

Arbeitsweise reziproker Frequenzzähler Pico-Fmeter2

Die Funktionen und Routinen zur Frequenzmessung beim Pico-Fmeter2 finden sich in der Datei „fm2-RP2040-3.c“. Die Kommentare zu den Programabschnitten erklären kleine Details in der Ausführung. Diese Beschreibung soll dazu dienen, die Arbeitsweise der Frequenzmessung zu beschreiben, um das Wie und das Warum deutlich zu machen.

zum Messverfahren:

Die Frequenz eines Signals ist definiert als Anzahl seiner Schwingungen pro Sekunde. Im einfachsten Fall zählt man für genau eine Sekunde die Schwingungen in einem Zähler und erhält direkt das Ergebnis in Hertz (Hz).

Bezogen auf ein Signal mit 10 MHz mit der Unsicherheit der Zählung um ± 1 können sich Zählerstände von 9999999, 10000000 und 10000001 ergeben, was einer 7-stelligen Auflösung entspricht. Bei höheren Frequenzen erhöht sich die Auflösung, bei niedrigen sinkt sie und wird deutlich schlechter wenn nur noch wenige Hz gemessen werden müssen.

So ist bei Messung der 50 Hz Netzfrequenz auch bei hochgenauer Messzeit von 1 s das Ergebnis von 49, 50 oder 51 mit diesem Messverfahren mit 2% nur noch auf zwei Stellen genau.

Das ist das „klassische“ Frequenzzähler Messverfahren, wie es noch ohne Unterstützung von μCs üblich war.

Benötigt man höhere Auflösung bei niedrigen Frequenzen ist es daher notwendig, nicht mehr die Ereignisse/Zeit zu zählen, sondern die Zeit für eine einzige Schwingung (Periode) zu messen und diese durch Kehrwertbildung auf die Frequenz umzurechnen.

Ausgehend von einer typischen Referenzfrequenz (Fref) von 10 MHz und einer Periodendauer (t) von 20 ms des 50 Hz Signals ergibt die Zeitmessung einen Zählerstand von 200000. Die Frequenz erhält man aus der Umrechnung: $f = 1/t * \text{Fref}$.

Da die Auflösung der Zeitmessung auch hier ± 1 beträgt, ergibt sich für die errechnete Frequenz eine Auflösung von gut 5 gültigen Stellen. Das ist eine deutliche Verbesserung gegenüber der ursprünglichen Messung mit 2-stelliger Auflösung.

Mit dem 1. Messverfahren kann man sehr hohe Frequenzen hochaufgelöst messen und mit dem 2. sehr niedrige. Selbst, wenn man bereit ist, je nach Eingangsfrequenz das Messverfahren umzuschalten, bekommt man bei ca. 3,162 kHz ($\sqrt{1} * 10e7$) mit beiden Verfahren nur eine Auflösung von rund 0,03%, was ein schlechter Wert ist.

reziproke Frequenzmessung:

Um eine hohe Auflösung unabhängig von der jeweiligen Eingangsfrequenz zu erhalten, kann man beide Verfahren derart kombinieren, daß zum einen die zuvor benötigte 1 s Messzeit in Anspruch genommen wird und zum anderen nicht nur eine einzelne Periode des Signals sondern so viele Perioden gezählt werden, bis 1 s Messdauer erreicht oder überschritten wurde.

Dafür werden jetzt zwei Zähler benötigt, von den einer die Anzahl der Perioden (N) des Signals zählt und ein zweiter Zähler die genaue Zeit (T) einer oder mehrerer Perioden. Ganz wichtig ist, daß der Zähler für die Zeitmessung synchron zu Beginn einer Periode gestartet und zum Ende der letzten Periode synchron gestoppt wird. Nach einer Sekunde Messdauer ist bei den angenommenen 10 MHz als Referenz für die Zeitmessung die Auflösung der Messung

jetzt $\geq 1 \cdot 10^7$ und das Ergebnis damit 7-stellig. Die Anzahl der gezählten Perioden ist immer ganzzahlig, wodurch die Auflösung nicht nachteilig beeinflusst wird.

Verwendet man schnelle Zähler für die Anzahl der Perioden und die Messzeit über alle Perioden muß der Messbereich nicht mehr eingestellt werden.

Das Messergebnis ergibt sich als: $f = N/T \cdot f_{ref}$.

Für den einfachsten Fall kann ein Messablauf etwa so aussehen:

1. Rücksetzen der Zähler für N und T
2. Warten auf Beginn einer Periode und beide Zähler synchron zählen lassen
3. Zähler zählen lassen, bis ≥ 1 s Messdauer abgelaufen ist: $T \geq 1 \cdot 10^7$
4. Warten auf Beginn einer Periode und synchron dazu beide Zähler stoppen

Die Periode, die zum Stopp der Messung führte, wird noch gezählt, ihre Periodendauer aber nicht mehr beachtet. Das ist deshalb korrekt, da für die 1. Periodendauer dieser Zähler ja noch auf ,0' stand. Die Anzahl der Perioden (N) sowie die gesamte Messzeit (T) passen somit zueinander und können ausgewertet werden.

Anschließend wird eine neue Messung mit gleichem Ablauf gestartet.

Jetzt kann man sich für den beschriebenen Ablauf eine elektronische Schaltung vorstellen, die mit separaten Zählern, Gattern und FlipFlops arbeitet. Soweit so gut zeigt sich bei dem beschriebenen Ablauf ein Haken: zwischen zwei Messungen wird mindestens eine Periode nicht berücksichtigt, die durch das Auslesen der Zählerstände entfällt. Bei höheren Frequenzen mag das belanglos sein, aber schon beim Abgleich eines Sekundenimpulses für eine Uhr, gibt es nur alle zwei Sekunden ein neues Ergebnis.

Da für die Frequenzberechnung ein Mikrocontroller notwendig ist, kann man sich für Frequenzmessungen bis in den unteren MHz-Bereich separate Zähler sparen, indem man geschickterweise die bereits μC integrierten Zähler nutzt. Diese Zähler verfügen zumeist über eine ,Capture'-Einheit, die den gegenwärtigen Zählerstand synchron zu einem externen oder internen Ereignis speichern kann. Man kann diese Funktion dazu nutzen, lückenlos alle Perioden eines Signals zu erfassen und auszuwerten. Anstatt Zähler zurückzusetzen, wird der letzte Zählerstand zu einem ,Capture'-Ereignis als ,0'-Referenz betrachtet und die Zählerwerte einer neuen Messung aus „neuer Wert“ – „alter Wert“ errechnet. Der letzte ,neuer Wert“ wird für die nächste Messung zum „alter Wert“.

Rechnet man mit vorzeichenlosen Zahlenwerten, müssen Überläufe der Zähler nicht beachtet werden. Einmalig gestartet laufen die Zähler einfach weiter.

Soweit die allgemeinen Voraussetzungen für eine reziproke Frequenzmessung.

Reziproke Frequenzmessung auf dem RP2040:

Der auf dem Pico-Board verwendete Controller RP2040 ist schnell, hat einen leistungsfähigen M0+ Kern (sogar doppelt, von denen hier nur einer verwendet wird) mit viel Speicher aber keinen einzigen Zähler mit einer ‚Capture‘-Einheit, die zwingend für eine reziproke Frequenzmessung erforderlich ist. Das ist zunächst ganz schlecht!

Der RP2040 hat jedoch zwei schnelle mikroprogrammierbare PIO-Einheiten, die man dazu nutzen kann, hinreichend schnelle Zähler mit Capture-Funktion nachzubilden. Wie zuvor beschrieben werden ein Zähler für die Periodendauer und ein Zähler für die Anzahl der Perioden benötigt.

Pro PIO gibt es 4 x SM-Einheiten („state machines“, Zustandsautomaten), von denen hier jeweils eine benötigt wird, um einen Zähler mit Capture-Funktion zu erhalten. Mit maximaler Frequenz betrieben, erfaßt eine SM-Einheit positive und negative Flanken des Eingangssignals wobei zusätzlich ein fortlaufend hochzählendes x-Register verwendet wird, um die genauen Zeitpunkte der Flankenwechsel zu erhalten.

Mit jeder fallenden Flanke wird der Zeitpunkt in ein Ausgaberegister zur Weiterverarbeitung geschrieben und damit für jede abgeschlossene Periode ein Zeitstempel erzeugt. Optimal programmiert wird das x-Register mit 1/4 der Taktfrequenz der SM erhöht und die gemessene Frequenz kann bei jeweils zwei Flankenwechseln pro Periode theoretisch 1/8 der SM-Taktfrequenz betragen. Der praktische Wert liegt knapp darunter, sodaß beispielsweise bei 133 MHz Prozessortakt und 33,25 MHz der SM-Einheit gut 15 MHz gemessen werden können.

Diese Zeitstempel liefern nun die genauen Zeitpunkte für jede einzelne Eingangsperiode.

Jetzt wird noch ein Zähler für die Anzahl der Perioden benötigt. Dafür bietet der RP2040 DMA-Einheiten, von denen eine benötigt wird, um die Zeitstempel aus dem PIO-Ausgangsregister fortlaufend an eine feste Adresse ins RAM zu schreiben und somit für die anschließende Auswertung verfügbar zu machen. Bei jedem DMA-Transfer wird der Transfer-Zähler des DMA-Kanals dekrementiert. Dieser Zähler ist 32 Bit breit, wird einmalig auf den maximalen Wert von 0xffffffff gesetzt und liefert beim Vergleich „alter Wert“ – „neuer Wert“ vorzeichenrichtig die Anzahl der gezählten Perioden. Um den aktuellen Zeitstempel und den aktuellen Wert des Transfer-Zählers auszulesen, muß der DMA-Kanal gestoppt werden. Nur so ist sichergestellt, daß beide Werte zueinander gehören. Würde der DMA-Kanal zu lange gestoppt werden, würden Zeistempel verloren gehen und das gesamte Messergebnis unbrauchbar machen.

Das Ausgangsregister eines PIO-SM-Kanals bietet einen Fifo-Speicher mit einer Tiefe von acht Werten. Ein verzögertes Auslesen bei gestopptem DMA-Kanal bremst die Erfassung neuer Zeitstempel daher nicht aus.

Bei 15 MHz max. Eingangsfrequenz und 8-facher Pufferung darf der DMA-Kanal folglich max. 0,53 µs angehalten werden, was bei 133 MHz Taktfrequenz 70 Prozessortakten entspricht. Bei ≤ 10 Prozessortakten, die für das Auslesen von DMA-Transfer-Zähler und Zeitstempel aus dem RAM benötigt werden, ist genug Sicherheit gegeben, die Messwerte fehlerfrei auszulesen.

Bis zu einer Eingangsfrequenz von 15 MHz stehen somit zu jede Eingangsperiode die genaue Anzahl sowie synchron dazu der genaue Zeitpunkt zur Verfügung.

Ein DMA-Kanal stoppt selbsttätig, wenn sein Transfer-Zähler auf 0 heruntergezählt ist. Per Interrupt wird er mit kurzer Unterbrechung wieder auf den Startwert gesetzt. Dieser kurze Stopp ist wie zuvor beschrieben durch die Fifo-Zwischenspeicherung kein Problem.

Die Auflösung der Zeitbasis ist wie erwähnt $\frac{1}{4}$ Prozessortakt und, um bei dem obigen Zahlenbeispiel zu bleiben, mit 1/33,25 MHz rund 30 ns. Möchte man eine flüssige Anzeige der Messwerte mit drei Messungen/s erhalten, ergibt sich die Auflösung für jede Messung von $0,3 \text{ s} / 3 \cdot 10^7 \text{ s} = 1 \cdot 10^{-7}$. Jedes Messergebnis liefert somit 7 gültige Stellen (oder anders ausgedrückt 0,1 ppm) unabhängig davon, ob nun 1 Hz oder 10 MHz gemessen werden.

Eine PIO-Einheit mit ihren 4 x SM-Einheiten in Verbindung mit 4 x DMA-Kanälen könnten gleichzeitig 4 unabhängige Frequenzmessungen durchführen. Benötigt man höhere Eingangsfrequenzen, so kann man einen 2. Messkanal programmieren, der mit einem festen Vorteiler betrieben wird und ebenfalls 7-stellige Ergebnisse / 0,3 s liefert. Damit ist die reziproke Messfunktion des Pico-Fmeters beschrieben.

Verbesserte Messauflösung bei höheren Eingangsfrequenzen:

Aus der Beschreibung zum Messverfahren wird deutlich: bei der Messung niedriger Frequenzen von zum Beispiel 1 Hz ist eine erhöhte Auflösung des Ergebnisses nur durch eine erhöhte Auflösung bei der Zeitmessung möglich. Das würde zusätzliche Hardware erfordern, was hier nicht vorgesehen ist und deshalb hier nicht weiter ausgeführt wird.

Vielfach wird eine höhere Auflösung zur Vermessung eines lokalen Frequenznormal von 10 MHz gewünscht bzw. benötigt. Bei einem leicht erhöhten Eingangssignal von 0,1 Hz bei 10 MHz liefern Messungen im Abstand von 1 s etwa $9 \times$ das Ergebnis 10,000000 MHz und dann $1 \times 10,000001$ MHz. Beide Werte sind „richtig“ da sie im Rahmen der ± 1 Auflösung liegen.

Wie oben beschrieben sind bei 10 MHz innerhalb einer Sekunde $1 \cdot 10^7$ Einzelwerte zu jeder Periode verfügbar, werden aber bei sekundlicher Auswertung nicht weiter beachtet, da hier nur die Messpunkte Startwert und Endwert beachtet werden.

Die Statistik bietet mit der linearen Regression ein Verfahren an, um bei vielen Einzelwerten eine Gerade durch die Messpunkte zu legen, die nicht mit dem Start- und Endpunkt der Einzelpunkte exakt übereinstimmen muß. Die Steigung der Geraden berücksichtigt dabei auch schwankende ± 1 Zwischenwerte und liefert eine höhere Auflösung der Messungen um den Faktor \sqrt{N} , wobei N die Anzahl der berücksichtigten Messwerte ist.

Bei 100 Einzelwerten ergibt sich mit Faktor 10 eine Stelle mehr Auflösung, bei 10000 schon Faktor 100 entsprechend zwei Stellen. Es bietet sich an, aus den verfügbaren Zeitstempeln einzelne Zwischen-Messwerte zu errechnen und statistisch zu bewerten. Um möglichst viele einzelne Messwerte in die Auswertung einfließen zu lassen, ist ein möglichst schneller μC notwendig. Der RP2040 ist auch in diesem Punkt eine gute Wahl.

Um für die lineare Regression möglichst viele Stützstellen zu erhalten, wird die Auswertung von 100000 Zeitstempeln/s (10 μ s Raster) angestrebt. Unterhalb von 100 kHz wird so jeder Zeitstempel verwendet und oberhalb 100 kHz das letzte Intervall von rund 10 μ s (synchronisiert mit den Zeitstempeln).

Während eine Messung läuft, werden die einzelnen Zeitstempel für die Endauswertung vorbereitet. Additionen und Multiplikationen der Einzelwerte müssen mit hoher Auflösung durchgeführt werden, um keine Überläufe bei den Berechnungen auftreten zu lassen. Hier werden 64 Bit Variablen verwendet, die eine gewisse Rechenzeit beanspruchen und die Auswerterate auf 100 kHz begrenzen. Der endgültige Messwert wird nach Ablauf der vorgegebenen Messdauer (typisch 1 s) errechnet.

Bei 1 Hz Signalfrequenz gibt es wie gehabt nur einen Start- und einen Endwert. Es bleibt bei der 7-stelligen Auflösung. Bei 10 MHz erhöht sich bei einer Messdauer von 1 s die Auflösung theoretisch um den Faktor 316 ($\sqrt{100000}$) auf 10-stellige Auflösung:

$$1 / (33,25 * 1e6 * 316) = 0,95 * 1e-10$$

Es wird weiterhin lückenlos gemessen und jede Periode des Eingangssignals berücksichtigt!

Mit dieser theoretischen Auflösung könnte man prima angeben, wenn da nicht die Praxis einige Schranken setzen würde. Die PIO-Einheiten innerhalb des RP2040 stören sich leicht, was erst bei dieser hohen Auflösung sichtbar wird.

Nur mit einem hochgenauen externen 10 MHz Referenzsignal ergibt diese erhöhte Auflösung überhaupt einen Sinn. Zudem muß die interne Auswertung von ext. Fref abgeschaltet werden. Es gibt noch weitere Einschränkungen, die hier aber nicht weiter ausgeführt werden. Es ist daher immer ratsam, die erreichte Auflösung durch Langzeitmessungen abzusichern.

Die Regressionsberechnung ist beim Pico-Fmeter2 permanent aktiv.

Höhere Eingangsfrequenzen:

Bislang wurde die Grundfunktion der Messung beschrieben, die bis etwa 15 MHz ohne Vorteiler auskommt. Für höhere Eingangsfrequenzen muß ein zusätzlicher Vorteiler verwendet werden. Beim Pico-Fmeter2 wird ein 4:1 Teiler bestehend aus 2 x D-FFs (74AUP2G80) verwendet, der laut Datenblatt bis 600 MHz funktioniert. Um weitere Hardware zu sparen, wird das 4:1 geteilte Signal noch einmal durch einen internen Zähler (PWM-Kanal) um 8:1 geteilt, sodaß das gesamte Teilverhältnis 32:1 beträgt. Mit zugeschaltetem Vorteiler muß das Ergebnis passend skaliert werden, Auflösung oder Genauigkeit werden dabei nicht reduziert!

Die Eingangsfrequenz eines als Vorteiler verwendeten PWM-Kanals beträgt:
maximale μ C-Taktfrequenz / 2.

Bei typisch 133 MHz μ C-Takt sind es real rund 63 MHz. Mit dem vorgeschalteten 4:1 Teiler erhöht sich die maximale Eingangsfrequenz somit auf rund 250 MHz.

Bei ersten Programmversionen wurden mit einer PIO und 2 x SM-Einheiten zum einen die Eingangsfrequenz direkt und zum anderen die hinter dem Vorteiler parallel gemessen und ab einer höheren Eingangsfrequenz das Ergebnis des Kanals mit Vorteiler angezeigt. Der Vorteil war keinerlei Umschaltverzögerung bei der Anzeige. Bei 7-stelliger Auflösung ist dieses Verfahren problemlos anwendbar.

Nach der Ergänzung um die Regressionsberechnung, wurden in den nun höher aufgelösten Ergebnissen Störungen sichtbar, die dazu führten, daß nur noch eine SM-Einheit verwendet werden konnte und für die Auswertung des vorgeteilten Signals der Eingangspin der SM-Einheit entsprechend umgeschaltet wird. Die Umschaltung auf den Kanal mit Vorteiler erfolgt ab ≥ 8 MHz und das Zurückschalten bei ≤ 4 MHz.

Diese Frequenzen wurden so hoch gewählt, daß eine Eingangsfrequenz von > 4 MHz bei aktivem Vorteiler vor der Regressionsberechnung noch 100000 Zeitstempel und damit maximale Auflösung liefern kann.

Um eine schnelle Umschaltzeit bei Erhöhung der Eingangsfrequenz in den oberen MHz-Bereich zu erhalten, wird parallel zu der normalen Messung noch eine klassische grobe Frequenzzählung (wie eingangs erläutert) mit einer festen Torzeit von 1 ms durchgeführt und permanent überwacht.

Nun gibt es zwei Möglichkeiten, mit dem Messwert bei einer Umschaltung des Vorteilers umzugehen. Man kann entweder das letzte Ergebnis komplett verwerfen und eine neue Messung starten oder aber auch bei der permanent aktiven Regressionsberechnung ein paar dieser 100 kHz Zeitstempel verwerfen, um beim Umschalten die Messrate nicht zu reduzieren. Beim Pico-Fmeter2 ist letztere Methode eingestellt. Bei sehr langsamer Frequenzänderung um die Umschaltswelle kann bei hoch eingestellter Auflösung eine Störung in 8. – 9. Stelle beobachtet werden. Verwirft man hingegen die komplette Messung, ergibt sich eine Verzögerung für einen neuen Messwert, der aber völlig störungsfrei ist. Eine Feinheit am Rande.

noch schneller:

Zur Messung noch höherer Eingangsfrequenzen kann man den Prozessortakt per Option verdoppeln: es wurde keine Fehlfunktion beobachtet.

Diese Option wird einmalig beim Einschalten getestet.

Wenn benötigt, kann sie verwendet werden, andernfalls beläßt man die Arbeitsfrequenz bei den laut Datenblatt angegebenen 133 MHz.