

Ulrich Bangert, DF6JB

Über die Stabilität von Oszillatoren und Frequenznormalen

Bei der kommenden AMSAT-Mars-Mission werden sowohl im Satelliten selber als auch bei den ihn steuernden Kommandostationen so genannte USOs zum Einsatz kommen. USO ist die Abkürzung für **Ultra Stable Oscillator**. Grund genug, sich einmal prinzipiell mit der Stabilität von Oszillatoren zu beschäftigen.

Die hohen Anforderungen an die Stabilität dieser Oszillatoren resultieren einerseits aus der projektierten Kommunikationstechnik. Angesichts der großen Entfernung zwischen Erde und Mars, die aufgrund der Umlaufbahnen der beiden Himmelskörper zwischen 55 und 400 Millionen Kilometern schwankt, ist die Kommunikationstechnik anspruchsvoll, um es einmal vorsichtig zu formulieren.

Insbesondere der Fall, dass die Richtantenne des Satelliten nicht korrekt auf die Erde ausgerichtet ist und die Kommandostationen den Satelliten über dessen Rundstrahler administrieren müssen, geht an die Grenze des mit Amateurmitteln technisch machbaren. Trotz der Verwendung von Turbo-Codes¹ wird die auf BPSK beruhende Kommunikation in diesem Fall mit maximal 5 Bit / s stattfinden, um mit minimaler Bandbreite im Empfänger arbeiten zu können. Diese geringe Übertragungsrate hat zur Folge, dass die PLL, die im Empfänger zur Rückgewinnung der Symbolrate dient, entsprechend große Zeitkonstanten haben muss. Dies wiederum hat zur Folge, dass, wie Karl Meinzer errechnet hat, die gesamte Frequenzaufbereitung in Sender und Empfänger keine größere Allan-Deviation² als 10^{-12} bei $\tau = 1$ s haben darf.

Andererseits kommt für die bei den Kommandostationen verwendeten Oszillatoren eine zweite Forderung hinzu, die mit der Steuerung des Satelliten zu tun hat. Für ein sicheres Einschwenken in den Marsorbit ist es erforderlich, die Position des Satelliten auf einige 10 m (kein Druckfehler!) genau bestimmen zu können. Die Positionsbestimmung wird über die Laufzeitmessung an einem Funksignal realisiert, welches von der Bodenstation an den Satelliten geschickt und von diesem (mit einer vermessenen und daher bekannten Verzögerung) beantwortet wird. Dass der Satellit (und auch die Erde) dabei nicht ruht sondern sich im Gegenteil mit hoher Geschwindigkeit bewegt, ist zwar eine Verkomplizierung der Situation, lässt sich aber mathematisch berücksichtigen.

Um eine Strecke von 10 m zurückzulegen, benötigt eine elektromagnetische Welle im Vakuum ungefähr 33 ns. Die Laufzeit der Welle von der Erde zum Satelliten und zurück liegt bei einer angenommenen Entfernung von 300 Millionen Kilometer bei ca. 2000 s. Daraus leitet sich ab, dass man bei der über ca. 2000 s dauernden Zeitmessung einen maximalen Fehler von 33 ns machen darf. Dies ist daher ein zweites Stabilitätskriterium, das der zum Messen benutzte Oszillator einhalten muss. Bislang geht man davon aus, dass dies zweckmäßigerweise der gleiche ist, der auch für die Frequenzaufbereitung maßgeblich ist, dies muss aber nicht unbedingt so sein.

Im Prinzip lässt sich auch dieses Kriterium in Form einer Allan-Deviation formulieren, die – diesmal für eine Beobachtungszeit $\tau = 2000$ s – nicht überschritten werden darf. Es zeigt sich jedoch, dass hier eine andere statistische Größe genannt **MTIE** - = **Maximum Time Interval Error** – hier noch besser anwendbar ist, da sie ganz unmittelbar die Problematik be-

¹ Hierbei handelt es sich um die derzeit effektivste bekannte „Forward Error Correction“- Methode

² Was das ist, darauf gehen wir gleich genauer ein

schreibt, die wir hier vorliegen haben: Den maximalen Fehler in den Messungen von Zeitintervallen zu ermitteln, wenn der fragliche Oszillator als Referenz benutzt wird.

In diesem Beitrag möchte ich versuchen, dem Leser mit möglichst wenig Mathematik zu erklären, wie die Stabilität eines Oszillators überhaupt definiert ist und wie die bei Stabilitätsmessungen angewandte Messtechnik beschaffen ist. Die Erkenntnisse, die wir dabei gewinnen werden, sind kein reiner Selbstzweck. Vielmehr wird sich sehr schnell zeigen, dass diese Erkenntnisse auf einmal auch ein ganz neues Licht auf den Bau von Frequenznormalen werfen werden und Antworten auf Fragen geben, wie sie in dieser Form meines Wissens nach noch nicht veröffentlicht worden sind.

Die Allan-Deviation

Wir wollen zunächst einmal annehmen, wir hätten eine fehlerlose Messeinrichtung zur Verfügung, mit der die Frequenz eines Oszillators innerhalb einer Messzeit von 1 s mit beliebiger Auflösung gemessen werden könnte. Dies ist eine Idealisierung, die nicht wirklich existiert, wir wollen aber so tun als ob, weil es hier zunächst nur um die **Weiterverarbeitung** der mit dieser Anordnung gewonnenen Werte geht. Mit dieser Messeinrichtung wollen wir nun einen Oszillator vermessen.

Diese Messeinrichtung würde also im Sekundentakt Messwerte der Frequenz dieses Oszillators produzieren, die wir statistisch auswerten können. Würden wir immer den gleichen Messwert erhalten, so hätten wir einen idealen unendlich stabilen Oszillator vorliegen. Wie sich bei Messungen an realen Oszillatoren zeigt, erhalten wir aber nicht immer den **gleichen** Messwert, sondern die Messwerte **schwanken** mit einer gewissen Breite um einen mittleren Wert herum. Da die Messeinrichtung selbst als fehlerlos angenommen wird, muss es sich bei diesen Schwankungen also um eine Eigenschaft des Prüflings handeln, eben um die Tatsache, dass er nur begrenzt „stabil“ ist.

Ich habe im vorletzten Satz bewusst den Begriff „Mittelwert“ vermieden und von einem „mittleren Wert“ geredet, weil der Mittelwert ja bereits ein feststehender Begriff in der Statistik ist. Tatsächlich würde aber jeder, der sich jemals mit statistischen Methoden beschäftigt hat, nun ganz unmittelbar auf die Idee kommen, aus den einzelnen Messwerten der Frequenz zunächst einen Mittelwert zu errechnen. Und wer jemals Statistik betreiben musste, der weiß, dass der nächste Schritt die Berechnung der Varianz bzw. der Standardabweichung sein muss.

Dies sind statistische Größen, die Aussagen über die Breite der statistischen Verteilung der Messwerte um den Mittelwert herum machen. Im Prinzip wäre die Standardabweichung schon ein ganz gutes Maß für die Stabilität des Oszillators, weil damit eine Aussage wie die folgende möglich wäre:

Wenn ich die Frequenz des Oszillators im Sekundentakt messe, so beobachte ich dabei statistisch verteilte Abweichungen vom Mittelwert von mal soundsoviel Hz und mal soundsoviel Hz und zwar in positive wie in negative Richtung. Der „**Mittelwert der Abweichungen**“, also wenn man so will, die „typische“ Schwankung von Sekunde zu Sekunde, beträgt dabei soundsoviel Hz. Die Anführungszeichen bei „Mittelwert der Abweichungen“ kommen daher, dass bei einer gewöhnlichen arithmetischen Mittelwertbildung sich positive und negative Abweichungen aufheben würden und genau das wollen wir hier gerade nicht. Positive und negative Abweichungen sollen sich nicht aufheben sondern im Gegenteil im gleichen Maße zum „typischen“ Wert der Abweichung beitragen. Deswegen summiert man bei der Berech-

nung der Standardabweichung. die **Quadrate der Abweichungen** auf und zieht aus der Summe wieder die Wurzel.

Die Standardabweichung wäre in der Tat ein geeignetes Maß für die Stabilität von Oszillatoren, wenn nicht der amerikanische Physiker David W. Allan herausgefunden hätte, dass es unter den Zufallsprozessen, die innerhalb eines Oszillators für Instabilitäten verantwortlich sind, auch solche gibt, für die die Standardabweichung kein adäquates Maß ist, weil bei Vorliegen dieser Art von Zufallsprozessen die Standardabweichung mit zunehmender Zahl der Messwerte nicht mehr auf einen endlichen Wert konvergiert sondern unendlich wird.

Wenn in den Naturwissenschaften Zahlen entweder unendlich groß oder unendlich klein werden, ist man entweder nahe daran, sich einen Nobelpreis zu verdienen, oder man ist einem groben Fehler in der Methode aufgesessen. Allan war der Ansicht, das zweite sei vermutlich der Fall und kam zu der Erkenntnis, dass aus technischen Gründen die Standardabweichung nicht geeignet sein kann, alle Formen von Zufallsprozessen in einem Oszillator richtig zu beschreiben.

Allan wandte einen Trick an, indem er die Formel für die Berechnung der Standardabweichung minimal so modifizierte, dass unter allen Umständen die Konvergenz der Berechnung auf einen endlichen Wert gewährleistet war. Diese neue Art der statistischen Berechnung wird ihm zu Ehren als Allan-Deviation, also als Allan-Abweichung bezeichnet. Gelegentlich werden Sie auch auf den Begriff der Allan-Varianz stoßen. So wie die Standardabweichung die Quadratwurzel aus der Varianz ist, so ist die Allan-Deviation die Quadratwurzel aus der Allan-Varianz.

Für den mathematisch Interessierten: Bei der Standardabweichung wird über die Quadrate der Abstände der einzelnen Messwerte vom Mittelwert summiert. Allan ersetzte dies durch eine Summation über die Quadrate der Abstände **aufeinander folgender Messwerte**. Wer noch mehr in die Mathematik eintauchen möchte, der findet im Internet Unmengen an Literatur darüber. Insbesondere [7] möchte ich jedem Interessierten empfehlen.

Damit diese Art der Berechnung vergleichbar mit der Standardabweichung wird, bedarf es zusätzlich eines Normierungsfaktors. Für die Praxis mag es genügen, wenn wir uns die Allan-Deviation als eine verbesserte Form der Standardabweichung vorstellen. Genau wie die Standardabweichung wird sie in Formeln durch ein griechisches Sigma σ dargestellt.

Nachdem wir nun wissen, wie Stabilität „berechnet“ wird, wollen wir uns nun mit der Frage beschäftigen, was eigentlich passiert wäre, wenn wir in unserem Beispiel nicht **jede** Sekunde die Frequenz gemessen hätten sondern vielleicht nur jede zweite oder vielleicht auch nur jede fünfte oder zehnte Sekunde.

Statt eine neue Messreihe zu machen und die Messungen wirklich in dem neuen Abstand durchzuführen können wir genauso gut hergehen und aus unserer ursprünglichen Messreihe z.B. jeden zweiten Wert streichen und so zu einer neuen Messreihe gelangen, die aus Messwerten im Zwei-Sekundentakt besteht.

Wenn wir dies tun, so ist zunächst klar, dass fast zwangsläufig der berechnete Wert der Allan-Deviation irgendwie anders sein muss als der vorher berechnete, einfach deswegen, weil wir ja mit **weniger** Zahlenmaterial rechnen. Andererseits würden wir aber nicht mit einem wirklich fundamentalen Unterschied rechnen sondern eher damit, dass die beiden Zahlen recht ähnlich sind.

Dass man aber durchaus große Unterschiede feststellen kann, die nichts mit der reinen Zahl der Messwerte zu tun haben, möge aber das folgende Beispiel verdeutlichen: Angenommen unser Oszillator hätte eine Störmodulation. Wir wollen annehmen, dass seine Ausgangsfrequenz mit einem ganz geringen Anteil (Hub) von Frequenzmodulation überlagert ist, die eine Frequenz von 0,5 Hz hat.

Egal, wie klein der Anteil auch sein mag, wir würden ihn bei den 1-Sekundenwerten **immer** feststellen können, weil wir dort ein hochgradig regelmäßiges Auf- und Ab der Messwerte beobachten. Dieses regelmäßige Auf- und Ab der Messwerte führt bei der Berechnung der Allan-Deviation (Sie erinnern sich: Es geht dabei um die Differenz aufeinander folgender Werte) dazu, dass jedes einzelne Wertepaar einen von 0 verschiedenen Anteil am Gesamtergebnis der Allan-Deviation liefert. Wie groß die Allan-Deviation in diesem Beispiel wirklich ist, hängt vom Hub der Störmodulation ab, ist aber für diese Überlegung nicht relevant. Wichtig ist: Die Störmodulation schlägt sich in der Allan-Deviation für die 1-Sekundenwerte nieder.

In den im 2-Sekundentakt gewonnenen Messwerten würden wir hingegen von dieser Störmodulation **überhaupt nichts** feststellen können, weil wir unsere Messungen in genau der gleichen Periodenlänge machen, welche die Störung selber hat. Wir erwischen bei den Messungen im Zweisekundentakt also die Störung sozusagen immer „an der gleichen Stelle“ ihrer Periodenlänge. Wir messen dadurch immer die Summe aus Mittenfrequenz plus Störung (an der gleichen Stelle) und damit eine konstante und (wegen der Störung) minimal falsche Frequenz, aber von der Tatsache, dass die Störung zu einer periodischen Änderung der Oszillatorfrequenz führt, davon sehen wir **nichts**.³ Jedes Wertepaar würde als Differenz „0“ liefern und wir erhalten (wenn keine anderen Störungen vorhanden sind) eine Allan-Deviation von „0“. Wir würden also bei 2-Sekundenmessungen den Oszillator für absolut stabil halten, obwohl er eine ausgeprägte Störmodulation aufweist und damit überhaupt nicht stabil ist!

Aus diesem Beispiel ergeben sich zwei **extrem** wichtige Schlussfolgerungen:

- Zum einen ist die Angabe einer Stabilität ohne die gleichzeitige Angabe der Beobachtungszeit für diese Stabilität zweck- und nutzlos. Angaben wie „Die Kurzzeitstabilität beträgt 1×10^{-11} “ sind technischer Unsinn, soweit man nicht genau spezifiziert, welche Beobachtungszeit man exakt unter „Kurzzeit“ verstehen möchte. Man kann auch nicht die Stabilität für einen ganzen **Zeitbereich** angeben sondern es ist eine Größe, die ganz konkret an **eine einzelne** Beobachtungszeit gekoppelt ist. Das gilt ganz analog natürlich auch für die oft zitierte „Langzeitstabilität“. Mit diesen beiden Begriffen ist in der Amateurfunkliteratur unglaublich viel Schindluder getrieben worden und Sie sollten den Sachverstand eines jeden, der Ihnen mit Hilfe solcher Begrifflichkeit etwas über Stabilität erzählen will, ohne gleichzeitig einen exakten Beobachtungszeitraum für diese Stabilität zu nennen, sehr kritisch hinterfragen. Der Beobachtungszeitraum ist dabei, um es noch einmal deutlich zu machen, nicht die gesamte Messzeit sondern der Zeitabstand derjenigen Messwerte, die man für die Berechnung der Allan-Deviation benutzt.
- Weil man a priori nicht weiß, bei **welchen** Frequenzen ein zu untersuchender Oszillator eine mögliche Störmodulation wie im Beispiel aufweist, ist die Berechnung eines

³ Diesen Effekt macht man sich in der Messtechnik seit Jahrzehnten zu nutze, in dem man bei integrierenden A/D-Wandlern (üblicherweise in digitalen Multimetern anzutreffen) die Integrationszeit auf die Periodenlänge der Netzfrequenz oder auf ein Vielfaches davon einstellt. Dabei heben sich alle netzsynchronen Fehler weg.

einzelnen Wertes der Allan-Deviation eigentlich witzlos, weil man unter Umständen wie im Beispiel mit zwei Sekunden einen Beobachtungszeitraum gewählt hat, der eine wichtige Informationen über die Eigenschaften des Oszillators genau vorenthält. Es müsste also so etwas her wie eine Darstellung der Allan-Deviation, aus der für **jede** sinnvolle Beobachtungszeit die dazugehörige Allan-Deviation abgelesen werden kann. Das gibt es tatsächlich und ist in der Horologie (der Wissenschaft der Zeit- und Frequenzmesstechnik) das Standardinstrument zur Charakterisierung der Stabilität von Oszillatoren. Es heißt: Sigma-Tau-Diagramm. Das Sigma ist das Kürzel der Allan-Deviation und das Tau kommt daher, dass in der Horologie der Beobachtungszeitraum gerne mit dem griechischen Buchstaben Tau τ dargestellt wird.

Das Sigma-Tau-Diagramm

Das Sigma-Tau-Diagramm stellt in einem doppelt-logarithmischen Diagramm über den interessierenden Bereich von Beobachtungszeiten die Allan-Deviation eines Oszillators auf der vertikalen Achse als Funktion der Beobachtungszeit Tau auf der horizontalen Achse dar.

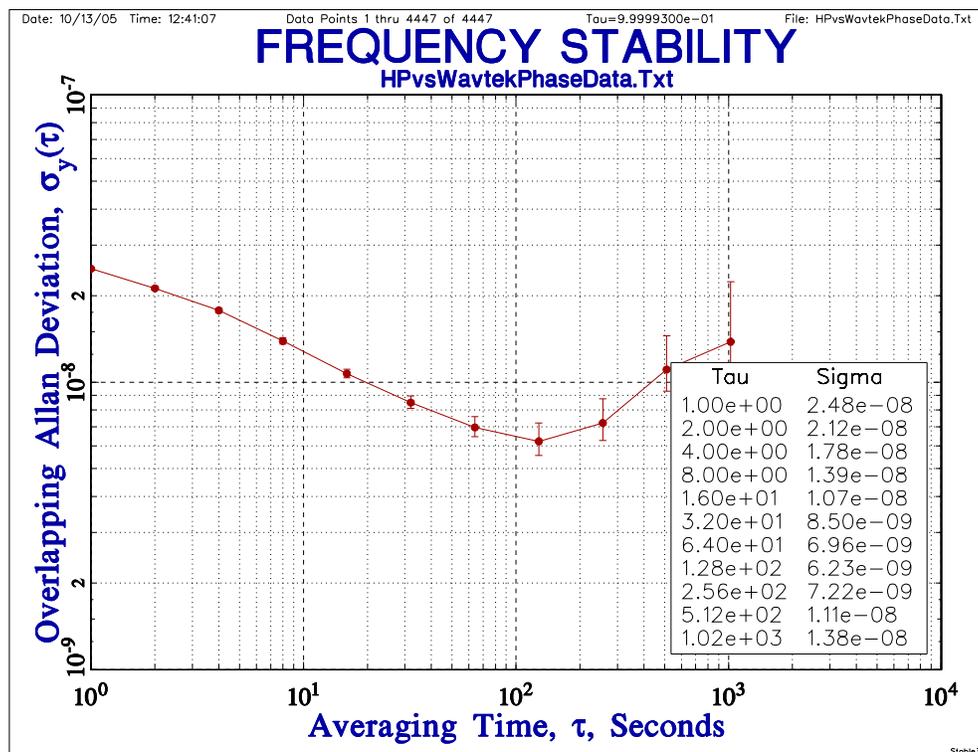


Abbildung 1: Sigma-Tau-Diagramm eines normalen Quarzoszillators

Ein Beispiel eines Sigma-Tau-Diagramms eines nicht ganz schlechten aber ansonsten gewöhnlichen Quarzoszillators⁴ ist in Abbildung 1⁵ gezeigt.

Wie man in dem Beispiel sieht, benutzt man als Beobachtungszeiten gerne solche, wo sich die jeweils nächste Beobachtungszeit als ein Vielfaches der vorherigen Beobachtungszeit ergibt, weil dies in der logarithmischen Darstellung zu äquidistanten Stützpunkten der Kurve führt. Fängt man bei 1 an und wählt als Faktor 2, so ergibt sich die im Beispiel dargestellte Abfolge von $\tau = 1, 2, 4, 8, 16 \dots 1024$ s. Zwischen den Stützpunkten interpoliert das Programm den zu erwartenden Kurvenverlauf.

Im Diagramm sind Fehlerbalken eingezeichnet, die mit zunehmender Beobachtungszeit größer werden. Der Grund dafür ist folgender: Wenn ich von vorn herein die Absicht habe, die Allan-Deviation für eine ganze Reihe von Beobachtungszeiten zu berechnen, also z.B. $\tau = 1, 2, 4, 8, 16 \dots 1024$ s, so werde ich tatsächlich so vorgehen, dass ich eine Messreihe im Sekundentakt durchführe und dann wie oben beschrieben für die Berechnung für $\tau = 2$ s jeden zweiten Wert wegstreiche, für $\tau = 4$ s von der verbleibenden Reihe jeden zweiten wegstreiche und so weiter und so fort. Durch das zunehmende Wegstreichen von Werten bleiben aber für die Berechnung der Allan-Deviation immer weniger Werte übrig, so dass die statistische Unsicherheit zu größeren Werten von τ zunimmt.

Der Vollständigkeit halber möchte ich erwähnen, dass das Diagramm nicht die „einfache“ sondern die „Overlapping“ Allan-Deviation zeigt. In der originalen Form der Allan-Deviation würde man für $\tau = 2$ s nur die Wertepaare W_1 und W_3 , W_3 und W_5 , W_5 und $W_7 \dots$ und so weiter miteinander verrechnen. Es gibt aber noch mehr Wertepaare, bei denen die Partner einen zeitlichen Abstand von 2 Sekunden haben, nämlich W_2 und W_4 , W_4 und $W_6 \dots$ und so weiter. Die Overlapping Allan-Deviation benutzt **alle** Wertepaare, bei denen die Partner den zeitlichen Abstand τ haben, für die Rechnung und kommt damit für eine vorgegebene Zahl von Messwerten zu statistisch genaueren Aussagen als die einfache Allan-Deviation.

Das Diagramm zeigt den für viele Oszillatoren typischen – einer Banane nicht unähnlichen – Verlauf. Man muss dieses Diagramm etwa so lesen: Bis zu einer für diesen Oszillator charakteristischen Beobachtungszeit τ von hier etwa 140-150 s überwiegen in dem Oszillator hinsichtlich seiner Stabilität Zufallsprozesse, deren statistische Eigenschaften so beschaffen sind, dass sie sich mit zunehmender Beobachtungszeit immer stärker „herausmitteln“ und die Stabilität somit **besser** wird.

Oberhalb von 150 s steht die Stabilität des Oszillators offensichtlich unter dem Regime von Prozessen, die sich mit zunehmender Beobachtungszeit nicht herausmitteln sondern – im Gegenteil – mit zunehmender Beobachtungszeit immer „schlimmer“ werden. In erster Linie sind dies die Reaktionen auf äußere Reize wie Temperatur- und Druckschwankungen aber eben auch Alterungsprozesse. Welche Prozesse dabei genau welche Rolle spielen, ist heute zwar ganz gut bekannt aber immer noch aktuelles Thema ganzer Forschungszeige. Für uns ist das nicht weiter relevant und wir sollten uns lediglich folgendes merken:

⁴ Es handelt sich um den Taktoszillator eines WAVETEK 278 Funktionsgenerators

⁵ Diese sowie einige weitere Abbildungen sind mit dem Programm STABLE32 von „Hamilton Technical Services“ errechnet und als Graphik erzeugt. STABLE32 ist **der Standard schlechthin** in der Horologie und wo immer Sie im Internet etwas zum Thema „Stabilität von Oszillatoren“ finden, sind Ergebnisse, die mit STABLE32 gerechnet und präsentiert wurden, nicht fern. William „Bill“ Riley, der Boss von Hamilton Technical Services, ist Gott sei Dank unter dem Call K2HRT aktiver Funkamateure und hat, nachdem ich ihm davon berichtet habe, dass ich mich im Auftrag von Karl Meinzer mit Stabilitätsuntersuchungen beschäftige, kostenlos eine Vollversion dieser ansonsten nicht ganz billigen Software zur Verfügung gestellt. Dafür möchte ich an dieser Stelle noch einmal herzlich danken.

- Bis zu einer für ihn typischen Zeit überwiegen für die Stabilität eines Quarzoszillators Prozesse, die sich mit zunehmender Beobachtungszeit immer besser herausmitteln.
- Oberhalb dieser typischen Zeit überwiegen Prozesse, welche die Stabilität des Oszillators mit zunehmender Beobachtungszeit immer schlechter werden lassen.
- Eine bessere Stabilität als diejenige am Umkehrpunkt der Kurve kann man mit diesem Oszillator nicht erreichen. Würde der vermessene Oszillator in einem Frequenzzähler als Zeitbasis eingesetzt werden, so täte man gut daran, dessen Gate-Zeit auf 150 s einzustellen (soweit dies technisch möglich ist, 100 s sind eher wahrscheinlich). Mit dieser Gate-Zeit wird man die stabilsten Messungen machen können. Das Diagramm gibt sogar die Aussage her, dass man in einer solchen Messanordnung mit einem statistischen Fehler von Messung zur nächsten Messung von etwa 6×10^{-9} rechnen darf. Wenn hier von den **stabilsten** Messungen die Rede ist, so sind diese nicht unbedingt gleichzeitig auch **genau**. Ich möchte an dieser Stelle auch mit dem weit verbreiteten Märchen Schluss machen, dass Stabilität und Genauigkeit eines Oszillators irgendwie miteinander zu tun haben, obwohl man sich in der Praxis natürlich beides wünscht. Wenn ein Oszillator nicht auf seiner Sollfrequenz schwingt, so ist er **ungenau**. Trotzdem kann er natürlich **auf dieser falschen Frequenz beliebig stabil arbeiten**. Ich hoffe, es wird klar, dass man diese beiden Begriffe „Genauigkeit“ und „Stabilität“ absolut nicht so miteinander vermengen darf, wie dies in [1] leider mehrfach der Fall ist. Damit der Unterschied zwischen Stabilität und Genauigkeit noch einmal ganz deutlich wird, habe ich eine Graphik gezeichnet, die in Abbildung 2 gezeigt wird. Von einem Frequenznormal würden wir natürlich fordern, dass es **genau** und **stabil** ist, wie rechts im Bild dargestellt. Wenn ich das Normal als Zeitbasis in einem Zähler benutzen möchte, was nutzt mir ein zwar genaues aber instabiles Normal, wie in Fall 2 dargestellt. Zwar werden, wenn ich sehr viele Messungen durchführe, die Messungen um den richtigen Wert herum liegen, aber der einzelnen Messung kann ich genauso wenig trauen, wie bei einer **ungenauen** Zeitbasis. Für ein Frequenznormal müssen daher Genauigkeit und Stabilität in einer vernünftigen Relation zueinander stehen. Leider wird gerade dieser Punkt allgemein zu wenig beachtet, wie sich am Ende des Beitrages noch genauer zeigen wird
- Für eine Beurteilung des Oszillators sollte man auch im Auge haben, wo die Kurve „als ganzes“ liegt. Bei einer Beobachtungszeit von 1 s zeigt unser Prüfling eine Allan-Deviation von etwa $2,5 \times 10^{-8}$. Das bedeutet nicht mehr und nicht weniger, als dass dieser (nicht so ganz schlechte) Quarzoszillator um weitaus mehr als **vier Zehnerpotenzen** von dem Ziel entfernt ist, das Karl Meinzer für den Mars-Missions-USO formuliert hat. Ich denke, dass **dieser** Vergleich uns ein Gefühl dafür geben kann, dass der Begriff „Ultra Stable“ bei weitem nicht zu weit hergeholt ist!

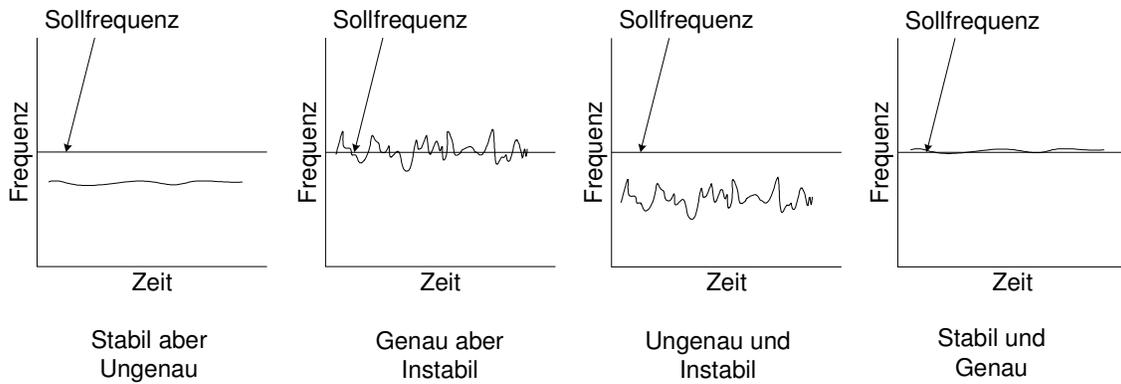


Abbildung 2: Genauigkeit und Stabilität sind nicht das gleiche!

Messmethoden für Stabilität

Die einfachste Methode der Stabilitätsmessung ist die so genannte „direkte“ Messmethode. Ihr Prinzip ist in Abbildung 3 gezeigt. Zusätzlich zu dem zu vermessenden Oszillator wird ein zweiter „Referenzoszillator“ benötigt. Die sinusförmigen Ausgangssignale beider Oszillatoren werden in Null-Durchgangs-Triggern zu digitalen Rechteckimpulsen umgewandelt und dem Start- bzw. Stoppeingang eines Zeitintervallmessgerätes (Time Interval Counter) zugeführt. In dieser Anordnung wird also gemessen, wie sich die Phasenlage der Nulldurchgänge von beiden Oszillatoren zeitlich zueinander entwickelt. Weil die Frequenz die zeitliche Ableitung der Phase ist, gibt es auch eine Formulierung für die Berechnung der Allan-Deviation, die nicht auf aufeinander folgenden Frequenzmessungen sondern auf aufeinander folgenden Phasenmessungen beruht.

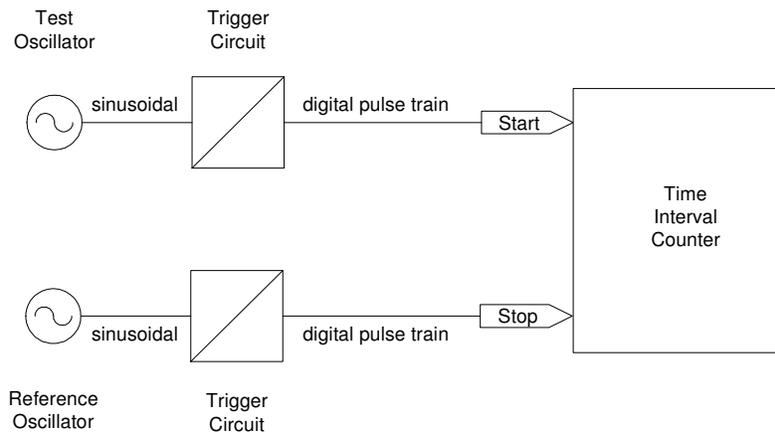


Abbildung 3: Prinzipbild der "direkten" Messmethode

Der kritische Leser wird einwenden wollen, dass die Qualität der Messergebnisse des Time Interval Counters doch maßgeblich auch von seiner **eigenen** Zeitbasis abhängig sind und somit ein dritter Oszillator ins Spiel kommt, der Einfluss auf die Messwerte hat. Dies ist im Prinzip richtig, die Situation kann jedoch dadurch bereinigt werden, dass der Referenzoszillator gleichzeitig als Zeitbasis für den Time Interval Counter dient.

Aber der Referenzoszillator: Der geht doch voll in die Messung ein, oder? Ja, richtig! Jedes Zeitintervall, welches wir mit dieser Anordnung zwischen den Nulldurchgängen von Prüfling und Referenz messen, ist immer gleichzeitig von **beiden Oszillatoren** abhängig, weil der eine das Start- und der andere das Stoppsignal liefert. Wenn Schwankungen beobachtet werden, so liefert die Anordnung keinen Hinweis darauf, ob die Schwankung am Anfang, am Ende oder an beiden Enden des Zeitintervalls verursacht wurde.

Ein mit dieser Anordnung bestimmter Wert der Allan-Deviation kann daher ohne weiteres nie eindeutig einem einzelnen der beiden Oszillatoren zugeordnet werden. Die Messergebnisse sind also immer als „Summe“ der Instabilitäten beider Oszillatoren zu verstehen („Summe“ in Anführungsstrichen, weil Instabilitäten sich nicht einfach algebraisch addieren. Die Sache ist letztlich ein wenig komplizierter, aber das braucht uns für unser Grundverständnis von der Sache nicht zu kümmern).

Auf diesen Tatbestand reagiert man in der Horologie auf zwei Weisen:

- Zum einen wählt man als Referenz die stabilste überhaupt verfügbare. Diese wird sich dann im Gegensatz zum Prüfling entsprechend wenig auf das Messergebnis auswirken. Für das Beispiel in Abbildung 1 wurde als Referenz ein „gut eingelaufener“ (=mehrere Monate ununterbrochener Betrieb) HP10811-Oszillator benutzt. Dieser wird von HP so spezifiziert, dass er bei $\tau = 1$ s eine Allan-Deviation von deutlich weniger als 10^{-11} hat. Wenn wir also bei $\tau = 1$ s eine Allan-Deviation von $2,5 \times 10^{-8}$ gemessen haben, so kann die Referenz dabei einen maximalen Beitrag geleistet haben, der weitaus kleiner als 1 Promille des Messergebnisses und damit vernachlässigbar ist. Dieser Messwert kann daher eindeutig dem Prüfling zugewiesen werden. Hätten wir hingegen z.B. $5,0 \times 10^{-11}$ gemessen, wäre diese Aussage so ohne weiteres nicht möglich. In dieser Situation setzt die Horologie auf die im nächsten Absatz beschriebene Methode.
- Hat man zwei Oszillatoren (einer davon die Referenz) vorliegen, die hinsichtlich ihrer Allan-Deviation dicht bei einander liegen, so kann man die Messergebnisse nicht eindeutig einem der Beiden zuordnen. Da jedoch **beide** zum Messergebnis beitragen, kann man die Aussage treffen, dass **jeder für sich alleine** besser als der gemessene Wert der Stabilität sein muss. Unter Umständen kommt man mit dieser Aussage ja schon weiter. Wenn man hingegen wirklich genau zuordnungsfähige Werte benötigt, so kann man folgendermaßen verfahren: Man besorgt sich einen dritten Oszillator mit möglichst gleichen Eigenschaften wie die beiden bereits vorhandenen. Benennt man diese mit O1, O2 und O3, so sind nun folgende Messreihen möglich O1 vs. O2, O1 vs. O3 und O2 vs. O3, Jeder gegen jeden, wenn man so will. Nehmen wir an, O1 ist der instabilste der drei Oszillatoren, dann wird sich das in der Messung O1 vs. O2 **und** in der Messung O1 vs. O3 niederschlagen. Mittels einer raffinierten „Three-Cornered-Hat“ (=Dreispitzhut) genannten statistischen Methode lässt sich aus den **drei** Messreihen die Allan-Deviation der **einzelnen** Oszillatoren rückrechnen.

Dies alles mag im ersten Moment sehr befremdlich klingen, ist aber letztendlich nichts anderes als die Manifestation der Tatsache, dass in der Physik ausnahmslos **jede** Messung ein Vergleichsvorgang mit einer Referenz ist, bei der die Referenz den gleichen Einfluss auf das Messergebnis hat wie der zu messende Wert selber.

Üblicherweise bleibt uns dies allerdings verborgen: Bei einer Spannungsmessung mit einem Multimeter denkt kaum einer daran, dass das Messgerät einen Vergleich zwischen der zu

messenden Spannung und der internen Referenz macht. Wir denken nicht daran, weil die Referenz dazu im Messgerät verbleibt und der Vergleich von der Elektronik darum herum geleistet wird. Außerdem sind die Referenzen meist sehr viel genauer als die zu messende Größe.

Die direkte Messmethode hat einen entscheidenden Nachteil: Der Phasenmessbereich der Anordnung ist wegen der Periodizität der Eingangssignale auf die Periodenlänge der Eingangssignale eingeschränkt. Laufen die Nulldurchgänge der Oszillatoren um mehr als eine Periodenlänge gegeneinander weg, so erhalten wir mehrdeutige Ergebnisse. Eine Anordnung, die diesen Fehler nicht hat, ist in Abbildung 4 gezeigt.

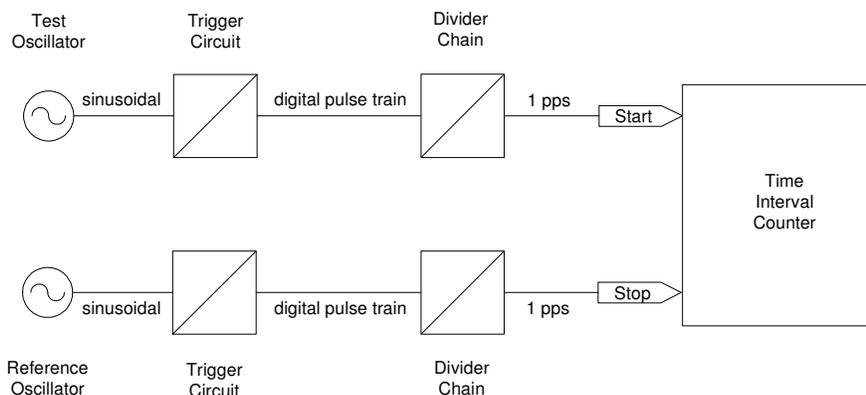


Abbildung 4: Prinzipbild der der "direkten" Messmethode mit erweitertem Messbereich

Hier wird das in eine Impulskette umgewandelte Oszillatorsignal nicht unmittelbar dem TIC (Time Interval Counter) zugeführt sondern durchläuft eine digitale Teilerkette, die am Ausgang einen einzelnen Impuls pro Sekunde liefert. In der Horologie wird ein solches Signal als „1 PPS“ also „1 Pulse Per Second“ bezeichnet. Dadurch reicht der Messbereich der Apparatur nun bis zu Phasenverschiebungen der Oszillatoren untereinander von 1 s, was angesichts der Tatsache, dass wir an der Untersuchung relativ stabiler Oszillatoren interessiert sind, einen sehr großen Messbereich darstellt. Der Vergleich von 1PPS-Signalen ist daher heute der Standard in der Stabilitäts-Messtechnik.

Bislang hat noch jeder, den ich erstmalig mit dieser Schaltung und dem Frequenzteiler konfrontiert habe, den Verdacht geäußert, dass der Frequenzteiler doch auch einen Einfluss auf die Phasenlage des herunter geteilten 1PPS-Signals haben müsste. Lassen sie uns deswegen dieser Frage nachgehen. In einen **synchronen** Frequenzteiler (den man dieser Stelle in der Tat benutzen muss!) werden **alle** Zähler-Flip-Flops vom gleichen Eingangs-Clock-Signal angesteuert. Das bedeutet, dass die Ausgänge **aller** Flip-Flops mit exakt **einem** typischen Propagation-Delay der benutzten Halbleitertechnologie nach einer Clock-Flanke den richtigen Zustand haben. Deswegen hat das 1PPS-Signal am Ausgang der Teilerkette auch genau **ein** typisches Propagation-Delay gegenüber dem Clock-Signal.

Das typische Propagation-Delay einer bestimmten Halbleitertechnik ist leider nicht absolut konstant sondern unter anderem von Temperatur und Betriebsspannung abhängig. Solange wir aber diese beiden Parameter halbwegs konstant halten, was nicht allzu schwer ist, können wir uns die synchrone Teilerkette als einen Auswahlmechanismus vorstellen, der unter allen Eingangs-Clock-Impulsen jeden z.B. 10-Millionsten aussortiert und mit einer fixen Verzögerung am Ausgang präsentiert. Mit anderen Worten: Die Teilerkette hat bis auf eine feste Verzögerung keinen Einfluss auf das Zeitverhalten des 1 PPS.

Haben wir mit dieser Anordnung nun ein ideales Messgerät für Stabilitätsmessungen zur Verfügung? Leider nein! Ein weiteres Problem kommt dadurch zustande, dass der TIC nur eine begrenzte Auflösung für die Zeitintervall-Messung hat. Eine Serie von unter Funkamateuren weit verbreiteten Universalzählern von RACAL-DANA (Model 1998, Model 1996), die seit einiger Zeit als Surplus preiswert erhältlich sind, hat z.B. eine Single-Shot Zeitauflösung von 1 ns.

Das bedeutet, dass wir für jedes Zeitintervall, welches wir mit diesem Zähler messen, nicht den „wahren“ Messwert erhalten, sondern einen auf die nächstliegende Nanosekunde gerundeten Wert. Durch das allgegenwärtige Quantisierungsrauschen werden wir daher, selbst wenn die zu messenden Zeitintervalle alle untereinander absolut identisch sind, Messergebnisse erhalten, die mal eine ns zu groß und mal eine ns zu klein sind.

Die Berechnung der Allan-Deviation „weiß“ natürlich nicht, dass dies von einer Eigenschaft der Messapparatur herrührt, sondern wird diese Sprünge als Instabilitäten der beteiligten Oszillatoren deuten. Das bedeutet, dass die Messeinrichtung dadurch, dass sie Zeitintervalle nur mit begrenzter Auflösung zu messen vermag, hinsichtlich der damit messbaren Allan-Deviation einen Noise-Floor bekommt, der unmittelbar von der Größe der Auflösung abhängig ist. Mit einer Nanosekunde Auflösung liegt der Noise-Floor bei 10^{-9} bei $\tau = 1$ s und fällt um eine Zehnerpotenz pro Verzehnfachung von τ . Im Sigma-Tau-Diagramm stellt er sich daher als Linie mit der Steigung -1 dar, die ihren Ursprung an der Stelle $X=1$ und $Y=10^{-9}$ hat. Allan-Deviation, die unter dieser Linie liegt, kann mit der Anordnung nicht gemessen werden.

Leider liegt dieser Noise-Floor damit um den Faktor 1000 höher, als das, was wir vom Mars-Missions-USO fordern. Den werden wir also in einer Apparatur mit **diesem** TIC nicht vermessen können. Eine Möglichkeit ist es, die Zeitauflösung zu verbessern. Um den Mars-Missions-USO zu vermessen, würde man eine Auflösung von deutlich unter 1 ps (10^{-12} s) benötigen.

Der beste derzeit von HP/Agilent verfügbare Zähler hat eine Zeitauflösung von 150 ps für Single-Shot Messungen. Den dürfen wir also auch abhaken. HP hat unter der Bezeichnung 5370A/B TICs gebaut, die über eine Auflösung von 20 ps verfügten und nicht mehr gebaut werden. Eine gigantische technologische Leistung, aber leider auch noch um mindestens einen Faktor 20 daneben. Der Vollständigkeit halber: Stanford Research baut auch heute noch unter der Bezeichnung SR620 einen TIC mit 20 ps Auflösung.

Es **gibt** heute Zeitmesstechnik im Sub-Picosekunden-Bereich, aber wir sollten als Funkamateure besser nicht die Hoffnung hegen, an Equipment dieser Art zu gelangen. Um noch mal ein Gefühl dafür zu bekommen, in welchen Dimensionen wir uns hier bewegen, möchte ich einen Zusammenhang zwischen diesen kurzen Zeiten und der höchsten Geschwindigkeit in unserem Universum, der Lichtgeschwindigkeit, herstellen. In einer Nanosekunde legt ein Lichtstrahl gerade mal noch 30 cm zurück, in einer Picosekunde demnach 0,3 mm (!), wir reden also schon über verflüchtete kurze Zeiten bzw. über extrem hohe Zeitauflösungen!

Wenn man – aus welchen Gründen auch immer – die Auflösung eines Messgerätes nicht ausreicht, um einen bestimmten physikalischen Effekt zu messen, so lautet die nächst beste Regel in der Physik: Dann **verstärke** den Effekt, bis er in eine Größenordnung gerät, wo er mit der begrenzten vorhandenen Auflösung gut gemessen werden kann. Wie könnte das aussehen? Werfen Sie dazu einen Blick auf Abbildung 5.

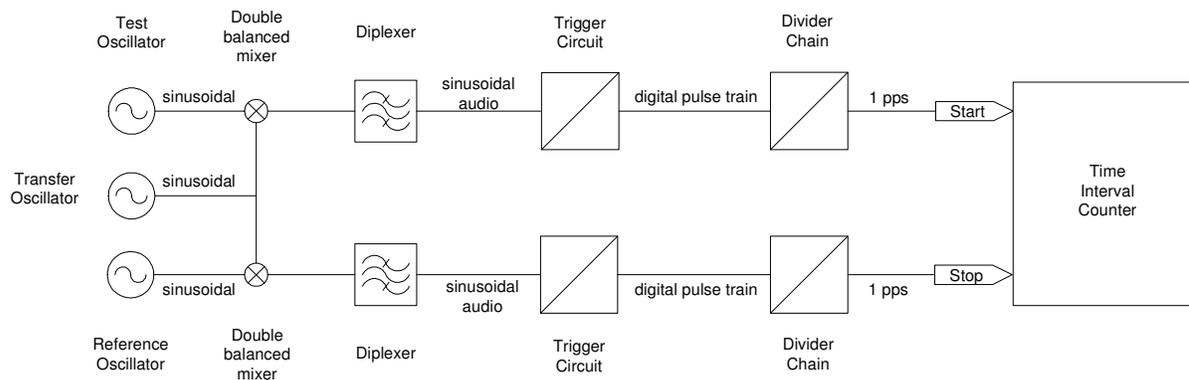


Abbildung 5: Prinzipbild der "Dual-Mixer"-Methode

In dieser Anordnung werden die beiden Oszillatorsignale in zwei Mixern mit der Frequenz eines dritten – Transfer Oszillator – genannten Oszillators, der z.B. 10 Hz Frequenzversatz gegenüber den beiden Oszillatoren hat, in den Audio-Bereich heruntergemischt. Ein nachfolgender Diplexer sorgt wie in einem Direct-Conversion-Empfänger dafür, dass das ZF-Tor des Mixers breitbandig mit 50 Ohm abgeschlossen ist. Am Ausgang des Mixers erscheinen im Idealfall nur die Summe und die Differenz aus HF- und LO-Signal. Wir verwenden hier den niederfrequenten Port des Diplexers und arbeiten mit dem Differenzsignal weiter. Der Rest der Anordnung ist wie bereits bekannt. Die Teilerkette teilt in diesem Fall natürlich nur noch durch den Faktor 10.

Was haben wir damit gewonnen? Wenn sich die Frequenz des Prüflings um 1 Hz verändert, so verändert sich auch die herunter gemischte Frequenz um 1 Hz. Dieses eine Hz macht auf der Originalfrequenz (hier mal 10 MHz vorausgesetzt) einen Effekt von 10^{-7} bezogen auf die Originalfrequenz aus. Bei 10 Hz Mischfrequenz macht dieses eine Hz aber einen 10 % (!) Effekt aus, der ungleich leichter zu detektieren ist als der 10^{-7} Effekt.

Noch frappierender ist die Wirkung der Schaltung, wenn man in der Phasendomäne denkt: Wenn sich auf der Originalfrequenz die Phasenlage zwischen den Oszillatoren um 1/100 der Periodenlänge verschiebt, so ist dies als Zeit ausgedrückt 1 ns und mit unserem Garden-Variety-Counter vielleicht gerade mal noch so feststellbar. Auch im herunter gemischten Signal verschiebt sich die Phase um 1/100 der Periodenlänge, aber um 1/100 der Periodenlänge des 10 Hz-Signals! Statt einer Verschiebung von 1 ns messen wir nun eine von 1 ms, eine um den Faktor 10^6 größere Zeit, die natürlich mit jedem Zähler einfach erfasst werden kann. Was ist mit den Eigenschaften des Transferoszillators? Nun, jeder Fehler des Transferoszillators gelangt ja in gleicher Phase in beide Messzweige. Deswegen ist die übliche Argumentation diejenige, dass bei der Bildung der Differenzzeiten zwischen den Messzweigen sich jeglicher Fehler des Transferoszillators herausheben muss.

Diese Schaltung werden Sie in dieser Form unter der Bezeichnung „Dual Mixer Method“ in unzähligen Veröffentlichungen im Internet finden und gilt als die Standardmethode der Stabilitätsmesstechnik.

Es gibt auch eine „Single-Mixer“-Methode, die aber der „Dual-Mixer“-Methode unterlegen ist. Bei der „Single-Mixer“-Methode wird der Prüfling direkt mit der Referenz gemischt. Für

die „Single-Mixer“-Methode muss daher der Referenz-Oszillator einen Offset gegenüber dem Prüfling haben.

Der größte Nachteil der „Single-Mixer“-Methode besteht darin, dass man nicht ihren „Über-Alles-Noise-Floor“ bestimmen kann. Wie wir bereits gelernt haben, werden wir nie in der Lage sein, Instabilitäten zu messen, die niedriger als diejenige des Referenz-Oszillators sind. Trotzdem können wir uns aber natürlich für die Frage interessieren, wie viel an am „Über-Alles-Noise-Floor“ vom Referenz-Oszillator beigetragen wird. Diese Frage wird sich z.B. dann stellen, wenn man wissen möchte, ob es sich lohnt, in eine bessere Referenz zu investieren. Dies lohnt sich natürlich nur, wenn der Rest des Messaufbaus mit der besseren Referenz mithalten kann. Aus der „Single-Mixer“-Methode können wir den Referenz-Oszillator nicht entfernen, weil es ohne ihn kein herunter gemischtes Signal gibt. Bei der „Dual-Mixer“-Methode können wir hingegen beiden Messzweigen das **gleiche** Oszillatorsignal zuführen! Da die Nutzsignale nun in beiden Messzweigen identisch sind, muss daher alles, was wir in dieser Situation als „Noise-Floor“ messen, zu 100% von der Messapparatur an sich verursacht werden. Diese Möglichkeit haben bei der „Single-Mixer“-Methode nicht.

Weniger ist bekannt, dass auch die „Dual-Mixer“-Methode bei weitem nicht so gut funktioniert, wie viele Autoren meinen. Dafür gibt es im Wesentlichen zwei Gründe:

Die Annahme, dass die Eigenschaften des Transferoszillators sich durch die Differenzmessung in den beiden Messzweigen aufheben, ist – zugegeben auf eine sehr subtile Art – leider unrichtig. Tatsächlich heben sich die vom Transferoszillator hervorgerufenen Fehler **wirklich** auf, wenn man das Geschehen in beiden Messzweigen **zum gleichen Zeitpunkt betrachtet**. Genau **dies** macht aber die „Dual-Mixer-Method“ **nicht!**

Wenn die beiden Oszillatorsignale auf 10 Hz herunter gemischt werden, so können die Nulldurchgänge der beiden 10 Hz Signale bis zu 50 ms auseinander liegen. Das bedeutet, ich sehe zum Zeitpunkt 1 (Nulldurchgang des Testoszillators) in den ersten Messzweig und irgendwann zwischen 0 und 50 ms später (Nulldurchgang der Referenz) in den zweiten Messzweig. In den vielleicht 50 ms hat der Transferoszillator 500000 komplette Schwingungen hinter sich gebracht und die Annahme, dass er sich hinsichtlich seiner Fehler in Frequenz und Phase **in genau dem gleichen Zustand** befinde wie 50 ms vorher, ist durch rein gar nichts gerechtfertigt.

Man mag zwar gewisse Hoffnungen darauf setzen, dass sich die Fehler des Transferoszillators im Laufe der Zeit nicht **zu schnell verändern**, so dass seine Fehler 50 ms später nicht komplett anders beschaffen sind sondern dem vorherigen Zustand zumindest so ähnlich, dass es ein **gewisses Maß** an Error-Cancellation zwischen den beiden Messzweigen tatsächlich gibt, aber exakt ist es eben nicht der Fall. Diese Tatsache wurde in der wissenschaftlichen Literatur erstmals in [2] in voller Konsequenz beschrieben und ebenfalls, wie man messtechnisch mit dieser Tatsache umgehen muss. Dies führt zu einem wesentlich komplexeren Messaufbau, der dann statt des TIC einen so genannten „Multi-Channel-Event-Counter“ benötigt.

Die „Dual-Mixer-Methode“ leidet jedoch auch an einem zweiten wenig bekanntem Phänomen, welches vom Autor von [3] folgendermaßen formuliert wird: „When it comes to timing measurements the Slope-to-noise-ratio is the figure of merit and **not** the Signal-to-noise-ratio“. Bei Zeitmessungen ist also die „Slope-to-noise-ratio“ die entscheidende Zahl und nicht die „Signal-to-noise-ratio“. Was hat es damit auf sich?

Wann immer wir an einem sinusförmigen Signal eine Zeitmessung vornehmen wollen, so müssen wir uns für Punkte auf der Kurvenform entscheiden, zwischen denen wir die Zeit messen wollen. Die Intuition sagt uns dabei, dass die Nulldurchgänge dafür am geeigneten sind, weil da das Signal dort die größte Steilheit hat und wir werden gleich sehen, **wie** wichtig dies tatsächlich ist. In den bisher gezeigten Prinzip-Bildern, wird die Aufgabe, den Nulldurchgang zu detektieren von den mit „Trigger-Circuit“ benannten Funktionsblöcken durchgeführt. Gerade weil ihnen eine große Bedeutung zukommt, sind sie als eigenständige Funktionsblöcke gezeichnet.

Überlegen wir zunächst, was ein rauschfreies sinusförmiges Signal am Nulldurchgang macht: Es schneidet die horizontale Achse (=den Trigger-Pegel des Nulldurchgangsdetektors) an **genau einer Stelle** und liefert daher ein „scharf definiertes“ Zeitsignal für den Nulldurchgang. Rauschfreie Signale sind leider eine Idealisierung, die es in der Praxis nicht gibt. In der Praxis haben wir immer Rauschen vorliegen, das sich in einem mehr oder weniger großen $(S+N)/N$ des Signals dokumentiert.

In Abbildung 6 habe ich einmal den Nulldurchgang eines stark verrauschten sinusförmigen Signals dargestellt. Bitte beachten Sie, dass ich dieses schlechte $(S+N)/N$ bewusst deswegen gewählt habe, damit Sie den Effekt graphisch erkennen können. Er tritt aber in der beschriebenen Form bei jedem $(S+N)/N$ auf, auch bei sehr gutem.

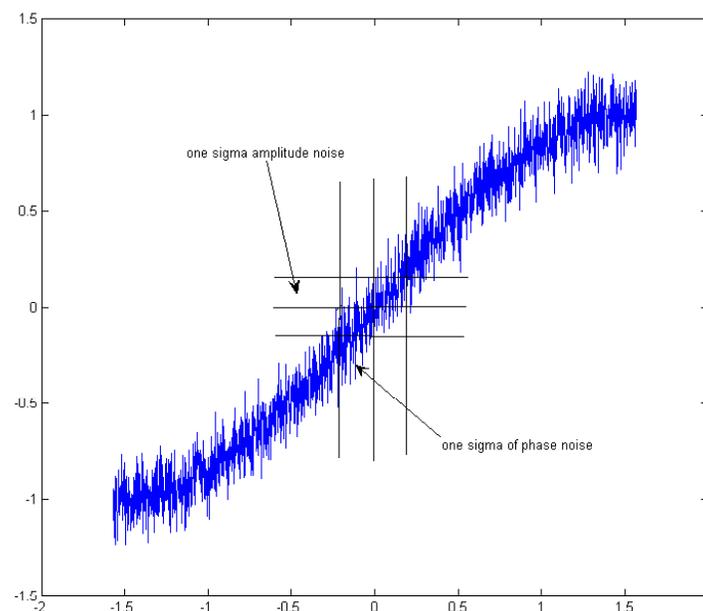


Abbildung 6: Übersetzen von Amplituden- in Phasenrauschen beim Nulldurchgang

Wie man sieht, gibt es nun wegen des Rauschens nicht mehr einen scharf definierten Zeitpunkt, an dem das Signal die horizontale Achse schneidet, sondern es entsteht ein ganzes „Zeitfenster“, in dem das Signal die Null-Linie schneiden könnte, unter Umständen sogar mehrfach in beide Richtungen.

Was wir hier beobachten ist ein Übersetzungsprozess, bei dem Amplitudenrauschen in Phasenrauschen bezüglich der für den Nulldurchgang gemessenen Zeiten übersetzt wird. Wie

man sich leicht überlegen kann, hängt der „Übersetzungsfaktor“ von Amplitudenrauschen in Phasenrauschen direkt von der Steilheit des Signals im Nulldurchgang ab.

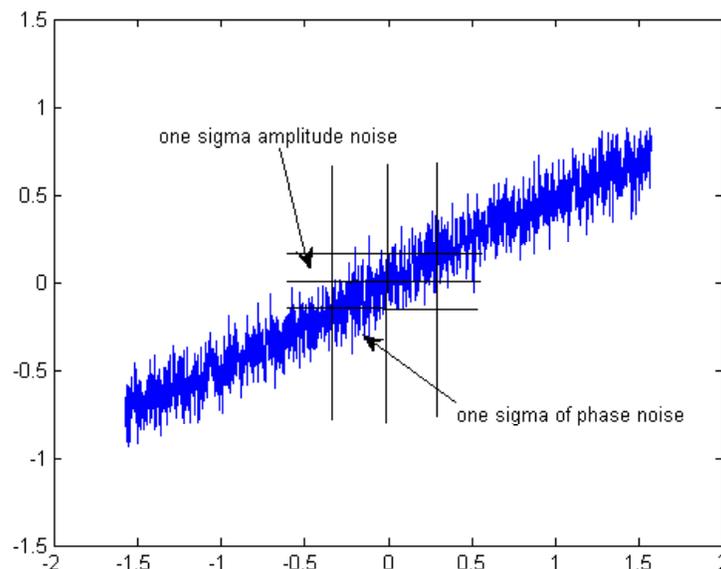


Abbildung 7: Geringere Steilheit am Nulldurchgang vergrößert das Phasenrauschen

In Abbildung 7 ist die Signalamplitude und das $(S+N)/N$ des Signals gleich den Bedingungen in Abbildung 6, lediglich die Frequenz des Signals wurde halbiert. Weil die Steilheit des Signals **nur halb so groß** ist, ist der mögliche Bereich für Schnittpunkte mit der horizontalen Achse **doppelt so groß** wie in Abbildung 6.

Das $(S+N)/N$ ist in beiden Fällen das gleiche und wir sehen in der Tat, dass die Slope-to-noise-ratio, also das Verhältnis von Signal-Steilheit und Rauschen wirklich das alles Entscheidende ist. Überschlüssig habe ich einmal ausgerechnet, dass auf 10 MHz ein $(S+N)/N$ von 20 dB ausreicht, um an einem sinusförmigen Signal eine Zeitmessung mit einer Unsicherheit von 1 ns durchzuführen. Bei 10 Hz mussten es hingegen 140 dB $(S+N)/N$ sein!

Die beschriebenen Effekte sind nicht akademisch-theoretischer Natur sondern können von jedem, der halbwegs geeignetes Mess-Equipment hat, leicht nachvollzogen werden. Man nehme einen Zähler mit ausreichend großer Messauflösung, z.B. den angesprochenen RACAL-DANA und synchronisiere ihn mit der besten Referenz, die man hat. Man nehme außerdem den besten und rauschärmsten Funktionsgenerator, den man bekommen kann und synchronisiere ihn ebenfalls mit der Referenz. Nun stelle man den Generator auf eine Frequenz von 10 Hz und mache eine Messreihe von Frequenz- oder Periodenlängen-Messungen. Die werte man anschließend statistisch aus. Ich verspreche: Das Staunen wird groß sein!

Und, um es noch einmal zu betonen: Womit wir es hier zu tun haben, ist **nicht** in irgendeiner Form **ein technischer Fehler**, der mit besseren Messgeräten nicht auftreten würde. Es handelt vielmehr um eine **prinzipielle Eigenschaft** des Signals selber, die sich auch mit beliebig gutem Equipment Wirkung zeigen würde.

Wer unsere ursprüngliche Intention, nämlich mit der „Dual-Mixer-Method“ kleine Allan-Deviations zu messen, weiter im Sinn hat, der wird möglicherweise schon sehen, welche fatale Auswirkung die oben beschriebenen Zusammenhänge auf unser Vorhaben haben:

Zwar haben wir mit dem Trick des Heruntermischens den zu messenden Effekt um den Faktor 10^6 verstärkt! Genau in dem gleichen Maße von 10^6 haben wir aber beim Heruntermischen von 10 MHz auf 10 Hz die Slope-to-noise-ratio **verschlechtert**! Mögen die Oszillatorsignale ein noch so großes $(S+N)/N$ haben und möge der Mischvorgang ein rauschfreier Prozess sein (eine sehr optimistische Annahme!): Die Slope-to-noise-ratio der herunter gemischten Signale wird um den Faktor 10^6 schlechter sein als die der Originalsignale, daran führt kein Weg vorbei und deswegen wird die „Effekt-Verstärkung“, die wir uns von dem Heruntermischen erhofft haben, **viel kleiner als erwartet** ausfallen.

Wenn dem so ist, warum macht man es dann überhaupt? Nun, es ist ja nicht so, als dass wir gar nichts gewonnen hätten: Allein die Tatsache, dass die beiden Oszillatorsignale nun in den niederfrequenten Audio-Bereich umgesetzt vorliegen, ist Ansatzpunkt für eine ganze Reihe von Maßnahmen, die darauf abzielen, das $(S+N)/N$ im Audio-Signal und damit letztlich auch die Slope-to-noise-ratio zu verbessern.

Die dafür notwendigen Filtermaßnahmen sind natürlich im Audio-Bereich vergleichsweise einfach zu bewerkstelligen und Sie dürfen mein Wort dafür nehmen, dass die „Schwarze Kunst“ bei der „Dual-Mixer-Method“ ausschließlich in der Signalkonditionierung des Audio-Signals besteht, bevor es auf den Nulldurchgangsdetektor gelangt. Wer hier ein einfaches Tiefpassfilter in den Signalweg einzeichnet, der will Sie entweder auf den Arm nehmen oder seine Firmengeheimnisse besser für sich behalten. Jedes dB im $(S+N)/N$, das man hier durch geeignete Konditionierungsmaßnahmen gewinnt, wird sich in einer verbesserten Zeitauflösung der Messanordnung und damit in einem niedrigeren Noise-Floor der messbaren Allan-Deviation niederschlagen.

Nachdem ich ungefähr diesen Kenntnisstand über die Vor- und Nachteile der „Dual-Mixer-Method“ erreicht hatte, stieß ich bei einer Recherche im Internet auf den hochinteressanten Artikel [5]. In diesem Beitrag⁶ schildert der bereits durch [2] bekannte Autor Greenhall, wie die NASA im Rahmen des Betriebes ihres DSN (Deep Space Network) auch Messanordnungen betreibt, in denen die herunter gemischten Signale nicht unmittelbar vermessen sondern mit einer gewissen Abtastrate abgetastet (gesampled) und für die spätere Weiterverarbeitung abgespeichert werden. Die komplette Mathematik, die für die nachträgliche Auswertung der gespeicherten Werte notwendig ist, das sind im Wesentlichen Algorithmen aus der digitalen Signalverarbeitung, ist freundlicherweise im Anhang des Artikels zu finden.

Die faszinierende Idee in diesem Artikel ist die folgende: Wenn ich die Nulldurchgänge zweier niederfrequenter Wellenzüge miteinander vergleiche, so findet da die bereits beschriebene Übersetzung von Amplituden- in Phasenrauschen mit allen ihren negativen Konsequenzen statt. Bei einem zeitdiskret abgetasteten Signal hat man aber z.B. pro Periode 4800 Samples zur Verfügung (10 Hz Signal mit 48000 Abtastungen / s). Jeder einzelne dieser Abtastwerte ist natürlich von Rauschen überlagert. Statt nun einfach die beiden Abtastungen zu suchen, zwischen denen der Nulldurchgang liegt, hat man stattdessen die Chance, mittels einer „nicht-linearen Regression“ genannten mathematischen Werkzeugs die abgetastete Kurvenform so zu rekonstruieren, dass sich das in den einzelnen Abtastwerten durchaus enthaltene Rauschen möglichst wenig auf das rekonstruierte Signal auswirkt. Mit anderen Worten: Die schiere Zahl der zur Verfügung stehenden Abtastwerte, gepaart mit ein wenig Mathematik, sollten zu einer eindrucksvollen Verbesserung des $(S+N)/N$ des **rekonstruierten** Signals führen.

⁶ Sie finden die im Literaturverzeichnis aufgeführten Dinge fast alle im Internet, wenn Sie mit Google nach dem Titel suchen.

In der digitalen Signalverarbeitung werden alle Frequenz- und Phasenwerte nicht als absolute sondern als auf die Abtastrate normierte Werte angegeben. Auch als Nicht-Fachmann für DSP kann man daher leicht nachvollziehen, dass, wenn wir Frequenz- und Phasenvergleiche mit Mitteln der DSP machen wollen, der Stabilität und Genauigkeit der Abtastate eine große Bedeutung zukommt. Der NASA steht an dieser Stelle eine professionelle 16-Bit-A/D-Wandlerkarte zur Verfügung, deren Abtastate sich an eine externe Referenz ankoppeln lässt. Die externe Referenz ist bei der NASA wie selbstverständlich ein Cäsium-Frequenznormal.

Dieser Artikel brachte bei mir die Idee hervor, das gleiche was Greenhall mit professionellem Equipment macht, mal mit einer Soundkarte zu probieren.

Zunächst einmal: Der Artikel war immerhin schon mehr als 10 Jahre alt und ich war einfach daran interessiert, ob das, was Greenhall als „zeitaufwendige und nur im Nachhinein am abgespeicherten Signal machbare Berechnung“ bezeichnet, nicht heute in „Echtzeit“ möglich wäre. Hier will ich das Ergebnis schon einmal vorausschicken: Selbst ein antiquierter Pentium II mit 400 MHz Taktfrequenz schafft dies!

Dann: 16 Bit Auflösung ist eigentlich gar nicht so viel und viele Soundkarten arbeiten heute bereits mit 24 Bit Auflösung.

Aber: Die Soundkarte leitet ihre Samplerate von einem „Canned“-Oszillator billigster Bauweise ab, der absolut keinen Vergleich mit einem Cäsium-Frequenznormal zulässt. Das konnte doch gar nicht funktionieren.

Andererseits: Viele Soundkarten benutzen Stereo-A/D-Wandler, die wirklich beide Stereokanäle absolut gleichzeitig abtasten und bedeutet nicht dies genau, dass sich alle Fehler in der Abtastate wieder aufheben müssen, wenn man wieder zu differenziellen Messungen zwischen den beiden Stereo-Kanälen kommt?

Es folgten Monate der Programmiererei und von Versuch und Irrtum. Wie sich im Laufe der Entwicklung herausgestellt hat, sind Fehler in der Abtastate der Soundkarte gar nicht **der** entscheidende Punkt, der Beachtung finden muss, sondern – vollkommen unerwartet von mir – die Übersprechdämpfung zwischen den beiden Stereo-Kanälen. Bei Soundkarten für den Hobby- und Spiel-Bereich findet man Übersprechdämpfungen (Cross-Talk) von 60 – 80 dB zwischen linkem und rechtem Stereo-Kanal. Das ist für alle „normalen“ Anwendungen jenseits von Gut und Böse. Für die Dinge, die ich mit der Soundkarte machen wollte, nämlich den Phasenvergleich an den herunter gemischten Oszillatorsignalen damit durchzuführen, hat sich dieses Übersprechen, auch wenn es um 60 – 80 dB abgedämpft ist, als eine Störquelle allergrößten Ausmaßes herausgestellt.

Das Übersprechen führt dazu, dass in jedem der beiden Kanäle das Nutzsignal von einer um 60 – 80 dB abgedämpften Kopie des Nutzsignals vom jeweils anderen Kanal überlagert ist. Überlagert bedeutet in diesem Fall: Die Amplituden von Nutz- und Störsignal überlagern sich. Da Nutz- und Störsignal nie wirklich exakt die gleiche Frequenz haben, führt diese Überlagerung zu einer Störmodulation, deren Frequenz die Beat-Frequenz der beiden Oszillatoren ist und die sowohl zu einer Amplituden- als auch zu einer Phasenmodulation des Nutzsignals führt. Auch wenn diese um 80 dB gegenüber dem Nutzsignal abgeschwächt ist: Das Sigma-Tau-Diagramm ist ein dermaßen empfindlicher Indikator für Dinge, die periodisch im Oszillator passieren, dass die so erhaltenen Sigma-Tau-Diagramme bei Beobachtungszeiten, die der Periodenlänge (oder Vielfachen davon) der Beatfrequenz entsprechen, extrem verzerrt werden.

Die Übersprechdämpfung von 60 – 80 dB entsteht auf der Soundkarte selber und kann durch äußere Beschaltung nur verschlimmert, aber nicht verbessert werden. Versuche, das Übersprechen mit DSP-Techniken nachträglich künstlich zu verbessern, waren anfangs viel versprechend, aber im Nachhinein doch nicht erfolgreich.

Mittlerweile sind jedoch Soundkarten verfügbar, die auf das Marktsegment des semiprofessionellen Musikers bzw. des ambitionierten Musikliebhabers abzielen. Diese warten nicht nur mit den in der Studioteknik üblichen symmetrischen Ein- und Ausgängen auf sondern können auch mit Übersprechdämpfungen bis zu 130 dB (!) glänzen. Dies ist keine Prospektangabe, sondern ein von mir nachgemessener Wert. Dass sich diese Soundkarten üblicherweise entweder über ein so genanntes „Word.Clock“-Signal oder über ein S/PDIF-Signal⁷ mit anderen Soundkarten synchronisieren lassen (und damit auch mit einer externen Referenz!), kommt als i-Tüpfelchen hinzu.

Ich betreibe in meinem Messaufbau mittlerweile eine „M-AUDIO Audiophile 2496“ Soundkarte, welche in Zusammenhang mit dem derzeit existierenden Softwarestand einen Noise-Floor von etwa 10^{-13} bei $\tau = 1$ s hat. Vorausgesetzt, man hat eine ausreichend stabile Referenz zur Verfügung, wäre mit diesem Messaufbau schon ein Vermessen des Mars-Missions-USO möglich. Schon der Vergleich von zweien dieser Oszillatoren würde Aufschluss darüber geben, ob man in der richtigen Größenordnung der geforderten Allan-Deviation liegt.

Vermutlich wird sich der Noise-Floor mit einigen Hardware-Modifikationen im Bereich der Mischer noch einmal deutlich verbessern lassen. Die Schaltungstechnik der Messanordnung soll Gegenstand eines eigenständigen Beitrages sein.

Die Abtastrate der Soundkarte ist über eine von mir entwickelte Schaltung phasenstarr an eine meiner 10 MHz Referenzen angehängt. Auch diese Schaltung soll an geeigneter Stelle veröffentlicht werden. Abbildung 8 zeigt die Programmoberfläche während einer Noise-Floor-Messung. Ich habe das Programm „DSP Frequency Stability Tester“ genannt.

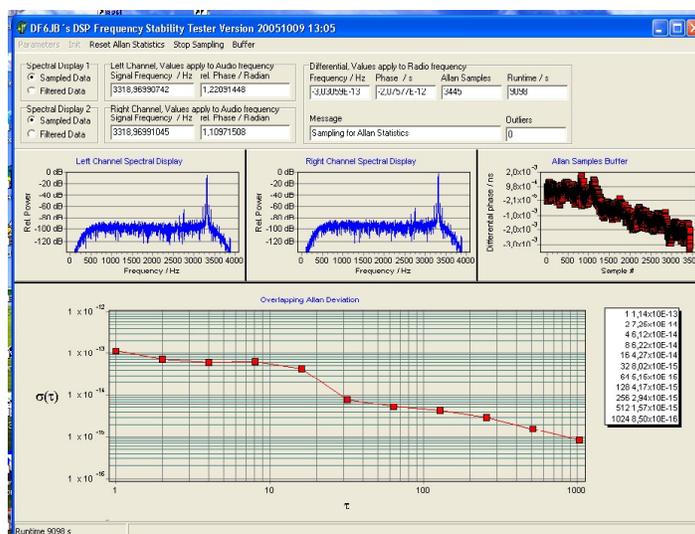


Abbildung 8: Bedienoberfläche des "DSP Frequency Stability Tester"

⁷ Sony / Philips Digital Interface

Die Stabilität und Genauigkeit von Frequenznormalen

Bislang war die Rede von Oszillatoren und ihrer Stabilität. In einem Frequenznormal für Funkamateure befindet sich ein lokaler Oszillator, der von einer externen Referenz nachgeregelt wird. Die Menschen, die Frequenznormale entwickeln, sprechen davon dass in ihren Frequenznormalen die hohe Kurzzeitstabilität des lokalen Oszillators aufs trefflichste mit der Langzeitstabilität der externen Referenz „verheiratet“ werden soll. Zu diesen Begriffen habe ich ja schon etwas gesagt. Gemeint ist, dass für kleine Beobachtungszeiträume der lokale Oszillator die Quelle von Stabilität sein soll, wohingegen die die externe Referenz für Stabilität bei langen Beobachtungszeiträumen **und** für Genauigkeit sorgen soll.

Als externe Referenzen werden entweder der Normalfrequenz- und Zeitzeichensender DCF77 in Mainflingen bei Frankfurt/Main oder aber neuerdings das weltweite GPS-Navigations-System benutzt. Andere externe Quellen sind denkbar, etwa das russische GLONASS System, eine Anbindung an die amerikanischen GOES Satelliten, aber auch terrestrische Navigationssysteme wie Loran-C oder Omega. In der deutschen Amateurfunkliteratur haben diese „Exoten“ allerdings bislang keine Rolle gespielt.

Lassen Sie uns zunächst einmal das Sigma-Tau-Diagramm eines für ein Frequenznormal geeigneten lokalen Oszillators betrachten. Dies ist in Abbildung 9 gezeigt. Es handelt sich um den mittlerweile auch in Amateurkreisen recht verbreiteten HP10811. Den hatten wir bereits dem „gewöhnlichen“ Quarzoszillator als Referenz beige stellt bei der Messung, die zu den Ergebnissen in Abbildung 1 führte. In diesem Fall war er um Zehnerpotenzen stabiler als der Prüfling. Wenn wir nun den HP10811 selber vermessen wollen, so sollten wir das im Prinzip auch mit einer zigfach stabileren Referenz tun.

Leider wird die Luft da sehr dünn! Es gibt BVA-Quarzoszillatoren, die in die Größenordnung von 10^{-13} bei $\tau = 1$ s kommen und es gibt tiefgekühlte „Whispering-Gallery-Saphire“-Oszillatoren, die nochmals besser sind, aber diese Dinge sind entweder noch auf den Tischen der Forschungslabors oder schlicht unbezahlbar. Das Beste, was mir als Referenz zur Verfügung steht, ist ein FTS1200 von „Frequency Time Systems“ (heute Symmetricon).

Bei dem Ofen dieses Oszillators wurden **alle** Register gezogen und eine große thermische Masse in einem aufwendigen Deware-Gefäß (ähnlich einer Thermos-Kanne) verpackt. Dieser Oszillator hat einmal in einem Cäsium-Frequenznormal als Flywheel-VCXO⁸ seinen Dienst getan, dessen Frequenz über Vervielfacherstufen direkt an den atomaren Resonanzprozess angekoppelt war. FTS gibt eine Allan-Deviation für $\tau = 1, 10, 100$ s von 10^{-12} an, aber es gibt Grund zu der Annahme, dass dies als konservativer Wert gelten darf.

Wenn man den HP10811 mit einer Referenz dieser Klasse vergleicht, so ergibt sich ein Bild wie in Abbildung 9 gezeigt. Dieses Diagramm ist in mehrfacher Hinsicht bemerkenswert. HP gibt für den 10811 bei $\tau = 1$ s eine Allan-Deviation von $5,0 \times 10^{-12}$ an und die Tatsache, dass wir mit unserer Messung tatsächlich ganz in der Nähe dieses Wertes liegen, ist eine unglaublich „vertrauensbildende“ Information, weil sie zeigt, dass unsere Messapparatur nicht ganz falsch sein kann, aber auch, dass die Allan-Deviation des FTS1200 sich wohl im Rahmen ihrer Spezifikation bewegt.

Die Kurve hat den schon bekannten Bananen-ähnlichen Verlauf und der Umkehrpunkt der Kurve liegt sogar bei einer kürzeren Zeit als bei dem gewöhnlichen Quarzoszillator, aber man

⁸ Dieser Begriff wird weiter unten erklärt

bedenke, dass die Kurve eben insgesamt um fast drei Zehnerpotenzen niedriger als in Abbildung 1 liegt. Insbesondere der sehr flache Anstieg der Kurve nach dem Umkehrpunkt zeugt von der hohen Qualität beider Oszillatoren.

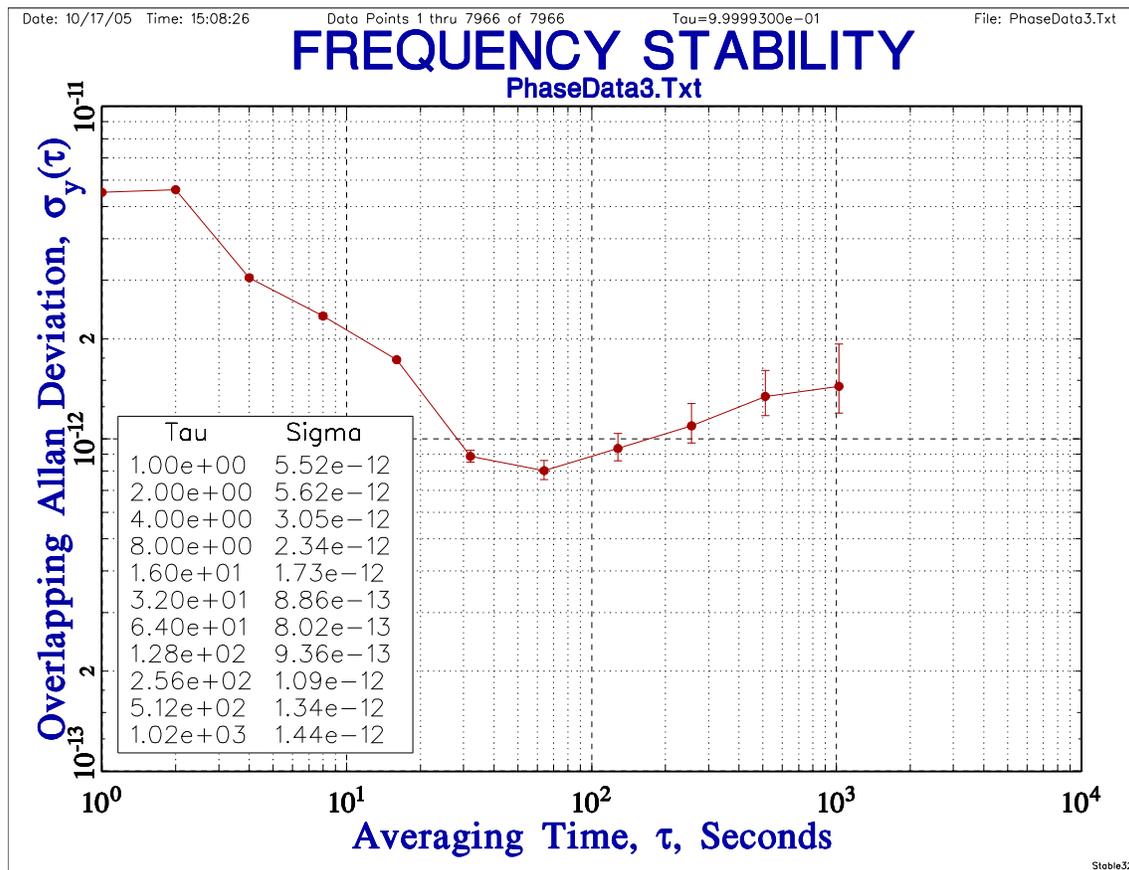


Abbildung 9: Sigma-Tau-Diagramm eines HP10811-Oszillators

Wenn man [5] Glauben schenken darf, so sollte der HP10811 eigentlich bei $\tau = 1$ s bis $\tau = 10$ s das Minimum seiner Sigma-Tau-Kurve haben und bei $\tau = 100$ s schon wieder auf einen Wert von 10^{-11} geklettert sein. Wenn wir hier ein wesentlich günstigeres Verhalten messen, dann führe ich das auf zwei Dinge zurück. Zum einen scheint es so zu sein, dass die Angaben von HP sich auf den Fall beziehen, dass der Oszillator gerade mal 24 h Stunden Zeit zum „Aufwärmen“ hatte, wohingegen er bei mir ununterbrochen seit Monaten in Betrieb ist.

Zum anderen betreibe ich den HP10811 aber auch nicht „einfach so“, sondern habe ihn in ein massives Aluminium-Übergehäuse mit 20 mm Wandstärke verpackt, dessen Wirkung man sich in etwa als „thermischen RC-Tiefpass“ vorstellen muss. Das Äquivalent zur Kapazität des Kondensators im RC-Glied bildet hier die Wärmekapazität des Alu-Blocks. Das Äquivalent zum Widerstand im RC-Glied bildet hier der Wärmewiderstand, der die thermische Ankopplung des Blocks an die umgebende Luft beschreibt. Eine überschlägige Rechnung zeigt, dass die Zeitkonstante dieses „thermischen RC-Glieds“ in der Größenordnung von einigen Tausend Sekunden liegen sollte, wobei allerdings der Wärmeübergang zur Umgebung ein großer Unsicherheitsfaktor ist, weil er stark unterschiedlich sein kann zwischen „stehender“ und turbulenter Umgebungsluft. Dieses Übergehäuse kann den Oszillator zwar nicht voll-

kommen von seiner Umgebung abkoppeln, bietet ihm aber eine thermisch „beruhigte“ Arbeitsumgebung.

Kommen wir nun zum Sigma-Tau-Diagramm eines für ein Frequenznormal besonders geeigneten GPS-Empfängers. Es handelt sich um den (leider abgekündigten) Motorola M12+ Empfänger. Dieser Empfänger ist von der gleichen Qualität wie der in [6] benutzte und etwa um einen Faktor 5 besser als der in [1] benutzte. Ich werde gleich noch darauf eingehen, worin genau dieses „besser“ besteht. Abbildung 10 zeigt uns das Sigma-Tau-Diagramm dieses Empfängers.

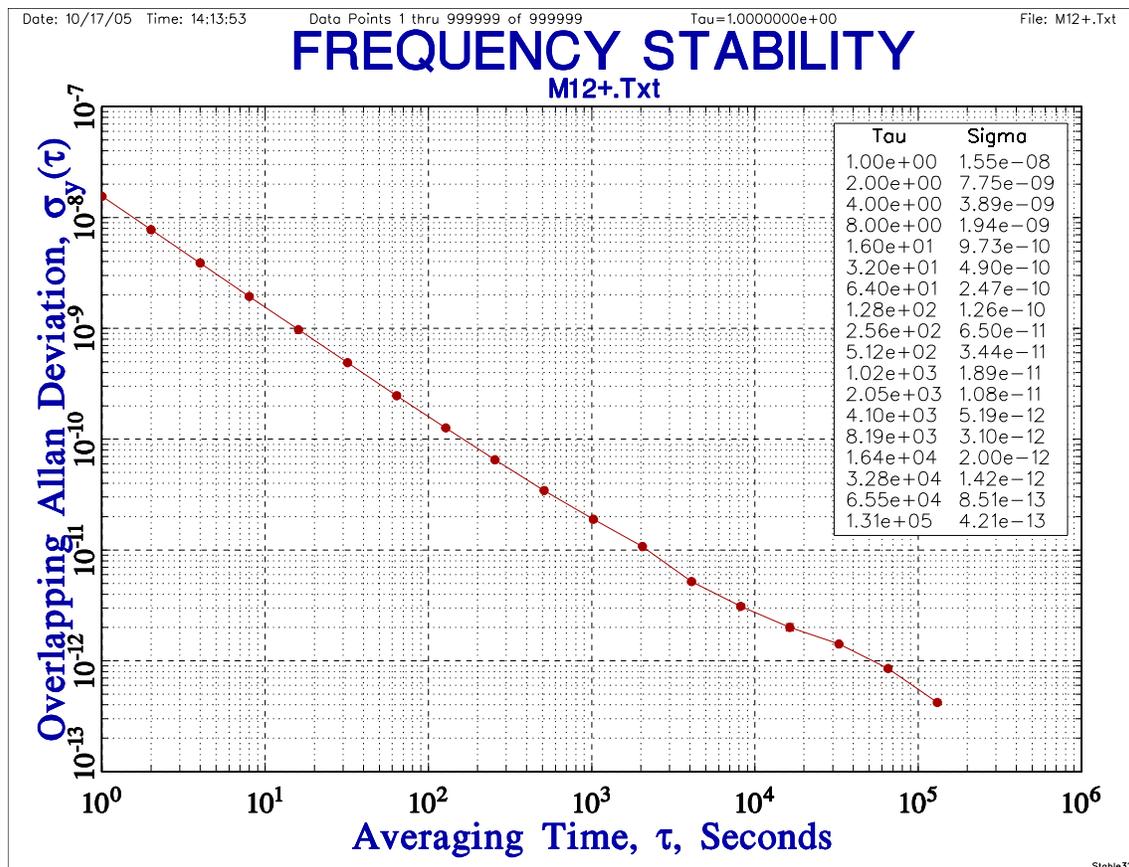


Abbildung 10: Sigma-Tau-Diagramm eines MOTOROLA M12+ GPS-Empfängers

Wie kommt dieses Diagramm zustande? Der 1 PPS eines jeden GPS-Empfängers weist bedingt durch seine Konstruktion (und nicht etwa bedingt durch Ausbreitungs-Effekte) einen Phasenjitter auf, der z.B. bei diesem, aber auch bei dem in [6] benutzten Empfänger in der Größenordnung von 15 - 20 ns liegt. Dies führt zu der Allan-Deviation von etwa $1,5 \times 10^{-8}$ bei $\tau = 1$ s. Die mit der Steigung -1 fallende Gerade zeigt an, dass dieser Jitter den Charakter weißen Rauschens hat und sich mit zunehmender Beobachtungszeit immer besser herausmittelt.

Die Kurve hat auch zu langen Beobachtungszeiten nicht den Umkehrpunkt, wie wir ihn von unseren Quarz-Oszillatoren kennen. Das liegt daran, dass diese Kurve **keine** Aussage über die dem GPS-Signal zugrunde liegende Referenz an sich macht, **sondern nur** darüber, mit welcher Stabilität der GPS-Empfänger die Referenz bei uns im Labor **reproduzieren** kann! Nach

dem Sigma-Tau-Diagramm der dem GPS-System zugrunde liegenden Referenz zu fragen, gibt hingegen deswegen keinen Sinn, weil die dahinter stehende Referenz einfach die „amtliche“ gesetzlich verbindliche ist. (zumindest was die USA angeht), weil die im GPS-System benutzten Uhren mit den nationalen Referenzen der USA synchronisiert werden.

Hingegen sieht man, wenn man genau hinguckt, zwischen $\tau = 10000$ s und $\tau = 100000$ s eine ganz leichte „Ausbuchtung“ der Kurve nach oben. Die Ursachen für diese Abweichungen von der idealen Gerade werden als so genannte „diurnale“ Effekte der Ionosphäre bezeichnet und kommen dadurch zustande, dass durch die unterschiedliche Elektronendichte in der Ionosphäre zwischen Tag und Nacht eine elektromagnetische Welle, welche die Ionosphäre durchläuft, unterschiedliche Ausbreitungsbedingungen vorfindet.

Was die Kurve uns sagt, ist, wie lange man über das GPS-Signal mitteln muss, um zu einer bestimmten Reproduzierbarkeit der entfernten Referenz zu gelangen. Dies ist übrigens vollkommen unabhängig davon, ob man den 1 PPS des Empfängers benutzt oder ein 10 KHz-Signal wie in [6], weil beide Signale entsprechend verjittert sind.

Wenn man die Sigma-Tau-Kurven vom HP10811 und dem M12+ in das gleiche Diagramm einzeichnet, so fällt auf, dass die beiden Kurven sich bei einer gewissen Beobachtungszeit τ schneiden würden, wenn die Messreihe für den HP10811 hinreichend lang gewesen wäre. Ich habe die Kurve „von Hand“ extrapoliert und es kommt für das Verständnis von dem, was jetzt folgt, nicht darauf an, ob der Schnittpunkt nun ein wenig höher oder tiefer liegt, als ich es gezeichnet habe. Dies ist in Abbildung 11 gezeigt.

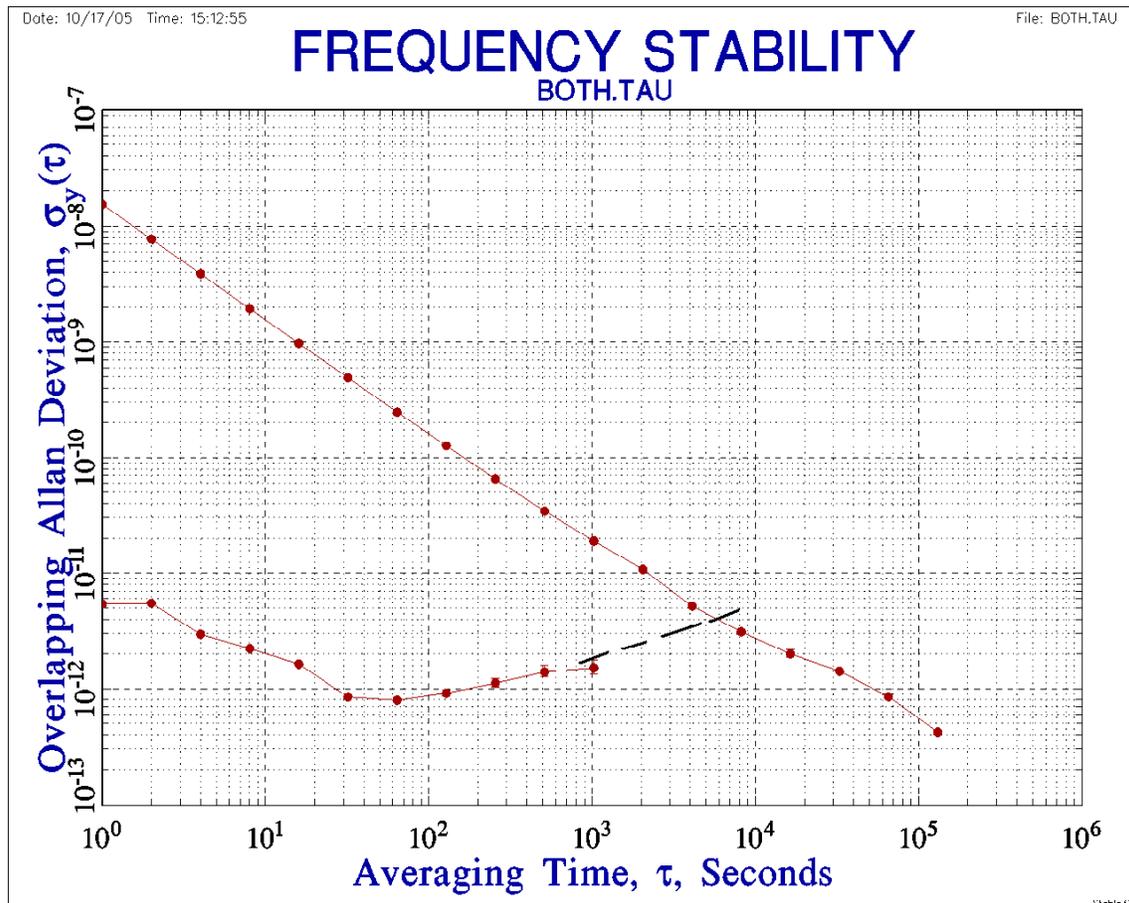


Abbildung 11: Sigma-Tau-Kurven von HP10811 und M12+ in einem gemeinsamen Diagramm

Dieser Schnittpunkt ist von fundamentaler Bedeutung. Für Beobachtungszeiten, die kleiner sind, als diejenige am Schnittpunkt der Kurven, weist der lokale Oszillator eine höhere Stabilität auf, als der GPS-Empfänger liefern kann. Für Beobachtungszeiten, die größer als diejenige beim Schnittpunkt, ist der GPS-Empfänger stabiler als der lokale Oszillator.

Es kann daher überhaupt keine Diskussion darüber geben, welche Regelzeitkonstante man in einem Frequenznormal für das Nachregeln des lokalen Oszillators benutzen **muss**, welches aus der Kombination aus einem Empfänger mit diesen Eigenschaften und einem lokalen Oszillator mit diesen Eigenschaften resultiert, nämlich genau diejenige, die am **Schnittpunkt** der beiden Kurven vorliegt.

Dazu ist auch die folgende kurze Überlegung hilfreich: Wenn ich die Regelzeitkonstante kleiner als diese Zeit mache, so versuche ich den lokalen Oszillator mit einem Signal nachzuregeln, das instabiler als das des Oszillators selber ist. Wenn ich die Regelzeitkonstante hingegen größer mache, so verschenke ich damit Stabilität, die der GPS-Empfänger ab dem Schnittpunkt der Kurven bereits liefern könnte. Wenn ich die Regelzeitkonstante aber auf den Schnittpunkt einstelle, so kombiniere ich damit das Beste aus beiden Welten. Das Sigma-Tau-Diagramm eines so eingestellten Frequenznormals ist also bis zur Regelzeitkonstanten mit dem des lokalen Oszillators identisch und ab dann mit dem des GPS-Empfängers.

Dies ist die **einzig** sinnvolle Begründung für eine Regelzeitkonstante im Bereich von einigen Stunden. Der in [6] hergestellte Zusammenhang zwischen der Benutzung von 1 PPS Signalen und sich angeblich daraus notwendigerweise ergebenden langen Regelzeitkonstanten ist technisch vollkommen hanebüchen und ich werde mich am Ende dieses Beitrages damit noch einmal auseinandersetzen.

Wenn man so will, so kann man Abbildung 11 durchaus auch als komplette Konstruktionsanleitung für Frequenznormale verstehen, wenn man sie nur zu lesen weiß. Die Graphik sagt nämlich: Für ein gutes Frequenznormal nehme man den **stabilsten** lokalen Oszillator, den man nur kriegen kann, weil für alle Beobachtungszeiten, die kleiner als diejenige am Schnittpunkt der Kurven sind, die Stabilität des Frequenznormals **ausschließlich** durch den lokalen Oszillator bestimmt wird und durch **nichts sonst**.

Für ein gutes Frequenznormal nehme man weiterhin den besten GPS-Empfänger, den man bekommen kann. „Den besten“ bedeutet in diesem Zusammenhang: Denjenigen, dessen Sigma-Tau-Kurve so niedrig wie möglich liegt. Der Sinn dieser Maßnahme besteht darin, den Schnittpunkt der beiden Kurven so weit wie möglich nach **links** und damit auch so weit wie möglich nach **unten** zu schieben. Denn es ist ja durchaus unerwünscht, dass die Allan-Deviation am Ausgang des Frequenznormals nach dem lokalen Minimum bei ca. $\tau = 40$ s zuerst wieder zunimmt um dann erst ab dem Schnittpunkt der Kurven wieder zu fallen. Immerhin ist aber auch so schon die Aussage möglich, dass ein so eingestelltes Frequenznormal für **alle Beobachtungszeiten** τ eine Allan-Deviation von besser 5×10^{-12} aufweisen kann.

Der in [1] benutzte μ Blox-Empfänger war schon zum damaligen Zeitpunkt in diesem Sinne nicht besonders gut geeignet, weil er eigentlich für Navigationszwecke und nicht für Anwendungen in der Horologie gebaut ist. Er ist dafür verantwortlich, dass die Autoren zu so großen Regelzeitkonstanten greifen mussten, weil seine Sigma-Tau-Kurve eben erheblich höher liegt. Der Schnittpunkt der beiden Kurven wandert damit halt entsprechend weiter nach „rechts“.

Um an dieser Stelle auch noch mal auf den Unterschied zwischen Stabilität und Genauigkeit aufmerksam zu machen: Nachdem die Regelzeitkonstante abgelaufen ist, kann man davon ausgehen, dass der lokale Oszillator im **Zeitmittel genau** ist, weil er vom Referenzsignal auf die richtige Frequenz gezogen wurde. Die Referenz ist jedoch nicht in der Lage, bei kurzen Beobachtungszeiten auch für Stabilität zu sorgen, dies ist ausschließlich Aufgabe des lokalen Oszillators. Die englischsprachige Literatur hat deswegen den leichtverständlichen Begriff des „Flywheel“-Oszillators, also des Schwungradoszillators für diese Aufgabe geprägt.

Wenn wir mit einem Frequenznormal bestehend aus HP10811 und MOTOROLA M12+ schon nicht die Stabilitätsforderung von 10^{-12} bei $\tau = 1$ s erfüllen können, so sollten wir nun prüfen, ob wenigstens das Ranging des Satelliten, welches ein MTIE von weniger als $3,3 \times 10^{-8}$ bei $\tau = 2000$ s erfordert, mit so einem Frequenznormal möglich ist. Dazu habe ich mit STABLE32 das MTIE einer Messreihe berechnen lassen und in Abbildung 12 dargestellt.

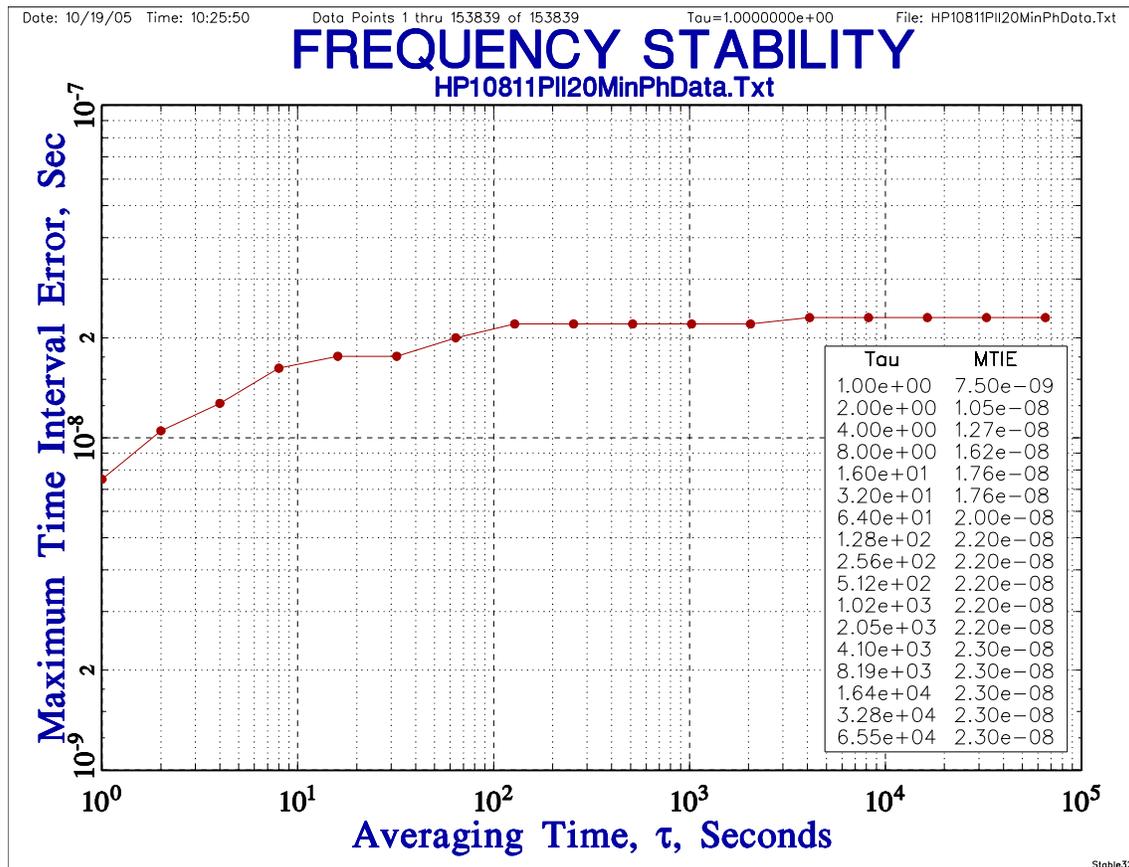


Abbildung 12: MTIE als Funktion von Tau für die Kombination HP10811/M12+

Wie man sieht, ist diese Aufgabe mit dieser Kombination von lokalem Oszillator und GPS-Empfänger leicht zu erfüllen. Es ist anzumerken, dass diese Messreihe mit einer PLL-Zeitkonstanten von 20 Minuten ermittelt wurde. Sie wurde zu einem Zeitpunkt aufgezeichnet, als der HP10811 noch nicht so gut „verpackt“ war und wesentlich schneller auf Umgebungsreize reagierte, daher lag der Schnittpunkt der Sigma-Tau-Kurven anders und es wurde 1200 Sekunden als optimale Regelzeitkonstante ermittelt. Mit dem besser verpackten HP10811 würde auch diese Kurve auch noch ein klein wenig günstiger ausfallen.

Offensichtlich in Unkenntnis der weiter oben betrachteten Zusammenhänge wird in [6] ein GPS-Empfänger mit einem normalen Quarzoszillator kombiniert. Der benutzte Empfänger ist dem M12+ durchaus gleichwertig. Der in [6] benutzten Ein-IC-Oszillatorschaltung möchte ich im positiven Sinne unterstellen, dass sie die gleichen Qualitäten hat, wie diejenige in meinem sündhaft teuren Funktionsgenerator, aber sicher keine höheren. Dann ergibt sich, wenn man die Sigma-Tau-Kurven von Empfänger und lokalem Oszillator in einem Diagramm kombiniert, in etwa die folgende Abbildung 13 für ein Frequenznormal nach [6].

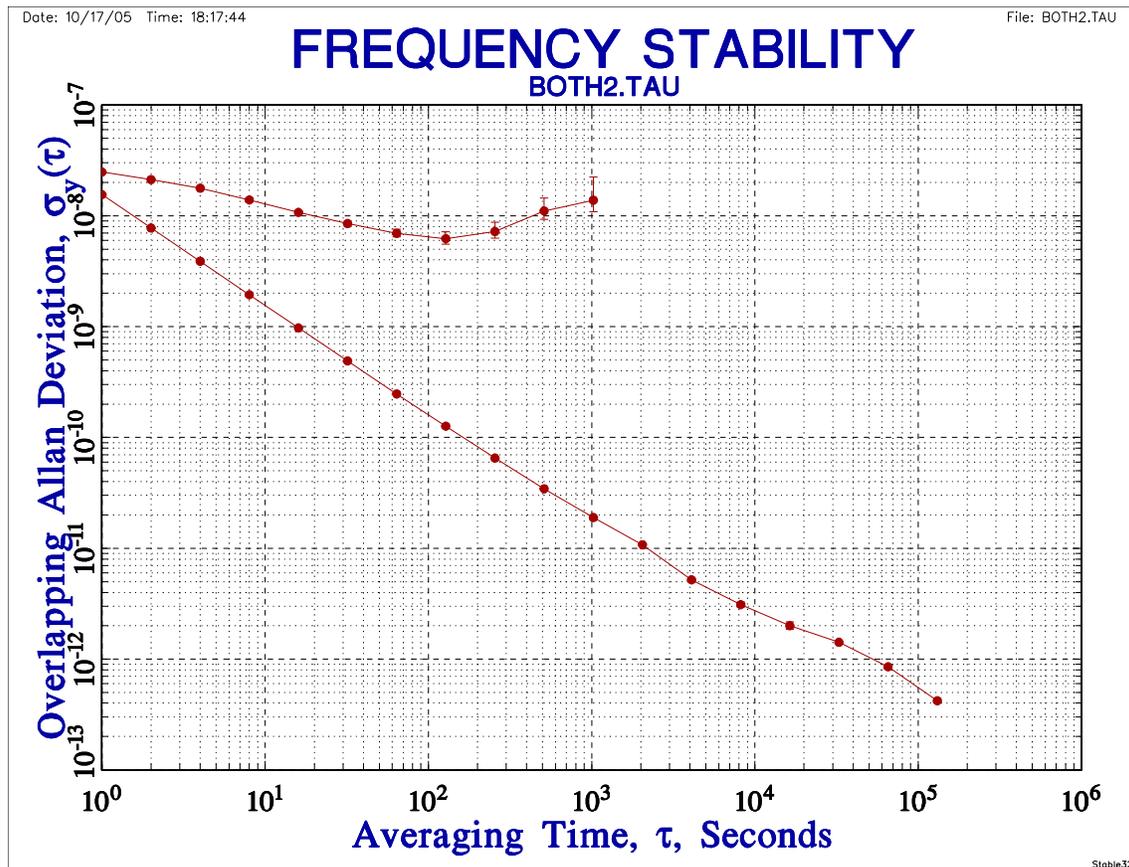


Abbildung 13: Sigma-Tau-Kurven von lokalem Oszillator und GPS-Empfänger in einem gemeinsamen Diagramm für ein Frequenznormal nach [6]

Wie man sofort sieht, haben die beiden Kurven hierbei keinen Schnittpunkt. Es gibt somit hier keine Beobachtungszeit τ , für die der lokale Oszillator stabiler als das GPS-Signal wäre und somit die Funktion eines ausgleichenden „Schwungrads“ übernehmen könnte. Deswegen hat man am Ausgang eines so aufgebauten Frequenznormals immer genau diejenige Stabilität vorliegen, die der GPS-Empfängers für die vorgegebene Beobachtungszeit hat.

Die für die Ausnutzung dieser Stabilität notwendige Mittelung über das GPS-Signal findet in diesem Frequenznormal allerdings **nicht statt** und muss daher von externer Elektronik geleistet werden. Dies kann z.B. ein Frequenzzähler tun. Wenn ich ein Frequenznormal gemäß [6] als Zeitbasis eines Frequenzzählers benutze und diesen auf eine Gate-Zeit von 1 s einstelle, dann darf ich von Messung zu Messung mit statistischen Schwankungen in der Größenordnung von $1,5 \times 10^{-8}$ rechnen, bei 10 s Gate-Zeit mit $1,5 \times 10^{-9}$ und so weiter und so fort, das heißt, der Frequenzzähler führt dadurch, dass wir seine Messzeit auf n Sekunden einstellen, gleichzeitig auch eine Mittelung über das GPS-Signal von $\tau = n$ Sekunden durch.

In diesen Situationen wird einem dann gerne vorgerechnet, dass z.B. bei der Messung eines 100 MHz Signals in einer Sekunde 10^8 ganze Schwingungen vorhanden sind und dass der Fehler von $1,5 \times 10^{-8}$ somit in der gleichen Größenordnung liege, wie der viel beschworene +/- 1 Digit-Fehler, der angeblich so typisch für jeden Frequenzzähler ist. Wer so argumentiert, hat ungefähr 30 Jahre Entwicklung im Bereich der Frequenzmesstechnik verschlafen. Heutige Zähler arbeiten nicht nur prinzipiell immer nach dem Reziprok-Verfahren, sondern sie sind durch raffinierte „Sub-Clock-Interpolation“-Mechanismen auch in der Lage, mit höherer Auflösung zu messen, als sich mancher denkt.

Aufgrund seiner 1 ns Single-Shot Zeitauflösung kann bereits mein betagter RACAL-DANA 1996 bei Frequenzmessungen unabhängig von der Eingangsfrequenz bei einer Gate-Zeit von 1 s **immer** 10 Anzeigestellen auflösen, von denen die letzte im Idealfall 10^{-10} der ersten Stelle und im ungünstigsten Fall 10^{-9} der ersten Stelle ausmacht. Selbstverständlich sind schon damit statistische Schwankungen von $1,5 \times 10^{-8}$ nicht nur nachweisbar, sondern auch störend! Bei 10 Sekunden Gate-Zeit verbessert sich der statistische Fehler zwar um den Faktor 10, der Zähler macht aber auch eine Stelle **mehr Auflösung**, so dass im Endeffekt die Situation die gleiche ist wie bei kleineren Gate-Zeiten.

Nun ist mein RACAL DANA 1996 ein „alter Kasten“ und bei einem modernen AGILENT 53132 verschlimmert sich die Situation nochmals um eine ganze Zehnerpotenz, weil dieser gar 11 Stellen Auflösung bei einer Gate-Zeit von 1 s macht (kraft seiner 150 ps Single-Shot Zeitauflösung). Wie man sieht, erfordert halbwegs moderne Messtechnik Stabilitäten, die eine Schaltung nach [6] nicht liefern kann.

Da sich wohl kaum ein Frequenzzähler wird auftreiben lassen, der über längere Gate-Zeiten als 100 s verfügt, abgesehen davon, dass dies auch einfach unpraktisch wäre, kann man mit Fug und Recht behaupten, dass die nutzbare Stabilität eines Frequenznormals nach [6] auf günstigstenfalls etwa 2×10^{-10} bei einer Gate-Zeit/Beobachtungs-Zeit von 100 s eingeschränkt ist.

Dies ist für viele Anwendungen sicher ein akzeptabler Wert, aber von vollmundig vorgetragenen Zahlen wie 10^{-13} für das GPS-System an sich eben doch fast drei Zehnerpotenzen entfernt ist. Um in den Bereich von 10^{-13} zu gelangen, wäre es notwendig, ca. 400000 s über das GPS-Signal zu mitteln, eine Zeit von mehreren Tagen, in der auch der beste verfügbare lokale Oszillator nicht mehr stabil genug ist.

Das Beispiel mit dem HP10811 hingegen zeigt, dass für alle sinnvollen Beobachtungszeiten eine Stabilität von 5×10^{-12} oder besser auch mit Amateurmitteln erzielbar ist. Auch von dieser Zahl bleibt das Konzept nach [6] noch fast zwei Zehnerpotenzen (!) entfernt.

Die PTB gibt übrigens Graphiken heraus, die im Prinzip das Sigma-Tau-Diagramm für den DCF77-Empfang in Abhängigkeit von der Entfernung vom Sender darstellen. Wegen der Effekte, die bei der Ausbreitung im Längstwellenbereich auftreten, liegen diese Kurven schon bei geringen Abständen vom Sender **erheblich** höher als die eines GPS-Empfängers. „Erheblich höher“ bedeutet, dass man, um zu einer vorgegebenen Stabilität zu gelangen, über das Signal des DCF77-Empfängers noch wesentlich länger mitteln muss, als dies beim GPS-Empfänger der Fall ist. Sinngemäß gilt diese Aussage auch für die exotischeren Möglichkeiten, die oben angesprochen wurden. Bis zu dem Zeitpunkt, wo das europäische GALILEO-System funktionsbereit ist, muss man daher das GPS-System für die Verteilung von Normalfrequenz und -Zeit derzeit als den Stand der Technik ansehen.

In [6] findet man die Formulierung „Häufig setzt man dabei z.B. recht teure OCXOs HP10811 oder Rubidiumoszillatoren ein, die über das so genannte 1-pps-Signal – ein Impuls von 1 s Länge – geregelt werden. Logischerweise dauern die eigentlichen Regelvorgänge dadurch sehr lange. Der verwendete OCXO muss daher hochstabil sein, da die Frequenz sonst in der Zwischenzeit schon wieder weggelaufen wäre.“

Da diese Formulierung dem Verfasser im Wesentlichen als Begründung dafür dient, warum er unbedingt einen normalen Quarzoszillator mit kurzer Zeitkonstante an das GPS-Signal anbinden will, lassen Sie uns ihren Wahrheitsgehalt noch einmal besonders genau untersuchen.

Zunächst einmal ist der 1PPS natürlich nicht „Ein Impuls von 1s Länge“ sondern je nach Empfänger zwischen 30 und 300 ms lang und wiederholt sich mit einer Rate von 1 / s. Motorola spezifiziert den 1 PPS z.B. so:

1PPS SIGNAL DEFINITION

0 to 5 V live pulse (0 to 3 V for M12+)

1 PPS time mark is synchronous with the mid point of the rising edge of the pulse rising from 0 V to 5 V

Rise time is approximately 20 to 30 ns

5 V pulse width is approximately 200 ms \pm 1 ms

The falling edge will occur approximately 200 ms after the rising edge

Mit einem gewissen Wohlwollen wird man „Ein Impuls von 1s Länge“ als Unschärfe in der Formulierung durchgehen lassen wollen. Hingegen ist die Behauptung „Logischerweise dauern die eigentlichen Regelvorgänge dadurch sehr lange“ schlichtweg falsch.

In der Mess- und Regeltechnik unterscheidet man zwischen kontinuierlichen und zeitdiskreten Regelsystemen. Zeitdiskrete Systeme sind solche, bei denen der Sollwert-Istwert-Vergleich und die daraus sich ergebende Veränderung des Stellwertes nicht kontinuierlich in der Zeit stattfindet sondern nur zu diskreten Zeitpunkten mit konstantem zeitlichem Abstand. Der Abstand wird dann (wie in der digitalen Signalverarbeitung) als Abtastrate bezeichnet. Im Sinne dieser Formulierung ist ein auf dem Vergleich von 1PPS-Signalen beruhendes Regelsystem sicherlich als zeitdiskret mit einer Abtastrate von 1 / s zu bezeichnen.

Eine gut bekannte Faustregel in der Mess- und Regeltechnik besagt, dass bei zeitdiskreten Regelsystemen je nach Anforderungen an die Präzision des Regelvorgangs die Periodenlänge des Abtastvorgangs nicht größer als 1/10 bis 1/100 der projektierten Regelzeitkonstanten sein soll. Im vorliegenden Fall können wir uns die Abtastrate natürlich nicht aussuchen, sie ist mit 1 / s fest vorgegeben. Hingegen können wir aus der Regel ableiten, dass selbst bei erhöhten Ansprüchen an den Regler Regelzeitkonstanten von 100 s problemlos realisiert werden können, wenn man einmal pro Sekunde den Vergleich zwischen Soll- und Istwert durchführt.

Ich hoffe, es ist weiter oben klar geworden, welch lange Zeitkonstanten man braucht, um durch Mitteln über das GPS-Signal zu einer brauchbaren Stabilität zu gelangen. Obwohl man es deshalb eigentlich **nicht** tun sollte und **nur** um zu demonstrieren, dass diese Faustregel keine reine Theorie ist und auch unter Verwendung von 1PPS Signalen vergleichsweise schnelle Regelungen möglich sind, habe ich in einem auf 1PPS-Vergleich beruhenden Frequenznormal⁹ die Regelzeitkonstante der PLL versuchsweise auf 60 s eingestellt und eine

⁹ Unveröffentlichte Eigenentwicklung des Autors

schlagartige Frequenzveränderung des lokalen Oszillators um ca. 0,026 Hz provoziert. Beachten Sie bitte, dass 0,026 Hz bezogen auf die Arbeitsfrequenz von 10 MHz immerhin schon eine Abweichung von $2,6 \times 10^{-9}$ ausmacht. Abbildung 14 zeigt das Ausregeln dieser Störung mit einem lehrbuchmäßigen Verlauf innerhalb weniger Minuten.

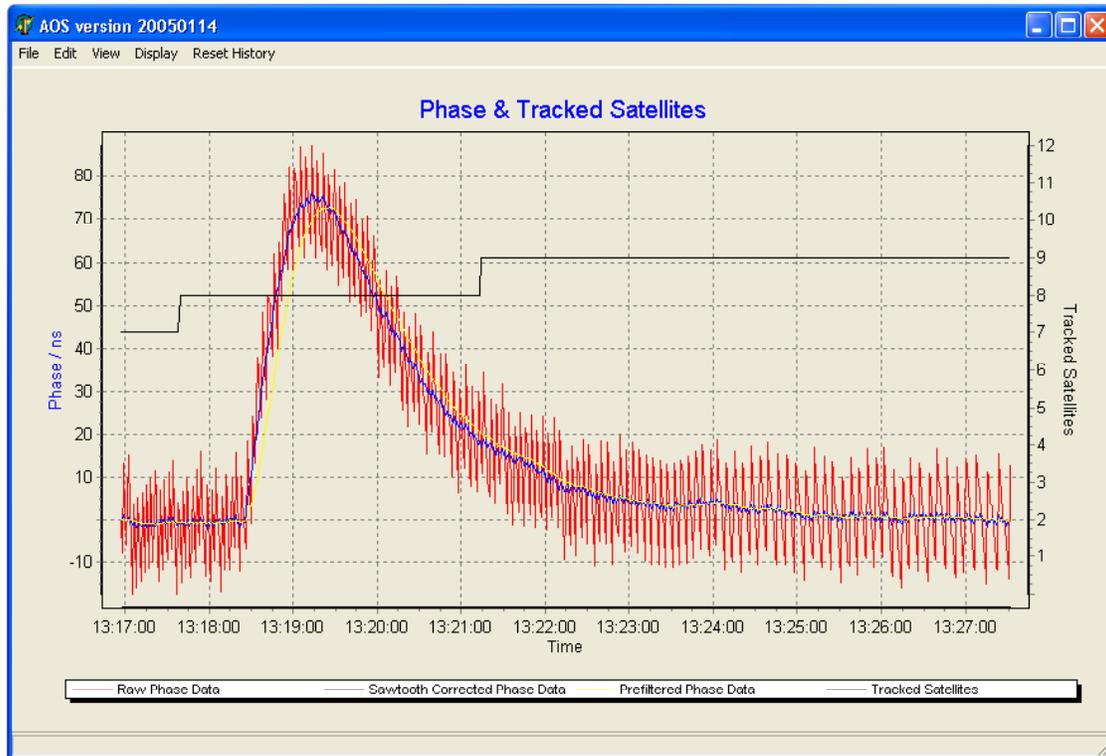


Abbildung 14: Schnelles Ausregeln einer Störung trotz PPS

In dieser Graphik stellt die rote Linie den Phasenversatz zwischen dem 1PPS des GPS-Empfängers und einem aus dem lokalen Oszillator abgeleiteten 1PPS dar. Der Phasenjitter von etwa 20 ns ist deutlich sichtbar.

In diesem Frequenznormal wird zusätzlich zu diesem „Roh-Phasenvergleich“ der so genannte „Sawtooth-Correction-Value“ des Empfängers berücksichtigt, mit dem er im seriellen Telegramm ankündigt, um wie viele ns seiner Ansicht nach der jeweils nächste 1PPS „daneben“ liegen wird. Berücksichtigt man diesen Korrekturwert, so gelangt man zu der blauen Linie. Die Frage, welche der beiden Linien man wohl eher als Istwert eines Regelkreises benutzen würde, erübrigt sich wohl.

Die gelbe Linie zeigt den Ausgang eines Vorfilters, welches seinerseits nochmals über etwa 10 Werte der blauen Linie mittelt, bevor diese Werte dann in die eigentliche PLL gefüttert werden. Die schwarze Linie zeigt die Zahl der empfangenen Satelliten an und bezieht sich auf die rechte Skala, alle anderen Linien benutzen die linke Skala.

Der Leser mag sich fragen, wo exakt er in diesem Diagramm die Zeitkonstante von 60 Sekunden wieder finden kann. Dies ist in der Tat gar nicht möglich. Stattdessen schaue man sich Abbildung 15 an. In Abbildung 15 ist der Verlauf des Stellwertes (blaue Linie, linke Skala) und der Temperaturverlauf im Frequenznormal (schwarze Linie, rechte Skala) dargestellt.

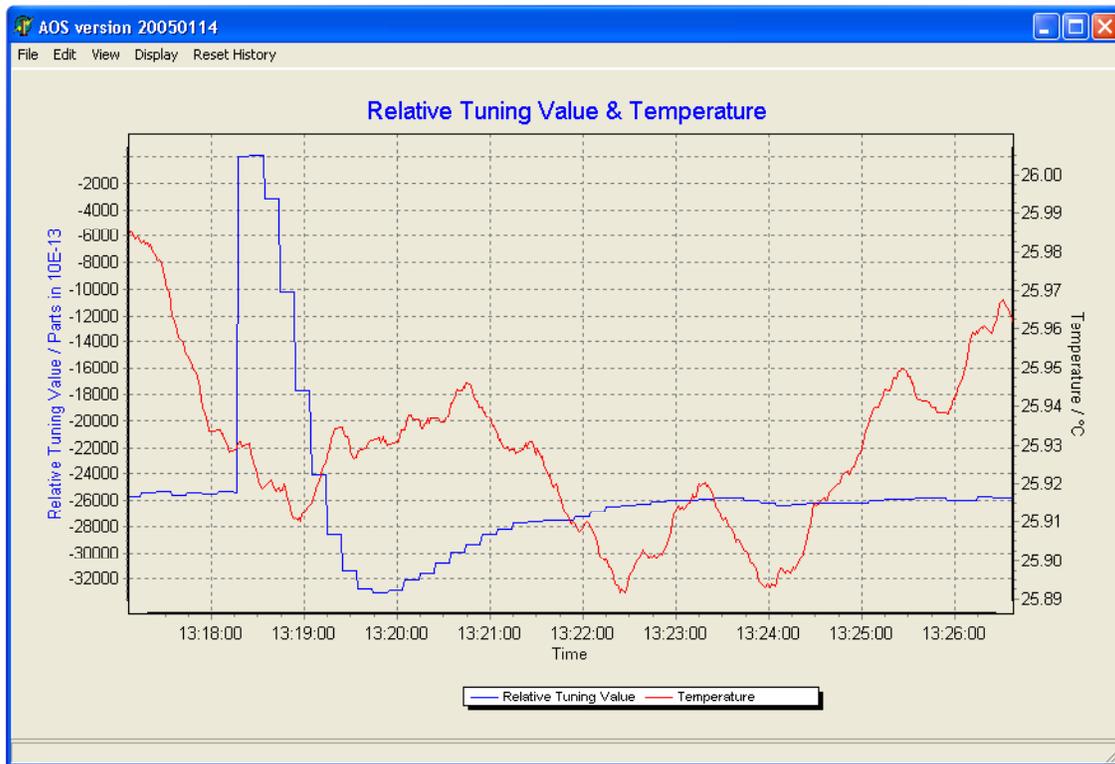


Abbildung 15: Verlauf des Stellwertes beim Ausregeln

Wie man an diesem Diagramm sieht, hat die Regelung den lokalen Oszillator genau eine Minute nach der plötzlichen Störung kurz nach 13:18 wieder auf den richtigen Wert nachgeregelt. Genau dieser Zeitraum, den die PLL benötigt, den lokalen Oszillator zunächst einmal wieder auf die Sollfrequenz zu bringen, bezeichnet man als die „natürliche“ Zeitkonstante der PLL.

Da aber in dieser einen Minute die Phase entsprechend weggelaufen ist, muss die Regelung die Frequenz des Oszillators nun einmal nach unten „durchschwingen“ lassen, um den Phasenversatz danach langsam wieder auf 0 zu bringen. Unmittelbar nach dem Versuch wurde die Zeitkonstante natürlich wieder auf den richtigen Wert von 2 h eingestellt.

Hiermit wurde auch der praktische Nachweis dafür erbracht, dass unter Benutzung von 1PPS-Signalen schnelle Regelungen möglich sind. Die tatsächlichen Gründe für die Benutzung hochqualitativer lokaler Oszillatoren und langen Regelzeitkonstanten sind, wenn dieser Beitrag sein Ziel auch nur halbwegs erreicht hat, bereits weiter oben deutlich geworden, die Begründung in [6] ist Unsinn.

Fazit: Wir haben gelernt, wie extrem hoch die Anforderungen an die USOs für die Mars Expedition tatsächlich sind. Zwar haben wir den Nachweis dafür erbracht, dass ein Frequenznormal, welches aus gängigen verfügbaren Mitteln gebaut ist, die Stabilitätsforderung für das Ranging des Satelliten erfüllt, wenn man bei seiner Konstruktion einige grundlegende Regeln beachtet. Von der „Kommunikations-Stabilität“ von 10^{-12} bei $\tau = 1$ s sind wir mit diesem Frequenznormal aber noch um den Faktor 5 entfernt. Ein Oszillator wie der FTS1200 würde aufgrund seiner Spezifikation schon zum USO für diese Mission taugen. Gegenwärtig liegen

in Marburg bei Karl Meinzer zwei weitere Oszillatoren der gleichen Klasse vor und es wird Aufgabe der nächsten Zeit sein, diese drei Oszillatoren untereinander zu vergleichen, um zu gesicherten Erkenntnissen über ihre Eigenschaften zu gelangen.

[1] Dipl.-Ing. Günter König, DJ8CY , Dipl.-Ing. Robert Tyrakowski, DK7NT , Dr. rer. nat. Horst Schütze, DF7HSA, Dipl.-Ing. Klaus Osterschek, DB4MP: Controlled Reference Oscillator (CRO) gesteuert durch GPS & ZDF, CQ-DL 9/2001 und CQ-DL 10/2001

[2] C. A. Greenhall, Common-Source Phase Error of a Dual-Mixer Stability Analyzer, TMO Progress Report 42-143 November 15, 2000

[3] Helmuth Spieler, Radiation Detectors and Signal Processing, Symposium Oct. 8 – Oct. 12, 2001; Univ. Heidelberg

[4] C. A. Greenhall, Digital Signal Processing in the Radio Science Stability Analyzer, TDA Progress Report 42-121, May 15, 1995

[5] Hewlett Packard, 10811A/B Quartz Crystal Oscillator, Operating & Service Manual

[6] Norbert Friedrich, DK6RX, Satelliten angezapft – GPS-Frequenznormal, Funkamateure 8/2005, 9/2005, 10/2005

[7] Stable32 User Manual, Download unter <http://www.wiley.com/Manual138.pdf>