Problems in solving fractional differential equations in a microcontroller implementation of an FOPID controller

MARIUSZ MATUSIAK, PIOTR OSTALCZYK

Lodz University of Technology
Institute of Applied Computer Science
Faculty of Electrical, Electronic, Computer and Control Engineering
Stefanowskiego 18/22, 90-924 Lodz, Poland
e-mail: mmatusiak@iis.p.lodz.pl, postalcz@p.lodz.pl

(Received: 23.11.2018, revised: 27.03.2019)

Abstract: The article focuses on the fractional-order backward difference, sum, linear time-invariant equation analysis, and difficulties of the fractional calculus microcontroller implementation with regard to designing a fractional-order proportional integral derivative (FOPID) controller. In opposite to the classic proportional integral derivative (PID), the FOPID controller is defined by five independent parameters. Hence, it is more customizable and, potentially, more precise on condition that the values of fractional integration and differentiation orders are properly selected. However, a number of operations and the time required to calculate the output signal continuously increase. This can be a significant problem considering the limitations of a microcontroller, including memory size and a constant sampling time of the set-up analog-to-digital (ADC) converters. In the article, three solutions are considered, and results obtained in the experiments are presented.

Key words: fractional calculus, Grünwald-Letnikov fractional-order backward difference/sum, FOPID, hardware implementation

1. Introduction

The fractional calculus [1, 2] has become more and more effective, and hence popular mathematical tool in the synthesis and analysis [3] of dynamic systems in numerous fields of science including signal processing, control, biomedical and electrical engineering [4–8]. In control applications, achieving better models of real objects and developing more sophisticated and precise
control procedures requires advanced computational tools. This statement also applies to the analysis of fractional-order system dynamics. The main concern brought by the implementation of these strategies is associated with a problem of the constantly, linearly growing over time number of multiplications and additions, known as the growing calculation tail [8]. This may be a source of considerable errors in real-time microcontroller-based applications that perform critical operations during a constant sampling time and save results in a limited memory [6]. Several approaches to implementing fractional-order operators have been proposed. Two popular are the Oustaloup approximation [9] and the Laguerre impulse response approximation (LIRA) [10]. Applications of the former were discussed among others in the works [5, 11–13] regarding non-integer order digital filters. The method allows one to approximate non-integer filters with a wider spectrum than LIRA; however, it is considered sensitive to the process of discretization at high sampling frequencies and may cause system instability. A step towards implementing Oustaloup approximation algorithms in a fractional-order PID controller [14] was made by Tepljakov in his praised work on a fractional-order modeling toolbox - FOMCON [15]. In the project, a performance comparison between an 8-bit AVR microcontroller and 32-bit ARM Cortex-M4 microcontroller, similar to the one used in the authors’ research, was also conducted.

In the following paper, an approach based on the Grünwald-Letnikov definition of the differ-integral and the Short Memory Principle optimization technique introduced by Podlubny [2] is considered. Usefulness of this method has been verified, among others, in air and plant heating control and robotic arm movement optimization in [4, 6, 8].

The structure of the paper is organized as follows: in Section 2, mathematical preliminaries including the Grünwald-Letnikov definition of the discrete fractional-order backward difference and a sum based on the oblivion function [2, 16] are introduced. Two solutions for the previously described problem are highlighted. In Section 3, the discrete fractional-order proportional-integral-derivative (FOPID) controller is considered [14]. An analysis of the implementation difficulties was conducted using the chosen Cortex-M7-based hardware testing platform. The methodology and the obtained results are described in Section 4. In Section 5, accuracies of the classic and fractional-order PID controllers [17] are compared. The impact of the difficulties on a control process is also discussed.

2. Grünwald-Letnikov definition of fractional-order backward difference/sum

Let us use the following notation $N_c = \{c, c+1, c+2, \ldots\}$, then $N := N_0 = \{0, 1, 2, 3, \ldots\}$. The discrete function of a variable $k \in N$ and a given fractional-order $\nu$, matching the inequality $0 < \nu \leq 1$, is defined by its values $a^{(\nu)}(k)$ (1):

$$a^{(\nu)}(k) = \begin{cases} 
1 & \text{for } k = 0 \\
(-1)^k \frac{\nu(\nu-1)\ldots(\nu-k+1)}{k!} & \text{for } k = N_1 
\end{cases}$$

The above formula is called the oblivion function and is a basis for evaluating the fractional-order backward difference and fractional-order backward sum. A recursive equivalent of Equa-
tion (1) is defined by:

\[
a^{(\nu)}(k) = \begin{cases} 
1 & \text{for } k = 0 \\
a^{(\nu)}(k-1) \left(1 - \frac{\nu+1}{k}\right) & \text{for } k = N_1 \end{cases}
\]  

(2)

The coefficients calculated for various positive \(\nu\) orders are presented in Fig. 1(a).

Using the oblivion function one can calculate fractional-order backward difference of an order \(\nu \in \mathbb{R}_+\) (3):

\[
GL\Delta_h^{(\nu)} f(t) = \sum_{k=0}^{\infty} a^{(\nu)}(k) f(t - kh) \quad \text{for } t, h \in \mathbb{R},
\]

where \(f(t)\) is the discrete-time real-valued function defined by its consecutive samples \(f(kh)\) and \(h\) denotes the sampling time. The fractional-order backward sum is described by a similar equation, with the difference that the fractional-order takes a negative value, e.g. \(\mu = -\nu\). A plot of \(\mu\)-order coefficients is shown in Fig. 1(b). One may also generalize Equations (1), (2) by evaluating the \(\nu\) order for each time instance \(kh\) as a result of a uniquely defined discrete function \(\nu(k)\). This generalization is referenced as the variable-, fractional-order backward difference/sum [8, 18].

The one-sided Z transform of the mentioned difference takes the following form (4):

\[
z \left\{GL\Delta_h^{(\nu)} f(t) \right\} = \left(1 - z^{-1}\right)^{\nu}.
\]

(4)

The left-sided Grünwald-Letnikov derivative is then defined by Formula (5) [19]:

\[
GLD_h^{(\nu)} f(t) = \lim_{h \to 0^+} \frac{GL\Delta_h^{(\nu)} f(t)}{h^{\nu}} \quad \text{for } t, h \in \mathbb{R}_+.
\]

(5)
Assuming that the sampling time $h$ is constant and as small as possible, (5) can be simplified to:

$$GLD_h^{(v)} f(t) = \frac{GLA_h^{(v)} f(t)}{h^v} \quad \text{for } t, h \in R_+.$$ (6)

If a positive value of the order $v$ is replaced in Equations (3), (5), (6) by a negative value of $\mu$ then the definitions of the Grünwald-Letnikov fractional-order sum and integral are considered.

The definitions are characterized by the growing number of calculations at every step. If $M$ represents the number of operations performed by a microcontroller in the first step then in the $k$-th step $kM$ operations should be executed. The issue arises when the time required to perform these operations is longer than the desired sampling period $h$ ($t_{kM} > h$). One of the possible solutions is to assume that for a specific value of $L$ for which $t_{LM} \geq h$ and $k > L$, the number of operations is limited to $LM$. Formula (1) then takes the form of:

$$a^{(v)}(k) = \begin{cases} 
0 & \text{for } k < 0 \\
1 & \text{for } k = 0 \\
(-1)^{k} \frac{v(v-1)\cdots(v-k+1)}{k!} & \text{for } 0 < k \leq L \\
(-1)^{L} \frac{v(v-1)\cdots(v-L+1)}{L!} & \text{for } k > L
\end{cases}$$ (7)

The second solution is based on the assumption that for small order values the coefficients drop rapidly to zero, as shown in Fig. 1, and further operations can be skipped on condition that a required precision of results has been reached:

$$a^{(v)}(k) = \begin{cases} 
0 & \text{for } k < 0 \\
1 & \text{for } k = 0 \\
(-1)^{k} \frac{v(v-1)\cdots(v-k+1)}{k!} & \text{for } 0 < k \leq L \\
0 & \text{for } k > L
\end{cases}$$ (8)

3. Fractional-order PID controller

The fractional-order PID controller for a closed-loop system (CLS) [14] is described by Equation (9):

$$y(kh) = K_P u(kh) + K_I \left( GLA_h^{(\mu)} \right) u(kh) + K_D \left( GLA_h^{(v)} \right) u(kh),$$ (9)

where $K_P$, $K_I$, and $K_D$ are the proportional, integral and derivative gains, $\mu < 0$, $v > 0$ are the fractional orders, and $u(kh)$ and $y(kh)$ are the input (error) and output signals, respectively.
Applying Formula (3) results in the form:

\[
y(kh) = K_P \left[ \begin{array}{c} u(kh) \\ u(kh - h) \\ \vdots \\ u(0h) \end{array} \right] + K_I \left[ \begin{array}{c} 1 \\ a(\omega)(1) \\ \vdots \\ a(\omega)(k) \end{array} \right] \left[ \begin{array}{c} u(kh) \\ u(kh - h) \\ \vdots \\ u(0h) \end{array} \right] + \\
+ K_D \left[ \begin{array}{c} 1 \\ a(\nu)(1) \\ \vdots \\ a(\nu)(k) \end{array} \right] \left[ \begin{array}{c} u(kh) \\ u(kh - h) \\ \vdots \\ u(0h) \end{array} \right].
\] (10)

The discrete transfer function of (10) is obtained by the one-sided Z-transform (assuming zero initial conditions):

\[
\frac{Y(z)}{U(z)} = \frac{K_P (1 - z^{-1})^{-\mu} + K_I + K_D (1 - z^{-1})^{-\nu}}{(1 - z^{-1})^{-\mu}}.
\] (11)

In an ideal controller, the degree of the numerator is higher than the degree of the denominator. In real applications, the differentiator strongly amplifies noise that is always present in a signal, so usually, the differentiator with inertia, acting as a discrete filter, is used. The transfer function then takes the form:

\[
\frac{Y(z)}{U(z)} = K_P + \frac{K_I}{(1 - z^{-1})^{-\mu}} + \frac{K_D (1 - z^{-1})^{\nu}}{(1 - z^{-1})^{\nu} + a_0}.
\] (12)

### 4. Hardware testing platform

For the analysis of fractional-order operator implementation, a hardware testing platform based on the ARM Cortex-M7 STM32F746ZG microcontroller [20] was prepared. Key features and peripherals of the device are listed in Table 1.

<table>
<thead>
<tr>
<th>Parameter name</th>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU main clock frequency</td>
<td>(F_{CPU})</td>
<td>up to 216 MHz</td>
</tr>
<tr>
<td>Memory</td>
<td></td>
<td>Flash, SRAM 1 024 MB (Flash) + 320 KB (SRAM)</td>
</tr>
<tr>
<td>Converters</td>
<td></td>
<td>ADC, DAC 3 × 12-bit 2.4 MSPS ADC, 2 × 12-bit DAC</td>
</tr>
<tr>
<td>Power supply</td>
<td>(V_{DD})</td>
<td>1.8–3.6 V (3.3 V set)</td>
</tr>
<tr>
<td>Other features</td>
<td></td>
<td>floating point unit real-time accelerator, DSP instructions</td>
</tr>
</tbody>
</table>

Several limitations must be taken into consideration when implementing the mentioned algorithms. First of all, the desired precision of numbers must be determined. For the fractional order
\( v \in (0, 1] \) and \( k \to \infty \), \( a^{(v)}(k) \) coefficients get closer to zero and for large values of \( k \), the coefficients might not be represented properly in a microcontroller memory. The GCC compiler for STM32F746ZG supports two ANSI C floating-point types: a single-precision 4-bytes \texttt{float} type and a double-precision 8-bytes \texttt{double} type [21]. Ranges are [1.17549435e-38, 3.40282347e+38] with 8 decimal points and [2.22507385850720138e-308, 1.79769313486231571e+308] with 17 decimal points respectively. The maximum positive value of the type is important for implementation of Formula (1) due to the growing factorial value and a possible risk of overflow. Otherwise, recursive Formula (2) needs to be used.

Choosing the more precise floating-point type implies consuming more memory which is the second important limitation to consider. If \( a^{(v)}(k) \) coefficients are not defined directly in a program source code, and thus not saved in a Flash section of memory, they need to be calculated at runtime and saved in RAM. In 320KB of the STM32F746ZG SRAM, up to 40 960 \texttt{double} type coefficients could be stored; however, the section is always shared by other parts of the program including the stack, heap, initialization routines, and other data. An exemplary C-language program for evaluating \texttt{double} type \( a^{(v)}(k) \) coefficients using recursive Formula (2) on the STM32F746ZG revealed that up to 8 084 coefficients could be calculated correctly, consuming 64 672 bytes of RAM.

The second hardware implementation problem, mentioned in Section 2, is the growing time of CPU calculations which in some circumstances can exceed the sampling time of the ADC converter. The maximum frequency of the ADC domain clock in the STM32F746ZG is \( f_{\text{ADC}} = 36 \) MHz [20], and the single 12-bit conversion takes \( c_{\text{ADC}} = 15 \) cycles to complete. Hence, the shortest possible sampling time can be estimated using Formula (13):

\[
T_{s_{\text{min}}} = \frac{1}{f_{\text{ADC}}} c_{\text{ADC}} = \frac{1}{36 \text{ MHz}} 15 = 0.417 \mu s. \tag{13}
\]

Usually, one of the available general-purpose timers is used to extend the time and to synchronize the conversions. To configure a timer one must know the exact time of critical operations performed during an ADC interrupt service routine (ISR). The ARM Cortex peripheral called Data Watchpoint and Trace (DWT) [20] allows one to measure the number of CPU cycles between two lines of the program. Knowing the CPU frequency, one can determine the time of specifics sets of operations (14):

\[
t_{\text{diff}} = \frac{c_{\text{after}} - c_{\text{before}}}{f_{\text{CPU}}}. \tag{14}
\]

This was tested with a C program running the algorithm of a discrete fractional-order differentiator (6) for 11 different coefficients lengths between 1 and 5 000. The results are presented in Fig. 2.

For \( k = 3 \ 000 \) coefficients already stored in the memory, the number of cycles required to calculate the response was equal to \( c_{\text{diff}} = 189 \ 345 \), and lasted for (15):

\[
t_{\text{diff}} = \frac{c_{\text{diff}}}{f_{\text{CPU}}} = \frac{189 \ 345}{216 \text{ MHz}} = 0.87 \text{ ms.} \tag{15}
\]

The time between subsequent ADC conversions has to be longer than the sum of \( t_{\text{diff}} \) and \( h_{\text{min}} \), so in this example, the timer configured to generate interrupt signals at 1 kHz fulfilled the condition.
To increase the overall microcontroller performance, one may consider rewriting critical sections of a program using inline assembly code. If a chosen device supports single instruction, multiple data (SIMD) extensions, then basic operations including multiply and accumulate (MAC) are executed in one clock cycle, significantly reducing time of complex algorithms. The performance of STM32 devices, using the optimized DSP instructions, was compared in the sheet [22].

5. Fractional-order PI\textsuperscript{\mu}D\textsuperscript{\nu} implementation

In order to implement discrete PID and PI\textsuperscript{\mu}D\textsuperscript{\nu} controllers [14], a plant had to be defined first. To indicate limitations of the microcontroller, unit step response of the controlled system had to have relatively wide transient characteristics. A DC motor, in general, is an example of a device that can meet that criterion; thus, the 6 V-powered DC motor with an encoder was used for the reference [23]. The unit step was simulated by turning the motor power supply on and measuring the number of encoder impulses every ten milliseconds. A graph of the normalized rotations per second is presented in Fig. 3 (thick red line).
Based on the diagram, an exemplary 1st-order dynamic plant with inertia was defined, described by transfer function (16):

\[
G(s) = \frac{k_M}{T_M s + 1} = \frac{0.98}{0.075s + 1},
\]

where \(k_M\) is the gain of the plant and \(T_M\) is the time constant. Using Tustin’s method [24] for the sampling time \(h = 0.01\) s, the above can be transformed into:

\[
G(z) = \frac{0.1223z^{-1}}{1 - 0.8752z^{-1}},
\]

which is the equivalent of the difference equation:

\[
y[n] = 0.1223x[n - 1] + 0.8752y[n - 1].
\]

The simulated 1st-order dynamic plant was implemented in a C program for the ARM Cortex-M3 core STM32L152RCT6 [25] microcontroller with ADC and DAC converters configured. The obtained unit step response is presented in Fig. 3 (thick purple line).

A general discrete PID controller is described by the following Formula (19):

\[
u[kh] = K_P(e[kh] + \frac{h}{T_I} \sum_{i=0}^{k} e[ih] + \frac{e[kh] - e[kh-h]}{h} T_D),
\]

where \(h\) is the sampling time, \(K_P\) is the proportional gain, and \(T_I\) and \(T_D\) are the integral and derivative times respectively. The implemented controller was tuned using the Ziegler-Nichols method [26]. Parameters are presented in Table 2.

<table>
<thead>
<tr>
<th>Parameter name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>(K_P)</td>
<td>2</td>
</tr>
<tr>
<td>(T_I)</td>
<td>0.028</td>
</tr>
<tr>
<td>(T_D)</td>
<td>0.007</td>
</tr>
</tbody>
</table>

Both the Matlab simulation and C program implementation of the controller were carried out. The target setpoint value of the regulator was set programmatically to 1 (corresponding to 1 V on the DAC output). The output of the STM32F746ZG DAC converter (PID) was connected directly to the input of the STM32L152RCT6 ADC converter (1st-order plant), and the STM32L152RCT6 DAC output as the negative feedback to the STM32F746ZG ADC (see Fig. 4). The converters were in sync with 100 Hz peripheral timers. Maximum (3.3 V) and minimum (0 V) control signal limits were also applied. Characteristics of the simulation and measured device output are presented in Fig. 5(a).

In a typical PID implementation, a total number of calculations performed at each step is constant. This is the outcome of Algorithm 1 where only two additional variables are saved in
memory. The first variable `errors_sum` represents the sum of all previous errors multiplied by the sampling time (the integral term), while the second `previous_error` keeps the value of the error from the last step (the derivative term).

![Connected STM32F746ZG (PID) and STM32L152RCT6 (plant) microcontrollers](image)

**Fig. 4.** Connected STM32F746ZG (PID) and STM32L152RCT6 (plant) microcontrollers

```plaintext
Algorithm 1. Classic discrete PID controller

Error_sum <- 0
previous_error <- 0
setpoint <- setpoint value
loop:
    error <- setpoint - measured_error
    errors_sum <- errors_sum + error * h
    derivative <- (error - previous_error) / h
    output <- Kp * error + Ki * errors_sum + Ku * derivative
    previous_error <- error
    sleep(n)
goto loop
```

In the fractional-order PID controller, according to Equation (10), both derivative and integral terms are sums of products of errors $e(k)$ and the corresponding backward difference coefficients, divided by the $v$-power of the sampling time $h$. As mentioned in Section 4, this may cause an undefined behavior of a microcontroller program if calculations are not properly limited.

To analyze the problem thoroughly, the fractional-order PID (FOPID) algorithm was implemented on the STM32F746ZG and several iterations of memory consumption and performance tests were executed. Fractional order values $v = 0.94$ and $u = -1.0$ were selected to minimize integral of the squared error (ISE criterion). The sampling time was set to $h = 0.01$ s (100 Hz). The $a^{(v)}(k)$ and $a^{(u)}(k)$ coefficients of type `float` were calculated during the initialization routine and stored in SRAM in variable-length arrays. The maximum possible length of the arrays equaled 10 200.

The computation time of the FOPID controller output signal as a function of the array length is shown in Fig. 6.

If a chosen sampling frequency $f_s$ is too high for the desired buffer length, e.g. $f_s = 1$ kHz for 4 000 samples, the controller program will skip upcoming ADC conversions, and, therefore, generate a wrong control signal. Calculation time can increase even further when the varying
fractional order $\nu(\cdot)$ in a variable-, fractional-order PID (VFOPID) controller is considered [8]. The oblivion coefficients are then recalculated in each ISR.

Depending on the selected fractional-order values, the application of solution (7) may cause a major problem related to shifting errors stored in memory. When the error buffer limit is reached, the first value, being also the largest at the start of the control, is overridden resulting in a temporary rapid drop of the signal, caused mainly by the integral term of the FOPID. An example of this issue registered for a buffer length of 30 is presented in Fig. 7.

At the expense of more complex implementation, the fractional-order PID control provides more tuning parameters; hence, much more flexibility in designing a controller best suitable for a particular system. Measured responses for three different control signals: unit step, classic PID and fractional-order PID are presented in Fig. 8. For the mentioned fractional-order values $\nu = 0.94$
and $\mu = -1.0$, integral of the squared error (ISE) for the simulation time of 1 s was slightly reduced compared to the classic PID defined by the same set of $K_p$, $T_I$, $T_D$ and $h$ parameters. Moreover, for $-0.9 > \mu > -1$ rise time of the plant response was even faster; however, a non-zero steady state error was then produced.

This problem may be eliminated with a VFOPID controller in which the value of the $\mu$-order gradually approaches $-1$. The stability criterion for the controller, tuned using the same parameters as the classic PID, is defined similarly [8]:

$$v(k) = \begin{cases} f(k) & \text{for } 0 \leq k < L < \infty, \\ 1 & \text{for } k > L \end{cases}. \quad (20)$$

Choosing the valid $\mu$ and $v$ values remains an open problem and strictly depends on properties of the considered closed-loop system.
6. Conclusions

Applying fractional calculus to the field of control engineering introduces more degrees of freedom [4] which, for accurately selected and adaptive values of fractional orders, provides the ability to design more precise and robust control algorithms. However, in order to eliminate the steady-state error present in a FOPID, an integer integration order value is eventually required. One must also be aware of facing additional difficulties when it comes to the microcontroller implementation, including memory size limits and the growing number of calculations. Some optimized algorithms affecting the controller accuracy might be necessary. The first possible approach would be to limit the number of operations or skip further calculations if the obtained results reach the desired precision. Depending on the selected fractional orders, the controller may then react sensitively on every rapid setpoint change. The second approach would be to increase the performance of a microcontroller by implementing a program using an inline assembly code and optimized DSP instructions. The question is particularly important when a variable-, fractional-order PID is considered.

Acknowledgements

This work was supported by Polish funds of the National Science Center under grant 2016/23/B/ST7/03686.

References

Problems in solving fractional differential equations


