Describing the speed of Morse code transmission

Douglas A. Kerr

Issue 1
November 27, 2022

ABSTRACT

From a traffic capacity and operator proficiency standpoint it is common to specify the *speed* of Morse code transmission in words per minute (WPM), and of course we immediately wonder “what counts as a ‘word’ ?“

From a transmission channel standpoint, engineers are interested in the *modulation rate* (denominated in bauds), the reciprocal of the time resolution of the serial signal (in seconds).

In this article I discuss how these are related under a certain standard convention.

An appendix presents several related topics.

1 THE MORSE CODE

1.1 Eponymy

The Morse code is named after Samuel Finley Breese Morse, considered (especially in America) as the “father of the telegraph”.

1.2 American and International Morse code

Wireline telegraph operation in the U.S. used Morse’s original code, now identified as the “American Morse” code. But wireline operation in Europe, and radiotelegraphy operation worldwide, used a code that differed for some characters, now identified as “International Morse”.

A difference is that in the International Morse code, the codes for the characters are sequences of pulses of only two lengths, while in the American Morse code, some characters involve pulses of additional lengths, some characters have an internal gap between pulses that is longer than normal, and some characters just have different codes.[1]

All the further discussions here assume the use of the International Morse code.

[1] What is now called the International Morse code was derived from the American Morse code by way an intermediate code by Friedrich Clemens Gerke of Germany.

1.3 Basic concept

Morse code is is a binary serial signal. Each character in the code alphabet is represented by a unique series of long and short pulses of one state against a background of the other state. The lengths of the characters vary, generally with those that occur most commonly in normal English text being the shortest, in the interest of most compactness of transmission.

1.4 Mark and space

In a binary telegraph channel (today we would say more generally "data channel") there are two logical states, which are called "mark" and "space" (sometimes, especially when they are used as adjectives, "marking and spacing").[2] In Morse operation the mark state corresponds to the key being down, and the space state to the key being up.

In the description that follows, I will use the terms "marking" and "spacing" to avoid confusion with the word "space" used in the familiar way (although as we will see shortly, the two are closely related).

The terms come from the early practice of having the arriving signal (over a wireline channel) make a recording on paper tape. Typically the tape was moved at a constant speed by a clockwork mechanism.

An electromagnet pushed an ink pen nib or just a blunt embossing stylus against the tape when there was current in the line. When there was current, a "mark" was made on the tape, and these marks of course were separated by "spaces" (when there was no current). Thus the terminology for the states of the telegraph channel.

1.5 Dot and dash, dit and dah

The short and long pulses ("elements") of which Morse characters are composed are traditionally known as "dot" and "dash", respectively. Folklore suggests that Morse himself conferred those names in honor of the nicknames of the two young girls who lived next door. I will not comment here on the credibility of that story.

But in practical work, the short and long elements are usually called "dit" and "dah", respectively, these mimicking the sounds of the two elements as received over a radio receiver in radiotelegraph operation.

---

[2] Often today the spacing state is considered equivalent to binary "0" and the marking state to binary "1".

However, in the interest of euphony, when "speaking" the code for a certain character, if a short element is followed by another element, it is elided to "di".

Thus, for the character "V", whose code is short-short-short-long, it is spoken as di-di-di-dah, while for the character "R", whose code is short-long-short, it it spoken as "di-dah-dit".

1.6 Physical transmission

On a "wireline" channel, marking might be represented by the flow of current and spacing by no flow of current. Or perhaps marking is conveyed by current flow in one direction and spacing by current flow in the other direction.

Or marking might be conveyed by the presence of a tone of some frequency over the channel, and spacing by no tone. Or marking by a tone of a certain frequency and spacing by a tone of a different, but nearby, frequency.

In radiotelegraph transmission, marking might be represented by the transmission of the carrier wave, and spacing by no transmission. Or perhaps marking is conveyed by the transmission of a carrier wave at one frequency and spacing by the transmission of a different, but nearby, frequency. Or audio tones might be sent (essentially as described above) by modulation of the carrier wave. And so forth.

2 TIME STRUCTURE

2.1 Introduction

The time structure described below is the "ideal" one. Obviously, when sending manually, via a traditional telegraph key, the actual timing will depart from the ideal. Nonetheless, the discussions to follow are predicated on the ideal time structure.

2.2 Figure

We can see the time definitions in figure 1 shown on a sample short phrase, "AM I".

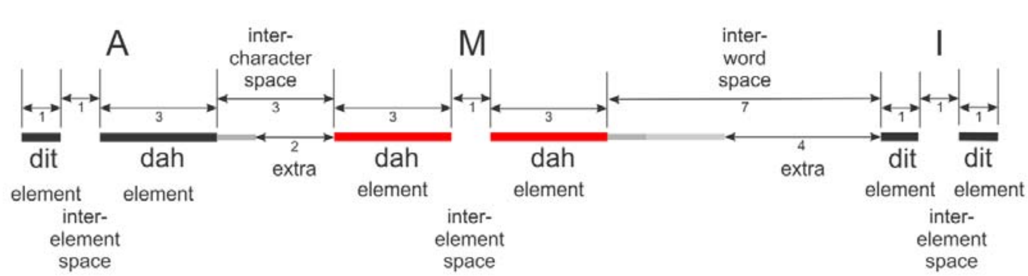

Figure 1. Morse code time interval definitions

Alternate characters are in black and red to help distinguish them. The gray bar after the dah of the “A” represents what would be the trailing inter-element space of that dah. The gray bars after the dah of the “M” represents what would be the trailing inter-element space of that dah and the rest of the trailing inter-character space of the “M”. Both of these help us to see the “extra” space added between characters and between words.

2.3 Unit time

The time descriptions to follow are normalized in terns of the “unit time”. Of course, the actual value of the unit time depends on the speed of transmission.

2.4 Dit and dah elements

A dit is a period of **marking** of length 1 unit, surrounded by periods of **spacing**.

A dah is a period of **marking** of length 3 units, surrounded by periods of **spacing**.

2.5 Spaces

Here “space” is not used as the opposite of “mark” but in the sense of a gap between constituents of the character’s code (although those two are really the same).

Within a multi-element character, the interval of **spacing** between its constituent elements (the “inter-element space”) is of length 1 unit.

Within a word, between characters, the interval of **spacing** between the adjacent elements (the “inter-character space”) is 3 units. If we think of each element as carrying after it its own “inter-element space”, then between characters we have an extra 2 units of **spacing** (which is the length of a dit with its associated inter-character space).

Within a message, between words, the interval of **spacing** between the adjacent elements (the “inter-word space”) is 7 units. If we think of each character as carrying after it, its own “inter-character space”, then between words we have an extra 4 units of **spacing** (which the length of a dah with its associated inter-character space).[3]

[3] At an earlier time (before 1949 in the international standard), the inter-word space was defined as 5 units.

3 SPEED OF TRANSMISSION

3.1 The basic concept

In traditional "telegram" services, the cost to send a message was based on the number of "words" in it. Of course if message were in actual text, the definition of a "word" was the familiar one.

3.2 Encrypted traffic

But if the message is of random-seeming letters (as if it were encrypted), that does not work out so well. For example, a customer might present this message to the telegraph operator:

```
AKDEPFKWMOWRTUVGEANT JEMEPRSLYHESOMJWORWP
BEHHAOMLWNOENGUOFLJZ UAWKYHGMSONUENUNUDZQ
POCEKNLESUGFMZQUPMFJ
```

and hope to be charged for five words (the usual "minimum" charge). So of course the telegraph agency would have special rules when the "words" were not (mostly) recognized English words.[4]

3.3 Operator proficiency

Another complication came about in the matter of the proficiency of the telegraph operator. At one time, applicants for most classes of FCC amateur radio licenses were required to show that they could send and receive Morse code at a certain speed. Applicants for FCC commercial radiotelegraph certificates (now called licenses) were required to do the same. These speeds were generally described in the regulations in terms of words per minute (WPM). But what might that mean when "calibrating" a test passage for the receiving test, or to be sent in the transmitting test?

I do not know what the precise premise was for defining transmission speed in WPM from the actual timing of the transmitted characters. It was generally said that this was based on "five characters count as a word". But of course we now know that this is only a partial definition.

It seems likely that the full definition is that described below in section 4 (based on the length of the model word "PARIS" as the length of a "word"), but maybe not.

In certain licensing situations, both text and quasi-random 5-letter groups were sent to the applicant. In one situation, it was said that for the "text" part of the test, the transmission speed was 20 WPM,

---

[4] Perhaps a wholly different rate schedule would apply, as it was tedious to accurately key such a message.

while for the "letter group" portion of the test, the transmission speed was said to be 16 WPM. (There is a further discussion of this in section 4.)

3.4 Modulation rate

In the matter of the transmitting channel (which of course did not understand anything about the Morse code), the "capacity" was generally known in terms of the maximum supportable modulation rate (the unit being the *baud*), which we can think of as the inverse of the time resolution (in seconds) of the binary "message" being sent. In terms of Morse code, this was the reciprocal of the unit time.

So if the maximum modulation rate of the channel were, say, 20 bauds, then how might the engineers explain to those responsible for the flow of telegraph traffic what was the capacity in words per minute?

3.5 Sorting this out

With respect to each of these issues, various definitions were adopted, some of them not very precise. For example, often a basic definition was that a "word" comprised five characters plus the following inter-word space.

But of course, in Morse code, the lengths of the characters varied widely. An "E" followed by its inter-element space was 2 units in length; a "J" followed by its inter-element space was 12 units in length. This was not really precisely handled in the matters of operator proficiency speed testing.

But in the case of the modulation rate, a quite rigorous definition was eventually adopted. I will speak of that here.

4 MODULATION RATE VS. WORDS PER MINUTE

4.1 The definition

It was eventually established by international regulation that the equivalence of a speed in words per minute to the modulation rate in bauds was based on a "word" that was 50 unit times in length (including its trailing inter-word space). That ratio was chosen based on a survey of the word length distribution in a large sample of telegraph messages (probably stretched to come out to a handy value).

4.2 Implication

That meant that for operation at 25 words per minute, the modulation rate was 20.83 bauds.

4.3 A mnemonic

To help us remember how this worked, we are reminded that the word "PARIS" in international Morse code is exactly 50 unit times in length (including its trailing inter-word space).

How this works out for that word is seen on figure 2. Again, here, alternate characters are in black and red to help distinguish them.

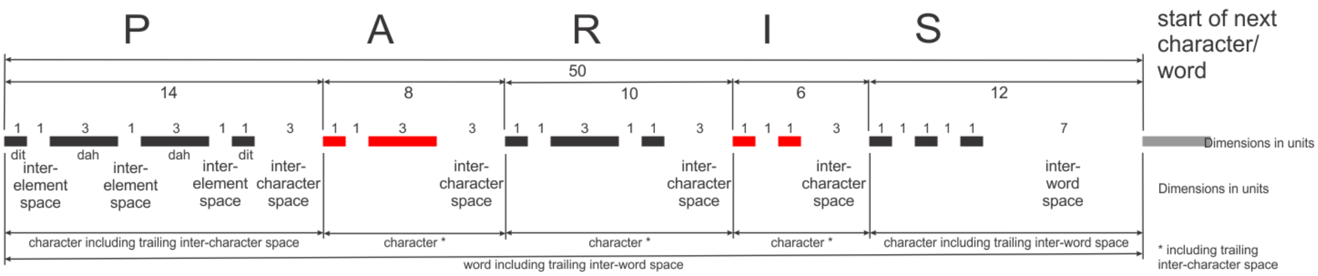

Figure 2. Standard reference word "PARIS"

Thus, for example, if we were to send "PARIS", using exactly the ideal time intervals, 25 times in a minute, the modulation rate would be 20.83 bauds.

4.4 An alternate

In a context where the traffic was typically of random or quasi-random characters (as for a channel used almost exclusively for encrypted traffic), where often the formulation was in terms of groups of five letters separated by spaces (each considered to be a "word"), the norm is often followed that a "word" is considered to have the length of 60 unit times.[5] This is said to have been derived from a model of quasi-random letter transmission in groups of five, with the standard inter-word space between the groups.

The mnemonic usually cited here is that the "word" "CODEX" (including its trailing inter-word space) has a length of exactly 60 unit times.

How this works out for that word is seen on figure 2.

---

[5] The theoretical value is 59.38 units.

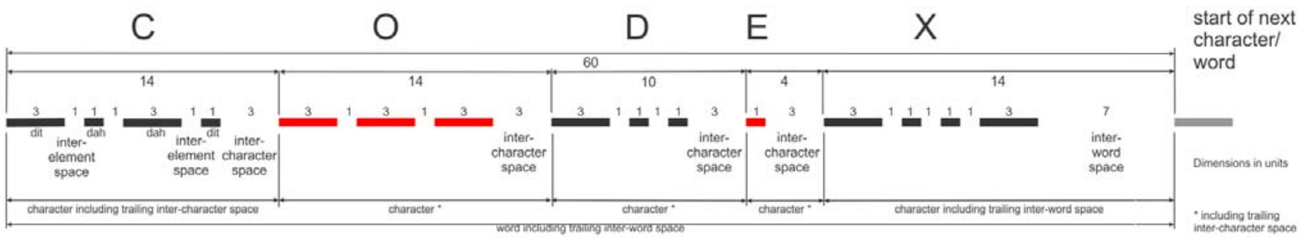

Figure 3. Standard reference word "CODEX"

Under this outlook, transmission at 25 words per minute would require a modulation rate of 25 bauds.

4.5 In certain license examinations

Through 2013, there were three classes of commercial radiotelegraph licenses. For all three, the Morse code test included a portion in which running text was sent and another section in which quasi-random five-letter groups were sent (the latter being considered equivalent to typical encrypted traffic). For the first-class license, the running text was sent at 25 WPM and the five-letter groups were sent at 20 WPM. For the other classes the running text portion was sent at 20 WPM, and the five-letter groups were sent at 16 WPM.

Would that mean that the microscopic character times (*e.g.*, the length of a dit or dah) was the same for both portions? That is seemingly the intent.

If we assume that the model "word" PARIS and the model letter group CODEX are used in this reckoning, then we would find that, if the running text were sent at 25 WPM, then for the same dit and dah lengths, the letter groups would have to be sent at 20.83 WPM. If the running test were sent at 20 WPM, then five-letter groups would have to be sent at 16.67 WPM.

It may in fact be that the five-letter group test sequences were actually sent at 16.67 WPM and 20.83 wpm, but are spoken of as "16 WPM" and "20 WPM"

-#-

Appendix A
Related topics

A.1 AUTOMATIC MORSE OPERATION

Many transoceanic Morse systems used automatic operation. Here, the operators key the message into perforators (reminiscent of teletypewriter perforators) which punch the Morse code sequences into a 2-track tape (a system attributed to Charles Wheatstone, another important telegraph pioneer). The tape is then fed into Wheatstone transmitters, which read the sequences from the tape and key the Morse telegraph circuit accordingly. The transmitters were set to run at a speed such that the baud rate generated was that suitable for the particular channel.

Accordingly, the traffic was sent at the maximum rate supported by the channels' baud rate capability, and was not dependent on the skill of the operator to send at exactly that rate. Additionally, the timing details were essentially ideal, again not dependent on the skill of the operator. This facilitated reception of the message at the receiving end, and in fact enhanced to the possibility that the signals would be received automatically, the messages perhaps being automatically printed on paper tape or even a page printer (if that was suitable for the message format).

The encoding scheme on the tape was very clever, but the details are beyond the scope of this article. In any cse, the scheme was that each step along the tape corresponded to 2 unit times. This a dit occupied 1 step along the tape and a dah 2 steps.

Accordingly, the word "PARIS" with its trailing inter-word space occupied 25 steps along the tape. Thus, for operation at 25 WPM, the tape would be advanced at about 10.42 steps per second.

A.2 THE FARNSWORTH SYSTEM

In training aspirants for a commercial radiotelegraph certificate/license or an amateur radio license to receive Morse code transmissions, it was recognized that the student should learn to grasp a character by its overall sound, not by a conscious recognition of its constituent dits and dahs.

But if the standard timing were used, then in the earlier stages of training, necessarily at low character rates, the character detail was so coarse that the student could not help but think of the character in terms of the individual dits and dahs, not the audible "gestalt" the student needed to recognize from the outset.

This led to an alternate method of training in which, in the early stages, the student was sent a transmission in which the composition of the characters (and their sound) was consistently as if for continuous transmission at a fairly high speed, but the rate at which those characters were sent was modest, gradually increasing as the student gained proficiency. This came to be known as "Farnsworth system" of Morse code training[6].

In fact, the American Radio Relay League (ARRL), the major U.S association of amateur radio operators, and the distributor of materials for the teaching of Morse code (including audio tapes with training messages at various speeds), in about 1990 (I think) converted all its Morse code training material to use the Farnsworth system.

In their version of the system, the character timing was consistently that which would have been occurred at 18 WPM under the concept described in section 4, based on the model "word" "PARIS". For lower speeds, the dit and dah length and the inter-element spacing were still what the 18 WPM speed would have called for.

However, both the inter-character space and the inter-word space were increased (still maintaining the ratio of 3:7 between them) to produce a *character rate* that would (again based on the model word "PARIS") be considered as being at the "WPM" rate for the current stage of the training.

At an actual speed of 18 WPM, the model word "PARIS" would (including its trailing interword space) last 3.33 seconds. The inter-character spaces would be 200 ms; the inter-word space would be 465 ms.

But for training at a speed of 10 WPM, "PARIS" (if repeated) would need to occur every 6.00 seconds. But, in the Farnsworth system, the word proper would still last only 3.33 seconds (just as it did at 18 WPM). Thus an additional 2.67 seconds of "dead space" would have to be inserted. It is equally divided between the 19 unit "slots" in the 4 inter-character spaces (3 slots each) and the inter-word space (7 slots).

The result would be that the inter-character spaces would grow to 421 ms and the inter-word space to 983 units (about twice the length they are for operation at 18 WPM with "normal" timing).[7]

---

[6] Said to have been invented by Donald R. (Russ) Farnsworth, W6TTB. Yet it is said that Farnsworth himself did not employ this system in his own work as a teacher of Morse code. So go figure.

[7] Thanks to Jonathan R. Bloom, KE3Z, of the ARRL Laboratory for his excellent description of how this works.

For transmissions at speeds above 18 WPM, the normal timing is used (just as at 18 WPM).

Note that the use of 18 WPM for the "character" speed isn't inherent in the Farnsworth plan. That is just the value chosen by ARRL for their Morse training materials (based on their extensive experience in that matter).

A.3 COMPARISON WITH TELETYPEWRITER OPERATION

A.3.1 Introduction

Of course, many readers my not have ever been involved with Morse code telegraph operation, but might be familiar with teletypewriter operation.

The character format here assumes an "idle" state for the telegraph channel of *marking*. Each character begins with a *start element*, one unit in length, of the *spacing* state. This is followed by five *information elements*, each one unit in length. These can each be *marking* or *spacing*, and carry the code for the character in what we would today call a "five-bit" code.

The character ends with a stop element, of the *marking* state, whose specified length is different under different transmission norms. We will first look at a norm in which the duration of the stop element is specified as 1.5 units. Thus, under that norm, the length of a character is (always) 7.5 units.

But if transmission is not continuous at the rate implied by the modulation rate in use (perhaps the characters are sent by manipulation of a keyboard, at an irregular pace), the stop element (*marking*) blends into the following "idle" period (also *marking*). Thus, the defined stop element can be thought of as the minimum permissible idle time between characters that are 6 units long.

A.3.2 An international norm

For comparison, one international norm for teletypewriter operation (pre-ASCII, using a "5-bit" code) provides for operation (on a "full-speed", or "cadence" basis) at 400 characters per minute, $6.\underline{66}$ characters per second.

The premise of the "word speed" definition is that a word comprises 5 characters plus a SPACE character, 6 characters overall. Thus this rate is considered to be $66.\underline{66}$ words per minute.

The characters are all 7.5 unit times in length. Thus the resulting modulation rate is 50 bauds. The unit time is exactly 20 ms.

When the messages are punched into 5-track perforated tape, each "step" of the tape represents 1 character. Thus, for operation at the speed mentioned just above, the tape is stepped at 6.<u>66</u> steps per second.

A.3.3 A North American norm

A comparable North American norm provides for operation at (theoretically) 367.56 characters per minute (usually cited as "368"), 6.126 characters per second. Under the same definition of a "word" mentioned above, this is 61.26 WPM. But it was spoken of as "60 WPM" operation in the roster of standard speeds.

The characters are all 7.42 unit times in length. Thus the resulting modulation rate is 45.45 bauds. The unit time is exactly 22 ms.

When the messages are punched into 5-track perforated tape, each "step" of the tape represents 1 character. Thus, for operation at the speed mentioned just above, the tape would be stepped at 6.126 steps per second.

Where these peculiar numbers came from is a fascinating story, beyond the scope of this article.

-#-