AN0002

Application note

PWM-based hardware DAC for mainstream STM32 microcontrollers

Author: Geraldo Lucas Pugialli de Paiva

Introduction

Most of the high performance STM32 microcontrollers embed 12-bit DAC (Digital to Analog Converter) modules, able to generate an analog output limited to the device's supply voltage capable to achieve up to 1 million samples per second. However, when using mainstream STM32 microcontrollers (cheaper and most common in online stores as prototyping boards) the DAC peripheral may not be available, limiting the capabilities of the embedded device designer.

Considering that several applications benefit from the microcontroller being able to control external devices or generate analog outputs, such as active industrial sensors with current or voltage analog output, low frequency signal generators and a complete range of applications, it is necessary to provide a mechanism to use this capability even in simpler and cheaper microcontrollers. This note demonstrates the theoretical background behind a Hardware DAC implementation using external operational amplifier circuits and suggests a prototyping circuit containing expressions to calculate the parameters according to the user application.

This document is applicable to every STM32 microcontroller containing a PWM output, including the ones with internal DAC peripherals, as a possible add-on to specific applications and a useful resource for prototyping.

Contents

List of tables

List of figures

1 Microcontroller DAC types

Signals are mainly classified in two types according to its variation in time: Analog and Digital. The information that we perceive in real world exists mostly in analog form while the digital devices such as computers, smartphones and others use discrete data signal in digital domain.

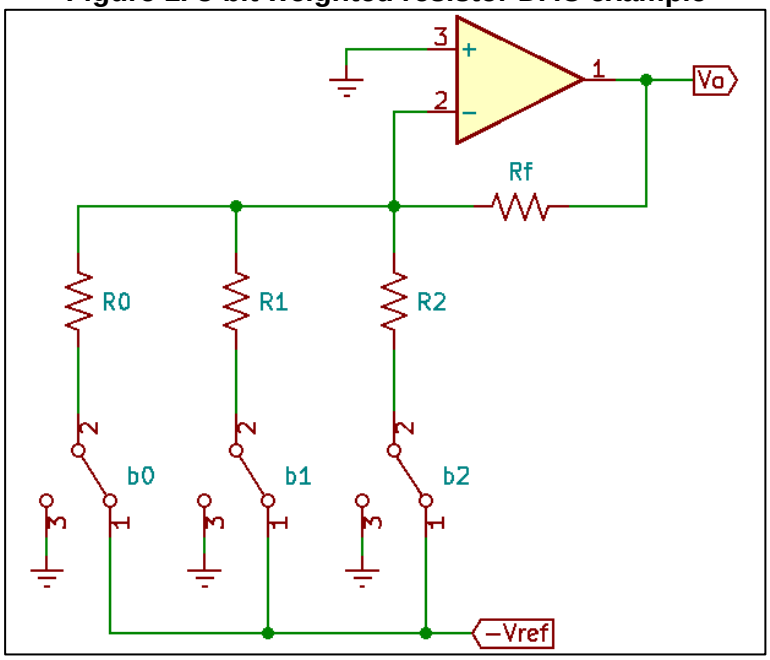
A Digital to Analog Converter (DAC) converts a multi bit (*n*) digital input signal into a single analog output signal. The digital signal is represented with a binary code, which is a combination of bits 0 and 1 and the output value is a variable voltage limited from 0 to the power supply voltage and with variations corresponding to the possible variations in the input (2^n).

In addition, there are several ways to build a DAC, depending on the technology available to design or application requirements, such as: using a summing amplifier (known as Weighted Resistor Method), using a combination of resistors (known as R-2R Ladder) or using a PWM base signal with appropriate conditioning circuits.

1.1 Weighted resistor

A weighted resistor DAC produces an analog output, which is approximate to the binary input by using binary weighted resistors in an inverting adder circuit. The circuit diagram of such device is shown in Figure 1, corresponding to a 3-bit topology.

Figure 1. 3-bit weighted resistor DAC example



Considering that the value present in the input is correspondent to $B_2B_1B_0$ (where B_x indicates the bit order) and each bit must assume a 0 or 1 value, the digital switches shown in Figure 1 will be connected to ground when the corresponding input bits are equal to 0 and will be connected to the negative reference voltage, $-V_{ref}$ when the corresponding input bits are equal to 1.To reproduce the bitwise behavior, the resistors associated to each bit input is a power of two multiple of a base resistor, located in the most significant bit (MSB). In the example above, $R_2 = 2^0R$, meaning that $R_1 = 2^1R$ and $R_0 = 2^2R$.

In the circuit, the non-inverting input terminal of an op-amp is connected to ground, which implies that zero volts is applied at the non-inverting input terminal of op-amp. Since the operational amplifier is displayed in a negative-feedback configuration, the virtual short concept applies, meaning that the voltage at both input terminals is the same. Then, it means the voltage at the inverting input terminal will also be equal to zero volts. The nodal equation at the inverting input terminal's node is expressed as:

$$\frac{0 + V_{ref}b_2}{2^0R} + \frac{0 + V_{ref}b_1}{2^1R} + \frac{0 + V_{ref}b_0}{2^2R} + \frac{0 + V_o}{R_f} = 0$$

$$V_o = \frac{V_{ref}R_f}{R}\left(\frac{b_2}{2^0} + \frac{b_1}{2^1} + \frac{b_0}{2^2}\right)$$

The base resistor is often selected as $R = 2R_f$, which makes the expression as:

$$V_o = \frac{V_{ref}}{2}\left(\frac{b_2}{2^0} + \frac{b_1}{2^1} + \frac{b_0}{2^2}\right)$$

The equation above represents the output voltage equation of the 3-bit weighted resistor DAC. Since the input has a 3-bit format, it is expected to obtain an output with 7 ($2^3 - 1$) possibilities, varying the binary input from 000 (equivalent to 0V) to 111 (equivalent to the maximum voltage $\frac{2^N-1}{N}V_{ref}$).

Based on this example, the generalized output voltage equation of an N-bit binary weighted resistor DAC as shown below:

$$V_o = \frac{V_{ref}}{2}\left(\frac{b_{N-1}}{2^0} + \frac{b_{N-2}}{2^1} + \cdots + \frac{b_0}{2^{N-1}}\right)$$

1.1.1 Advantages and disadvantages

Based on this type of DAC design, it is important to mention that it contains a reduced number of components required to build the circuit for a reduced input size, which grows linearly with the resolution increase, but with a unitary rate.

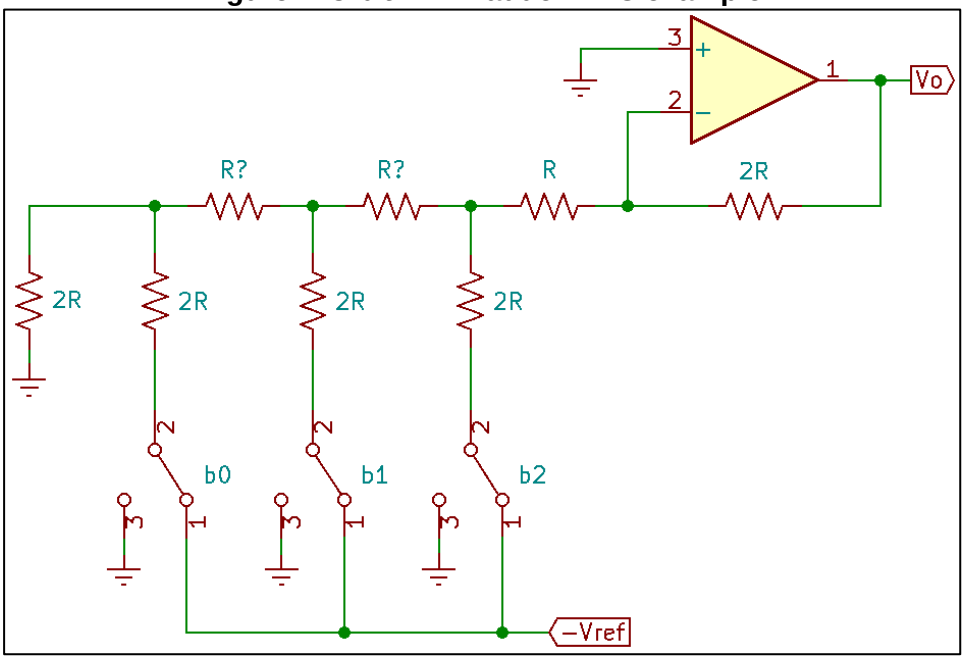
On the other side, the difference between the resistance values corresponding to LSB & MSB will increase as the number of bits present in the digital input increases, which means that a wider range of components is required. Furthermore, it is difficult to design more accurate resistors as the number of bits

present in the digital input increases, since commercial values do not follow the power of two rule.

1.2 R-2R ladder

A R-2R ladder DAC produces an analog output, which is approximate to the binary input by using two resistor values (a base one and a multiple of 2) in an inverting adder circuit. It improves some of the disadvantages of a weighted resistor DAC, such as the increasing difficulty to find a wide range of resistance values. The circuit diagram of such device is shown in Figure 2, corresponding to a 3-bit topology.

Figure 2. 3-bit R-2R ladder DAC example



Considering that the value present in the input is correspondent to $B_2B_1B_0$ (where B_x indicates the bit order) and each bit must assume a 0 or 1 value, the digital switches shown in Figure 1 will be connected to ground when the corresponding input bits are equal to 0 and will be connected to the negative reference voltage, $-V_{ref}$ when the corresponding input bits are equal to 1.To reproduce the bitwise behavior, the resistors associated to each bit input is a power of two multiple of a base resistor, located in the most significant bit (MSB). In the example above, $R_2 = 2^0R$, meaning that $R_1 = 2^1R$ and $R_0 = 2^2R$.

In the circuit, the non-inverting input terminal of an op-amp is connected to ground, which implies that zero volts is applied at the non-inverting input terminal of op-amp. Since the operational amplifier is displayed in a negative-feedback configuration, the virtual short concept applies, meaning that the voltage at both input terminals is the same. Then, it means the voltage at the inverting input

terminal will also be equal to zero volts. The nodal equation at the inverting input terminal's node is not possible to deduce as before. It is necessary to perform a Thevenin equivalent circuit analysis for each bit and superpose each configuration. Then, the final equation for the output voltage is:

$$V_o = V_{ref}\left(\frac{b_2}{2^1} + \frac{b_1}{2^2} + \frac{b_0}{2^3}\right)$$

The equation above represents the output voltage equation of the 3-bit R-2R ladder DAC. Since the input has a 3-bit format, it is expected to obtain an output with 7 ($2^3 - 1$) possibilities, varying the binary input from 000 (equivalent to 0V) to 111 (equivalent to the maximum voltage $\frac{2^N-1}{N}V_{ref}$).

Based on this example, the generalized output voltage equation of an N-bit binary weighted resistor DAC as shown below:

$$V_o = V_{ref}\left(\frac{b_{N-1}}{2^1} + \frac{b_{N-2}}{2^2} + \cdots + \frac{b_0}{2^N}\right)$$

1.2.1 Advantages and disadvantages

Based on this type of DAC design, it is important to mention that it contains more components than the weighted binary version to build the circuit for a reduced input size, which grows linearly with the resolution increase, but with a double rate. Also, this topology uses only two values of resistances, making easier to select the components available commercially, it is easily scalable, and its output impedance is always R.

On the other side, this circuit, as the previous one, require considerable input terminals to define de output value, requiring the microcontroller to provide sufficient pins for each output, often unavailable in simpler devices with a reduced pin count.

1.3 PWM-based

A PWM-based DAC produces an analog output, which is approximate to the filtered single input signal, transformed by the appropriate conditioning circuits. Pulse Width Modulation (or PWM) is a method of varying the average power of a signal by varying its duty cycle, also measured in the percentage of turn on time of the signal over the signal total period. For example, a 10% duty cycle corresponds to a pulse which is kept at the on value for 10% of the time and 90% at the off value.

1.3.1 PWM spectrum

A pulse-width modulated signal such as the one output by a microcontroller is essentially a rectangular wave with a fixed period (therefore a fixed frequency). It

oscillates between two levels, high and low (corresponding to on and off). For the scope of this analysis, it will be assumed that the on voltage is called V_+ and the off voltage is zero, as might be the scenario present when using microcontrollers.

Considering any time variable signal, a PWM signal can be considered as the sum of an infinite amount of sine and cosine waves of varying amplitude and frequency. Then, it is possible to obtain expressions for the amplitudes of these individual waves by computing the Fourier series for the signal. In essence, the original signal can be expressed as:

$$f(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} a_n \cos\left(\frac{n\pi t}{L}\right) + \sum_{n=1}^{\infty} a_n \sin\left(\frac{n\pi t}{L}\right)$$

In this case, the signal period is $2L$ and the coefficients are given by:

$$a_0 = \frac{1}{L}\int_{-L}^{L} f(t)\, dt$$

$$a_n = \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{n\pi t}{L}\right)\, dt, \;\; n = 1,2,3,\ldots$$

$$b_n = \frac{1}{L}\int_{-L}^{L} f(t)\sin\left(\frac{n\pi t}{L}\right)\, dt\,, \;\; n = 1,2,3,\ldots$$

The coefficient a_0 corresponds to the DC offset component of the signal. The a_1 and b_1 coefficients describe the amplitudes of the fundamental frequency sine and cosine waves, while coefficients with higher indices relate to harmonics of the fundamental frequency. Considering $L = 1$ and defining the PWM signal as:

$$f(t) = \begin{cases} 0, & -1 < t < -D \\ V_+, & -D < t < D \\ 0, & D < t < 1 \end{cases}$$

With these assumptions, the coefficients are given by:

$$a_0 = \frac{1}{1}\int_{-1}^{1} f(t)\, dt = \int_{-D}^{D} V_+\, dt = 2V_+ \cdot D$$

$$a_n = \int_{-1}^{1} f(t)\cos(n\pi t)\; dt = \int_{-D}^{D} V_+\cos(n\pi t)\; dt = \frac{V_+}{n\pi}[\sin(n\pi D) - \sin(-n\pi D)]$$

$$a_n = 2\frac{V_+}{n\pi}\sin(n\pi D)$$

$$b_n = \int_{-1}^{1} f(t)\sin(n\pi t)\; dt = \int_{-D}^{D} V_+\sin(n\pi t)\; dt = \frac{V_+}{n\pi}[\cos(n\pi D) - \cos(-n\pi D)]$$

$$b_n = 0$$

This way, the PWM waveform's spectrum is composed by the signal mean value and the carrier fundamental frequency and its harmonics (which are integer multiples of the fundamental). In this kind of analysis, the carrier frequency is assumed to be considerably higher than the message maximum frequency. Then, the message is considered as a continuous signal represented only by the mean value constant of the Fourier series.

1.3.2 PWM signal spectrum examples

To observe the spectrum from the output signal generated by the PWM-based DAC, two example waveforms were designed with a 50% duty cycle and with a 30% duty cycle, both based on a 1kHz carrier, to show the differences between them. The magnitude spectrums of both scenarios are shown in the Figures 4 and 5, respectively.

Figure 3. 50% duty cycle 1kHz PWM theoretical spectrum

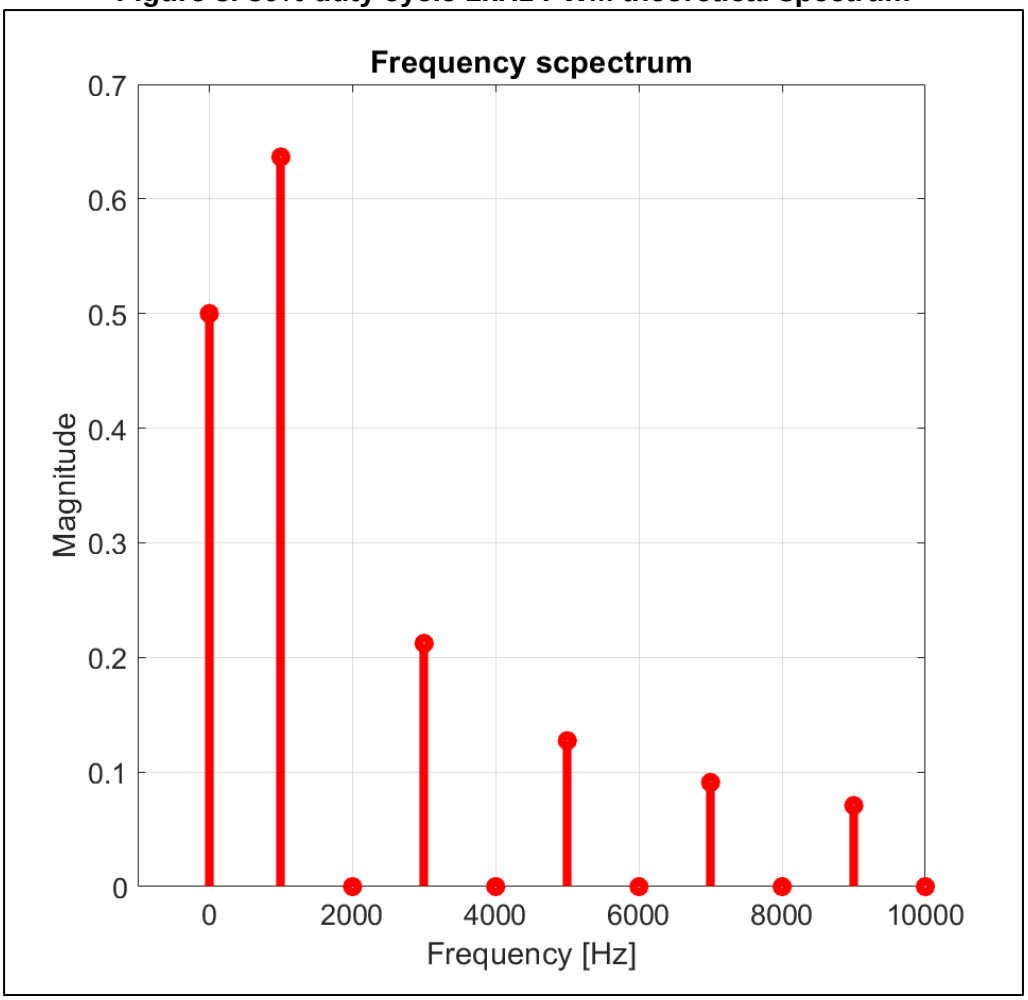
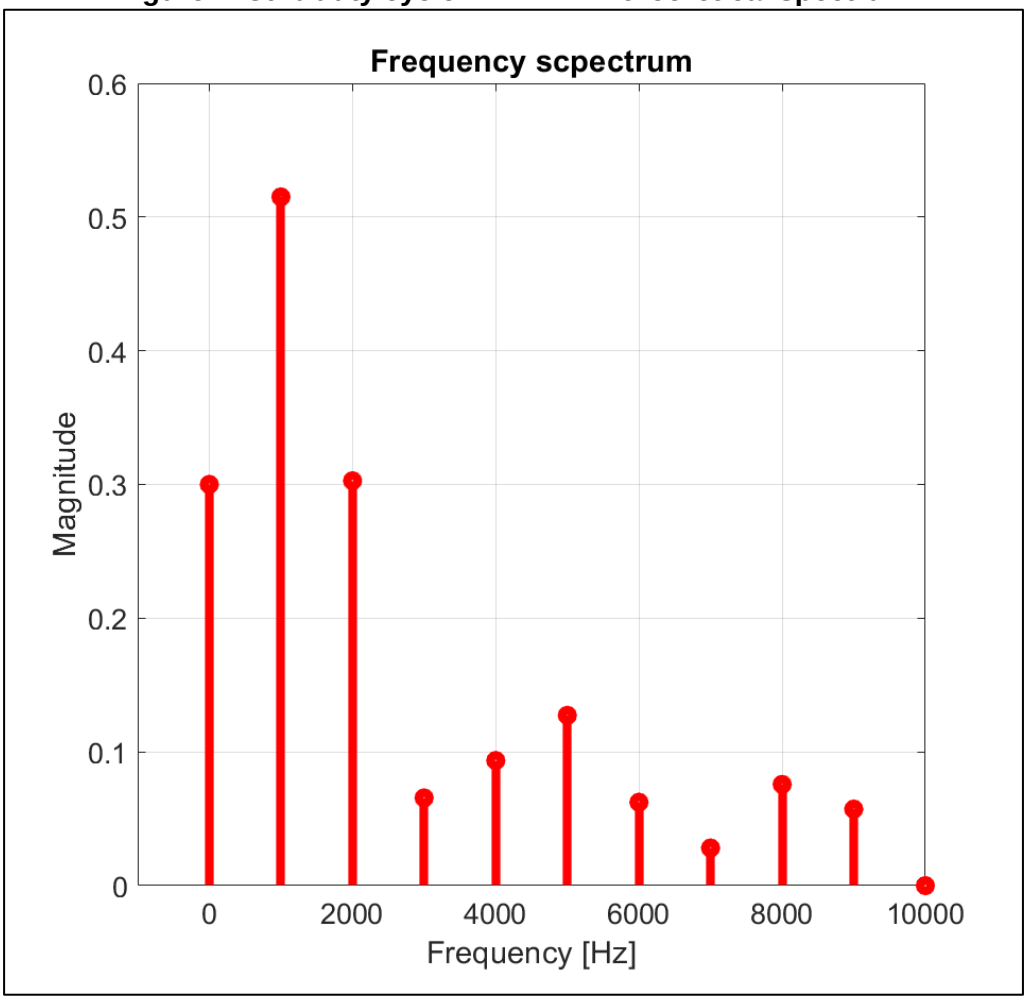


Figure 4. 30% duty cycle 1kHz PWM theoretical spectrum



In the first one, it is observed that at 0Hz (corresponding to the DC mean value) the amplitude has the exact same 50% percentage as defined for the duty cycle. As the time waveform is perfectly symmetrical, the PWM spectrum is composed by only the fundamental frequency and its odd-harmonics with a variable but descending decay. For the second one, it is observed that at 0Hz the amplitude has the exact same 30% percentage as defined for the duty cycle. As the time waveform is not symmetrical, the PWM spectrum has the contribution of the fundamental and all the harmonics.

Based on both examples, theoretically, there is no frequency between the message (0Hz) and the carrier (1kHz, in this case) in the spectrum, meaning that the filter may be tuned remove the fundamental carrier frequency, which will also attenuate the harmonics due to the inherent characteristic of the low pass filter.

1.3.3 Active low-pass filter

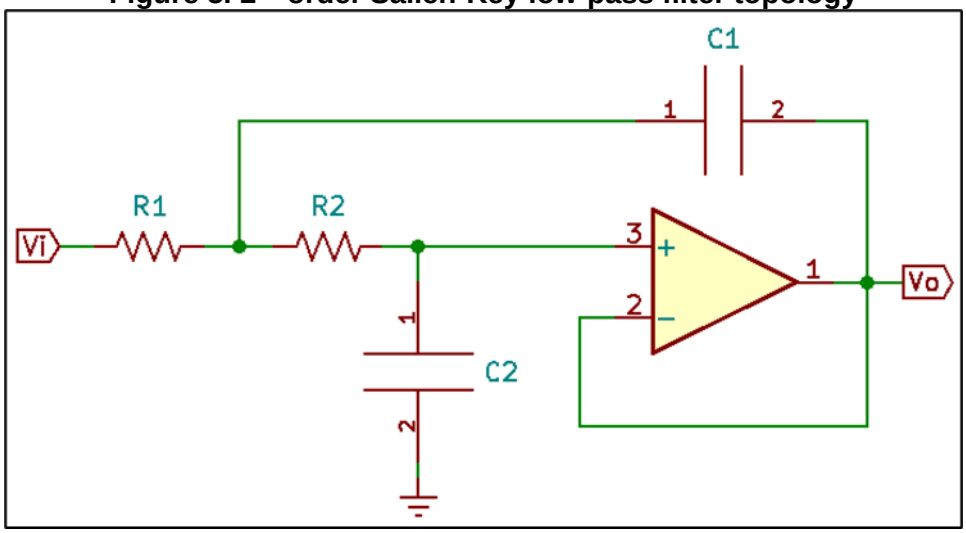
To suppress the carrier frequency and its harmonics and keep only the message signal it is necessary to design a low-pass filter to attenuate the undesired high frequencies. The performance of filtered PWM signal as a DAC depends heavily on the choice of the analog low-pass filter. In general, active filters avoid the impedance loading issues suffered by passive filters, where the upstream or downstream impedances surrounding the filter can change the filter properties. Passive filters can offer lower cost and reduced design complexity.

With active filters, one must also consider the gain bandwidth of the op-amps used. The gain bandwidth represents the upper frequency that the op-amp can effectively handle when used in a closed-loop circuit configuration with small signal input. In terms of active low-pass filters, input signal components with frequency above the gain bandwidth will be attenuated since the op-amp will not have the ability to handle such frequencies. As a practical matter, it is desirable to use an op-amp with gain bandwidth at least 5 to 10 times greater than the highest expected input frequency so that the gain bandwidth of the op-amp does not influence the circuit.

Two important filter properties in the PWM-based DAC application are the bandwidth and the stopband roll-off rate. The filter bandwidth is defined as the frequency at which the filter frequency response magnitude equals 0.707 (-3 dB). The filter bandwidth relates to the maximum signal frequency that the DAC will effectively handle. The stopband roll-off rate is the slope of the frequency response magnitude at high frequency. Combined with the bandwidth, the roll-off rate determines the amount of harmonic ripple that will be seen in the output of the filter. The higher the order of the filter, most attenuation is provided after the bandwidth frequency.

In this specific application, for general purpose analog signal generation with mainstream microcontrollers, a simple 2^{nd} order Sallen-Key low-pass filter tuned at a specific cutoff frequency can isolate the PWM message signal from the high frequency carrier and its harmonic components. This filter topology is displayed in Figure 3.

Figure 5. 2nd order Sallen-Key low-pass filter topology



The expression for the transfer function of the filter present above is:

$$G(s) = \frac{V_o(s)}{V_i(s)} = \frac{\frac{1}{R_1 C_1 R_2 C_2}}{s^2 + \left(\frac{1}{R_2 C_1} + \frac{1}{R_1 C_1}\right)s + \frac{1}{R_1 C_1 R_2 C_2}}$$

Considering that the transfer function of a second-order filter is given by:

$$H(s) = \frac{k_0}{s^2 + k_1 \cdot s + k_0} = \frac{\omega_0^2}{s^2 + 2\zeta\omega_0 \cdot s + \omega_0^2}$$

Then it is possible to determine the cut-off frequency f_0:

$$f_0 = \frac{\omega_0}{2\pi} \rightarrow \boxed{f_0 = \frac{1}{2\pi}\frac{1}{\sqrt{R_1 C_1 R_2 C_2}}}$$

The damping ratio is defined as:

$$\zeta = \frac{k_1}{2\omega_0}$$

And the quality factor is:

$$Q = \frac{1}{2\zeta}$$

With the previous equations it is possible to design the output filter, being suggested to select the capacitors first and then calculate the resistance values that fit your application, given there are variable resistors available such as potentiometers (coarse value setting) and trimmer potentiometers (more precise value setting).

1.3.4 Advantages and disadvantages

Based on this type of DAC design, its greatest advantages are related to the output resolution, which is causally related to the PWM registers resolution and, depending on the microcontroller selection, is higher than any other DAC peripheral format; other key point is that each channel only requires 1 dedicated pin to generate the analog waveform, only being necessary to constantly update the register values associated with each PWM channel. Then, it makes possible for smaller microcontrollers or ones with fewer resources to provide an analog output feature for simpler designs and costing fewer, depending on the application.

On the other side, as each PWM channel require its own filter to generate an analog output, which requires more space to place these circuits around the circuit board. Another problem arises when it is necessary to generate high frequency signals, given it is necessary to design a PWM signal with a higher carrier frequency when compared to the message signal, it may be impossible to generate such configuration due to the microcontroller oscillator limitations.

Besides that, comparing to other DAC designs, the PWM-based DAC is a very flexible and powerful hardware implementation to provide analog output in your design.

2 Example

To demonstrate a simple PWM-based DAC, an example was designed, following the standard procedure for specifying the appropriate components and circuits. For this implementation, it was selected a STM32F103RB microcontroller, available in the NUCLEO-F103RB development board, alongside with an external filter circuit assembled and connected to it.

2.1 Software implementation

In this example, it is intended to create a PWM signal with variable duty cycle and a 1kHz carrier, designed to generate an analog output voltage with a PWM-based DAC. Using the internal HSI RC oscillator with the development board, the maximum clock frequency available using the microcontroller's PLL is 64MHz. The clock configuration based on the STM32F1 registers and bits is present in Table 1, with the respective values (here, only the non-reset values are defined).

Table 1. Register and bit values for RCC configuration

Register	Bit	Value
RCC_CR	PLLON [bit24]	0b1
RCC_CFGR	SW [bits 1-0]	0b10
	PPRE1 [bits 10-8]	0b100
	PLLMUL [bits 21-18]	0b1110
RCC_APB2ENR	AFIOEN [bit 0]	0b1
	IOPAEN [bit 2]	0b1
	IOPBEN [bit 3]	0b1
	IOPCEN [bit4]	0b1
	IOPDEN bit 5]	0b1
	TIM1EN [bit 11]	0b1
RCC_APB1ENR	PWREN [bit 28]	0b1

With these settings, the microcontroller is configured to run with a 64MHz system clock using the internal oscillator and the PLL, also enabling clock at the GPIO pins on every port, enabling the timer functionality to Timer 1 and enabling the power interface clock.

After that, it is necessary to configure the Timer 1 for its PWM functionality, in this case set to the Channel 1. It is performed by configuring the following registers and bits present in Table 2.

Table 2. Register and bit values for TIM1 CH1 PWM configuration

GPIOA_CRH	MODE8 [bits 1-0]	0b10
	CNF8 [bits 3-2]	0b10
	CNF13 [bits 23-22]	0b10
	CNF14 [bits 27-26]	0b10
TIM1_CR1	CEN [bit 0]	0b1
TIM1_CCMR1	OC1PE [bit 3]	0b1
	OC1M [bits 6-4]	0b110
TIM1_ARR		0x0000F9FF
TIM1_DCR	DBA [bits 4-0]	0b00001
TIM1_CCMR2	CC1E [bit 0]	0b1

With these settings, the microcontroller's Timer 1 counter and preload are enabled, Output compare 1 mode is set to PWM mode 1 and the OC1 signal is output on the corresponding output pin. Also, the auto-reload register is set to 63999 to provide a 1kHz PWM signal, with a voltage resolution of 51,56µV. At last, the PA8 pin is configured to output the PWM signal, also configuring the debug pins (PA13 and PA14).

With these parameters configured, the last step is to define the duty cycle of the PWM signal generated, which corresponds to the mean percentage of the output voltage generated by the PWM-based DAC. It is performed by setting the TIM1_CCR1 counter register 16-bit value. As the auto-reload register value is set to 63999, the maximum possible value expected for the TIM1_CCR1 register is 63999, that is equivalent of setting a 100% duty cycle (or a 3.3V voltage at the output). The minimum value is 0, equivalent of setting a 0% duty cycle and, therefore, a null voltage at the output. As the counter has a linear relationship, the expression for calculating its value based on the duty cycle ($0 \leq D \leq 1$) is:

$$CCR1 = D \cdot (63999 - 1)$$

In this case, as the registers only accept integer values, it is necessary to round or truncate the result to the nearest integer.

For the matter of an example application, the PWM duty cycle is configured to 50% and 30% values, to compare both outputs when connected to the filter design in the next section. Following the expression above:

$$CCR1_{50\%} = 0{,}5 \cdot (63999 - 1) = 31999$$

$$CCR1_{30\%} = 0{,}3 \cdot (63999 - 1) = 19199$$

2.2 Hardware implementation

Based on the circuit present in the section 1.3.3 and on the frequency configured for the PWM signal carrier in the section 2.1, the 2nd order low-pass filter was designed for a cut-off frequency of 200Hz, which will allow around -30dB attenuation at 1kHz, sufficient to remove the PWM carrier contribution in the output signal. Considering nominal values for the components and a 1% tolerance, the selected ones are displayed in Table 3.

Table 3. 2nd order low-pass filter components

Component	Value
C_1	220nF
C_2	100nF
R_1	3,9kΩ
R_2	7,5kΩ

Based on these values, the cut-off frequency can be calculated, as:

$$f_0 = \frac{1}{2\pi}\frac{1}{\sqrt{R_1 C_1 R_2 C_2}} = \frac{1}{2\pi}\frac{1}{\sqrt{3{,}9 \times 10^3 \cdot 7{,}5 \times 10^3 \cdot 220 \times 10^{-9} \cdot 100 \times 10^{-9}}}$$

$$f_0 = \frac{1}{2\pi}\frac{1}{\sqrt{643{,}5 \times 10^{-9}}} \rightarrow \boxed{f_o = 198{,}401Hz}$$

This cut-off frequency value error is lower than the 1% tolerance admitted, making the parameters adequate to the application.

3 Measurements

The output signals for 50% duty cycle and 30% duty cycle measured from the implemented circuit and software are presented in Figures 6 and 7.

Figure 6. 50% duty cycle 1kHz PWM experimental signal

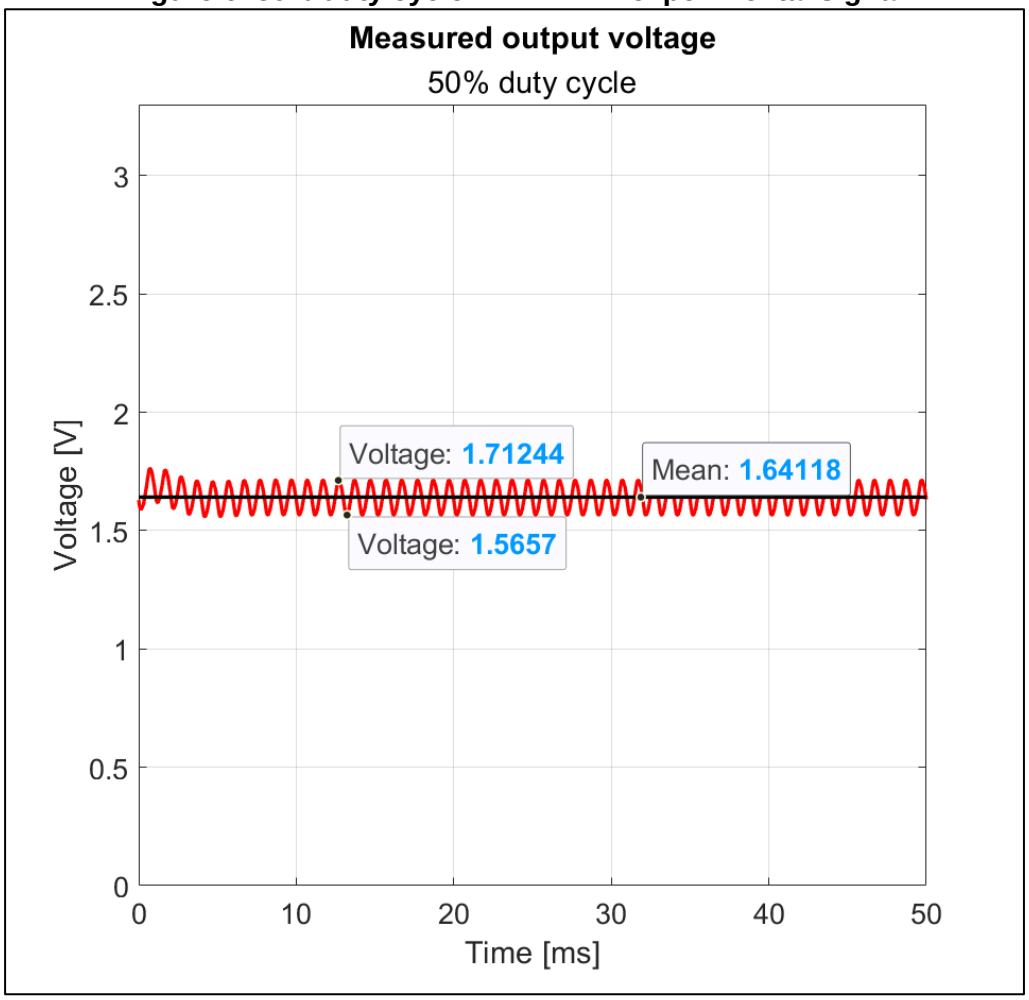
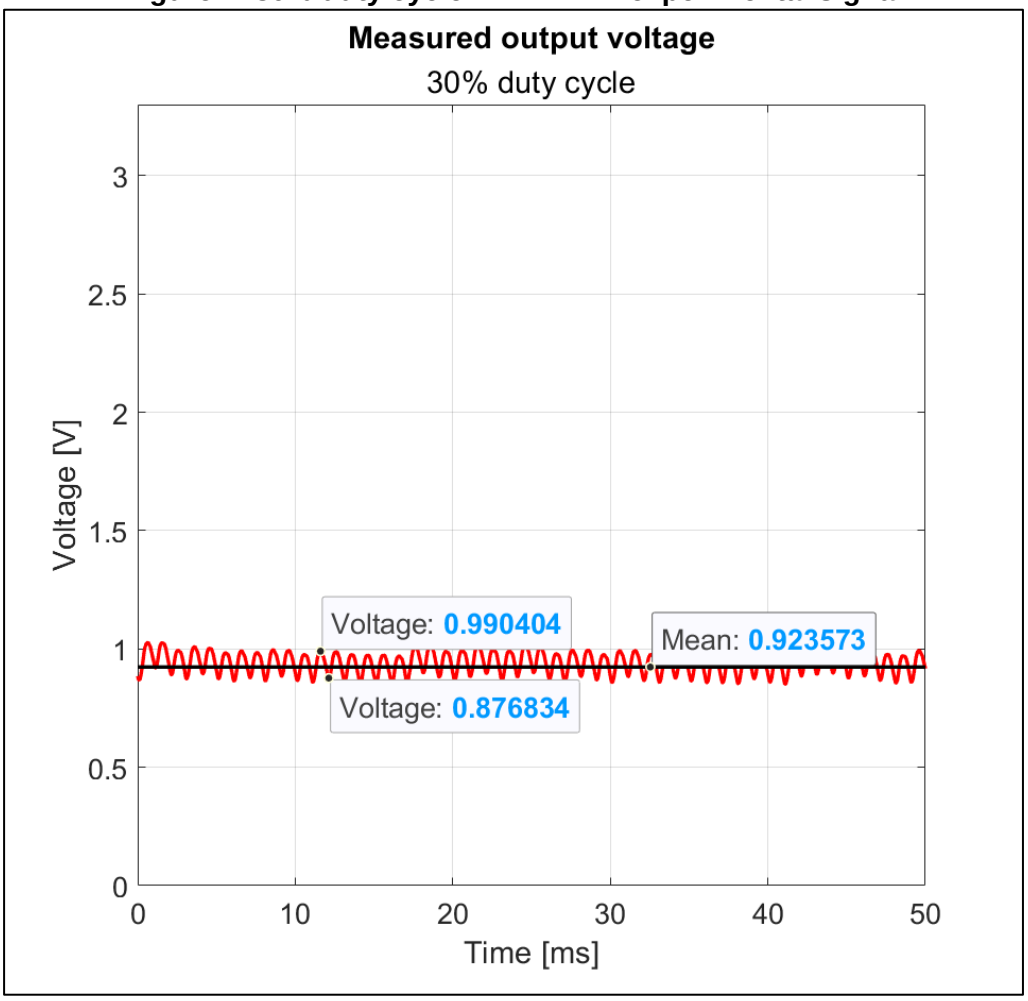


Figure 7. 30% duty cycle 1kHz PWM experimental signal



As seen on the images above, considering that the microcontroller output on voltage is 3,3V and the reference value is equal to the mean voltage present in the output signal, the value when selecting 50% duty cycle was expected to be 1,65V and it was measured 1,641118V, which corresponds to a 0,538% error. In the 30% duty cycle scenario, it was expected a 0,99V output and the value measured was 0,923573V, corresponding to a 6,71% error.

4 Conclusions

The PWM-based DAC developed for the STM32F103RB microcontroller (which do not provide a regular DAC peripheral) was able to perform as expected, operating from a regular PWM signal, and generating a variable analog value. Considering the error associated to the expected value and de measured voltage, it may be stated that the mean voltage value is closer to the reference when using higher duty cycle values, increasing as the duty cycle is reduced.

Additional remarks:

- by selecting a smaller register period, it is possible to increase the PWM oscillating frequency and increase the analog output signal frequency as well.
- by using an on-chip ADC feedback it is possible to perform an output control to ensure the correct output value, even with disturbances present.

5 Revision history

Table 4. Document revision history

Date	Revision	Changes
01-Mar-2021	1	Initial release.