanwhile, Pauli associates them with the concept of the perpetual incarnation. [17, 18] At the same time, both Jung and Pauli link the perpetual creation and perpetual incarnation with the concept of synchronicity[6]. Thus, based on the above, we can posit the hypothesis that the fine structure constant, in symbolic and allegorical form in background physics and other symbolic languages, closely connects via an Acausal Background Creative Intelligence with the crown symbolism and with the delicate ornamental construction of actual royal crowns. According to our hypothesis, under the "impetus" of earlier traditions, the algebraic and geometric

[6] See Jung: "*As incarnatio continua, it is synonymous with the creatio continua and actually means the materialization of a potentially available reality.*" [17] See further Pauli's view in footnote 9.

contemplation of the time (the middle ages), and the probable synchronicities, a few great minds, "pressured" by "Acausal Background Creative Intelligence" and with Jung-like intuitive mythical imagination, were able to recognise the universal significance of the ancient form of a much-later concept[7]. That is what we would like to demonstrate for the patient reader.

We would like to stress in advance that our approach is basically *empirical* and *heuristic,* and it is much more a question of discovery than of philosophical legitimacy.

2 The Problem and the Hypothesis

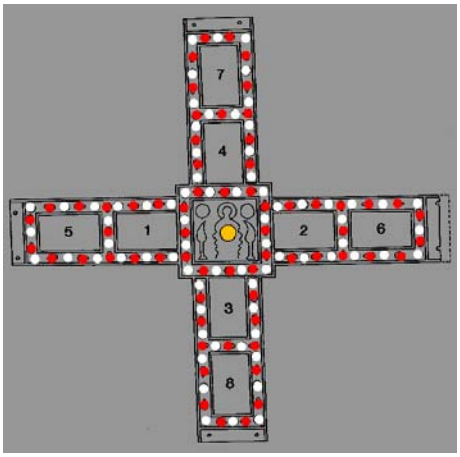
We will show that the historical origin of the concept and theory of the fine structure constant can with confidence be linked back to works of art designed and crafted in the court of King Saint Stephen of Hungary. Among these works of art, the image of Emperor Constantine on the Greek crown (*corona graeca*) of the Holy Crown of Hungary occupies an honoured place, where the special variable structure of the number 137 provides rich semantic variants to the attentive interpreter. Our narrower hermeneutic interpretation can be stated, in brief, that the name *Constantinus* carries within it the "*Constans*", which truly presents us on a fine structure the number 137, or that is, the number 1/137. The enamel picture in Latin satisfies the Latin concept of "*structura subtilis"*; and thus from *Constantinus* it follows that the "*constans*", together with the interpretation of the 1/137 visible on the image, can be referred to as *subtilis structurae constans*, which can be taken as an accurate Latin translation for the fine structure constant.[8] The structuring of the enamel pictures, with the various enamel colours and bits of gold separating them, is a fortunate symbolic expression of both the different outward forms of light and the spectrum lines. As we will see, on the Constantine enamel picture, we can define, or interpret, seven 137 structures, while on the entire Holy Crown, in addition to those on the Constantine image, we can find at least seven more 137 depictions. These connections definitively link the Holy Crown with the number 137. Furthermore, in the various strings of Judaic mysticism, as we have discussed, it has been well demonstrated that the wheel of

[7] *"As a consequence of the rationalistic attitude of scientists since the eighteenth century, the (creative) background processes that accompany the development of the natural sciences, although present, as always and of decisive effect, remained to a large extent unheeded, that is to say, confined to the unconscious. On the other hand, in the Middle Ages down to the beginning of modern times, we have no natural science in the present-day sense, but merely the pre-scientific stage, just mentioned, of a magical-symbolical description of nature."[20]*

[8] The dimension of the constant (FSC) is the "unit". See "Constans unit" = "Constantinus".

Ezekiel's vision, the *ofen* (*echod*), with its number-value of 137, is the archetypal concept and pattern for the Hebrew word for crown, the *atara*, whose Greek translation στεφανος (the Latin *Stephanus*) is the great Hungarian king's name. The depiction of the notion which we have so far expressed is well suited to the symbolic language of Pauli's Background Physics, a concept which Pauli himself introduced. In his physical symbolic background language, the "duplication" of the spectrum lines always refers to the fine structure constant (see e.g. [18]). As the physical background language conception of the duplication, he chooses the divine twins, whom he usually refers to as the Dioscuri, while elsewhere he calls them by their individual names, Castor and Pollux[9].

Among the 7+7 occurrences of the 137 structure on the Holy Crown, several of them very clearly, and as the most important variational version in the 68+68+1 *form*, bear the number 137. The most important central number 137 on the Latin crown is composed of 72 white pearls and 64 gems, with the central leaning cross. The double hoop visible in Fig. 1 and Fig. 2 contains the number 2x68 in a 2x36+32 structure.



Figures 1-2
Illustration of the 137 structures of white pearls and red gems on the Latin crown [28]

The 36 and 32 as Lulav (in Hebrew lo lev, 36 32 = **לולב**) is the frond of the date palm, whose name in Hebrew is Tamar. (This taken together can also refer to the 36 true men on the 32 paths of wisdom). Because on the feast of Sukkot, *Lulav* is

[9] See Pauli: "*...the Dioscuri myth on the one hand and doublet splitting of spectral lines and isotope separation on the other" and "Isotope separation is familiar to me as a symbol of the individuation process (doubling motifs, cf. the brothers Castor and Pollux, Christ is God and Man), which always appear when progress is being made in the development of consciousness and is connected with the incarnation of an archetype.*" [17]

the symbol of the Messiah king (while the *Etrog* symbolises the bride, Israel), the 2x36-32 clearly refers to Tamar's messianic twins, Pharez and Zerah. They "compose" the "sign" of the redeemer as 137, because Pharez is the forefather of David and Christ. Thus we get the symbolic composition, according to Pauli's physical background language, of the doubled spectrum lines, with the allegorical image of the messianic twins, where the twins give us the dominant denominator of the 137 fine structure constant.

In the great book of the proto-Kabbalah, the Book of Bahir, in passage 197, Pharez and Zerah are mystically treated, based on their names, as the Moon and "breaking through", and as the Sun and radiance. In other places, the treatment as the Sun and the Moon, as the morning and the evening, and as the day and the night exactly matches the traditional interpretation of the Olympos and the Hades, the daytime and night of Castor and Pollux (see [1]).

Thus, perhaps Saint Stephen (and his court), with a certain kind of pre-cognition *sensed* the concept of the fine structure constant some 900 year prior to its conceptual birth, while Pauli with retro-cognition perhaps discovered the constant's historical origin. This long-term synchronistic concord is perhaps most significantly demonstrated by this twin nature of Pharez and Zerah or by that of Castor and Pollux, with the number 137 connected with them.

Thus we could summarise the first part of our hypothesis as follows: that the number 137 on the Holy Crown and particularly on the image of Constantine, based on the physical background language of the Acausal Background Creative Intelligence (of the Pauli sort), gives with almost complete isomorphism the symbolic and mythical range of ideas and concepts of the connection between the fine structures, the light spectrum, the constant, the duplication of the spectrum lines and the fine structure constant.

3 The Base Structure of the 137 and the Kabbalistic Interpretation

It is also significant to note that the white colour of the 72 white pearls on the Latin crown of the Holy Crown expresses the Hebrew word for mercy, *Chesed* (the numeric value of which is 72), while the red colour of the 64 gems symbolises the Hebrew word for severe judgement, the attribute of divine power, the *Din* (numeric value = 64). These two concepts appear as the personification of God's left and right arm, as the 4^{th} and 5^{th} Sephirah of the Kabbalah's Sephirot system. Although there are other name variations, Scholem indicates them in exactly this manner in his Sephirotic tree (*Chesed*, *Din*, (see in [25]). The representation on the Holy Crown precedes the later Kabbalah colour conception by at least 200 years. Thus, this reinforces the likelihood of the origin from partially Khazarian sources

of the Proto-Kabbalah in Saint Stephen's court. The number of the 72 white pearls refers to Tamar in the Book Bahir and to the Hebrew word for pearl, "*Margalit*", in which both the name "Tamar" and the word "revelation" can be found, and which also represents the 10th Sephira, the Kingdom or the Crown (Atarah), in the Bahir as well (§ 72). The red gems are the messianic symbol for Judah's lion and the invincible Sun. The Hebrew word for stone is also an allusion to the messianic king. The conceptualisation of mercy and judgement as God's right and left arms played an important role in the Talmud and in the mystical Midrashim. In these allegorical depictions, King Solomon personifies the Lord, under the name *Melech Slomo*, the Talmudic interpretation of which is "the king to whom peace belongs". For example, in the *Shir Hashirim* (Song of Songs) *Midrash Rabbah*, the right arm is personified by the Archangel Michael, while the left arm is personified by the Archangel Gabriel. On the Greek part of the Holy Crown, this kind of the "Trinity" (Deesis) appears exactly in this manner, with a specific and shared epaulet edge system. Emperor Constantine is depicted here as King Solomon, with an angelic face, between Archangels Michael and Gabriel. (This is exactly the same manner of depiction as on the Pala d'oro. Figs. 3 and 4.)

Figures 3-4
(3) The archangels and the emperor with Solomon's angelic face on the Holy Crown [28]
(4) Solomon and the archangels on the Pala d'oro

This fact is very important to our hypothesis because on the right side of the Emperor, the number of white pearl decorations taken together with the Archangel Michael represent the number 72, while the same decorations on his left side taken

together with Archangel Gabriel shows 64 entities. The largest central white pearl is on the spiral on the Emperor's elbow, as on the mantle of King Solomon (the Hebrew name for Solomon also means mantle or cloak). Thus, the 72 white pearls of the C*hesed* and the 64 red gems of the *Din* of the Latin part of the Holy Crown is explained, according to Judaic mysticism, by the clear and unambiguous mantel decoration of the Emperor appearing as King Solomon with the two archangels. This structure completely matches the 72+64+1 construction of the pearls and gems. This connection, in addition to the likely origin of the Proto-Kabbalah, is important in the background language interpretation of the fine structure constant.

4 Some Historical Comments

Based on some of our former research, the image of Constantine was probably originally located on the front of the Holy Crown. Next to this image is Saint Stephen's father, with the name *Geobitzas (Γεωβιτζας, with a potential pun of "ζη Τωβιας"),* looking at the Emperor. Above them is the arch-shaped enamel picture of Christ. The picture of the Emperor has multiple symbolic meanings. On one hand it matches to King Solomon, as the equivalent image of King Solomon appears on the Venetian Pala d'Oro. On the other hand, given the appearance of Saint Stephen's father, Géza (usually Geysa or Geusa in the Latin texts), it refers to Constantine as being "the Purple-born." What is more, we have the situation of Saint Stephen's grandfather, Géza's father-in-law, Rex Iulus in Latin, as a Roman Patrician and Constantine's adopted son, adopting Byzantine Christianity and returning from the Byzantine Empire (cca. 950) to spread Christianity with his daughter and son-in-law. Taken as a whole, and with the angelic face – in the context of other works of art – this may depict the Emperor Constantine the Great with the "T"-letter labarum. Because the "Purple-born" Constantine, Constantine VII, was the seventh Constantine up to the time of Saint Stephen, here he may represent all seven of the emperors Constantine. The same can be said about the name *Geobitzas*, since István (Stephen) was the Christian name of Saint Stephen's grandfather, according to the old Church Slavic sources. It is generally accepted, and can be seen on ancient coins, that Géza's (St Stephen's father's) name was also Stephen, just like his son's name. So, symbolically speaking, the Byzantine and Pannonian concept of Rome is expressed in unity with the numbers three and seven, and we thus have the numbers 1, 3 and 7, an allusion to the number 137. Let us not forget that the Holy Roman Emperor Henry II, who was without a male heir, was the brother-in-law of Saint Stephen and thus the uncle of Stephen's son, Saint Emeric, who could therefore have been heir apparent to that (Holy Roman) throne. If we accept the potential, according to known historical resources, for a Byzantine wife for Emeric, then we could have witnessed at least the symbolic and genealogical union of the three Romes. And according to our hypothesis, the Holy Crown was prepared by St Stephen of Hungary for Emeric's coronation.

The unit 1 and this system (3+7) is also visible on the enamel picture of the Last Supper on the Pala d'Oro, where the special shape of the table takes on the symbolic shape of the Hagia Sophia, which was founded by Constantine. (see Fig. 5 [28])

Figure 5
The Emperor KωN (σταντινος), that is Constantine with the reconstructed Holy Crown with the Emperor and the King (see the details in [33])

Because of the quotation from Proverbs 9.1 which appears at the same place on the image of Solomon on the Pala d'oro, we can interpret this to mean that Sophia is building her house, and for it seven columns. Above the 7 columns of the table image, on the sketched Hagia Sophia picture, there are 3 edges visible. The name for these architectural elements in Latin is usually *corona*, while in Greek it is of course *stephanos*. The word "constans", or the name "Constantinus", suits well to the idea of the stability of the columns. The number of the ornaments of the edge of the table is 137 (Fig. 6).

According to our theory, these enamel pictures on the Pala d'Oro were made at the court of Saint Stephen, and the Holy King may have sent them to Venice for the occasion of the restoration of the house of Orseolo, the Venetian dynasty (in 1029), whose crown prince and heir to the throne was Stephen's sister's son and later King of Hungary, Peter Orseolo. The unity (1), the 3 and the 7 can be a symbol for the "*On-ogur*" (*Ungarus*) – the Ungaria, at that time without the "H" – to which came the 10 tribes, in other words, to the unity (1) of the 3 Kabar and the 7 Magyar tribes. Also linked to the numbers 1, 3, and 7 is the well-known 3 + 7 division of the Sephiroth system, along with the postulation of its divine unity. The work is consciously alive with these numbers. Further, it is certain that the court of Saint Stephen was knowledgeable of Arabic numerals. After all, it was

Pope Sylvester II who "crowned" Stephen in 1000, and it was the same Sylvester who before becoming Pope, as Gerbert, a famous mathematician, was a teacher to members of the dynasty of Saint Stephen's wife, Gizella, for example, to Otto III. It is well known that Sylvester was a vigorous propagator of Arabic numerals in Europe. [27] Thus, the 1, 3, and 7 numbers themselves, with respect to the decimal number system, refer to the number 137 automatically.

Figure 6
The "Last Supper" on the Pala d'oro and the reconstructed Holy Crown with the "Nikopoia"
(The references on Pala d'oro's pictures see in detail [29])

5 The Constantine Picture and the Leviathan

On the Constantine picture, along with the other pictures on the Pala d'Oro and the Holy Crown of Hungary, we can observe a special and decisively important interpretation of the number 137. This interpretation is able to resolve the conflicting accounts in the Gospels of Luke and Matthew of Jesus's genealogy and the virgin conception through the physical and spiritual deduction of ancestors up to Christ who brought about the incarnation. Thus, if we take together the first 20 ancestors from Adam to Abraham (which appear in Luke) with those from Matthew, we come to 62 forefathers (20 plus the original 42) who are transmitted to Mary by the Holy Spirit.

Based on the well-known tradition, Luke's genealogical tree in the Vulgate contains 76 names, from Adam to Christ through Mary. The ancestors from Adam (Abraham) to David are the same in both genealogies, while after David they diverge, coming through Solomon and through Nathan. Counting the only Christ, we get 137 out of the 138 entities[10]. Thus, via this interpretation, the incarnation of God is also indicated by the number 137. And in the age of Saint Stephen, on the occasion of their crowning and anointing, these priestly kings (*rex sacerdos*) spiritually received Christ (and so in effect the 137 incarnation. [5, 35] The Hebrew word for "receiving" is Kabbalah, whose number value is famously 137. Thus, the incarnation of God and the anointing and coronation of the Holy Roman king is given by the number 137. According to Jung, the equivalent of this receiving or incarnation in modern psychological terms is making conscious the self (*selbst*) in the process of individuation, and so in actuality, 137 is the *Selbst* archetype as well (see [30, 34]). As we shall see, this can be observed on the Constantine image, representing the fine structure constant in a "symbolic physical background language".[11]

---

[9] We have the same (original [32]) "system" on the Pala d'oro with the central 2+2 prophets whose names (Moses, Solomon on right side + David, Isaiah on the left side) written in Greek (the names of the other prophets are in Latin). Here the number of the characters 62+76=138 including 2 crosses (a Latin and a Greek). This interpretation is supported by the fact that the number of the letters is also 137. The 62 items are on Solomon's side (with Lex of Moses!) and the 76 items are on the Virgin's side (!) with Isaiah's text, the famous "virgo concipiet et pariet filium", can be considered a perfect isomorphic structure of the double genealogical system of Christ above. (see [32])

[11] Perhaps, it is worth mentioning Jung's opinion of this tradition: *"I'm rather certain that the sefirot tree contains the whole symbolism of Jewish development parallel to the Christian idea (concerning the incarnation of God). The characteristic difference is that God's incarnation is understood to be a historical fact in the Christian belief, while in the Jewish Gnosis it is an entirely pleromatic process symbolized by the concentration of the Supreme triad of Kether, Hokhmah and Binah in the Figure of Tifereth. Being the equivalent of the son and the Holy Ghost, he is the Sponsus bringing about the great solution through his union with the Malkuth (Atarah). This*

On the other hand, the Constantine image, along with the other images on the Holy Crown and the Pala d'Oro, contain the idea of the capture of the Leviathan, based on the line of Christ's ancestors, or in our case, on the number 137. We can see the same image in the Leviathan picture from the Hortus Deliciarum, which can be found in the illustrations of the Pauli dream series in Jung's famous *Psychology and Alchemy*. The Leviathan is the symbol of this unreflecting primordial cosmos, and its capture, marked with the number 137, is the symbol of the cognition of the cosmos through the number 137, the fine structure constant. (This questions has been discussed in detail in [36]).

The documented historical premises of the above conception of the Leviathan go back to Origen of Alexandria's description in *Contra Celsum*. [8, 27] Here, according to a Jewish-Christian sect, the Leviathan is composed of 10 concentric circles, which form a unit of 3 + 7 components. Thus, we find the one, the three and the seven in the ten. So, in a synchronistic sense, in a decimal system, the numbers 1, 3 and 7 can signify the number 137. Of course, that is precisely the structure of the Kabbalah, which appears nearly 1000 years later. Considering Biblical tradition, it seems that the Leviathan contains within itself all possible animal species, and in this sense can be considered a universal all-encompassing "genus memory". At least that is how Maimonides in the 12th Century conceived of the Leviathan, almost as an abstract concept. On the other hand, in the mystical Christian interpretation, the Leviathan contains within it all potential human souls, and the souls of the dead return to within the Leviathan until the time of Christ's redemption; in this way, the Leviathan symbolises both death and the afterlife. With Christ's redemption and the conquering of death ("*Mors perit in mort*e" [29]), the Leviathan releases the redeemed souls of the dead to the afterlife of Christ's kingdom.[12]

In the Christian depictions of the Leviathan, in addition to the fish and the snake motifs, there also exist those of the bird with wings and of the earthly being with strong legs. If we take the snake to be a fiery being (and the Christian allegories identify the red snake from the Book of Revelations with the Leviathan), with

*union is equivalent to the Assumptio Beatae Virginis, but definitely more comprehensive then the letter as it seems to include even the extraneous world of the Kelipoth. X (*almost surely Scholem, author's note) *is certainly all wet when he thinks that the Jewish Gnosis contains nothing of the Christian Mystery. It contains practically whole of it but in its unrevealed pleromatic state."* [12]

12 Here it is worth noting the Latin text on Hortus Deliciarum's illustration of the Leviathan. *"Postquam primus parens per lignum in pelagus hujus saeculi quasi in verticem naufragü corruit, atque avidus* ***Leviathan*** *seva morte totum genus humanum absorbuit, placuit redemptori nostro vexillum sancte crucis erigere, et hamo carnis sue squamea hostis guttura constringere, ut cuspide vitalis ligni perfossus evomeret quos per vetitum lignum improbus predo devorasset. Hec sancta crux est nobis lampas lucis eterne in hujus vite caligine, que suos sequaces ducit ad celestia, suis amatoribus confert gaudia angelica." [26]*

these various descriptions of the Leviathan, we can interpret all four elements of the world – water, earth, air and fire. Thus, the Leviathan here can be the material world's *prima materia*. The ten circles in the already mentioned 1, 3, 7 system, together with the material world circle, and thus symbolically with pi, refers to the Leviathan's geometric character. The alchemistic conception of the *prima materia*, with macrocosmic and microcosmic interpretations, could mean that the circle symbolises the atom as much as it symbolises the spatial position and motion of the stars. This latter is evidenced by the fact the Hebrew name for the Leviathan can mean (for the mystic interpreter – **לו יתן**) that he will give the 36, the 36 decans, just men who rise over the ten days, one by one. The three faces of the 12 zodiac, the 12 months, the 360 days and the year itself, namely in astronomical sense the space-time construction, can be brought into close relationship with the Leviathan. Thus, the Leviathan doubly contains the four ontological layers, since it carries within itself the intellectual and spiritual essence of humanity, while also carrying those creatures with soul from the animal world. In addition to this, it can be referred to as the material world and the transcendent background of all three ontological layers. At the same time, as a soaring winged bird flying in the air, the Leviathan is also a spiritual being. As an aquatic and land creature, it carries within itself spiritual and material substance. As a snake, the Leviathan can be comprehended as the personification of the transcendent world, as the allegorical pre-figuring of the Christian alchemical Mercurius.

Thus, we return to the number 137. The Leviathan contains within the 10, as 1, 3, and 7, just as Christ's genealogy contains the number 137; and the capture of the Leviathan is the physical cosmos, and the knowing and know-ability of its most important number, 137. At the same time, this goes together with knowing and know-ability of the universal spiritual world, the number of which – as the number of Christ's incarnation – is also the number 137.

On the enamel picture of Constantine, the 36 (LV-**לו**) hearts, the 50 (N-**ן**) white pearls, and, on the epaulet's edge, the letter T (**ת**) of the Labarum along with its 8+2=10 (I-**י**) decorative elements give – if we interpret the numbers as Hebrew letters – the word Leviathan in Hebrew (LVITN - **לויתן**). The 50 white pearls on the ornamental edge show a drawing of an anchor in reverse view. The Latin word for hook, *uncus*, in literary style can also mean anchor. Thus the hook-anchor and the hidden name, Leviathan, together symbolically represent the capture of the Leviathan, the conquering of death and the redemption (Fig. 5). In relation to Constantine I the Great, this refers to the concept of Roma Libera, while in the case of Constantine VII Porphirogenite "the Purple-born", it can be identified as the conversion of the "Magyar" people to Christianity (with the concept of Pannonia-Roma-Libera).

6 A Detailed Description and Interpretation of the Constantine Picture

As we have already analysed, the picture of Constantine can be seen as the form of King Solomon with the face of an angel. He is at the same time the symbol of the Lord, to whom peace belongs. His church is, naturally, the Sophia, the heavenly elevated Virgin Mary, the house of Wisdom, as we can see on the Pala d'Oro, and as we can similarly see on the mosaic picture of the Sophia's apse in Kiev. On the picture, next to the hidden Hebrew name of the Leviathan (as has already been discussed), there is the anchor, and the cross can be seen together with the 36 heart-shaped ornaments of Constantine's mantle. Thus, because the church of Sophia is attached to Solomon, the cross, the anchor and the heart, and the three attributes of Sophia (faith, hope, love), represent a very clear allusion to the Saint Sophia Church of the Christian Rome, linked with Constantine the Great. The identity preserved in the tradition of Constantine the Great and Solomon is also shown by the fact that Emperor Justinian, after having completed the reconstruction of Hagia Sophia, exclaimed, "Solomon (and of course Constantine), I have surpassed you!". On the Holy Crown, the holy depiction of the four contemporary Diocletian martyrs makes reference to Constantine the Great. Thus, as we have seen, Constantine VII "the Purple-born", as the seventh, along with the three Stephens, contains the system of 1,3 and 7 in the ten. According to a reconstruction undertaken in one of our previous works, the 10 pictures on the Greek section of the Holy Crown consist of 3 sections devoted to victory:

1 Mary with the two archangels, symbolising the incarnation and birth of the God-child. The Greek name of this section, Nikopoia, also contains the word victory.

2 Christ with the Roman Constantine ruler and with the king of the people of "Roma Pannonia", who are switching to Christianity (Pannonia-Roma Libera), which together represent a natural, victorious ensemble.

3 Between these two sections mentioned above, there are the four holy martyrs from the age of Diocletian, bringing forth through their sacrifice the Roma Liberia, which here represents Constantine the Great's victorious section. The 10 is the basic symbol of the victory in Kabbalah because (according to this mythologem) in the Red Sea at the moment of victory God revealed the Ten Sephiroth for the people of Israel. It is of course the pre-image of the Christian Easter which can be considered the greatest victory.

Another interesting structural feature on the Holy Crown can be obtained from the letters of the 8 + 8 = **16** "persons" of the Greek and Latin Crowns, respectively. This composition is probably related, at least partly, to the "fact" that the combination of pairs from 16 equals to **136** and the number of the sum 1+2+…+16, naturally, also equals to **136.**

Since the eight Latin names composed of 72 letters, taken together with the 65 letters of the names on the Greek crown, give us the number 137[13, 14].

Perhaps, it is worth noting that the 36 hearts by themselves symbolise the *Lulav*, the frond of the date palm tree, which explains the 2x36+32 system of the Latin section of the Holy Crown[15].

7 The 137 Genealogical System on the Enamel Picture of the Two Archangels and the Emperor

As can be seen on Fig. 7, on the edges of the shoulders and arms of Archangel Michael, on the edge there are **33** white, round, pearl-like decorations. On the emperor's mantle, the similar white pearl-like decorations, together with the central spots counted on the right side of the mantle, give us a sum of **41** items. On the left, on the shoulder and arms of Archangel Gabriel, the same **28** ornaments obviously symbolise the 28 names which follow David's in Matthew's Gospel account of the lineage of Christ. The number of the remaining decorations on the left side of the emperor's mantle is **34**, which represents the 34 forefathers from Adam to David, according to the Gospel of Luke. The single large pearl at the

[13] In the Holy Crown of Hungary there are **137** letters of the names: 8 Greek and 8 Latin names on the enamel pictures of the Crown: **ΟΑΡΧ ΜΙΧ ΟΑΡΧ ΓΑΒΡΙΗΛ Ο ΓΕΩΡΓΙΟΣ Ο ΔΗΜΗΤΡΙΟΣ Ο ΚΟΣΜΑΣ Ο ΔΑΜΙΑΝΟΣ ΚΩΝ ΓΕΩΒΙΤΖΑΣ (65 letters). SCS PETRVS SCS IOhS SCS PAVLUS SCS IACOBVS SCS ANDREAS ARTHOLO (sic!) SCS PHILIPVS SCS THOMAS (72 letters) i.e. 137 letters.**

[14] The separate texts on the emperor and king pictures contain 30+20 complete letters, to which are added the 4 letters each of the Lord and of the God-parent. The three hidden letters, the hypothetical letter "éta" of the word "Métér" (Μητηρ), the letter "nű" of the word "Porphyrogennétos" (Πορφυρογεννητος) and the letter "T" of the cross (both from the Constantine image) form a letter-triad, which in Greek equals to the number-statement of 358. This in Hebrew is the number value of the words Messiah and snake. Concealed within the letter "T" and the word "born" on the Constantine picture, as well as being contained within "mother of God", 358 very clearly refers to the birth of the Messiah. So this 137, 58, +3 system precisely matches the depiction system of the seven-column ornamental elements of the Hagia Sophia on the Pala d'Oro. The ornament system of the cupola, with the three edges or the crown, also leads to the number 137. Either of the two versions, 71+66 or 72+65 (GLGL ADVNI or GLGVL ADNI – **גלגל אדוני גלגול אדני**), means the reincarnation of the Lord.

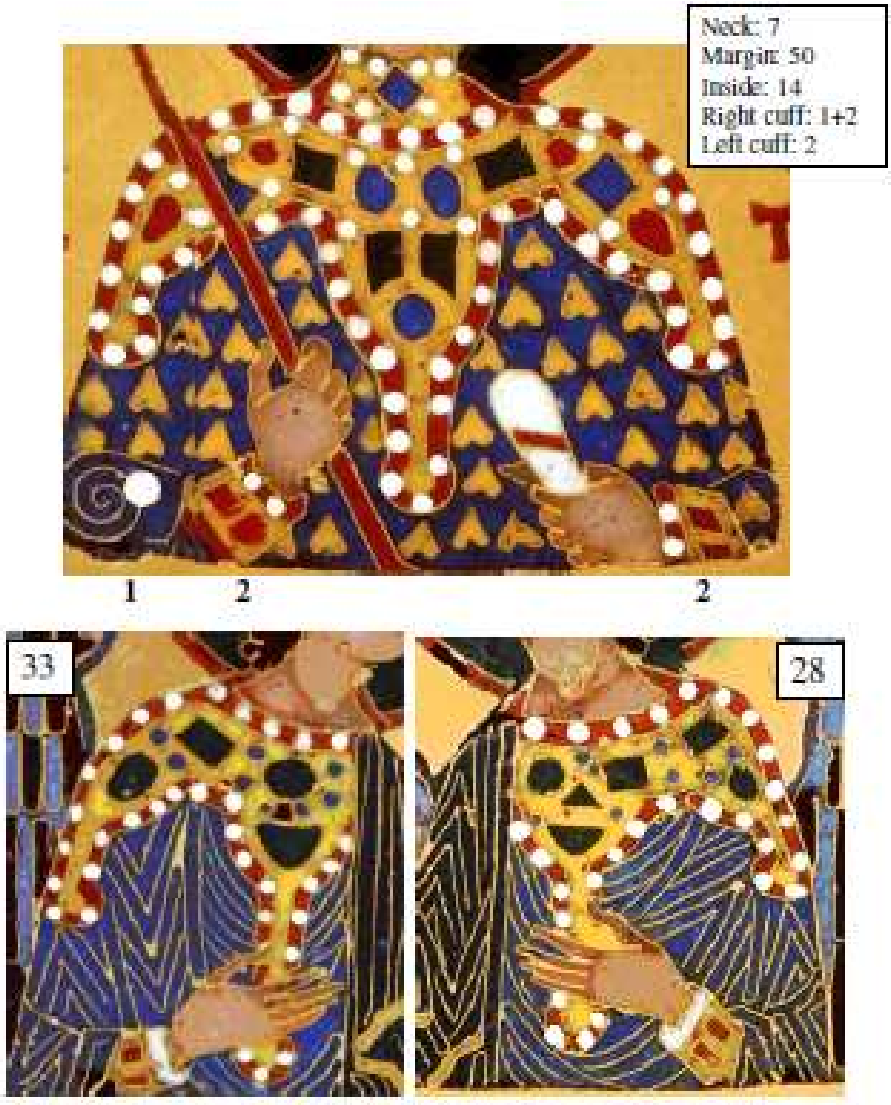
[15] The decoration system of the upper part of Constantine's mantle (on image of Fig. 7), precisely shows the name David in Hebrew, of which the central letter "vav" (which means 6) shows the diagram equivalent of David's star. This is surrounded by the 36 hearts; and the number value of the word "heart" (32) clearly refers to the 32 paths of wisdom.

emperor's right spiral[16] elbow, which has already been mentioned, can also be linked to the 33 or 41 decorations – which is obviously based on conscious planning (see Fig. 7). Basically, we can see the same thing here. If we consider the pearl to be David and add to this the 33, then we can interpret the image in one way as representing the genealogy (34) from Adam to David. If, on the other hand, we take the white pearl as representing Christ, then in another interpretation, we have the white pearl plus 41, representing the 42 forefathers starting from Nathan after David and through Mary to Christ. Thus, in the ambivalent 137 or 138 system, on the right until Christ in a 34+42 division, we can interpret 76 people altogether in Christ's lineage. On the left, similarly in a 34+28 composition, counting Joseph and Christ represented in the Holy Spirit, we can identify 62 forefathers altogether. This system ends the conflict between the two accounts of the lineage of Christ, "realizing" or "satisfying" the dogma of the virgin birth of Christ from Mary. Thus the right side means (in mirror-symmetrical depiction) the corporeal and merciful side of the incarnation. And thus the corporeal, merciful and forgiving lineage forms a unit. This line from Adam to David may represent the bodily and merciful attributes of the ancestors. On the other side, the Holy Spirit, representing the spiritual aspects of the 62 forefathers, paired together with the Law, allows God to expand His impact through the Holy Spirit in pursuing His divine "plan of incarnation". In this sense, mercy prevails in the incarnation through the body of Mary, while the law is realised through the spirit of Joseph with the mediation of the Holy Spirit. So the merciful/corporeal process of incarnation and the spiritual law incarnation process from Adam to David divide into two after the messiah king, David. And from David, the incarnation comes through Solomon and through Nathan to Christ, which means the 137th number of the incarnation.

We would like to emphasise again that the multiple contradictions regarding the virgin birth conception and between the two genealogies are fully resolved in this interpretation; and this system leads to a consistent, complex receiving of the incarnation. And, as we have mentioned several times and as is well known, the number value of the Hebrew word, Kabbalah, the meaning of which is "receiving", is 137. Certainly, the same is true for the Hebrew word for "crucifixion" – *celibah* – that which ensured the redemption (CLIBH, 137=**צליבה**). We would like here to emphasise again that the system described above in all details matches that of the Hortus Deliciarum's capture of the Leviathan, with the help of the line of Christ's incarnation lineage. (see [36]) The same also applies to Constantine the Great, who with the symbolic capturing of the Leviathan, or in other words, the defeat of evil, set free the church. And the same returns as well on the 1000th anniversary of Christ's redemption (which is at the same time the "*Resurrectio Prima*" according to St. John's Revelations), at the coronation of the

[16] It is an appropriate symbol here as the singular 137th for the snake Leviathan and the birth of the Messiah as well. (see Jung's comments on the dream 32 in [8])

new king, together with the capture of the Leviathan and the symbolic reincarnation of God.



$$7+50+14+5 = 76;\ 33+28 = 61;\ 76+61 = 137$$

Figure 7
Illustration for the 137 structures of the "white pearls" on the mantels of the archangels and the emperor

8 The 137 Structures on the Enamel Picture of Constantine and the two Archangels

8.1 A Short Summary of the Analysis to this Point

As we have seen, the 137 structure on the Holy Crown of the 72 white pearls and 64 red gems together with the leaning cross defines the base structure. Because this is located symmetrically on the two bands together with the centre, it can be denoted as a 68+68+1 structure. Looking in more detail, (36+32)+1+(36+32)=137. Adding the conjunctive unit (X) to the number 72 or the number 64, we get two variants: 72+65=137 or 73+64=137. The former means the mercy of the Lord; that is, the Lord's 72-letter name, or the Lord's 72 structure. The latter, meanwhile, means the (Solomonic) judgement of wisdom (64+73 = 137 = **דין חכמה**). (See Bahir § 64 [1].)

As we have already mentioned, the 72 Latin letters (of the eight apostles) and the 65 Greek letters (of the names of the 8 individuals on the Greek part of the crown) reinforce the base scheme 72+65=137. The mentioned 59+8+2+59+8 system of ornamentation on the front upper part on the Greek crown, which separates the two parts of the centre, repeats the order of 68+68 (+1) = 137.

If we look again at the edges of the shared system of mantles of Constantine and the archangels (priestly Ephod's edge-SNI - **שני**), the white pearls of the system, if we consider Constantine and the angels separately, also show the 76+61=137 system.

Let us consider the three pictures divided into left and right sides, with an even central ornament separation. In this case, we reach the 72+64+1=**137** system, which matches the structure of the 72 white pearls and the 64 red gems of the Latin crown. The 72, the *Chesed* (**חסד**=72), the white colour of mercy, is here being replaced with the right side (with God's right arm), according to the well-known Midrashic tradition and the broad interpretation of the Kabbalah. The 64, the *Din* **(דין**=64), the red colour of strict judgement, is replaced with the left side (the left arm of God), according to the same tradition. At the single, individual white pearl at the emperor's elbow, as we mentioned, a spiral decoration can be found. If we count the pearls of the archangels (61) together with the neck items of the emperor's mantle (7) we may identify a 68+1+68 construction again. Counting Michael's 32+1 (individual) pearls separately from the "Emperor-Gabriel construction" we can identify a 104 +32+1=137 structure, where the 104 is the number of the Hebrew word *kokavon* (**כוכבון**), which is the Asterisk (αστερισκος)! (A similar letter-number structure can be found in § 112 of Book Bahir.)

Let us consider for the sake of the completeness once again the left/right division of the white pearls of the "pontifical mantles" of Emperor Constantine and the

Archangels based on the dominance of the right side. In this case, we count the ornamental items in the centre as being on the right side. And in this case, therefore, we can count as being on the right side 75 white pearls, on the left 62. A more detailed partition, with the addition of the individual pearl, is 34+41 on the right side, 34+28 on the left side. This brings us to the two-pronged incarnation genealogy of Christ, one of which is the corporeal line up to Mary (34 from Adam to David, and 41 from David's son, Nathan, to Mary (Joseph). Adding the individual white pearl, we get either 34 from Adam to David or 42 after David's son Nathan to Jesus. On the left, with the addition up to Adam, we get the spiritual genealogy of 34 from Adam to David and 28 from David's son Solomon to Jesus Christ.

8.2 Additional 137 Structures

From this point forward, we will concentrate only on the complete system of decoration or ornamentation of the picture of Constantine. In addition to the white pearls, we will consider the pearl and "stone decoration" of the mantle, the labarum and the crown of the Emperor.

Here we focus separately on the decoration of the 36 golden hearts seen on the blue base of the mantle. As we have seen earlier, taking into consideration the double twin-like depiction order of the hearts, we interpret the 36 doubly, and in the same way we interpret the dual meaning of the Hebrew word for heart, which also means the value 32. By itself, then, this allows us a 68+68 = (36+32) + (36+32) interpretation, which together with the imperial red-blue sign could lead to the 137 interpretation.

The additional decorations, aside from the 76-pearl decoration already mentioned, are complemented by the 31-white-pearl decoration of the crown of the Emperor together with the 8-white-pearl decoration of the labarum's letter *Tau* (T). Further, we find column-like gem decorations on the imperial mantle (14), on the cuffs (2+2=4) on the crown (1+1=2) and on the labarum (2). Thus, aside from the 76 white pearls, we can count together exactly another 61 ornamental items.

If we consider these separately, we get the 76+61=137 structure, in manner similar to the – already analysed – triple "archangel-emperor mantle-edge" system.

If we approach with a left-right division, with the central items paired equally with both sides, we also achieve the 75+1+61=137 construction. To be more precise, we can distinguish on the right side a 34+42 (34+41+1) structure and on the left side a 34+27 structure, with right side dominance for the pearls of the mantle and a left side dominance for the remaining items (that is, with the central decorative elements counted on the right or on the left side). This structure is in mirror-symmetry with the triple system, that is, with the structural order which means God's incarnation.

If we consider the left-right division, mirror-symmetrically in accordance with the triple system, with left-side dominance (that is, with the central decorative elements counted on the left side), and with the distinction of the individual pearl at the imperial crown, then once again we achieve the 72+1+64=137 system. If we count the two individual pearls (at the elbow and the crown on the left side, we obtain the 71+66=137 construction which can be seen on the Pala d'oro (see above).

Thus, if we consider the mantle-system of the three pictures together or the decorative system of the Constantine picture alone, we can find 3+3+1 different versions of 137, a fact which in itself is reference to 7, or to Constantine VII, whose epithet was "Purple-born" independent of the original concept. From the analysis above, we can also establish a base case that, based on the three enamel pictures, we can find three 137 systems, the differentiation of which gives us seven variant interpretations.

On the pictures of the two archangels, the additional, as-of-yet unused decorative elements form a 32 system. This system is comprised of the central, red, circle-like ornament of the hair bands – looking from up to down – and the inner decorative elements of the mantles, which are supplemented by the decorative elements of the cuffs (two each) to equal 32. On Michael's mantle, we can see 8 blue circles and 4 column-like decorative elements. On Gabriel's mantle we can also identify 8 blue circles and 6 column-like elements. Thus, with Michael we get the 8+2=10 and the 4+1=5, while with Gabriel we get the 8+1=9. Synthesising these, with the numbers 9, 6, 2, 10 and 5 we get precisely the Hebrew name Tobias (**טוביה**), the numerical value of which is of course 32. If we consider that 137 is the number value of the Hebrew for wheel, *ofen*, which at the same time always means "crown", as the wheel is the archetype for the crown, then we get a clear connection between the name Tobias and the crown. (The significance and various interpretations of this, along with the Biblical background and tradition, we have dealt with in detail in previous papers [32, 35].) It can explain the poetic, creatively deformed, name of St Stephen's father. Here the transformation of the usual written form of Geysa (in Latin) or Γειτζας (in Greek) into Γε**ωβ**ιτζας is an anagrammatic allusion to the expression Ζη Τωβιας (where η = ε + γ – 8=5+3), i. e. "Live Tobias!" which is a sign of victory over death, i.e. capturing the great fish, the Leviathan. This name could be very important for the medieval author and interpreter because, beside being the hero who captures "Leviathan", in Vulgate (Zach. 6, 10-14.) Tobias is a maker of the double – gold and silver (kingly and priestly) – crown as well (see still in [35]).

In addition to the decorative system containing 32 items discussed above, we can also consider as decorative items the 3 paired bands decorating the hair of the archangels, along with the circular red ornament in the middle band above the forehead. These bands, together with the red ornament, form a quite frequent (menorah-like) 3+1+3=7 decorative system. We can also take into consideration ripidon's (fan's) colour decorative item consisting of three elements (see Fig. 6).

Thus, in addition to the 32 Michael and Gabriel items discussed so far, we can also consider another 6+6+3+3 decorative elements. Since the number of the decorative elements of the imperial enamel painting is, as we have seen, 137+36 (36 gold hearts on the blue mantle background), or a total of 173, and since furthermore the decorative items of the two archangels is 61+32+18=111, then we can define a total of 284 decorative items altogether. The total number of 284 is definitely an allusion to the number value of the Hebrew word for crown, *atarah (284=***עטרה***)*, which, as we have seen, always equals the number archetype 137 and the name Stephanus as well. Thus, the decorative system of the three enamel pictures refers to the royal crown. The Greek translation of the Hebrew word *atarah* is ***Στεφανος*** and the Latin translation is *Corona*.

To conclude this section, we point out that the construction of the system of symmetry of the decorative structure allows us to observe some isomorphic relationships. On the decorative edge of the *efod*-like shoulder epaulets of the emperor and the two archangels, we can count **111** white pearl decorations. (This in Hebrew is *Alef*, which can mean either "one" or "thousand". In the Bahir, *Alef* is the Kether (**כתר**), i.e. the upper crown of the Creator, the first Sephira. The number value of *Alef* -**אלף** is 111.) As we have seen, the number of the archangels' total decorative elements is also 111. The number of decorative items of the edges of the *efod*/epaulets of just the archangels is altogether 61, while in the picture of the emperor, in addition to the 76 white pearl decorative elements on the mantle, we can find another 61 decorative elements. On the archangels' mantles, not counting the elements on the edges, the number of decorative element is **26**. On the mantle of the emperor, again not counting the items on the edge, the number of decorative elements is also 26. Both of the previous **111** decorative systems take the 61+50 structure. The **111+26** (=137) i.e. Alef YHWH (137= **אלף יהוה**) could mean the Crown of the Lord and at the same time the Millennium of Lord, i.e. the Coronation in the year of "*Resurrectio Prima*", thus 1032/33 according to the well-known medieval chronology based upon the Gospel of Luke. (See some further details concerning "111+26=137" in the Appendix, especially in the footnote 18.)

The 284 system, the 32 system which refers to Tobias and the 137 system are all similarly visible in the structure system of the chapter titles in the *Admonitions to Emeric* [35].

9 Comments on Methodology

The Pauli-type unconscious, creative background processes, together with the connected imaginative and synchronicity series, can be interpreted as Pauli's Acausal Background Creative Intelligence, according to the approach of intelligence systems researchers. This intelligence entered the unconsciousness of

our “modern” age, according to Pauli, and the “undetached observer” can identify its “dynamics” only through the systematic hermeneutic interpretation of his dreams, visions, active imagination and synchronicity phenomena, applying the symbolical background languages. With the fine structure constant and in connection with its crown symbolism, this function of symbolic background languages, according to us, can be seen almost identically in the symbols and hermeneutical allegories of the Holy Crown of Hungary (discussed above) and in the events of the Jung-Pauli cooperation, as well as in their imaginative, reflective and psychic contents.

We judge as significant the 137 number archetype structural symbolism of the incarnation because, according to well-known Christian tradition, the incarnation is considered the second creation. (Jung often emphasis the significance of this.) Similarly, Pauli, in the symbolic language of background physics, describes the duplication of the spectrum lines together with the number 137 belonging to them, while further describing the interpretation of the isotopes as a “process of incarnation” or a “process of creation”. As we saw in the introduction, MacGregor, together with other physicists, talks about the “physical” creative power of the alpha, that is the FSC. Thus, in the course of analysing the work of art, in this sense, the incarnation motifs also refer to the theological and physical symbolic background language, as we have shown in other studies, together with the psychological symbolic background language. The presence of the guiding and symbolic hermeneutical background language is obvious, both from the artwork’s aesthetic aspect and from its well-known governing (that is, guiding) aspect of the Holy Crown of Hungary as an allusion to the symbolic background control language.

Based on the above, we may suppose that these creative background processes can in part be identified on the basis of the partially observed output and the synchronistic processes. In the case of a system identification, in a broader sense, when searching for typical patterns and symmetrical structures in the dynamic structure of the background process, based on Jung’s concept of synchronicity, we can identify historically documentable events and works. Naturally, this latter hypothesis was also introduced by Jung.

In this case, the question arises whether the complex data lines appearing in the observer’s dreams and visions and their connected synchronistic series call to the observer’s mind the patterns and pictures of created works of art which had come into existence earlier in history, when these background processes were moving along the border of the conscious and the unconscious in their early, vigorous and unreflected or weakly reflected differentiation. Since the main observer and gauge for us is primarily Pauli, along with Jung, one of the main characters in the background processes leading to the fine structure constant is the concept of the incarnation. For Pauli, the incarnation is connected to the divine or messianic twins, and, at the same time, as we mentioned above, it “belongs to” the doubling

of the fine structure's lines and the isotope phenomenon. For Jung and Pauli, the *incarnatio continua* is always and at the same time the *creatio continua* too.

We consider now only two examples:

(**1**) In 1927 Jung knows nothing of the FSC or of Pauli, who was at this time in Hamburg as a professor of physics. However, his surely most famous and important Mandala (the "Window into Eternity") was composed in early 1927 on the basis of his also famous "Liverpool dream" related to the death of his "twin brother friend", Hermann Sigg (see in detail [37]). The structure of this mandala (see Fig. 8) is an ideal composition of the number archetype 137 and at the same time it is entirely isomorphic with the main structures of 137 on the Holy Crown of Hungary. It has a 68+1+68 and 72+1+64 structure, as well as 68, which can be composed by the sum of 36 and 32, etc. It is really a twin structure of 137 and a perfect manifestation of honour and memory for a loving twin brother relationship in the case of a suddenly deceased very close friend. Furthermore, this 137 mandala's structure corresponds to Eddington's later concept of "128+8+1=137" as the "true expression" of the fine structure constant.

(**2**) Jung's mandala above can be compared with Pauli's "crowning dream image" visualized five years later. As we have shown in our papers in detail, Pauli's World Clock vision as a space-temporal mandala is a properly manifold complete representation of FSC and 137 as a number archetype of the Self. From the Pauli–Jung letters [17], it is clear that the Black Bird is supported by the female symbolic figure of *anima*, where the anima is adorned with the number '7'. In this way, the black bird (1), the rotating discs and ring (3), and the figure of the anima (7) altogether connect to the number '137'.

Furthermore, the temporal structure (rotation scale) itself is double 32, and the spatial structure is also evidently 2×(32+4) (the four little men with pendulums – horizontal disc, and four cardinal positions of pointers – vertical disc), that altogether results in 1+2×32+2×36=**137.**[17]

[17] Pauli's dreams No. 39 and 51 from Jung's *Psychology and Alchemy*, naturally and evidently contained in the structure of the World Clock vision (four little men, four colours create an isomorphic map with the four eyes and four colours of the bear). Therefore, the three power-like rhythms with the four space-like quaternio (like a multiplicator) is naturally and evidently an isomorphic map of the $4\pi^3+\pi^2+\pi=137,036\ldots$ Consequently, the structure of the World Clock vision (together with dreams No. 39 and 51) is a perfect isomorphic structure of the above "abstract" formula of the fine-structure constant and its discussed isomorphic interpretations, if $\pi\sim4$ or $\pi\sim2$ (see the details in [8]).

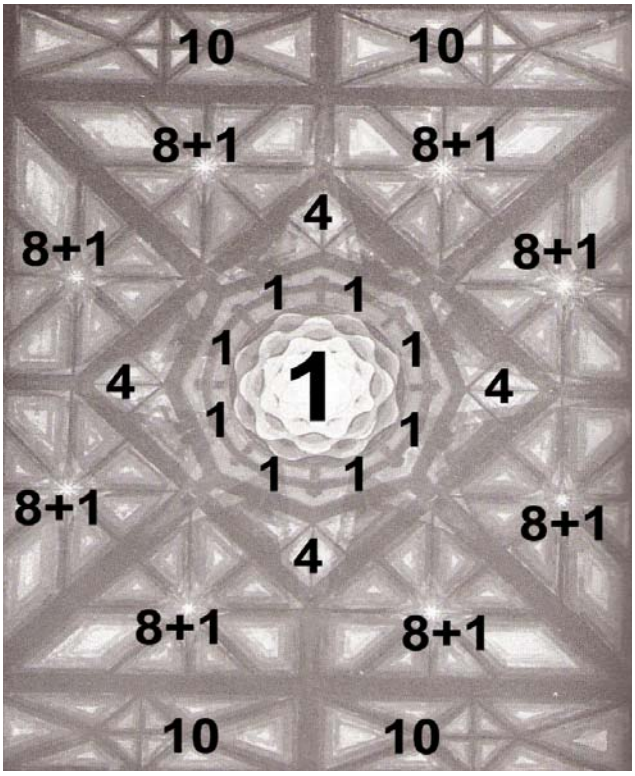


Figure 8
The „Liverpool Mandala of Jung from 1927 [10] illustrating the number of sections (quarters according to Jung)

Now, if we compare the structures of the "World Clock" and the Holy Crown as well as their number-systems we may recognize an almost complete isomorphy between them. The vertical two semicircles of the Latin crown can correspond to the vertical disc of the "Clock". Similarly, the horizontal circle of the Greek crown can correspond to the horizontal disc and golden ring of the "World Clock". The basic number structure is also the same, i.e. $1+2\times32+2\times36=$**137.**

Furthermore, the whole double system of the World Clock can be considered as a Crown of the Black Eagle, which is a symbol of the cosmic Anthropos and the King as well. In his book Jung interpreted and illustrated exactly this meaning of the "image" as a convincing synchronicity. (The fact that footnote number of this dream in the English edition is 137 is another synchronicity) Here, the twin Black Eagles as the temporal and eternal kings are bearing the double crown. So, both the first and the last dream image contain the archetypal twin crown linked strongly to the concept and number of FSC. The whole image of these symbols can be considered an almost perfect representation of the Holy Crown of Hungary

together with the hidden unconscious recognition of FSC in St Stephen's crowning achievement.

Figure 9

(a) The World Clock. An impression generated by W. Byer Brown based on accounts of Pauli's dream. (b) The alchemical figure of the rising double or "twin" black eagle is hatched wearing the spiritual and temporal crowns as cosmic Anthropos and eternal King. Codex Palatinus Latinus (15th Cent.). An illusration chosen by Jung for Pauli's "Great Vision" [8, 11]

MacGregor, bringing in the allegorical crown symbolism, connects the fine structure constant – along with other people – to the concept of *creatio continua* and *incarnatio continua*. Therefore, as regards the FSC's background processes, it is necessary to look for the living theurgical motif of the *incarnatio continua* and the *creatio secunda,* which truly existed in the royal theology and the concept of the ideal king in the 10th and 11th Centuries. [5, 35] Considering the perspective of synchronicity, it is worth noting as well the Latin expression for the fine structure, which through the words *subtilis* (*subtilitas*) and *structura* is strongly linked with medieval hermeneutics. Here it is valuable to consider a work with aesthetic characteristics that include equally the theological, philosophical, legal and artistic hermeneutics of that age.

In this case, the *Constans*, that is the *Constantinus*, of Saint Stephen's exceptionally fine-structured enamel pictures, standing in the centre of the culture of the number 137, together with the represented 137 structure-numbers, shows us the ancient concept of the fine structure constant hidden in the creative background processes. Here, the number 137 of the messianic incarnation connects, or rather appears, in an inseparable unit with the number 137 of the physical background's fine structure. What gives decisive momentum to this theory is the fact that the name of the person on the picture who represents the psychological, hermeneutical (*subtilitas*) and physical background of the symbolic theology is Constantine, who, as Christ's monarch, can be linked to the symbolic background language of systems and controls. The other side of the synchronicity is hidden in the proto-Kabbalistic meaning system of the picture, which is thus the

anticipation (pre-cognition, or *retro*-cognition according to our hypothesis above) of Sommerfeld's conceptualisation of FSC, which is, as is well-known, strongly affected by the Kabbalah.

It is also well-known that Jung identified psychological manifestations of the self (the *Selbst*) with the mandala, and para-psychological manifestations with synchronicity. In another of our articles, we demonstrate that the Jung-Pauli dreams and their own interpretations of these dreams represent an acausal background and control theory, or more universally, an Acausal Background Creative Intelligence, which is manifested in the mandala and in synchronicity [37]. Thus, instead of the para-psychological considerations, such as pre-cognition, retro-cognition, etc., we can also evaluate from a control theory aspect the series of manifested psychic contents and synchronicities. Based on the above, we were led to the conclusion that a hypothetical Acausal Background Creative Intelligence in the potential of the synchronicity series expresses itself as an acausal synchronistic control system together with its own 137 structure [54]. This at the same time matches the mandala dreams discussed as the psychological manifestations of the *Selbst* and the para-psychological and synchronistic manifestations of Selbst by the number 137, via its own (self) number archetype. According to our hypothesis and hopes, we can speak about some sort of similar "phenomenon" here, at least in the hermeneutical circle opened in our paper and now closed.

Conclusion

In our discussions above, we have intended to show some crucial patterns of the archetypal concept of FSC and Number archetype constructions of 137-structures in the long-range synchronicities between the hermeneutical system of the Holy Crown of Hungary and the spontaneous dream images and their interpretations manifested from the unconscious in the course of the Pauli-Jung collaboration. Our entire endeavour focused on identifying some characteristic features of Pauli's creative Background Processes, reformulated as a hypothetical acausal background creative intelligence (ABC intelligence), which operates in the unconscious mind.

Appendix 1
The Schematic Enumeration and some Details of the *137* Structures on the Holy Crown

A1.1 The 7 Additional *137* Structures

1 The 72 white pearls and the 64 red gems with the slanting cross found on the Latin crown of the Holy Crown.

2 The 72 Latin letters of the names of 8 apostles on the Latin crown, together with the 65 letters of the names of the 8 people on the Greek crown.

ΟΑΡΧ ΜΙΧ ΟΑΡΧ ΓΑΒΡΙΗΛ Ο ΓΕΩΡΓΙΟΣ Ο ΔΗΜΗΤΡΙΟΣ Ο ΚΟΣΜΑΣ ΔΑΜΙΑΝΟΣ ΚΩΝ ΓΕΩΒΙΤΖΑΣ (65 letters).

SCS PETRVS SCS IOhS SCS PAVLUS SCS IACOBVS SCS ANDREAS ARTHOLO SCS PHILIPVS SCS THOMAS (72 letters).

3 The 59+59+8+8+2 decorative elements of the bands decoration, forming together with the cross a 137 structure.

4 The "leaning cross" equals the letter *Aleph*, containing God's name (IVI=IHVH), thus representing both the number 26 and 111 (111+26=137). This representation and interpretation can be found in § 70 of Book Bahir on the basis of Micah 2, 13. Now let us see Kaplan's comments on this part of Bahir which corresponds to Scholem's interpretation of the "Alef equals to the Tetragrammaton YHVH": *"As already discussed, Alef represents Keter-Crown (see 15, 17, 26, 117, 140)" The Alef also consists of a Yud on the upper right, a Yud on the lower left, and a diagonal line separating the two Yud's (see Fig. 10)*

Figure 10
The name of God as YHVH (IVI) and Alef as well as the schema of **X** as old Hebrew Tav and Christ's monogram (after Kaplan [1])

(The upper Yud represents the second and third Sephira while the Yud on the lower left represents the Kingdom i.e. the tenth Sephira.) *The line has the shape of a Vav, which represents the six intervening Sefirot. The two Yuds each have a numerical value of ten, while the Vav has a value of six. Therefore the two Yuds and the Vav have a total numerical value of twenty –six, which is also the numerical value of the Tetragrammaton YHVH. The verse (Micah 2, 13), "YHVH at their head (Alef)," is therefore introduced." [1]*

As we have seen Alef+YHVH= 111+26=**137**. But the Aleph means as a word the number 1000 in Hebrew. This kind of Aleph equals to **X** which is at the same time the old Hebrew TAV, i.e. the sign of redemption. Naturally it can mean Christ's monogram and in Latin it may mean the number ten (which is valid for the kingdom and the Sephiroth as well) and in Greek the number thousand, respectively. Thus, together with the 32 decorative entities on the upper picture of Pantocrator, it is a strong allusion to the intended (coronation) year of 1032/33 (Resurrectio Prima) as well.[18]

5 The picture of the apostle James, on the right and left side, showing a naturally interpretable number system consisting of 68+68 decorative elements. The symbolism of the apostle looking upwards to the cross, together with the cross, indicates the 137 structure.

---

[18] The cited original Hebrew verse of Micah 2.13, in the hermeneutical context of Bahir, can afford a good opportunity for a mystical interpretation of the coronation of a King together with the symbolic "137" incarnation of the Lord into Him (according to idea found in [5]):

עלה הפרץ לפניהם פרצו ויעברו שער ויצאו בו ויעבר מלכם לפניהם ויהוה בראשם

Namely, the Hebrew noun and verb peretz (to break breaker – פרץ פרצו) naturally can be understood as the name of Pharez (פרץ), son of Tamar and Judah. From another point of view, the Hebrew word עבר in the verse has a connotation of a kind of reincarnation or overshadowing of the Righteous, as "ibbur" in the (later) Kabbalah. The ויהוה בראשם (which, as we have seen, leads us to the conclusion of **137** = אלף יהוה) can be understood partly as a coronation by the Alef=Kether, partly as an incarnation of the Lord into the Head (i.e. the King) of his people (the Lord-YHVH- is in their King or in the their Head). Furthermore, according to the § 70 in Bahir the Alef is not only the Kether but the Holy Palace (shrine) and number 1000 as well. The essence of the allegorical pattern is the personification of the denary Sephirotic system by the first ten princeps of the tribe of Judah. The first "sephira" is naturally Pharez (son of Judah) the "Breaker through", the Pioneer, while the tenth is David completing the joint work of his (9) forefathers with the foundation of the "eternal" Kingdom. The transfer through the Gate may correspond to the entry into the (Sophia's) Apse of the Coronation. In this case the mystical interpretation of Micah 2.13 in the context of Bahir, is as follows: "Pharez (who breaks through) will go up before them, His (God's) Pharez and they pass the gate. Inside they satisfy their duty (the coronation ceremony – here we used another meaning of the verb יצא'). He (Pharez) overshadows their King who is before them. The Lord is in their Head (=King) by the Crown of the Lord in the 1000th year of the Lord's redemption". Naturally, it means the incarnation through 137. Prophet Micah has a distinguished place on St Stephen's Casula (or mantle) above the King himself as Sephanvs Rex.

6 The enamel picture of the apostle Peter also indicates a similar system, consisting of 68+68 decorative elements.

7 On the picture of the apostle Andrew, we can clearly identify a 72+1+64=137 structure.

A1.2 A Detailed Review of the 137 System of Decoration of the Bands of the Greek Crown as well as on the Pictures of James, Peter and Andrew of the Latin Crown

1 **A detailed description of the decoration of the bands**. On the left and right sides of the front of the Greek crown, we can observe a symmetrical system of triangle and arch-shaped decorations, 4 each. The frames of the two triangles and two arched decorations, together with the gems found on their tops, we can consider as 2x(4+4) elements. The frame of the picture of Christ in the middle with gems on its top we can also consider as 1+1 elements. The Lord, himself, or the letter X of his name means the centre. From the front, the outermost arched decoration on each side contains 14 elements. On the next 3 decorations on the bands, moving forward, we can count 7 elements. On the following, larger arch-shaped decoration we can count 23 decorative elements on both sides. The triangular decorations surrounding the Christ show 15 elements. Thus, the entire system gives (14+7+23+15+4+4+1) +XP(1) +(1+4+4+15+23+7+14) = 68+1+68 = 137 decorative elements. We can see that inside the parenthesis we can apply the 14+7+15=36 synthesis, while the number of the remaining elements is 23+4+4+1=32. Thus, the structural system can be described, according to the system of the white pearls and red gems on the Latin crown, by the 72+1+64=137 structure.

2 **The enamel picture of James**. On the enamel picture of James, we can find at the top 4 white quaternion decorations, on the right side 4 and on the left side 5. Thus, their total number is 16, or 20. The number of red decorative elements on the apostolic image is 3x6=18, on the right and left sides equally. The number of the white decorative dividing elements are 20 each. The number of the additional elements on the upper part on the right side is 6+3, and on the left side 4+3. On the right the “tree”, likely referring to James’s staff, contains 3 decorative elements. On the left side, we can find a decorative element in the middle, a rhombus-shaped decoration above and a circle-shaped one on the lower left part of the mantle. Thus, we get a system consisting of 68+68 elements in total.

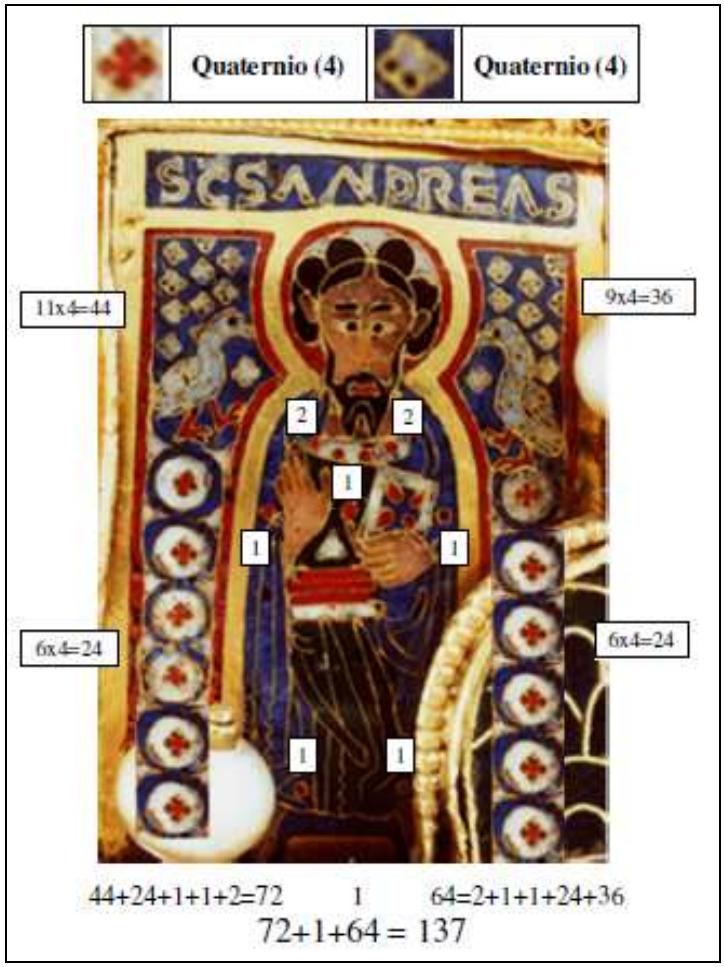


Figure 11
The illustration of the 137 structure on the reconstructed enamel picture of Andrew of the Latin crown

3 **The enamel picture of Peter**. Here uppermost we can find 12 quaternion decorations on the right side, while on the left we can find 11. Thus, the number of decorative elements is 48, or 44. The number of the square and circle-shaped red decorative elements is 12 each. The number of the decorative elements under the quaternions on the right side is 6, while on the left side it is 9. The number of elements on the body on the right side, counting the double centre on the right, is 2, while on the left side the number is 3) taking into consideration the flame of the candle). Thus, we can observe 68 elements each on the left and right sides.

4 **The enamel picture of Andrew** (see Fig. 11). Above we can find on the right side 11 white quaternions, on the left side 9. Thus, the number of all the elements

here considered is 44 or 36. On each of the right and left sides of the apostle there are 6 red quaternions. (The total sum of them is 24+24) On each of the right and left sides we can also find 4 red circle-shaped ornaments, two on each side of the neck, and one at each hand and foot. The same sort of gem-like ornament can be found in the centre, to which the index finger of the apostle's left hand is drifting. Thus, we can determine on the right side in total 72 elements, on the left side 64 (44+24+4=72 and 36+24+4=64). And so we determine again a 72+1+64=137 system, according to decoration system of the white pearls and red gems on the Latin crown.[19]

The picture of Emperor Constantine, as we have demonstrated, on its own and together with the two archangels also includes in ***seven*** separate variants the several meaningful ***137 structures***.

Appendix 2
The Structures of 137 on the Coronation Mantle (Casula) and the Royal Sceptre

A2.1 The Royal Priestly Casula (The Coronation Mantle)

1 The four mandorla showing the letter *Tau* from behind contain 137 letters in total in a 72+65 division.

EMICAT IN CELO SANCTAE GENITRICIS IMAGO / DAT SVMMO REGI FAMVLATVM CONCIO CELI (65) HOSTIBVS EN XPISTVS PROSTRATIS EMICAT ALT X / SESSIO REGNANTEM NOTAS ET XPM DOMINANTEM (72)

2 The circular inscription contains on the right side 72 characters, on the left 64, together with the central image of the cross (or the interpretation of the double ® giving the usual 137 structure.

®EGINA HOC CASVLA OPERATA ~ ET DATA ECCLESIAE SANCTAЄ MARIAE SITAE IN CIVITATE ALBA : (72 characters)

ANNO INCARNACIONIS XPI : M : XXXI : INDICCIONE XIIII : A STEPHANO REGE ET GISLA (64 characters) ®

[19] Taking into account the tradition that Saint Andrew was the chief priest and apostle of Scythia, the Latin-Greek symmetry is a hint to the union of the western and eastern Dardanian Royal Houses using the number archetype of 137 as "Coding Archetype." (see [32, 35])

3 The Coronation Mantle, or to be more precise, the entire shape system of the conical *casula* displays 68 shapes each, apart from the centre. In the centre, the four phases of the Salvation story are visible (incarnation & birth, crucifixion & resurrection, Christ's triumph over his enemies, and the ruling Christ at his throne), which forms a unit in Christ. Thus, again we get to the 68+1+68=137 structure. (see Fig. 12)

Figure 12

(a) The illustration of the 137 structure of the Shapes on the reconstructed Casula of St. Stephen of Hungary (b) The four mandorlas with 137 letters (c) The place of the round inscription with the 137 characters on the reconstructed Casula

In the given mystical tradition, discussed above, these 137 structures can ensure the Incarnation of Christ in the Crown Prince by the theurgical procedure using the Holy Crown, the (Holy) Casula and the Royal Sceptre.[20]

[20] St. Stephen of Hungary, as can be traced in his reliquies (and according to his "Vita" in Latin) was a master of the languages, symmetries and symbolical mathematics similarly to Pauli and Jung.

A2.2 The Royal Sceptre

According to § 4 of Appendix 1.1 in the **Royal Sceptre** (see Fig. 13 [28]) we can identify the same pattern of YVY=YHVH=Alef=Kether-Crown (**יוי** = **יהוה**=**אלף** = **כתר**) which as we have seen above contains the number archetype **137** as well. Namely we can see two denary (10) structures which are connected a structure consisting of six (6) decorative entities (see Fig. 10). This structure as a unit embraces three Lion pictures which can represent the three attributes of Christ and David as the Lion of Judah.

Figure 13
The Royal Sceptre [28] with the structure of 10+6+10 (137= **יוי אלף**)

The whole symbolism through the "Alef YHVH" as Crown of the Lord represents the three crowns of the Messiah (royal, priestly and prophetic or Torah) according to the § 152 of the Book Bahir (**ג כתרים ישראל**). In this case, the meaning of 137 is the incarnation of Christ and David in the Crown Prince through his three attributes or crowns reinforcing this act by the power of the (Holy) Royal Sceptre too.

References

[1] The Book Bahir (ed. A. Kaplan) Samuel Weiser, INC., York Beach, Maine, 1989

[2] Dan, J.: Three Types of Ancient Jewish Mysticism. In: 7th Rabbi Feinberg Memorial Lec. in Judaic Studies. University of Cincinnati (1984)

[3] Dan, J.: 'The Emergence of Mystical Prayer', Studies in Jewish Mysticism, Proceedings of Regional Conferences held at the University of California, Los Angeles, and McGill University, eds. Joseph Dan and Frank Talmage, Cambridge Mass: Association for Jewish Studies 1982, pp. 85-120

[4] Dan, J.: **תורת הסוד של חסידי אשכנז ירושלים**. 1968, pp. 119-122

[5] Gerics J., Ladányi E.: The Idea of King – St. Stephan of Hungary – Europe, Levéltári Szemle, 54. évf. 2. sz., 2004, 3-14. o. (in Hungrian)

[6] Heisenberg, W. „Wolfgang Paulis Philosophische Auffassungen", in Ztschr. für Parapsychologie und Grenzgebiete der Psychologie. III. Nr. 2/3, 1960, p. 127 [69]

[7] Idel, M.: Kabbala New Perspectives. Yale Univ. Press, 1988

[8] Jung, C. G.: Psychologie und Alchemy. Walter Verlag, Olten (1972)

[9] Jung, C. G: Mysterium Conjunctionis. Princeton Univ. Press., 1977

[10] Jung, C. G., Pauli, W.: Naturerklärung und Psyche. Rascher Verlag, Zürich (1952). In English The Interpretation of Nature and the Psyche, New York, 1955

[11] Jung, C. G.: Collected Works, Bolingen, Vols. 8, 9, 11, 12, 14, Princeton University Press (1973)

[12] Jung, C. G. (ed. G. Adler, A Jaffe) Letters I-II. Princeton Univ. P (1973)

[13] Kalman, R., et al.: Mathematical System Theory, chap. On Invariants, Canonical Forms, Moduli for Linear Constant Final Dimensional Dynamical Systems. Springer Verlag (1971)

[14] Lindorff, D.: Pauli and Jung: The Meeting of Two Great Minds. Quest Books (2004)

[15] MacGregor, Malcolm H., (2007) The Power of Alpha, World Scientific. Singapore

[16] Margalioth.:(The Book Bahir), **ירושלים, ספר הבהיר מרגליות מוסד הרב קוק**

[17] Meier, C. (ed.): Atom and Archetype: The Pauli/Jung Letters, 1932-1958, Routledge, London (2002)

[18] Miller, I. A. Deciphering the Cosmic Number the Strange Friendship of W. Pauli and C. G. Jung, W. W. Norton, New York London, 2009

[19] Pauli, W.: Phenomenon and Physical Reality. Dialectica 11, 35-48 (1957). Introduction to a Symposium on the occasion of the International Congress of Philosophers in Zürich, 1954

[20] Pauli, W. (eds. Enz, C., Meyenn, K. V.): Writings on Physics and Philosophy. Springer (1994)

[21] Popper, K.: Conjectures and Refutations: The Growth of Scientific Knowledge. Routledge and Kegan Paul, London, New York (1992)

[22] Ricoeur, P.: Structure et herméneutique. In: Le Conflit des Interprétations, pp. 31-63, Seuil (1969)

[23] Scholem, G.: Origins of the Kabbalah. Princeton Univ. Press, 1990

[24] Scholem, G.: On Mystical Shape of Godhead. Schocken, New York, 1996

[25] Scholem, G.: Major Trends in Jewish Mysticism, Schoken, New York, 1983

[26] Straub, A., Keller, G.: Hortus Deliciarum, Strassbourg. 1879-1900 (see still Herrad of Hohenbourg Hortus Deliciarum, (Commentary, Reconstruction) (ed. R Green), London, 1979

[27] Thorndike, L.: A History of Magic and Experimental Sciences. Columbia Univ. Press., 1952

[28] Tóth E., Szelényi K.: The Holy Crown of Hungary, Budapest, 2002 (in Hungarian)

[29] Kóczy, L., Várlaki, P.: A Comparative Study of Pictures from Pala d'oro in St. Mark Catherdral of Venice and from the Holy Crown of Hungary. In: Proc. of Intl. Conference on Genealogy and Heraldry. Hungary (2006) pp. 131-171 (in Hungarian)

[30] Várlaki, P., Nádai, L., Bokor, J.: Number Archetypes and "Background" Control Theory Concerning the Fine Structure Constant. Acta Polytechnica Hungarica, Vol. 5, No. 2, pp. 71-104 (2008)

[31] Várlaki, P., Nádai, L., Background Control and Number Archetype in Perspective of the Pauli–Jung Correspondence, In Proc. of Workshop On System and Control Theory, In Honor of J. Bokor on 60th birthday, Budapest, pp. 195-229, 2009

[32] Várlaki P., Kóczy L. T.: Genealogical Myth and Allegory in the Pala d'Oro of Venice and Saint Stephen's Royal Mirror. Internat. Conf. Of the Hungarian Heraldic and Genealogical Society, pp. 89-124 (in Hungarian)

[33] Várlaki, P., Kóczy, L., Kiss G.: Analysis of the Enamel Pictures on the Holy Crown of Hungary, Internat. Conf. of the Hungarian Heraldic and Genealogical Society, pp. 101-130 (in Hungarian)

[34] Várlaki, P., Rudas, I.: Twin Concept of Fine Structure Constantas the 'Self Number-Archetype'in Perspective of the Pauli-Jung Correspondence. Part I. Observation, Identification and Interpretation, and Part II. Cognition, Imagination and Background Process, Acta Polytechnica Hungarica, 2009, Vol. 6, No. 2, pp. 77-108 and pp. 109-137

[35] Várlaki P., Bokor J.: Number Archetypes, Symbolic Coding Letters and "Background Communication Theory" in Saint Stephen's Royal Mirror. In:

IEEE 7th International Conference on Computational Cybernetics (ICCC 2009) Palma de Mallorca, Spain, November 26-29, 2009, pp. 201-211

[36] Várlaki, P., Bokor J., Rudas I. J.: "System Identification" and Hermeneutics for Long Run Series of Synchronicities in the Pauli-Jung Relationship. In: IEEE 7th International Conference on Computational Cybernetics (ICCC 2009) Palma de Mallorca, Spain, November 26-29, 2009, pp. 129-140

[37] Várlaki, P., Rudas I., Nádai L.*:* System Hermeneutics and Acausal Background Creative Intelligence in "Observations" and Synchronicities of the Pauli–Jung Relationship (under publication) 2010