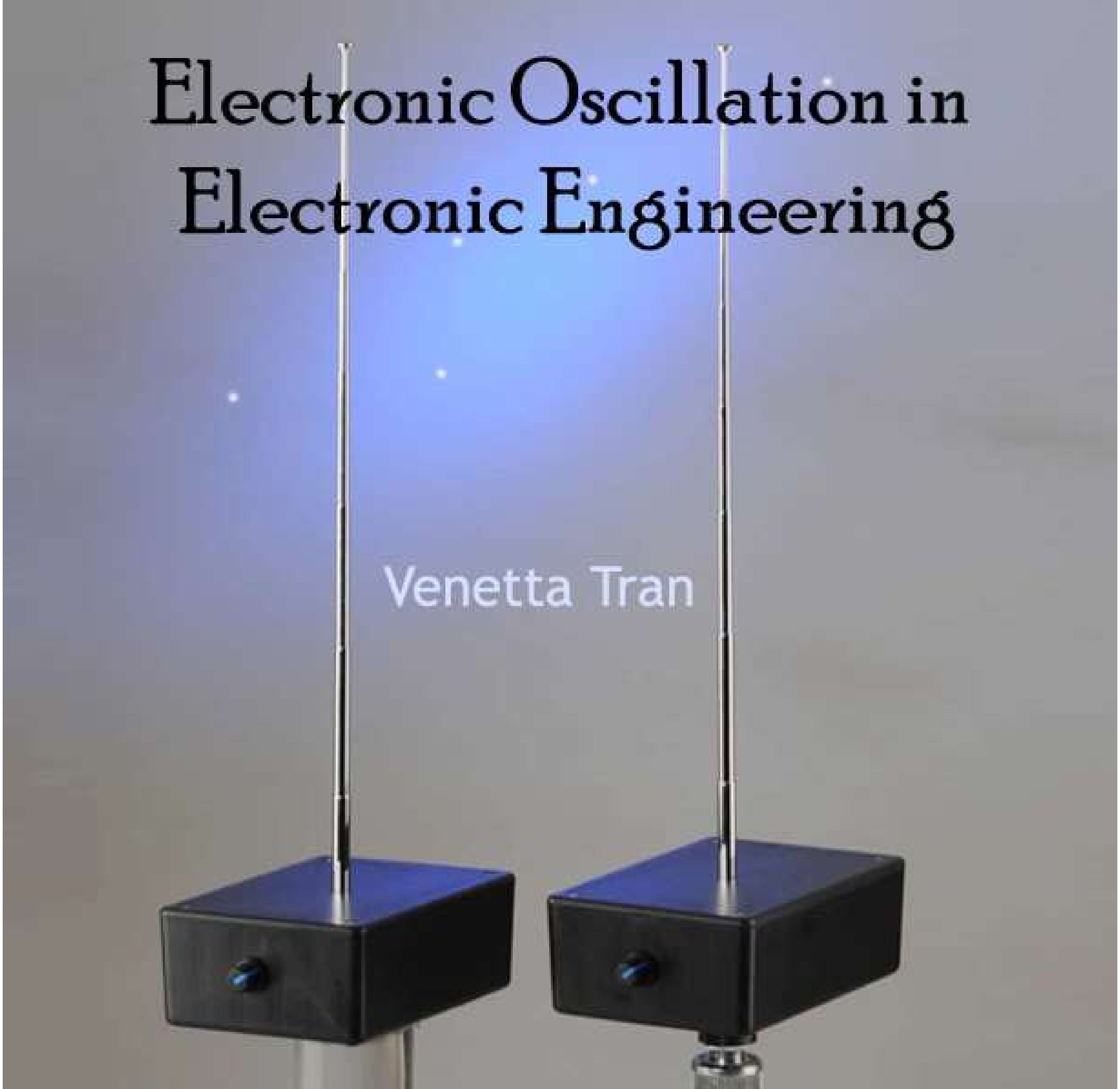

Electronic Oscillation in
Electronic Engineering
Venetta Tran

First Edition, 2012

ISBN 978-81-323-2852-0

WT

© All rights reserved.

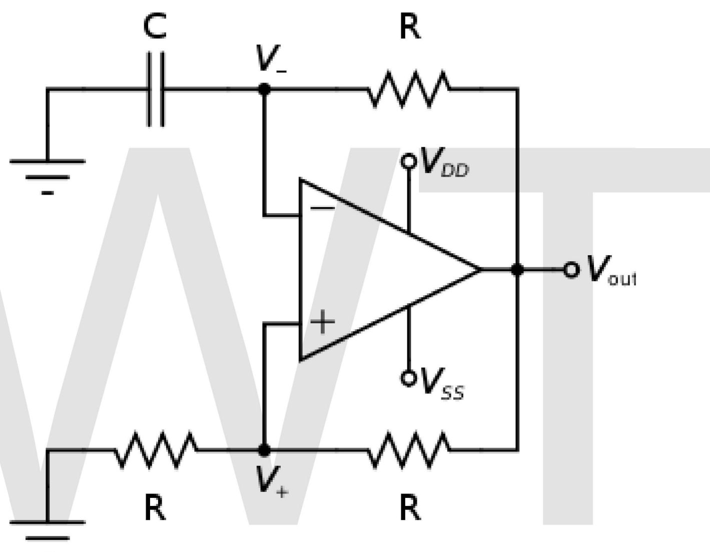
Published by:
Orange Apple
4735/22 Prakashdeep Bldg,
Ansari Road, Darya Ganj,
Delhi - 110002
Email: info@wtbooks.com

Table of Contents

WT

Chapter- 1

Electronic Oscillator



A popular op-amp relaxation oscillator.

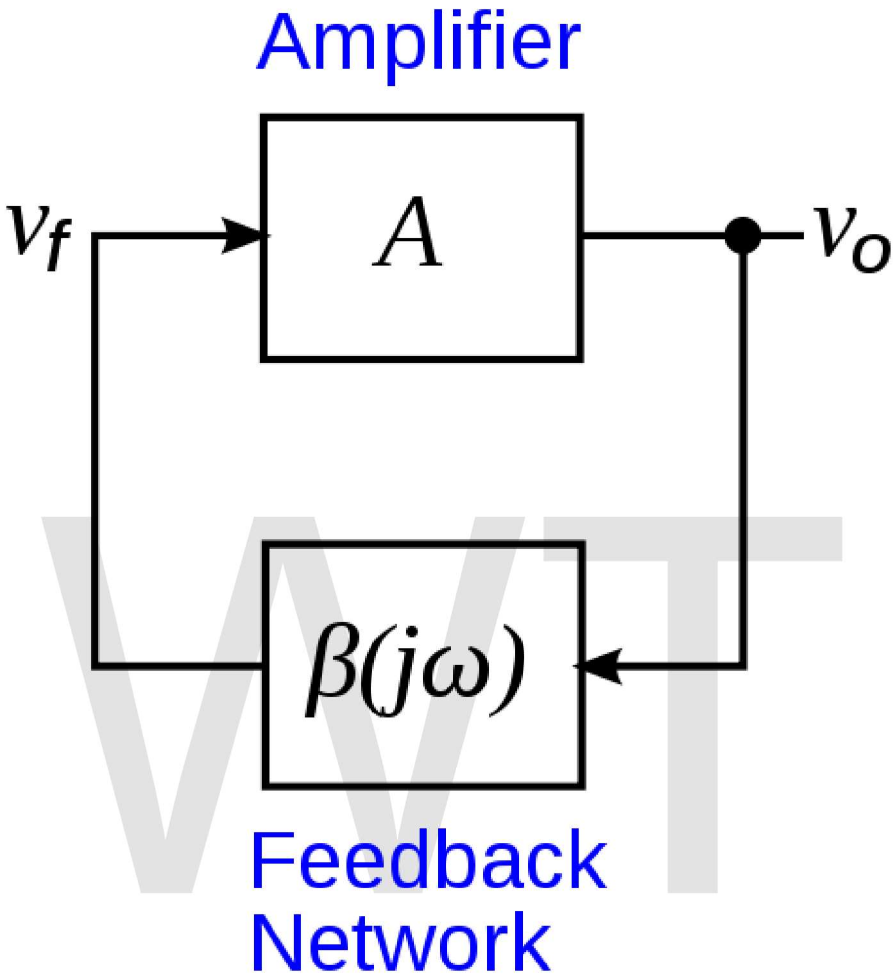
An **electronic oscillator** is an electronic circuit that produces a repetitive electronic signal, often a sine wave or a square wave. They are widely used in innumerable electronic devices. Common examples of signals generated by oscillators include signals broadcast by radio and television transmitters, clock signals that regulate computers and quartz clocks, and the sounds produced by electronic beepers and video games.

A low-frequency oscillator (LFO) is an electronic oscillator that generates an AC waveform at a frequency below ≈20 Hz. This term is typically used in the field of audio synthesizers, to distinguish it from an audio frequency oscillator.

Oscillators designed to produce a high-power AC output from a DC supply are usually called inverters.

There are two main types of electronic oscillator: the harmonic oscillator and the relaxation oscillator.

Harmonic oscillator



Block diagram of a harmonic oscillator; an amplifier *A* with its output v_o fed back into its input v_f through an electronic filter circuit, $\beta(j\omega)$.

The harmonic, or *linear*, oscillator produces a sinusoidal output.

The basic form of a harmonic oscillator is an electronic amplifier connected in a feedback loop, with its output fed back into its input through an electronic filter. When the power supply to the amplifier is first switched on, the amplifier's output consists only of noise. The noise travels around the loop, being filtered and re-amplified until it increasingly resembles the desired signal. Very quickly the signal in the loop becomes a sine wave at a single frequency.

100khz 1968 stikls

14,9895KHz

In inductive-capacitive or *LC oscillators*, the filter is a tuned circuit (often called a *tank circuit*) consisting of an inductor (L) and capacitor (C) connected together. Charge flows back and forth between the capacitor's plates through the inductor, so the tuned circuit can store electrical energy oscillating at its resonant frequency. The feedback from the amplifier creates a negative resistance that compensates for the internal resistance of the LC circuit, sustaining the oscillations. LC oscillators are typically used when a tunable frequency source is necessary, such as in signal generators, tunable radio transmitters and the local oscillators in radio receivers. Typical LC oscillator circuits are the Hartley, Colpitts and Clapp circuits. On-chip inductors usually don't have a high enough Q-factor to use in the tuned circuit.

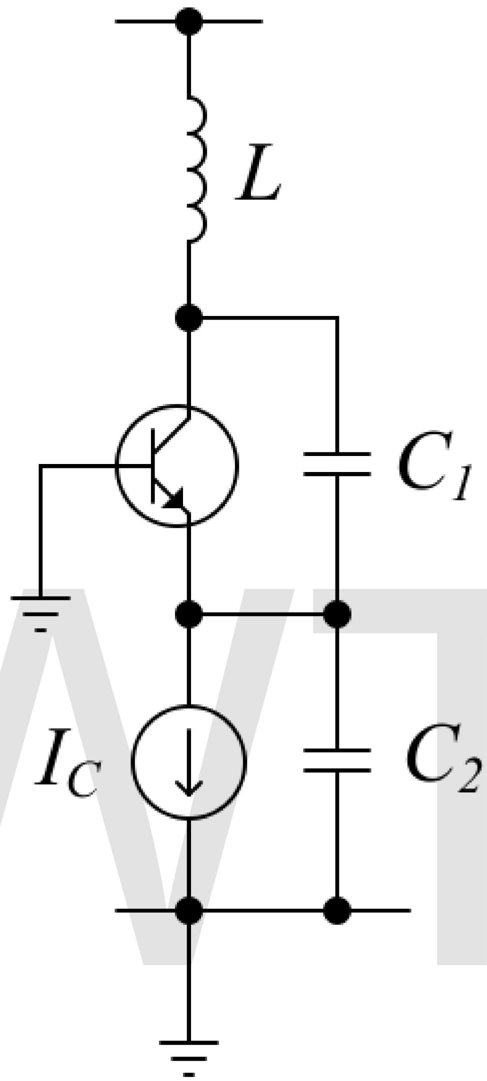
A piezoelectric crystal (commonly quartz) may take the place of the filter to stabilise the frequency of oscillation, this is called a crystal oscillator. These kinds of oscillators contain quartz crystals that mechanically vibrate as resonators, and their vibration determines the oscillation frequency. Crystals have very high Q-factor and also better temperature stability than tuned circuits, so crystal oscillators have much better frequency stability than LC or RC oscillators. They are used to stabilize the frequency of most radio transmitters, and to generate the clock signal in computers. The Pierce oscillator circuit is often used for crystal oscillators. Because the crystal is an off-chip component, it adds some cost and complexity to the system design, but the crystal itself is generally quite inexpensive.

Surface acoustic wave (SAW) devices are a kind of crystal oscillator, but achieve much higher frequencies by establishing standing waves on the surface of the quartz crystal. These are more expensive than crystal oscillators, and are used in specialized applications which require a direct and very accurate high frequency reference, for example, in cellular telephones.

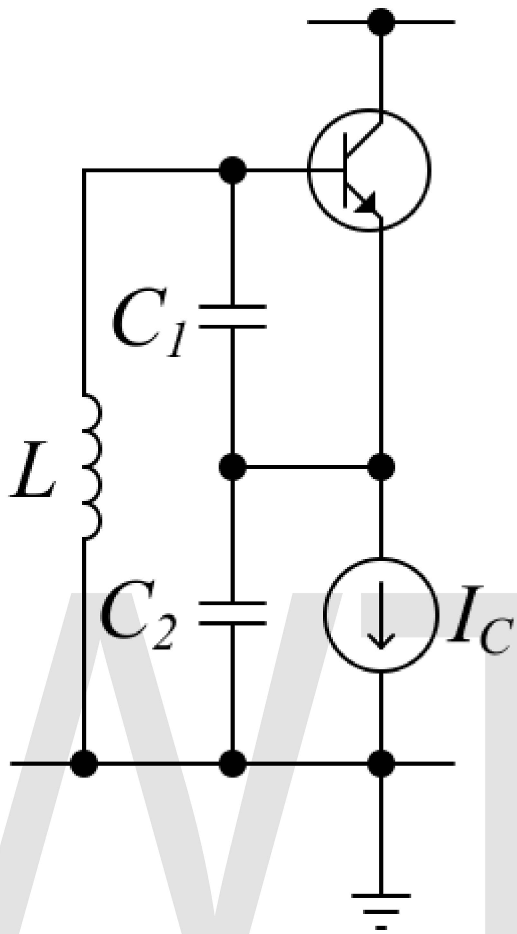
There are many ways to implement harmonic oscillators, because there are different ways to amplify and filter. Some of the different circuits are:

- Armstrong oscillator
- Hartley oscillator
- Colpitts oscillator
- Clapp oscillator
- Delay line oscillator
- Pierce oscillator (crystal)
- Phase-shift oscillator
- RC oscillator (Wien Bridge and "Twin-T")
- Cross-coupled LC oscillator
- Vackář oscillator
- Opto-Electronic Oscillator.

WT



Cb colp



WT

Cc colp2

Relaxation oscillator

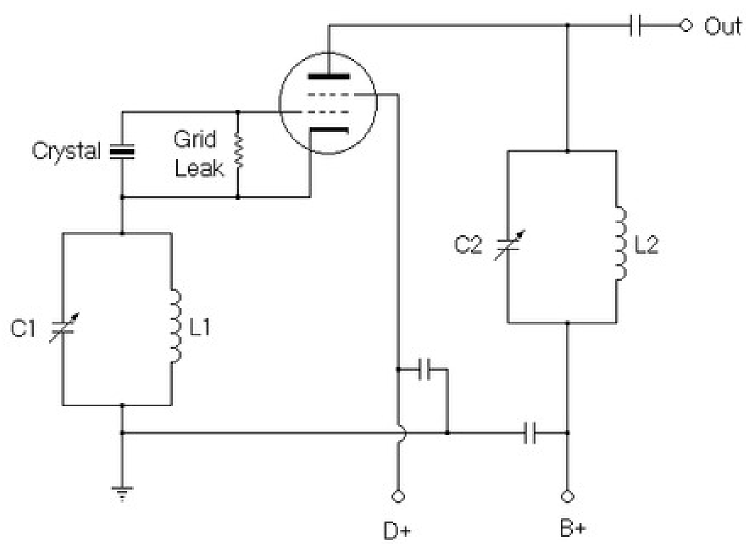
A relaxation oscillator produces a non-sinusoidal output, such as a square, sawtooth or triangle wave. It contains an energy-storing element (a capacitor or, more rarely, an inductor) and a nonlinear trigger circuit (a latch, Schmitt trigger, or negative resistance element) that periodically charges and discharges the energy stored in the storage element thus causing abrupt changes in the output waveform.

Square-wave relaxation oscillators are used to provide the clock signal for sequential logic circuits such as timers and counters, although crystal oscillators are often preferred for their greater stability. Triangle wave or sawtooth oscillators are used in the timebase circuits that generate the horizontal deflection signals for cathode ray tubes in analogue oscilloscopes and television sets. In function generators, this triangle wave may then be further shaped into a close approximation of a sine wave.

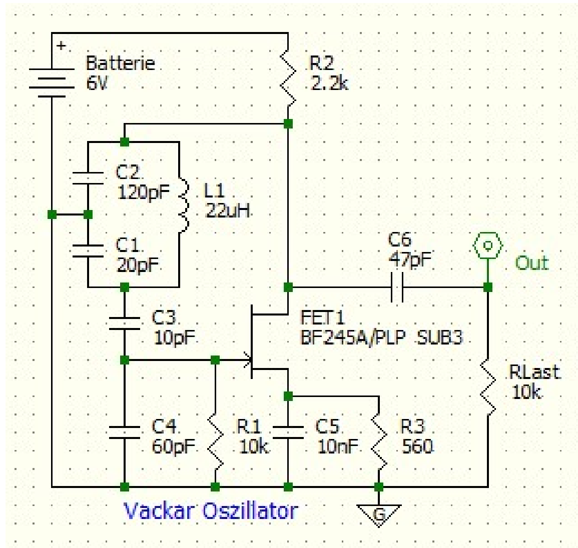
Ring oscillators are built of a ring of active delay stages. Generally the ring has an odd number of inverting stages, so that there is no single stable state for the internal ring voltages. Instead, a single transition propagates endlessly around the ring.

Types of relaxation oscillator circuits include:

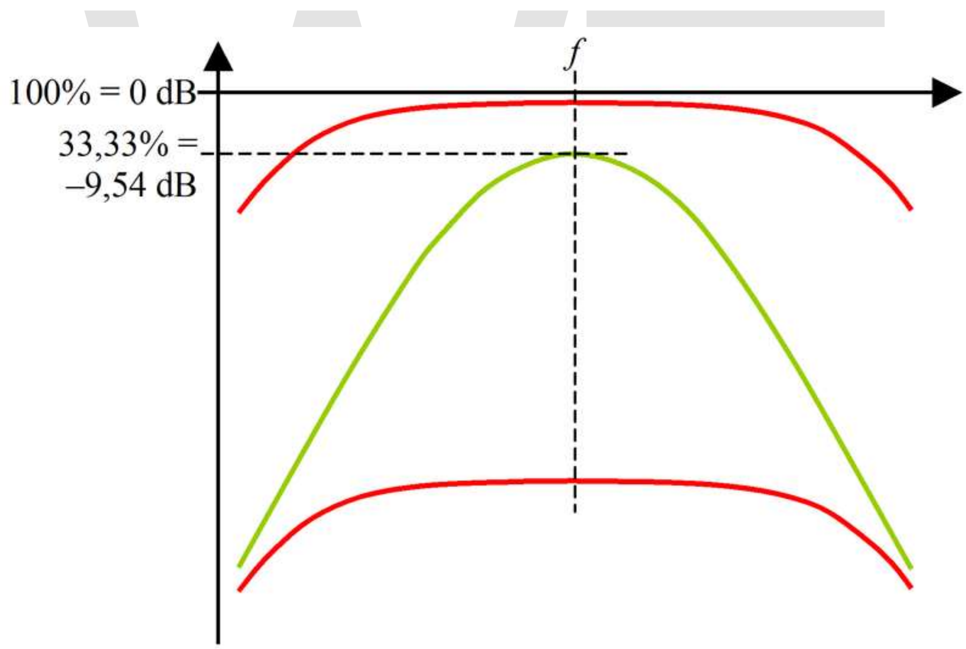
- multivibrator
- ring oscillator
- delay line oscillator
- rotary traveling wave oscillator.



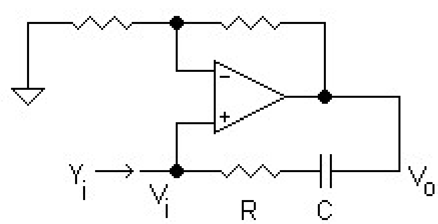
Tri-Tet Oscillator Schematic



Vackar Oszillator FET



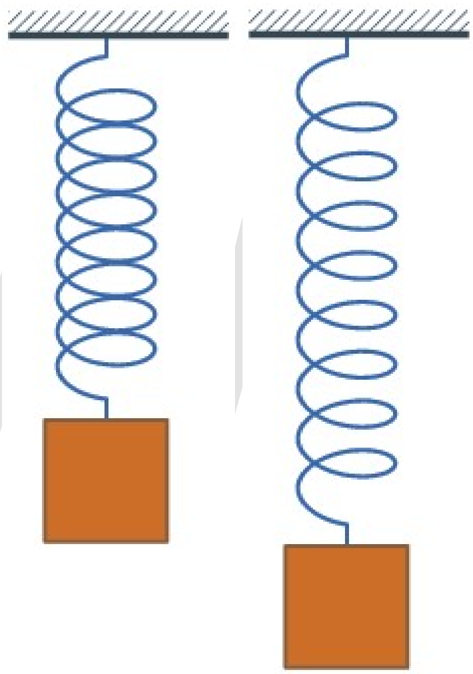
Wien bridge bode plot



Wien bridge yin

Chapter- 2

Harmonic Oscillator



An undamped spring-mass system is a simple harmonic oscillator.

In classical mechanics, a **harmonic oscillator** is a system that, when displaced from its equilibrium position, experiences a restoring force, *F*, proportional to the displacement, *x*:

$$F = -kx$$

where *k* is a positive constant.

If *F* is the only force acting on the system, the system is called a **simple harmonic oscillator**, and it undergoes simple harmonic motion: sinusoidal oscillations about the equilibrium point, with a constant amplitude and a constant frequency (which does not depend on the amplitude).

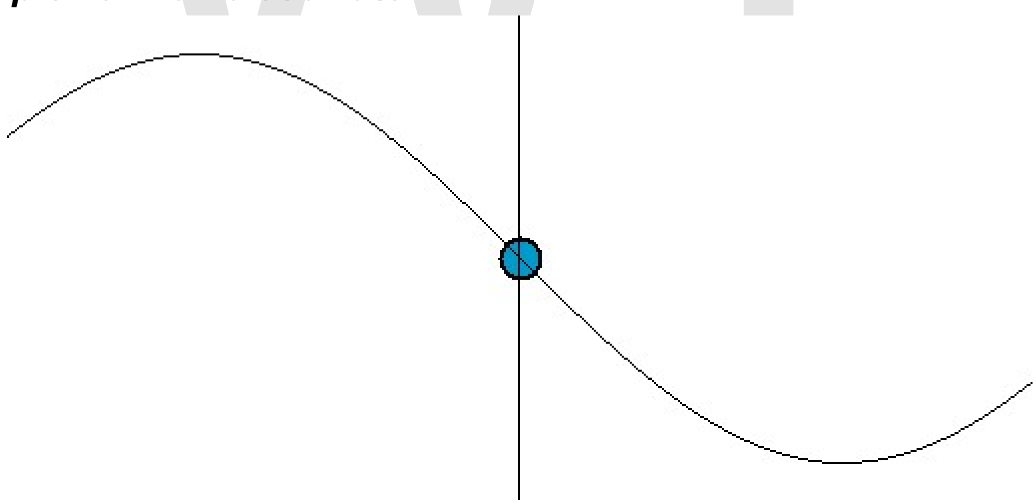
If a frictional force (damping) proportional to the velocity is also present, the harmonic oscillator is described as a **damped oscillator**. Depending on the friction coefficient, the system can:

- Oscillate with a frequency smaller than in the non-damped case, and an amplitude decreasing with time (*underdamped* oscillator).
- Decay exponentially to the equilibrium position, without oscillations (*overdamped* oscillator).

If an external time dependent force is present, the harmonic oscillator is described as a **driven oscillator**.

Mechanical examples include pendula (with small angles of displacement), masses connected to springs, and acoustical systems. Other analogous systems include electrical harmonic oscillators such as RLC circuits. The harmonic oscillator model is very important in physics, because any mass subject to a force in stable equilibrium acts as a harmonic oscillator for small vibrations. Harmonic oscillators occur widely in nature and are exploited in many manmade devices, such as clocks and radio circuits. They are the source of virtually all sinusoidal vibrations and waves.

Simple harmonic oscillator



Simple harmonic motion.

A simple harmonic oscillator is an oscillator that is neither driven nor damped. It consists of a mass m, which experiences a single force, F, which pulls the mass in the direction of the point x=0 and depends only on the mass's position x and a constant k. Newton's second law for the system is

$$F = ma = m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx.$$

Solving this differential equation, we find that the motion is described by the function

$$x(t) = A\cos(2\pi ft + \phi),$$

where

$$f = \frac{1}{2\pi}\sqrt{\frac{k}{m}} = \frac{1}{T}.$$

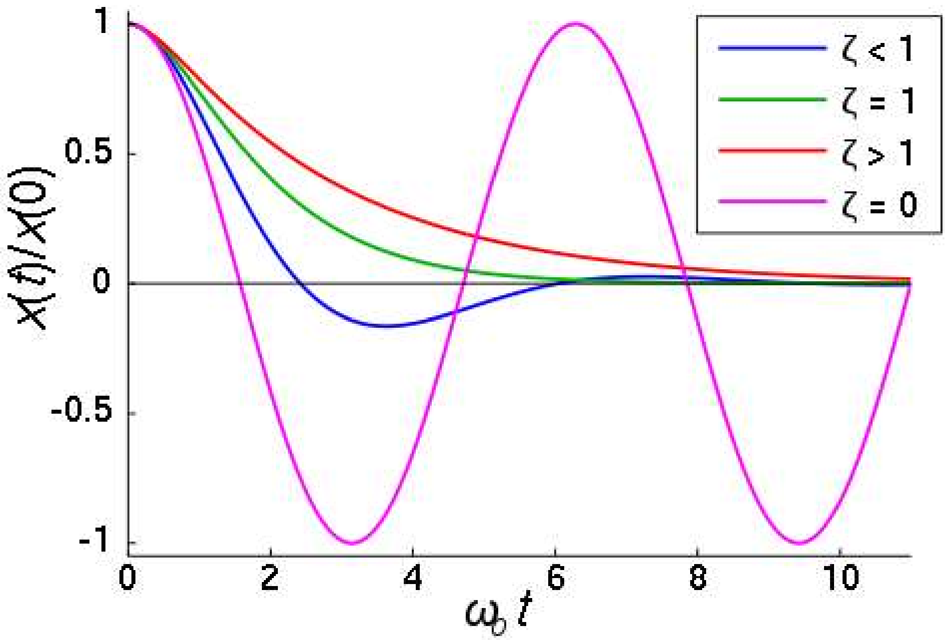
The motion is periodic— repeating itself in a sinusoidal fashion with constant amplitude, A. In addition to its amplitude, the motion of a simple harmonic oscillator is characterized by its period T, the time for a single oscillation or its frequency $f = {}^1\!/_T$, the number of cycles per unit time. The position at a given time t also depends on the phase, φ, which determines the starting point on the sine wave. The period and frequency are determined by the size of the mass m and the force constant k, while the amplitude and phase are determined by the starting position and velocity.

The velocity and acceleration of a simple harmonic oscillator oscillate with the same frequency as the position but with shifted phases. The velocity is maximum for zero displacement, while the acceleration is in the opposite direction as the displacement.

The potential energy stored in a simple harmonic oscillator at position x is

$$U = \frac{1}{2}kx^2.$$

Damped harmonic oscillator



Dependence of the system behavior on the value of the damping ratio ζ.

In real oscillators friction, or **damping**, slows the motion of the system. In many vibrating systems the frictional force F_f can be modeled as being proportional to the velocity v of the object: $F_f = -cv$, where c is called the *viscous damping coefficient*.

Newton's second law for damped harmonic oscillators is then

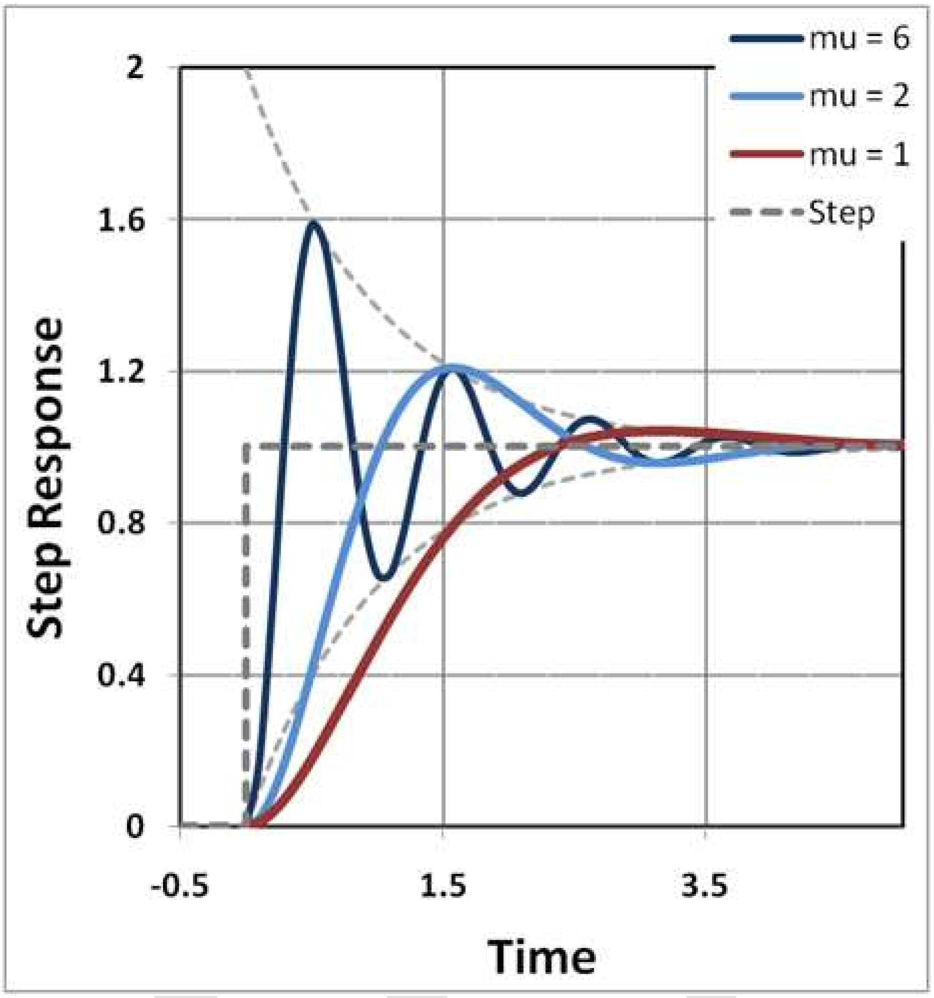
$$F = -kx - c\frac{dx}{dt} = m\frac{d^2x}{dt^2}.$$

This is rewritten into the form

$$\frac{d^2x}{dt^2} + 2\zeta\omega_0\frac{dx}{dt} + \omega_0^2 x = 0,$$

where

$\omega_0 = \sqrt{\frac{k}{m}}$ is called the 'undamped angular frequency of the oscillator' and
$\zeta = \frac{c}{2m\omega_0}$ is called the 'damping ratio'.



Step-response of a damped harmonic oscillator; curves are plotted for three values of $\mu = \omega_1 = \omega_0\sqrt{1-\zeta^2}$. Time is in units of the decay time $\tau = 1/(\zeta\omega_0)$.

The value of the damping ratio ζ critically determines the behavior of the system. A damped harmonic oscillator can be:

- *Overdamped* ($\zeta > 1$): The system returns (exponentially decays) to equilibrium without oscillating. Larger values of the damping ratio ζ return to equilibrium slower.
- *Critically damped* ($\zeta = 1$): The system returns to equilibrium as quickly as possible without oscillating. This is often desired for the damping of systems such as doors.
- *Underdamped* ($\zeta < 1$): The system oscillates (with a slightly different frequency than the undamped case) with the amplitude gradually decreasing to zero. The angular frequency of the underdamped harmonic oscillator is given by

$$\omega_1 = \omega_0\sqrt{1-\zeta^2}.$$

The Q factor of a damped oscillator is defined as

$$Q = 2\pi \times \frac{\text{Energy stored}}{\text{Energy lost per cycle}}.$$

Q is related to the damping ratio by the equation $Q = \frac{1}{2\zeta}$.

Driven harmonic oscillators

Driven harmonic oscillators are damped oscillators further affected by an externally applied force *F*(*t*).

Newton's second law takes the form

$$F(t) - kx - c\frac{dx}{dt} = m\frac{d^2x}{dt^2}.$$

It is usually rewritten into the form

$$\frac{d^2x}{dt^2} + 2\zeta\omega_0\frac{dx}{dt} + \omega_0^2 x = \frac{F(t)}{m}.$$

This equation can be solved exactly for any driving force using the solutions *z*(*t*) to the unforced equation, which satisfy:

$$\frac{d^2z}{dt^2} + 2\zeta\omega_0\frac{dz}{dt} + \omega_0^2 z = 0,$$

which can be expressed as a damped sinusoidal oscillation:

$$z(t) = Ae^{-\zeta\omega_0 t}\ \sin\left(\sqrt{1-\zeta^2}\ \omega_0 t + \varphi\right),$$

in the case where $\zeta \leq 1$. The amplitude *A* and phase φ determine the behavior needed to match the initial conditions.

Step input

In the case $\zeta < 1$ and a unit step input with $x(0) = 0$:

$$F(t) = \begin{cases} \omega_0^2 & t \geq 0 \\ 0 & t < 0 \end{cases}$$

the solution is:

$$x(t) = 1 - e^{-\zeta\omega_0 t} \frac{\sin\left(\sqrt{1-\zeta^2}\, \omega_0 t + \varphi\right)}{\sin(\varphi)} \ ,$$

with phase φ given by

$$\cos\varphi = \zeta.$$

This behavior is found in (for example) feedback amplifiers, where the amplifier design is adjusted to obtain the fastest step response possible without undue overshoot or undershoot and with an adequate settling time.

The time an oscillator needs to adapt to changed external conditions is of the order $\tau=1/(\zeta\omega_0)$. In physics, the adaptation is called relaxation, and τ is called the relaxation time.

In electrical engineering, a multiple of τ is called the *settling time*, i.e. the time necessary to insure the signal is within a fixed departure from final value, typically within 10%. The term *overshoot* refers to the extent the maximum response exceeds final value, and *undershoot* refers to the extent the response falls below final value for times following the maximum response.

Sinusoidal driving force

In the case of a sinusoidal driving force:

$$\frac{d^2x}{dt^2} + 2\zeta\omega_0\frac{dx}{dt} + \omega_0^2 x = \frac{1}{m}F_0\sin(\omega t) \ ,$$

where $F_0$is the driving amplitude and ω is the driving frequency for a sinusoidal driving mechanism. This type of system appears in AC driven RLC circuits (resistor-inductor-capacitor) and driven spring systems having internal mechanical resistance or external air resistance.

The general solution is a sum of a transient solution that depends on initial conditions, and a steady state that is independent of initial conditions and depends only on the driving amplitude F_0, driving frequency, ω , undamped angular frequency ω_0, and the damping ratio ζ.

The steady-state solution is proportional to the driving force with an induced phase change of ϕ:

$$x(t) = \frac{F_0}{mZ_m\omega}\sin(\omega t + \phi)$$

where

$$Z_m = \sqrt{\left(2\omega_0\zeta\right)^2 + \frac{1}{\omega^2}\left(\omega_0^2 - \omega^2\right)^2}$$

is the absolute value of the impedance or linear response function and

$$\phi = \arctan\left(\frac{2\omega\omega_0\zeta}{\omega^2 - \omega_0^2}\right)$$

is the phase of the oscillation relative to the driving force.

For a particular driving frequency called the resonance frequency $\omega_r = \omega_0\sqrt{1 - 2\zeta^2}$, the amplitude (for a given F_0) is maximum. For underdamped systems the value of the amplitude can become quite large near the resonance frequency.

The transient solutions are the same as the unforced ($F_0 = 0$) damped harmonic oscillator and represent the systems response to other events that occurred previously. The transient solutions typically die out rapidly enough that they can be ignored.

Parametric oscillators

A parametric oscillator is a harmonic oscillator whose parameters oscillate in time. For example, a well known parametric oscillator is a child on a swing where periodically changing the child's center of gravity causes the swing to oscillate. The varying of the parameters drives the system. Examples of parameters that may be varied are its resonance frequency ω and damping β.

Parametric oscillators are used in many applications. The classical varactor parametric oscillator oscillates when the diode's capacitance is varied periodically. The circuit that varies the diode's capacitance is called the "pump" or "driver". In microwave electronics, waveguide/YAG based parametric oscillators operate in the same fashion. The designer varies a parameter periodically to induce oscillations.

Parametric oscillators have been developed as low-noise amplifiers, especially in the radio and microwave frequency range. Thermal noise is minimal, since a reactance (not a resistance) is varied. Another common use is frequency conversion, e.g., conversion from audio to radio frequencies. For example, the Optical parametric oscillator converts an input laser wave into two output waves of lower frequency (ω_s,ω_i).

Parametric resonance occurs in a mechanical system when a system is parametrically excited and oscillates at one of its resonant frequencies. Parametric excitation differs from forcing, since the action appears as a time varying modification on a system parameter. This effect is different from regular resonance because it exhibits the instability phenomenon.

Universal oscillator equation

The equation

$$\frac{\mathrm{d}^2q}{\mathrm{d}\tau^2} + 2\zeta\frac{\mathrm{d}q}{\mathrm{d}\tau} + q = 0$$

is known as the **universal oscillator equation** since all second order linear oscillatory systems can be reduced to this form. This is done through nondimensionalization.

If the forcing function is $f(t) = \cos(\omega t) = \cos(\omega t_c \tau) = \cos(\omega\tau)$, where $\omega = \omega t_c$, the equation becomes

$$\frac{\mathrm{d}^2q}{\mathrm{d}\tau^2} + 2\zeta\frac{\mathrm{d}q}{\mathrm{d}\tau} + q = \cos(\omega\tau).$$

The solution to this differential equation contains two parts, the "transient" and the "steady state".

Transient solution

The solution based on solving the ordinary differential equation is for arbitrary constants c_1 and c_2

$$q_t(\tau) = \begin{cases} e^{-\zeta\tau}\left(c_1 e^{\tau\sqrt{\zeta^2-1}} + c_2 e^{-\tau\sqrt{\zeta^2-1}}\right) & \zeta > 1 \text{ (overdamping)} \\ e^{-\zeta\tau}(c_1 + c_2\tau) = e^{-\tau}(c_1 + c_2\tau) & \zeta = 1 \text{ (critical damping)} \\ e^{-\zeta\tau}\left[c_1\cos\left(\sqrt{1-\zeta^2}\tau\right) + c_2\sin\left(\sqrt{1-\zeta^2}\tau\right)\right] & \zeta < 1 \text{ (underdamping)} \end{cases}$$

The transient solution is independent of the forcing function.

Steady-state solution

Apply the "complex variables method" by solving the auxiliary equation below and then finding the real part of its solution:

$$\frac{d^2q}{d\tau^2} + 2\zeta\frac{dq}{d\tau} + q = \cos(\omega\tau) + i\sin(\omega\tau) = e^{i\omega\tau}.$$

Supposing the solution is of the form

$$q_s(\tau) = Ae^{i(\omega\tau+\phi)}.$$

Its derivatives from zero to 2nd order are

$$q_s = Ae^{i(\omega\tau+\phi)},\ \frac{dq_s}{d\tau} = i\omega Ae^{i(\omega\tau+\phi)},\ \frac{d^2q_s}{d\tau^2} = -\omega^2 Ae^{i(\omega\tau+\phi)}.$$

Substituting these quantities into the differential equation gives

$$-\omega^2 Ae^{i(\omega\tau+\phi)} + 2\zeta i\omega Ae^{i(\omega\tau+\phi)} + Ae^{i(\omega\tau+\phi)} = (-\omega^2 A + 2\zeta i\omega A + A)e^{i(\omega\tau+\phi)} = e^{i\omega\tau}.$$

Dividing by the exponential term on the left results in

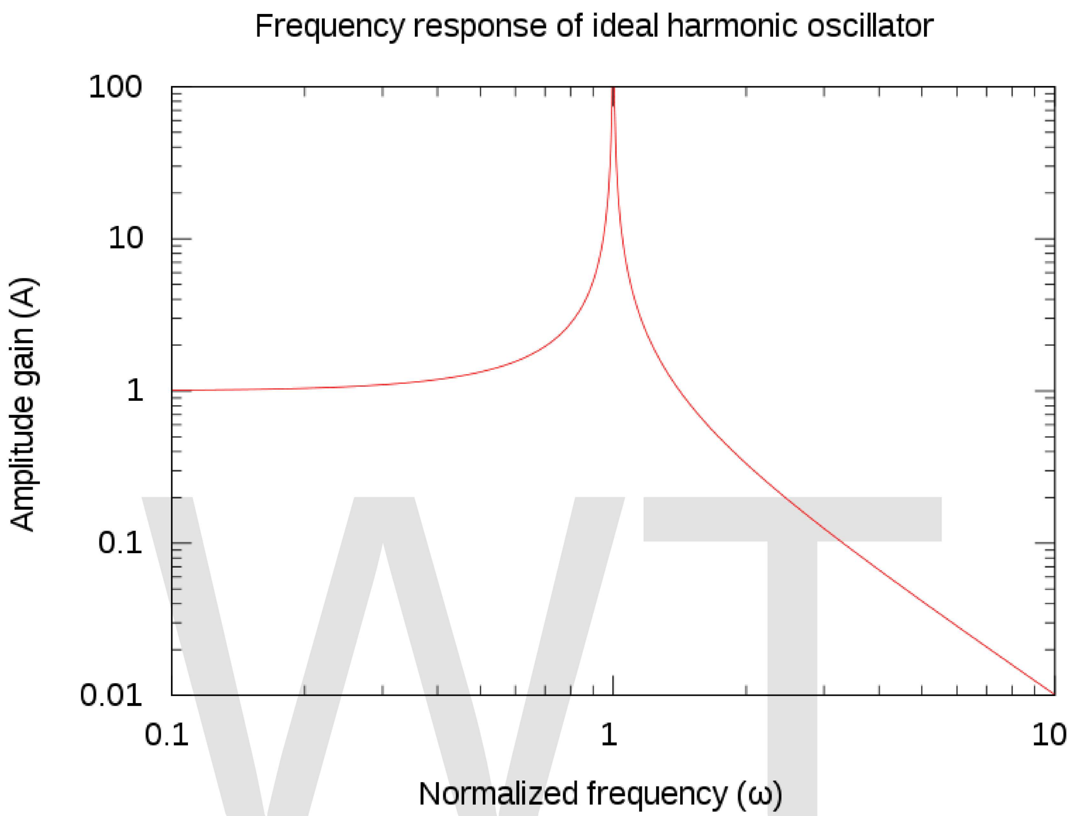
$$-\omega^2 A + 2\zeta i\omega A + A = e^{-i\phi} = \cos\phi - i\sin\phi.$$

Equating the real and imaginary parts results in two independent equations

$$A(1-\omega^2) = \cos\phi \qquad 2\zeta\omega A = -\sin\phi.$$

WT

Amplitude part



Bode plot of the frequency response of an ideal harmonic oscillator.

Squaring both equations and adding them together gives

$$\left.\begin{aligned} A^2(1-\omega^2)^2 &= \cos^2\phi \\ (2\zeta\omega A)^2 &= \sin^2\phi \end{aligned}\right\} \Rightarrow A^2[(1-\omega^2)^2 + (2\zeta\omega)^2] = 1.$$

By convention the positive root is taken since amplitude is usually considered a positive quantity. Therefore,

$$A = A(\zeta, \omega) = \frac{1}{\sqrt{(1-\omega^2)^2 + (2\zeta\omega)^2}}.$$

Compare this result with the theory section on resonance, as well as the "magnitude part" of the RLC circuit. This amplitude function is particularly important in the analysis and understanding of the frequency response of second-order systems.

Phase part

To solve for φ, divide both equations to get

$$\tan\phi = -\frac{2\zeta\omega}{1-\omega^2} = \frac{2\zeta\omega}{\omega^2-1} \Rightarrow \phi \equiv \phi(\zeta,\omega) = \arctan\left(\frac{2\zeta\omega}{\omega^2-1}\right).$$

This phase function is particularly important in the analysis and understanding of the frequency response of second-order systems.

Full solution

Combining the amplitude and phase portions results in the steady-state solution

$$q_s(\tau) = A(\zeta,\omega)\cos(\omega\tau + \phi(\zeta,\omega)) = A\cos(\omega\tau+\phi).$$

The solution of original universal oscillator equation is a superposition (sum) of the transient and steady-state solutions

$$q(\tau) = q_t(\tau) + q_s(\tau).$$

Equivalent systems

Harmonic oscillators occurring in a number of areas of engineering are equivalent in the sense that their mathematical models are identical. Below is a table showing analogous quantities in four harmonic oscillator systems in mechanics and electronics. If analogous parameters on the same line in the table are given numerically equal values, the behavior of the oscillators—their output waveform, resonant frequency, damping factor, etc.—are the same.

Translational Mechanical	Torsional Mechanical	Series RLC Circuit	Parallel RLC Circuit
Position x	Angle θ	Charge q	Voltage e
Velocity $\frac{dx}{dt}$	Angular velocity $\frac{d\theta}{dt}$	Current $\frac{dq}{dt}$	$\frac{de}{dt}$
Mass M	Moment of inertia I	Inductance L	Capacitance C
Spring constant K	Torsion constant μ	Elastance $1/C$	Susceptance $1/L$
Friction γ	Rotational friction Γ	Resistance R	Conductance $1/R$

Drive force $F(t)$	Drive torque $\tau(t)$	e	di/dt
Undamped resonant frequency f_n:			
$\frac{1}{2\pi}\sqrt{\frac{K}{M}}$	$\frac{1}{2\pi}\sqrt{\frac{\mu}{I}}$	$\frac{1}{2\pi}\sqrt{\frac{1}{LC}}$	$\frac{1}{2\pi}\sqrt{\frac{1}{LC}}$
Differential equation:			
$M\ddot{x}+\gamma\dot{x}+Kx=F$	$I\ddot{\theta}+\Gamma\dot{\theta}+\mu\theta=\tau$	$L\ddot{q}+R\dot{q}+q/C=e$	$C\ddot{e}+\dot{e}/R+e/L=\dot{i}$

Application to a conservative force

The problem of the simple harmonic oscillator occurs frequently in physics, because a mass at equilibrium under the influence of any conservative force, in the limit of small motions, behaves as a simple harmonic oscillator.

A conservative force is one that has a potential energy function. The potential energy function of a harmonic oscillator is:

$$V(x)=\frac{1}{2}kx^2$$

Given an arbitrary potential energy function $V(x)$, one can do a Taylor expansion in terms of x around an energy minimum ($x = x_0$) to model the behavior of small perturbations from equilibrium.

$$V(x)=V(x_0)+(x-x_0)V'(x_0)+\frac{1}{2}(x-x_0)^2V^{(2)}(x_0)+O(x-x_0)^3$$

Because $V(x_0)$ is a minimum, the first derivative evaluated at x_0 must be zero, so the linear term drops out:

$$V(x)=V(x_0)+\frac{1}{2}(x-x_0)^2V^{(2)}(x_0)+O(x-x_0)^3$$

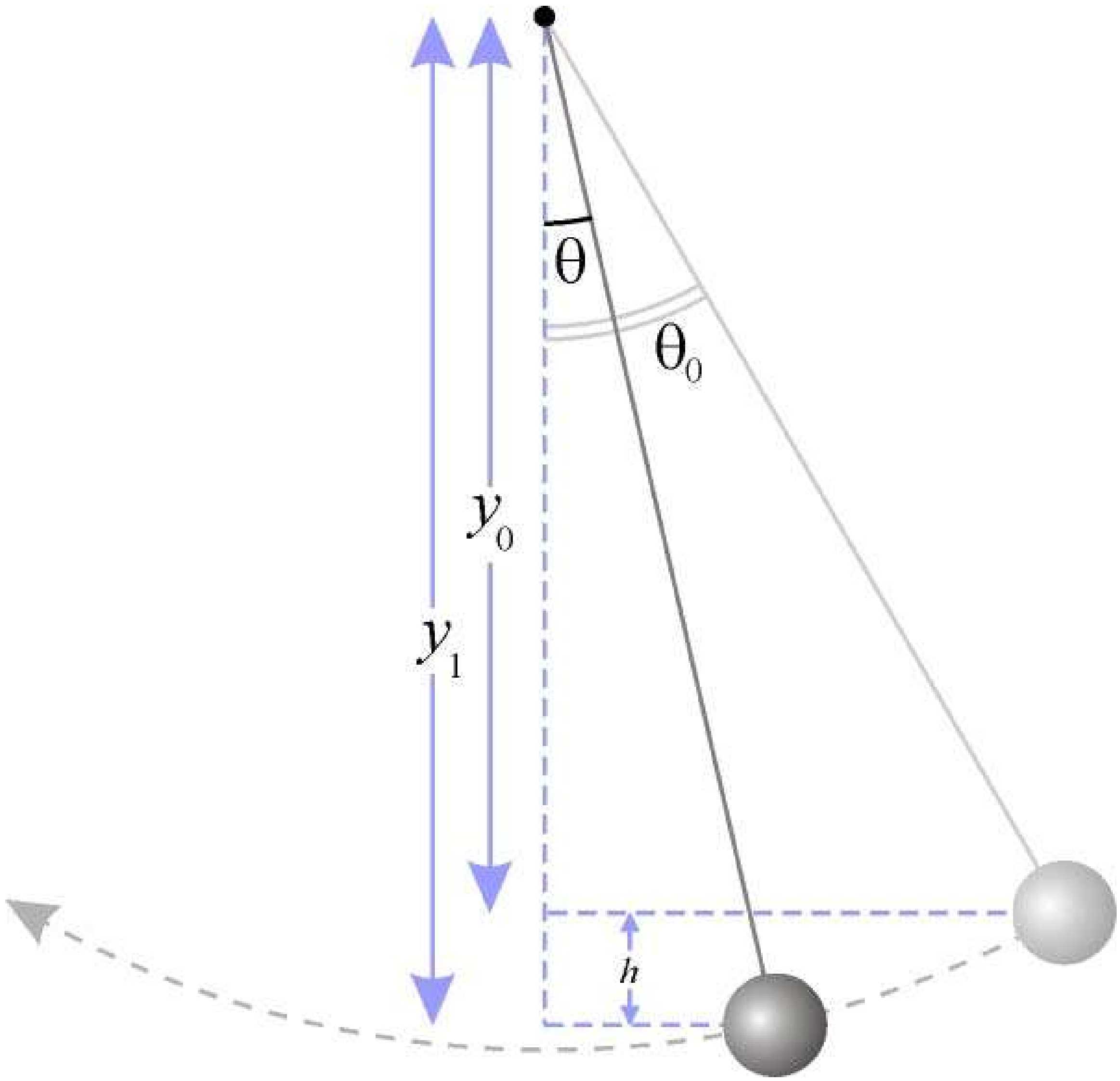
The constant term $V(x_0)$ is arbitrary and thus may be dropped, and a coordinate transformation allows the form of the simple harmonic oscillator to be retrieved:

$$V(x)\approx\frac{1}{2}x^2V^{(2)}(0)=\frac{1}{2}kx^2$$

Thus, given an arbitrary potential energy function $V(x)$ with a non-vanishing second derivative, one can use the solution to the simple harmonic oscillator to provide an approximate solution for small perturbations around the equilibrium point.

Examples

Simple pendulum



A simple pendulum exhibits simple harmonic motion under the conditions of no damping and small amplitude.

Assuming no damping and small amplitudes, the differential equation governing a simple pendulum is

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \frac{g}{\ell}\theta = 0.$$

The solution to this equation is given by:

$$\theta(t) = \theta_0 \cos\left(\sqrt{\frac{g}{\ell}}t\right) \qquad |\theta_0| \ll 1$$

where θ_0 is the largest angle attained by the pendulum. The period, the time for one complete oscillation, is given by 2π divided by whatever is multiplying the time in the argument of the cosine ($\sqrt{\frac{g}{\ell}}$ here).

$$T_0 = 2\pi\sqrt{\frac{\ell}{g}} \qquad |\theta_0| \ll 1.$$

Pendulum swinging over turntable

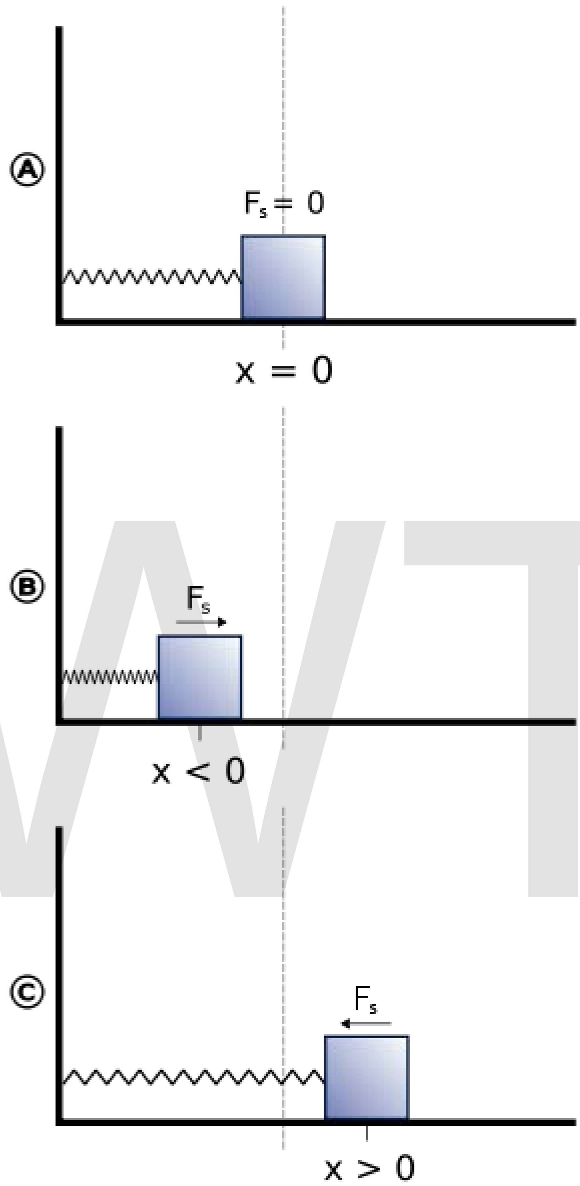
Simple harmonic motion can in some cases be considered to be the one-dimensional projection of two-dimensional circular motion. Consider a long pendulum swinging over the turntable of a record player. On the edge of the turntable there is an object. If the object is viewed from the same level as the turntable, a projection of the motion of the object seems to be moving backwards and forwards on a straight line. When the frequency of rotation of the turntable has perfect synchronization with the motion of the pendulum, the angular speed of the turntable is known as the pulsation of the pendulum. Note that the pulsation of the pendulum is not the angular speed of the pendulum itself, but it is the angular speed of the corresponding circular motion.

In general, the pulsation of straight-line simple harmonic motion is the angular speed of the corresponding circular motion. Mathematically, a motion with period T and frequency $f=1/T$ has pulsation

$$\omega = 2\pi f = \frac{2\pi}{T}.$$

WT

Spring–mass system



Spring–mass system in equilibrium (A), compressed (B) and stretched (C) states.

When a spring is stretched or compressed by a mass, the spring develops a restoring force. Hooke's law gives the relationship of the force exerted by the spring when the spring is compressed or stretched a certain length:

$$F(t) = -kx(t)$$

where F is the force, k is the spring constant, and x is the displacement of the mass with respect to the equilibrium position. This relationship shows that the distance of the spring is always opposite to the force of the spring.

By using either force balance or an energy method, it can be readily shown that the motion of this system is given by the following differential equation:

$$F(t) = -kx(t) = m\frac{\mathrm{d}^2}{\mathrm{d}t^2}x\left(t\right) = ma.$$

...the latter evidently being Newton's second law of motion.

If the initial displacement is A, and there is no initial velocity, the solution of this equation is given by:

$$x\left(t\right) = A\cos\left(\sqrt{\frac{k}{m}}t\right).$$

Given an ideal massless spring, m is the mass on the end of the spring. If the spring itself has mass, its effective mass must be included in m.

Energy variation in the spring–damping system

In terms of energy, all systems have two types of energy, potential energy and kinetic energy. When a spring is stretched or compressed, it stores elastic potential energy, which then is transferred into kinetic energy. The potential energy within a spring is determined by the equation $U = kx^2 / 2$.

When the spring is stretched or compressed, kinetic energy of the mass gets converted into potential energy of the spring. By conservation of energy, assuming the datum is defined at the equilibrium position, when the spring reaches its maximum potential energy, the kinetic energy of the mass is zero. When the spring is released, it tries to return to equilibrium, and all its potential energy converts to kinetic energy of the mass.

WT

Chapter- 3

Relaxation Oscillator

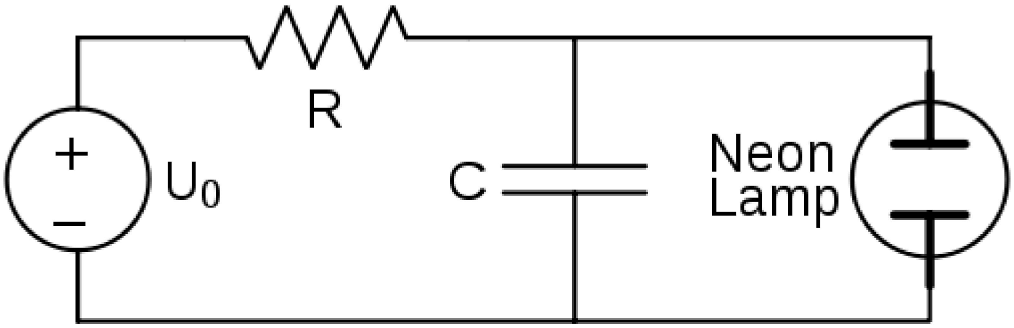
A **relaxation oscillator** is an oscillator based upon the behavior of a physical system's return to equilibrium after being disturbed. That is, a dynamical system within the oscillator continuously dissipates its internal energy. Normally the system would return to its natural equilibrium; however, each time the system reaches some threshold sufficiently close to its equilibrium, a mechanism disturbs it with additional energy. Hence, the oscillator's behavior is characterized by long periods of dissipation followed by short impulses. The period of the oscillations are set by the time it takes for the system to relax from each disturbed state to the threshold that triggers the next disturbance.

Implementation

Many electronic relaxation oscillators store energy in a capacitor and then dissipate that energy repeatedly to setup the oscillations. For example, the capacitor can be charged toward a positive power supply until it reaches a threshold voltage sufficiently close to the supply. At that instant, the capacitor can be quickly discharged (e.g., shorted). Alternatively, when the capacitor reaches each threshold, the charging source can be switched from the positive power supply to the negative power supply or vice versa. In all such capacitor-based relaxation oscillators, the period of the oscillations is set by the dissipation rate(s) of the capacitor. Implementations of these two types of relaxation oscillators are shown here, but relaxation oscillators need not be electronic in general. Any oscillator whose oscillations are driven by a system that almost always is dissipating energy may be called a relaxation oscillator.

WT

Pearson–Anson electronic relaxation oscillator



Circuit diagram of a capacitive relaxation oscillator with a neon lamp threshold device

This example can be implemented with a capacitive or resistive-capacitive integrating circuit driven respectively by a constant current or voltage source, and a threshold device with hysteresis (neon lamp, thyratron, diac or unijunction transistor) connected in parallel to the capacitor. The capacitor is charged by the input source causing the voltage across the capacitor to rise. The threshold device does not conduct at all until the capacitor voltage reaches its threshold (trigger) voltage. It then increases heavily its conductance in an avalanche-like manner because of the inherent positive feedback, which quickly discharges the capacitor. When the voltage across the capacitor drops to some lower threshold voltage, the device stops conducting and the capacitor begins charging again, and the cycle repeats ad infinitum.

If the threshold element is a neon lamp, the circuit also provides a flash of light with each discharge of the capacitor. This lamp example is depicted below in the typical circuit used to describe the Pearson–Anson effect. The discharging duration can be extended by connecting an additional resistor in series to the threshold element. The two resistors form a voltage divider; so, the additional resistor has to have low enough resistance to reach the low threshold.

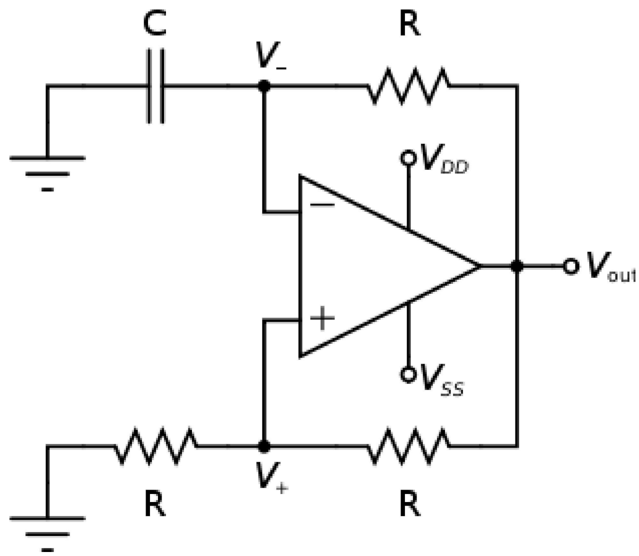
Alternative implementation with 555 timer

A similar relaxation oscillator can be built with a 555 timer (acting in astable mode) that takes the place of the neon bulb above. That is, when the capacitor reaches a certain value, comparators within the 555 timer cause the latch to activate a transistor switch that discharges the capacitor through a resistor to ground. At the instant the capacitor falls to a sufficiently low value, the switch deactivates to let the capacitor charge again.

Comparator–based electronic relaxation oscillator

Alternatively, when the capacitor reaches each threshold, the charging source can be switched from the positive power supply to the negative power supply or vice versa. This case is shown in the comparator-based implementation here.

WVT



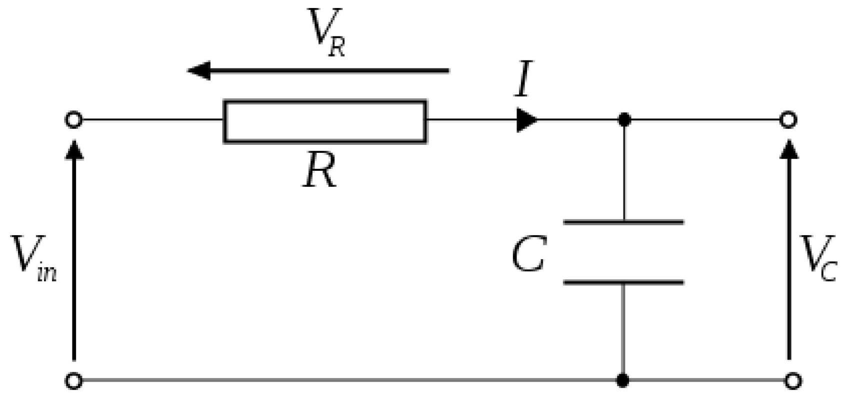
A comparator-based hysteretic oscillator.

This relaxation oscillator is a hysteretic oscillator, named this way because of the hysteresis created by the positive feedback loop implemented with the comparator (similar to, but different than, an op-amp). A circuit that implements this form of hysteretic switching is known as a Schmitt trigger. Alone, the trigger is a bistable multivibrator. However, the slow negative feedback added to the trigger by the RC circuit causes the circuit to oscillate automatically. That is, the addition of the RC circuit turns the hysteretic bistable multivibrator into an astable multivibrator.

General Concept

The system is in unstable equilibrium if both the inputs and outputs of the comparator are at zero volts. The moment any sort of noise, be it thermal or electromagnetic noise brings the output of the comparator above zero (the case of the comparator output going below zero is also possible, and a similar argument to what follows applies), the positive feedback in the comparator results in the output of the comparator saturating at the positive rail.

In other words, because the output of the comparator is now positive, the non-inverting input to the comparator is also positive, and continues to increase as the output increases, due to the voltage divider. After a short time, the output of the comparator is the positive voltage rail, V_{DD}.



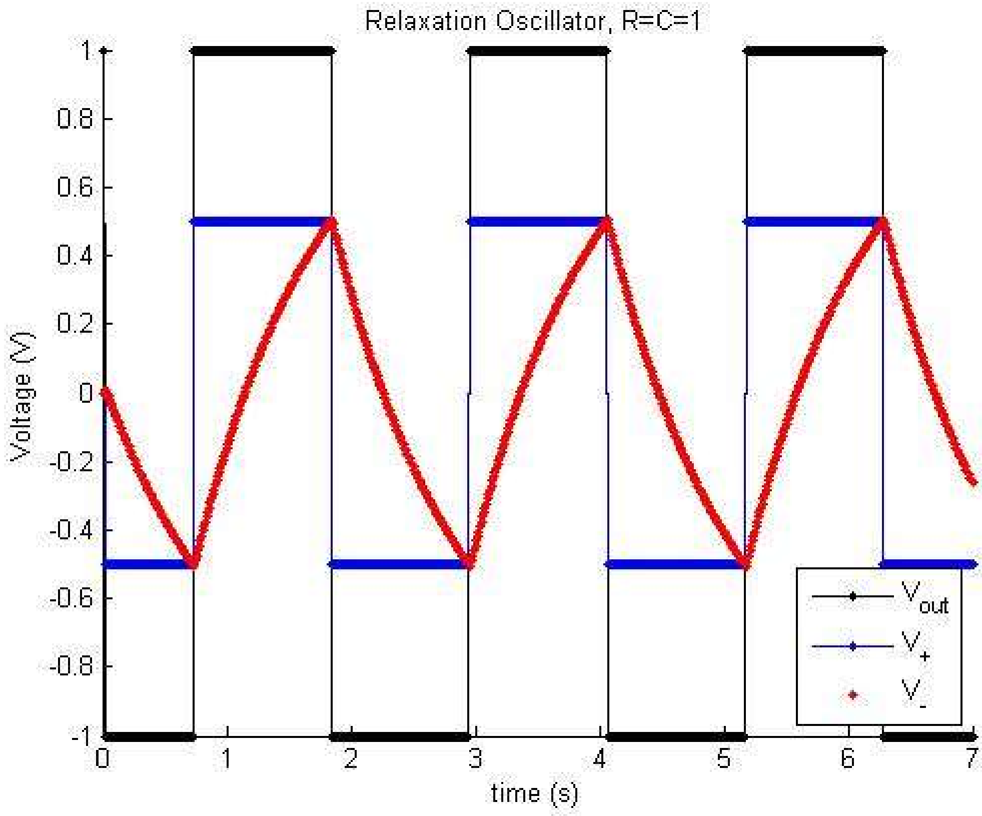
Series RC Circuit

The inverting input and the output of the comparator are linked by a series RC circuit. Because of this, the inverting input of the comparator asymptotically approaches the comparator output voltage with a time constant RC. At the point where voltage at the inverting input is greater than the non-inverting input, the output of the comparator falls quickly due to positive feedback.

This is because the non-inverting input is less than the inverting input, and as the output continues to decrease, the difference between the inputs gets more and more negative. Again, the inverting input approaches the comparator's output voltage asymptotically, and the cycle repeats itself once the non-inverting input is greater than the inverting input, hence the system oscillates.

WT

Example: Differential Equation Analysis of comparator-based Relaxation Oscillator



Transient analysis of an comparator-based relaxation oscillator.

$V_+$is set by V_{out}across a resistive voltage divider:

$$V_+ = \frac{V_{out}}{2}$$

$V_-$is obtained using Ohm's law and the capacitor differential equation:

$$\frac{V_{out} - V_-}{R} = C\frac{dV_-}{dt}$$

Rearranging the V_-differential equation into standard form results in the following:

$$\frac{dV_-}{dt} + \frac{V_-}{RC} = \frac{V_{out}}{RC}$$

Notice there are two solutions to the differential equation, the driven or particular solution and the homogeneous solution. Solving for the driven solution, observe that for this particular form, the solution is a constant. In other words, $V_- = A$ where A is a constant and $\frac{dV_-}{dt} = 0$.

$$\frac{A}{RC} = \frac{V_{out}}{RC}$$
$$A = V_{out}$$

Using the Laplace transform to solve the homogeneous equation $\frac{dV_-}{dt} + \frac{V_-}{RC} = 0$ results in

$$V_- = Be^{\frac{-1}{RC}t}$$

V_- is the sum of the particular and homogeneous solution.

$$V_- = A + Be^{\frac{-1}{RC}t}$$
$$V_- = V_{out} + Be^{\frac{-1}{RC}t}$$

Solving for B requires evaluation of the initial conditions. At time 0, $V_{out} = V_{dd}$ and $V_- = 0$. Substituting into our previous equation,

$$0 = V_{dd} + B$$
$$B = -V_{dd}$$

Frequency of Oscillation

First let's assume that $V_{dd} = -V_{ss}$ for ease of calculation. Ignoring the initial charge up of the capacitor, which is irrelevant for calculations of the frequency, note that charges and discharges oscillating between $\frac{V_{dd}}{2}$ and $\frac{V_{ss}}{2}$. For the circuit above, V_{ss} must be less than 0. Half of the period (T) is the same as time that V_{out} switches from V_{dd}. This occurs when V_- charges up from $-\frac{V_{dd}}{2}$ to $\frac{V_{dd}}{2}$.

$$V_- = A + Be^{\frac{-1}{RC}t}$$
$$\frac{V_{dd}}{2} = V_{dd}\left(1 - \frac{3}{2}e^{\frac{-1}{RC}\frac{T}{2}}\right)$$
$$\frac{1}{3} = e^{\frac{-1}{RC}\frac{T}{2}}$$

$$\ln\left(\frac{1}{3}\right) = \frac{-1}{RC}\frac{T}{2}$$
$$T = 2\ln(3)RC$$
$$f = \frac{1}{2\ln(3)RC}$$

When V_{ss} is not the inverse of V_{dd} we need to worry about asymmetric charge up and discharge times. Taking this into account we end up with a formula of the form:

$$T = (RC)[\ln(\frac{2V_{ss} - V_{dd}}{V_{ss}}) + \ln(\frac{2V_{dd} - V_{ss}}{V_{dd}})]$$

Which reduces to the above result in the case that $V_{dd} = -V_{ss}$.

Practical examples of the use of the relaxation oscillator

This type of circuit was used as the time base in early oscilloscopes and television receivers. Variants of this circuit find use in stroboscopes used in machine shops and nightclubs. Electronic camera flashes are a monostable version of this circuit, generating *one cycle* of the sawtooth, the rising edge as the flash capacitor is charged and the rapid falling edge as the capacitor is discharged and the flash is produced upon receiving the firing signal from the shutter button. Use as a timebase in oscilloscopes was discontinued when the much more linear Miller Integrator timebase circuit using "hard" valves, (vacuum tubes) as a constant current source, was developed.

WT

Chapter- 4

Multivibrator

A **multivibrator** is an electronic circuit used to implement a variety of simple two-state systems such as oscillators, timers and flip-flops. It is characterized by two amplifying devices (transistors, electron tubes or other devices) cross-coupled by resistors or capacitors. The name "multivibrator" was initially applied to the free-running oscillator version of the circuit because its output waveform was rich in harmonics. There are three types of multivibrator circuit depending on the circuit operation:

- **astable**, in which the circuit is not stable in either state —it continually switches from one state to the other. It does not require an input such as a clock pulse.
- **monostable**, in which one of the states is stable, but the other state is unstable (transient). A trigger causes the circuit to enter the unstable state. After entering the unstable state, the circuit will return to the stable state after a set time. Such a circuit is useful for creating a timing period of fixed duration in response to some external event. This circuit is also known as a **one shot**.
- **bistable**, in which the circuit is stable in either state. The circuit can be flipped from one state to the other by an external event or trigger.

Multivibrators find applications in a variety of systems where square waves or timed intervals are required. For example, before the advent of low-cost integrated circuits, chains of multivibrators found use as frequency dividers. A free-running multivibrator with a frequency of one-half to one-tenth of the reference frequency would accurately lock to the reference frequency. This technique was used in early electronic organs, to keep notes of different octaves accurately in tune. Other applications included early television systems, where the various line and frame frequencies were kept synchronized by pulses included in the video signal.

History

The classic multivibrator circuit (also called a *plate-coupled multivibrator*) is first described by H. Abraham and E. Bloch in *Publication 27* of the French *Ministère de la Guerre*, and in *Annales de Physique 12, 252 (1919)*. It is a predecessor of Eccles-Jordan trigger derived from this circuit a year later on.

Astable multivibrator

An astable multivibrator is a regenerative circuit consisting of two amplifying stages connected in a positive feedback loop by two capacitive-resistive coupling networks. The amplifying elements may be junction or field-effect transistors, vacuum tubes, operational amplifiers, or other types of amplifier. The example diagram shows bipolar junction transistors.

The circuit is usually drawn in a symmetric form as a cross-coupled pair. Two output terminals can be defined at the active devices, which will have complementary states; one will have high voltage while the other has low voltage, (except during the brief transitions from one state to the other).

Operation

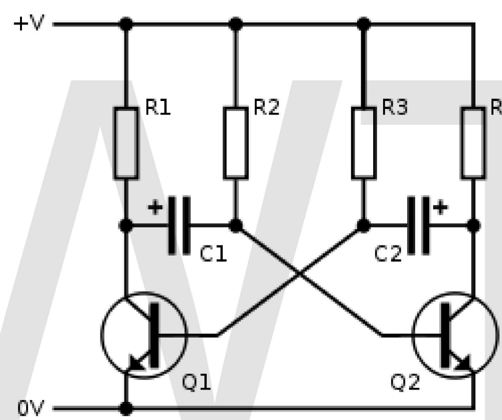


Figure 1: Basic BJT astable multivibrator

The circuit has two stable states that change alternatively with maximum transition rate because of the "accelerating" positive feedback. It is implemented by the coupling capacitors that instantly transfer voltage changes because the voltage across a capacitor cannot suddenly change. In each state, one transistor is switched on and the other is switched off. Accordingly, one fully charged capacitor discharges (reverse charges) slowly thus converting the time into an exponentially changing voltage. At the same time, the other empty capacitor quickly charges thus restoring its charge (the first capacitor acts as a time-setting capacitor and the second prepares to play this role in the next state). The circuit operation is based on the fact that the forward-biased base-emitter junction of the switched-on bipolar transistor can provide a path for the capacitor restoration.

State 1 (Q1 is switched on, Q2 is switched off):
In the beginning, the capacitor C1 is fully charged (in the previous State 2) to the power supply voltage *V* with the polarity shown in Figure 1. Q1 is *on* and connects the left-hand positive plate of C1 to ground. As its right-hand negative plate is connected to Q2 base, a maximum negative voltage (-*V*) is applied to Q2 base that keeps Q2 firmly *off*. C1 begins discharging (reverse charging) via the high-resistive base resistor R2, so that the voltage

of its right-hand plate (and at the base of Q2) is rising from below ground (-*V*) toward +*V*. As Q2 base-emitter junction is backward-biased, it does not impact on the exponential process (R2-C1 integrating network is unloaded). Simultaneously, C2 that is fully discharged and even slightly charged to 0.6 V (in the previous State 2) quickly charges via the low-resistive collector resistor R4 and Q1 forward-biased base-emitter junction (because R4 is less than R2, C2 charges faster than C1). Thus C2 restores its charge and prepares for the next State 2 when it will act as a time-setting capacitor. Q1 is firmly saturated in the beginning by the "forcing" C2 charging current added to R3 current; in the end, only R3 provides the needed input base current. The resistanse R3 is chosen small enough to keep Q1 (not deeply) saturated after C2 is fully charged.

When the voltage of C1 right-hand plate (Q2 base voltage) becomes positive and reaches 0.6 V, Q2 base-emitter junction begins diverting a part of R3 charging current. Q2 begins conducting and this starts the avalanche-like positive feedback process as follows. Q2 collector voltage begins falling; this change transfers through the fully charged C2 to Q1 base and Q1 begins cutting off. Its collector voltage begins rising; this change transfers back through the almost empty C1 to Q2 base and makes Q2 conduct more thus sustaining the initial input impact on Q2 base. Thus the initial input change circulates along the feedback loop and grows in an avalanche-like manner until finally Q1 switches off and Q2 switches on. The forward-biased Q2 base-emitter junction fixes the voltage of C1 right-hand plate at 0.6 V and does not allow it to continue rising toward +*V*.

State 2 (Q1 is switched off, Q2 is switched on):
Now, the capacitor C2 is fully charged (in the previous State 1) to the power supply voltage *V* with the polarity shown in Figure 1. Q2 is *on* and connects the right-hand positive plate of C2 to ground. As its left-hand negative plate is connected to Q1 base, a maximum negative voltage (-*V*) is applied to Q1 base that keeps Q1 firmly *off*. C2 begins discharging (reverse charging) via the high-resistive base resistor R3, so that the voltage of its left-hand plate (and at the base of Q1) is rising from below ground (-*V*) toward +*V*. Simultaneously, C1 that is fully discharged and even slightly charged to 0.6 V (in the previous State 1) quickly charges via the low-resistive collector resistor R1 and Q2 forward-biased base-emitter junction (because R1 is less than R3, C1 charges faster than C2). Thus C1 restores its charge and prepares for the next State 1 when it will act again as a time-setting capacitor...and so on... (the next explanations are a mirror copy of the second part of Step 1).

Multivibrator period (frequency)

Derivation

The duration of state 1 (low output) will be related to the time constant R_2C_1 as it depends on the charging of C1, and the duration of state 2 (high output) will be related to the time constant R_3C_2 as it depends on the charging of C2. Because they do not need to be the same, an asymmetric duty cycle is easily achieved. The extended content below contains a derivation of the multivibrator period (frequency).

Summary

The total period of oscillation is given by:

$T = t_1 + t_2 = \ln(2)R_2\, C_1 + \ln(2)R_3\, C_2$

$$f = \frac{1}{T} = \frac{1}{\ln(2)\cdot(R_2C_1 + R_3C_2)} \approx \frac{1}{0.693\cdot(R_2C_1 + R_3C_2)}$$

where...

- f is frequency in hertz.
- R_2 and R_3 are resistor values in ohms.
- C_1 and C_2 are capacitor values in farads.
- T is the period (In this case, the sum of two period durations).

For the special case where

- $t_1 = t_2$ (50% duty cycle)
- $R_2 = R_3$
- $C_1 = C_2$

$$f = \frac{1}{T} = \frac{1}{\ln(2)\cdot 2RC} \approx \frac{0.721}{RC}$$

The duration of each state also depends on the initial state of charge of the capacitor in question, and this in turn will depend on the amount of discharge during the previous state, which will also depend on the resistors used during discharge (R1 and R4) and also on the duration of the previous state, *etc*. The result is that when first powered up, the period will be quite long as the capacitors are initially fully discharged, but the period will quickly shorten and stabilise.

The period will also depend on any current drawn from the output and on the supply voltage.

Output pulse shape

The output voltage has a shape that approximates a square waveform. It is considered below for the transistor Q1.

During State 1, Q2 base-emitter junction is backward-biased and the capacitor C1 is "unhooked" from ground. The output voltage of the switched-on transistor Q1 changes rapidly from high to low since this low-resistive output is loaded by a high impedance load (the series connected capacitor C1 and the high-resistive base resistor R2).

During State 2, Q2 base-emitter junction is forward-biased and the capacitor C1 is "hooked" to ground. The output voltage of the switched-off transistor Q1 changes exponentially from low to high since this relatively high resistive output is loaded by a low impedance load (the capacitance C1). This is the output voltage of R_1C_1 integrating circuit.

To approach the needed square waveform, the collector resistors have to be low resistance. The base resistors have to be low enough to make the transistors saturate in the end of the restoration ($R_B < \beta . R_C$).

Initial power-up

When the circuit is first powered up, neither transistor will be switched on. However, this means that at this stage they will both have high base voltages and therefore a tendency to switch on, and inevitable slight asymmetries will mean that one of the transistors is first to switch on. This will quickly put the circuit into one of the above states, and oscillation will ensue. In practice, oscillation always occurs for practical values of *R* and *C*.

However, if the circuit is temporarily held with both bases high, for longer than it takes for both capacitors to charge fully, then the circuit will remain in this stable state, with both bases at 0.6 V, both collectors at 0 V, and both capacitors charged backwards to −0.6 V. This can occur at startup without external intervention, if *R* and *C* are both very small.

Frequency divider

An astable multivibrator can be synchronized to an external chain of pulses. A single pair of active devices can be used to divide a reference by a large ratio, however, the stability of the technique is poor owing to the variability of the power supply and the circuit elements; a division ratio of 10, for example, is easy to obtain but not dependable. Chains of bistable flip-flops provide more predictable division, at the cost of more active elements.

Protective components

While not fundamental to circuit operation, diodes connected in series with the base or emitter of the transistors are required to prevent the base-emitter junction being driven into reverse breakdown when the supply voltage is in excess of the V_{eb} breakdown voltage, typically around 5-10 volts for general purpose silicon transistors. In the monostable configuration, only one of the transistors requires protection.

Monostable multivibrator circuit

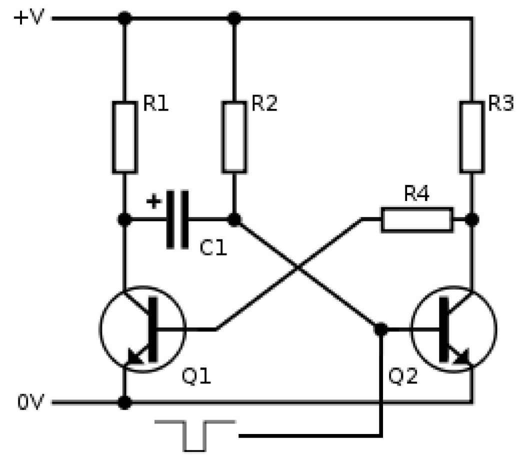


Figure 2: Basic BJT monostable multivibrator.

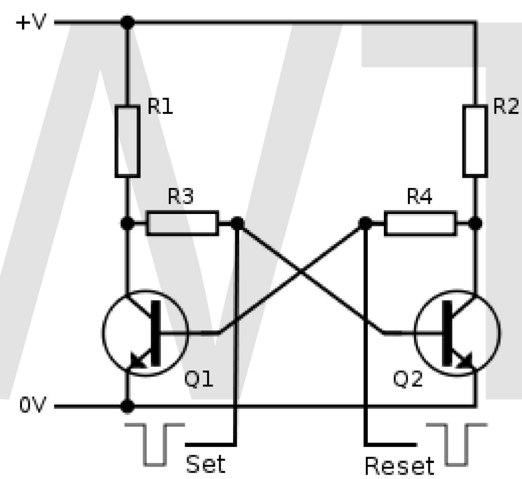


Figure 3: Basic BJT bistable multivibrator (suggested values: *R1, R2* = 1 kΩ *R3, R4* = 10 kΩ).

In the monostable multivibrator, the one resistive-capacitive network (C_2-R_3 in figure 1) is replaced by a resistive network (just a resistor). The circuit can be thought as a 1/2 astable multivibrator. Q2 collector voltage is the output of the circuit (in contrast to the astable circuit, it has a perfect square waveform since the output is not loaded by the capacitor).

When triggered by an input pulse, a monostable multivibrator will switch to its unstable position for a period of time, and then return to its stable state. The time period monostable multivibrator remains in unstable state is given by $t = \ln(2)R_2C_1$. If repeated application of the input pulse maintains the circuit in the unstable state, it is called a *retriggerable* monostable. If further trigger pulses do not affect the period, the circuit is a *non-retriggerable* multivibrator.

For the circuit in Figure 2, in the stable state Q1 is turned off and Q2 is turned on. It is triggered by zero or negative input signal applied to Q2 base (with the same success it can be triggered by applying a positive input signal through a resistor to Q1 base). As a result, the circuit goes in State 1 described above. After elapsing the time, it returns to its stable initial state.

Bistable multivibrator circuit

In the bistable multivibrator, both the resistive-capacitive network are replaced by resistive networks (just resistors or direct coupling).

This latch circuit is similar to an astable multivibrator, except that there is no charge or discharge time, due to the absence of capacitors. Hence, when the circuit is switched on, if Q1 is on, its collector is at 0 V. As a result, Q2 gets switched off. This results in more than half $+V$ volts being applied to R4 causing current into the base of Q1, thus keeping it on. Thus, the circuit remains stable in a single state continuously. Similarly, Q2 remains on continuously, if it happens to get switched on first.

Switching of state can be done via Set and Reset terminals connected to the bases. For example, if Q2 is on and Set is grounded momentarily, this switches Q2 off, and makes Q1 on. Thus, Set is used to "set" Q1 on, and Reset is used to "reset" it to off state.

WT

Chapter- 5

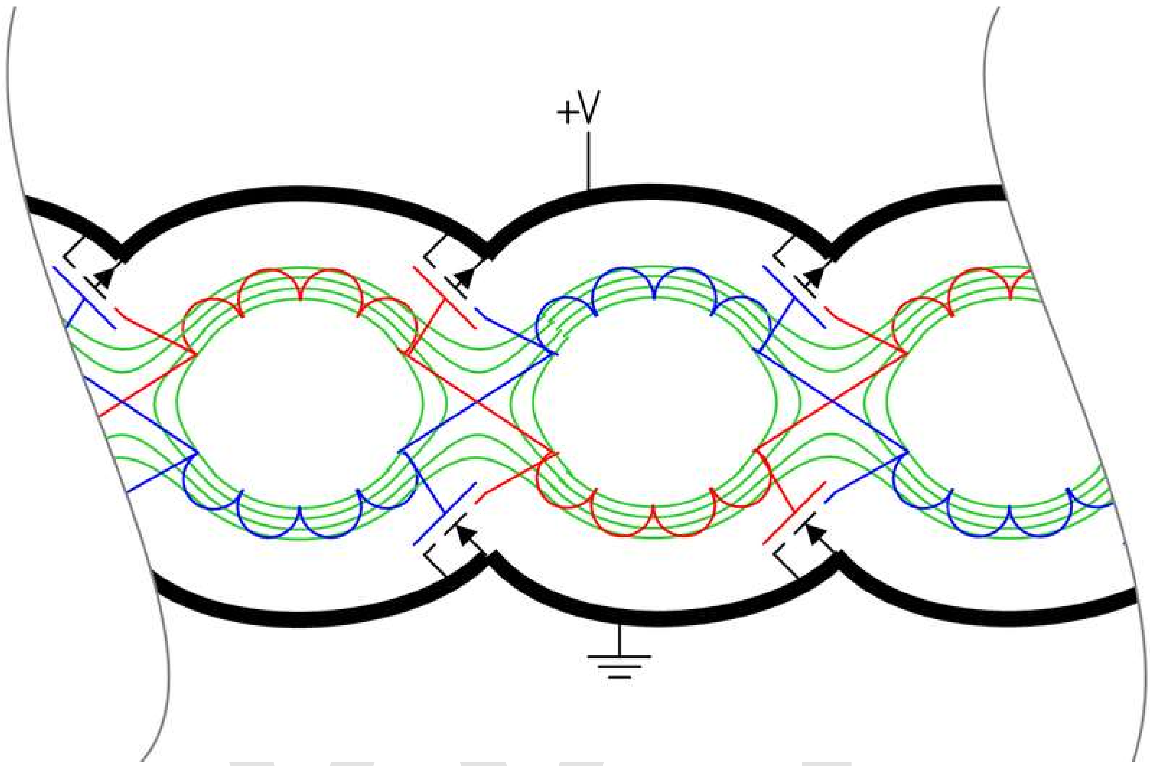
Time to Digital Converter and Ring Oscillator

Time to digital converter

In electronic instrumentation and signal processing, a **time to digital converter** (abbreviated **TDC**) is a device for converting a signal of sporadic pulses into a digital representation of their time indices. In other words, a TDC outputs the time of arrival for each incoming pulse. Because the magnitudes of the pulses are not usually measured, a TDC is used when the important information is to be found in the timing of events. In practice, a TDC usually follows a discriminator. TDCs are most often used in applications where measurement events happen infrequently, such as high energy physics experiments, where the sheer number of data channels in most detectors ensures that each channel will be excited only infrequently by particles such as electrons, photons, and ions.

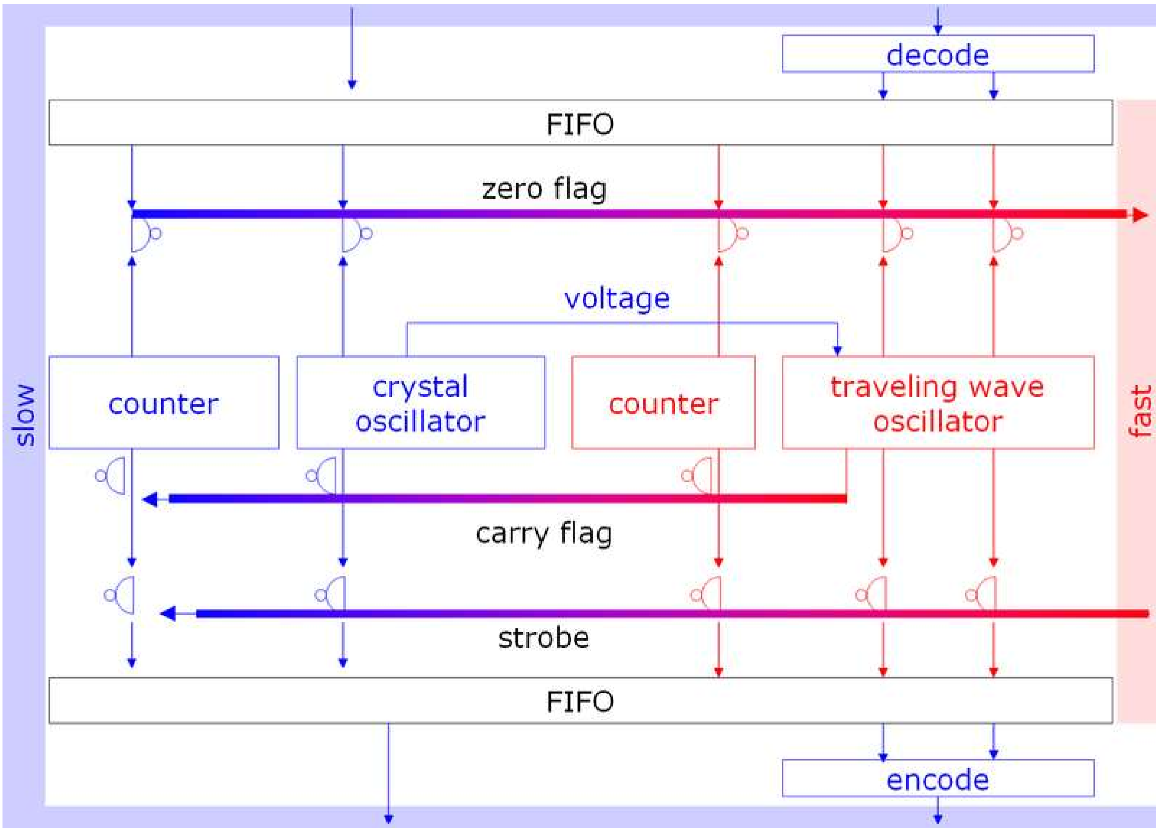
In its simplest implementation, a TDC is simply a high-frequency counter with a buffered output.

Implementation



A CMOS (rotary) traveling wave oscillator or delay line or distributed amplifier runs at a flip-flip compatible frequency, but has sharper edges and sub-edge resolution

VVI



Similarity between a **TDC** (bottom) and a **Delay Generator** (top, but needs bottom for trigger). The strobe is gated by the oscillator to avoid a race with the carry bit

Typically a TDC has a crystal oscillator, which has good long term stability, but oscillates too slowly (80 MHz in 2007). A faster overtone crystal oscillator or a SAW oscillator could be used for the reference, but the TDC also has a VCO for the ultimate frequency needed. To be fast this is a loop of transistors (series of logic gates, namely inverters). This also acts as a Johnson counter, which divides this frequency down to a slow frequency. A phase-locked loop is used to lock this low frequency to the frequency of the crystal oscillator. A slow synchronous counter, counts the slow oscillations.

For every stop pulse a copy of the timer is stored. Some TDCs start the counter after a start pulse and stop after a stop pulse. But this is incompatible with the TDC and allows only one stop pulse.

The speed of counters fabricated in CMOS-technology is limited by the capacity between the gate and the channel and by the resistance of the channel and the signal traces. The product of both is the cut-off-frequency. Modern chip technology allows multiple metal layers and therefore coils with a large number of windings to be inserted into the chip. This allows to peak the device for a specific frequency, which may lie well above the cut-off-frequency of the original transistor. The counter is then called a prescaler. The frequency of the voltage-controlled oscillator, for which the use of a coil is more common, has to be matched to the prescaler. A peaked variant of the Johnson counter is the traveling-wave counter which also achieves sub-cycle resolution. Other methods to achieve sub-cycle resolution include analog-to-digital converters and vernier Johnson counters.

Delay generator

This is a digital to time converter. Whereas the TDC measures the time between a start and a stop pulse, the **delay generator** gets a start pulse at its inputs, then counts down and outputs a stop pulse. For low jitter the synchronous counter has to feed a zero flag from the most significant bit down to the least significant bit and then combine it with the output from the Johnson counter.

A digital-to-analog converter (DAC) could be used to achieve sub-cycle resolution, but it is easier to either use vernier Johnson counters or traveling-wave Johnson counters.

The delay generator can be used for pulse width modulation, e.g. to drive a MOSFET to load a Pockels cell within 8 ns with a specific charge.

The output of a delay generator can gate a digital-to-analog converter and so pulses of a variable height can be generated. This allows matching to low levels needed by analog electronics, higher levels for ECL and even higher levels for TTL. If a series of DACs is gated in sequence, variable pulse shapes can be generated to account for any transfer function.

Discriminator

Even if the input is binary, after sampling with the oscillator signal the signal is analog (due to the finite edge width). Therefore an analog-to-digital converter is always employed after the sampler, the same is true for a simple comparator which can be implemented as a cascade of differential amplifiers, where the latter stages are driven into saturation, that means either into the low or the high state (1-bit ADC). At every leading edge of the time-discrete and voltage-discrete signal the time is fed into the FIFO.

Constant fraction discriminator

As a TDC is mainly a logic device and has problems with voltages between low and high state. A constant fraction discriminator differentiates a copy of the input signal. A look at Fourier transformation shows, that this can be accomplished by a 90° phase shift for all frequency components. Omitting the $1/f$ has some benefits. This signal is fed into a second comparator to get the maximum. At the end the signals of both comparators is send into an AND gate. A constant fraction discriminator reduces jitter when the single particles produce short pulses of different height and the measurement device blurs them into long pulses with a constant shape.

Utilities

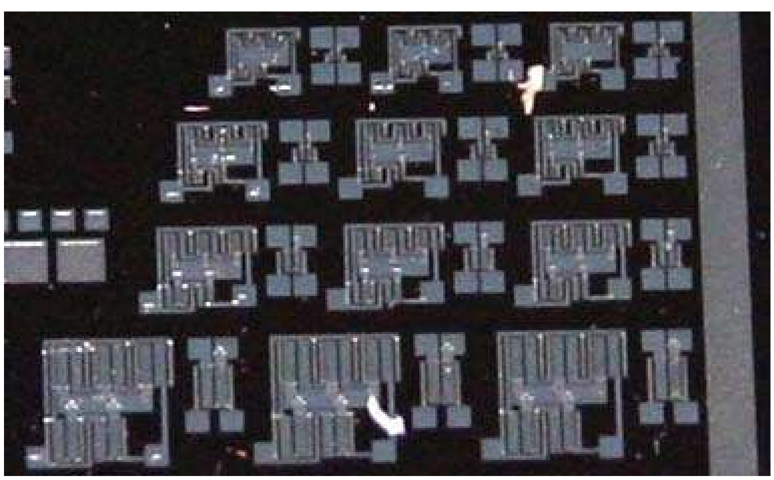
A flip flop can be used to convert pulses into edges and vice versa.

A gate has an analog and a logic input. When the logic input is low, the output is zero. A typical sampling pulse of 1 ns width and a 1Msamples ADC and a signal-to-noise ratio of 100 means that the analog signal has to be suppressed by a factor of 100000 (low leakage). If the gate is used in conjunction with an integrator, the transmission should be constant over the integration interval within 0.001.

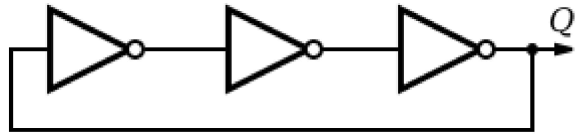
In a sampling oscilloscope an edge of a signal is sharpened by a diode pair, then fed through a delay generator, then converted into a pulse, and then gates the signal. The output of the gate is held in a capacitor and converted by an ADC. A multi-channel analyzer derives the gate pulse from a const-fraction discriminator and uses a fixed delay for a third copy of the original signal to account for the delay in the discriminator. If a series of gates is opened in a short sequence, the pulse shape can be sampled. No const-fraction discriminator is needed then and the system is very flexible concerning the pulse shape.

Often the sampled charge could be stored in the capacitor formed by the gate and channel of a MOSFET. The any current drawn by the ADC will not reduce the charge. Furthermore by drawing the current a long time a large charge amplification is possible. This is called a sample and hold circuit or also a track and hold circuit. The binary variant is called a buffer.

Ring oscillator



Ring oscillators fabricated on silicon using p-type MOSFETs.



A schematic of a simple 3-inverter ring oscillator whose output frequency is 1/(6×inverter delay).

A **ring oscillator** is a device composed of an odd number of NOT gates whose output oscillates between two voltage levels, representing *true* and *false*. The NOT gates, or inverters, are attached in a chain; the output of the last inverter is fed back into the first.

Details

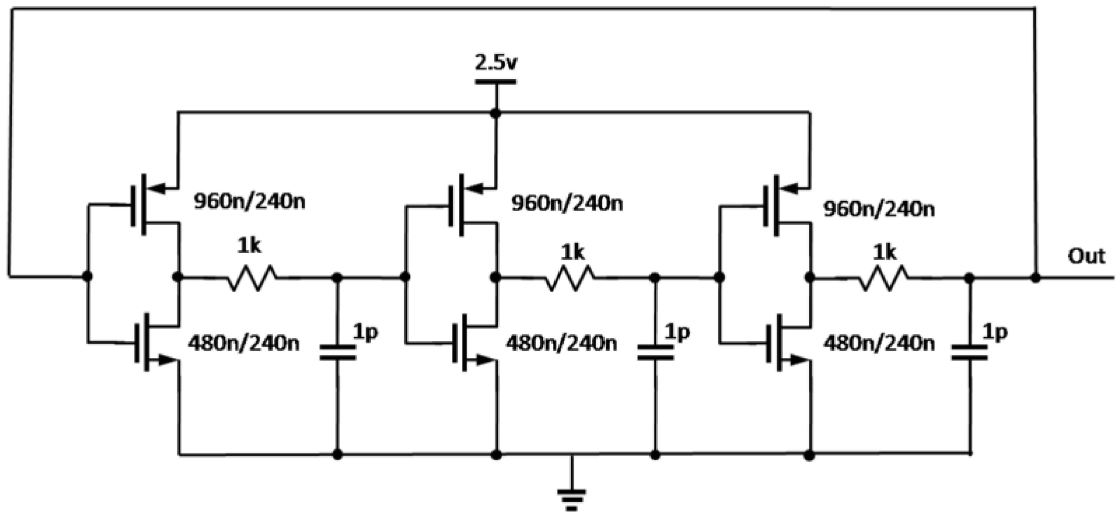
Because a single inverter computes the logical NOT of its input, it can be shown that the last output of a chain of an odd number of inverters is the logical NOT of the first input. This final output is asserted a finite amount of time after the first input is asserted; the feedback of this last output to the input causes oscillation.

A circular chain composed of an even number of inverters cannot be used as a ring oscillator; the last output in this case is the same as the input. However, this configuration of inverter feedback can be used as a storage element; it is the basic building block of static random access memory, or SRAM.

A real ring oscillator only requires power to operate; above a certain threshold voltage, oscillations begin spontaneously. To increase the frequency of oscillation, two methods may be used. Firstly, the applied voltage may be increased; this increases both the frequency of the oscillation and the power consumed, which is dissipated as heat. The heat dissipated limits the speed of a given oscillator. Secondly, a smaller ring oscillator may be fabricated; this results in a higher frequency of oscillation given a certain power consumption.

Operation

To understand the operation of a ring oscillator, one must first understand gate delay. In a physical device, no gate can switch instantaneously; in a device fabricated with MOSFETs, for example, the gate capacitance must be charged before current can flow between the source and the drain. Thus, the output of every inverter of a ring oscillator changes a finite amount of time after the input has changed. From here, it can be easily seen that adding more inverters to the chain increases the total gate delay, reducing the frequency of oscillation.



A transistor level schematic of a three-stage ring oscillator with delay in a .25u CMOS process.

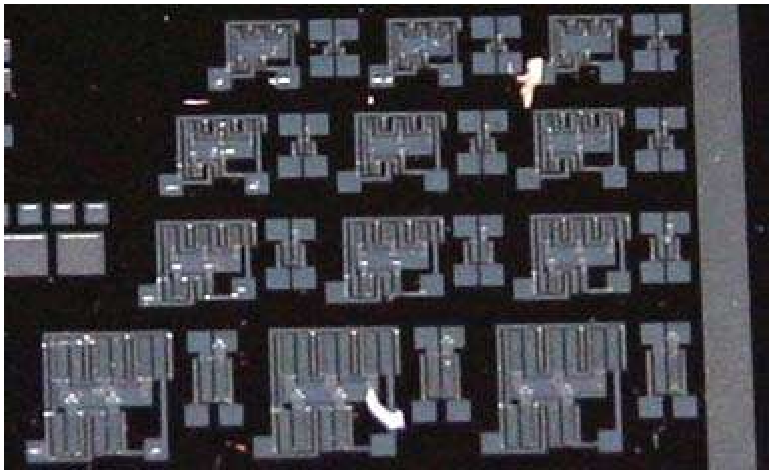
The ring oscillator is a member of the class of time delay oscillators. A time-delay oscillator consists of an inverting amplifier with a delay element between the amplifier output and its input. The amplifier must have a gain of greater than 1.0 at the intended oscillation frequency. Consider the initial case where the amplifier input and output voltages are momentarily balanced at a stable point. A small amount of noise can cause the amplifier output to rise slightly. After passing through the time-delay element, this small output voltage change will be presented to the amplifier input. The amplifier has a negative gain of greater than 1, so the output will change in the direction opposite to this input voltage. It will change by an amount larger than the input value, for a gain of greater than 1. This amplified, reversed signal propagates from the output through the time-delay and back to the input, where it is amplified and inverted again. The result of this sequential loop is a square-wave signal at the amplifier output with the period of each half of the square wave equal to the time delay. The square wave will grow until the amplifier output voltage reaches its limits, where it will stabilize. A more exact analysis will show that the wave that grows from the initial noise may not be square as it grows, but it will become square as the amplifier reaches its output limits.

The ring oscillator is a distributed version of the delay oscillator. The ring oscillator uses an odd number of inverters to give the effect of a single inverting amplifier with a gain of greater than one. Rather than having a single delay element, each inverter contributes to the delay of the signal around the ring of inverters, hence the name ring oscillator. Adding pairs of inverters to the ring increases the total delay and thereby decreases the oscillator frequency. Changing the supply voltage changes the delay through each inverter, with higher voltages typically decreasing the delay and increasing the oscillator frequency.

Applications

- The voltage-controlled oscillator in most phase-locked loops is built from a ring oscillator.
- A ring oscillator is often used to demonstrate a new hardware technology, analogous to the way a hello world program is often used to demonstrate a new software technology.
- Many wafers include a ring oscillator as part of the scribe line test structures. They are used during wafer testing to measure the effects of manufacturing process variations.
- Ring oscillators can also be used to measure the effects of voltage and temperature on a chip.

Oscylator typu ring



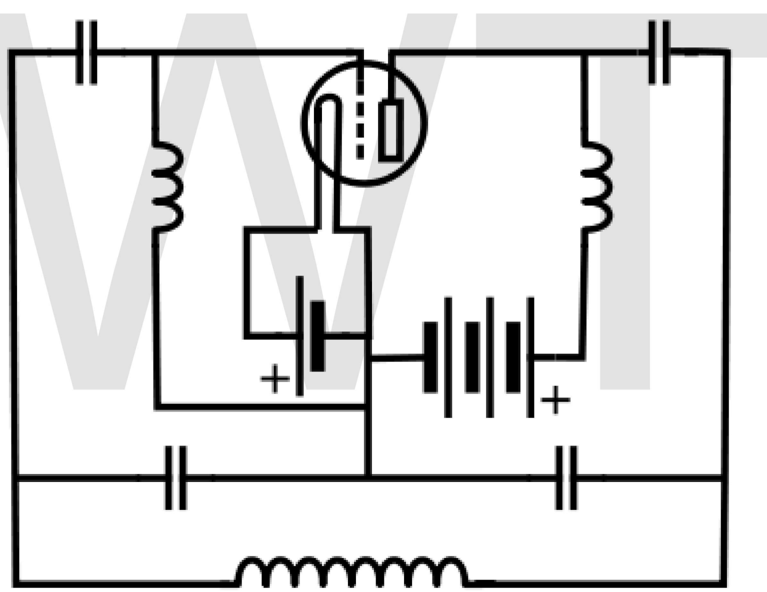
Pmos ring oscillator

WT

Chapter- 6

Colpitts Oscillator and Clapp Oscillator

Colpitts oscillator



Historic schematic using a vacuum tube

A **Colpitts oscillator**, named after its inventor Edwin H. Colpitts, is one of a number of designs for electronic oscillator circuits using the combination of an inductance (L) with a capacitor (C) for frequency determination, thus also called LC oscillator. The distinguishing feature of the Colpitts circuit is that the feedback signal is taken from a voltage divider made by two capacitors in series. One of the key features of this type of oscillator is its simplicity (needs only a single inductor) and robustness. The left picture shows the schematic as used in the first publication. Colpitts obtained US Patent 1624537 for this circuit.

The frequency is generally determined by the inductor and the two capacitors at the bottom of the drawing.

Implementation

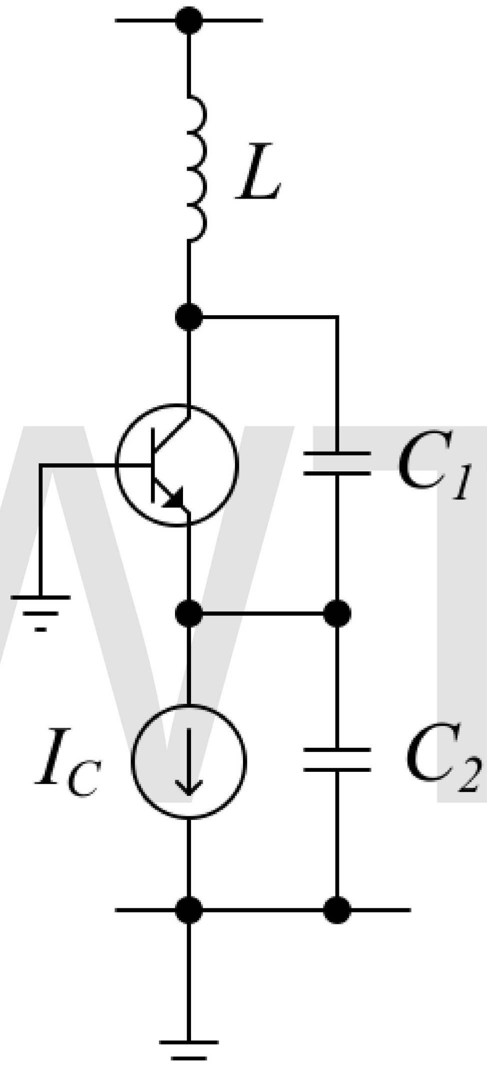


Figure 1: Simple common base Colpitts oscillator (with simplified biasing)

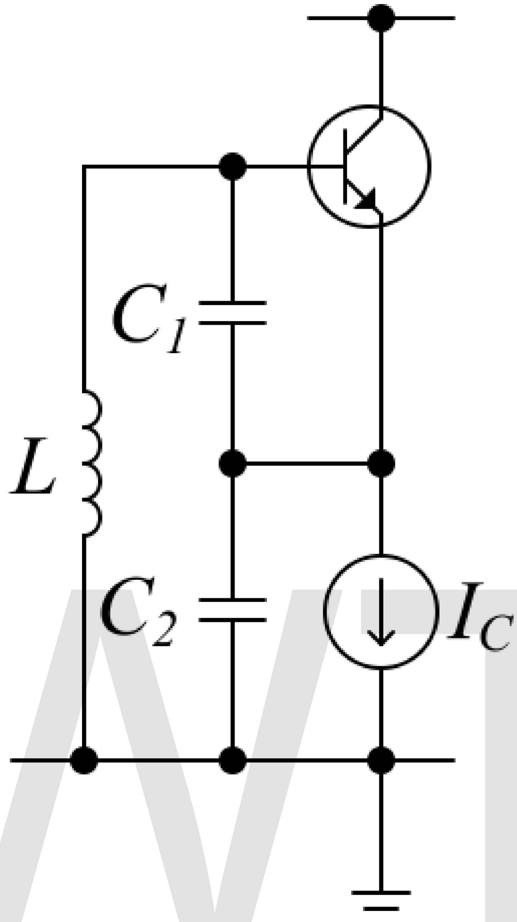


Figure 2: Simple common collector Colpitts oscillator (with simplified biasing)

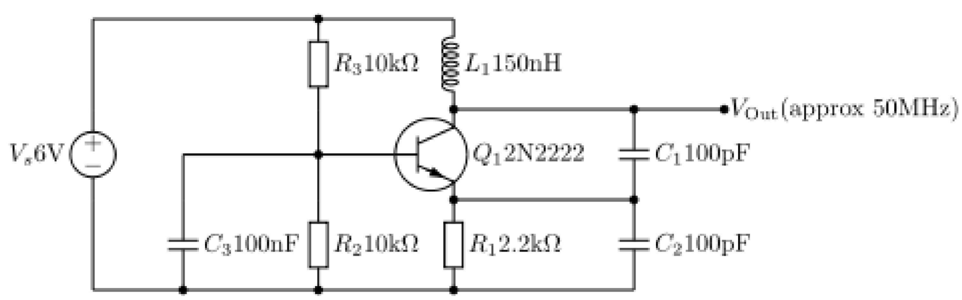


Figure 3: Practical common base Colpitts oscillator (with an oscillation frequency of ~50 MHz)

A Colpitts oscillator is the electrical dual of a Hartley oscillator. Fig. 1 shows the basic Colpitts circuit, where two capacitors and one inductor determine the frequency of oscillation. The feedback needed for oscillation is taken from a voltage divider made of two capacitors, whereas in the Hartley oscillator the feedback is taken from a voltage divider made of two inductors (or a single, tapped inductor).

As with any oscillator, the amplification of the active component should be marginally larger than the attenuation of the capacitive voltage divider, to obtain stable operation. Thus, a Colpitts oscillator used as a variable frequency oscillator (VFO) performs best when a variable inductance is used for tuning, as opposed to tuning one of the two capacitors. If tuning by variable capacitor is needed, it should be done via a third capacitor connected in parallel to the inductor (or in series as in the Clapp oscillator).

Fig. 2 shows an often preferred variant, where the inductor is also grounded (which makes circuit layout easier for higher frequencies). Note that feedback energy is fed into the connection between the two capacitors. This amplifier provides current, not voltage, amplification.

Fig. 3 shows a working example with component values. Instead of bipolar junction transistors, other active components such as field effect transistors or vacuum tubes, capable of producing gain at the desired frequency, could be used.

Theory

Oscillation frequency

The ideal frequency of oscillation for the circuits in Figures 1 and 2 are given by the equation:

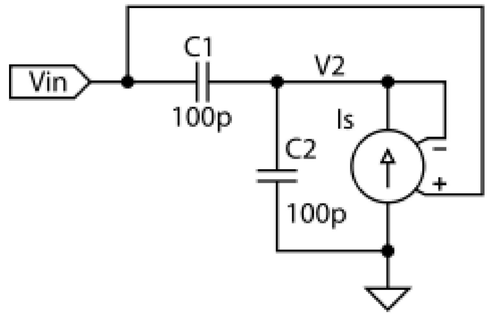
$$f_0 = \frac{1}{2\pi\sqrt{L \cdot \left(\frac{C_1 \cdot C_2}{C_1 + C_2}\right)}}$$

where the series combination of C1 and C2 creates the effective capacitance of the LC tank.

Real circuits will oscillate at a slightly lower frequency due to junction capacitances of the transistor and possibly other stray capacitances.

WT

Instability criteria



Colpitts oscillator model used in analysis at left.

One method of oscillator analysis is to determine the input impedance of an input port neglecting any reactive components. If the impedance yields a negative resistance term, oscillation is possible. This method will be used here to determine conditions of oscillation and the frequency of oscillation.

An ideal model is shown to the right. This configuration models the common collector circuit in the section above. For initial analysis, parasitic elements and device non-linearities will be ignored. These terms can be included later in a more rigorous analysis. Even with these approximations, acceptable comparison with experimental results is possible.

Ignoring the inductor, the input impedance can be written as

$$Z_{in} = \frac{v_1}{i_1}$$

Where v_1 is the input voltage and i_1 is the input current. The voltage v_2 is given by

$$v_2 = i_2 Z_2$$

Where Z_2 is the impedance of C_2. The current flowing into C_2 is i_2, which is the sum of two currents:

$$i_2 = i_1 + i_s$$

Where i_s is the current supplied by the transistor. i_s is a dependent current source given by

$$i_s = g_m \left(v_1 - v_2\right)$$

Where g_m is the transconductance of the transistor. The input current i_1 is given by

$$i_1 = \frac{v_1 - v_2}{Z_1}$$

Where Z_1 is the impedance of C_1. Solving for v_2 and substituting above yields

$$Z_{in} = Z_1 + Z_2 + g_m Z_1 Z_2$$

The input impedance appears as the two capacitors in series with an interesting term, R_{in} which is proportional to the product of the two impedances:

$$R_{in} = g_m \cdot Z_1 \cdot Z_2$$

If Z_1 and Z_2 are complex and of the same sign, R_{in} will be a negative resistance. If the impedances for Z_1 and Z_2 are substituted, R_{in} is

$$R_{in} = \frac{-g_m}{\omega^2 C_1 C_2}$$

If an inductor is connected to the input, the circuit will oscillate if the magnitude of the negative resistance is greater than the resistance of the inductor and any stray elements. The frequency of oscillation is as given in the previous section.

For the example oscillator above, the emitter current is roughly 1 mA. The transconductance is roughly 40 mS. Given all other values, the input resistance is roughly

$$R_{in} = -30\ \Omega$$

This value should be sufficient to overcome any positive resistance in the circuit. By inspection, oscillation is more likely for larger values of transconductance and/or smaller values of capacitance. A more complicated analysis of the common-base oscillator reveals that a low frequency amplifier voltage gain must be at least four to achieve oscillation. The low frequency gain is given by:

$$A_v = g_m \cdot R_p \geq 4$$

If the two capacitors are replaced by inductors and magnetic coupling is ignored, the circuit becomes a Hartley oscillator. In that case, the input impedance is the sum of the two inductors and a negative resistance given by:

$$R_{in} = -g_m \omega^2 L_1 L_2$$

In the Hartley circuit, oscillation is more likely for larger values of transconductance and/or larger values of inductance.

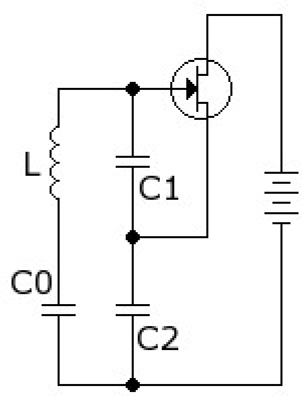
Oscillation amplitude

The amplitude of oscillation is generally difficult to predict, but it can often be accurately estimated using the describing function method.

Clapp oscillator

The **Clapp oscillator** is one of several types of electronic oscillator constructed from a transistor (or vacuum tube) and a positive feedback network, using the combination of an inductance (L) with a capacitor (C) for frequency determination, thus also called LC oscillator.

It was published by James Kilton Clapp in 1948. According to Vačkář, oscillators of this kind were independently developed by several inventors, and one developed by Gouriet had been in operation at the BBC since 1938.



Clapp Oscillator (direct-current biasing network not shown)

Referring to the notional circuit in the figure, the network comprises a single inductor and three capacitors. Capacitors C1 and C2 form a voltage divider that determines the amount of feedback voltage applied to the transistor input. The Clapp oscillator is a Colpitts oscillator that has an additional capacitor placed in series with the inductor. The oscillation frequency in hertz (cycles per second) for the circuit in the figure, which uses a field-effect transistor (FET), is

$$f_0 = \frac{1}{2\pi}\sqrt{\frac{1}{L}\left(\frac{1}{C_0}+\frac{1}{C_1}+\frac{1}{C_2}\right)}\ .$$

A Clapp circuit is often preferred over a Colpitts circuit for constructing a variable frequency oscillator (VFO). In a Colpitts VFO, the voltage divider contains the variable capacitor (either C1 or C2). This causes the feedback voltage to be variable as well, sometimes making the Colpitts circuit less likely to achieve oscillation over a portion of the desired frequency range. This problem is avoided in the Clapp circuit by using fixed capacitors in the voltage divider and a variable capacitor (C0) in series with the inductor.

WT

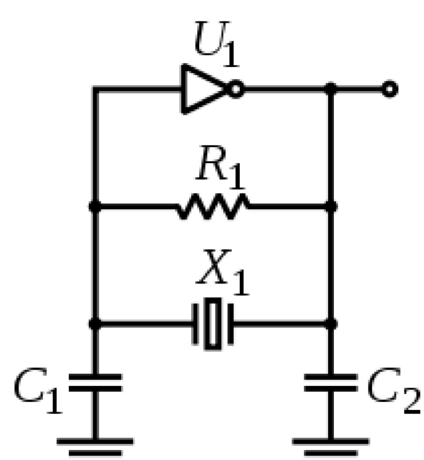
Chapter- 7

Pierce Oscillator, Phase-shift Oscillator, Hartley Oscillator and RC Oscillator

Pierce oscillator

The **Pierce oscillator** is a type of electronic oscillator particularly well-suited for use in piezoelectric crystal oscillator circuits. Named for its inventor, George W. Pierce (1872-1956), the Pierce oscillator is a derivative of the Colpitts oscillator. Virtually all digital IC clock oscillators are of Pierce type, as the circuit can be implemented using a minimum of components: a single digital inverter, two resistors, two capacitors, and the quartz crystal, which acts as a highly selective filter element. The low manufacturing cost of this circuit, and the outstanding frequency stability of the quartz crystal, give it an advantage over other designs in many consumer electronics applications.

Operation



Simple Pierce oscillator

Biasing resistor

R_1 acts as a feedback resistor, biasing the inverter in its linear region of operation and effectively causing it to function as a high gain inverting amplifier. To see this, assume the inverter is ideal, with infinite input impedance and zero output impedance. The resistor forces the input and output voltages to be equal. Hence the inverter will neither be fully on nor fully off, but will operate in the transition region where it has gain.

Resonator

The crystal in combination with C_1 and C_2 forms a pi network band-pass filter, which provides a 180 degree phase shift and a voltage gain from the output to input at approximately the resonant frequency of the crystal. To understand the operation, note that at the frequency of oscillation, the crystal appears inductive. Thus, it can be considered a large, high Q inductor. The combination of the 180 degree phase shift (i.e. inverting gain) from the pi network, and the negative gain from the inverter, results in a positive loop gain (positive feedback), making the bias point set by R_1 unstable and leading to oscillation.

Isolation resistor

Ruan Lourens strongly recommends a series resistor Rs between the output of the inverter and the crystal. The series resistor Rs reduces the chance of overtone oscillation and can improve start-up time. A second resistor could be used between the output of the inverter and the crystal to isolate the inverter from the crystal network. This would also add additional phase shift to C_1.

Load capacitance

The total capacitance seen from the crystal looking into the rest of the circuit is called the "load capacitance". When a manufacturer makes a "parallel" crystal, a technician adjusts a Pierce oscillator with a variable capacitor (often 18 or 20 pF) to trim the crystal to oscillate at exactly the frequency written on its package.

To assure operation at the correct frequency, one must make sure the capacitances in the circuit match the value specified on the crystal's data sheet. Load capacitance C_L can be calculated from the series combination of C_1 and C_2, taking into account C_i and C_o, the input and output capacitance of the inverter, and C_s, the stray capacitances from the oscillator, PCB layout, and crystal case (typically 3-9 pF):

$$C_L = \frac{[C_1 + C_i] \times [C_2 + C_o]}{[C_1 + C_i + C_2 + C_o]} + C_S$$

WT

When a "series" crystal is used in a Pierce oscillator, the Pierce oscillator (as always) drives the crystal at nearly its parallel resonance frequency. But that frequency is few kilohertz higher than the series resonant frequency printed on the package of a "series" crystal. Increasing the "load capacitance" slightly decreases the frequency generated by a Pierce oscillator, but never enough to reduce it all the way down to the series resonant frequency.

Phase-shift oscillator

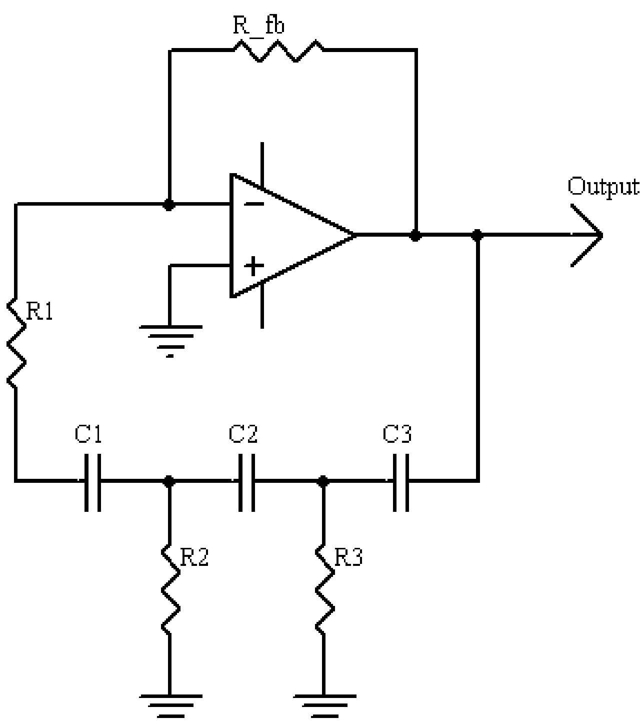
A **phase-shift oscillator** is a simple electronic oscillator. It contains an inverting amplifier, and a feedback filter which 'shifts' the phase of the amplifier output by 180 degrees at the oscillation frequency.

The filter produces a phase shift that increases with frequency. It must have a maximum phase shift of considerably greater than 180° at high frequencies, so that the phase shift at the desired oscillation frequency is 180°.

The most common way of achieving this kind of filter is using three identical cascaded resistor-capacitor filters, which together produce a phase shift of zero at low frequencies, and 270 degrees at high frequencies. At the oscillation frequency each filter produces a phase shift of 60 degrees and the whole filter circuit produces a phase shift of 180 degrees.

WT

Op-amp implementation



A simple example of a phase-shift oscillator

One of the simplest implementations for this type of oscillator uses an operational amplifier (op-amp), three capacitors and four resistors, as shown in the diagram.

The mathematics for calculating the oscillation frequency and oscillation criterion for this circuit are surprisingly complex, due to each R-C stage loading the previous ones. The calculations are greatly simplified by setting all the resistors (except the negative feedback resistor) and all the capacitors to the same values. In the diagram, if R1 = R2 = R3 = *R*, and C1 = C2 = C3 = *C*, then:

$$f_{\text{oscillation}} = \frac{1}{2\pi RC\sqrt{6}}$$

and the oscillation criterion is:

$R_{\text{feedback}} = 29(R)$

Without the simplification of all the resistors and capacitors having the same values, the calculations become more complex:

$$f_{\text{oscillation}} = \frac{1}{2\pi\sqrt{R_2R_3(C_1C_2+C_1C_3+C_2C_3)+R_1R_3(C_1C_2+C_1C_3)+R_1R_2C_1C_2}}$$

Oscillation criterion:

$$R_{\text{feedback}} = 2(R_1+R_2+R_3)+\frac{2R_1R_3}{R_2}+\frac{C_2R_2+C_2R_3+C_3R_3}{C_1}$$
$$+\frac{2C_1R_1+C_1R_2+C_3R_3}{C_2}+\frac{2C_1R_1+2C_2R_1+C_1R_2+C_2R_2+C_2R_3}{C_3}$$
$$+\frac{C_1R_1^2+C_3R_1R_3}{C_2R_2}+\frac{C_2R_1R_3+C_1R_1^2}{C_3R_2}+\frac{C_1R_1^2+C_1R_1R_2+C_2R_1R_2}{C_3R_3}$$

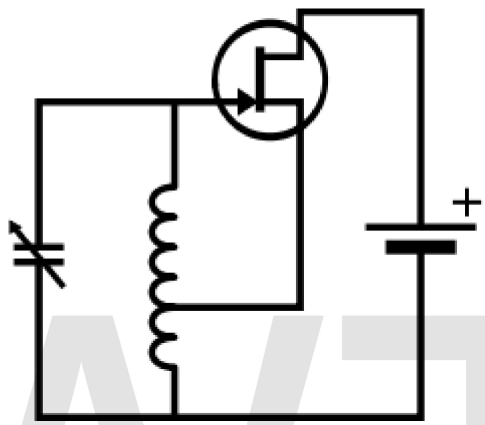
A version of this circuit can be made by putting an op-amp buffer between each R-C stage which simplifies the calculations. The voltage gain of the inverting channel is always unity.

When the oscillation frequency is high enough to be near the amplifier's cutoff frequency, the amplifier will contribute significant phase shift itself, which will add to the phase shift of the feedback network. Therefore the circuit will oscillate at a frequency at which the phase shift of the feedback filter is less than 180 degrees.

Hartley oscillator

The **Hartley oscillator** is an electronic oscillator circuit that uses an inductor and a capacitor in parallel to determine the frequency. Invented in 1915 by American engineer Ralph Hartley, the distinguishing feature of the Hartley circuit is that the feedback needed for oscillation is taken from a tap on the coil, or the junction of two coils in series.

Operation



Schematic diagram

A Hartley oscillator is essentially any configuration that uses two series-connected coils and a single capacitor. Although there is no requirement for there to be mutual coupling between the two coil segments, the circuit is usually implemented this way.

It is made up of the following:

- Two inductors in series, which need not be mutual
- One tuning capacitor

Advantages of the Hartley oscillator include:

- The frequency may be adjusted using a single variable capacitor
- The output amplitude remains constant over the frequency range
- Either a tapped coil or two fixed inductors are needed

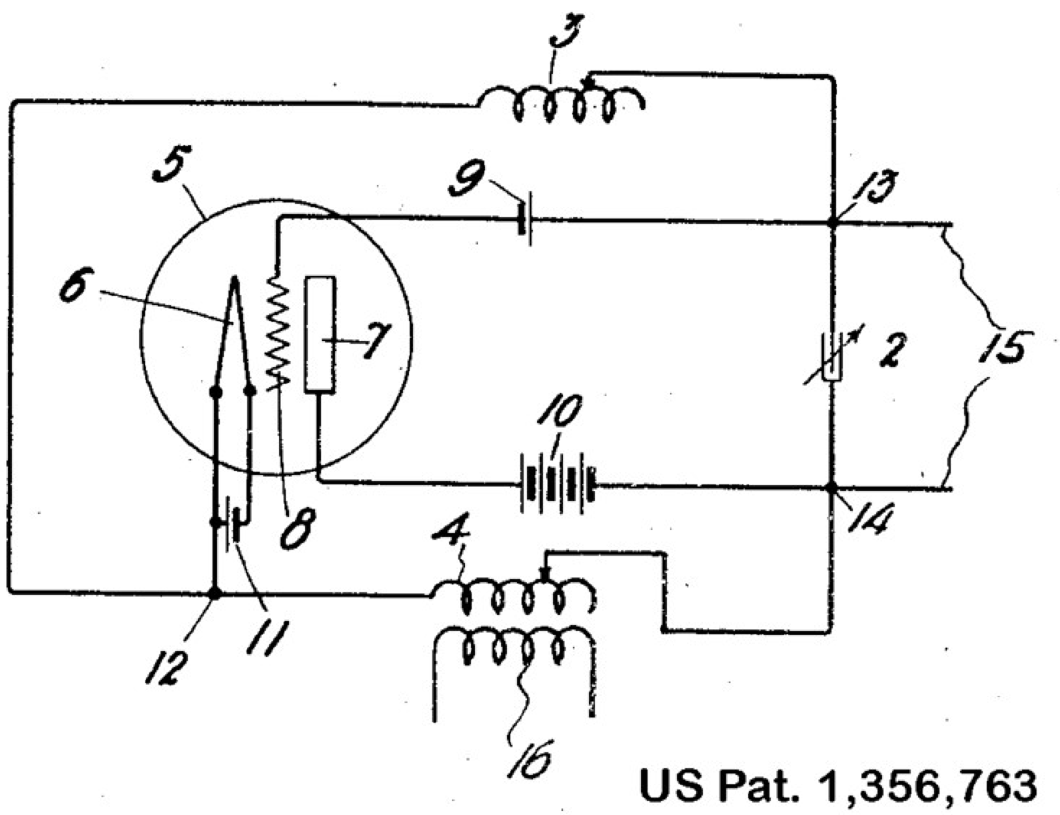
Disadvantages include:

- Harmonic-rich content if taken from the amplifier and not directly from the LC circuit.

Note that, if the inductance of the two partial coils L_1 and L_2 is given (e.g. in a simulator), the total effective inductance that determines the frequency of the oscillation is (coupling factor k):

$$L_0 = L_1 + L_2 + k * \sqrt{L_1 * L_2}$$

History

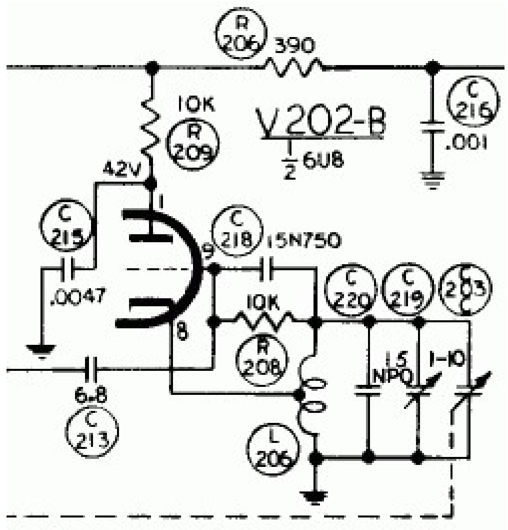


Original Patent Drawing.

The Hartley oscillator was invented by Ralph V.L. Hartley while he was working for the Research Laboratory of the Western Electric Company. Hartley invented and patented the design in 1915 while overseeing Bell System's transatlantic radiotelephone tests; it was awarded patent number 1,356,763 on October 26, 1920. Note that the above basic schematic is essentially the same as in the patent drawing, except that the tube is replaced by a J-FET, and that the battery for a negative grid bias is not needed.

In 1946 Hartley was awarded the IRE medal of honor "For his early work on oscillating circuits employing triode tubes and likewise for his early recognition and clear exposition of the fundamental relationship between the total amount of information which may be transmitted over a transmission system of limited band-width and the time required."(The second half of the citation refers to Hartley's work in information theory which largely paralleled Harry Nyquist.)

Applications



Part of Scott 310E circuit diagram

The Hartley oscillator was extensively used on all broadcast bands including the FM 88-108MHz band. An example is given of the Scott 310E RF oscillator for its FM section.

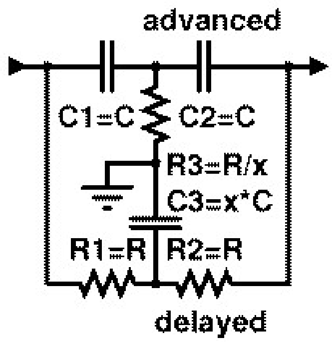
RC oscillator

Linear electronic oscillator circuits, which generate a sinusoidal output signal, are composed of an amplifier and a frequency selective element, a filter. An oscillator circuit which uses an RC network, a combination of resistors and capacitors, for its frequency selective part is called an **RC oscillator**.

Sine wave oscillators

Two configurations are common. One is called a Wien bridge oscillator. In this circuit, two RC circuits are used, one with the RC components in series and one with the RC components in parallel. The Wien Bridge is often used in audio signal generators because it can be easily tuned using a two-section variable capacitor or a two section variable potentiometer (which is more easily obtained than a variable capacitor suitable for generation at low frequencies). The archetypical HP 200 audio oscillator is a Wien Bridge oscillator.

WT



Twin-T filter

The second common design is called a "Twin-T" oscillator as it uses two "T" RC circuits operated in parallel. One circuit is an R-C-R "T" which acts as a low-pass filter. The second circuit is a C-R-C "T" which operates as a high-pass filter. Together, these circuits form a bridge which is tuned at the desired frequency of oscillation. The signal in the C-R-C branch of the Twin-T filter is advanced, in the R-C-R - delayed, so they may cancel one another for frequency $f = \frac{1}{2\pi RC}$ if $x = 2$; if it is connected as a negative feedback to an amplifier, and x>2, the amplified becomes an oscillator.

Another common design is the phase-shift oscillator.

If they are to produce an undistorted sine wave, RC oscillators usually require some form of amplitude control. Many common designs simply use an incandescent lamp or a thermistor in the feedback circuit. These oscillators take advantage of the fact that the resistance of the tungsten filament of the lamp increases in proportion to its temperature, a thermistor works in a similar fashion. Operated well below the point at which the filament actually illuminates, increased amplitude of the feedback signal causes increased current flow in the filament thereby increasing the resistance of the filament. The increased resistance of the filament reduces the feedback signal, limiting the oscillator's signal to the linear range. (That is, clipping is prevented.) The HP 200 oscillator introduced this technique.

More-sophisticated oscillators measure the output level and use this as feedback to control the gain of the voltage-controlled amplifier within the oscillator. If the amplitude detector has a flat frequency response, then this negative feedback of the amplitude measurement will ensure that the oscillator has a constant output amplitude no matter what frequency it is set to generate.

Non-sine wave oscillators

Many designs exist for RC oscillators that are not required to produce a sine wave. Square waves are the most common. Multivibrators are one approach. The 555 timer IC is another very common approach. Most non-sine wave RC oscillators require only a single RC network.

Chapter- 8

Numerically-controlled Oscillator

A **numerically-controlled oscillator** (**NCO**) is a digital signal generator which creates a synchronous (i.e. clocked), discrete-time, discrete-valued representation of a waveform, usually sinusoidal. NCOs are often used in conjunction with a digital-to-analog converter (DAC) at the output to create a direct digital synthesizer (DDS).

Numerically-controlled oscillators offer several advantages over other types of oscillators in terms of agility, accuracy, stability and reliability. NCOs are used in many communications systems including digital up/down converters used in 3G wireless and software radio systems, digital PLLs, radar systems, drivers for optical or acoustic transmissions, and multilevel FSK/PSK modulators/demodulators.

Operation

An NCO generally consists of two parts:

- A *phase accumulator*, which adds to the value held at its output a frequency control value at each clock sample.
- A *phase-to-amplitude converter*, which uses the phase accumulator output word (phase word) as an index into a waveform look-up table to provide a corresponding amplitude sample.

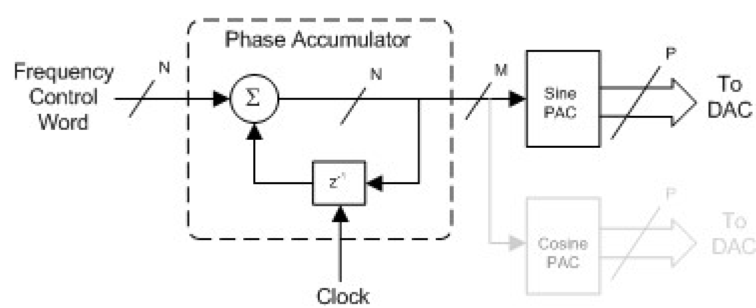


Figure 1: Numerically controlled oscillator with optional quadrature output

When clocked, the phase accumulator (PA) creates a modulo-2^N sawtooth waveform which is then converted by the phase-to-amplitude converter (PAC) to a sampled sinusoid, where N is the number of bits carried in the phase accumulator. N sets the NCO frequency resolution and is normally much larger than the number of bits defining the memory space of the PAC look-up table. If the PAC capacity is 2^M, the PA output word must be truncated to M bits as shown in Figure 1. The truncation of the phase output word does not affect the frequency accuracy but produces a time-varying periodic phase error which is a primary source of spurious products. Another spurious product generation mechanism is finite word length effects of the PAC output (amplitude) word.

The frequency accuracy relative to the clock frequency is limited only by the precision of the arithmetic used to compute the phase. NCOs are phase- and frequency-agile, and can be trivially modified to produce phase-modulated or frequency-modulated by summation at the appropriate node, or provide quadrature outputs as shown in the figure.

Phase accumulator

A binary phase accumulator consists of an N-bit binary adder and a register configured as shown in Figure 1. Each clock cycle produces a new N-bit output consisting of the previous output obtained from the register summed with the frequency control word (FCW) which is constant for a given output frequency. The resulting output waveform is a staircase with step size ΔF, the integer value of the FCW. In some configurations, the phase output is taken from the output of the register which introduces a one clock cycle latency but allows the adder to operate at a higher clock rate.

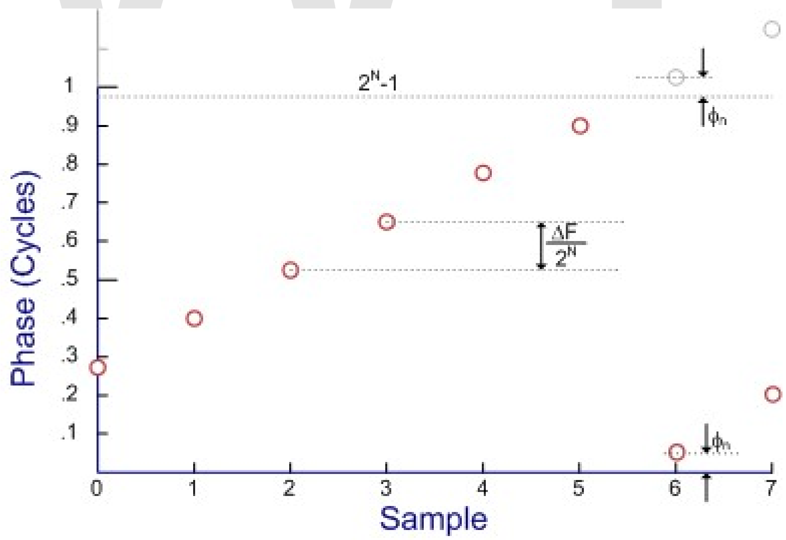


Figure 2: Normalized phase accumulator output

The adder is designed to overflow when the sum of the absolute value of its operands exceeds its capacity (2^N-1). The overflow bit is discarded so the output word width is always equal to its input word width. The remainder φ_n, called the residual, is stored in the register and the cycle repeats, starting this time from φ_n. Since a phase accumulator is a finite state machine, eventually the residual at some sample K must return to the initial value φ_0. The interval K is referred to as the grand repetition rate (GRR) given by

$$\mathrm{GRR} = \frac{2^N}{\mathrm{GCD}(\Delta F, 2^N)}$$

where GCD is the greatest common divisor function. The GRR represents the true periodicity for a given ΔF which for a high resolution NCO can be very long. Usually we are more interested in the *operating frequency* determined by the average overflow rate, given by

$$F_{out} = \frac{\Delta F}{2^N} F_{clock} \quad (1)$$

The *frequency resolution*, defined as the smallest possible incremental change in frequency, is given by

$$F_{res} = \frac{F_{clock}}{2^N} \quad (2)$$

Equation (1) shows that the phase accumulator can be thought of as a programmable non-integer frequency divider of divide ratio $\Delta F / 2^N$.

Phase-to-amplitude converter

The phase-amplitude converter creates the sample-domain waveform from the truncated phase output word received from the PA. The PAC can be a simple read only memory containing 2^M contiguous samples of the desired output waveform which typically is a sinusoid. Oftentimes though, various tricks are employed to reduce the amount of memory required. This include various trigonometric expansions, trigonometric approximations and methods which take advantage of the quadrature symmetry exhibited by sinusoids. Alternately, the PAC may consist of random access memory which can be filled as desired to create an arbitrary waveform generator.

Spurious products

Spurious products are the result of harmonic or non-harmonic distortion in the creation of the output waveform due to non-linear numerical effects in the signal processing chain. Only numerical errors are covered here. For other distortion mechanisms created in the digital-to-analog converter see the corresponding section in the direct-digital synthesizer article.

Phase truncation spurs

The number of phase accumulator bits of an NCO, N is usually between 24 and 64. If the PA output word were used directly to index the PAC look-up table an untenably high storage capacity in the ROM would be required. As such, the PA output word must be truncated to span a reasonable memory space. Truncation of the phase word cause phase modulation of the output sinusoid which introduces non-harmonic distortion in proportion to the number of bits truncated. The number of spurious products created by this distortion is given by:

$$n_W = \frac{2^N}{\mathrm{GCD}(\Delta F, 2^W)} \qquad (3)$$

where W is the number of bits truncated.

In calculating the Spurious-free dynamic range, we are interested in the spurious product with the largest amplitude relative to the carrier output level given by:

$$\zeta_{max} = 2^{-M} \frac{\pi \mathrm{GCD}(\Delta F, 2^W)}{\sin \pi \cdot 2^{-P} \mathrm{GCD}(\Delta F, 2^W)}$$

where P is word width of the DAC. For W >4,

$$\zeta_{max} \approx -6.02 \cdot P \text{ dBc}.$$

Another related spurious generation method is the slight modulation due to the GRR outlined above. The amplitude of these spurs is low for large N and their frequency is generally too low to be detectable but they may cause issues for some applications.

Amplitude truncation spurs

Another source of spurious products is the amplitude quantization of the sampled waveform contained in the PAC look up table(s). If the number of DAC bits is P, the AM spur level is approximately equal to −6.02 P − 1.76 dBc.

Mitigation techniques

Phase truncation spurs can be reduced substantially by the introduction of white gaussian noise prior to truncation. The so-called dither noise is summed into the lower W+1 bits of the PA output word to linearize the truncation operation. Often the improvement can be achieved without penalty because the DAC noise floor tends to dominate system performance. Amplitude truncation spurs can not be mitigated in this fashion. Introduction of noise into the static values held in the PAC ROMs would not eliminate the cyclicality of the trunction error terms and thus would not achieve the desired effect.

Chapter- 9

Voltage-controlled Oscillator

A **voltage-controlled oscillator** or **VCO** is an electronic oscillator designed to be controlled in oscillation frequency by a voltage input. The frequency of oscillation is varied by the applied DC voltage, while modulating signals may also be fed into the VCO to cause frequency modulation (FM) or phase modulation (PM); a VCO with digital pulse output may similarly have its repetition rate (FSK, PSK) or pulse width modulated (PWM).

A microwave (12-18 GHz) Voltage Controlled Oscillator

Types of VCOs

VCOs can be generally categorized into two groups based on the type of waveform produced: 1) harmonic oscillators, and 2) relaxation oscillators.

Harmonic oscillators generate a sinusoidal waveform. They consist of an amplifier that provides adequate gain and a resonant circuit that feeds back signal to the input. Oscillation occurs at the resonant frequency where a positive gain arises around the loop. Some examples of harmonic oscillators are crystal oscillators and LC-tank oscillators. When part of the resonant circuit's capacitance is provided by a varactor diode, the voltage applied to that diode varies the frequency.

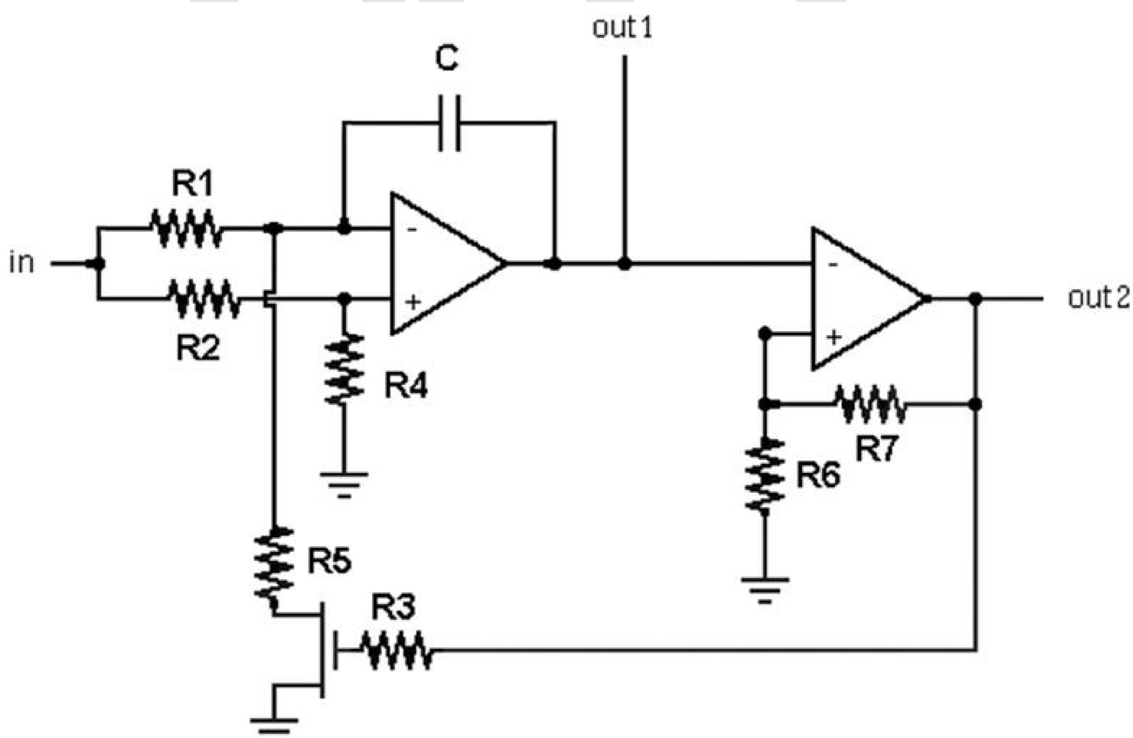
Relaxation oscillators can generate a sawtooth or triangular waveform. They are commonly used in monolithic integrated circuits (ICs). They can provide a wide range of operational frequencies with a minimal number of external components. Relaxation oscillator VCOs can have three topologies: 1) grounded-capacitor VCOs, 2) emitter-coupled VCOs, and 3) delay-based ring VCOs. The first two of these types operate similarly. The amount of time in each state depends on the time for a current to charge or discharge a capacitor. The delay-based ring VCO operates somewhat differently however. For this type, the gain stages are connected in a ring. The output frequency is then a function of the delay in each of stages.

Harmonic oscillator VCOs have these advantages over relaxation oscillators.

- Frequency stability with respect to temperature, noise, and power supply is much better for harmonic oscillator VCOs.
- They have good accuracy for frequency control since the frequency is controlled by a crystal or tank circuit.

A disadvantage of harmonic oscillator VCOs is that they cannot be easily implemented in monolithic ICs. Relaxation oscillator VCOs are better suited for this technology. Relaxation VCOs are also tunable over a wider range of frequencies.

Control of frequency in VCOs



Voltage-controlled oscillator schematic - audio

A voltage-controlled capacitor is one method of making an LC oscillator vary its frequency in response to a control voltage. Any reverse-biased semiconductor diode displays a measure of voltage-dependent capacitance and can be used to change the frequency of an oscillator by varying a control voltage applied to the diode. Special-purpose variable capacitance varactor diodes are available with well-characterized wide-ranging values of capacitance. Such devices are very convenient in the manufacture of voltage-controlled oscillators For low-frequency VCOs, other methods of varying the frequency (such as altering the charging rate of a capacitor by means of a voltage controlled current source) are used.

The frequency of a ring oscillator is controlled by varying either the supply voltage or the capacitive loading on each stage.

Voltage-controlled crystal oscillators

A **voltage-controlled crystal oscillator** (**VCXO**) is used when the frequency of operation needs to be adjusted only finely. The frequency of a voltage-controlled crystal oscillator can be varied only by typically a few tens of parts per million (ppm), because the high Q factor of the crystals allows "pulling" over only a small range of frequencies.

There are two reasons for using a VCXO:

- To adjust the output frequency to match (or perhaps be some exact multiple of) an accurate external reference.
- Where the oscillator drives equipment that may generate radio-frequency interference, adding a varying voltage to its control input can disperse the interference spectrum to make it less objectionable.

A **temperature-compensated VCXO** (**TCVCXO**) incorporates components that partially correct the dependence on temperature of the resonant frequency of the crystal. A smaller range of voltage control then suffices to stabilize the oscillator frequency in applications where temperature varies, such as heat buildup inside a transmitter.

VCO time-domain equations

$$f_{tuning}(t) = K_o \cdot \ v_{in}(t)$$
$$\int f_{tuning}(t)\, dt = \theta_{out}(t)$$

- K_o is called the oscillator gain. Its units are hertz per volt.
- $f_{tuning}(t)$ is the symbol for the time-domain waveform that is the VCO's tunable frequency component.
- $\theta_{out}(t)$ is the symbol for the time-domain waveform that is the VCO's output phase.
- $v_{in}(t)$ is the time-domain symbol of the control (input) voltage of the VCO; it is sometimes also represented as $v_{tune}(t)$

VCO freq-domain equations

$$F_{tuning}(s) = K_o \cdot \ V_{in}(s)$$
$$\frac{F_{tuning}(s)}{s} = \Theta_{out}(s)$$

VCO design and circuits

Tuning range, tuning gain and phase noise are the most important factors of the basic design of a VCO. Generally low phase noise is preferred in the VCO. The important elements that determine the phase noise of an oscillator are the material, transistor's flicker noise corner frequency, the loaded Q of the resonator and the final signal to noise ratio.

Most commonly used VCO circuits are the Clapp and Colpitts oscillators. The more widely used oscillator of the two is Colpitts and these oscillators are very similar in configuration.

VCOs generally have the lowest Q-factor of the used oscillators, and so suffer more jitter than the other types. The jitter can be made low enough for many applications (such as driving an ASIC), in which case VCOs enjoy the advantages of having no off-chip components (expensive) or on-chip inductors (low yields on generic CMOS processes). These oscillators also have larger tuning ranges than the other kinds, which improves yield and is sometimes a feature of the end product (for instance, the dot clock on a graphics card which drives a wide range of monitors).

Applications

VCOs are used in:

- Electronic jamming equipment
- Function generators,
- The production of electronic music, to generate variable tones,
- Phase-locked loops,
- Frequency synthesizers used in communication equipment.

Voltage-Controlled Crystal Oscillator as a Clock Generator

A clock generator is an oscillator that provides a timing signal to synchronize operations in digital circuits. VCXO clock generators are used in many areas such as digital TV, modems, transmitters and computers. Design parameters for a VCXO clock generator are tuning voltage range, center frequency, frequency tuning range and the timing jitter of the output signal. Jitter is a form of phase noise that must be minimised in applications such as radio receivers, transmitters and measuring equipment.

The tuning range of a VCXO is typically a few ppm (parts per million) corresponding to a control voltage range of typically 0 to 3 volts. When a wider selection of clock frequencies is needed the VCXO output can be passed through digital divider circuits to obtain lower frequency(ies) or be fed to a PLL (Phase Locked Loop). ICs containing both a VCXO (for external crystal) and a PLL are available. A typical application is to provide clock frequencies in a range from 12 kHz to 96 kHz to an audio digital to analog converter.

WT

Chapter- 10

Phase-locked Loop

A **phase-locked loop** or **phase lock loop** (PLL) is a control system that tries to generate an output signal whose phase is related to the phase of the input "reference" signal. It is an electronic circuit consisting of a variable frequency oscillator and a phase detector. This circuit compares the phase of the input signal with the phase of the signal derived from its output oscillator and adjusts the frequency of its oscillator to keep the phases matched. The signal from the phase detector is used to control the oscillator in a feedback loop.

Frequency is the derivative of phase. Keeping the input and output phase in lock step implies keeping the input and output frequencies in lock step. Consequently, a phase-locked loop can track an input frequency, or it can generate a frequency that is a multiple of the input frequency. The former property is used for demodulation, and the latter property is used for indirect frequency synthesis.

Phase-locked loops are widely used in radio, telecommunications, computers and other electronic applications. They may generate stable frequencies, recover a signal from a noisy communication channel, or distribute clock timing pulses in digital logic designs such as microprocessors. Since a single integrated circuit can provide a complete phase-locked-loop building block, the technique is widely used in modern electronic devices, with output frequencies from a fraction of a hertz up to many gigahertz.

Practical analogies

Automobile race analogy

For a practical idea of what is going on, consider an auto race. There are many cars, and each of them wants to go around the track as fast as possible. Each lap corresponds to a complete cycle, and each car will complete dozens of laps per hour. The number of laps per hour (a speed) is a frequency, but the number of laps (a distance) corresponds to a phase. At one instant, car 3 may have gone 37.23 laps.

During most of the race, each car is on its own and is trying to beat every other car on the course. However, if there is an accident, a pace car comes out to set a safe speed. None of the race cars is permitted to pass the pace car (or the race cars in front of them), but each of the race cars wants to stay as close to the pace car as it can. While it is on the track, the pace car is a reference, and the race cars become phase-locked loops. Each driver will measure the phase difference (a distance in laps) between him and the pace car. If the driver is far away, he will increase his engine speed to close the gap. If he's too close to the pace car, he will slow down. The result is all the race cars lock on to the phase of the pace car. The cars travel around the track in a tight group that is a small fraction of a lap.

Clock analogy

Phase can be proportional to time, so a phase difference can be a time difference. Clocks are, with varying degrees of accuracy, phase-locked (time-locked) to a master clock.

Left on its own, each clock will mark time at slightly different rates. A wall clock, for example, might be fast by a few seconds per hour compared to the reference clock at NIST. Over time, that time difference would become substantial.

To keep his clock in synch, each week the owner compares the time on his wall clock to a more accurate clock (a phase comparison), and he resets his clock. Left alone, the wall clock will continue to diverge from the reference clock at the same few seconds per hour rate.

Some clocks have a timing adjustment (a fast-slow control). When the owner compared his wall clock's time to the reference time, he noticed that his clock was too fast. Consequently, he could turn the timing adjust a small amount to make the clock run a little slower. If things work out right, his clock will be more accurate. Over a series of weekly adjustments, the wall clock's notion of a second would agree with the reference time (within the wall clock's stability).

An early mechanical version of a phase-locked loop was used in 1921 in the Shortt-Synchronome clock.

History

Automatic synchronization of electronic oscillators was described in 1923. Earliest research towards what became known as the phase-locked loop goes back to 1932, when British researchers developed an alternative to Edwin Armstrong's superheterodyne receiver, the Homodyne or direct-conversion receiver. In the homodyne or synchrodyne system, a local oscillator was tuned to the desired input frequency and multiplied with the input signal. The resulting output signal included the original modulation information. The intent was to develop an alternative receiver circuit that required fewer tuned circuits than the superheterodyne receiver. Since the local oscillator would rapidly drift in frequency, an automatic correction signal was applied to the oscillator, maintaining it in the same phase and frequency as the desired signal. The technique was described in 1932, in a paper by Henri de Bellescize, in the French journal *L'Onde Électrique*.

In analog television receivers since at least the late 1930s, phase-locked-loop horizontal and vertical sweep circuits are locked to synchronization pulses in the broadcast signal.

When Signetics introduced a line of monolithic integrated circuits that were complete phase-locked loop systems on a chip in 1969, applications for the technique multiplied. A few years later RCA introduced the "CD4046" CMOS Micropower Phase-Locked Loop, which became a popular integrated circuit.

Structure and function

Phase-locked loop mechanisms may be implemented as either analog or digital circuits. Both implementations use the same basic structure. Both analog and digital PLL circuits include four basic elements:

- Phase detector,
- low-pass filter
- Variable frequency oscillator, and
- feedback path (which may include a frequency divider).

Variations

There are several variations of PLLs. Some terms that are used are analog phase-locked loop (APLL) also referred to as a linear phase-locked loop (LPLL), digital phase-locked loop (DPLL), all digital phase-locked loop (ADPLL), and software phase-locked loop (SPLL).

Analog or Linear PLL (LPLL)
: Phase detector is an analog multiplier. Loop filter is active or passive. Uses a Voltage-controlled oscillator (VCO).

Digital PLL (DPLL)
: An analog PLL with a digital phase detector (such as XOR, edge-trigger JK, phase frequency detector). May have digital divider in the loop.

All digital PLL (ADPLL)
: Phase detector, filter and oscillator are digital. Uses a numerically-controlled oscillator (NCO).

Software PLL (SPLL)
: Functional blocks are implemented by software rather than specialized hardware.

Performance parameters

- Type and order
- Lock range: The frequency range the PLL is able to stay locked. Mainly defined by the VCO range.

WT

- Capture range: The frequencie range the PLL is able to lock-in, starting from unlocked condition. This range is usually smaller than the lock range and will depend e.g. on phase detector.
- Loop bandwidth: Defining the speed of the control loop.
- Transient response: Like overshoot and settling time to a certain accuracy (like 50ppm).
- Steady-state errors: Like remaining phase or timing error
- Output spectrum purity: Like sidebands generated from a certain VCO tuning voltage ripple.
- Phase-noise: Defined by noise energy in a certain frequency band (like 10kHz offset from carrier). Highly dependent on VCO phase-noise, PLL bandwidth, etc.
- General parameters: Such as power consumption, supply voltage range, output amplitude, etc.

Applications

Phase-locked loops are widely used for synchronization purposes; in space communications for coherent demodulation and threshold extension, bit synchronization, and symbol synchronization. Phase-locked loops can also be used to demodulate frequency-modulated signals. In radio transmitters, a PLL is used to synthesize new frequencies which are a multiple of a reference frequency, with the same stability as the reference frequency.

Other applications include:

- Demodulation of both FM and AM signals
- Recovery of small signals that otherwise would be lost in noise (lock-in amplifier)
- Recovery of clock timing information from a data stream such as from a disk drive
- Clock multipliers in microprocessors that allow internal processor elements to run faster than external connections, while maintaining precise timing relationships
- DTMF decoders, modems, and other tone decoders, for remote control and telecommunications

Clock recovery

Some data streams, especially high-speed serial data streams (such as the raw stream of data from the magnetic head of a disk drive), are sent without an accompanying clock. The receiver generates a clock from an approximate frequency reference, and then phase-aligns to the transitions in the data stream with a PLL. This process is referred to as clock recovery. In order for this scheme to work, the data stream must have a transition frequently enough to correct any drift in the PLL's oscillator. Typically, some sort of redundant encoding is used; 8B10B is very common.

Deskewing

If a clock is sent in parallel with data, that clock can be used to sample the data. Because the clock must be received and amplified before it can drive the flip-flops which sample the data, there will be a finite, and process-, temperature-, and voltage-dependent delay between the detected clock edge and the received data window. This delay limits the frequency at which data can be sent. One way of eliminating this delay is to include a deskew PLL on the receive side, so that the clock at each data flip-flop is phase-matched to the received clock. In that type of application, a special form of a PLL called a delay-locked loop (DLL) is frequently used.

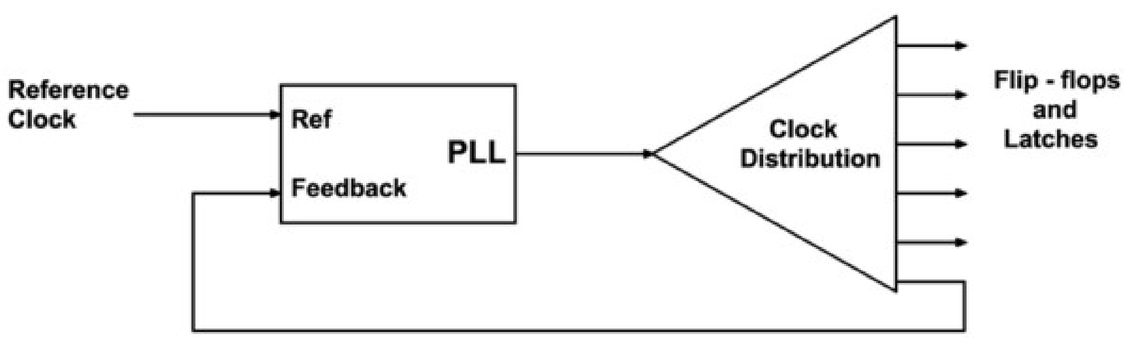
Clock generation

Many electronic systems include processors of various sorts that operate at hundreds of megahertz. Typically, the clocks supplied to these processors come from clock generator PLLs, which multiply a lower-frequency reference clock (usually 50 or 100 MHz) up to the operating frequency of the processor. The multiplication factor can be quite large in cases where the operating frequency is multiple gigahertz and the reference crystal is just tens or hundreds of megahertz.

Spread spectrum

All electronic systems emit some unwanted radio frequency energy. Various regulatory agencies (such as the FCC in the United States) put limits on the emitted energy and any interference caused by it. The emitted noise generally appears at sharp spectral peaks (usually at the operating frequency of the device, and a few harmonics). A system designer can use a spread-spectrum PLL to reduce interference with high-Q receivers by spreading the energy over a larger portion of the spectrum. For example, by changing the operating frequency up and down by a small amount (about 1%), a device running at hundreds of megahertz can spread its interference evenly over a few megahertz of spectrum, which drastically reduces the amount of noise seen on broadcast FM radio channels, which have a bandwidth of several tens of kilohertz.

Clock distribution



Typically, the reference clock enters the chip and drives a phase locked loop (**PLL**), which then drives the system's clock distribution. The clock distribution is usually balanced so that the clock arrives at every endpoint simultaneously. One of those endpoints is the PLL's feedback input. The function of the PLL is to compare the distributed clock to the incoming reference clock, and vary the phase and frequency of its output until the reference and feedback clocks are phase and frequency matched.

PLLs are ubiquitous—they tune clocks in systems several feet across, as well as clocks in small portions of individual chips. Sometimes the reference clock may not actually be a pure clock at all, but rather a data stream with enough transitions that the PLL is able to recover a regular clock from that stream. Sometimes the reference clock is the same frequency as the clock driven through the clock distribution, other times the distributed clock may be some rational multiple of the reference.

Jitter and noise reduction

One desirable property of all PLLs is that the reference and feedback clock edges be brought into very close alignment. The average difference in time between the phases of the two signals when the PLL has achieved lock is called the **static phase offset** (also called the **steady-state phase error**). The variance between these phases is called **tracking jitter**. Ideally, the static phase offset should be zero, and the tracking jitter should be as low as possible.

Phase noise is another type of jitter observed in PLLs, and is caused by the oscillator itself and by elements used in the oscillator's frequency control circuit. Some technologies are known to perform better than others in this regard. The best digital PLLs are constructed with emitter-coupled logic (ECL) elements, at the expense of high power consumption. To keep phase noise low in PLL circuits, it is best to avoid saturating logic families such as transistor-transistor logic (TTL) or CMOS.

Another desirable property of all PLLs is that the phase and frequency of the generated clock be unaffected by rapid changes in the voltages of the power and ground supply lines, as well as the voltage of the substrate on which the PLL circuits are fabricated. This is called substrate and supply noise rejection. The higher the noise rejection, the better.

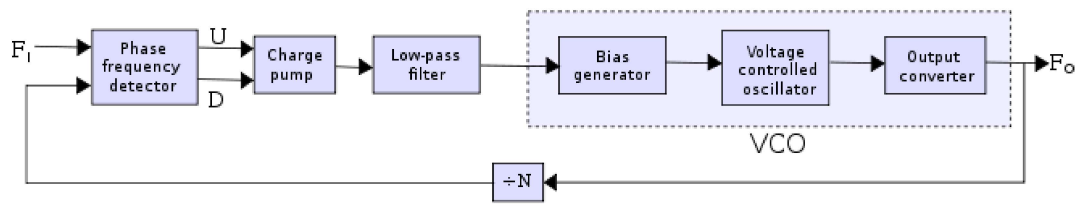
To further improve the phase noise of the output, an injection locked oscillator can be employed following the VCO in the PLL.

Frequency Synthesis

In digital wireless communication systems (GSM, CDMA etc.), PLLs are used to provide the local oscillator for up-conversion during transmission and down-conversion during reception. In most cellular handsets this function has been largely integrated into a single integrated circuit to reduce the cost and size of the handset. However, due to the high performance required of base station terminals, the transmission and reception circuits are built with discrete components to achieve the levels of performance required. GSM local oscillator modules are typically built with a frequency synthesizer integrated circuit and discrete resonator VCOs.

Frequency synthesizer manufacturers include Analog Devices, National Semiconductor and Texas Instruments. VCO manufacturers include Sirenza, Z-Communications, Inc. (Z-COMM).

Phase-locked loop block diagram



Digital phase-locked loop block diagram

A phase detector compares two input signals and produces an error signal which is proportional to their phase difference. The error signal is then low-pass filtered and used to drive a VCO which creates an output phase. The output is fed through an optional divider back to the input of the system, producing a negative feedback loop. If the output phase drifts, the error signal will increase, driving the VCO phase in the opposite direction so as to reduce the error. Thus the output phase is locked to the phase at the other input. This input is called the reference.

Analog phase locked loops are generally built with an analog phase detector, low pass filter and VCO placed in a negative feedback configuration. A digital phase locked loop uses a digital phase detector; it may also have a divider in the feedback path or in the reference path, or both, in order to make the PLL's output signal frequency a rational multiple of the reference frequency. A non-integer multiple of the reference frequency can also be created by replacing the simple divide-by-N counter in the feedback path with a programmable pulse swallowing counter. This technique is usually referred to as a fractional-N synthesizer or fractional-N PLL.

The oscillator generates a periodic output signal. Assume that initially the oscillator is at nearly the same frequency as the reference signal. If the phase from the oscillator falls behind that of the reference, the phase detector changes the control voltage of the oscillator so that it speeds up. Likewise, if the phase creeps ahead of the reference, the phase detector changes the control voltage to slow down the oscillator. Since initially the oscillator may be far from the reference frequency, practical phase detectors may also respond to frequency differences, so as to increase the lock-in range of allowable inputs.

Depending on the application, either the output of the controlled oscillator, or the control signal to the oscillator, provides the useful output of the PLL system.

WT

Elements

Phase detector

The two inputs of the phase detector are the reference input and the feedback from the VCO. The PD output controls the VCO such that the phase difference between the two inputs is held constant, making it a negative feedback system. There are several types of phase detectors in the two main categories of analog and digital.

Different types of phase detectors have different performance characteristics.

For instance, the frequency mixer produces harmonics that adds complexity in applications where spectral purity of the VCO signal is important. The resulting unwanted (spurious) sidebands, also called "reference spurs" can dominate the filter requirements and reduce the capture range and lock time well below the requirements. In these applications the more complex digital phase detectors are used which do not have as severe a reference spur component on their output. Also, when in lock, the steady-state phase difference at the inputs using this type of phase detector is near 90 degrees. The actual difference is determined by the DC loop gain.

A **bang-bang** charge pump phase detector must always have a **dead band** where the phases of inputs are close enough that the detector detects no phase error. For this reason, bang-bang phase detectors are associated with significant minimum peak-to-peak jitter, because of drift within the dead band. However these types, having outputs consisting of very narrow pulses at lock, are very useful for applications requiring very low VCO spurious outputs. The narrow pulses contain very little energy and are easy to filter out of the VCO control voltage. This results in low VCO control line ripple and therefore low FM sidebands on the VCO.

In PLL applications it is frequently required to know when the loop is out of lock. The more complex digital phase-frequency detectors usually have an output that allows a reliable indication of an out of lock condition.

Filter

The block commonly called the PLL loop filter (usually a low pass filter) generally has two distinct functions.

The primary function is to determine loop dynamics, also called stability. This is how the loop responds to disturbances, such as changes in the reference frequency, changes of the feedback divider, or at startup. Common considerations are the range over which the loop can achieve lock (pull-in range, lock range or capture range), how fast the loop achieves lock (lock time, lock-up time or settling time) and damping behavior. Depending on the application, this may require one or more of the following: a simple proportion (gain or attenuation), an integral (low pass filter) and/or derivative (high pass filter). Loop parameters commonly examined for this are the loop's gain margin and phase margin.

Common concepts in control theory including the PID controller are used to design this function.

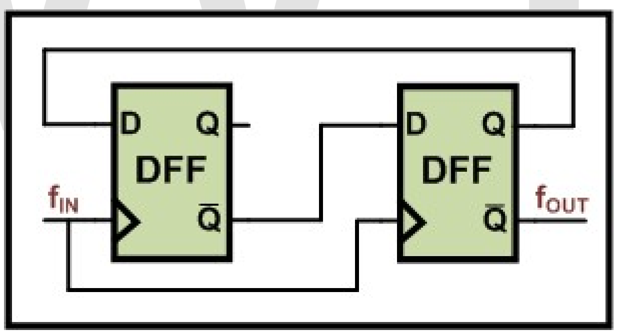
The second common consideration is limiting the amount of reference frequency energy (ripple) appearing at the phase detector output that is then applied to the VCO control input. This frequency modulates the VCO and produces FM sidebands commonly called "reference spurious". The low pass characteristic of this block can be used to attenuate this energy, but at times a band reject "notch" may also be useful.

The design of this block can be dominated by either of these considerations, or can be a complex process juggling the interactions of the two. Typical trade-offs are: increasing the bandwidth usually degrades the stability or too much damping for better stability will reduce the speed and increase settling time. Often also the phase-noise is affected.

Oscillator

All phase-locked loops employ an oscillator element with variable frequency capability. This can be an analog VCO either driven by analog circuitry in the case of an APLL or driven digitally through the use of a digital-to-analog converter as is the case for some DPLL designs. Pure digital oscillators such as a numerically-controlled oscillator are used in ADPLLs.

Feedback path and optional divider



An Example Digital Divider (by 4) for use in the Feedback Path of a Multiplying PLL

PLLs may include a divider between the oscillator and the feedback input to the phase detector to produce a frequency synthesizer. A programmable divider is particularly useful in radio transmitter applications, since a large number of transmit frequencies can be produced from a single stable, accurate, but expensive, quartz crystal–controlled reference oscillator.

Some PLLs also include a divider between the reference clock and the reference input to the phase detector. If the divider in the feedback path divides by N and the reference input divider divides by M, it allows the PLL to multiply the reference frequency by N / M. It might seem simpler to just feed the PLL a lower frequency, but in some cases the reference frequency may be constrained by other issues, and then the reference divider is useful.

Frequency multiplication in a sense can also be attained by locking the PLL to the 'N'th harmonic of the signal.

It should also be noted that the feedback is not limited to a frequency divider. This element can be other elements such as a frequency multiplier, or a mixer. The multiplier will make the VCO output a sub-multiple (rather than a multiple) of the reference frequency. A mixer can translate the VCO frequency by a fixed offset. It may also be a combination of these. An example being a divider following a mixer; this allows the divider to operate at a much lower frequency than the VCO without a loss in loop gain.

Modeling

Time domain model

The equations governing a phase-locked loop with an analog multiplier as the phase detector may be derived as follows. Let the input to the phase detector be $x_c(t)$ and the output of the VCO is $x_r(t)$ with frequency $\omega_r(t)$, then the output of the phase detector $x_m(t)$ is given by

$$x_m(t) = x_c(t) \cdot x_r(t)$$

the VCO frequency may be written as a function of the VCO input $y(t)$ as

$$\omega_r(t) = \omega_f + g_v y(t)$$

where g_v is the *sensitivity* of the VCO and is expressed in Hz / V.

Hence the VCO output takes the form

$$x_r(t) = A_r \cos\left(\int_0^t \omega_r(\tau)\, d\tau\right) = A_r \cos(\omega_f t + \varphi(t))$$

where

$$\varphi(t) = \int_0^t g_v y(\tau)\, d\tau$$

The loop filter receives this signal as input and produces an output

$$x_f(t) = F_{\text{filter}}(x_m(t))$$

where F_{Filter} is the operator representing the loop filter transformation.

When the loop is closed, the output from the loop filter becomes the input to the VCO thus

$$y(t) = x_f(t) = F_{\text{filter}}(x_m(t))$$

We can deduce how the PLL reacts to a sinusoidal input signal:

$$x_c(t) = A_c \sin(\omega_c t).$$

The output of the phase detector then is:

$$x_m(t) = A_c \sin(\omega_c t) A_r \cos(\omega_f t + \varphi(t)).$$

This can be rewritten into sum and difference components using trigonometric identities:

$$x_m(t) = \frac{A_c A_f}{2} \sin(\omega_c t - \omega_f t - \varphi(t)) + \frac{A_c A_f}{2} \sin(\omega_c t + \omega_f t + \varphi(t))$$

As an approximation to the behaviour of the loop filter we may consider only the difference frequency being passed with no phase change, which enables us to derive a small-signal model of the phase-locked loop. If we can make $\omega_f \approx \omega_c$, then the $\sin(\cdot)$can be approximated by its argument resulting in: $y(t) = x_f(t) \simeq -A_c A_f \varphi(t)/2$. The phase-locked loop is said to be *locked* if this is the case.

Linearized phase domain model

Phase locked loops can also be analyzed as control systems by applying the Laplace transform. The loop response can be written as:

$$\frac{\theta_o}{\theta_i} = \frac{K_p K_v F(s)}{s + K_p K_v F(s)}$$

Where

- θ_o is the output phase in radians
- θ_i is the input phase in radians
- K_p is the phase detector gain in volts per radian
- K_v is the VCO gain in radians per volt-second
- $F(s)$ is the loop filter transfer function (dimensionless)

The loop characteristics can be controlled by inserting different types of loop filters. The simplest filter is a one-pole RC circuit. The loop transfer function in this case is:

$$F(s) = \frac{1}{1 + sRC}$$

The loop response becomes:

$$\frac{\theta_o}{\theta_i} = \frac{\frac{K_p K_v}{RC}}{s^2 + \frac{s}{RC} + \frac{K_p K_v}{RC}}$$

This is the form of a classic harmonic oscillator. The denominator can be related to that of a second order system:

$$s^2 + 2s\zeta\omega_n + \omega_n^2$$

Where

- ζ is the damping factor
- ω_n is the natural frequency of the loop

For the one-pole RC filter,

$$\omega_n = \sqrt{\frac{K_p K_v}{RC}}$$

$$\zeta = \frac{1}{2\sqrt{K_p K_v RC}}$$

The loop natural frequency is a measure of the response time of the loop, and the damping factor is a measure of the overshoot and ringing. Ideally, the natural frequency should be high and the damping factor should be near 0.707 (critical damping). With a single pole filter, it is not possible to control the loop frequency and damping factor independently. For the case of critical damping,

$$RC = \frac{1}{2K_p K_v}$$

$$\omega_c = K_p K_v \sqrt{2}$$

A slightly more effective filter, the lag-lead filter includes one pole and one zero. This can be realized with two resistors and one capacitor. The transfer function for this filter is

$$F(s) = \frac{1 + sCR_2}{1 + sC(R_1 + R_2)}$$

This filter has two time constants

$\tau_1 = C(R_1 + R_2)$
$\tau_2 = CR_2$

Substituting above yields the following natural frequency and damping factor

$$\omega_n = \sqrt{\frac{K_p K_v}{\tau_1}}$$
$$\zeta = \frac{1}{2\omega_n \tau_1} + \frac{\omega_n \tau_2}{2}$$

The loop filter components can be calculated independently for a given natural frequency and damping factor

$$\tau_1 = \frac{K_p K_v}{\omega_n^2}$$
$$\tau_2 = \frac{2\zeta}{\omega_n} - \frac{1}{K_p K_v}$$

Real world loop filter design can be much more complex e.g. using higher order filters to reduce various types or source of phase noise.

WT

Chapter- 11

Variable-frequency Oscillator

A **variable frequency oscillator (VFO)** in electronics is a oscillator whose frequency can be tuned (i.e. varied) over some range. It is a necessary component in any tunable radio receiver or transmitter that works by the superheterodyne principle, and controls the frequency to which the apparatus is tuned.

Purpose

In a simple superhet radio receiver, the incoming radio frequency signal (at frequency f_{IN}) from the antenna is *mixed* with the VFO output signal tuned to f_{LO}, producing an intermediate frequency (*IF*) signal that can be processed downstream to extract the modulated information. The IF signal frequency is chosen to be either the sum of the two frequencies at the mixer inputs (up-conversion), $f_{IN} + f_{LO}$ or more commonly, the difference frequency (down-conversion), $f_{IN} - f_{LO}$, depending on the receiver design.

In addition to the desired *IF* signal and its unwanted image (the mixing product of opposite sign above), the mixer output will also contain the two original frequencies, f_{IN} and f_{LO} and various harmonic combinations of the input signals. These undesired signals are rejected by the IF filter. If a double balanced mixer is employed, the input signals appearing at the mixer outputs are greatly attenuated, reducing the required complexity of the IF filter.

The advantage of using a VFO as a hetrodyning oscillator is that only a small portion of the radio receiver (all sections before the mixer such as the preamplifier) need to have a wide bandwidth. The rest of the receiver can be finely tuned to the IF frequency.

In a direct-conversion receiver, the VFO is tuned to the same frequency as the incoming radio frequency and $f_{IF} = 0$ Hz. Demodulation takes place at baseband using low-pass filters and amplifiers.

In a radio frequency (RF) transmitter, VFOs are often used to tune the frequency of the output signal, often indirectly through a hetrodyning process similar to that described above. Other uses include chirp generators for radar systems where the VFO is swept rapidly through a range of frequencies, timing signal generation for oscilloscopes and time domain reflectometers, and variable frequency audio generators used in musical instruments and audio test equipment.

There are two main types of VFO in use: analog and digital.

Analog VFO

An analog VFO is an electronic oscillator where the value of at least one of the passive components is adjustable under user control so as to alter its output frequency. The passive component whose value is adjustable is usually a capacitor, but could be a variable inductor.

Tuning Capacitor

The variable capacitor is a mechanical device in which the separation of a series of interleaved metal plates is physically altered to vary its capacitance. Adjustment of this capacitor is sometimes facilitated by a mechanical step-down gearbox to achieve fine tuning.

Varactor

A reversed-biased semiconductor diode exhibits capacitance. Since the width of its non-conducting depletion region depends on the magnitude of the reverse bias voltage, this voltage can be used to control the junction capacitance. The varactor bias voltage may be generated in a number of ways and there may need to be no significant moving parts in the final design. Varactors have a number of disadvantages including temperature drift and aging, electronic noise, low Q factor and non-linearity.

Digital Crystal VFO

Modern radio receivers and transmitters usually use some form of digital frequency synthesis to generate their VFO signal. The advantages including smaller designs, lack of moving parts, and the ease with which preset frequencies can be stored and manipulated in the digital computer that is usually embedded in the design for other purposes.

It is also possible for the radio to become extremely frequency-agile in that the control computer could alter the radio's tuned frequency many tens, thousands or even millions of times a second. This capability allows communications receivers effectively to monitor many channels at once, perhaps using digital selective calling (DSC) techniques to decide when to open an audio output channel and alert users to incoming communications. Pre-programmed frequency agility also forms the basis of some military radio encryption and stealth techniques. Extreme frequency agility lies at the heart of spread spectrum techniques that are currently gaining mainstream acceptance in computer wireless networking such as Wi-Fi.

There are disadvantages to digital synthesis such as the inability of a digital synthesiser to tune smoothly through all frequencies, but with the channelisation of many radio bands, this can also be seen as an advantage in that it prevents radios from operating in between two recognised channels.

Digital frequency synthesis relies on stable crystal controlled reference frequency sources. Crystal controlled oscillators are more stable than inductively and capacitively controlled oscillators. Their disadvantage is that changing frequency (more than a small amount) requires changing the crystal, but frequency synthesizer techniques have made this unnecessary in modern designs.

Digital Frequency Synthesis

The electronic and digital techniques involved in this include:

Direct Digital Synthesis (DDS)
: Enough data points for a mathematical sine function are stored in digital memory. These are recalled at the right speed and fed to a digital to analog converter where the required sine wave is built up.

Direct Frequency Synthesis
: Early channelized communication radios had multiple crystals - one for each channel on which they could operate. After a while this thinking was combined with the basic ideas of heterodyning and mixing described under purpose above. Multiple crystals can be mixed in various combinations to produce various output frequencies.

Phase Locked Loop (PLL)
: Using a varactor-controlled or voltage-controlled oscillator (VCO) (described above in varactor under analog VFO techniques) and a phase detector, a control-loop can be set up so that the VCO's output is frequency-locked to a crystal controlled reference oscillator. The phase detector's comparison is made between the outputs of the two oscillators after frequency division by different divisors. Then by altering the frequency-division divisor(s) under computer control, a variety of actual (undivided) VCO output frequencies can be generated.

The PLL technique dominates most radio VFO designs today.

Performance

The quality metrics for a VFO include frequency stability, phase noise and spectral purity. All of these factors tend to be inversely proportional to the tuning circuit's Q factor. Since in general, the tuning range is also inversely proportional to Q, these performance factors generally degrade as the VFO's frequency range is increased.

Stability

Stability is the measure of how far a VFO's output frequency drifts with time and temperature. To mitigate this problem, VFOs are generally "phase locked" to a stable reference oscillator. PLLs use negative feedback to correct for the frequency drift of the VFO allowing for both wide tuning range and good frequency stability.

Repeatability

Ideally, for the same control input to the VFO, the oscillator should generate exactly the same frequency. A change in the calibration of the VFO can change receiver tuning calibration; periodic re-alignment of a receiver may be needed. VFO's used as part of a phase-locked loop frequency synthesizer have less stringent requirements since the system is as stable as the crystal-controlled reference frequency.

Purity

A plot of a VFO's amplitude vs. frequency may show several peaks, probably harmonically related. Each of these peaks can potentially mix with some other incoming signal and produce a *spurious* response. These *spurii* (sometimes spelled *spuriae*) can result in increased noise or two signals detected where there should only be one. Additional components can be added to a VFO to suppress high-frequency parasitic oscillations, should these be present.

In a transmitter, these spurious signals are generated along with the one desired signal. Filtering may be required to ensure the transmitted signal meets regulations for bandwidth and spurious emissions.

Phase noise

When examined with very sensitive equipment, the pure sine-wave peak in a VFO's frequency graph will most likely turn out not to be sitting on a flat noise-floor. Slight random 'jitters' in the signal's timing will mean that the peak is sitting on 'skirts' of phase noise at frequencies either side of the desired one.

These are also troublesome in crowded bands. They allow through unwanted signals that are fairly close to the expected one, but because of the random quality of these phase-noise 'skirts', the signals are usually unintelligible, appearing just as extra noise in the received signal. The effect is that what should be a clean signal in a crowded band can appear to be a very noisy signal, because of the effects of strong signals nearby.

The effect of VFO phase noise on a transmitter is that random noise is actually transmitted either side of the required signal. Again, this must be avoided for legal reasons in many cases.

WT

Crystal control

In all performances cases, crystal controlled oscillators are better behaved than the semiconductor- and LC-based alternatives. They tend to be more stable, more repeatable, have fewer and lower harmonics and lower noise than all the alternatives in their cost-band. This in part explains their huge popularity in low-cost and computer-controlled (i.e. PLL and synthesizer-based) VFOs.

WT

Chapter- 12

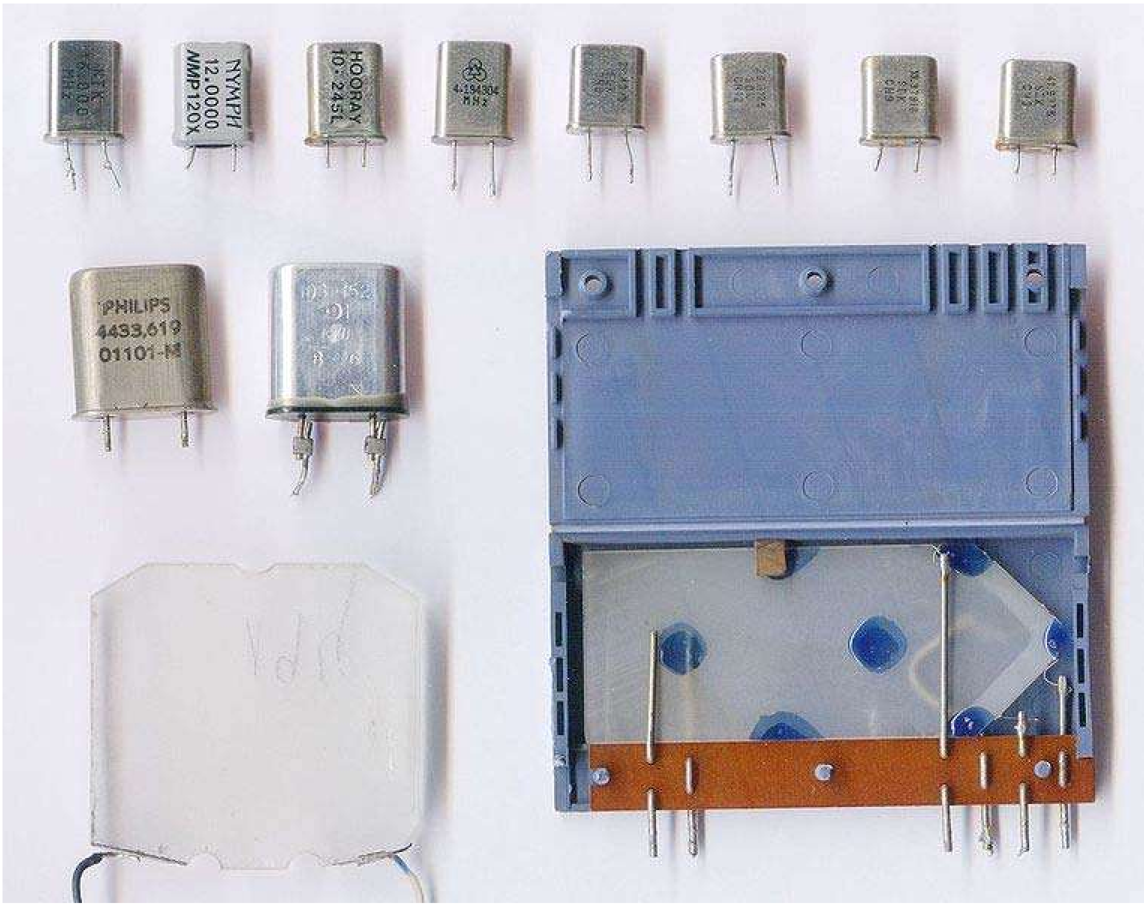
Crystal Oscillator

A miniature 4 MHz quartz crystal enclosed in a hermetically sealed HC-49/US package, used as the resonator in a crystal oscillator.

A **crystal oscillator** is an electronic oscillator circuit that uses the mechanical resonance of a vibrating crystal of piezoelectric material to create an electrical signal with a very precise frequency. This frequency is commonly used to keep track of time (as in quartz wristwatches), to provide a stable clock signal for digital integrated circuits, and to stabilize frequencies for radio transmitters and receivers. The most common type of piezoelectric resonator used is the quartz crystal, so oscillator circuits designed around them became known as "crystal oscillators."

Quartz crystals are manufactured for frequencies from a few tens of kilohertz to tens of megahertz. More than two billion (2×10^9) crystals are manufactured annually. Most are used for consumer devices such as wristwatches, clocks, radios, computers, and cellphones. Quartz crystals are also found inside test and measurement equipment, such as counters, signal generators, and oscilloscopes.

History



Collection of several older crystals

Piezoelectricity was discovered by Jacques and Pierre Curie in 1880. Paul Langevin first investigated quartz resonators for use in sonar during World War I. The first crystal-controlled oscillator, using a crystal of Rochelle salt, was built in 1917 and patented in 1918 by Alexander M. Nicholson at Bell Telephone Laboratories, although his priority was disputed by Walter Guyton Cady. Cady built the first quartz crystal oscillator in 1921. Other early innovators in quartz crystal oscillators include G. W. Pierce and Louis Essen.

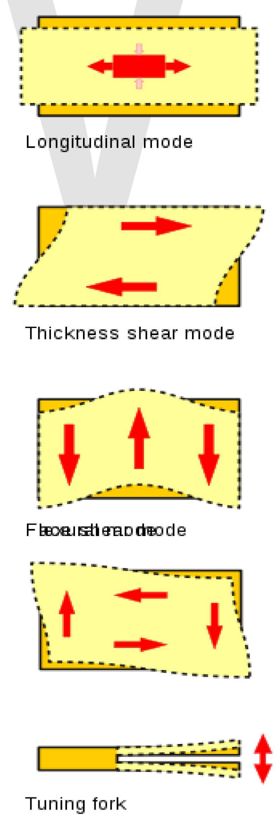
Quartz crystal oscillators were developed for high-stability frequency references during the 1920s and 1930s. By 1926 quartz crystals were used to control the frequency of radio broadcasting stations and were popular with amateur radio operators. In 1928, Warren Marrison (of Bell Telephone Laboratories) developed the first quartz crystal clock. This invention replaced the escapement and pendulum (as the timing reference), relying instead on the natural vibrations occurring in the quartz crystal as the oscillator. This improved timing accuracies to 1 sec in 30 years (or 30 ms/year).

A number of firms started producing quartz crystals for electronic use during this time. Using what are now considered primitive methods, about 100,000 crystal units were produced in the United States during 1939. During World War II, the demand for accurate frequency control of military and naval radios and radars spurred rapid development of the quartz crystal manufacturing industry. Suitable types of quartz became a critical war material, and much of it was eventually supplied by Brazil.

In 1968, Juergen Staudte invented a process for manufacturing quartz crystal oscillators while working at North American Aviation (now Rockwell). Staudte patented his invention, which used a photolithographic process that is similar to the way integrated circuits are made. In 1970 he left North American Aviation to start his own company, Statek, in Orange, California. Statek began manufacturing and marketing the quartz oscillators in 1971.

Although crystal oscillators still most commonly use quartz crystals, devices using other materials are becoming more common, such as ceramic resonators.

Operation



Crystal oscillation modes

A crystal is a solid in which the constituent atoms, molecules, or ions are packed in a regularly ordered, repeating pattern extending in all three spatial dimensions.

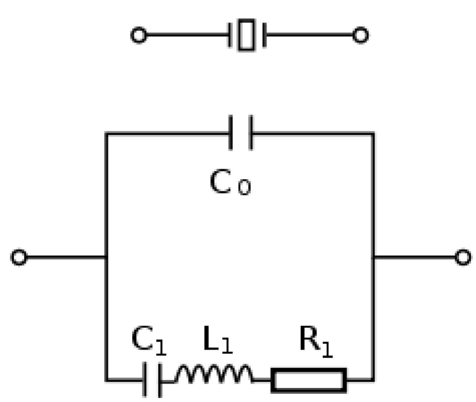
Almost any object made of an elastic material could be used like a crystal, with appropriate transducers, since all objects have natural resonant frequencies of vibration. For example, steel is very elastic and has a high speed of sound. It was often used in mechanical filters before quartz. The resonant frequency depends on size, shape, elasticity, and the speed of sound in the material. High-frequency crystals are typically cut in the shape of a simple, rectangular plate. Low-frequency crystals, such as those used in digital watches, are typically cut in the shape of a tuning fork. For applications not needing very precise timing, a low-cost ceramic resonator is often used in place of a quartz crystal.

When a crystal of quartz is properly cut and mounted, it can be made to distort in an electric field by applying a voltage to an electrode near or on the crystal. This property is known as piezoelectricity. When the field is removed, the quartz will generate an electric field as it returns to its previous shape, and this can generate a voltage. The result is that a quartz crystal behaves like a circuit composed of an inductor, capacitor and resistor, with a precise resonant frequency.

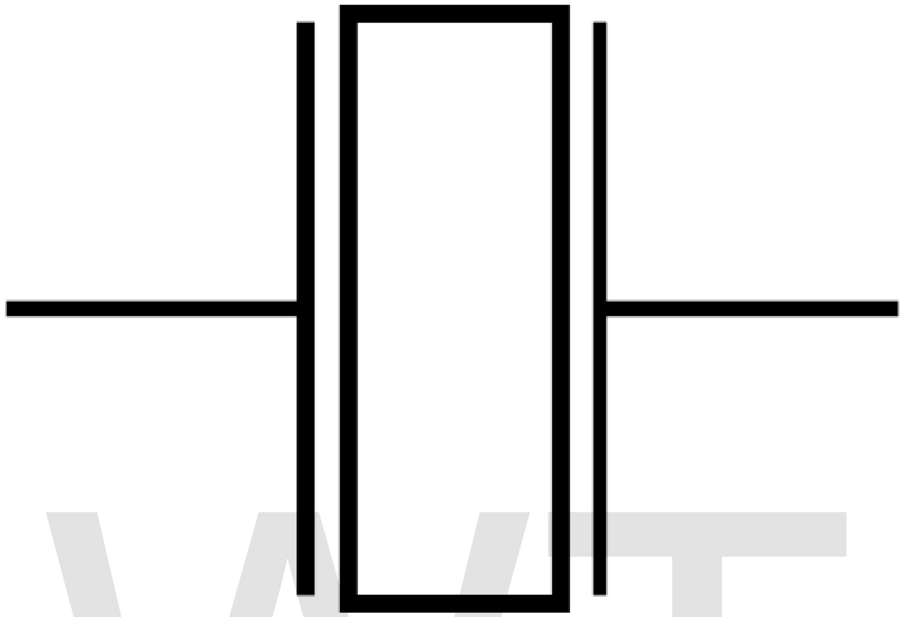
Quartz has the further advantage that its elastic constants and its size change in such a way that the frequency dependence on temperature can be very low. The specific characteristics will depend on the mode of vibration and the angle at which the quartz is cut (relative to its crystallographic axes). Therefore, the resonant frequency of the plate, which depends on its size, will not change much, either. This means that a quartz clock, filter or oscillator will remain accurate. For critical applications the quartz oscillator is mounted in a temperature-controlled container, called a crystal oven, and can also be mounted on shock absorbers to prevent perturbation by external mechanical vibrations.

Modeling

Electrical model



Schematic symbol and equivalent circuit for a quartz crystal in an oscillator



Electronic symbol for a piezoelectric crystal resonator

A quartz crystal can be modeled as an electrical network with a low impedance (series) and a high impedance (parallel) resonance point spaced closely together. Mathematically (using the Laplace transform) the impedance of this network can be written as:

$$Z(s) = \left(\frac{1}{s \cdot C_1} + s \cdot L_1 + R_1\right) || \left(\frac{1}{s \cdot C_0}\right)$$

or,

$$Z(s) = \frac{s^2 + s\frac{R_1}{L_1} + {\omega_s}^2}{(s \cdot C_0)[s^2 + s\frac{R_1}{L_1} + {\omega_p}^2]}$$

$$\Rightarrow \omega_s = \frac{1}{\sqrt{L_1 \cdot C_1}}, \quad \omega_p = \sqrt{\frac{C_1 + C_0}{L_1 \cdot C_1 \cdot C_0}} = \omega_s\sqrt{1 + \frac{C_1}{C_0}} \approx \omega_s\left(1 + \frac{C_1}{2C_0}\right) \quad (C_0 \gg C_1)$$

where s is the complex frequency ($s = j\omega$), ω_s is the series resonant frequency in radians per second and ω_p is the parallel resonant frequency in radians per second.

Adding additional capacitance across a crystal will cause the parallel resonance to shift downward. This can be used to adjust the frequency at which a crystal oscillates. Crystal manufacturers normally cut and trim their crystals to have a specified resonance frequency with a known 'load' capacitance added to the crystal. For example, a crystal intended for a 6 pF load has its specified parallel resonance frequency when a 6.0 pF capacitor is placed across it. Without this capacitance, the resonance frequency is higher.

Resonance modes

A quartz crystal provides both series and parallel resonance. The series resonance is a few kilohertz lower than the parallel one. Crystals below 300 MHz are generally operated between series and parallel resonance, which means that the crystal appears as an inductive reactance in operation. Any additional circuit capacitance will thus pull the frequency down. For a parallel resonance crystal to operate at its specified frequency, the electronic circuit has to provide a total parallel capacitance as specified by the crystal manufacturer.

Crystals above 30 MHz (up to >200 MHz) are generally operated at series resonance where the impedance appears at its minimum and equal to the series resistance. For these crystals the series resistance is specified (<100 Ω) instead of the parallel capacitance. To reach higher frequencies, a crystal can be made to vibrate at one of its overtone modes, which occur at multiples of the fundamental resonant frequency. Only odd numbered overtones are used. Such a crystal is referred to as a 3rd, 5th, or even 7th overtone crystal. To accomplish this, the oscillator circuit usually includes additional LC circuits to select the desired overtone.

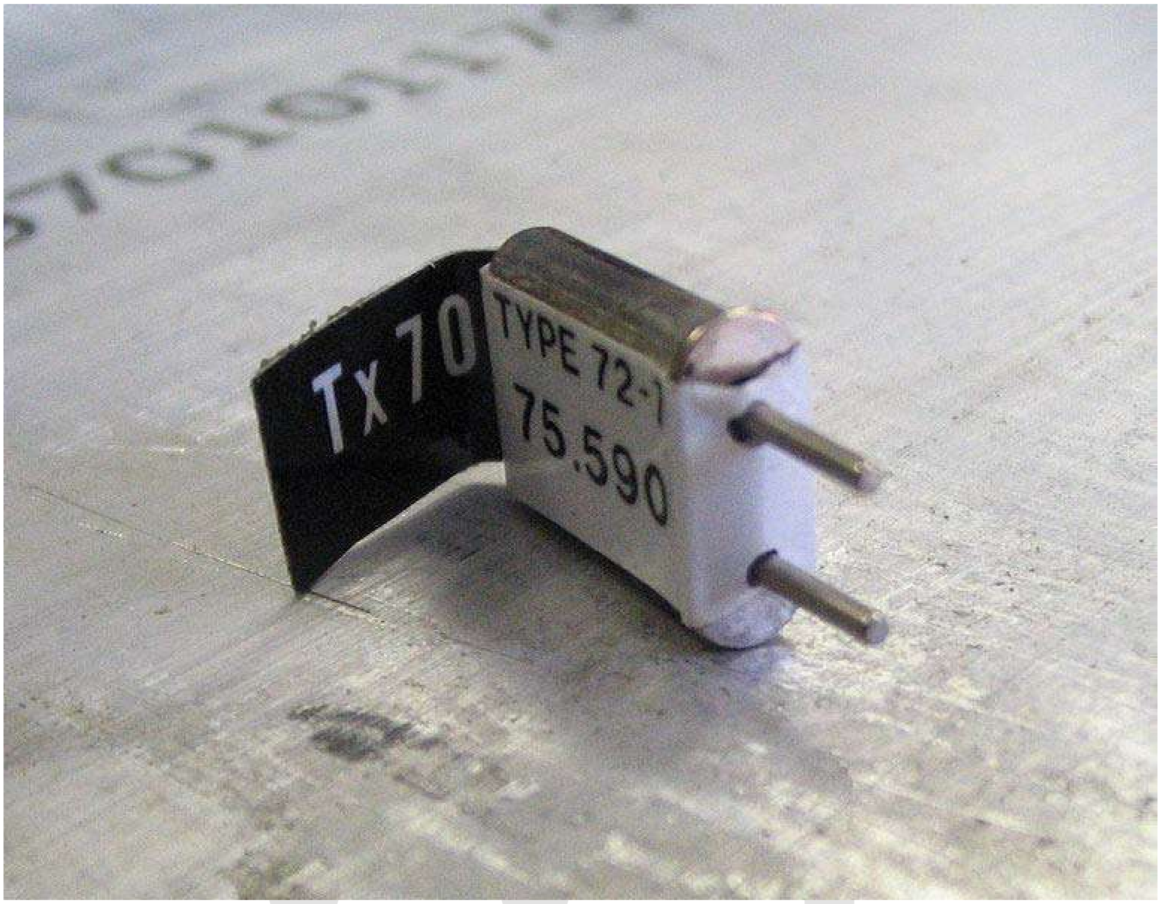
Temperature effects

A crystal's frequency characteristic depends on the shape or 'cut' of the crystal. A tuning fork crystal is usually cut such that its frequency over temperature is a parabolic curve centered around 25 °C. This means that a tuning fork crystal oscillator will resonate close to its target frequency at room temperature, but will slow down when the temperature either increases or decreases from room temperature. A common parabolic coefficient for a 32 kHz tuning fork crystal is −0.04 ppm/°C².

$$f = f_0[1 - 0.04\ \mathrm{ppm}(T - T_0)^2]$$

In a real application, this means that a clock built using a regular 32 kHz tuning fork crystal will keep good time at room temperature, lose 2 minutes per year at 10 degrees Celsius above (or below) room temperature and lose 8 minutes per year at 20 degrees Celsius above (or below) room temperature due to the quartz crystal.

Electrical oscillators



A crystal used in hobby radio control equipment to select frequency.

The crystal oscillator circuit sustains oscillation by taking a voltage signal from the quartz resonator, amplifying it, and feeding it back to the resonator. The rate of expansion and contraction of the quartz is the resonant frequency, and is determined by the cut and size of the crystal. When the energy of the generated output frequencies matches the losses in the circuit, an oscillation can be sustained.

An oscillator crystal has two electrically conductive plates, with a slice or tuning fork of quartz crystal sandwiched between them. During startup, the circuit around the crystal applies a random noise AC signal to it, and purely by chance, a tiny fraction of the noise will be at the resonant frequency of the crystal. The crystal will therefore start oscillating in synchrony with that signal. As the oscillator amplifies the signals coming out of the crystal, the signals in the crystal's frequency band will become stronger, eventually dominating the output of the oscillator. The narrow resonance band of the quartz crystal filters out all the unwanted frequencies.

The output frequency of a quartz oscillator can be either the fundamental resonance or a multiple of the resonance, called an overtone frequency. High frequency crystals are often designed to operate at third, fifth, or seventh overtones.

A major reason for the wide use of crystal oscillators is their high Q factor. A typical Q value for a quartz oscillator ranges from 10^4 to 10^6, compared to perhaps 10^2 for an LC oscillator. The maximum Q for a high stability quartz oscillator can be estimated as $Q = 1.6 \times 10^7/f$, where f is the resonance frequency in megahertz.

One of the most important traits of quartz crystal oscillators is that they can exhibit very low phase noise. In many oscillators, any spectral energy at the resonant frequency will be amplified by the oscillator, resulting in a collection of tones at different phases. In a crystal oscillator, the crystal mostly vibrates in one axis, therefore only one phase is dominant. This property of low phase noise makes them particularly useful in telecommunications where stable signals are needed, and in scientific equipment where very precise time references are needed.

Environmental changes of temperature, humidity, pressure, and vibration can change the resonant frequency of a quartz crystal, but there are several designs that reduce these environmental effects. These include the TCXO, MCXO, and OCXO (defined below). These designs (particularly the OCXO) often produce devices with excellent short-term stability. The limitations in short-term stability are due mainly to noise from electronic components in the oscillator circuits. Long term stability is limited by aging of the crystal.

Due to aging and environmental factors (such as temperature and vibration), it is difficult to keep even the best quartz oscillators within one part in 10^{10} of their nominal frequency without constant adjustment. For this reason, atomic oscillators are used for applications requiring better long-term stability and accuracy.

Spurious frequencies

For crystals operated in series resonance, significant (and temperature-dependent) spurious responses may be experienced. These responses typically appear some tens of kilohertz above the wanted series resonance. Even if the series resistances at the spurious resonances appear higher than the one at wanted frequency, the oscillator may lock at a spurious frequency (at some temperatures). This is generally avoided by using low impedance oscillator circuits to enhance the series resistance differences.

Spurious frequencies are also generated by subjecting the crystal to vibrations and are offset to the resonance frequency by the frequency of the vibrations. SC-cut crystals are designed to have their internal stresses compensated and are therefore less sensitive to vibrations. Mechanical properties of the mounting assembly are however more significant than the crystal cut.

Commonly used crystal frequencies

Crystal oscillator circuits are often designed around relatively few standard frequencies, such as 3.579545 MHz, 4.433619 MHz, 10 MHz, 14.318182 MHz, 17.734475 MHz, 20 MHz, 33.33 MHz, and 40 MHz. The popularity of the 3.579545 MHz crystals is due to low cost since they are used for NTSC color television receivers. Using frequency dividers, frequency multipliers and phase locked loop circuits, it is practical to derive a wide range of frequencies from one reference frequency. 14.318182 MHz (four times 3.579545 MHz) is used in computer video displays to generate a bitmapped video display for NTSC color monitors, such as the CGA used with the original IBM PC. (The IBM PC used 14.318182 MHz, divided by three, as its 4.77 MHz clock source, using one crystal for two purposes.) The 4.433619 MHz and 17.734475 MHz values are used in PAL color television equipment and devices intended to produce PAL signals.

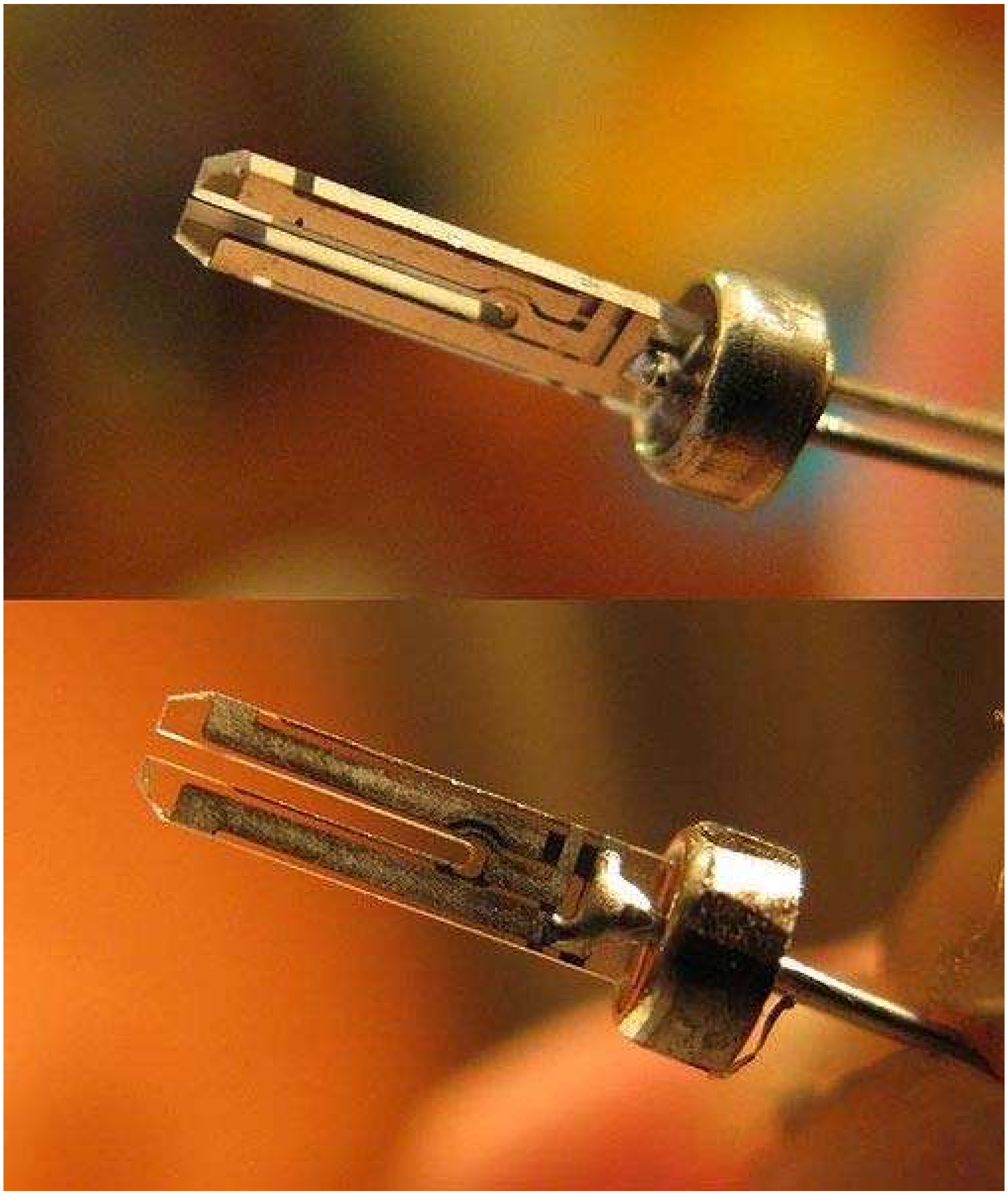
Crystals can be manufactured for oscillation over a wide range of frequencies, from a few kilohertz up to several hundred megahertz. Many applications call for a crystal oscillator frequency conveniently related to some other desired frequency, so hundreds of standard crystal frequencies are made in large quantities and stocked by electronics distributors.

WT

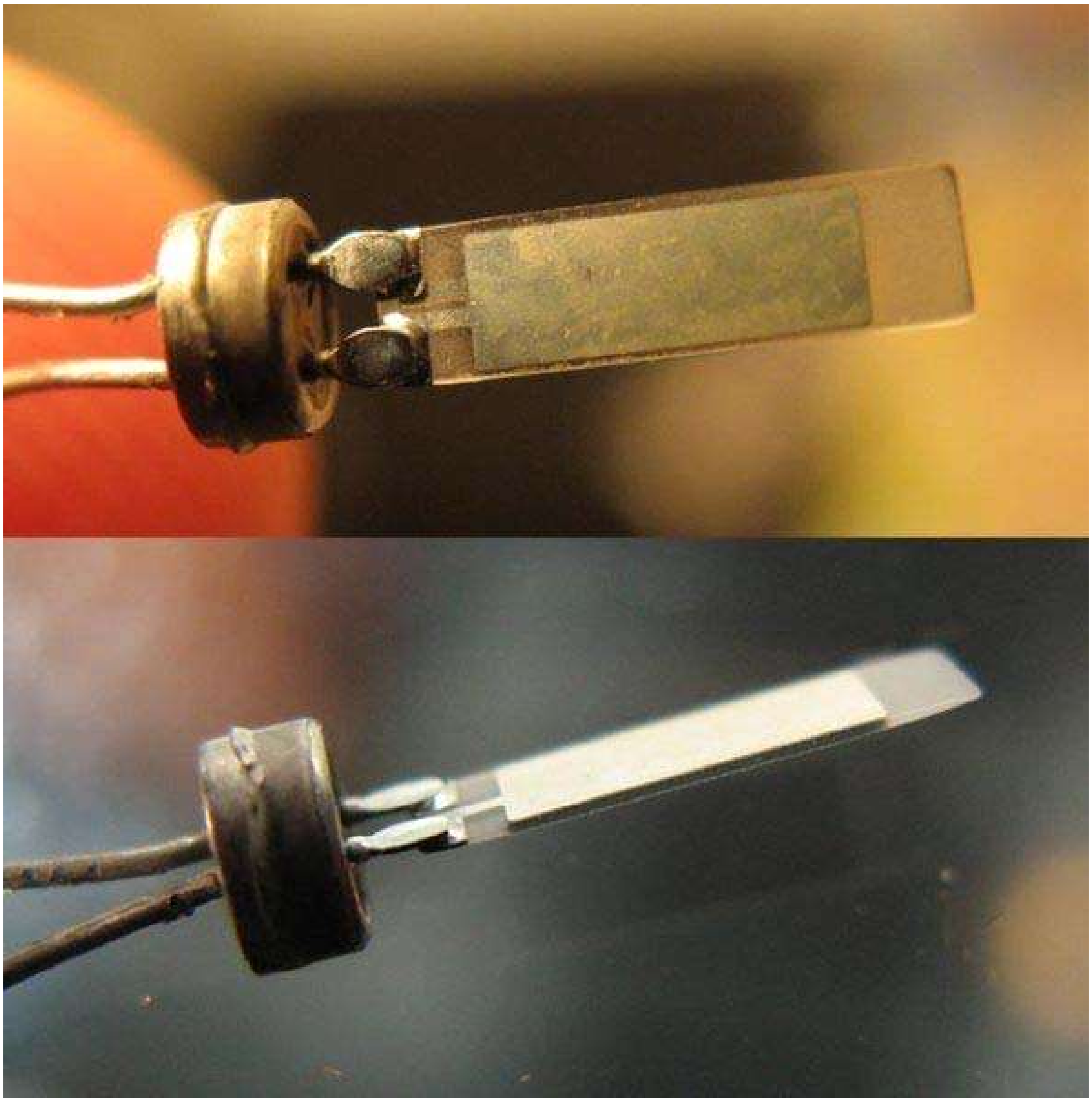
Crystal structures and materials

Cluster of natural quartz crystals

A synthetic quartz crystal grown by the hydrothermal synthesis, about 19 cm long and weighing about 127 grams.

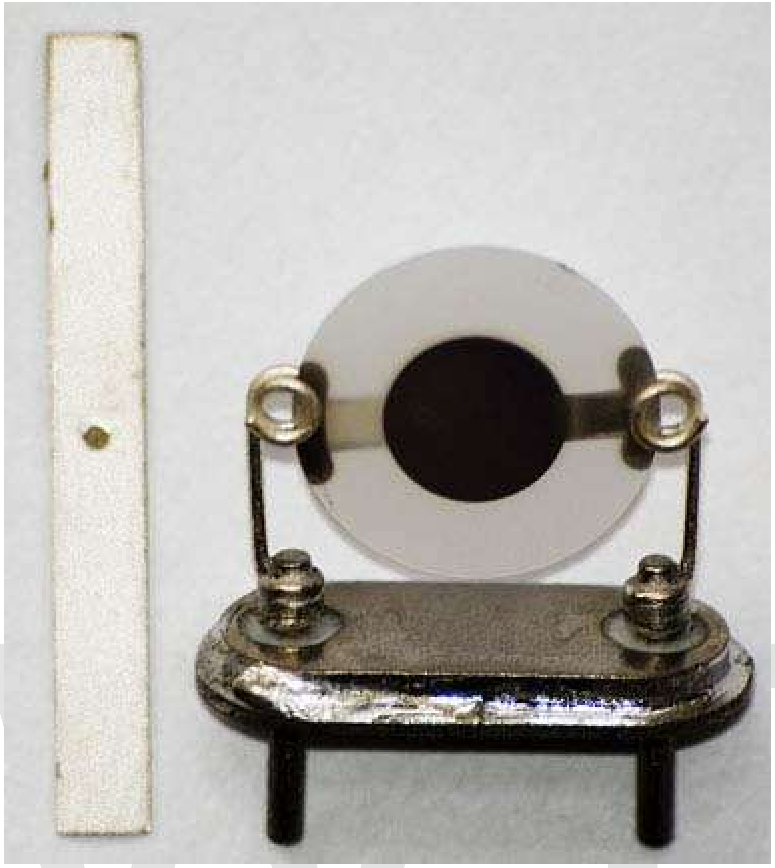


Tuning fork crystal



Simple quartz crystal

Inside construction of a modern high performance HC-49 package quartz crystal



Flexural and thickness shear crystals

The most common material for oscillator crystals is quartz. At the beginning of the technology, natural quartz crystals were used; now synthetic crystalline quartz grown by hydrothermal synthesis is predominant due to higher purity, lower cost, and more convenient handling. One of the few remaining uses of natural crystals is for pressure transducers in deep wells. During World War II and for some time afterwards, natural quartz was considered a strategic material by the USA. Large crystals were imported from Brazil. Raw "lascas", the source material quartz for hydrothermal synthesis, are imported to USA or mined locally by Coleman Quartz. The average value of as-grown synthetic quartz in 1994 was USD60/kg.

Two types of quartz crystals exist: left-handed and right-handed, differing in the optical rotation but identical in other physical properties. Both left and right-handed crystals can be used for oscillators, if the cut angle is correct. In manufacture, right-handed quartz is generally used. The SiO_4 tetrahedrons form parallel helixes; the direction of twist of the helix determines the left- or right-hand orientation. The helixes are aligned along the z-axis and merged together, sharing atoms. The mass of the helixes forms a mesh of small and large channels parallel to the z-axis; the large ones are large enough to allow some mobility of smaller ions and molecules through the crystal.

Quartz exists in several phases. At 573 °C at 1 atmosphere (and at higher temperatures and higher pressures) the α-quartz undergoes quartz inversion, transforms reversibly to β-quartz. The reverse process however is not entirely homogeneous and crystal twinning occurs. Care has to be taken during manufacture and processing to avoid the phase transformation. Other phases, e.g. the higher-temperature phases tridymite and cristobalite, are not significant for oscillators. All quartz oscillator crystals are the α-quartz type.

Infrared spectrophotometry is used as one of the methods for measuring the quality of the grown crystals. The wavenumbers 3585, 3500 and 3410 cm^{-1} are commonly used. The measured value is based on the absorption bands of the OH radical and the infrared Q value is calculated. The electronic grade crystals, grade C, have Q of 1.8 million or above; the premium grade B crystals have Q of 2.2 million, and special premium grade A crystals have Q of 3.0 million. The Q value is calculated only for the z region; crystals containing other regions can be adversely affected. Another quality indicator is the etch channel density; when the crystal is etched, tubular channels are created along linear defects. For processing involving etching, e.g. the wristwatch tuning fork crystals, low etch channel density is desirable. The etch channel density for swept quartz is about 10-100 and significantly more for unswept quartz. Presence of etch channels and etch pits degrades the resonator's Q and introduces nonlinearities.

Quartz crystals can be grown for specific purposes.

Crystals for AT-cut are the most common in mass production of oscillator materials; the shape and dimensions are optimized for high yield of the required wafers. High-purity quartz crystals are grown with especially low content of aluminium, alkali metal and other impurities and minimal defects; the low amount of alkali metals provides increased resistance to ionizing radiation. Crystals for wrist watches, for cutting the tuning fork 32768 Hz crystals, are grown with very low etch channel density.

Crystals for SAW devices are grown as flat, with large X-size seed with low etch channel density.

Special high-Q crystals, for use in highly stable oscillators, are grown at constant slow speed and have constant low infrared absorption along the entire Z axis. Crystals can be grown as Y-bar, with a seed crystal in bar shape and elongated along the Y axis, or as Z-plate, grown from a plate seed with Y-axis direction length and X-axis width. The region around the seed crystal contains a large number of crystal defects and should not be used for the wafers.

Crystals grow anisotropically; the growth along the Z axis is up to 3 times faster than along the X axis. The growth direction and rate also influences the rate of uptake of impurities. Y-bar crystals, or Z-plate crystals with long Y axis, have four growth regions usually called +X, -X, Z, and S. The distribution of impurities during growth is uneven; different growth areas contain different level of contaminants. The z regions are the purest, the small occasionally present s regions are less pure, the +x region is yet less pure, and the -x region has the highest level of impurities. The impurities have negative impact on radiation hardness, susceptibility to twinning, filter loss, and long and short term stability of the crystals. Different-cut seeds in different orientations may provide other kinds of growth regions. The growth speed of the -x direction is slowest due to the effect of adsorption of water molecules on the crystal surface; aluminium impurities suppress growth in two other directions. The content of aluminium is lowest in z region, higher in +x, yet higher in -x, and highest in s; the size of s regions also grows with increased amount of aluminium present. The content of hydrogen is lowest in z region, higher in +x region, yet higher in s region, and highest in -x. Aluminium inclusions transform to color centers with a gamma ray irradiation, causing darkening of the crystal proportional to the dose and level of impurities; presence of regions with different darkness reveals the different growth regions.

The dominant type of defect of concern in quartz crystals is the substitution of Al(III) for Si(IV) atom in the crystal lattice. The aluminium ion has an associated interstitial charge compensator present nearby, which can be a H^+ ion (attached to the nearby oxygen and forming a hydroxyl group, called Al-OH defect), Li^+ ion, Na^+ ion, K^+ ion (less common), or an electron hole trapped in a nearby oxygen atom orbital. The composition of the growth solution, whether it is based on lithium or sodium alkali compounds, determines the charge compensating ions for the aluminium defects. The ion impurities are of concern as they are not firmly bound and can migrate through the crystal, altering the local lattice elasticity and the resonant frequency of the crystal. Other common impurities of concern are e.g. iron(III) (interstitial), fluorine, boron(III), phosphorus(V) (substitution), titanium(IV) (substitution, universally present in magmatic quartz, less common in hydrothermal quartz), and germanium(IV) (substitution). Sodium and iron ions can cause inclusions of acnite and elemeusite crystals. Inclusions of water may be present in fast-grown crystals; interstitial water molecules are abundant near the crystal seed. Another defect of importance is the hydrogen containing growth defect, when instead of a Si-O-Si structure a pair of Si-OH HO-Si groups is formed; essentially a hydrolyzed bond. Fast-grown crystals contain more hydrogen defects than slow-grown ones. These growth defects source as supply of hydrogen ions for radiation-induced processes and forming Al-OH defects. Germanium impurities tend to trap electrons created during irradiation; the alkali metal cations then migrate towards the negatively charged center and form a stabilizing complex. Matrix defects can be also present; oxygen vacancies, silicon vacancies (usually compensated by 4 hydrogens or 3 hydrogens and a hole), peroxy groups, etc. Some of the defects produce localized levels in the forbidden band, serving as charge traps; Al(III) and B(III) typically serve as hole traps while electron vacancies, titanium, germanium, and phosphorus atoms serve as electron traps. The trapped charge carriers can be released by heating; their recombination is the cause of thermoluminescence.

The mobility of interstitial ions depends strongly on temperature. Hydrogen ions are mobile down to 10 K, but alkali metal ions become mobile only at temperatures around and above 200 K. The hydroxyl defects can be measured by near-infrared spectroscopy. The trapped holes can be measured by electron spin resonance. The Al-Na^+ defects show as an acoustic loss peak due to their stress-induced motion; the Al-Li^+ defects do not form a potential well so are not detectable this way. Some of the radiation induced defects during their thermal annealing produce thermoluminescence; defects related to aluminium, titanium, and germanium can be distinguished.

Swept crystals are crystals that undergone a solid-state electrodiffusion purification process. Sweeping involves heating the crystal above 500 °C in a hydrogen-free atmosphere, and the voltage gradient of at least 1 kilovolt/cm, for several (usually over 12) hours. The migration of impurities and the gradual replacement of alkali metal ions with hydrogen (when swept in air) or electron holes (when swept in vacuum) causes a weak electric current through the crystal; decay of this current to a constant value signals end of the process. The crystal is then left to cool, while the electric field is maintained. The impurities are concentrated at the cathode region of the crystal, which is cut off afterwards and discarded. Swept crystals have increased resistance to radiation, as the dose effects are dependent on the level of alkali metal impurities; they are suitable for use in devices exposed to ionizing radiation, e.g. for nuclear and space technology. Sweeping under vacuum at higher temperatures and higher field strengths yields yet more radiation-hard crystals. The level and character of impurities can be measured by infrared spectroscopy. Quartz can be swept in both α and β phase; sweeping in β phase is faster, but the phase transition may induce twinning. Twinning can be mitigated by subjecting the crystal to compression stress in the X direction, or an AC or DC electric field along the X axis while the crystal cools through the phase transformation temperature region.

Sweeping can be also used to introduce one kind of an impurity into the crystal. Lithium, sodium, and hydrogen swept crystals are used for e.g. studying quartz behavior.

Very small crystals for high fundamental mode frequencies can be manufactured by photolithography.

Crystals can be adjusted to exact frequency by laser trimming. A technique used in the world of amateur radio for slight decrease of the crystal frequency may be achieved by exposing crystals with silver electrodes to vapors of iodine, which causes a slight mass increase on the surface by forming a thin layer of silver iodide; such crystals however had problematic long-term stability. Another method commonly used is electrochemical increase or decrease of silver electrode thickness by submerging resonator in lapis solved in water, citric acid in water, or water with salt, and using resonator as one electrode, and small silver electrode as another.

By choosing direction of current, one can either increase, or decrease mass of electrodes. Details were published in "Radio" magazine (3/1978) by UB5LEV.

Raising frequency by scratching off parts of the electrodes is advised against, as this may damage the crystal and lower its Q factor. Capacitor trimmers can be also used for frequency adjustment of the oscillator circuit.

Some other piezoelectric materials than quartz can be employed; e.g. single crystals of lithium tantalate, lithium niobate, lithium borate, berlinite, gallium arsenide, lithium

WT

tetraborate, aluminium phosphate, bismuth germanium oxide, polycrystalline zirconium titanate ceramics, high-alumina ceramics, silicon-zinc oxide composite, or dipotassium tartrate; some materials may be more suitable for specific applications. An oscillator crystal can be also manufactured by depositing the resonator material on the silicon chip surface. Crystals of gallium phosphate, langasite, langanite and langanate are about 10 times more pullable than the corresponding quartz crystals, and are used in some VCXO oscillators.

Stability and aging

The frequency stability is determined by the crystal's Q. It is inversely dependent on the frequency, and on the constant that is dependent on the particular cut. Other factors influencing Q are the overtone used, the temperature, the level of driving of the crystal, the quality of the surface finish, the mechanical stresses imposed on the crystal by bonding and mounting, the geometry of the crystal and the attached electrodes, the material purity and defects in the crystal, type and pressure of the gas in the enclosure, interfering modes, and presence and absorbed dose of ionizing and neutron radiation.

Temperature influences the operating frequency; various forms of compensation are used, from analog compensation (TCXO) and microcontroller compensation (MCXO) to stabilization of the temperature with a crystal oven (OCXO). The crystals possess temperature hysteresis; the frequency at a given temperature achieved by increasing the temperature is not equal to the frequency on the same temperature achieved by decreasing the temperature. The temperature sensitivity depends primarily on the cut; the temperature compensated cuts are chosen as to minimize frequency/temperature dependence. Special cuts can be made with a linear temperature characteristics; the LC cut is used in quartz thermometers. Other influencing factors are the overtone used, the mounting and electrodes, impurities in the crystal, mechanical strain, crystal geometry, rate of temperature change, thermal history (due to hysteresis), ionizing radiation, and drive level.

Crystals tend to suffer anomalies in their frequency/temperature and resistance/temperature characteristics, known as activity dips. These are small downward (in frequency) or upward (in resistance) excursions localized at certain temperatures, with their temperature position dependent on the value of the load capacitors.

Mechanical stresses also influence the frequency. The stresses can be induced by mounting, bonding, and application of the electrodes, by differential thermal expansion of the mounting, electrodes, and the crystal itself, by differential thermal stresses when there is a temperature gradient present, by expansion or shrinkage of the bonding materials during curing, by the air pressure that is transferred to the ambient pressure within the crystal enclosure, by the stresses of the crystal lattice itself (nonuniform growth, impurities, dislocations), by the surface imperfections and damage caused during manufacture, and by the action of gravity on the mass of the crystal; the frequency can therefore be influenced by position of the crystal. Other dynamic stress inducing factors are shocks, vibrations, and acoustic noise. Some cuts are less sensitive to stresses; the SC

(Stress Compensated) cut is an example. Atmospheric pressure changes can also introduce deformations to the housing, influencing the frequency by changing stray capacitances.

Atmospheric humidity influences the thermal transfer properties of air, and can change electrical properties of plastics by diffusion of water molecules into their structure, altering the dielectric constants and electrical conductivity.

Other factors influencing the frequency are the power supply voltage, load impedance, magnetic fields, electric fields (in case of cuts that are sensitive to them, e.g. SC), the presence and absorbed dose of γ-particles and ionizing radiation, and the age of the crystal.

Crystals undergo slow gradual change of frequency with time, known as aging. There are many mechanisms involved. The mounting and contacts may undergo relief of the build-in stresses. Molecules of contamination either from the residual atmosphere, outgassed from the crystal, electrodes or packaging materials, or introduced during sealing the housing can be adsorbed on the crystal surface, changing its mass; this effect is exploited in quartz crystal microbalances. The composition of the crystal can be gradually altered by outgassing, diffusion of atoms of impurities or migrating from the electrodes, or the lattice can be damaged by radiation. Slow chemical reactions may occur on or in the crystal, or on the inner surfaces of the enclosure. Electrode material, e.g. chromium or aluminium, can react with the crystal, creating layers of metal oxide and silicon; these interface layers can undergo changes in time. The pressure in the enclosure can change due to varying atmospheric pressure, temperature, leaks, or outgassing of the materials inside. Factors outside of the crystal itself are e.g. aging of the oscillator circuitry (and e.g. change of capacitances), and drift of parameters of the crystal oven. External atmosphere composition can also influence the aging; hydrogen can diffuse through nickel housing. Helium can cause similar issues when it diffuses through glass enclosures of rubidium standards.

Gold is a favored electrode material for low-aging resonators; its adhesion to quartz is strong enough to maintain contact even at strong mechanical shocks, but weak enough to not support significant strain gradients (unlike chromium, aluminium, and nickel). Gold also does not form oxides; it adsorbs organic contaminants from the air, but these are easy to remove. However, gold alone can undergo delamination; a layer of chromium is therefore sometimes used for improved binding strength. Silver and aluminium are often used as electrodes; however both form oxide layers with time that increases the crystal mass and lowers frequency. Silver can be passivated by exposition to iodine vapors, forming a layer of silver iodide. Aluminium oxidizes readily but slowly, until about 5 nm thickness is reached; increased temperature during artificial aging does not significantly increase the oxide forming speed; a thick oxide layer can be formed during manufacture by anodizing. Exposition of silver-plated crystal to iodine vapors can be also used in amateur conditions for lowering the crystal frequency slightly; the frequency can be also increased by scratching off parts of the electrodes, but that carries risk of damage to the crystal and loss of Q.

A DC voltage bias between the electrodes can accelerate the initial aging, probably by induced diffusion of impurities through the crystal. Placing a capacitor in the series with the crystal and a several megaohms resistor in parallel can minimize such voltages.

Crystals suffer from minor short-term frequency fluctuations as well. The main causes of such noise are e.g. thermal noise (which limits the noise floor), phonon scattering (influenced by lattice defects), adsorption/desorption of molecules on the surface of the crystal, noise of the oscillator circuits, mechanical shocks and vibrations, acceleration and orientation changes, temperature fluctuations, and relief of mechanical stresses. The short-term stability is measured by four main parameters: Allan variance (the most common one specified in oscillator datasheets), phase noise, spectral density of phase deviations, and spectral density of fractional frequency deviations. The effects of acceleration and vibration tend to dominate the other noise sources; surface acoustic wave devices tend to be more sensitive than bulk acoustic wave (BAW) ones, and the stress-compensated cuts are even less sensitive. The relative orientation of the acceleration vector to the crystal dramatically influences the crystal's vibration sensitivity. Mechanical vibration isolation mountings can be used for high-stability crystals.

Crystals are sensitive to shock. The mechanical stress causes short-time change in the oscillator frequency due to the stress-sensitivity of the crystal, and can introduce a permanent change of frequency due to shock-induced changes of mounting and internal stresses (if the elastic limits of the mechanical parts are exceeded), desorption of contamination from the crystal surfaces, or change in parameters of the oscillator circuit. High magnitudes of shocks may tear the crystals off their mountings (especially the case of large low-frequency crystals suspended on thin wires), or cause cracking of the crystal. Crystals free of surface imperfections are highly shock-resistant; chemical polishing can produce crystals able to survive tens of thousands g.

Phase noise plays significant role in frequency synthesis systems using frequency multiplication; a multiplication of a frequency by N increases the phase noise by N^2. A frequency multiplication by 10 times multiplies the phase error by 100 times. This can be disastrous for systems employing e.g. PLL or FSK technologies.

Crystals are somewhat sensitive to radiation damage. Natural quartz is much more sensitive than artificially grown crystals, and sensitivity can be further reduced by sweeping the crystal - heating the crystal to at least 400 °C in hydrogen-free atmosphere in electric field of at least 500 V/cm for at least 12 hours. Such swept crystals have very low response to steady ionizing radiation. Some Si(IV) atoms are replaced with Al(III) impurities, each having a compensating Li^+ or Na^+ cation nearby. Ionization produces electron-hole pairs; the holes are trapped in the lattice near the Al atom, the resulting Li and Na atoms are loosely trapped along the Z axis; the change of the lattice near the Al atom and the corresponding elastic constant then causes a corresponding change in frequency. Sweeping removes the Li^+ and Na^+ ions from the lattice, reducing this effect. The Al^{3+} site can also trap hydrogen atoms. All crystals have transient negative frequency shift after exposition to an X-ray pulse; the frequency then shifts gradually back; natural quartz reaches stable frequency after 10–1000 seconds, with negative offset to pre-irradiation frequency, artificial crystals return to frequency slightly lower or higher than pre-irradiation, swept crystals anneal virtually back to original frequency. The annealing is faster at higher temperatures. Sweeping under vacuum at higher temperatures and field strength can further reduce the crystal's response to X-ray pulses. Series resistance of unswept crystals increases after an X-ray dose, and anneals back to a somewhat higher value for a natural quartz (requiring a corresponding gain reserve in the circuit) and back to pre-irradiation value for synthetic crystals. Series resistance of swept crystals is unaffected. Increase of series resistance degrades Q; too high increase can stop the oscillations. Neutron radiation induces frequency changes by introducing dislocations into the lattice by knocking out atoms, a single fast neutron can produce many defects; the SC and AT cut frequency increases roughly linearly with absorbed neutron dose, while the frequency of the BT cuts decreases. Neutrons also alter the temperature-frequency characteristics. Frequency change at low ionizing radiation doses is proportionally higher than for higher doses. High-intensity radiation can stop the oscillator by inducing photoconductivity in the crystal and transistors; with a swept crystal and properly designed circuit the oscillations can restart within 15 microseconds after the radiation burst. Quartz crystals with high level of alkali metal impurities lose Q with irradiation; Q of swept artificial crystals is unaffected. Irradiation with higher doses (over 10^5 rad) lowers sensitivity to subsequent doses. Very low radiation doses (below 300 rad) have disproportionally higher effect, but this nonlinearity saturates at higher doses. At very high doses, the radiation response of the crystal saturates as well, due to finite number of impurity sites that can be affected.

Magnetic fields have low effect on the crystal itself, as quartz is diamagnetic; eddy currents or AC voltages can however be induced into the circuits, and magnetic parts of the mounting and housing may be influenced.

After the power-up, the crystals take several seconds to minutes to "warm up" and stabilize their frequency. The oven-controlled OCXOs require usually 3–10 minutes for heating up and reaching thermal equilibrium, the oven-less oscillators stabilize in several seconds as the few milliwatts dissipated in the crystal cause a small but noticeable level of internal heating.

Crystals have no inherent failure mechanisms; some are operating in devices for decades. Failures may be however introduced by faults in bonding, leaky enclosures, corrosion, frequency shift by aging, breaking the crystal by too high mechanical shock, or radiation induced damage when nonswept quartz is used. Crystals can be also damaged by overdriving.

The crystals have to be driven at the appropriate drive level. While AT cuts tend to be fairly forgiving, and only their electrical parameters, stability and aging characteristics are degraded when overdriven, low-frequency crystals, especially flexural-mode ones, may fracture at too high drive levels. The drive level is specified as the amount of power dissipated in the crystal. The appropriate drive levels are about 5 microwatts for flexural modes up to 100 kHz, 1 microwatt for fundamental modes at 1-4 MHz, 0.5 microwatts for fundamental modes 4-20 MHz, and 0.5 microwatts for overtone modes at 20-200 MHz. Too low drive level may cause problems with starting the oscillator. Low drive levels are better for higher stability and lower power consumption of the oscillator. Higher drive levels, in turn, reduce the impact of noise by increasing the signal-to-noise ratio.

The stability of AT cut crystals decreases with increasing frequency. For more accurate higher frequencies it is better to use a crystal with lower fundamental frequency, operating at an overtone.

Aging decreases logarithmically with time, the highest changes occurring shortly after manufacture. Artificially aging the crystal by its prolonged storage at between 85-125 °C can be done for increasing long-term stability.

A badly designed oscillator circuit may suddenly become oscillating on an overtone; in 1972, a train in Fremont, CA crashed. An inappropriate value of the tank capacitor caused the crystal in a control board to be overdriven, jumping to an overtone, and causing the train to speed up instead of slowing down.

Crystal cuts

The resonator plate can be cut from the source crystal in many different ways. The orientation of the cut influences the crystal's aging characteristics, frequency stability, thermal characteristics, and other parameters. These cuts operate at bulk acoustic wave (BAW); for higher frequencies, surface acoustic wave (SAW) devices are employed.

Image of several crystal cuts

Cut	Frequency range	Mode	Angles	Description
AT	0.5-300 MHz	thickness shear (c-mode, slow quasi-shear)	35°15', 0° (<25 MHz) 35°18', 0°(>10 MHz)	The most common cut, developed in 1934. The plate contains the crystal's x axis and is inclined by 35°15' from the z (optic) axis. The frequency-temperature curve is a sine-shaped curve with inflection point at around 25-35 °C. Has frequency constant 1.661 MHz-mm. Most (estimated over 90%) of all crystals are this variant. Used for oscillators operating in wider temperature range, for range of 0.5 to 200 MHz; also used in oven-controlled oscillators. Sensitive to mechanical stresses, whether caused by external forces or by temperature gradients. Thickness-shear crystals typically

				operate in fundamental mode at 1-30 MHz, 3rd overtone at 30-90 MHz, and 5th overtone at 90-150 MHz; according to other source they can be made for fundamental mode operation up to 300 MHz, though that mode is usually used only to 100 MHz and according to yet another source the upper limit for fundamental frequency of the AT cut is limited to 40 MHz for small diameter blanks. Can be manufactured either as a conventional round disk, or as a strip resonator; the latter allows much smaller size. The thickness of the quartz blank is about (1.661 mm)/(frequency in MHz), with the frequency somewhat shifted by further processing. The third overtone is about 3 times the fundamental frequency; the overtones are higher than the equivalent multiple of the fundamental frequency by about 25 kHz per overtone. Crystals designed for operating in overtone modes have to be specially processed for plane parallelism and surface finish for the best performance at a given overtone frequency.
SC	0.5-200 MHz	thickness shear	35°15', 21°54'	A special cut (Stress Compensated) developed in 1974, is a double-rotated cut (35°15' and 21°54') for oven-stabilized oscillators with low phase noise and good aging characteristics. Less sensitive to mechanical stresses. Has faster warm-up speed, higher Q, better close-in phase noise, less sensitivity to spatial orientation against the vector of gravity, and less sensitivity to vibrations. Its frequency constant is 1.797 MHz-mm. Coupled modes are worse than the AT cut, resistance tends to be higher; much more care is required to convert between overtones. Operates at the same frequencies as the AT cut. The frequency-temperature curve is a third order downward parabola with

WT

				inflection point at 95 °C and much lower temperature sensitivity than the AT cut. Suitable for OCXOs in e.g. space and GPS systems. Less available than AT cut, more difficult to manufacture; the order-of-magnitude improvement of parameters is traded for an order of magnitude tighter crystal orientation tolerances. Aging characteristics are 2 to 3 times better than of the AT cuts. Less sensitive to drive levels. Far fewer activity dips. Less sensitive to plate geometry. Requires an oven, does not operate well at ambient temperatures as the frequency rapidly falls off at lower temperatures. Has several times lower motional capacitance than the corresponding AT cut, reducing the possibility to adjust the crystal frequency by attached capacitor; this restricts usage in conventional TCXO and VCXO devices, and other applications where the frequency of the crystal has to be adjustable. The temperature coefficients for the fundamental frequency is different than for its third overtone; when the crystal is driven to operate on both frequencies simultaneously, the resulting beat frequency can be used for temperature sensing in e.g. microcomputer-compensated crystal oscillators. Sensitive to electric fields. Sensitive to air damping, to obtain optimum Q it has to be packaged in vacuum. Temperature coefficient for b-mode is -25 ppm/°C, for dual mode 80 to over 100 ppm/°C.
BT	0.5-200 MHz	thickness shear (b-mode, fast quasi-shear)	-49°8', 0°	A special cut, similar to AT cut, except the plate is cut at 49° from the z axis. Operates in thickness shear mode, in b-mode (fast quasi-shear). It has well known and repeatable characteristics. Has frequency constant 2.536 MHz-mm. Has poorer temperature characteristics than the AT cut. Due to the higher

				frequency constant, can be used for crystals with higher frequencies than the AT cut, up to over 50 MHz.
IT		thickness shear		A special cut, is a double-rotated cut with improved characteristics for oven-stabilized oscillators. Operates in thickness shear mode. The frequency-temperature curve is a third order downward parabola with inflection point at 78 °C. Rarely used. Has similar performance and properties to the SC cut, more suitable for higher temperatures.
FC		thickness shear		A special cut, a double-rotated cut with improved characteristics for oven-stabilized oscillators. Operates in thickness shear mode. The frequency-temperature curve is a third order downward parabola with inflection point at 52 °C. Rarely used. Employed in oven-controlled oscillators; the oven can be set to lower temperature than for the AT/IT/SC cuts, to the beginning of the flat part of the temperature-frequency curve (which is also broader than of the other cuts); when the ambient temperature reaches this region, the oven switches off and the crystal operates at the ambient temperature, while maintaining reasonable accuracy. This cut therefore combines the power saving feature of allowing relatively low oven temperature with reasonable stability at higher ambient temperatures.
AK		thickness shear		a double rotated cut with better temperature-frequency characteristics than AT and BT cuts and with higher tolerance to crystallographic orientation than the AT, BT, and SC cuts (by factor 50 against a standard AT cut, according to calculations). Operates in thickness-shear mode.
CT	300–900 kHz	face shear	38°, 0°	The frequency-temperature curve is a downward parabola.
DT	75–	face shear	-52°, 0°	Similar to CT cut. The frequency-

	800 kHz			temperature curve is a downward parabola. The temperature coefficient is lower than the CT cut; where the frequency range permits, DT is preferred over CT.
SL		face-shear	-57°, 0°	
GT	0.1-3 MHz	width-extensional	51°7'	Its temperature coefficient between -25..+75 °C is near-zero, due to cancelling effect between two modes.
E, **5°X**	50–250 kHz	longitudal		Has reasonably low temperature coefficient, widely used for low-frequency crystal filters.
MT	40–200 kHz	longitudal		
ET			66°30'	
FT			-57°	
NT	8–130 kHz	length-width flexure (bending)		
XY, **tuning fork**	3–85 kHz	length-width flexure		The dominant low-frequency crystal, as it is smaller than other low-frequency cuts, less expensive, has low impedance and low Co/C1 ratio. The chief application is the 32.768 kHz RTC crystal. Its second overtone is about six times the fundamental frequency.
H	8–130 kHz	length-width flexure		Used extensively for wideband filters. The temperature coefficient is linear.
J	1–12 kHz	length-thickness flexure		J cut is made of two quartz plates bonded together, selected to produce out of phase motion for a given electrical field.
RT				A double rotated cut.
SBTC				A double rotated cut.
TS				A double rotated cut.
X 30°				A double rotated cut.
LC		thickness shear	11.17°/9.39°	A double rotated cut ("Linear Coefficient") with a linear temperature-frequency response; can be used as a sensor in crystal thermometers. Temperature coefficient is 35.4 ppm/°C.
AC			31°	Temperature-sensitive, can be used as a sensor. Single mode with steep

WT

		frequency-temperature characteristics. Temperature coefficient is 20 ppm/°C.
BC	-60°	Temperature-sensitive.
NLSC		Temperature-sensitive. Temperature coefficient is about 14 ppm/°C.
Y		Temperature-sensitive, can be used as a sensor. Single mode with steep frequency-temperature characteristics. The plane of the plate is perpendicular to the Y axis of the crystal. Also called **parallel** or **30-degree**. Temperature coefficient is about 90 ppm/°C.
X		Used in one of the first crystal oscillators in 1921 by W.G. Cady, and as a 50 kHz oscillator in the first crystal clock by Horton and Marrison in 1927. The plane of the plate is perpendicular to the X axis of the crystal. Also called **perpendicular**, **normal**, **Curie**, **zero-angle**, or **ultrasonic**.

The T in the cut name marks a temperature-compensated cut, a cut oriented in a way that the temperature coefficients of the lattice are minimal; the FC and SC cuts are also temperature-compensated.

The high frequency cuts are mounted by their edges, usually on springs; the stiffness of the spring has to be optimal, as too stiff could transfer mechanical shocks to the crystal and cause its breaking, and too little stiffness may allow the crystal to collide with the inside of the package when subjected to a mechanical shock, and break. Strip resonators, usually AT cuts, are smaller and therefore less sensitive to mechanical shocks. At the same frequency and overtone, the strip will have less pullability, higher resistance, and higher temperature coefficient.

The low frequency cuts are mounted at the nodes where they are virtually motionless; thin wires are attached at such points on each side between the crystal and the leads. The large mass of the crystal suspended the thin wires makes the assembly sensitive to mechanical shocks and vibrations.

The crystals are usually mounted in hermetically sealed glass or metal cases, filled with a dry and inert atmosphere, usually vacuum, nitrogen, or helium. Plastic housings can be used as well, but those are not hermetic and another secondary sealing has to be built around the crystal.

Several resonator configurations are possible, in addition to the classical way of directly attaching leads to the crystal. E.g. the **BVA resonator** (Boîtier à Vieillissement

Amélioré, Enclosure with Improved Aging), developed in 1976; the parts that influence the vibrations are machined from a single crystal (which reduces the mounting stress), and the electrodes are deposited not on the resonator itself but on inner sides of two condenser discs made of adjacent slices of the quartz from the same bar, forming a three-layer sandwich with no stress between the electrodes and the vibrating element. The gap between the electrodes and the resonator act as two small series capacitors, making the crystal less sensitive to circuit influences. The architecture eliminates the effects of the surface contacts between the electrodes, the constraints in the mounting connections, and the issues related to ion migration from the electrodes into the lattice of the vibrating element. The resulting configuration is rugged, resistant to shock and vibration, resistant to acceleration and ionizing radiation, and has improved aging characteristics. AT cut is usually used, though SC cut variants exist as well. BVA resonators are often used in spacecraft applications.

In 1930s to 1950s it was fairly common for people to adjust the frequency of the crystals by manual grinding. The crystals were ground using a fine abrasive slurry, or even a toothpaste, to increase their frequency. A slight decrease by 1–2 kHz when the crystal was overground was possible by marking the crystal face with a pencil lead, for the price of lowering the Q.

The frequency of the crystal is slightly adjustable ("pullable") by modifying the attached capacitances. A varactor, a diode with capacitance depending on applied voltage, is often used in voltage-controlled crystal oscillators, VCXO. The crystal cuts are usually AT or rarely SC, and operate in fundamental mode; the amount of available frequency deviation is inversely proportional to the square of the overtone number, so a third overtone will have only one ninth of the pullability of the fundamental mode. SC cuts, while more stable, are significantly less pullable.

Circuit notations and abbreviations

On electrical schematic diagrams, *crystals* are designated with the class letter *Y* (Y1, Y2, etc.) Oscillators, whether they are crystal *oscillators* or other, are designated with the class letter *G* (G1, G2, etc.) On occasion, one may see a crystal designated on a schematic with *X* or *XTAL*, or a crystal oscillator with *XO*, but these forms are deprecated.

Crystal oscillator types and their abbreviations:

- **ATCXO** — Analog temperature controlled crystal oscillator
- **CDXO** — Calibrated dual crystal oscillator
- **DTCXO** — Digital temperature compensated crystal oscillator
- **EMXO** — Evacuated miniature crystal oscillator
- **GPSDO** — Global positioning system disciplined oscillator
- **MCXO** — Microcomputer-compensated crystal oscillator
- **OCVCXO** — oven-controlled voltage-controlled crystal oscillator
- **OCXO** — Oven-controlled crystal oscillator

- **RbXO** — Rubidium crystal oscillators (RbXO), a crystal oscillator (can be an MCXO) synchronized with a built-in rubidium standard which is run only occasionally to save power
- **TCVCXO** — Temperature-compensated voltage-controlled crystal oscillator
- **TCXO** — Temperature-compensated crystal oscillator
- **TMXO -** Tactical miniature crystal oscillator
- **TSXO** — Temperature-sensing crystal oscillator, an adaptation of the TCXO
- **VCTCXO** — Voltage-controlled temperature-compensated crystal oscillator
- **VCXO** — Voltage-controlled crystal oscillator

WT