



SNS COLLEGE OF ENGINEERING

(Autonomous)

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING



MICROPROCESSORS & ADVANCED MICRO CONTROLLERS

Fixed vs Floating point Processors

Prepared by,

K.Sangeetha

Assistant Professor/ECE

SNS College of Engineering



Introduction

Digital Signal Processing can be divided into two categories,

1. **Fixed point** and
2. **Floating point.**

These refer to the format used to store and manipulate numbers within the devices. Fixed point DSPs usually represent each number with a minimum of 16 bits, although a different length can be used.

In comparison, floating point DSPs typically use a minimum of 32 bits to store each value. This results in many more bit patterns than for fixed point, $2^{32} = 4,294,967,296$ to be exact. A key feature of floating point notation is that the represented numbers are *not* uniformly spaced.



DIGITAL SIGNAL PROCESSOR

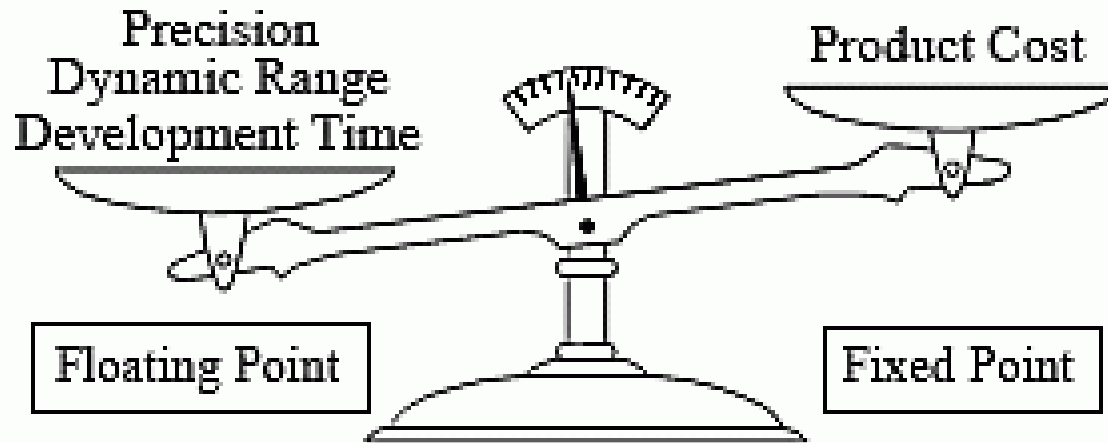
All floating point DSPs can also handle fixed point numbers, a necessity to implement counters, loops, and signals coming from the ADC and going to the DAC.

Fixed point math will be carried out as quickly as the floating point operations; it depends on the internal architecture.

For instance, the SHARC DSPs are optimized for both floating point and fixed point operations, and executes them with equal efficiency.



KEY DIFFERENCE



1
Fixed versus floating point. Fixed point DSPs are generally cheaper, while floating point devices have better precision, higher dynamic range, and a shorter development cycle.



Floating Vs Fixed Point Processors

- The fixed point arithmetic is much faster than floating point in general purpose computers
- The internal architecture of a floating point DSP is more complicated than for a fixed point device.
- All the registers and data buses must be 32 bits wide instead of only 16; the multiplier and ALU must be able to quickly perform floating point arithmetic, the instruction set must be larger (so that they can handle both floating and fixed point numbers), and so on.
- Floating point (32 bit) has better precision and a higher dynamic range than fixed point (16 bit) .
- In addition, floating point programs often have a shorter development cycle, since the programmer doesn't generally need to worry about issues such as overflow, underflow, and round-off error.
- fixed point DSPs have traditionally been cheaper than floating point devices
- **signal-to-noise ratio is more in Fixed point processors**
- FIR filter in fixed point is used to eliminate SNR.



Fixed vs. Floating Point

Characteristic	Floating point 32-bit	Fixed point 16-bit
Dynamic range	much larger	smaller
Resolution	comparable	comparable
Ease of programming	comparable	comparable
Compiler efficiency	much easier	more difficult
Power consumption	more efficient	less efficient
Chip cost	comparable	comparable
System cost	comparable	comparable
Design cost	less	more
Time to market	faster	slower

Thank you