

AEM 617

Software

# Floating Point Numbers in a binary computer

Base 10:

$$\underbrace{d_0.d_1d_2\dots d_p}_{\text{significand } p \text{ digits}} \times \underbrace{\beta^e}_{\text{base}} \Rightarrow 3.1415 \times 10^0 \\
 = 3 \times 10^0 + 1 \times 10^{-1} + 4 \times 10^{-2} + 1 \times 10^{-3} + 5 \times 10^{-4} \\
 = d_0 \times \beta^0 + d_1 \times \beta^{-1} + d_2 \times \beta^{-2} + d_3 \times \beta^{-3}$$

Base 2:

$$b.bbb \times \beta^e = b.bbb\dots b \times 2^e$$

16 bit significand with  $e=1$

$$\begin{array}{cccccccccccccccc}
 1 & 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\
 2^0 & 2^{-1} & 2^{-2} & 2^{-3} & & & & & & & & & & & & 2^{-15}
 \end{array}$$

$$1 + \frac{1}{2} + \frac{1}{16} + \frac{1}{128} + \frac{1}{4096} + \frac{1}{8192} + \frac{1}{16384} + \frac{1}{32768} = \cancel{3.1415} \\
 1.570679$$

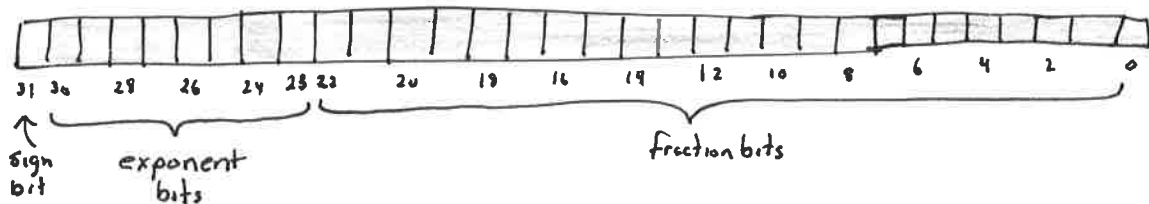
$$\text{Value} = 1.570679 \times 2^1 = 3.141357$$

## IEEE 754 standard

Single precision: "float" in C, "real" in Fortran 32 bits

$$\beta = 2, p = 24, e = 8 \text{ bits}$$

Actually not this simple, we need a sign ( $\pm$ ) bit and  $2^0$  implied bit



$$\text{Value} = (-1)^{\text{sign}} \times \left( 1 + \sum_{i=1}^{23} b_{23-i} \cdot 2^{-i} \right) \times 2^{(e-127)}$$

## Special #s

- Two zeros are possible  $0, -0$   
exponents are  $00_{Hex}$  and fraction bits are  $0x000000$   
The sign bit can be 0 or 1.
- Smaller "denormal #s"  
exponents are  $0x00$  and fraction bits not 0.  
Value =  $(-1)^{sign} \times 2^{-126} \times 0, \text{ fraction bits}$   
Adds extra values near zero. "underflow"
- $\infty$  "infinity"  
exponents are  $0xFF$  and fraction bits  $0x000000$
- NaN "Not a number"  
exponents are  $0xFF$  and fraction bits  $\neq 0$   
NaN is contagious; any operation with an NaN makes the result NaN.

Range: (watch out for the special #s!)

Minimum absolute value:

$$\text{sign bit} = 1$$

$$\text{exponents: } 0xF7 = 254 \Rightarrow e-127 = 127$$

$$\text{fraction bits: } 0xFFFFFFFF \text{ (if we had 24 bits, but we only have 23 stored)}$$

$$\sum_{i=0}^{23} 2^{-i} = 2$$

$$\text{Value} = -3.4028 \times 10^{38}$$

Max value

$$\text{Value} = 3.4028 \times 10^{38}$$

Smallest representable #.

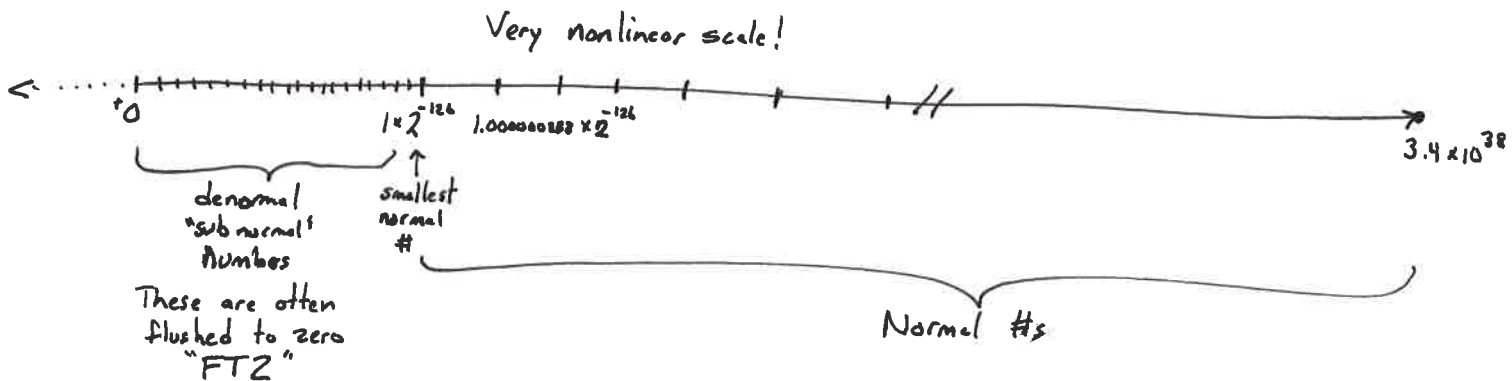
denormal normal #

$$V = 2^{-126} \approx 1.1 \times 10^{-38}$$

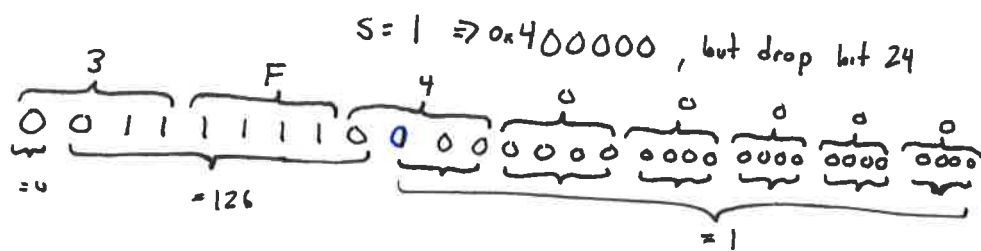


What #s can be represented with IEEE 754 single precision?

Q: Can every number be represented? A: No!



• Represent  $\frac{1}{2} = 1 \times 2^{-1} \Rightarrow$  Sign = 0  $e^{-127} = 1 \Rightarrow e = 126 = 0x7E$



exactly representable! 0x3FB00000

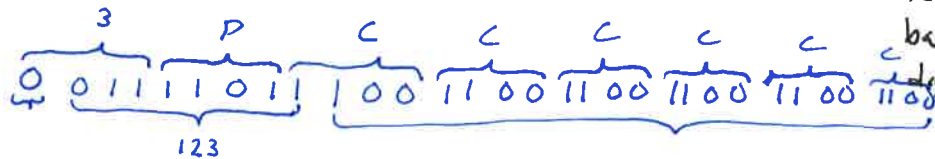
• Represent  $\frac{1}{10} = 1.6 \times 2^{-4} = 1.6 \times 0.0625$ . Can we represent 1.6 as  $1 + \sum_{i=1}^{23} a_i 2^{-i}$ ?

No. 1100, 1100, 1100, 1100, 1100, 1100  
gets us 1.599999905

Error for S.p.  $\approx 9.5 \times 10^{-8}$

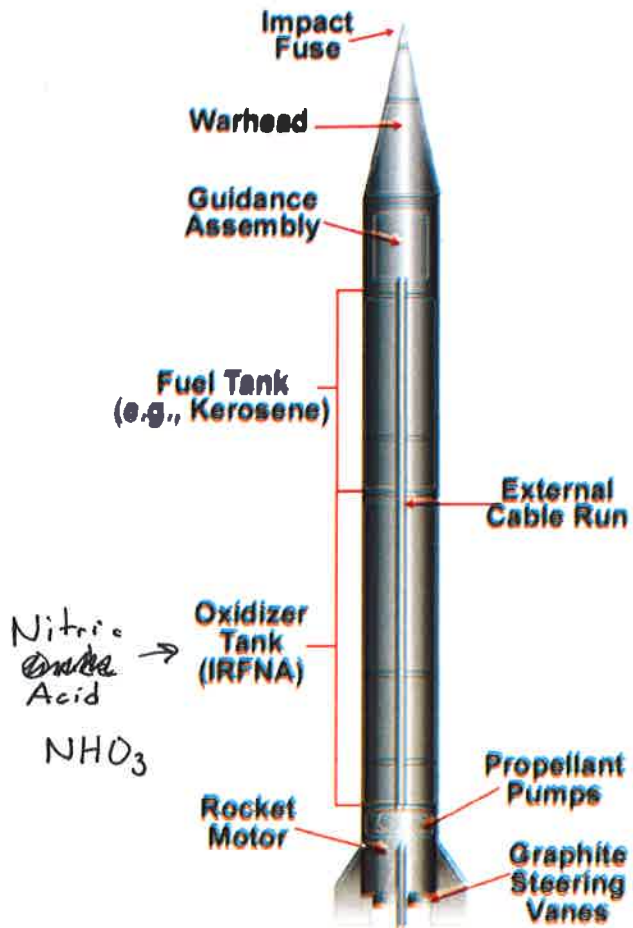
← Notice a pattern?  
We will NEVER  
get exactly 1.6.

$\frac{1}{10} \approx 0.99999994039$



Exactly for the same  
reason that  $\frac{1}{3}$  in  
base 10 has repeating  
digits = 0.3333333333333333

# SCUD



Range  $\approx$  400 miles

CEP  $\approx$  1000 m

500 kg payload



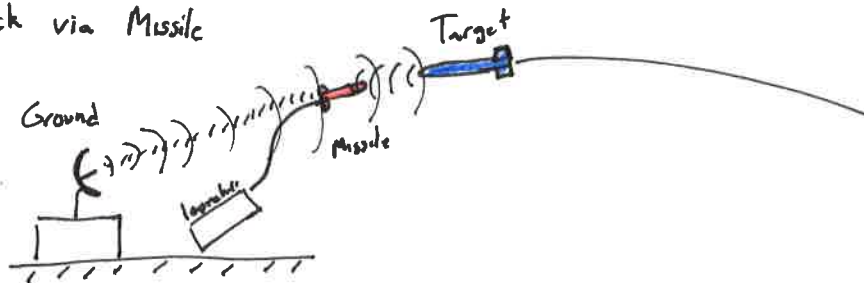


# Patriot Air Defense Missile Failure

GAO Report:

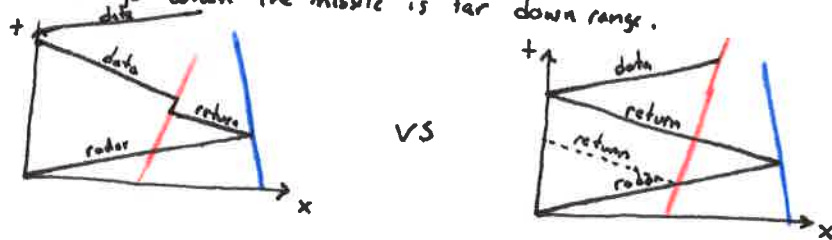
"On Feb 25, 1991, a Patriot missile defense system operating at Dhahran, Saudi Arabia, during Operation Desert Storm failed to track and intercept an incoming Scud. This Scud subsequently hit an Army barracks, killing 28 Americans."

Track via Missile



- 1) Radar signal transmitted from ground station (includes missile guidance commands)
- 2) Radar return from ~~target~~ target received by missile
- 3) Missile transmits data to ground
- 4) Ground ~~provides~~ calculates missile guidance command for next transmit cycle.

Notice the huge advantage when the missile is far down range.



However, the battery (ground) and missile must be synchronized well.

The Patriot system at Dhahran used a ~~rough~~ precision  $\Delta t$  of  $\frac{1}{10}$  seconds expressed as an integer (0, 1, 2, 3, 4, ...)

Unfortunately, the system was left on  $\approx 100$  hrs. (3.6 million time steps)

Unfortunately, the system used a 24 bit floating point number ( $\beta=2, p=20$ )

$$\Delta t \text{ error} = 9.5 \times 10^{-8}$$

$$\text{total time offset} = 9.5 \times 10^{-8} \cdot 3.6 \times 10^6 = 0.34 \text{ s}$$

Unfortunately, a Scud travels at 4000 mph during intercept.  $L = \Delta t \cdot V \approx 2000 \text{ ft}$



Figure 3: Correctly Calculated Range Gate

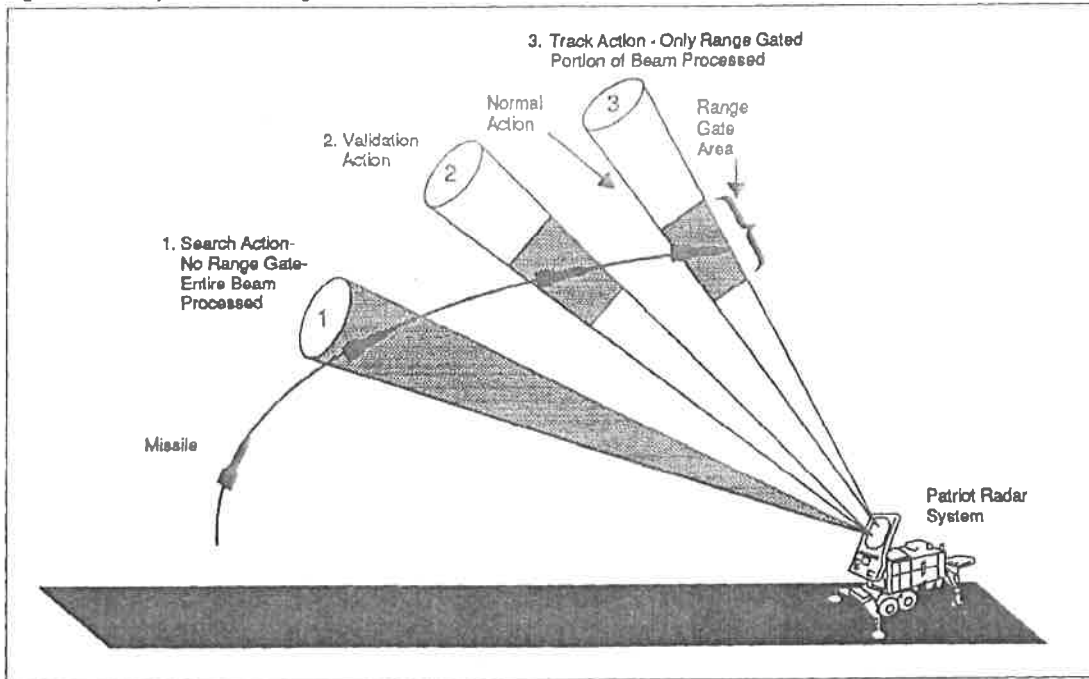
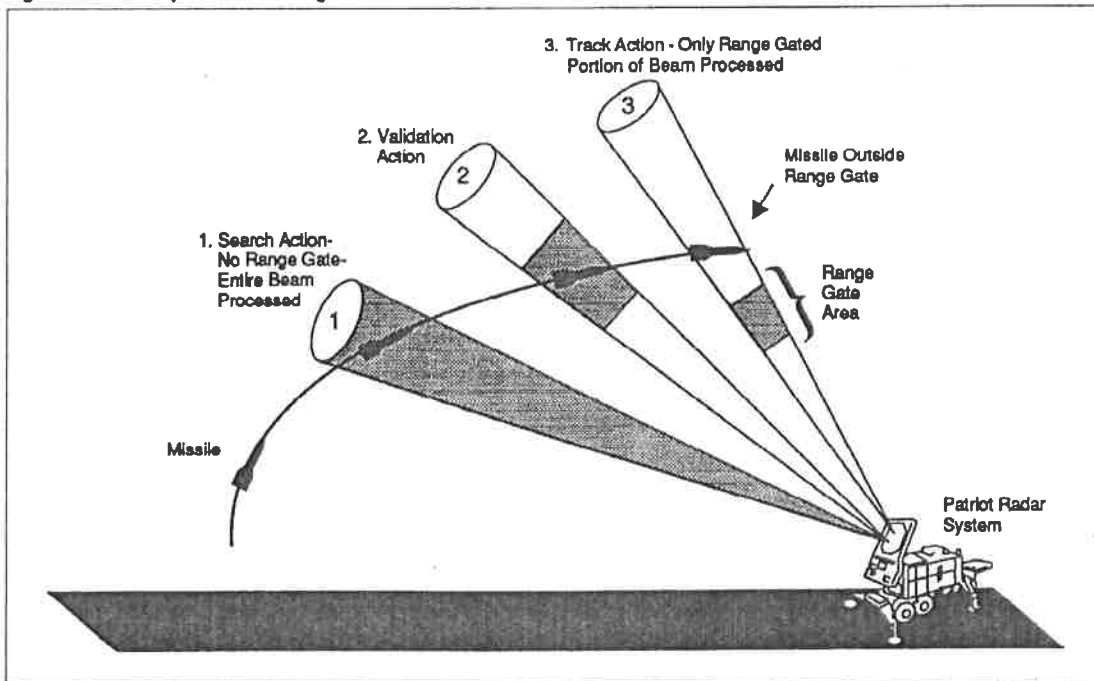


Figure 5: Incorrectly Calculated Range Gate



But wait, the time offset shouldn't matter!

Correct!

Unfortunately, an updated software package corrected the time offset in some of the code .... but not everywhere.

What is the true failure root cause?

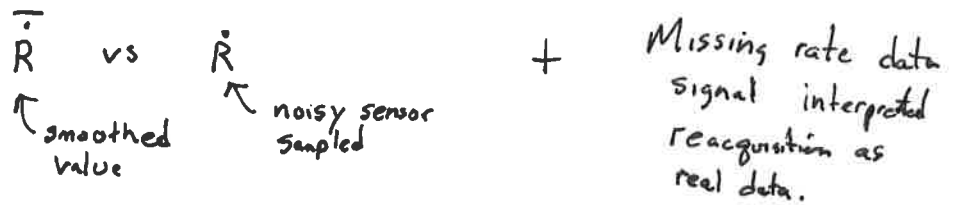
Asides

- Updated software fixing this error arrived one day late to Dhahran. (Patriot Project Office, Huntsville, AL)
- The Israeli military informed the U.S. Army of this issue.
- 4 days before, PPO sent message warning not to "operate for very long run times" but failed to mention what is long.

Being late and imprecise can kill.

# Mars I (1962)

Guidance system failure:



- NOT caused by the often repeated myth of the Fortran DO loop

DO 5 K = 1 . 3       $\Rightarrow$   $\underbrace{DO 5 K = 1.3}_{\text{variable}}$

rather than

DO 5 K = 1, 3       $\Rightarrow$  loop k = 1, 2, 3

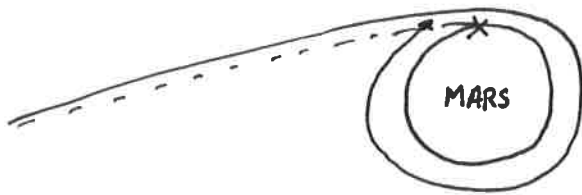
# Mars Climate Orbiter (1999)

\$125M

Units error in thrust performance data.

Data in lbf, programmers assumed Newtons

The space craft entered the Martian atmosphere and burned up rather than establishing orbit.



# USSR Gas Pipeline

## Enemy Action!!

Contrary to popular belief, Soviet Union development of advanced technology was ~~not~~ relatively weak and relied on bought or stolen western technology.

e.g.

- B-29 copy (TU-4) exact down to rivets.  
Tires bought at U.S surplus!
- Space Shuttle vs Buran
- A Soviet guest to Boeing applied adhesive to his shoes to gather metal samples.
- In 1972, the Soviets bought 25% of wheat by intercepting phone calls from market players (allegedly).  
The U.S. govt subsidized wheat prices and thus paid the Soviets to take the grain!
- Nixon's National Security Decision Memorandum (NSDM) 1974 restricted computer technology sales to the Soviets. (15 yrs behind in 1984 according to Gorbachev's adviser)  
Moore's law double 18 months =  $2^{10}$  = 1000 slower
- MiG-15 engine directly copied from British design sold to Soviets.

What would you do?

The CIA "allowed" the Soviet network to purchase micro controllers. These systems were <sup>in part</sup> used for controlling a pipeline network in Siberia (with the aim of supplying gas to Europe). The  $\mu$  controllers were ~~purposely~~ <sup>purposely</sup> "designed" to pass tests and then fail in a dangerous mode after some time.

Result:

The largest non-nuclear man-made explosion in known history.

No EMP, seismography. All purchasers suspect, Soviets/Russians never acknowledged blast. Nobody killed!

# Ariane 5 Rocket

([tiny.cc/AEM617Ariane](http://tiny.cc/AEM617Ariane))

Failure Report Forward, " On 4 June 1996, the maiden flight of the Ariane 5 launcher ended in a failure. Only about 40 seconds after initiation of the flight sequence, at an altitude of about 3700m, the launcher veered off its flight path, broke up and exploded."

" Investigation... showed ...

- Nominal behaviour ... up to  $310 + 36$  seconds
- Failure of the back up Inertial Reference System followed immediately by failure of the active Inertial Reference System
- Swivelling into the extreme position of the nozzles of the two solid boosters and slightly later, of the Vulcain engine, causing the launcher to veer abruptly.
- Self destruction of the launcher correctly triggered by rupture of the links between the solid boosters and the core stage."

Post-flight analysis of telemetry has shown a number of anomalies which have been reported to the Board. They are mostly of minor significance and such as to be expected on a demonstration flight.

One anomaly which was brought to the particular attention of the Board was the gradual development, starting at  $H_0 + 22$  seconds, of variations in the hydraulic pressure of the actuators of the main engine nozzle. These variations had a frequency of approximately 10 Hz.

There are some preliminary explanations as to the cause of these variations, which are now under investigation.

After consideration, the Board has formed the opinion that this anomaly, while significant, has no bearing on the failure of Ariane 501.

## **2. ANALYSIS OF THE FAILURE**

### **2.1 CHAIN OF TECHNICAL EVENTS**

In general terms, the Flight Control System of the Ariane 5 is of a standard design. The attitude of the launcher and its movements in space are measured by an Inertial Reference System (SRI). It has its own internal computer, in which angles and velocities are calculated on the basis of information from a "strap-down" inertial platform, with laser gyros and accelerometers. The data from the SRI are transmitted through the databus to the On-Board Computer (OBC), which executes the flight program and controls the nozzles of the solid boosters and the Vulcain cryogenic engine, via servovalves and hydraulic actuators.

In order to improve reliability there is considerable redundancy at equipment level. There are two SRIs operating in parallel, with identical hardware and software. One SRI is active and one is in "hot" stand-by, and if the OBC detects that the active SRI has failed it immediately switches to the other one, provided that this unit is functioning properly. Likewise there are two OBCs, and a number of other units in the Flight Control System are also duplicated.

The design of the Ariane 5 SRI is practically the same as that of an SRI which is presently used on Ariane 4, particularly as regards the software.

Based on the extensive documentation and data on the Ariane 501 failure made available to the Board, the following chain of events, their inter-relations and causes have been established, starting with the destruction of the launcher and tracing back in time towards the primary cause.

- The launcher started to disintegrate at about  $H_0 + 39$  seconds because of high aerodynamic loads due to an angle of attack of more than 20 degrees that led to separation of the boosters from the main stage, in turn triggering the self-destruct system of the launcher.
- This angle of attack was caused by full nozzle deflections of the solid boosters and the Vulcain main engine.

- These nozzle deflections were commanded by the On-Board Computer (OBC) software on the basis of data transmitted by the active Inertial Reference System (SRI 2). Part of these data at that time did not contain proper flight data, but showed a diagnostic bit pattern of the computer of the SRI 2, which was interpreted as flight data.
- The reason why the active SRI 2 did not send correct attitude data was that the unit had declared a failure due to a software exception.
- The OBC could not switch to the back-up SRI 1 because that unit had already ceased to function during the previous data cycle (72 milliseconds period) for the same reason as SRI 2.
- The internal SRI software exception was caused during execution of a data conversion from 64-bit floating point to 16-bit signed integer value. The floating point number which was converted had a value greater than what could be represented by a 16-bit signed integer. This resulted in an Operand Error. The data conversion instructions (in Ada code) were not protected from causing an Operand Error, although other conversions of comparable variables in the same place in the code were protected.
- The error occurred in a part of the software that only performs alignment of the strap-down inertial platform. This software module computes meaningful results only before lift-off. As soon as the launcher lifts off, this function serves no purpose.
- The alignment function is operative for 50 seconds after starting of the Flight Mode of the SRIs which occurs at H0 - 3 seconds for Ariane 5. Consequently, when lift-off occurs, the function continues for approx. 40 seconds of flight. This time sequence is based on a requirement of Ariane 4 and is not required for Ariane 5.
- The Operand Error occurred due to an unexpected high value of an internal alignment function result called BH, Horizontal Bias, related to the horizontal velocity sensed by the platform. This value is calculated as an indicator for alignment precision over time.
- The value of BH was much higher than expected because the early part of the trajectory of Ariane 5 differs from that of Ariane 4 and results in considerably higher horizontal velocity values.

The SRI internal events that led to the failure have been reproduced by simulation calculations. Furthermore, both SRIs were recovered during the Board's investigation and the failure context was precisely determined from memory readouts. In addition, the Board has examined the software code which was shown to be consistent with the failure scenario. The results of these examinations are documented in the Technical Report.

Therefore, it is established beyond reasonable doubt that the chain of events set out above reflects the technical causes of the failure of Ariane 501.

## **2.2 COMMENTS ON THE FAILURE SCENARIO**

In the failure scenario, the primary technical causes are the Operand Error when converting the horizontal bias variable BH, and the lack of protection of this conversion which caused the SRI computer to stop.

It has been stated to the Board that not all the conversions were protected because a maximum workload target of 80% had been set for the SRI computer. To determine the vulnerability of